

Lawrence Berkeley National Laboratory

LBL Publications

Title

Field testing and validation of a low-cost MPC for demand flexibility for grid-interactive K-12 schools

Permalink

<https://escholarship.org/uc/item/7676p1xs>

Journal

Energy and Buildings, 360

ISSN

0378-7788

Authors

Ham, Sang Woo
Hariri, Mohamad El
Kim, Donghun
[et al.](#)

Publication Date

2026-06-01

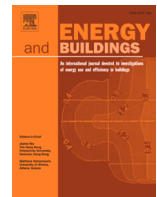
DOI

10.1016/j.enbuild.2026.117378

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial License, available at <https://creativecommons.org/licenses/by-nc/4.0/>

Peer reviewed



Field testing and validation of a low-cost MPC for demand flexibility for grid-interactive K-12 schools

Sang Woo Ham^{a,*}, Mohamad El Hariri^b, Donghun Kim^a, Tanya Barham^b

^a Building & Industrial Energy Systems Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

^b Community Energy Labs, Portland, OR, USA

ARTICLE INFO

Keywords:

Model predictive control
School buildings
Demand flexibility
Peak demand reduction
Grid reliability
Technology transfer

ABSTRACT

K-12 school buildings account for the highest energy consumption within the public sector. Implementing advanced HVAC controls in grid-interactive K-12 schools could bring substantial economic advantages and grid flexibility. Our previous study demonstrated that a low-cost model predictive control (MPC) solution, which coordinates multiple packaged units, can enable demand flexibility without major hardware upgrades. However, a significant gap remains between academic pilots and market-ready scalable solutions. This paper extends the previous single-site pilot to a multi-site demonstration involving three school campuses (95 total units) through a commercial technology transfer process. Addressing the challenge of verifying performance with sparse field data, we present a new statistical approach using Bayesian methods to estimate the MPC's effect on peak demand. Unlike traditional methods, this approach robustly quantifies uncertainty in non-normal, limited datasets. The results confirm the solution's replicability, achieving a 21.6–38.9% reduction in HVAC peak demand (10.8–22.1% at the site-level) with >98% probability across diverse locations. Finally, we document critical barriers to scaling software-as-a-service (SaaS) solutions—such as API instability and diverse legacy systems—and offer practical strategies to accelerate the commercial adoption of grid-interactive efficient buildings.

1. Introduction

1.1. Background and motivation

Electricity grids are increasingly challenged by reliability and resilience issues due to rising electricity demand, especially during peak summer months. Over the next decade, peak demand is expected to grow significantly, driven by factors such as the expansion of data centers and the broader adoption of electrification. This projected increase is further intensified by planned generator retirements, which threaten the adequacy of available resources [1,2]. Additionally, the surge in demand is compounded by transmission limitations and the integration of intermittent renewable energy sources. Addressing these challenges requires innovative grid management strategies. In this context, grid-interactive Efficient Buildings (GEBs), which offer demand flexibility applications such as load shifting and shedding, are an active research area because they can contribute to reducing grid stress and enhance overall system resilience in a cost-effective way [3].

Historically, building control research has focused on optimizing the energy efficiency of heating, ventilation, and air-conditioning (HVAC) systems [4]. However, more recent studies have shifted their focus to-

wards developing advanced control methods that enable buildings to provide grid services through dynamic HVAC operations, including load shifting and peak demand management [5]. These advanced controls can integrate the operations of various resources, such as HVAC systems, solar photovoltaics, and batteries, while considering factors like grid conditions, weather, and occupancy. To achieve this, researchers have employed techniques like model-based optimal control, including model predictive control (MPC) [6], as well as artificial intelligence approaches, such as reinforcement learning [7]. Studies have demonstrated the potential of these advanced control strategies to reduce energy costs [7] and increase the utilization of renewable energy sources [8] in field demonstrations.

K-12 schools, from kindergarten to the 12th grade, possess substantial potential as Grid-Interactive Efficient Buildings (GEBs); nonetheless, they have been largely overlooked in this domain. Specifically, there are approximately 100,000 K-12 schools in the United States [9]. HVAC modernization represents the highest priority among all building systems requiring renovation, with 41% of districts nationwide reporting the need for comprehensive system updates [10].

Beyond hardware retrofits, a crucial area of research focuses on implementing dynamic HVAC operations that respond to grid signals,

* Corresponding author.

E-mail address: sham@lbl.gov (S.W. Ham).

<https://doi.org/10.1016/j.enbuild.2026.117378>

Received 15 September 2025; Received in revised form 27 February 2026; Accepted 23 March 2026

Available online 26 March 2026

0378-7788/© 2026 Elsevier B.V. All rights reserved, including those for text and data mining, AI training, and similar technologies.

particularly during periods of high grid stress or peak demand, through strategies such as adjusting HVAC setpoints or modifying HVAC operational schedules. Because school buildings represent the largest energy consumers in the public sector [11], with HVAC systems accounting for 46% of their total energy consumption [12], this presents significant opportunities for demand flexibility, such as managing peak demand, and lowering energy costs from demand charge or time-of-use rate. In addition, schools implementing grid-interactive operations offer unique advantages beyond typical buildings, serving as educational tools for students, fostering a deeper understanding of energy management and grid reliability [13]. Finally, the standardized nature of HVAC configurations across K-12 schools—typically featuring single- or multi-staged rooftop units with dedicated thermostats—presents an opportunity to develop a universal control solution applicable to these facilities.

1.2. Literature review

Model Predictive Control (MPC) has been extensively studied for decades to optimize dynamic HVAC and building system operations [14]. This approach leverages physics-based building models and forecasted disturbances (e.g., weather conditions) to optimize HVAC performance within specified constraints. When integrated with grid price signals—including Time-of-Use (TOU) rates, and real-time pricing—MPC can effectively manage flexible loads and provide essential grid services [15], particularly in areas of load shifting and peak demand reduction.

Over the past few decades, researchers have extensively studied MPC in building applications. However, most studies have primarily relied on simulation-based analyses, with only a few actual implementations in real buildings. A comprehensive review by Khabbazi et al. [16] analyzed 104 field demonstrations of MPC and reinforcement learning for residential and commercial HVAC systems, revealing that only 4% of nearly 3000 published papers on advanced building control focus on actual field demonstrations. This limited evidence of MPC demonstrations raises concerns among building owners and industry professionals regarding its tangible benefits and technological maturity. Such uncertainty presents a significant obstacle to the broader adoption of MPC technology in the building sector. To understand these barriers, it is essential to review the history and implications of field demonstration research.

Field demonstrations in residential buildings have primarily focused on energy cost minimization through load shifting based on price signals or energy-efficient operation. Pedersen et al. [17] conducted a long-term demonstration in four Danish homes with heat pump radiant floor heating systems, achieving 9% cost savings by shifting electricity usage from high- to low-price periods based on day-ahead market prices—pioneering MPC for demand flexibility. Lindelof et al. [18] tested an MPC controller on 10 residential buildings in Germany and Switzerland with various heating sources during the 2013–2014 heating season, achieving $28 \pm 4\%$ energy savings by optimizing hot water temperature and overriding the existing outdoor air temperature-responsive controller, which reduced distribution losses and prevented overheating. Zong et al. [19] implemented an MPC on a 3-floor residential building at DTU Denmark with water-based radiators during winter, achieving 3–11.3% cost savings based on Nord Pool hourly prices through pre-heating during low-price periods and load shifting to match wind production. Afram and Janabi-Sharifi [20] deployed an MPC on a 5-zone residential building in Ontario, Canada, with multi-zone air handling units, radiant floor heating systems, geo-source heat pumps, and buffer tanks, achieving 8.5–56.1% cost savings (depending on baseline) using Ontario TOU pricing by pre-cooling/pre-heating during off-peak hours and leveraging passive thermal storage. Bunning et al. [21] implemented Data Predictive Control (DPC) using machine learning on a residential unit in Switzerland with water-based ceiling cooling panels, achieving 24.9% cooling energy savings by providing exactly the required cooling (maintaining temperature near the upper comfort boundary). They subsequently compared different modeling strategies (physics-informed

ARMAX, Random Forests, and Input Convex Neural Networks) [22]. Thorsteinsson et al. [23] implemented hierarchical mixed-integer MPC on an occupied 142 m² near-zero emission single-family house in Denmark with an air-to-water heat pump and floor heating during a 97-day winter period (Nov 2022–Mar 2023), achieving 2.3–32.8% cost savings based on Nord Pool hourly spot prices and Danish TOU tariffs by shifting evening peak load to lower-price periods. Jiang et al. [24] implemented MPC-based optimal precooling on 9 single-family homes in Oklahoma and Florida using existing smart thermostats during summer, achieving from -13.5 to 87.1% cost savings through precooling during off-peak hours. Pergantis et al. [25] implemented MPC on an occupied all-electric single-family house in Indiana with an air-to-air heat pump during January–March 2023, achieving 19% energy savings by reducing backup heater use. The same group subsequently incorporated humidity [26] and maximum electricity capacity constraints [27]. In summary, while field demonstrations confirm MPC's ability to minimize costs in individual homes, a significant gap remains in demonstrating demand flexibility (e.g., peak demand reduction) through multi-building coordination. Furthermore, as implementations shift from sensor-heavy setups to thermostat-based solutions, the impact of resident override behaviors [28] and control conflicts [5] on scalable deployments remains inadequately investigated.

Commercial buildings have been a particularly active area for MPC research and field demonstrations. Given the extensive literature, we focus on key research studies with sufficient field trial periods to demonstrate practical viability. MPC applications in commercial buildings initially began in university and research institution buildings, where researchers had easier access to building systems and data. Liao and Dexter [29] demonstrated significant energy cost savings in research institution buildings by optimally controlling radiant heating with thermostatic valves. Building on this foundation, several studies explored thermally activated building systems (TABS), which utilize hydronic ceiling radiant heating systems for energy and cost savings [30–32]. Similarly, radiant floor-based systems in open-plan office spaces have received considerable attention, with researchers developing distributed MPC approaches for both energy optimization [33,34] and personalized comfort delivery through self-tuned controllers that learn occupant preferences online [35].

Office buildings represent the largest category of MPC field demonstrations. Early implementations showed substantial energy savings while maintaining occupant thermal satisfaction [36,37]. Several European studies followed, reporting energy savings ranging from 17–30% compared to conventional control strategies [38–40]. Drgona et al. [41] demonstrated one of the most comprehensive white-box MPC implementations in a GEOTABS office building, achieving substantial energy savings and thermal comfort improvement during challenging transitional seasons. Freund et al. [42] further validated MPC performance in a large-scale, fully occupied office building, while Sampaio et al. [43] used grey-box models calibrated through black-box multiobjective optimization. Lolli et al. [44] confirmed that MPC operation maintained comparable thermal comfort compared to business-as-usual operation, an important finding for user acceptance.

Beyond office buildings, MPC has been successfully applied to other commercial building types, including schools [45] and small commercial buildings with behind-the-meter photovoltaics and batteries [46]. District energy systems have also benefited from MPC implementations, with demonstrations showing energy cost reductions, efficiency improvements, and greenhouse gas emissions reductions in campus-level district cooling systems [8,47] and multi-boiler district heating systems [48]. Yue et al. [49] successfully integrated data-driven MPC into the existing Building Automation System of an airport terminal, achieving substantial energy consumption reduction and thermal comfort improvement.

Despite these successes, most commercial demonstrations rely on custom engineering or complex BMS integrations, highlighting a critical need for scalable, low-cost solutions for the fragmented

small-to-medium commercial building (SMCB) sector. These buildings vary widely in their characteristics, including differences in metadata and BAS configurations, such as naming conventions for measurable and controllable variables and database structures. Implementing MPC in a specific building typically requires highly skilled building and control engineers, as well as advanced BAS platforms that support control customization and time-series databases for model development and real-time control.

The implementation process involves numerous manual, site-specific tasks, such as understanding the building's design, determining the model structure and order, selecting an identification algorithm with suitable initial guesses, obtaining high-quality datasets, and using appropriate software tools [50]. These challenges persist even when advanced modeling software [51] and toolchains [52] are available.

While building-specific solutions are viable for large commercial buildings with substantial financial investments and retrofits, they are often impractical for small and medium commercial buildings (SMCBs) due to limited human resources and insufficient funds for custom programming of the hardware and software. Typically, these buildings use packaged HVAC systems with thermostats controlled by 7-day setback schedules. Although heuristic controls, such as morning warm-ups, can be implemented, advanced supervisory-level controls that optimize resource operation and manage site-level demand in a predictive manner based on grid conditions require significant expertise and experience [50]. Developing predictive building thermal models is particularly challenging with standard BAS or thermostat data points, as they often lack various unmeasured variables—including internal heat gains and air infiltration. This limitation is considered a major hurdle for model development in predictive control research [53].

Kim et al. (2022) [6] have developed a low-cost MPC solution specifically designed for SMCBs with ON/OFF rooftop units (RTUs). Specifically, this solution addresses key implementation barriers by requiring only internet-connected thermostats, thereby eliminating the need for additional sensor installations of unmeasured disturbances such as internal heat gains. The innovation lies in incorporating unmeasured variables [53], such as occupant-generated heat, directly into the building's thermal model through disturbance modeling. This approach enables accurate performance prediction using only thermostat temperature readings and HVAC operational data.

The following literature review provides the history of the evolution of the low-cost MPC solution and its demonstrations. Kim et al. [54] developed a plug-and-play style building model using an auto-regressive with exogenous input (ARX) model and applied it to the MPC problem to coordinate multiple RTUs, thereby reducing peak energy consumption. During a demonstration in a gymnasium building, they achieved 8% energy reductions and 40% peak demand reductions. Kim and Braun [55] extended this approach by using a building gray-box model for the MPC formulation. Their MPC algorithm controls multiple ON/OFF RTUs in SMCBs to reduce energy costs and peak energy consumption. This approach can be applied to off-the-shelf internet-connected thermostats with packaged RTUs without additional sensors in the system identification of the gray-box model. They conducted a multi-month trial demonstration, resulting in 12% energy savings and 18% peak demand reduction via optimized RTU coordination.

Further advancing this research, they [6] introduced a hierarchical MPC (HMPC) approach for controlling ON/OFF staged packaged units in a laboratory setting. This advanced algorithm, designed to optimize both load shifting and peak demand management, resulted in substantial cost benefits: 30% savings in demand costs while maintaining overall cost savings of 10%. Based on the HMPC's results in the laboratory building, we applied the same algorithm to real K-12 school buildings equipped with six rooftop units (RTUs) [5]. The control solution achieved significant peak demand reductions of 24% during summer operations.

Finally, the technology was further expanded to address different system types by slightly modifying the MPC formulation: a dual-fuel heating system consisting of heat pumps and a gas furnace [56], and a

variable refrigerant flow (VRF) system [57]. For the dual-fuel system, the control strategy was applied to a small commercial building located in NY, U.S., successfully eliminating the need for gas consumption while maintaining occupant comfort. Furthermore, the MPC for the VRF system was applied to an office building located in CA, U.S., for a heating week with a dynamic pricing signal. The results showed that it achieved a 32% peak demand reduction and a 3% energy cost reduction through 24% peak-time load shifting to non-peak time hours.

The field demonstration results provide evidence of the efficacy of the technology, but this does not mean the technology itself is market-ready. Specifically, the adoption readiness level (ARL) framework has been introduced to complement the technology readiness level (TRL) when evaluating deployment potential [58]. ARL encompasses broader considerations, including value proposition, market acceptance, resource availability, and license to operate. In transitioning the research outcomes to the market product, many aspects of ARL for the research outcomes depend on the commercial partner who sells the technology to the market. Specifically, it can be successful when the research-level technology is successfully transferred to the commercial partner, enabling them to develop those strategies based on a deep understanding of the technology, including customization. Furthermore, beyond the technical capabilities, successful market deployment requires careful consideration of multiple stakeholders' needs and perspectives, such as market analysis and operational troubleshooting [59].

1.3. Research gap

Despite the extensive body of MPC research and demonstrations reviewed above, several critical gaps remain that hinder the practical deployment of MPC technology for demand flexibility in the building sector:

- **Scale and deployment gap:** As highlighted by Khabbazi et al. [16], only 4% of nearly 3,000 published papers on advanced building control focus on actual field demonstrations, with most being single-building, short-term pilots. Coordinated MPC demonstrations at the campus scale involving dozens of RTUs in buildings remain extremely rare, and long-term deployments spanning multiple years are virtually absent from the literature.
- **Building type gap:** While large commercial buildings with sophisticated building automation systems have received considerable attention, K-12 schools—representing approximately 100,000 buildings in the U.S. with substantial HVAC modernization needs [10]—have been largely overlooked as potential grid-interactive efficient buildings.
- **Technology transfer gap:** Existing demonstrations have primarily focused on technical performance while neglecting the broader adoption readiness considerations [58], including commercial partnership, technology transfer processes, market deployment strategies, and operational troubleshooting in real-world settings. The path from research prototype to commercial product remains poorly documented.
- **Demand flexibility focus gap:** Most residential MPC demonstrations have emphasized energy cost minimization through load shifting based on price signals, while commercial building demonstrations have prioritized energy efficiency. However, coordinated peak demand reduction across multiple buildings—increasingly critical for grid reliability—has received limited attention, particularly in the SMCB sector.
- **Statistical evaluation gap:** Field demonstrations often face challenges with limited sample sizes and non-normal distributions due to practical constraints (weather variations, operational schedules, utility requirements). However, existing literature lacks robust statistical methodologies specifically designed to quantify MPC performance impacts under these real-world conditions.

This paper addresses these gaps by presenting multi-site, multi-year MPC demonstrations in K-12 schools, documenting the commercial

Table 1
Comparison between previous study and current work.

Aspect	Previous Study [5]	Current Work
Scale and Scope		
Number of sites	1 school campus	3 school campuses
Number of RTUs	6 RTUs	15–42 RTUs per site (95 total)
Demonstration periods	1 initial demonstration	5 demonstration periods
Building sector coverage	Single site	Multiple K-12 schools
Primary Focus		
Main emphasis	Algorithm development and technical details	Multi-site deployment and scalability
Deployment stage	Initial field demonstration	Commercial deployment and technology transfer
Methodology		
Statistical approach	Standard performance metrics	Novel Bayesian methodology for limited samples
Performance analysis	Demonstration results	Statistical quantification with non-normal distributions
Contributions		
Technical contributions	MPC algorithm and implementation	Large-scale validation and practical insights Technology transfer pathway, deployment lessons learned,
Practical aspects	Limited discussion	commercial partnership dynamics
Generalizability	Single-site validation	Multi-site evidence of sustained performance

technology transfer and deployment process, and introducing a novel statistical approach for evaluating demand flexibility performance.

1.4. Contributions

This paper bridges the gap between MPC research and commercial deployment through multi-site demonstrations in K-12 schools. Building on our previous work [5], we expand from 1 to 3 school campuses with 5 demonstration periods, introduce a novel statistical methodology, and document practical deployment considerations. The difference from the previous study is summarized in Table 1.

The main contributions are:

- **Large-scale field demonstration:** MPC deployment at three K-12 school campuses, ranging from 15 to 42 RTUs per site, demonstrating scalability in an understudied but high-potential building sector.
- **Novel statistical methodology:** A Bayesian approach for quantifying MPC impacts on peak demand reduction with limited samples and non-normal distributions.
- **Practical deployment insights:** Evidence of sustained performance across different sites, addressing industry concerns about technology maturity and reliability.
- **Technology transfer pathway:** Documentation of the process from research prototype to commercial product, including partnership dynamics and operational lessons learned—rarely discussed in academic literature but critical for adoption.

1.5. Structure of the paper

The paper first presents a summary of the low-cost MPC technology developed in our previous work. Next, we outline the building and HVAC characteristics of the deployment sites, followed by an explanation of the new statistical approach and the measurement and verification (M&V) process. Finally, we present results from the experimental cases and conclude with lessons learned during the technology transfer and deployment.

2. Hierarchical MPC

2.1. Overview of hierarchical MPC technology

Fig. 1 illustrates a schematic diagram of a low-cost MPC and its associated controlling devices in buildings with disturbances. The MPC is designed for typical SMCB configurations, where one RTU serves a single web-enabled thermostat without additional sensor retrofits, minimizing hardware costs [53]. By integrating the disturbance model into a control-oriented gray-box model, this approach enables the optimization to be formulated as mixed-integer linear programming (MILP). Using standard control-oriented MILP reduces computational costs, making the overall solution a low-cost approach to demand-flexibility control at scale [6]. However, this design is tailored to the specific SMCB configuration and is not applicable to other systems, such as central variable-air-volume configurations. By utilizing future price signals, weather forecasts, and building thermal models, the MPC provides optimal operations for each RTU and sends the corresponding setpoints to each thermostat to reduce electricity peak demand and energy cost via load shifting by coordinating the operations of multiple RTUs at a supervisory level.

Specifically, the MPC has a hierarchical structure (HMPC) [5,6], and it is composed of an upper-level MPC (UMPC) and a lower-level MPC (LMPC)). Fig. 2 illustrates the data flows between components in the hierarchical MPC. Based on the weather forecast, price signals, and building thermal model (gray-box model), the LMPC sends optimal setpoints to the thermostats to control RTUs based on the decision made by the UMPC. The UMPC provides the optimal setpoints for RTUs to minimize energy costs and peak demand while maintaining room temperatures within comfort ranges based on weather forecasts and energy price signals for the relatively long-term prediction horizon (i.e., 24 hours) and control timestep (i.e., 30 min). Electricity prices are high during times of grid congestion, and the UMPC provides demand flexibility by increasing and decreasing energy consumption during low and high price periods (i.e., load shifting), respectively. While the UMPC decides the optimal setpoints for the relatively long-term prediction horizon with a longer control timestep, the LMPC limits the peak demand by reducing the number of simultaneous operations of devices (RTUs) within the UMPC's control timestep with a relatively short-term prediction horizon

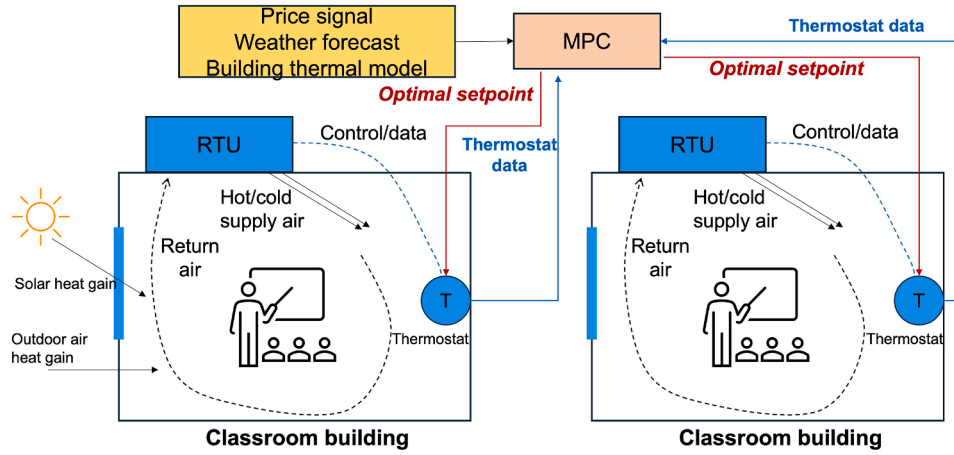


Fig. 1. Schematic diagram of MPC and its controlling devices in buildings with disturbances.

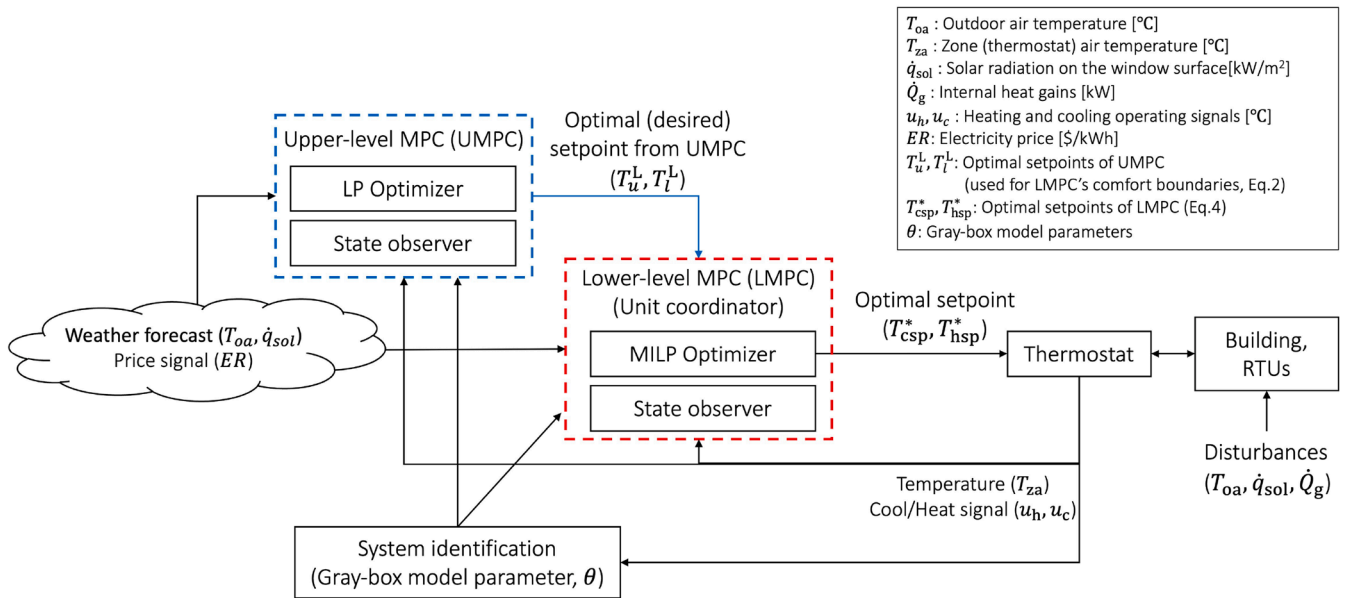


Fig. 2. Schematic diagram of a hierarchical MPC for optimal load shifting and peak demand reduction (see Eqs. (1)–(4) for variables.

(i.e., 15 min) and a shorter control timestep (i.e., 5 min or the minimum device ON/OFF time). A typical single-level MPC would optimize both load shifting and peak demand reduction simultaneously using a single control timestep. While the UMPC includes a demand term to reduce peak demand over that interval, not the actual number of simultaneously operating devices. This can result in multiple RTUs activating at the same time within the 30 min window, creating short-term peaks. Our two-stage hierarchical approach addresses this limitation by decoupling the objectives: the UMPC handles long-term load shifting (24-hour horizon with 30 min timesteps) based on price signals through continuous optimization, while the LMPC enforces short-term operational coordination (15 min horizon with 5 min timesteps) to prevent simultaneous RTU activation through discrete binary optimization with hard constraints on the number of operating devices. This separation allows the system to guarantee peak demand limits through the LMPC's device-level binary controls while still achieving cost-effective load shifting through the UMPC.

Fig. 3 illustrates the operation of the HMPC using a conceptual diagram as an example. The UMPC's optimal cooling strategy, represented by the dashed line indicating the desired temperature, pre-cools the

room below typical comfort boundaries (i.e., the cooling setpoint of non-MPC operation) before peak pricing periods (Fig. 3 (a)). This strategy reduces energy consumption during peak pricing times by allowing the temperature to gradually rise without using cooling energy until the effect of the pre-cooling dissipates. Subsequently, the LMPC determines the scheduling of RTU usage to minimize simultaneous operations while keeping the room temperature within the UMPC's desired range. For instance, when two RTUs have overlapping ON/OFF cycles, the LMPC staggers these cycles. This method provides the same cooling energy during the LMPC's control timestep while reducing peak power demand, as shown in Fig. 3 (b).

To preserve occupant thermal comfort while achieving demand flexibility, the HMPC incorporates an override mechanism that allows occupants to manually adjust thermostat setpoints or schedules at any time. When an override is detected, MPC control of the affected unit is suspended for 3 hours. During this period, the overridden unit's operation is incorporated as a constraint in the optimization, allowing the system to calculate optimal control strategies for the remaining units that accommodate the override while still pursuing the original demand management objectives. This ensures individual comfort preferences take precedence while maintaining coordinated control of the remaining units.

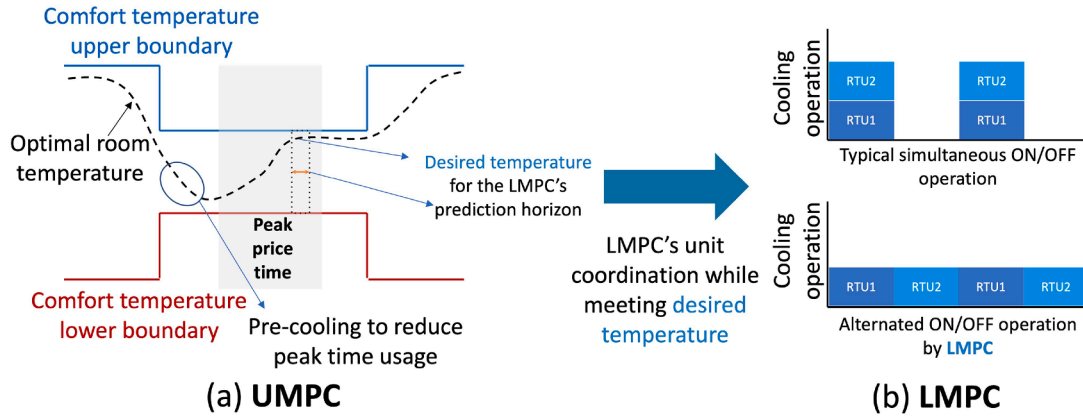


Fig. 3. Conceptual diagram of the MPC for load shifting (UMPC) and peak demand reduction (LMPC) (adopted from [5]).

Technical implementation details are provided in [5], and the rationale for this temperature-based comfort approach is discussed in Section 6.4.

2.2. Problem formulation

2.2.1. Optimization problem

In this section, we present a summary of the mathematical framework of the HMPC. The comprehensive mathematical details can be found in the following articles: system identification methods in [5,53] and HMPC formulation in Ham et al. [5], Kim and Braun [6].

At each UMPC's sampling interval, the UMPC provides the optimal runtime fraction (RTF) trajectories of RTUs ($\bar{u}_{hc,i}(k+j-1)$) by solving the optimization problem in Eq. (1). Specifically, the RTF represents the fraction of time each RTU operates within the control timestep on a 0-1 scale (e.g., RTF = 0.5 for a 30 min timestep indicates 15 min ON and 15 min OFF across all cycling intervals). The objective function is to minimize the total energy cost for each RTU ($ER(k+j-1)P_{RTU,i}(k+j-1)\bar{u}_{hc,i}^U(k+j-1)$) for all RTUs ($i = 1 \dots n$), over the prediction time horizon ($j = 1 \dots N_p^U$) while including peak power in the prediction horizon (δ) and comfort violations ($\Gamma_{l,i}$ and $\Gamma_{u,i}$) as soft constraints. The optimal RTF refers to the RTFs of each RTU for each timestep that minimize the objective function.

$$\min_{\substack{\bar{u}_{hc,i}^U(k+j-1) \\ \delta, \Gamma_l, \Gamma_u}} \sum_{j=1}^{N_p^U} \sum_{i=1}^n ER(k+j-1)P_{RTU,i}(k+j-1)\bar{u}_{hc,i}^U(k+j-1) + \omega_d \delta + \omega_l \Gamma_l + \omega_u \Gamma_u \quad (1)$$

$$s.t. \quad T_{l,i}^U(k+j) - \Gamma_l \leq \mathbb{E}(\bar{y}_i(k+j)|D_k) \leq T_{u,i}^U(k+j) + \Gamma_u \quad (\forall i \in \{1, \dots, n\})$$

$$\sum_{i=1}^n P_{RTU,i}(k+j-1)\bar{u}_{hc,i}^U(k+j-1) \leq \delta$$

$$0 \leq \bar{u}_{hc,i}^U(k+j-1) \leq 1 \quad (\forall j \in \{1, \dots, N_p^U\}),$$

where n is the number of RTUs, N_p^U is the prediction horizon of the UMPC, i is the index for the i^{th} RTU (and corresponding thermostat), j is the control timestep from the current timestep k , $\bar{(\cdot)}$ denotes control timestep moving-averaged variables, $\bar{y}_i(k+j)$ is the zone thermostat temperature for the i^{th} RTU at the timestep $k+j$, ER is the electricity price (\$/kWh), $P_{RTU,i}$ is the electric power use of the i^{th} RTU at full-speed, $\bar{u}_{hc,i}^U(k+j)$ is the moving-averaged heating or cooling RTF (normalized 0-1 scale) for the i^{th} unit at the $(k+j)^{\text{th}}$ timestep of the UMPC, $\mathbb{E}(\bar{y}_i(k+j)|D_k)$ is the j -step temperature prediction via the building model (see the following Section 2.2.2) given the data and optimal decisions ($D_k = \{y(k-1), y(k-2), \dots, u_{hc}^U(k+j-1), u_{hc}^U(k+j-2), \dots\}$) using a Kalman filter, $T_{l,i}^U(k+j)$, $T_{u,i}^U(k+j)$ are the desired (average) comfortable temperature ranges (i.e., the default heating setpoint

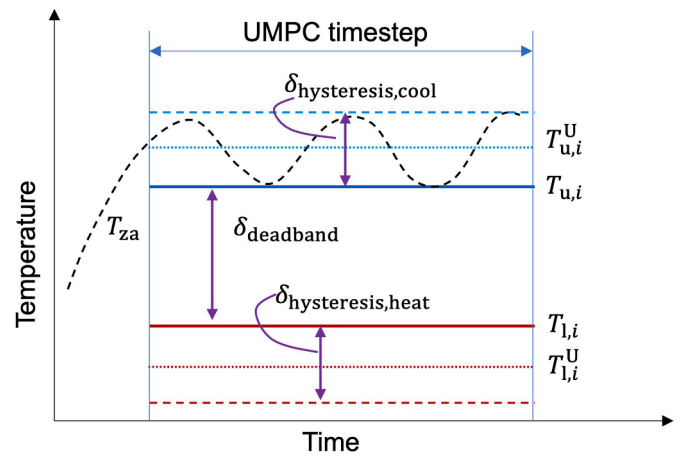


Fig. 4. Desired temperature ($T_{u,i}^U$ and $T_{l,i}^U$) as comfort boundaries for UMPC formulation considering the hysteresis ($\delta_{\text{hysteresis,cool}}$ and $\delta_{\text{hysteresis,heat}}$) and setpoints ($T_{u,i}$ and $T_{l,i}$) during the UMPC timestep due to ON/OFF cycling operation.

minus half the heating hysteresis ($T_{l,i}(k+j) - \Delta_{\text{heat,hysteresis}}/2$) or the default cooling setpoint plus half the cooling hysteresis ($T_{u,i}(k+j) + \Delta_{\text{cool,hysteresis}}/2$) during the UMPC's timestep as shown in Fig. 4. This adjustment is necessary because the temperature fluctuates within the hysteresis band through ON/OFF operation, resulting in the average temperature of $T_{l,i}^U(k+j)$ and $T_{u,i}^U(k+j)$, respectively) for the i^{th} RTU at time $k+j$ (in other words, the resulting average temperature for the timestep is $T_{l,i}^U(k+j)$ or $T_{u,i}^U(k+j)$ when we denote the heating or cooling setpoint as $T_{l,i}(k+j)$ or $T_{u,i}(k+j)$ for continuous heating or cooling operation.), $\omega_l, \omega_u \in \mathbb{R}^+$ and $\omega_d \in \mathbb{R}^+$ are weights for the variables of $\Gamma_l, \Gamma_u \in \mathbb{R}^+$ and $\delta \in \mathbb{R}^+$, $\delta \in \mathbb{R}^+$ is the peak power demand slack variable over the prediction horizon, Γ_l and Γ_u are comfort violation variables. The optimizer finds the optimal sequence of RTF for RTUs over the prediction horizon, δ , Γ_l , and Γ_u , which forms a linear programming problem.

During the optimization, the predicted temperatures ($\mathbb{E}(\bar{y}_i^*(k+j)|D_k)$) are obtained based on the optimal RTF trajectories, and the predictions for the next timestep (i.e., the first values of the predicted temperatures, $\bar{y}_{1,n}^*(k+1)$) are used as the desired temperature setpoints. The LMPC, presented in the following section (Eq. (3)), will use the desired setpoints for its temperature boundaries after conversion via Eq. (2).

$$T_{u,i}^L(k) = \begin{cases} \bar{y}_i^*(k+1) + \Delta_{\text{cool,hysteresis}}/2, & \text{if } \bar{u}_{c,i}^U(k) > 0 \\ T_{u,i}(k) + \Delta_{\text{cool,hysteresis}}/2, & \text{if } \bar{u}_{c,i}^U(k) = 0 \end{cases} \quad (2)$$

$$T_{l,i}^L(k) = \begin{cases} \bar{y}_i^*(k+1) - \Delta_{\text{heat,hysteresis}}/2, & \text{if } \bar{u}_{h,i}^{U*}(k) > 0 \\ T_{l,i}(k) - \Delta_{\text{heat,hysteresis}}/2, & \text{if } \bar{u}_{h,i}^{U*}(k) = 0 \end{cases}$$

where $\bar{u}_{h,i}^{U*}(k)$ and $\bar{u}_{c,i}^{U*}(k)$ are the optimal heating and cooling RTFs of the i^{th} unit at the k^{th} timestep of the UMPC, $\bar{y}_i^*(k+1)$ is the predicted indoor temperature with $\bar{u}_{h,i}^{U*}(k)$ and $\bar{u}_{c,i}^{U*}(k)$, and $\Delta_{\text{heat,hysteresis}}$ and $\Delta_{\text{cool,hysteresis}}$ are the hysteresis values for heating and cooling operations (i.e., temperature fluctuations due to the ON/OFF operations based on the thermostat setpoints). When the optimal RTFs are zero, the boundaries of the UMPC (i.e., default setpoint schedules) are used for the LMPC. Otherwise, the desired temperature profiles are used as the upper or lower bounds of the LMPC for cooling or heating operation, respectively.

The optimization problem of the LMPC can be written as Eq. (3). The main purpose of the LMPC is to coordinate the operations of RTUs for a relatively short-term prediction horizon (i.e., 15 min) to limit the total number of ON RTUs (i.e., $\sum_{i=1}^n P_{\text{RTU},i}(k+j-1)u_{h,c,i}(k+j-1) \leq \delta$ ($\forall j \in \{1, \dots, N_p^L\}$)) while maintaining the temperature within the desired range presented in Eq. (2). This optimization problem determines the optimal ON/OFF sequences for the RTU operations (i.e., binary signals), which is the main difference compared to the UMPC, which uses RTFs as optimization variables.

$$\begin{aligned} \min_{\substack{u_{h,c,i}^{L*}(k+j-1) \\ \delta, \Gamma_l, \Gamma_u}} \quad & \sum_{j=1}^n \sum_{i=1}^n P_{\text{RTU},i}(k+j-1)ER(k+j-1)u_{h,c,i}^L(k+j-1) + \omega_d \delta + \omega_l \Gamma_l + \omega_u \Gamma_u \quad (3) \\ \text{s.t.} \quad & T_{l,i}^L(k+j) - \Gamma_l \leq \mathbb{E}(y_i(k+j)|D_k) \leq T_{u,i}^L(k+j) + \Gamma_u \quad (\forall i \in \{1, \dots, n\}) \\ & \sum_{i=1}^n P_{\text{RTU},i}u_{h,c,i}^L(k+j-1) \leq \delta \quad (\forall j \in \{1, \dots, N_p^L\}), \end{aligned}$$

where n is the number of RTUs, N_p^L is the prediction horizon of the LMPC, i is the index for the i^{th} RTU (and corresponding thermostat), j is the time step for control changes from the current measured value of k , ER is the electricity price (\$/kWh), $P_{\text{RTU},i}$ is the electric power use of the i^{th} RTU at full-speed, $u_{h,i}^L(k+j)$ and $u_{c,i}^L(k+j)$ are the RTU unit ON/OFF status of heating and cooling operations (binary) for the i^{th} unit at the $(k+j)^{\text{th}}$ timestep, $y_i(k+j)$ is the thermostat temperature of the i^{th} room at the $k+j^{\text{th}}$ timestep, $\mathbb{E}(y_i(k+j)|D_k)$ is the optimal j -step temperature prediction from the building model (Section 2.2.2) given the data $D_k = \{y(k-1), y(k-2), \dots, u_{h,c}^L(k+j-1), u_{h,c}^L(k+j-2), \dots\}$ via a Kalman filter, $(T_{l,i}^L(k+j), T_{u,i}^L(k+j))$ are the desired setpoints obtained from UMPC for the i^{th} RTU at time $k+j$, $\omega_l, \omega_u (\in \mathbb{R}^+)$ and $\omega_d (\in \mathbb{R}^+)$ are weights for the variables of $\Gamma_l, \Gamma_u (\in \mathbb{R}^+)$ and $\delta (\in \mathbb{R}^+)$, $\delta (\in \mathbb{R}^+)$ is the peak power demand slack variable over the prediction horizon, Γ_l and Γ_u are comfort violation variables. The optimizer finds the optimal sequence of RTU ON/OFF statuses over the prediction horizon, δ , Γ_l , and Γ_u , which forms a mixed integer linear programming problem.

To effectively manage peak demand reduction in the LMPC framework, a constraint is implemented that limits the maximum electrical demand (represented by δ) throughout the prediction horizon. Our previous work [55] presents comprehensive details about this approach along with extensive experimental validation results.

There are three tuning hyperparameters in the MPC: (ω_d , ω_l and ω_u). The weights for comfort violations (ω_l and ω_u) are set to 1000, which is sufficiently large to prioritize thermal comfort as a soft constraint. The tuning of the weight for demand (ω_d) is more subjective and can be determined based on the demand reduction goal (δ). For example, in this research, we set the upper boundary for δ at 70% of the combined RTU power capacity ($0.7 \times \sum_{i=1}^n P_{\text{RTU},i}$), based on our previous field experiment [6] as an achievable target. Although this limits δ , setting ω_d too high could result in excessive pre-heating or pre-cooling (i.e., excessive overall energy increase for load shifting) with a reduction of δ more than desired. Therefore, ω_d was empirically chosen by monitoring the initial experimental results, and setting it to 10 has produced

reasonably acceptable results so far. Furthermore, to comply with standard RTU anti-cycling requirements and prevent compressor damage, the LMPC control actions are executed at 5 min intervals, ensuring adequate minimum off-time periods.

The optimization problems for UMPC and LMPC are linear and mixed-integer linear programming problems, respectively, and solved via an open-source mixed-integer linear programming solver [60]. The control sequence begins with UMPC calculations at 30 min intervals, followed by LMPC optimizations every 5 min. The LMPC operates within thermal comfort boundaries (Eq. (2)) established by the UMPC, which remain the same until the next UMPC update. The RTU operation is controlled through thermostat setpoint adjustments based on the LMPC's optimal control signals. For the thermostat control, we implement optimal setpoints by adjusting the setpoints with a specific offset (i.e., $\Delta_{\text{LMPCOffset}} = 3^\circ\text{C}$) to turn ON or OFF as shown in Eq. (4). For example, during the cooling operation, we set the temperature higher than the current zone temperature to turn the RTU OFF, and lower to turn it ON. The opposite rule applies to the heating operation.

$$\begin{aligned} T_{\text{csp},i}^*(k) &= \begin{cases} y_i(k) - \Delta_{\text{LMPCOffset}}, & \text{if } u_{c,i}^{L*}(k) > 0 \\ y_i(k) + \Delta_{\text{LMPCOffset}}, & \text{if } u_{c,i}^{L*}(k) = 0 \end{cases} \quad (4) \\ T_{\text{hsp},i}^*(k) &= \begin{cases} y_i(k) + \Delta_{\text{LMPCOffset}}, & \text{if } u_{h,i}^{L*}(k) > 0 \\ y_i(k) - \Delta_{\text{LMPCOffset}}, & \text{if } u_{h,i}^{L*}(k) = 0 \end{cases} \end{aligned}$$

where $u_{c,i}^{L*}(k)$ and $u_{h,i}^{L*}(k)$ are the optimal cooling and heating binary decisions of the i^{th} unit at the k^{th} timestep of the LMPC, $y_i^*(k)$ is the current indoor temperature of the i^{th} unit at the k^{th} timestep, $T_{\text{csp},i}^*(k)$ and $T_{\text{hsp},i}^*(k)$ are the LMPC optimal cooling and heating setpoints, respectively, to the i^{th} unit thermostat.

2.2.2. Building model

Gray-box modeling is widely adopted in building control applications to predict indoor air temperatures using control signals and measured disturbances such as weather conditions. By balancing the trade-off between black-box and white-box models, this approach offers several advantages. While its parameters have physical interpretations derived from prior knowledge and fundamental principles of energy conservation, the parameter estimation process remains data-driven.

In the context of SMCBs, system identification (SYSID) faces critical challenges. The primary obstacle is the limited availability of sensors to measure various heat sources, such as occupant heat generation, equipment usage, and air exchange through infiltration/exfiltration. Previous studies by [53,61] have shown that these unmeasured thermal loads significantly affect SYSID accuracy, particularly when they correlate with control variables or measured environmental parameters, regardless of the chosen modeling approach and SYSID algorithm.

Although additional sensor deployment could resolve these measurement limitations, such solutions are typically not economically viable for SMCBs. Our approach addresses this challenge by implementing the lumped disturbance (LD) method, which has demonstrated effectiveness in handling significant unmeasured thermal inputs during system identification. The LD approach incorporates all unmeasured heat gains as a single lumped disturbance term within a discretized state-space model framework. This term is modeled as a filtered white noise process, enabling the model to account for unmeasured thermal loads such as occupancy, equipment usage, and infiltration without requiring dedicated sensors. The resulting state-space model maps control inputs (RTU run-time fractions) and measured disturbances (outdoor temperature, solar radiation) to predicted zone temperatures. Model parameters are identified through the prediction error method, with technical details of the state-space formulation, disturbance modeling, and parameter identification algorithms are provided in [53,61].

For the state-space model of the LD approach, a 2R2C gray-box model structure was selected to model a classroom with a single thermostat [5]. For open spaces containing multiple thermostats, a multi-zone

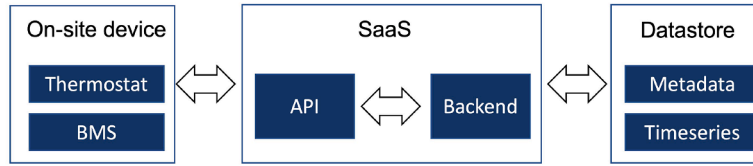


Fig. 5. Overview of dataflow in software as a service (SaaS).

model (Fig. 10 in [5]) is typically generated by augmenting the system matrix as described in Kim et al. [61]. However, in this study, classrooms were either physically detached from one another or not strongly thermally coupled; consequently, all spaces were treated as single-zone models. The gray-box model maps the inputs [T_{oa} (outdoor air temperature), $\dot{q}_{sol}$ (solar radiation on the window surface), $\dot{u}_h$ (heating RTF), $\dot{u}_c$ (cooling RTF)] to the output [T_{za} (thermostat air temperature)], and data from three weekends—utilizing system perturbations based on a PRBS signal—were used for model training [62,63]. To model RTU power consumption, we assume a constant COP, making power a function of the runtime fraction (RTF), denoted as $\bar{u}_{hc}$, where power is calculated as the rated RTU power (P_{RTU}) multiplied by RTF (i.e., $P_{RTU} \times \bar{u}_{hc}$). This simplified assumption is reasonable for our demonstration sites on the U.S. West Coast, where mild weather results in relatively small COP variations within the prediction horizon. While this introduces some modeling error compared to a weather-dependent COP model, it enables practical deployment by eliminating the need to collect RTU power data under various outdoor air conditions. However, if weather significantly impacts power consumption (e.g., due to an electrical auxiliary heater), it is advisable to model power as a function of outdoor air temperature or treat the auxiliary heating mode as a separate operation, as described in Ham et al. [56]. Due to limited space and the large number of models (approximately 95), we omit the system identification results in this study, however, all the models were trained using data containing PRBS signals and validated by comparing measured and estimated cooling and heating runtime fractions as described in Fig. 13 in Ham et al. [5]. Further details on system identification and training procedures are provided in Ham et al. [5].

3. Technology transfer and field demonstration sites

3.1. Implementation via technology transfer

The MPC solution is implemented via a three-step procedure: gathering building metadata and operational signals, performing system identification and model validation, and finally deploying the control strategy for performance assessment [5]. Once the MPC is ready for deployment, it is distributed as a Software as a Service (SaaS); the data flow architecture is shown in Fig. 5. On-site devices such as internet-connected thermostats or Building Management Systems (BMS) continuously communicate with the SaaS platform via bidirectional application programming interfaces (APIs). In the data collection phase, the SaaS API retrieves operational data (e.g., zone temperatures, heating/cooling signals) from on-site devices and stores them in the datastore, which maintains both time-series data and metadata. When the MPC service is activated, the SaaS backend reads the necessary historical and real-time data from the datastore, executes the optimization algorithm to generate optimal setpoints, and subsequently writes these control signals back to the on-site devices via the same API. This bidirectional communication enables continuous monitoring and real-time control adjustment.

The successful market adoption of our technology depends on effectively transferring the complete implementation process (research outcomes) to our commercial partner. While the software is designed for automatic configuration when input data follows specified formats, implementation may require modifications or additional features for non-standard scenarios. To facilitate this technology transfer, we have

Table 2

Summary of demonstration cases.

Case	Site	Tariff	# Devices	Season	Period	# Days (BS/MPC)	T_{oa} Range [°C]
1	1	Flat, DC	38 RTUs	Heating	Dec, 2023–Feb, 2024	14/17	- 5.9–11.7
2	1	Flat, DC	38 RTUs	Cooling	Aug–Oct, 2023	22/19	11.7–26.4
3	2	Flat, DC	15 RTUs	Heating	Jan–Feb, 2024, 2024	6/5	3.9–11.5
4	2	Flat, DC	15 RTUs	Cooling	Sep–Oct, 2023	11/9	6.7–18.3
5	3	ToU, DC	42 RTUs	Heating	Jan–Feb, 2024	4/4	11.3–16.6

DC: Demand Charge. Site 1: High school in WA; Site 2: Elementary school in WA; Site 3: High school in CA. T_{oa} : daily average outdoor air temperature.

adopted a collaborative learning approach. During initial implementation, we provide comprehensive guidance to the commercial partner through video conferences, explaining each parameter of the low-cost MPC setup in detail. The commercial partner then documents this process for the internal scaling. When encountering special cases, we develop additional parameters or case studies and provide guidance on necessary source code modifications through video conferences. This approach empowers the commercial partner to independently adapt and implement the technology for future applications.

3.2. Field demonstration sites

The MPC has been implemented in three K-12 school campuses (Fig. 6) in California (CA) and Washington (WA), U.S. for cooling and heating seasons. Based on the site and season, we classified the results as 5 cases as shown in Table 2. More details about the characteristics of each site are described in the following sections.

Site 1 is a high school located in WA, U.S. The MPC system is implemented for 38 heat pumps (HPs) serving 38 classrooms, with a nominal total power of 285 kW for heating or cooling operations and 380 kW for electric auxiliary heating. The temperature setpoints are maintained at 18.9–23.3°C (66–74°F) during occupied hours (6:00 AM to 4:00 PM, Monday to Friday) and 16.7–29.4°C (62–85°F) during unoccupied hours. This school operates under a fixed-rate pricing structure for electrical usage, which includes a demand charge; therefore, the primary focus of the MPC is to reduce peak demand and minimize these charges.

Site 2 is an elementary school located in WA, U.S. This site features 15 heat pumps serving 15 classrooms across two buildings. The total nominal power of all heat pumps is 150 kW for both heating and cooling operations. The heat pumps operate on an occupied schedule from 6:00 AM to 5:00 PM, Monday through Friday, with temperature setpoints ranging from 18.9 to 23.3°C (66 to 74°F). During unoccupied hours, the setpoints are adjusted to 16.7–29.4°C (62–85°F). Similar to Site 1, this site also operates under a fixed-rate pricing structure with a demand charge, making the primary focus of the MPC the reduction of peak demand.

Site 3 is a high school located in CA, U.S. The HVAC system includes 45 heat pumps serving 45 classrooms, with a total nominal power of 585 kW for heating and cooling operations. The occupancy schedule is from 6:00 AM to 4:00 PM, Monday through Friday, with pre-occupied hours from 5:00 AM to 6:00 AM and 4:00 PM to 5:00 PM. The temperature setpoints are maintained at 20.0–22.2 °C (68–72°F) during occupied hours,

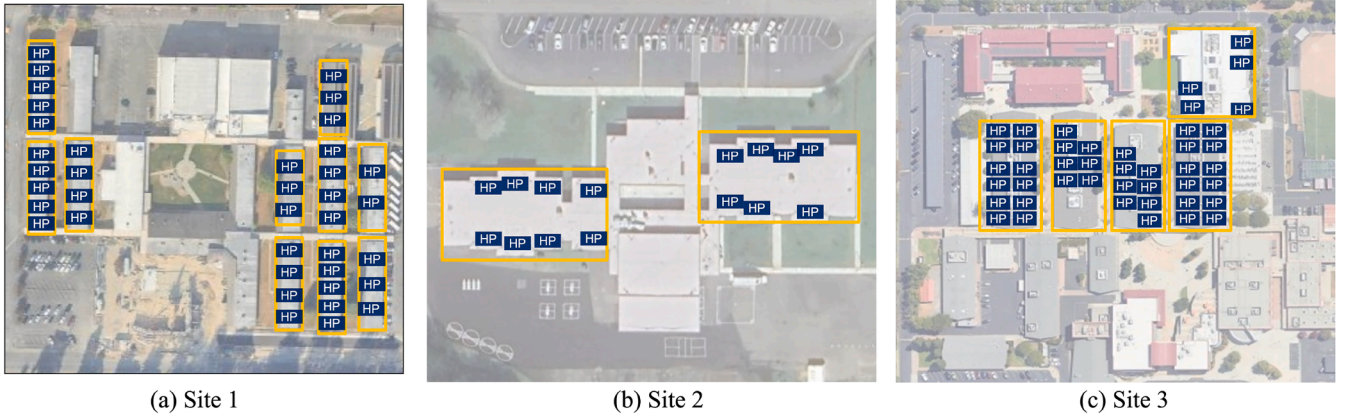


Fig. 6. Demonstration buildings for each site.

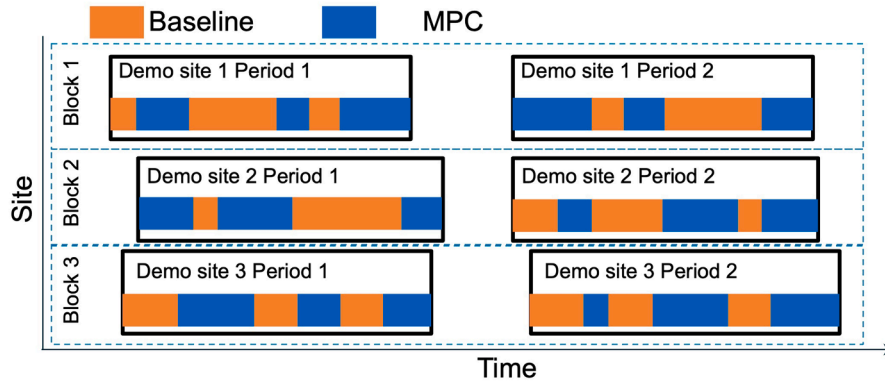


Fig. 7. Conceptual diagram of the randomized block schedule (Adopted from [5]).

while they are adjusted to 16.7–29.4°C (62–85°F) during pre-occupied and unoccupied hours. This site operates under a Time-of-Use rate structure, which includes a demand charge. The peak pricing period is from 8:30 AM to 9:30 PM, Monday through Friday. Given the similarity between peak and non-peak rates and the extended duration of the peak pricing period, the MPC has a primary focus on reducing peak demand.

3.3. Design of experiment

We applied a specific experimental design for performance evaluation that differs from that one used for the traditional measurement and verification (M&V), which typically relies on comparing pre- and post-installation periods. Our methodology incorporates a randomized block design that alternates between Baseline (BS) and MPC operations on a daily basis. To achieve a statistically rigorous and balanced comparison, we implemented blocking [64] to ensure equal distribution of operational days between the two control strategies within designated block periods (see Fig. 7).

Based on the experimental schedule, either BS or MPC control is enabled each day. BS represents the original weekly thermostat setpoint schedule programmed by facility operators according to the school district's policy. As described in Section 3.2, these setpoint schedules are already programmed into the thermostats, so no additional action is required on BS days. On MPC days, the MPC system is activated and overrides the programmed setpoint schedules by sending optimal setpoints to the thermostats.

As shown in Table 2, the number of experimental days varies across cases, with Cases 3 and 5 having notably fewer BS and MPC days (6/5 and 4/4 days, respectively) compared to Cases 1, 2, and 4. This limitation arose from several field deployment challenges encountered during the experiments. Specifically, interruptions due to third-party API ser-

vice instability (particularly weather data/forecast services), code execution inefficiencies that exceeded the designated time threshold as the number of HPs increased, and experimental suspensions requested by schools during severe weather alerts or special events resulted in the loss of experimental days that could not be classified as either BS or MPC days (see Section 6.3 for detailed discussion). Despite these constraints, our Bayesian statistical framework (Section 4) provides robust effect estimation by incorporating uncertainty quantification and leveraging the hierarchical peak demand model, enabling reliable inference even with limited sample sizes.

4. Measurement and verification

This section presents the measurement and verification results for the field demonstrations, based on our performance metrics and statistical measures. Since the primary focus of this study is to compare overall performance across multiple sites, a detailed analysis of the system identification results and a day-by-day comparison of load profiles are omitted due to space limitations. Readers may refer to our previous study for a more comprehensive analysis [5].

4.1. Performance metrics

The performance of MPC over the Baseline is quantified using two performance metrics: daily peak demand and daily electricity cost. Since both the HVAC operational schedule and weather profiles exhibit daily patterns, it is common practice to characterize these metrics on a daily basis. However, as weather conditions vary each day, we developed a model that normalizes the performance metrics with respect to weather conditions, allowing us to quantify the effect of MPC versus the baseline under different weather scenarios (see Section 4.2). Thermal comfort is

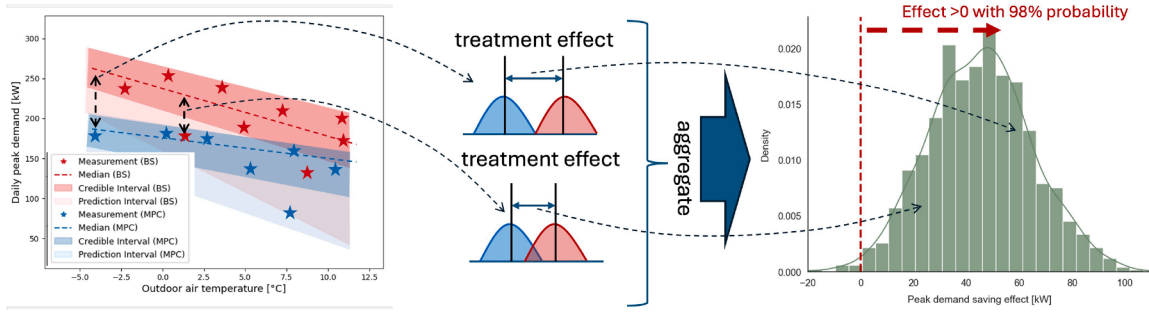


Fig. 8. Conceptual diagram of treatment effect estimation with uncertainties from the model (the numbers in this Figure is all artificial to explain the concept effect estimation).

not included among our performance metrics because the MPC's thermal comfort boundary is defined by the thermostat's default schedule. The MPC will revert to these default setpoints when there is no heating or cooling requirement or in case of any failures. In other words, both the Baseline and MPC are expected to maintain room temperatures within the default schedule boundaries.

Typically, utility demand charges are based on the highest electricity demand recorded within a 15- or 30 min window for the billing unit (e.g., a school campus) over a month. Given that electricity usage follows daily patterns, the day with the highest daily peak demand can represent the monthly peak demand for the demand charge. The daily peak demand is estimated by identifying the maximum value of the daily electricity demand (in kW) using a 15 min moving average. Additionally, the daily electricity cost is defined as the sum of electricity consumption multiplied by the electricity rate (i.e., $ER \times P$ in Eqs. (1) and (3)).

4.2. Effect estimation model for MPC vs. baseline

Quantifying the performance difference between MPC and Baseline in field experiments is challenging because varying weather conditions confound direct comparisons between experimental periods. Simply comparing raw performance metrics can lead to misleading conclusions if weather conditions differ substantially between MPC and BS days. Additionally, the limited experimental duration results in small sample sizes that make traditional statistical approaches unreliable. To address these challenges, we developed an effect estimation model that normalizes performance metrics with respect to weather conditions. The model employs a Bayesian framework to quantify the MPC effect and its associated uncertainty.

4.2.1. Concept of effect estimation

We developed a new approach to estimate the intervention effect of MPC over BS based on experimental data using a Bayesian framework. The conceptual diagram of this approach is presented in Fig. 8. The first step is to build a model that explains the measured target variable (i.e., daily peak demand or energy cost). Fig. 8 (left) displays daily peak demands versus outdoor air temperatures with the developed model as an example. The data points are marked as stars (BS: red and MPC: blue), while the dotted lines represent the estimated models and shaded areas represent the 95% credible intervals (uncertainty ranges) obtained via Bayesian inference [65]. In the Bayesian approach, the predictive distribution for each data point is represented by a large number of posterior samples generated via the Markov Chain Monte Carlo (MCMC) method. For example, when there are 15-day experimental data with 10000 samples, the output's dimension is [10000x15] (number of samples x number of data points).

To estimate the intervention effect with uncertainty, we use different approaches for the peak demand and energy cost. Energy cost is billed by usage, so the difference between the average values of MPC and BS

for each posterior sample shows the average daily cost savings. By aggregating these values for all samples, we can estimate the average daily cost reduction with uncertainty.

On the other hand, peak demand is charged based on the highest monthly peak demand. Assuming the experimental data periods can represent a typical summer or winter season month, the difference between the maximum values predicted across all data points for each posterior sample can be viewed as the peak demand reduction effect for that specific instance. Collecting these differences across all samples generates the posterior distribution of the peak demand reduction effect. Specifically, we calculate the saving effect for each individual sample drawn from the posterior model. The collection of these individual effects forms the histogram in the Fig. 8 (right), where the spread represents the full range of probable outcomes.

Fig. 8 (right) shows the distribution of the aggregated intervention effect. Since the effect represents the difference between BS and MPC, values greater than 0 indicate MPC's peak demand or cost reduction compared to BS. Conversely, values less than 0 indicate that MPC shows higher peak demand or energy cost than BS. The aggregated distribution, from which the credible interval is derived, provides the range of the intervention effect while accounting for uncertainties. While traditional hypothesis testing only indicates whether the difference is statistically significant, the Bayesian framework directly provides the effect size in absolute numbers with associated probabilities. For example, the proportion of posterior samples greater than 0 indicates the probability of superiority—i.e., the certainty that MPC reduces peak demands.

4.2.2. Peak demand model

Fig. 9 illustrates the model structure for the daily peak demand of BS and MPC. We developed the model in a Bayesian framework to encode the unique characteristic of the peak demand and make the inference robust when there is limited data.

The core idea of this model is straightforward: we want to predict daily peak demand based on outdoor temperature while capturing two key physical phenomena. First, peak demand increases as outdoor temperature becomes more extreme. When it's very cold or very hot outside, the heating or cooling system has to run longer to maintain comfortable indoor temperatures. In both cases, more RTUs are running simultaneously, resulting in higher peak demand. Second, peak demand is more variable during mild weather. When outdoor conditions are moderate, individual RTUs turn on and off intermittently based on the specific needs of each room. This makes it harder to predict exactly when peak demand will occur. In contrast, during extreme weather, most RTUs run continuously to meet the large heating or cooling load, which makes peak demand more consistent and predictable.

Fundamentally, the daily peak demand is modeled as a linear function of outdoor temperature. Both the daily peak demand and outdoor air temperature are normalized to a range from -1 to 1 using min-max normalization [66]. Here, y_i and x_i represent the min-max normalized daily peak demand and outdoor air temperature, respectively. The daily

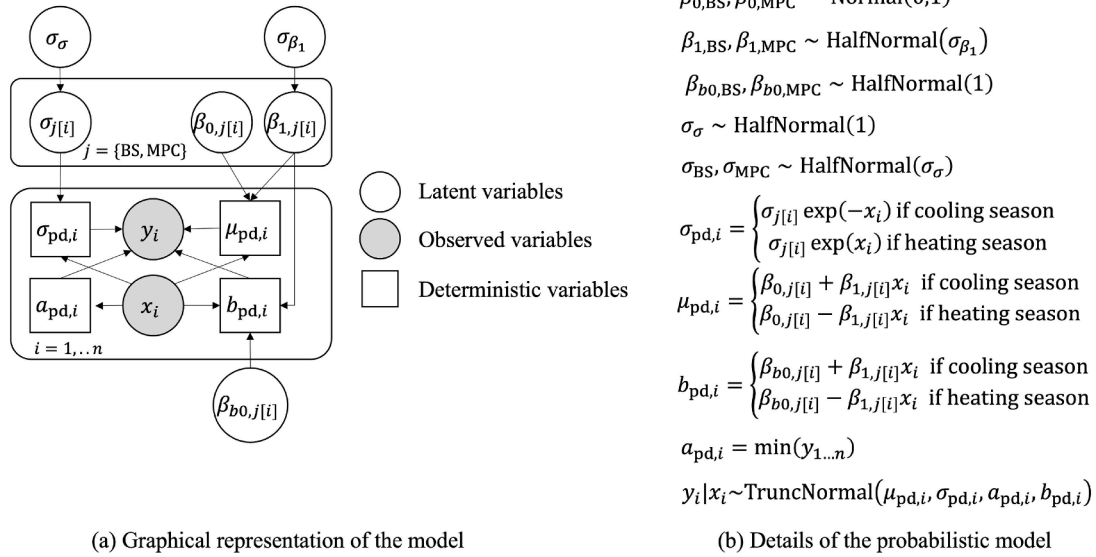


Fig. 9. Graphical representation (a) and probabilistic model structure (b) for a peak demand estimation model.

peak demands of BS and MPC are modeled using two linear regression lines. However, daily peak demands exhibit unique characteristics. The theoretical maximum daily peak demand occurs when all the RTUs operate simultaneously. Typically, the daily peak demand correlates with outdoor temperature (i.e., heating and cooling loads) because it is related to the total number of RTUs that are ON, with more operations occurring as cooling or heating loads increase, leading to a higher likelihood of simultaneous operation. While typical daily peak demands are correlated with outdoor air temperature, the variations in daily peak demands for different days with the same cooling or heating load can increase when cooling or heating loads decrease. This is because the overlap of the RTU ON operations are more stochastic when the heating or cooling loads decrease (i.e., sparser ON/OFF sequences).

To model this behavior, we first represent the uncertainty ranges of the regression line, bounded by upper ($b_{pd,i}$) and lower ($a_{pd,i}$) limits, using a truncated normal distribution [67] (i.e., $a_{pd,i} \leq \mu_{pd,i} \leq b_{pd,i}$). Intuitively, this truncated normal distribution ensures that predicted peak demands stay within physically realistic bounds—they cannot exceed the maximum possible demand (all RTUs on) or fall below the minimum observed demand, while most predictions cluster around the mean trend line. Here, i indicates the data index from 1 to n , while $j[i]$ denotes either BS or MPC for the i th data point. The mean value ($\mu_{pd,i}$) captures the general trends of daily peak demand versus outdoor air temperature, and the upper bound ($b_{pd,i}$) represents the potential maximum daily peak demands. In this study, we assume that the truncated normal distribution is negatively skewed, with a concentration at the mean value and a tail extending toward zero. This assumption is based on observed data, which shows that daily peak demands exhibit similar values under the same weather conditions when there are sufficient heating or cooling loads, with some variations in a negative direction, particularly when heating or cooling loads are low. While the theoretical maximum daily peak demand is the sum of all RTU powers, this scenario rarely occurs unless all RTUs are activated due to high cooling and heating loads. For this reason, we modeled the upper bound ($b_{pd,i}$) as a regression line with the same mean value ($\mu_{pd,i}$) but with a different intercept to provide some variation near the mean value (i.e., the same slope parameter $\beta_{1,j[i]}$). The lower bound ($a_{pd,i}$) is set to the minimum value of all the measured $y_{i,n}$.

One important characteristic is that variations increase when cooling or heating loads decrease because the simultaneous operation of the RTUs is less likely to occur. For this reason, the standard deviation

parameter of the truncated normal distribution ($\sigma_{pd,i}$) is modeled as a function of the outdoor air temperature using an exponential function to ensure positive values.

Finally, we designed the model in a hierarchical structure to incorporate common prior information for specific parameters. While the parameters of BS and MPC are different, their slopes ($\beta_{1,BS}$ and $\beta_{1,MPC}$) and standard deviations (σ_{BS} and σ_{MPC}) could have similar values. In the Bayesian framework, we can assign common hyperpriors (σ_{σ} and σ_{β_1}) for these parameters. This hierarchical structure acts as a form of information sharing between BS and MPC data—since both control strategies operate the same physical RTUs under similar conditions, their temperature-demand relationships should be similar even if their absolute peak demands differ. This sharing is particularly valuable when we have limited experimental data, as it prevents the model from making unrealistic predictions based on sparse observations. This approach prevents overfitting, especially when there is a limited number of observations, by regularizing the parameters using both BS and MPC data. This model has been implemented using PyMC [67], with the differential evolution Metropolis sampler used for sampling. The priors in Fig. 9 (b) are set to cover a wide range of parameter spaces after normalizing the input data between -1 and 1.

4.2.3. Energy cost reduction

A similar modeling approach was applied to the energy cost model, as shown in Fig. 10. The normalized energy cost (x_i) is modeled as a function of the outdoor air temperature (y_i), which is a commonly used methodology in building energy M&V. The energy cost models for BS and MPC are modeled independently (i.e., two independent regression lines), but we use a single parameter for the standard deviation (σ_{ec}) since the variations due to measurements or other stochastic factors, such as human heat gains, are similar regardless of the RTU control. Similar to the peak demand model, PyMC [67] is used to build the model with the differential evolution Metropolis sampler. The priors, as shown in Fig. 10 (b), are set to cover a wide range of parameter spaces after normalizing the input data between -1 and 1.

4.3. M&V variable

The electricity consumption data for the entire school campus was collected through power sensors and utility APIs. Our control strategy focuses solely on the heating and cooling operations of the

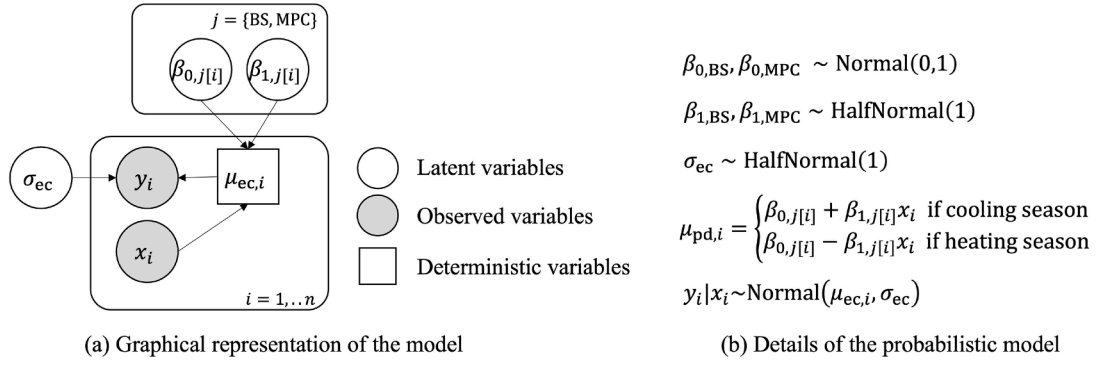


Fig. 10. Graphical representation (a) and probabilistic model structure (b) for an energy cost estimation model.

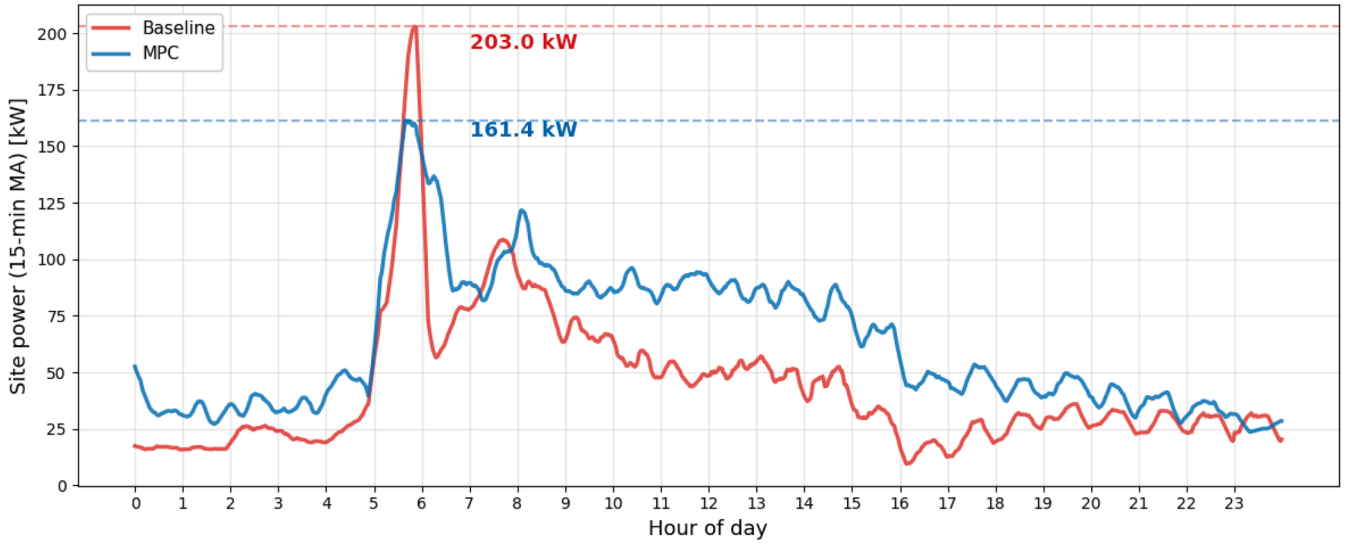


Fig. 11. Comparison of daily power profiles (Baseline vs. MPC) for Case 1 (Site 1) on heating days with similar weather.

RTUs (or HPs) in classrooms. However, these typically account for the majority of the site's electricity consumption, particularly for the daily peak demand. Since utility costs are calculated based on total site electricity usage, we selected this variable as our primary M&V variable.

During periods of mild weather conditions, when heating or cooling demands were low, only a subset of RTUs (or HPs) was in operation. In these instances, the RTU energy consumption could be lower than non-HVAC loads, so the RTUs' energy consumption was used for the M&V analysis. Due to the limited installation of power sensors on individual RTUs, we calculated energy consumption by multiplying the rated power by the operation time of each unit.

5. Results

In this study, our commercial partner implemented the low-cost MPC solution at the demonstration sites through technology transfer. Unlike research experiments conducted in a laboratory setting, the experiments were carried out entirely by the commercial partner to ensure performance in a real-world scenario. To avoid biasing with the experiments by checking interim results, we conducted M&V only after all the experiments were completed. Consequently, we focused on performance verification through a statistical approach rather than presenting individual daily comparisons for each site.

5.1. Case 1: Site 1 trial for heating season

To illustrate the coordination strategy and peak load reduction mechanism, Fig. 11 presents daily site power profiles comparing Baseline and MPC control on days with similar weather conditions. The data are averaged over 15 min intervals, as the utility company charges peak demand based on a 15 min window. In the Baseline operation, the power profile peaks at 6 AM because all RTUs activate simultaneously at the start of the occupancy period. This is primarily because the school district utilizes identical comfort schedules for all classrooms based on their policy (Section 3.2), resulting in a high peak load of 203 kW. Conversely, the MPC initiates heating for specific classrooms prior to 6 AM and staggers the operation of different RTUs. Consequently, only a subset of classroom RTUs are active at 6 AM, which reduces the peak load by 161.4 kW. As it is difficult to visualize the operation of all 38 RTUs at this site, readers can refer to [5] for details on how the MPC coordinates a smaller subset of 6 RTUs.

Fig. 12 compares the HVAC electricity peak demands for Baseline (BS) and MPC operations in Case 1 (Site 1 heating season). The data reveal a clear correlation where daily peak demands increase as outdoor air temperatures decrease, primarily due to heating start-up sequences that trigger significant simultaneous RTU operations. The MPC consistently achieves lower peak demands than the Baseline, particularly during cold weather ($\leq 0^\circ\text{C}$). Notably, the Baseline exhibits unnecessarily high peaks even during mild weather (e.g., reaching nearly 190 kW around 6°C), whereas the MPC effectively mitigates these spikes. However, BS and MPC also show considerable overlap in mild weather

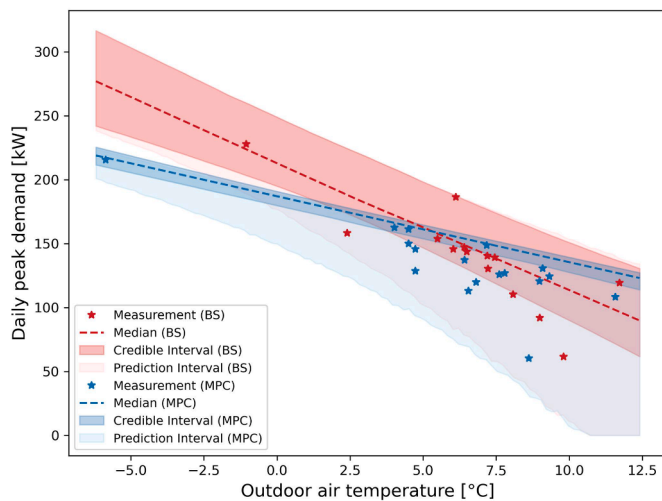


Fig. 12. Comparison of peak demands (HVAC electricity) for Baseline and MPC in Case 1 (Site 1).

conditions. This phenomenon occurs because peak demand depends on simultaneous RTU operations, which exhibit stochastic behavior that becomes more pronounced during mild weather. When outdoor temperatures are mild, not all classrooms require heating simultaneously; some units cycle on briefly while others turn off, creating unpredictable temporal overlaps in equipment operation. This stochastic pattern results in higher variability and occasional coincidental peaks in the Baseline case, where multiple RTUs may activate simultaneously without coordination. These observations highlight the stochastic nature of peak demand—where the maximum potential load correlates with temperature, but individual daily peaks vary—validating the structure of our peak demand model (Section 4.2.2).

The quantitative effectiveness of the MPC is illustrated in Fig. 13, which details the effect estimation derived from the differences between posterior predictive distributions (as conceptually shown in Fig. 8). The analysis reveals a median peak demand reduction of 59.1 kW (21.6%). The 95% credible interval ranges from 21.1 kW (8.9%) to 99.5 kW (32.3%). Furthermore, the MPC demonstrates a positive effect (peak demand reduction) with a 99.9% probability, providing strong evidence of its effectiveness.

It is important to note that these HVAC-specific reductions directly translate to the site-level electricity profile, as HVAC operations are the dominant contributor to the facility's peak demand. The site-level data mirror these trends, showing similar reduction patterns and statistical significance. Detailed comparisons and effect estimations for the total site electricity consumption are provided in the Appendix (Figs. A.1 and A.2).

5.2. Case 2: Site 1 trial for cooling season

In Site 1, experiments were also conducted during the cooling season, and Fig. 14 presents the comparison of the peak demand models of BS and MPC. Unlike the heating season, only a few RTUs were activated throughout the experimental period because the site had small cooling demands, as it is located in WA with mild summer weather. Therefore, we used the estimated RTU electricity consumption for the comparison described in Section 4.3.

In contrast to the heating season, cooling demands are proportional to the increase in outdoor air temperature, and both the BS and MPC data show that trend. Unlike the heating season (where the morning heating peak happens at start-up time), the daily peak demand has more variation in similar weather because the cooling peak is typically observed during the daytime, with more randomness in simultaneous operations. Despite this variability, the MPC consistently maintains lower

daily peak demands, with the exception of the region with sparse data (outdoor air temperature $\geq 20^\circ\text{C}$).

The quantitative intervention effect of the MPC, derived from posterior predictive distributions, is detailed in Appendix Fig. A.3. The analysis indicates a median peak demand reduction of 6.4 kW (28.9%), with the 95% credible interval ranging from 4 kW (20.4%) to 8 kW (36.1%). Furthermore, the probability of a positive reduction is $\geq 99.9\%$, demonstrating high effectiveness even during mild cooling conditions.

5.3. Case 3: Site 2 trial for heating season

In Site 2, we conducted a heating experiment (Case 3). Due to the site schedule and software commissioning periods to replace unstable 3rd party API for weather data (see Section 6.3), we have only a few measurement data points that have clean MPC or BS operations. Similar to Site 1, the daily heating peaks (Fig. 15) were observed in the early morning and are inversely proportional to outdoor air temperature. Despite the small number of observations, it can be seen that the MPC has a lower daily peak demand than the BS.

The quantitative effect of the MPC is visualized in Appendix Fig. A.4. The median effect size is a 25.4 kW (38.9%) reduction, with the 95% credible interval ranging from 12.2 kW (20.1%) to 34.6 kW (48.5%). Consequently, the MPC yields a peak demand reduction with a 99.9% probability.

The site-level electricity analysis shows similar trends to the HVAC results. The detailed comparison and effect estimation for site electricity are provided in the Appendix (Figs. A.5 and A.6).

5.4. Case 4: Site 2 trial for cooling season

As shown in Fig. 16, the cooling season experiment was conducted at Site 2 (Case 4). Unlike the heating operation, the non-HP energy consumption was not negligible, and the total number of HPs in operation was small during the cooling season. Therefore, we also used estimated HP power consumption (Section 4.3) for the M&V. Similar to Case 2, the daily peak demands show a proportional relationship with the outdoor air temperature, with mild slopes and some vertical variations. This can be attributed to the stochastic nature of the cooling peak demand, as observed in Case 2.

The quantitative effect estimation for Case 4 is provided in the Appendix (Fig. A.7). The analysis shows a median reduction of 20.6 kW (29.5%) with a $>99.9\%$ probability of a positive effect.

5.5. Case 5: Site 3 trial for heating season

For Site 3, we conducted only a heating season experiment. Due to scheduling constraints with school policy and code refactoring to resolve the inefficient data query when handling dozens of data endpoints (detailed in Section 6.3), only a limited number of measurements were obtained (Fig. 17). The daily peak demands show an inversely proportional relationship with outdoor air temperatures. Due to the limited number of measurements, the inference mainly relies on the peak demand model. For the cold temperature region ($\leq 13.5^\circ\text{C}$), the BS shows higher peak demands than the MPC, but they exhibit overlapping data points in the mild weather region.

The quantitative effect of the MPC on HVAC peak demand savings is visualized in Appendix Fig. A.8. The median estimated reduction is 24.8 kW (29.1%), with the 95% credible interval ranging from 2.8 kW (4.2%) to 46.1 kW (50.3%). Furthermore, the analysis indicates a 98.3% probability that the MPC yields a peak demand reduction.

5.6. Summary of the results

In this section, we summarize the HVAC (Table 3) and site (Table 4) peak demand reductions for all demonstration cases. The MPC solution showed 21.6–38.9% HVAC peak demand reductions, with a mean of

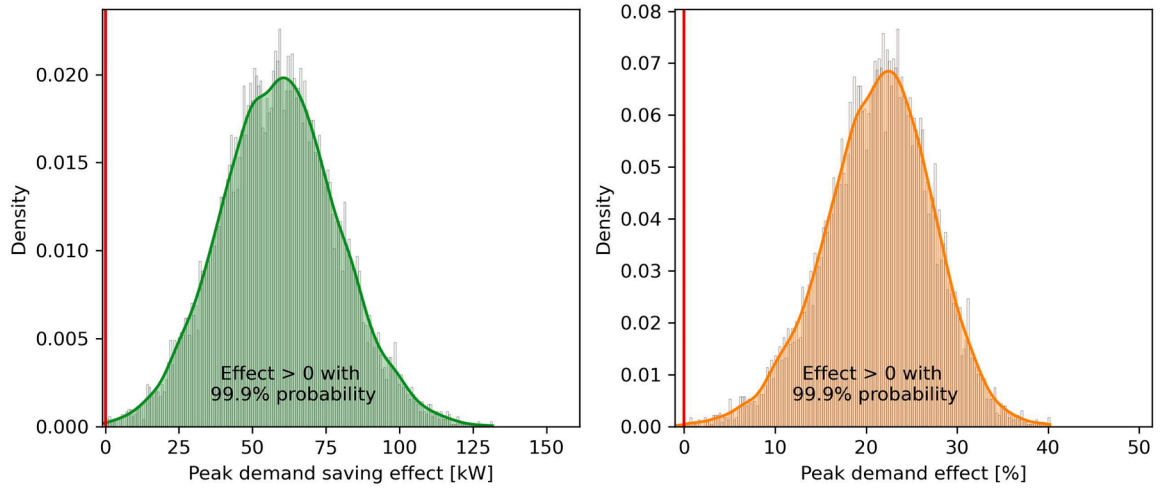


Fig. 13. Peak demand reduction effects (HVAC electricity) in [kW] (left) and in [%] (right) of Case 1 (Site 1).

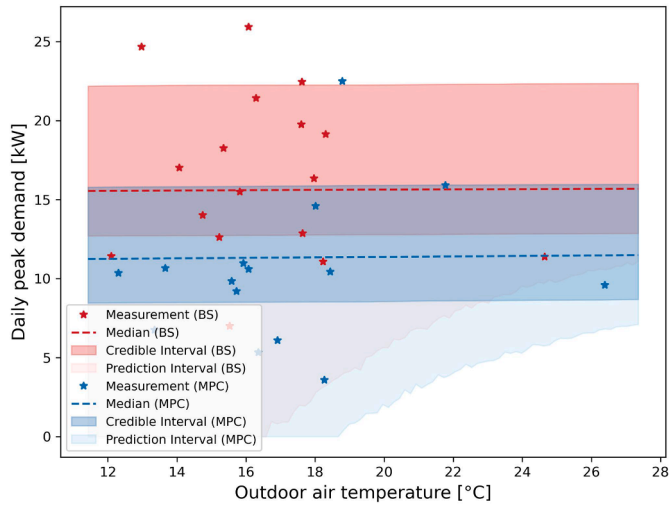


Fig. 14. Comparison of peak demands (HVAC electricity) for Baseline and MPC in Case 2 (Site 1).

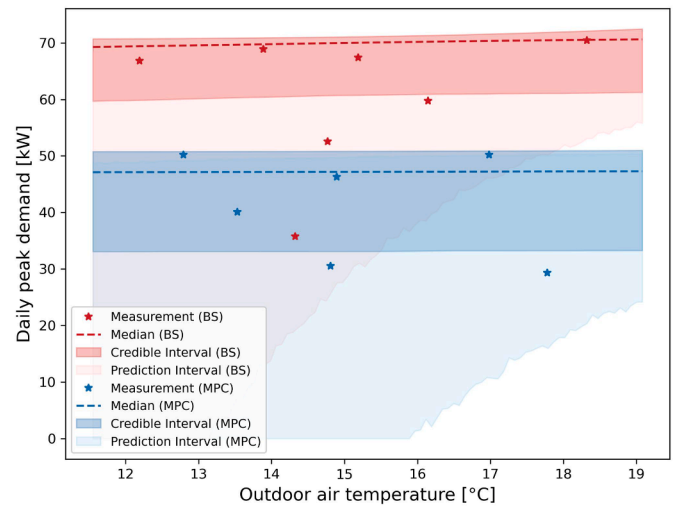


Fig. 16. Comparison of peak demands (Site electricity) for Baseline and MPC in Case 4 (Site 2).

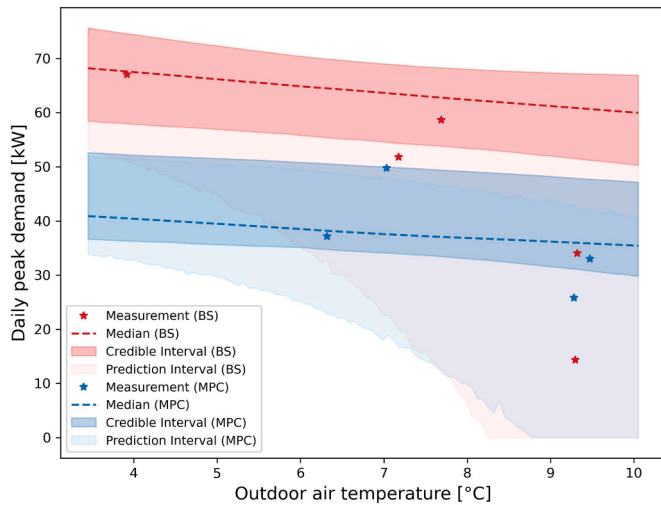


Fig. 15. Comparison of peak demands (HVAC electricity) for Baseline and MPC in Case 3 (Site 2).

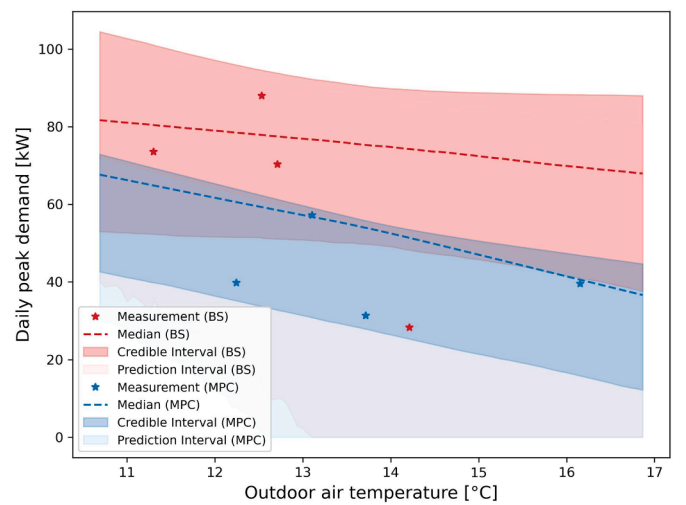


Fig. 17. Comparison of peak demands (HVAC electricity) for Baseline and MPC in Case 5 (Site 3).

Table 3
Summary of HVAC peak demand reduction.

Case ID	Site ID	M&V variable ^a	Peak reduction in [kW] (50%, [2.5%,97.5%]) ^b	Peak reduction in [%] 50%, [2.5%,97.5%] ^b	$P(\text{reduction} > 0)$ ^c [%]	Season
Case 1	Site1	HVAC elec.	59.1, [21.1,99.5]	21.6, [8.9,32.3]	99.9%	Heating
Case 2	Site1	HVAC elec.	6.4, [4.0, 8.0]	28.9, [20.4, 36.1]	100%	Cooling
Case 3	Site2	HVAC elec.	25.4, [12.2, 34.6]	38.9, [20.1,48.5]	99.9%	Heating
Case 4	Site2	HVAC elec.	20.6, [16.2, 28.2]	29.5, [24.5, 40.3]	100%	Cooling
Case 5	Site3	HVAC elec.	24.8, [2.8,46.1]	29.1, [4.2,50.3]	98.3%	Heating

^a A variable used for M&V. (Site elec.: school campus net electricity, HVAC elec.: HVAC electricity inferred from heating/cooling operating signals (see Section 4.3).)

^b 50, [2.5,97.5] percentiles of the credible intervals in the treatment effect.

^c Posterior probability that MPC reduces peak demand compared to Baseline (i.e., percentage of posterior distribution above zero).

29.6% across all cases. This value aligns with our previous study [5], which reported about a 30% HVAC peak demand reduction. For the site peak demand reduction, it shows reductions of 10.8–22.1%.

5.7. Computational performance

To demonstrate the feasibility of real-time field deployment, we benchmarked the UMPC and LMPC computation times across a range of RTU counts on a workstation equipped with an AMD Ryzen 7 5800X processor, and 48 GB DDR4 RAM, running Ubuntu 22.04 under WSL 2.0. The optimization problem sizes scale linearly with the number of zones N_z : the LMPC has $8N_z + 1$ decision variables and $4N_z + 18$ constraints, while the UMPC has $98N_z + 1$ decision variables and $144N_z + 48$ constraints. The substantially larger problem size of the UMPC is primarily attributable to its much longer prediction horizon—48 steps (24 hours at 30 min timesteps) versus only 3 steps (15 min at 5 min timesteps) for the LMPC—which requires a full sequence of decision variables across the entire horizon. Despite the LMPC being the harder problem class (MILP) and the UMPC being an LP, this difference in horizon length dominates the computational scaling behavior observed at high N_z values. Each benchmark was repeated 10 times and a single representative calculation step is reported. Fig. 18 shows the LMPC computation time as a function of the number of RTUs. As a MILP problem, the LMPC exhibits some fluctuation in solve times; however, computation remains well within the 300 s control timestep across the RTU counts encountered in our demonstrations (15–42 RTUs per site). Despite acceptable average-case performance, field deployments with approximately 40 RTUs occasionally experienced solver delays that caused individual control actions to be skipped. To mitigate this, we recommend three complementary strategies: (1) imposing a solver time limit (e.g., 60 s) so that the best feasible solution found within the allotted time is accepted rather than waiting for a provably optimal solution—this is practical because near-optimal RTU coordination still achieves the peak demand reduction objective; (2) implementing a timer-based fallback that skips or retries an optimization cycle when the solver exceeds the time limit; and (3) decomposing the LMPC into smaller independent subproblems (e.g., four groups of 10 RTUs), which preserves the peak-demand coordination objective—distributing RTU operations within each timestep—while substantially reducing the size of each MILP instance. All three strategies can be used in combination, providing layered protection against solver delays in large-scale deployments.

Fig. 19 shows the UMPC computation time. As a linear programming problem, the UMPC scales efficiently and completes within a few seconds for all demonstration sites. Computation time remains acceptable up to approximately 70 RTUs (around 60 s), which is well within the 30 min UMPC control interval. Beyond approximately 85 RTUs ($98 \times 85 + 1 = 8,331$ decision variables and $144 \times 85 + 48 = 12,288$ constraints), solve times increase sharply and solver termination or indefinite hangs become likely. For deployments at or above this scale, we recommend adopting HiGHS [68] as an alternative open-source LP solver, which offers improved numerical stability and scalability.

Table 5 presents the daily average comfort violation hours for both Baseline and MPC operations across all demonstration cases. Comfort violations are defined as periods when room temperatures fall outside the comfort boundaries established by the baseline thermostat setpoints and associated hysteresis bands (deadband). The results demonstrate that both control strategies maintain room temperatures predominantly within the comfort region, with minimal violations observed across all cases. MPC maintains comparable or lower violation hours than Baseline in four out of five cases while simultaneously achieving peak demand reductions, as the MPC formulation constrains zone temperatures to remain within baseline comfort bounds during optimization. The one exception (Case 5) shows a marginal increase of 0.15 hours/day under MPC. Cases 3 and 5 (Sites 2 and 3) show particularly low violation hours due to their single-setpoint control strategy, which operates within a narrow deadband that naturally limits temperature excursions.

6. Discussion

6.1. Energy cost saving

In this paper, we propose both peak demand and energy cost models for estimating the effect of MPC intervention. In our study, MPC reduces energy costs by preheating or precooling the space before peak price periods—storing thermal energy in the air and building mass—and then releasing it during peak price times without relying on mechanical heating or cooling (i.e., load shifting). Therefore, the peak/non-peak ratio and the overlap between peak heating/cooling demand and peak price periods are critical for achieving tangible energy cost savings via MPC. Unfortunately, only one site (Site 3) has a ToU rate structure, meaning that no energy cost savings are expected for Sites 1 and 2. Moreover, the peak/non-peak ratio of ToU rates for commercial buildings in the CA region is very low (e.g., a factor of 1.2), making it difficult to observe clear energy cost savings from load shifting. This dynamic may change in the future with the advancement of electrification and the evolution of dynamic pricing structures.

For example, Figs. 20 and 21 show the energy cost comparison and effect distribution plots for Case 5. Observing the median profiles of the MPC and BS in Fig. 20, one might claim there is an energy cost reduction (\$3.5/day) by calculating the difference between the two lines. However, there are large overlaps between the uncertainty ranges, and the probability of a positive effect is shown as 77%, indicating uncertainty regarding a clear cost reduction due to MPC. This aligns with our expectations as discussed in the previous paragraph.

6.2. Technology transfer

The technology transfer and development of the low-cost MPC solution were conducted concurrently through close collaboration with engineers from our commercial partner. The implementation process was facilitated through comprehensive documentation, real-time communication channels, and regular video conferences. During this pro-

Table 4
Summary of site energy peak demand reduction.

Case ID	Site ID	M&V variable ^a	Peak reduction in [kW] (50%, [2.5%,97.5%]) ^b	Peak reduction in [%] 50%, [2.5%,97.5%] ^b	$P(\text{reduction} > 0)$ ^c [%]	Season
Case 1	Site1	Site elec.	62.8, [24.4, 101.5]	22.1, [9.3, 33.7]	99.2%	Heating
Case 2	Site1	NA	NA	NA	NA	NA
Case 3	Site2	Site elec.	10.7, [3.6, 15.7]	11.0, [3.8, 15.7]	99.7%	Heating
Case 4	Site2	NA	NA	NA	NA	NA
Case 5	Site3	Site elec.	14.6, [6.5, 20.4]	10.8, [4.9, 14.8]	99.8%	Heating

^a A variable used for M&V. (Site elec.: school campus net electricity, HVAC elec.: HVAC electricity inferred from heating/cooling operating signals (see Section 4.3).)

^b 50, [2.5,97.5] percentiles of the credible intervals in the treatment effect.

^c Posterior probability that MPC reduces peak demand compared to Baseline (i.e., percentage of posterior distribution above zero).

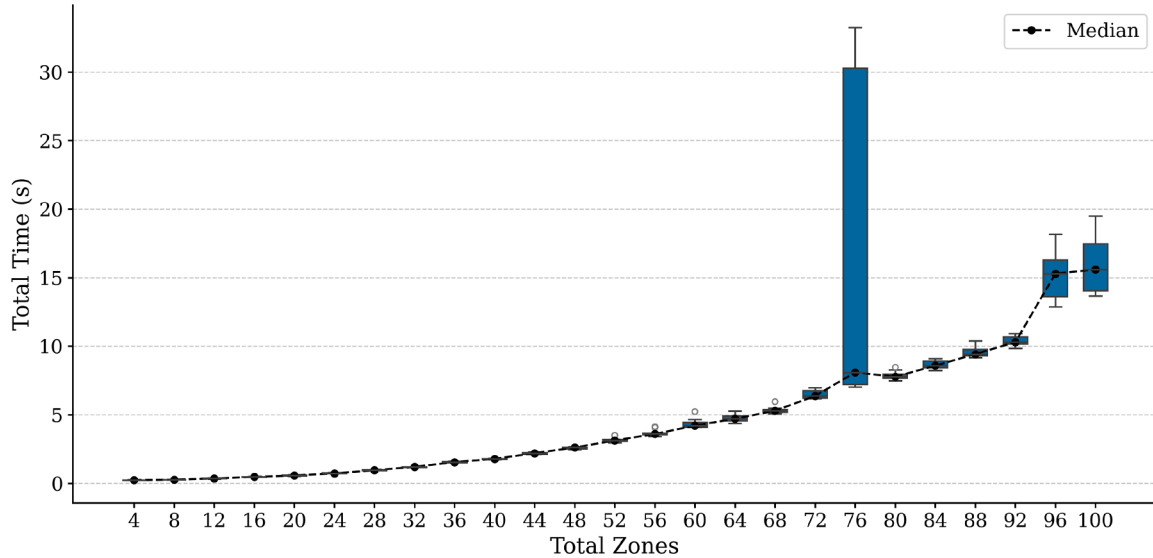


Fig. 18. LMPC computation time as a function of the number of zones (i.e., RTUs). Each data point represents a single calculation step from 10 repeated benchmark runs on the reference hardware. The dashed horizontal line indicates the 300 s LMPC control timestep.

Table 5
Summary of daily average comfort violation hours.

Case ID	Site ID	Baseline [hours/day]	MPC [hours/day]
Case 1	Site 1	1.7	0.6
Case 2	Site 1	0.5	0.3
Case 3	Site 2	0.0	0.0
Case 4	Site 2	0.7	0.4
Case 5	Site 3	0.0	0.15

cess, the lead engineer from our commercial partner developed detailed implementation guidelines and efficiently distributed tasks among team members for multiple deployments.

A significant challenge encountered during implementation was the iterative nature of the process. The diverse range of thermostat controllers necessitated frequent modifications to configuration files and source code to ensure accurate setpoint control and achieve the desired outcomes. When source code modifications were required, we conducted detailed video conferences to demonstrate the implementation of new features, enabling the commercial partner's team to independently manage future updates and modifications.

6.3. Field deployment challenge

During field deployment, several unexpected circumstances delayed and constrained the experiments. One challenge was the interruption of experiments due to instability in third-party API services, particularly

the weather data/forecast service. To prevent such disruptions in future implementations, we recommend decoupling weather data acquisition from the MPC computation loop by implementing a dedicated weather data collection agent. This agent would continuously retrieve and store both historical and forecast data into the main database in real-time, allowing the MPC agent to query reliable local data rather than depending on external API calls during critical optimization cycles. Additionally, the weather agent can be configured with multiple API endpoints for redundancy, distributing usage across providers to minimize service interruptions while managing costs through load balancing.

All computational and API communication errors were systematically logged in the cloud service for troubleshooting purposes. As a conservative approach, we implemented baseline setpoint schedules for emergencies to ensure continuous HP operation during any system issues, such as API service errors. While this design choice ensured safe operation and thermal comfort, these periods could not be classified as either baseline or MPC days. To enhance system reliability monitoring, future implementations should incorporate a heartbeat health-check mechanism within the MPC framework. This feature would periodically verify the operational status of all critical components—including data acquisition agents, optimization solvers, and communication interfaces—and automatically trigger alarms when anomalies are detected. Such proactive monitoring enables rapid identification and resolution of system failures before they impact building operations, while also providing operators with real-time visibility into system health status.

Additionally, as the number of heat pumps (HPs) increased, code execution time exceeded the designated threshold (1-2 min). This in-

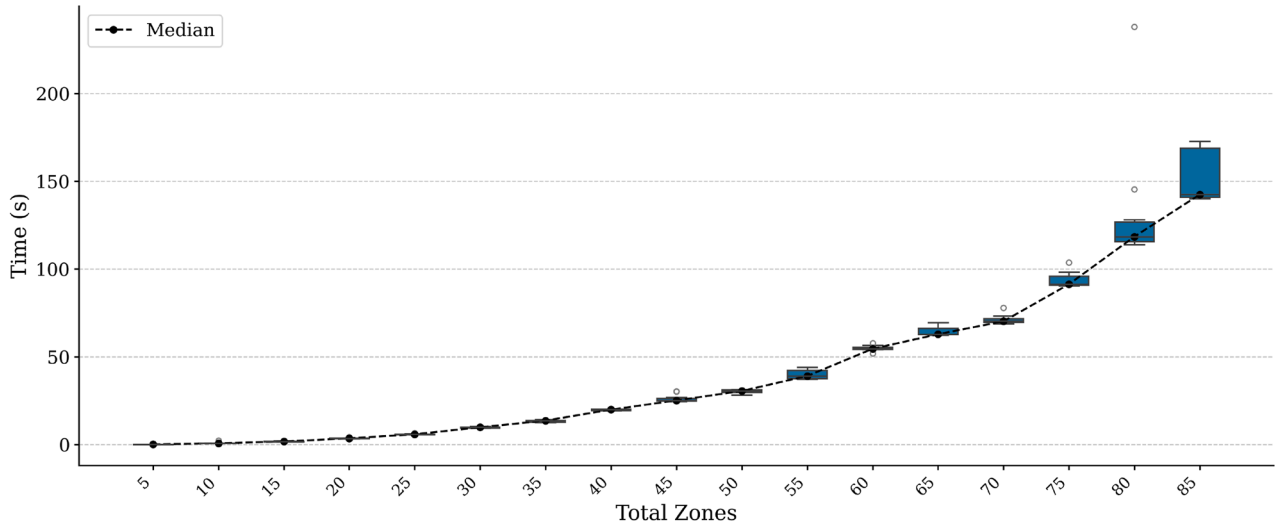


Fig. 19. UMPC computation time as a function of the number of zones (i.e., RTUs). Each data point represents a single calculation step from 10 repeated benchmark runs on the reference hardware. Computation time increases sharply beyond approximately 85 RTUs, where solver termination or indefinite hangs become likely.

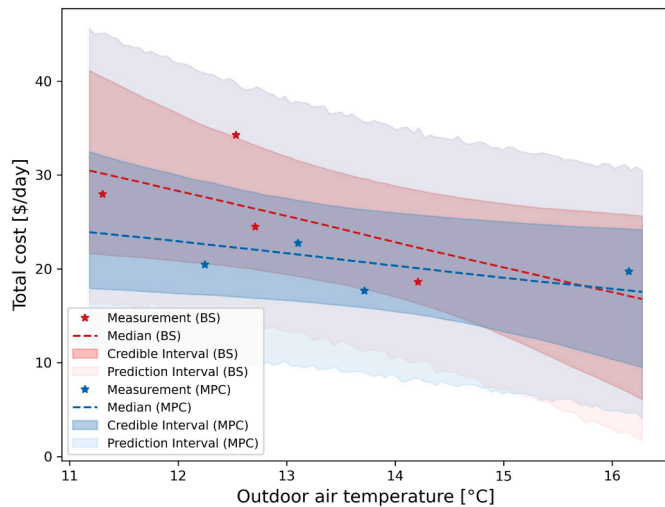


Fig. 20. Comparison of energy costs (site electricity) for Baseline and MPC in Case 5 (Site 3).

efficiency was partially attributed to the software architecture, which queried common information such as weather forecasts and aggregated data for each HP to formulate the optimization problem. Although this issue was eventually resolved, it resulted in the loss of several experimental days. This experience highlights the importance of scalable database architecture design. Future implementations should support a query across multiple endpoints with appropriate time-interval buffering, enabling efficient retrieval of input variables and outputs required for optimization. Since MPC requires recent historical data for state initialization, computational overhead can be further reduced by implementing caching mechanisms that store frequently accessed data in memory. Ideally, deploying an edge computing solution at each site—where both the database and MPC computation reside locally—would eliminate data query latency via cloud API, ensuring consistent real-time performance regardless of system scale. Furthermore, experimental operations were occasionally suspended at the schools' request due to severe weather alerts or special events. While such situations required careful coordination with facility operators, miscommunication sometimes led to unnecessary experimental interruptions.

These challenges particularly affected Cases 3 and 5, resulting in limited clean measurement data points for these experiments. Despite these limitations, the statistical approach used in our M&V methodology (Section 4) allows for robust effect estimation even with smaller sample sizes by leveraging the peak demand model and Bayesian inference.

Finally, we implemented several practical measures, including establishing weekly meetings at the experiment's onset and distributing experimental schedules and instructional materials to teachers regarding proper thermostat usage. For example, teachers initially kept overriding thermostat setpoints when they entered classrooms and submitted work orders because they confused the thermostat's setpoint with the actual room temperature. Through communication and data analysis, we discovered that this behavior was habitual rather than a response to thermal discomfort [5]. Furthermore, the issue with incorrect work orders was resolved by educating the teachers on how to accurately read the thermostat status. Then, we distributed educational materials such as fliers before we initiated the MPC.

6.4. Balancing demand flexibility with occupant thermal comfort

We adopted temperature-based comfort constraints rather than more comprehensive metrics like the Predicted Mean Vote (PMV) index for several practical reasons. First, PMV requires monitoring multiple variables—mean radiant temperature, humidity, clothing levels, and metabolic rates—that are difficult to measure reliably in operational buildings. Second, incorporating PMV increases computational complexity that could compromise real-time control performance. Third, PMV cannot fully address the substantial variability in individual thermal preferences among occupants. Our approach maintains parity with current practice, as baseline systems also rely on temperature-based schedules with manual override capability, while adding demand flexibility. Importantly, when occupants override thermostat settings, the overridden unit's operation is incorporated as a constraint in the optimization, and the system calculates control strategies for the remaining units to accommodate the override while still pursuing demand management objectives. This design ensures individual that comfort preferences take precedence while maintaining system-level coordination. The strategy achieved substantial demand reduction in our school building deployments (Section 5), demonstrating that effective demand flexibility can be achieved within practical implementation constraints.

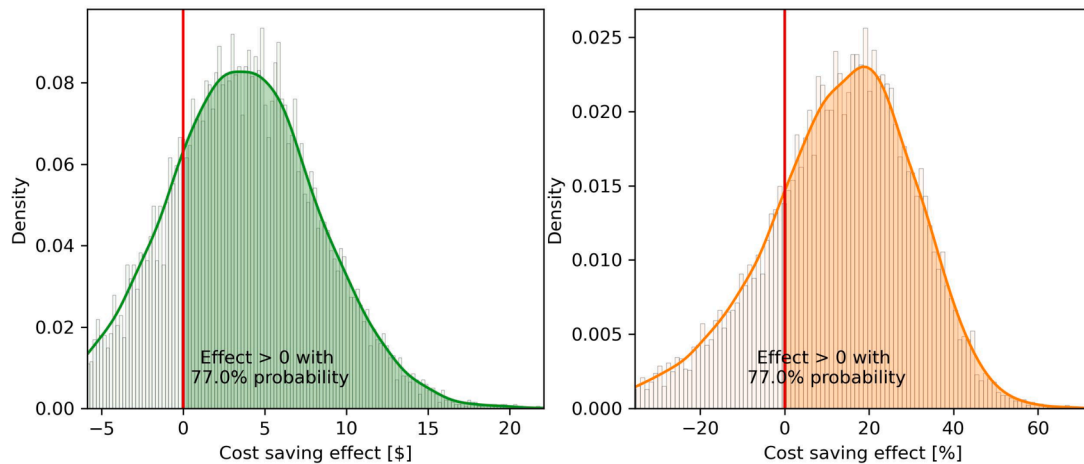


Fig. 21. Energy cost reduction effects (site electricity) in [\$] (left) and in [%] (right) of Case 5 (Site 3).

6.5. Controller scalability and replicability

A key advantage of the HMPC approach is its scalability and transferability across different building sites. The same tuning parameters ($\omega_d = 10$, $\omega_l = \omega_u = 1000$) were applied successfully across all five schools in this study without site-specific adjustments, demonstrating the generalizability of the control strategy. When more aggressive demand reduction is desired, users can increase the demand term weight (ω_d). From a computational perspective, as quantified in Section 5.7, the LP-based UMPC scales efficiently and completes within a few seconds for RTU counts encountered in this study (15–42 RTUs per site), and remains acceptable up to approximately 70 RTUs (around 60 s), which is well within the 30 min UMPC control interval. Beyond approximately 85 RTUs, solver termination or indefinite hangs become likely, and adopting HiGS [68] is recommended for larger-scale deployments. The MILP-based LMPC is computationally more demanding due to its binary decision variables; practical mitigation strategies for larger deployments are discussed in Section 5.7. From a computational perspective, the MILP formulation enables real-time optimization, scaling to buildings with up to approximately 50 RTU zones on standard hardware. Our implementation completes optimization within one minute (optimization calculation only), well within the 5 min LMPC control interval, providing adequate computational headroom even during peak load conditions. The primary deployment requirements—web-enabled thermostats with API capability and reliable weather forecasts—represent infrastructure that is increasingly common in modern buildings. Importantly, by leveraging existing web-enabled thermostats without additional sensor installations, the approach significantly reduces commissioning time and eliminates hardware retrofit costs, lowering barriers to adoption. These characteristics suggest that the HMPC framework can be readily replicated across school districts and other SMCB portfolios without extensive site-specific calibration, making it a practical solution for large-scale demand flexibility programs. However, a climate-dependent limitation exists: in regions with extreme climates requiring near-continuous heating or cooling (approaching 100% RTU capacity utilization), the absolute potential for demand flexibility becomes inherently limited, as there is minimal operational margin for load shifting and coordination for peak demand reduction.

7. Conclusion

In this study, we extended the implementation of our previously developed low-cost MPC solution [5] to multiple K-12 school campuses for grid-interactive HVAC operations. While the low-cost MPC solution was designed for universal application in school buildings equipped with packaged systems (e.g., Rooftop Units or Heat Pumps) through auto-

mated software toolchains, widespread market adoption relies on successful technology transfer to commercial partners. The primary objective of this study is to present demonstration results achieved through technology transfer (implemented by the commercial partner) and provide statistical validation of performance outcomes. The MPC solution was deployed across three school campuses, encompassing five distinct experimental cases.

We developed a novel statistical approach utilizing Bayesian methods to estimate the MPC's impact on peak demand and energy cost reductions, specifically designed for sparse and limited datasets. The results demonstrate that MPC consistently reduces peak demand across all sites with high probability (98–100% posterior probability of positive effect), achieving a 10.8–22.1% reduction in site energy consumption and a 21.6–38.9% reduction in HVAC energy consumption, consistent with our previous findings [5].

However, scaling up the demonstrations revealed several challenges, including diverse thermostat configurations, computational delays from multiple queries, and communication inefficiencies in managing special event operations and troubleshooting. These challenges are thoroughly examined as lessons learned in the discussion section (Section 6). Nevertheless, the low-cost MPC solution represents a compelling case study of successful research-to-market technology transfer, facilitating the development of grid-interactive school buildings to enhance grid resilience through scalable demand flexibility.

CRediT authorship contribution statement

Sang Woo Ham: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization; **Mohamad El Hariri:** Writing – review & editing, Software, Investigation, Data curation; **Donghun Kim:** Writing – review & editing, Supervision, Software, Project administration, Methodology, Funding acquisition, Conceptualization; **Tanya Barham:** Writing – review & editing, Supervision, Funding acquisition.

Data availability

The authors do not have permission to share data.

Declaration of competing interest

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

Acknowledgement

This work was supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, Building Technologies Office of the [Department of Energy](#) under Contract No. [DE-AC02-05CH11231](#), through the CRADA project entitled “Get K-12 Ready for Future Utility Tariff Changes and Load Flexibility: Wholistic Control Package”.

Appendix A. Appendix

Case 1: Site power comparisons

[Fig. A.1](#) presents the comparison of site-level electricity peak demands, obtained from the whole-campus power meter (encompassing both HVAC and non-HVAC loads). Overall, the site-level profiles mirror the HVAC trends discussed in the main text, though there is greater overlap between BS and MPC during warmer weather. This is expected, as the MPC controls only the HVAC system (RTUs), and non-HVAC loads

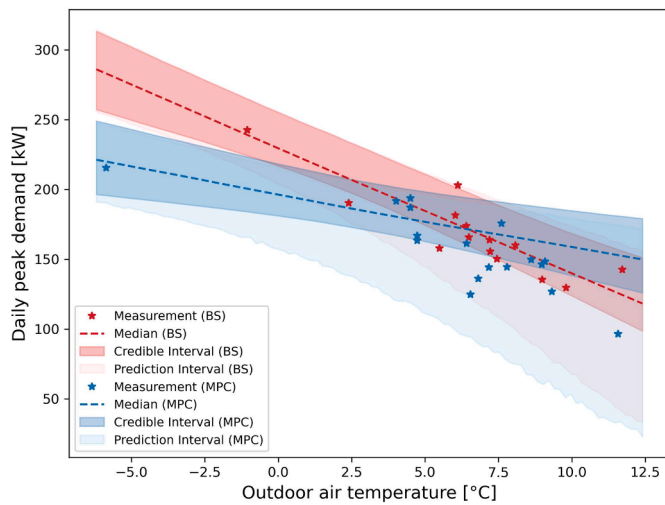


Fig. A.1. Comparison of peak demands (Site electricity) for Baseline and MPC in Case 1 (Site 1).

become more dominant when heating demand is low. Nevertheless, the similarity between site and HVAC profiles confirms that HVAC operation is the primary driver of site-level peak demand.

The site-level savings are visualized in [Fig. A.2](#). The median effect shows a reduction of 62.8 kW (22.1%), with the effect size ranging from 24.4 kW (9.3%) to 101.5 kW (33.7%) within the 95% credible intervals. Consistent with the HVAC-specific results, the MPC demonstrates a site-level peak demand reduction with a 99.9% probability.

Case 2: Effect estimation

[Fig. A.3](#) visualizes the peak demand reduction effects for the cooling season in Case 2 (Site 1).

Case 3: Site power comparisons and effect estimation

[Fig. A.4](#) visualizes the peak demand reduction effects (HVAC electricity) for Case 3 (Site 2), supporting the quantitative analysis presented in [Section 5.3](#).

The site peak demands of the BS and the MPC for Case 3 are compared in [Fig. A.5](#). The site peak demands decrease as the outdoor air temperature increases, and this trend aligns with the trend of the HVAC peak demands, with fewer variations similar to Case 1.

The distribution of the MPC effect on site electricity is visualized in [Fig. A.6](#). While the median effect is a 10.7 kW (11.0%) reduction, the 95% credible intervals range between 3.6 kW (3.8%) and 15.7 kW (15.7%). Additionally, the effect is positive with a 99.7% probability, within a range of 4.2 kW to 15.4 kW.

Case 4: Effect estimation

[Fig. A.7](#) shows the distribution of the MPC effect for Case 4. The median effect is a reduction of 20.6 kW (29.5%), and the 95% credible interval ranges from 16.2 kW (24.5%) to 28.2 kW (40.3%). As a result, the probability of the effect being greater than 0 is 100%, with a range of 16.2 kW to 28.2 kW.

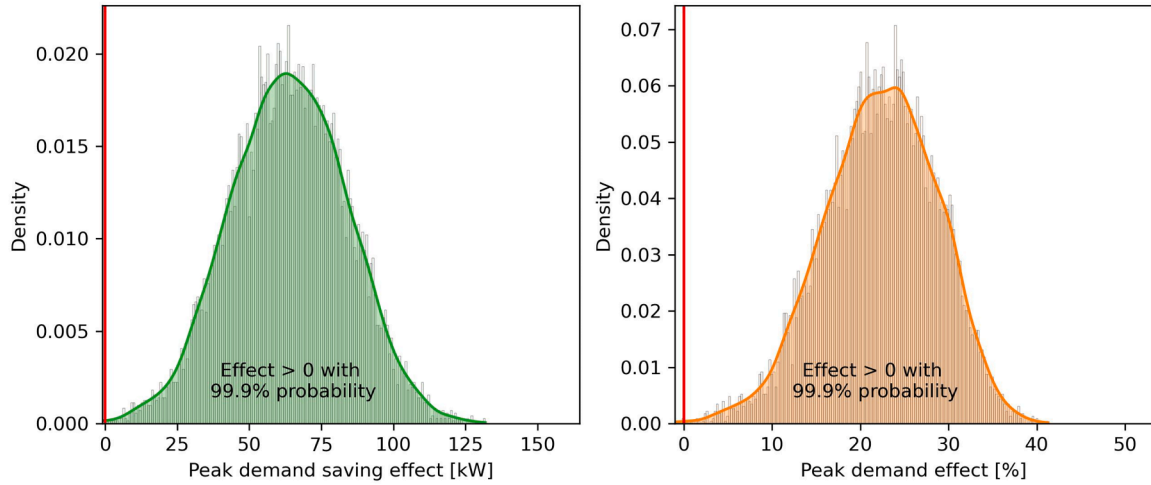


Fig. A.2. Peak demand reduction effects (Site electricity) in [kW] (left) and in [%] (right) of Case 1 (Site 1).

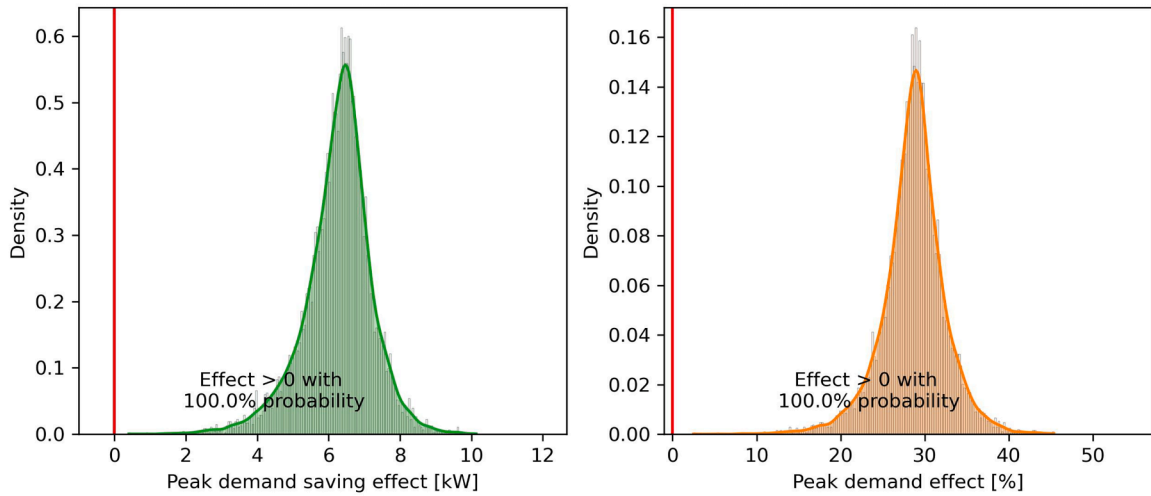


Fig. A.3. Peak demand reduction effects (HVAC electricity) in [kW] (left) and in [%] (right) of Case 2 (Site 1).

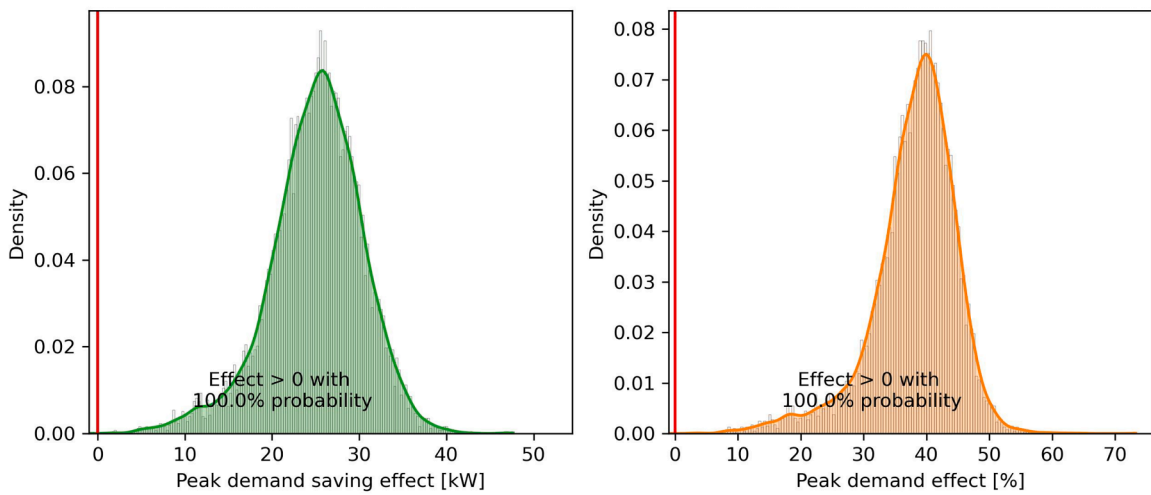


Fig. A.4. Peak demand reduction effects (HVAC electricity) in [kW] (left) and in [%] (right) of Case 3 (Site 2).

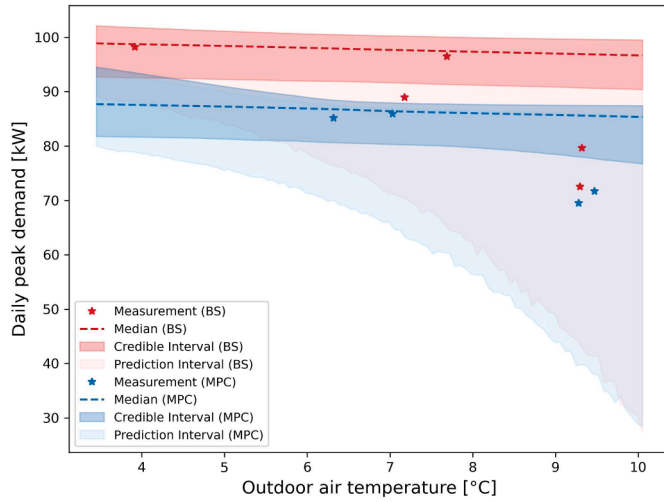


Fig. A.5. Comparison of peak demands (Site electricity) for Baseline and MPC in Case 3 (Site 2).

Case 5: Site power comparisons and effect estimation

Fig. A.8 visualizes the peak demand reduction effects (HVAC electricity) for Case 5 (Site 3), supporting the quantitative analysis presented in Section 5.5.

We also compared the site peak demands of BS and MPC (Fig. A.9). Similar to other winter cases, site peak demands are negatively correlated with outdoor air temperature and exhibit a profile similar to the HVAC peak demands, though with less variation between individual trials.

Fig. A.10 illustrates the distribution of the estimated MPC effect on site electricity. The median effect is a reduction of 14.6 kW (10.8%), with 95% credible intervals ranging from 6.5 kW (4.9%) to 20.4 kW (14.8%). The probability of a positive reduction (effect > 0) is 99.8%.

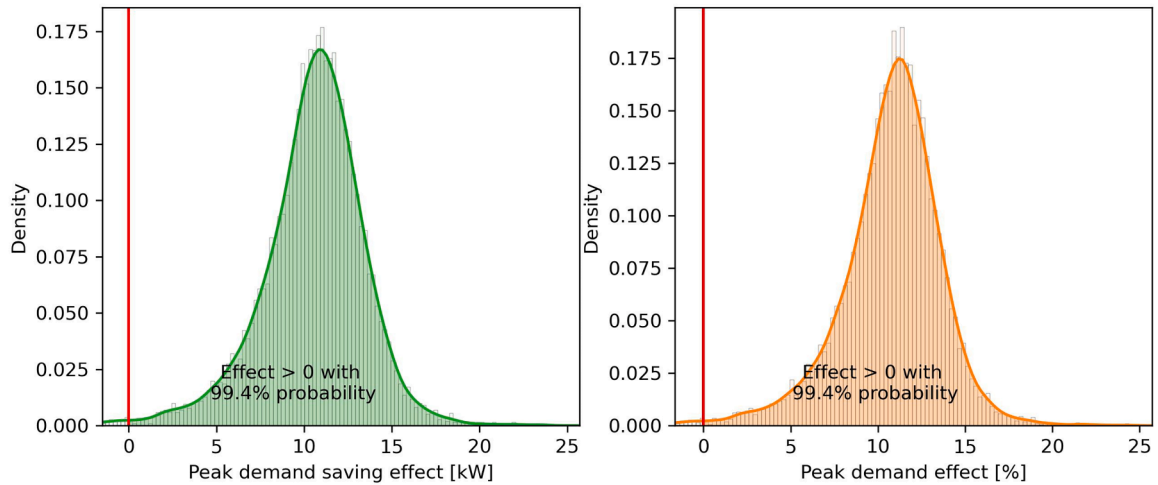


Fig. A.6. Peak demand reduction effects (site electricity) in [kW] (left) and in [%] (right) of Case 3 (Site 2).

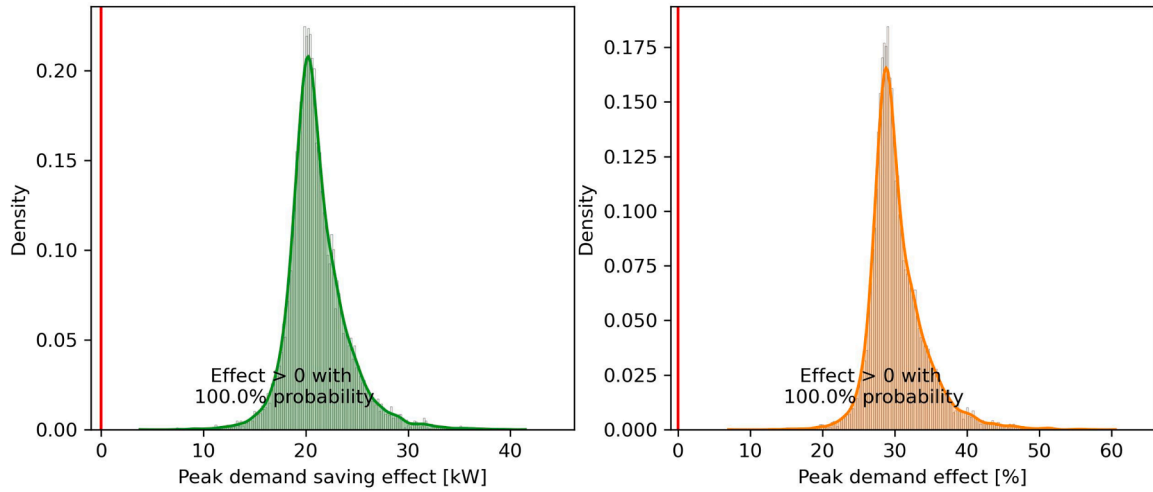


Fig. A.7. Peak demand reduction effects (site electricity) in [kW] (left) and in [%] (right) of Case 4 (Site 2).

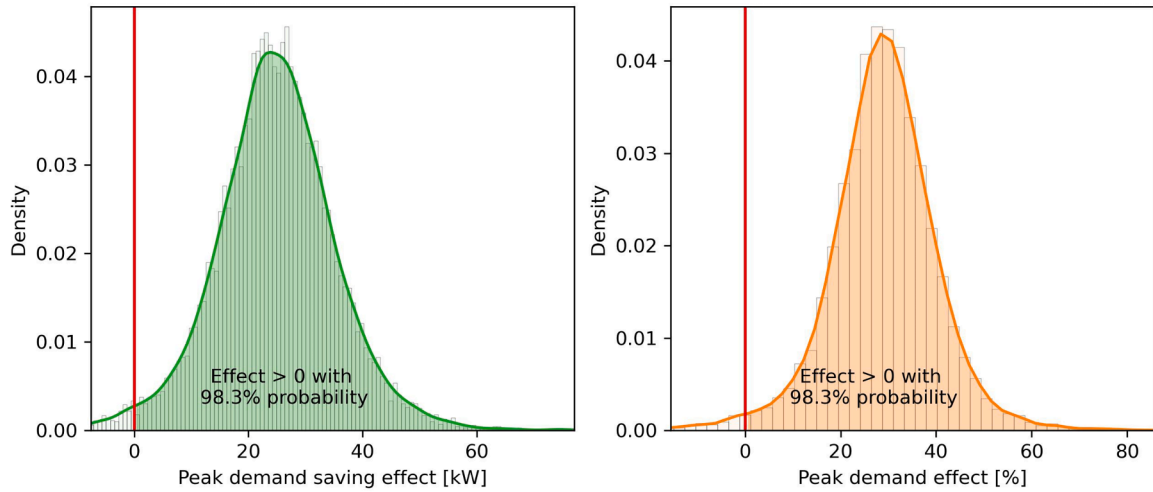


Fig. A.8. Peak demand reduction effects (HVAC electricity) in [kW] (left) and in [%] (right) of Case 5 (Site 3).

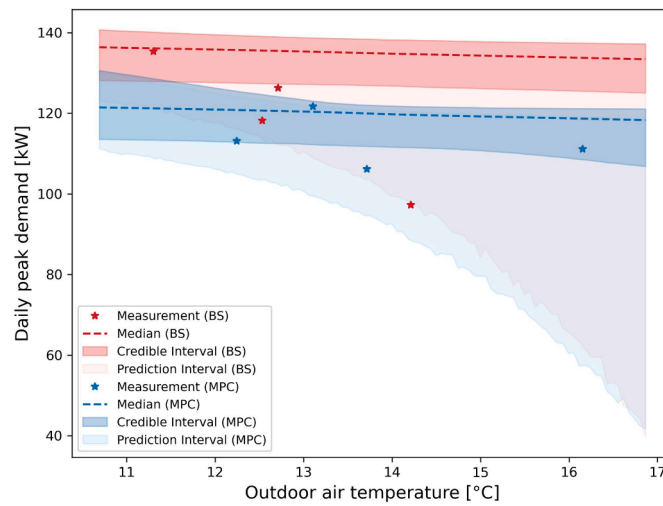


Fig. A.9. Comparison of peak demands (Site electricity) for Baseline and MPC in Case 5 (Site 3).

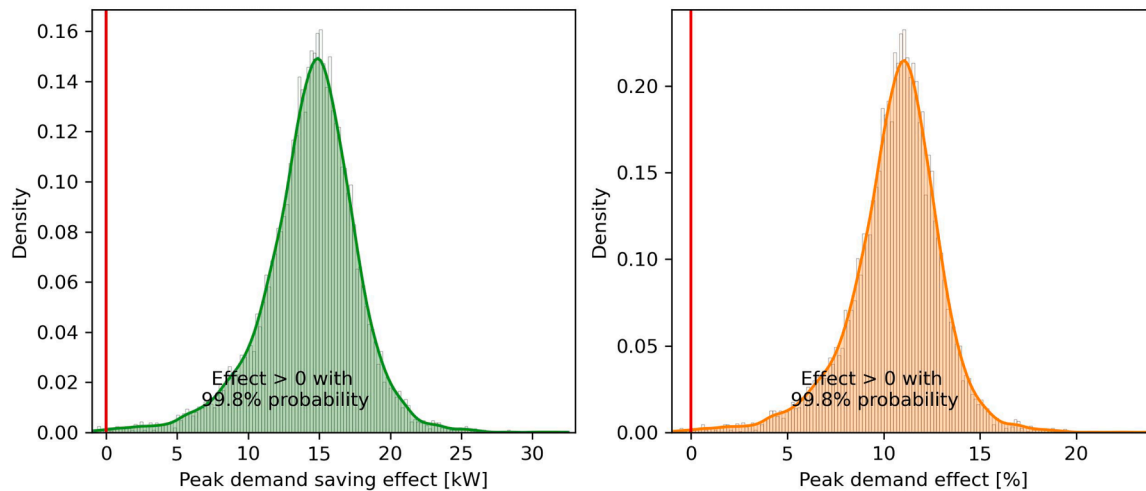


Fig. A.10. Peak demand reduction effects (Site electricity) in [kW] (left) and in [%] (right) of Case 5 (Site 3).

References

- NERC, Long-term reliability assessment, 2024, https://www.nerc.com/pa/RAPA/ra/Reliability%20Assessments%20DL/NERC_Long%20Term%20Reliability%20Assessment_2024.pdf.
- NERC, Urgent need for resources over 10-year horizon as electricity demand growth accelerates, LTRA Finds, 2024, https://www.nerc.com/news/Headlines%20DL/17_2024%20LTRA%20Announcement%20final.pdf.
- M. Neukomm, V. Nubbe, R. Pares, Grid-interactive efficient buildings technical report series: overview of research challenges and gaps, 2019, <https://doi.org/10.2172/1577966>.
- ASHRAE, ASHRAE Guideline 14-2021, high performance sequences of operation for HVAC systems, Atlanta: American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc, 2021.
- S.W. Ham, D. Kim, T. Barham, K. Ramseyer, The first field application of a low-cost MPC for grid-interactive K-12 schools: Lessons-learned and savings assessment 296 (2023) 113351. <https://doi.org/10.1016/j.enbuild.2023.113351>.
- D. Kim, J.E. Braun, MPC solution for optimal load shifting for buildings with ON/OFF staged packaged units: Experimental demonstration, and lessons learned 266 (2022) 112118. <https://doi.org/10.1016/j.enbuild.2022.112118>.
- S. Touzani, A.K. Prakash, Z. Wang, S. Agarwal, M. Pritoni, M. Kiran, R. Brown, J. Granderson, Controlling distributed energy resources via deep reinforcement learning for load flexibility and energy efficiency 304 (2021) 117733. <https://doi.org/10.1016/j.apenergy.2021.117733>.
- D. Kim, Z. Wang, J. Brugger, D. Blum, M. Wetter, T. Hong, M.A. Piette, Site demonstration and performance evaluation of MPC for a large chiller plant with TES for renewable energy integration and grid decarbonization 321 (2022) 119343. <https://doi.org/10.1016/j.apenergy.2022.119343>.
- National center for education statistics, Digest of education statistics - table 105.50. Number of educational institutions, by level and control of institution: 2009-10 through 2019-20, 2022, https://nces.ed.gov/programs/digest/d21/tables/dt21_105.50.asp.
- U.S. Government Accountability Office, K-12 Education: School Districts Frequently Identified Multiple Building Systems Needing Updates or Replacement, 2020, <https://www.gao.gov/products/gao-20-494>.
- U.S. energy information administration (EIA), Commercial buildings energy consumption survey (CBECS) Data: Table PBA3. Sum of major fuel consumption totals and gross energy intensities by building activity subcategories, 2012, 2012.
- U.S. energy information administration (EIA), Commercial buildings energy consumption survey (CBECS) Data: Table B38-B46, 2012.
- New buildings institute, Why K-12 should feature in America's national climate strategy, 2021, <https://newbuildings.org/resource/why-k-12-should-feature-in-americas-national-climate-strategy/>.
- J. Drgoňa, J. Arroyo, I. Cupeiro Figueroa, D. Blum, K. Arendt, D. Kim, E.P. Ollé, J. Oravec, M. Wetter, D.L. Vrabie, L. Helsen, All you need to know about model predictive control for buildings 50 (2020) 190–232. <https://doi.org/10.1016/j.arcontrol.2020.09.001>.
- A. Satchwell, M. Piette, A. Khandekar, J. Granderson, N. Frick, R. Hledik, A. Faruqui, L. Lam, S. Ross, J. Cohen, K. Wang, D. Urigwe, D. Delurey, M. Neukomm, D. Nemtsov, A national roadmap for grid-interactive efficient buildings, 2021, <https://www.osti.gov/servlets/purl/1784302/>.
- A.J. Khabbazi, E.N. Pergantis, L.D. Reyes Premer, P. Papageorgiou, A.H. Lee, J.E. Braun, G.P. Henze, K.J. Kircher, Lessons learned from field demonstrations of model predictive control and reinforcement learning for residential and commercial HVAC: a review 399 (2025) 126459. <https://doi.org/10.1016/j.apenergy.2025.126459>.
- T.S. Pedersen, P. Andersen, K.M. Nielsen, Central control of heat pumps for smart grid purposes tested on single family houses, in: 2013 10th IEEE International Conference on Networking, Sensing and Control (ICNSC), IEEE, 2013, pp. 118–123. <https://doi.org/10.1109/icnsc.2013.6548722>.
- D. Lindelöf, H. Afshari, M. Alisafaei, J. Biswas, M. Caban, X. Mocellin, J. Viane, Field tests of an adaptive, model-predictive heating controller for residential buildings 99 (2025) 292–302. <https://doi.org/10.1016/j.enbuild.2015.04.029>.
- Y. Zong, G.M. Böning, R.M. Santos, S. You, J. Hu, X. Han, Challenges of implementing economic model predictive control strategy for buildings interacting with smart energy systems 114 (2017) 1476–1486. <https://doi.org/10.1016/j.applthermaleng.2016.11.141>.
- A. Afram, F. Janabi-Sharifi, Supervisory model predictive controller (MPC) for residential HVAC systems: implementation and experimentation on archetype sustainable house in toronto 154 (2017) 268–282. <https://doi.org/10.1016/j.enbuild.2017.08.060>.
- F. Bünnig, B. Huber, P. Heer, A. Aboudonia, J. Lygeros, Experimental demonstration of data predictive control for energy optimization and thermal comfort in buildings 211 (2020) 109792. <https://doi.org/10.1016/j.enbuild.2020.109792>.
- F. Bünnig, B. Huber, A. Schalbetter, A. Aboudonia, M. Hudoba de Badyn, P. Heer, R.S. Smith, J. Lygeros, Physics-informed linear regression is competitive with two machine learning methods in residential building MPC 310 (2022) 118491. <https://doi.org/10.1016/j.apenergy.2021.118491>.
- S. Thorsteinsson, A.A.S. Kalaei, P. Vogler-Finck, H.L. Stærnmoose, I. Katic, J.D. Bendtsen, Long-term experimental study of price responsive predictive control in a real occupied single-family house with heat pump 347 (2023) 121398. <https://doi.org/10.1016/j.apenergy.2023.121398>.
- Y. Jiang, A. Ejenakevwe, J. Wang, C.Y. Tang, L. Song, Development, implementation, and impact analysis of model predictive control-based optimal precooling using smart home thermostats (2023) 113790. <https://doi.org/10.1016/j.enbuild.2023.113790>.
- E.N. Pergantis, Priyadarshan, N.A. Theeb, P. Dhillon, J.P. Ore, D. Ziviani, E.A. Groll, K.J. Kircher, Field demonstration of predictive heating control for an all-electric house in a cold climate 360 (2024) 122820. <https://doi.org/10.1016/j.apenergy.2024.122820>.
- E.N. Pergantis, P. Dhillon, L.D.R. Premer, A.H. Lee, D. Ziviani, K.J. Kircher, Humidity-aware model predictive control for residential air conditioning: a field study 266 (2024) 112093. <https://doi.org/10.1016/j.builtenv.2024.112093>.
- E.N. Pergantis, L.D. Reyes Premer, A.H. Lee, Priyadarshan, H. Liu, E.A. Groll, D. Ziviani, K.J. Kircher, Protecting residential electrical panels and service through model predictive control: A field study 386 (2025) 125528. <https://doi.org/10.1016/j.apenergy.2025.125528>.
- R. Yang, M.W. Newman, Learning from a learning thermostat: lessons for intelligent systems for the home, in: Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing, UbiComp '13, Association for Computing Machinery, 2013, pp. 93–102. <https://doi.org/10.1145/2493432.2493489>.
- Z. Liao, A.L. Dexter, An inferential model-based predictive control scheme for optimizing the operation of boilers in building space-heating systems 18 (2010) 1092–1102. <https://doi.org/10.1109/tcst.2009.2033667>.
- S. Privara, Z. Vana, D. Gyalistras, J. Cigler, C. Sagerschnig, M. Morari, L. Ferkl, Modeling and identification of a large multi-zone office building, in: 2011 IEEE International Conference on Control Applications (CCA), IEEE, 2011, pp. 55–60. <https://doi.org/10.1109/CCA.2011.6044402>.
- J. Široký, F. Oldewurtel, J. Cigler, S. Privara, Experimental analysis of model predictive control for an energy efficient building heating system 88 (2011) 3079–3087. <https://doi.org/10.1016/j.apenergy.2011.03.009>.
- M. Killian, M. Kozek, Implementation of cooperative Fuzzy model predictive control for an energy-efficient office building 158 (2018) 1404–1416. <https://doi.org/10.1016/j.enbuild.2017.11.021>.
- J. Joe, P. Karava, X. Hou, Y. Xiao, J. Hu, A distributed approach to model-predictive control of radiant comfort delivery systems in office spaces with localized

- thermal environments 175 (2018) 173–188. <https://doi.org/10.1016/j.enbuild.2018.06.068>
- [34] J. Joe, P. Karava, A model predictive control strategy to optimize the performance of radiant floor heating and cooling systems in office buildings (2018). <https://www.sciencedirect.com/science/article/pii/S0306261919306191>
- [35] S. Lee, J. Joe, P. Karava, I. Bilonis, A. Tzempelikos, Implementation of a self-tuned HVAC controller to satisfy occupant thermal preferences and optimize energy use 194 (2019) 301–316. <https://doi.org/10.1016/j.enbuild.2019.04.016>
- [36] S.C. Bengae, A.D. Kelman, F. Borrelli, R. Taylor, S. Narayanan, Implementation of model predictive control for an HVAC system in a mid-size commercial building 20 (2014) 121–135. <https://doi.org/10.1080/10789669.2013.834781>
- [37] S.R. West, J.K. Ward, J. Wall, Trial results from a model predictive control and optimisation system for commercial building HVAC 72 (2014) 271–279. <https://doi.org/10.1016/j.enbuild.2013.12.037>
- [38] M. Castilla, J.D. Álvarez, J.E. Normey-Rico, F. Rodríguez, Thermal comfort control using a non-linear MPC strategy: a real case of study in a bioclimatic building 24 (2014) 703–713. <https://doi.org/10.1016/j.jprocont.2013.08.009>
- [39] Z. Váňa, J. Cigler, J. Široký, E. Žáčková, L. Ferkl, Model-based energy efficient control applied to an office building 24 (2014) 790–797. <https://doi.org/10.1016/j.jprocont.2014.01.016>
- [40] R. De Coninck, L. Helsén, Practical implementation and evaluation of model predictive control for an office building in Brussels 111 (2016) 290–298. <https://doi.org/10.1016/j.enbuild.2015.11.014>
- [41] J. Drgoňa, D. Picard, L. Helsén, Cloud-based implementation of white-box model predictive control for a GEOTABS office building: a field test demonstration 88 (2020) 63–77. <https://doi.org/10.1016/j.jprocont.2020.02.007>
- [42] S. Freund, G. Schmitz, Implementation of model predictive control in a large-sized, low-energy office building 197 (2021) 107830. <https://doi.org/10.1016/j.buildenv.2021.107830>
- [43] P.R. Sampaio, R. Salvazet, P. Mandel, G. Becker, D. Chenu, Simulation and optimal control of heating and cooling systems: a case study of a commercial building 246 (2021) 111102. <https://doi.org/10.1016/j.enbuild.2021.111102>
- [44] N. Lolli, E. Gorantonaki, J. Clauß, Predictive heating control and perceived thermal comfort in a Norwegian office building 17 (2024) 3719. <https://doi.org/10.3390/en17153719>
- [45] M. Schmidt, M.V. Moreno, A. Schülke, K. Macek, K. Mařík, A.G. Pastor, Optimizing legacy building operation: the evolution into data-driven predictive cyber-physical systems 148 (2017) 257–279. <https://doi.org/10.1016/j.enbuild.2017.05.002>
- [46] K. Zhang, A. Prakash, L. Paul, D. Blum, P. Alstone, J. Zoellick, R. Brown, M. Pritoni, Model predictive control for demand flexibility: real-world operation of a commercial building with photovoltaic and battery systems 7 (2022) 100099. <https://doi.org/10.1016/j.adapen.2022.100099>
- [47] Y. Ma, F. Borrelli, B. Hency, B. Coffey, S. Bengae, P. Haves, Model predictive control for the operation of building cooling systems 20 (2012) 796–803. <https://doi.org/10.1109/tcst.2011.2124461>
- [48] E. Saloux, J. Runge, K. Zhang, Operation optimization of multi-boiler district heating systems using artificial intelligence-based model predictive control: Field demonstrations 285 (2023) 129524. <https://doi.org/10.1016/j.energy.2023.129524>
- [49] B. Yue, B. Su, F. Xiao, A. Li, K. Li, S. Li, R. Yan, Q. Lian, A. Li, Y. Li, X. Fang, X. Liang, Energy-oriented control retrofit for existing HVAC system adopting data-driven MPC - methodology, implementation and field test 295 (2023) 113286. <https://doi.org/10.1016/j.enbuild.2023.113286>
- [50] D.H. Blum, K. Arendt, L. Rivalin, M.A. Piette, M. Wetter, C.T. Veje, Practical factors of envelope model setup and their effects on the performance of model predictive control for building heating, ventilating, and air conditioning systems 236 (2019) 410–425. <https://doi.org/10.1016/j.apenergy.2018.11.093>
- [51] M. Wetter, W. Zuo, T.S. Nouidui, X. Pang, Modelica buildings library 7 (2014) 253–270. <https://doi.org/10.1080/19401493.2013.765506>
- [52] F. Jorissen, W. Boydens, L. Helsén, TACO, an automated toolchain for model predictive control of building systems: implementation and verification 12 (2019) 180–192. <https://doi.org/10.1080/19401493.2018.1498537>
- [53] D. Kim, J. Cai, K.B. Ariyur, J.E. Braun, System identification for building thermal systems under the presence of unmeasured disturbances in closed loop operation: Lumped disturbance modeling approach 107 (2016) 169–180. <https://doi.org/10.1016/j.buildenv.2016.07.007>
- [54] D. Kim, J.E. Braun, J. Cai, D.L. Fugate, Development and experimental demonstration of a plug-and-play multiple RTU coordination control algorithm for small/medium commercial buildings 107 (2015) 279–293. <https://doi.org/10.1016/j.enbuild.2015.08.025>
- [55] D. Kim, J.E. Braun, Development, implementation and performance of a model predictive controller for packaged air conditioners in small and medium-sized commercial building applications 178 (2018) 49–60. <https://doi.org/10.1016/j.enbuild.2018.08.019>
- [56] S.W. Ham, L. Paul, D. Kim, M. Pritoni, R. Brown, J. Feng, Decarbonization of heat pump dual fuel systems using a practical model predictive control: Field demonstration in a small commercial building 361 (2024) 122935. <https://doi.org/10.1016/j.apenergy.2024.122935>
- [57] S.W. Ham, D. Kim, L. Paul, Design and experimental performance of practical MPC for multi-zone VRF system for small and medium commercial buildings, in: *International High Performance Buildings Conference*, 2024. <https://docs.lib.purdue.edu/ihpbc/474>
- [58] L. Tian, J. Mees, V. Chan, W. Dean, Commercial adoption readiness assessment tool (CARAT), 2023. https://www.energy.gov/sites/default/files/2023-06/CARAT-R10_6-2-23.pdf
- [59] S.W. Ham, L. Paul, A. Casillas, A. Prakash, D. Kim, R. Brown, M. Pritoni, P. Grant, Practical challenges of model predictive control (MPC) for grid-interactive small and medium commercial buildings, in: *Proceedings of the 2024 ACEEE Summer Study on Energy Efficiency in Buildings*, American Council for an Energy-Efficient Economy, 2024 <https://escholarship.org/uc/item/9jv9p8rj>
- [60] B. Boyle, PYGLPK: updated fork of T. finley's PyGLPK module, 2014, <https://github.com/bradfordboyle/pyglpk>
- [61] D. Kim, J. Cai, J.E. Braun, K.B. Ariyur, System identification for building thermal systems under the presence of unmeasured disturbances in closed loop operation: theoretical analysis and application 167 (2018) 359–369. <https://doi.org/10.1016/j.enbuild.2017.12.007>
- [62] H. Madsen, J.M. Schultz, Short Time Determination of the Heat Dynamics of Buildings, Technical University of Denmark, Department of Civil Engineering, 1993 <p>BYG-rapport; Report no. 243 </p>
- [63] S.W. Ham, D. Kim, Field demonstration of low-cost and scalable MPC at a K-12 school and practical challenges 129 (2023) 857–864. Name - department of energy; copyright - copyright American society of heating, refrigeration and air conditioning engineers, Inc. 2024; last updated - 2024-06-06; SubjectsTermNotLitGenreText - United States-US, <https://www.proquest.com/scholarly-journals/field-demonstration-low-cost-scalable-mpc-at-k-12/docview/2973217852/se-2?accountid=14496>
- [64] R.N. Edmondson, Multi-level block designs for comparative experiments 25 (2020) 500–522. <https://doi.org/10.1007/s13253-020-00416-0>
- [65] A. Gelman, Bayesian data analysis, Chapman & Hall/CRC Texts in Statistical Science, CRC Press, third edition edition, 2014
- [66] Wikipedia contributors, Feature Scaling, 2024, https://en.wikipedia.org/w/index.php?title=Feature_scaling&oldid=1241939671
- [67] O. Abril-Pla, V. Andreani, C. Carroll, L. Dong, C.J. Fonnesbeck, M. Kochurov, R. Kumar, J. Lao, C.C. Luhmann, O.A. Martin, M. Osthege, R. Vieira, T. Wiecki, R. Zinkov, PyMC: a modern, and comprehensive probabilistic programming framework in Python 9 (2023) e1516. <https://doi.org/10.7717/peerj-cs.1516>
- [68] Q. Huangfu, J.A.J. Hall, Parallelizing the dual revised simplex method 10 (2018) 119–142. <https://doi.org/10.1007/s12532-017-0130-5>