Energy Conversion and Management 340 (2025) 119927

Contents lists available at ScienceDirect

Energy Conversion and Management

journal homepage: www.elsevier.com/locate/enconman

Research Paper

Optimizing building energy systems for grid-interactivity, comfort and resilience

Wanfu Zheng [a,b], Ziqi Hu [a,b], Dan Wang [a,c], Zhe Wang [a,b,*]

[a] Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Hong Kong, China
[b] HKUST Shenzhen-Hong Kong Collaborative Innovation Research Institute, Futian, Shenzhen, China
[c] Beijing Key Lab of Heating, Gas Supply, Ventilating and Air Conditioning Engineering, Beijing University of Civil Engineering and Architecture, Beijing, China

ARTICLE INFO

Dataset link: https://github.com/wfzheng/2nd-place-solution-neurips-citylearn2023-control

Keywords:
Community energy management
Model predictive control
Transfer learning
Demand response
Decarbonization

ABSTRACT

With the proliferation of renewable energy sources such as solar photovoltaics, managing the complexity of building energy systems while ensuring grid stability, occupant comfort, and resilience to power outages has become increasingly challenging. To address this challenge, this study proposes a hierarchical control framework that optimally coordinates battery energy storage, heat pumps, and domestic hot water (DHW) systems across multiple residential buildings. Forecasting models for disturbances are developed using linear regression, k-nearest neighbors regression, and LightGBM. At the building level, a data-driven model predictive control (MPC) strategy optimally regulates heat pump operations to ensure occupant comfort, complemented by a rule-based controller for DHW storage scheduling. At the microgrid level, a physics-based MPC dispatches battery energy to achieve grid-level objectives such as peak shaving and emission reduction. Coordination between the two levels is achieved through a bottom-up structure: building-level controllers estimate their future electricity demand, which is passed as a disturbance input to the upper-level battery dispatch optimization. The framework performed effectively in the 2023 NeurIPS CityLearn Challenge, securing second place overall and achieving the best performance in public buildings across comfort, emissions, grid efficiency, and resilience metrics. This work provides an effective solution for community-scale energy management, emphasizing the importance of multi-level coordination between building systems and microgrids to support sustainable and resilient energy operations. Source code are available at: https://github.com/wfzheng/2nd-place-solution-neurips-citylearn2023-control

1. Introduction

Buildings account for a substantial portion of global energy consumption, representing about 34% of total energy demand and 37% of energy-related CO_2 emissions [1]. In buildings, heating, ventilation, and air conditioning (HVAC) systems are the main contributors to the overall energy use [2]. However, they also offer substantial opportunities for energy savings if operated optimally. As renewable resources such as photovoltaics (PV) and wind energy become more prevalent, effectively managing their intermittency while maintaining occupant comfort remains a significant challenge. At the same time, recent perspectives on grid-efficient efficient buildings underscore a shift from passive energy consumption to active grid participation through integrated, responsive, and bidirectional energy management [3].

Demand Response (DR) programs have emerged as key strategies to address these complexities [4]. By incentivizing end users to adjust their energy consumption patterns, DR facilitates load change, balances energy supply and demand, and improves the reliability of the power system [5]. Recent studies indicate that medium-scale commercial buildings (e.g. hotels) with peak daytime loads can substantially benefit from DR and photovoltaic battery microgrids, achieving reductions in cost and peak demand despite battery degradation and comfort constraints [6]. Advances in control technologies further reinforce the role of DR in building energy efficiency and flexibility. Traditional open-loop control schemes, such as Rule-Based Control (RBC), are widely used for their simplicity, robustness, and fast real-time execution. However, these approaches often rely on predefined schedules and heuristic rules derived from expert knowledge without adjusting decisions based on real-time feedback from system states or external disturbances [7]. As energy systems grow more complex and data-intensive, adaptive and intelligent control methods such as Model Predictive Control (MPC) and Reinforcement Learning (RL) are gaining momentum because they can optimize system performance under uncertainty. MPC predicts

* Corresponding author at: Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Hong Kong, China.
E-mail address: cezhewang@ust.hk (Z. Wang).

https://doi.org/10.1016/j.enconman.2025.119927
Received 19 February 2025; Received in revised form 15 April 2025; Accepted 8 May 2025
Available online 29 May 2025

Nomenclature

Variables, parameters, and indices

$\mathcal{I}$	Index set of buildings.
$\mathcal{T}$	Index set of time steps.
T^{set}	Indoor temperature setpoint.
T_{init}	The initial measured indoor temperature.
$\widetilde{T}$	Controlled indoor temperature.
u, Q^{hp}	Cooling demand of the heat pump.
a	Actual control command to the heat pump.
d	Disturbances.
COP	Coefficient of performance of the heat pump.
$P^{\text{hp}}_{\text{nom}}$	Nominal power rating of the heat pump.
η^{tech}	Technical efficiency of the heat pump.
B^{em}	Base of carbon emission.
B^{grid}	Base of grid ramping cost.
CI	Carbon intensity.
w^{factor}	Weight factor.
$P^{\text{bat}}_{\text{nom}}$	Nominal power rating of the battery.
C^{bat}	Capacity of the battery.
L	Load of the building.
E^{thres}	Threshold value for consumption.
$E^{\text{init\_agg}}$	Initial aggregate consumption.
$S^{\text{bat}}_{\text{init}}$	Initial state of charge of the battery.
S^{dod}	Depth-of-discharge related SOC limit of the battery.
$\eta^{\text{bat,ch}}$	Charging efficiency.
$\eta^{\text{bat,dis}}$	Discharging efficiency.
$\eta^{\text{bat,rt}}$	Round-trip efficiency factor.
x	Charging action of the battery.
y	Discharging action of the battery.
E^{net}	Net electricity consumption from the grid.
E^{ns}	Non-shiftable load.
E^{pv}	Solar generation.
E^{hp}	Cooling electricity consumption.
E^{dhw}	DHW electricity consumption.
E^{agg}	Aggregate energy consumption (sum of all buildings).
E^{max}	Maximum total consumption over the control horizon.
E^{over}	Amount of consumption exceeding E^{thres}.
W	Grid ramping.
E^{avg}	Average total consumption over the control horizon.
S^{bat}	State of charge of the battery.
$S^{\text{bat, 1}}$, $S^{\text{bat, 2}}$	Sub-components of S^{bat} (needed for piecewise modeling).

future system states over a specified horizon and iteratively solves an optimization problem. Through a receding horizon principle, MPC generates an optimal control sequence, implements the first control action and re-solves the optimization problem at each time step using updated system states and forecasts [8]. Multiple objectives such as minimizing energy costs, reducing carbon emissions, and improving grid efficiency—can be incorporated into MPC by customizing the objective functions [9]. In [10], the authors compare an open-loop control strategy with the MPC approach. Their results show that MPC significantly increases renewable energy penetration and reduces energy curtailment, highlighting the advantages of closed-loop optimization that incorporates forecasting into real-time decision-making. Meanwhile, RL emerges as a powerful model-free approach for optimal control. A recent review [11] highlights that model-based RL offers greater generalization and training stability when accurate physical models are available, such as for grid voltage control, whereas model-free RL provides more flexibility in highly uncertain contexts (e.g., occupant behavior or electricity pricing). However, challenges in interpretability, training safety and multi-agent coordination remain significant barriers to practical deployment.

The effectiveness of MPC has been demonstrated across various building energy systems and configurations. For instance, in a commercial building equipped with onsite PV and batteries, Zhang et al. [12] employed MPC to manage electricity usage and storage, resulting in a 12% annual cost reduction and a 34% decrease in peak demand. In a grid-connected residential household, Wamalwa et al. [13] adopted MPC to minimize grid import costs and battery wear, yielding a 42% decrease in annual grid energy expenditures compared to a rule-based controller, with an estimated payback period of seven years. In an off-grid rural clinic powered by a photovoltaic-diesel-battery system, Zhu et al. [14] introduced a switched MPC framework that incorporates online battery efficiency estimation and switching constraints, effectively minimizing diesel fuel consumption, particularly during summer solar

profiles and showing greater resilience to disturbances compared to heuristic or open-loop methods.

Recent progress in machine learning has further widened MPC's capabilities by enabling data-driven control strategies. In contrast to conventional MPC methods based on physical models (e.g., resistance–capacitance or RC thermal models), data-driven MPC leverages machine learning algorithms to forecast system dynamics. Aruta et al. [15] developed an MPC framework powered by artificial neural networks (ANN) using a nonlinear autoregressive network (NARX) model for predicting indoor temperature and heating load. With the aid of genetic algorithms (GA) for optimization, this framework delivered notable energy savings, including heating cost reductions of up to €41.22 over a 120-day heating season relative to traditional methods. Stoffel et al. [16] showed that MPC using ANN-, Gaussian Process Regression (GPR)-, and Multiple Linear Regression (MLR)-based system models can be swiftly deployed and improved via online learning. While the MLR-based MPC achieved the highest computational efficiency, the ANN-based model had the most accurate predictions, although the performance gap during MPC operations was relatively small. Zheng et al. [17] compared a standard RC-based MPC with an ANN-based counterpart, demonstrating comparable energy efficiency and comfort outcomes. Although physical, RC-based MPC tends to be computationally efficient and requires minimal training data, ANN-based MPC offers greater flexibility and demands less domain-specific knowledge, albeit at the cost of requiring ample data for robust generalization.

Distributed energy resources (DERs), such as batteries, domestic hot water (DHW) storage, and heat pumps play a critical role in achieving operational flexibility. Batteries can store excess solar energy and reduce peak loads [18], DHW storage can serve as a thermal buffer [19], and heat pumps can shift or modulate heating and cooling demands [20]. Realizing the full potential of these technologies depends on accurate forecasts of exogenous variables such as weather conditions and building loads. For instance, Peter et al. [21] reported that weather forecast errors led to a 13.3% rise in energy costs under an MPC regime, whereas bias correction mitigated much of this impact. Underscoring the significance of load forecasting, Li et al. [22] explored how inaccuracies in building load predictions affect PV-battery energy storage system optimization, concluding that underestimating loads is more detrimental than overestimation. Forecast uncertainty is similarly crucial during the design phase of building energy systems [23], where decisions about component selection, system sizing, and configuration are made. In this vein, Hu et al. [24] proposed an uncertainty-aware design approach for zero energy buildings, employing Monte Carlo simulation and a grid interaction metric to capture variability in cooling demand and renewable generation; compared to deterministic design, their method reduced costs by 9%, cut emissions by 6.8%, and improved grid stability. A number of studies have also examined how MPC can optimize various DER components. Joe et al. [25] designed an MPC approach for hydronic radiant floor heating and cooling systems, realizing 34% cost savings during cooling periods and a 16% decrease in energy use during heating seasons. Deng et al. [26] employed dynamic programming and mixed-integer linear programming to optimize electrical chillers and thermal energy storage in campus cooling plants; their linearization strategy boosted computational efficiency, reducing electricity costs by 9.70%. Ra et al. [27] introduced a deep neural network (DNN)-based MPC for a factory HVAC system, lowering cooling energy use by 35.1% with no loss of occupant comfort. Grid-interactive efficient buildings incorporating DERs, such as PV arrays and battery storage, have similarly gained prominence for their capacity to optimize energy use and enhance grid stability. For example, Deepranjan et al. [28] developed an MPC-based battery scheduling mechanism that effectively accounted for forecast uncertainties and hardware constraints to decrease peak loads. Meanwhile, Abraham et al. [29] explored a community battery dispatch strategy using MPC under varying forecast horizons, revealing that even shorter control horizons can yield meaningful monthly cost savings, ranging from $11.47 using synchronous average predictions to $67.91 with perfect predictions—compared to a baseline electricity bill of $227.35.

Despite notable progress in MPC and DER integration, current energy management strategies rarely unifiedly address grid interactivity, occupant comfort, and resilience. Existing studies tend to focus on one or two of these aspects at a time. Many studies [12,15,16,25–27,29,30] concentrate on single buildings or subsystems (e.g., HVAC or PV-battery systems) and do not provide unified frameworks for coordinating multiple subsystems simultaneously. Regarding occupant comfort, traditional approaches [17] commonly maintain indoor temperatures within prescribed ranges, such as 21–24 °C when occupied and 15–30 °C when unoccupied. It simplifies reality and ignores the variability of occupant behavior. Moreover, these approaches generally presume that accurate thermal models of buildings are readily available without exploring how controllers can be adapted for new buildings with unknown thermal characteristics. While energy efficiency remains a primary objective, the emergence of grid-interactive operations demands broader considerations, such as regulating ramping rates and peak demand without sacrificing occupant comfort. Few studies have addressed how to integrate these requirements across multiple buildings under diverse conditions, including grid outages and frameworks that accommodate both normal and emergency operating modes remain scarce. To address these issues, this paper introduces and evaluates a multi-layered MPC framework for a series of residential buildings equipped with diverse energy resources. The key contributions are summarized as follows:

- **Multi-layer integrated control framework.** We propose a two-layer system that combines conventional MPC, data-driven MPC, and RBC. MPC and RBC strategies manage HVAC operations at the building level, while a centralized MPC optimizes battery dispatch at the microgrid level. Our design achieves significant peak-load reduction and improved energy efficiency compared to conventional approaches targeting a single device or building.
- **Data-driven MPC and transfer learning.** An LSTM-based thermal dynamics model is incorporated into the MPC scheme through gradient-based optimization. We employ transfer learning to enhance adaptability by fine-tuning thermal models trained on known buildings, enabling rapid deployment in new buildings with similar characteristics.
- **Enhanced resilience measures.** We address unexpected grid outages by maintaining a higher minimum battery state of charge (SOC), adjusting indoor temperature setpoints dynamically, and prioritizing critical loads. These strategies ensure essential services and occupant comfort, even under adverse grid conditions.
- **Empirical validation in the NeurIPS CityLearn Challenge 2023.** The experimental results demonstrate strong performance in improving grid efficiency, reducing emissions and managing outages. The proposed framework achieved second place in the NeurIPS CityLearn Challenge 2023. These outcomes offer practical insights supporting its potential for real-world deployment. To facilitate further research, we have also made our implementation code open-source, which is accessible at: https://github.com/wfzheng/2nd-place-solution-neurips-citylearn2023-control.

This paper is organized as follows: Section 2 overviews the system and problem statement. Section 3 presents the overall control framework, Section 4 presents and discusses the results during the testing period, and Section 5 summarizes key findings and explores future research directions.

2. System description and problem statement

The community is established using CityLearn [31], an open-source OpenAI Gym environment designed for DR and urban energy management research, supporting benchmarks for RBC, MPC, and Reinforcement Learning Control (RLC) while incorporating stochastic power

outages modeled on U.S. distribution-system reliability data. The community is represented by a neighborhood of six distinct single-family dwellings, as shown in Fig. 1. The neighborhood includes six single-family dwellings interconnected through a microgrid that interfaces with the local power grid. All component models used in this study, including building thermal dynamics, heat pumps, batteries, PV systems, electric heaters, and domestic hot water (DHW) storage tanks, are predefined by the CityLearn environment. These models are developed and validated in the work by Vazquez-Canteli et al. [31] and are summarized in Appendix A for reference. Fig. 2 depicts the energy flow within a single-family home energy management system. The microgrid and PV system jointly supply power, with the battery capable of charging and discharging electricity, and the heat pump uses electrical energy to generate cooling loads, providing chilled air through a fan coil unit. The electric heater supplies heat to meet the DHW demand, ensuring the availability of hot water either directly or by discharging from the DHW tank. Various non-shiftable loads, such as lighting, refrigerators, and ventilation systems, consume electrical energy from the distributed energy system. In each home energy management system, we utilize one electrical energy buffer and two thermal energy buffers. The battery represents the electrical energy buffer, the built-in thermal mass of the cooling system, and the DHW storage tank serves as the thermal energy buffer. The net electricity consumption of the building from the grid can be expressed as follows:

$$E^{\text{net}} = E^{\text{ns}} - E^{\text{pv}} + E^{\text{hp}} + E^{\text{bat}} + E^{\text{dhw}} \tag{1}$$

where E^{net} denotes the net electricity consumption of the building from the grid, E^{ns} indicates the electricity demand of non-shiftable loads in the building, E^{hp} represents the electricity used by the heat pump, which contributes to the cooling demand of the building, E^{pv} refers to the electricity generated by the PV system, E^{bat} represents the net electricity of the battery, where a positive value indicates that the battery is charging, and a negative value indicates discharging. E^{dhw} represents the electricity consumed to meet the hot water demand, which can be provided directly by the electric heater, from the DHW storage, or a combination of both, as illustrated in the upper-right corner of Fig. 2.

The CityLearn Challenge 2023 requires participants to design single- or multi-agent control policies for managing battery systems, DHW storage, and heat pumps in six single-family residential buildings. It aims to achieve four primary objectives: (i) maintaining thermal comfort, (ii) reducing CO_2 emissions, (iii) promoting grid efficiency, and (iv) preserving resilience during power outages. Performance is measured through four cost components: comfort (S_c), emission (S_e), grid cost (S_g), and resilience (S_r), and each is normalized against a no-control baseline, with a weighted sum yielding the total control score:

$$\text{Score} = 0.3S_c + 0.1S_e + 0.3S_g + 0.3S_r, \tag{2}$$

S_c quantifies how often indoor temperatures deviate from occupant comfort bands during occupied hours; S_e is calculated by multiplying the controller's net electricity consumption E^{net} by the carbon intensity CI and normalized against the baseline emissions. S_g is a composite indicator capturing ramping, load factor, daily peak, and all-time peak. S_r assesses the system's ability to maintain temperature setpoints and meet critical loads during stochastic outages. It should be noted that the weights are determined by the competition organizers to ensure consistency in evaluating the strategies. In practical applications, the selection of appropriate weights may need to take into account the current specific context or user preferences. For details on how the metrics are calculated, please refer to Appendix A.

Beyond these four overarching objectives, another important aspect of this research is to explore how thermal dynamics models can be transferred to others that lack explicit thermal models. In the competition, Buildings 1, 2, and 3 possess known thermal models, while Buildings 4, 5, and 6 do not. During the training phase, the controller is developed using a one-month dataset and evaluated on a separate dataset covering a three-month period. The testing dataset is derived from distinct weather profiles spanning those three months. As a result, the submitted agent must be capable of generalizing to previously unseen weather conditions, occupant behaviors, and building characteristics. Each dataset contains 31 state variables at each hourly time step (summarized in Table 1). The objective is to enable effective control across all six buildings without individually modeling the thermal dynamics of each one, thereby facilitating controller generalization and reducing time-consuming system identification efforts for buildings with unknown thermal parameters. The developed controller is first assessed in the three buildings with known thermal dynamics, after which its overall performance across all six buildings is evaluated.

3. Methodology

3.1. Overview

The control framework integrates several modules to optimize building energy management, as shown in Fig. 3. The left part is the prediction module, which consists of three machine learning models: linear regression, K nearest neighbors (KNN) regression, and LightGBM [32]. These models predict key variables for the next 24 h, aligning with the control horizon of MPC controllers, and include factors such as non-shiftable load, indoor temperature setpoint, carbon intensity, solar radiation, occupant count, etc. Our choice of linear regression, KNN regression, and LightGBM balances predictive accuracy and computational efficiency for real-time control. We employed linear regression to forecast weather-related disturbances (e.g., outdoor temperature and solar radiation) and occupant count, as it delivers decent accuracy while remaining simple and computationally efficient. To predict indoor temperature setpoints, we initially used a KNN regressor during the first week of deployment, as only limited historical data were available to build features. KNN effectively captures short-term patterns under sparse data by matching current conditions with similar past instances. Once sufficient data were accumulated to enable richer feature construction, we transitioned to LightGBM, which was also employed to forecast non-shiftable loads and PV generation. LightGBM was chosen due to its strong empirical performance in various Kaggle competitions and real-world applications. Its gradient-boosting framework handles nonlinear relationships effectively while offering efficient training. We also explored an LSTM-based approach but found that it required extensive hyperparameter tuning. In our tests, it was less flexible in incorporating large input features and underperformed compared to LightGBM. Moreover, to enhance the adaptability of the forecasting models and ensure their long-term robustness, we implement a dynamic data buffer that retains the most recent real-time observations, as shown at the bottom of Fig. 3. The buffer is a continuously evolving dataset, enabling periodic fine-tuning of the forecasting models during deployment. By incorporating newly acquired data into the training cycle, the models undergo self-adaptive updates, thereby sustaining high predictive accuracy. This mechanism allows the control system to maintain robust performance amid temporal variability in seasonal trends, weather conditions, and occupant behaviors.

The control module utilizes the predictions, which operate at two hierarchical levels: the HVAC system level and the microgrid level. At the HVAC system level, each building has a local controller with a data-driven MPC for managing the heat pump and an RBC for DHW storage. The data-driven MPC employs a neural network to model indoor thermal dynamics and uses gradient-based optimization to regulate the heat pump, ensuring thermal comfort. The RBC for DHW storage follows a time-based strategy to shift heating loads to off-peak periods, thus contributing to peak shaving. At the microgrid level, a centralized physics-based MPC controller dispatches the community's battery energy storage. This controller relies on individual buildings' estimated 24-hour electricity consumption profiles, which are generated by the

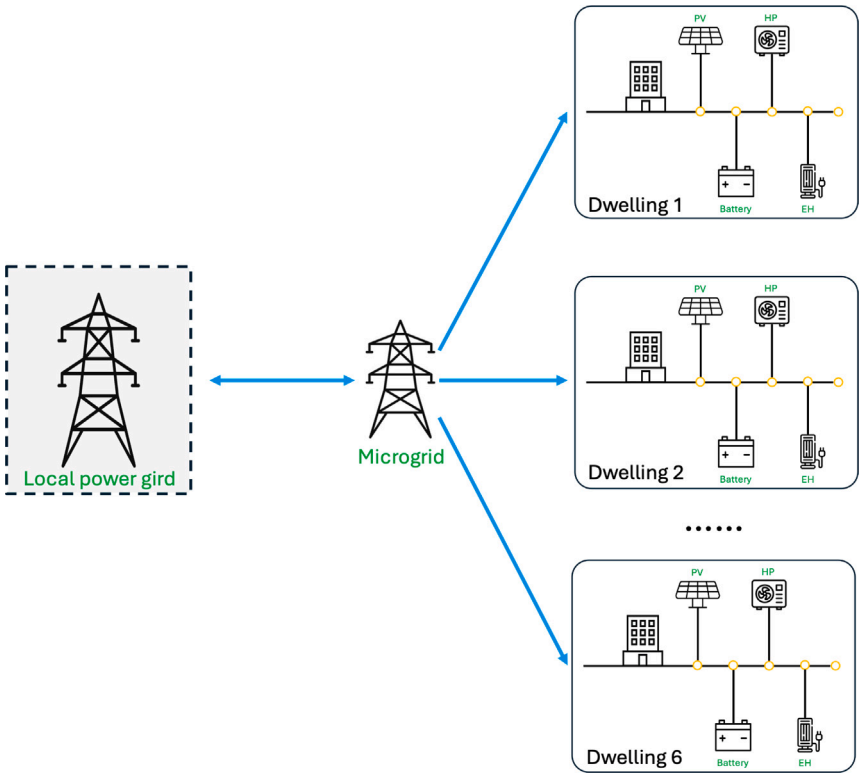

Fig. 1. Diagram illustrating the energy community modeled in CityLearn.

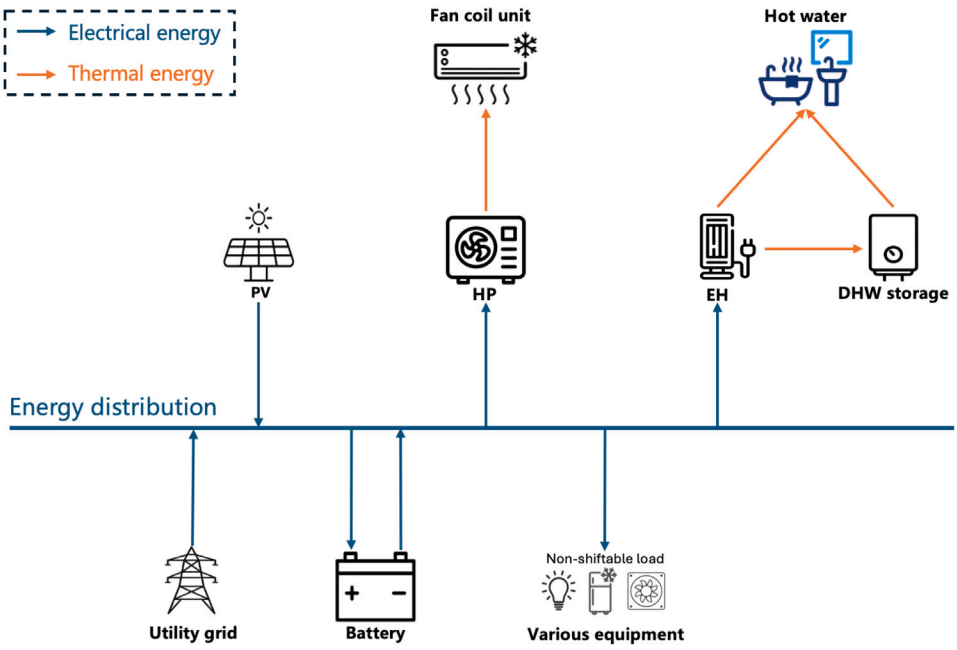

Fig. 2. Energy flow for a single-family home energy management system.

building-level MPC and RBC modules. These consumption profiles, including the projected power usage of heat pumps and DHW systems, are treated as disturbance inputs in the upper-layer optimization problem.

The coordination mechanism works in a bottom-up manner. The bottom layer (HVAC) first optimizes operations and outputs preliminary control decisions and forecasts. The top layer (microgrid) then

Table 1
Summary of state variables and their units for each home energy management system.

Symbol	Description	Units
M	Month	–
D	Day of week (1=Monday, 2=Tuesday, …, 7=Sunday)	–
H	Hour	–
T^{out}	Outdoor dry bulb temperature	°C
$T^{\text{out, 6h}}$	Outdoor dry bulb temperature predicted 6 h ahead	°C
$T^{\text{out, 12h}}$	Outdoor dry bulb temperature predicted 12 h ahead	°C
$T^{\text{out, 24h}}$	Outdoor dry bulb temperature predicted 24 h ahead	°C
I^{dif}	Diffuse solar irradiance	W/m^2
$I^{\text{dif, 6h}}$	Diffuse solar irradiance predicted 6 h ahead	W/m^2
$I^{\text{dif, 12h}}$	Diffuse solar irradiance predicted 12 h ahead	W/m^2
$I^{\text{dif, 24h}}$	Diffuse solar irradiance predicted 24 h ahead	W/m^2
I^{dir}	Direct solar irradiance	W/m^2
$I^{\text{dir, 6h}}$	Direct solar irradiance predicted 6 h ahead	W/m^2
$I^{\text{dir, 12h}}$	Direct solar irradiance predicted 12 h ahead	W/m^2
$I^{\text{dir, 24h}}$	Direct solar irradiance predicted 24 h ahead	W/m^2
CI	Carbon intensity	kg CO_2/kWh
$CI^{6\text{h}}$	Carbon intensity predicted 6 h ahead	kg CO_2/kWh
$CI^{12\text{h}}$	Carbon intensity predicted 12 h ahead	kg CO_2/kWh
$CI^{24\text{h}}$	Carbon intensity predicted 24 h ahead	kg CO_2/kWh
E^{ns}	Non-shiftable load	kWh
E^{pv}	Solar generation from PV panels	kWh
Q^{hp}	Cooling demand	kWh
Q^{dhw}	Domestic hot water demand	kWh
E^{net}	Net electricity consumption	kWh
S^{dhw}	SOC of DHW storage tank	–
S^{bat}	SOC of battery	–
N^{occ}	Occupant count	–
T	Indoor dry bulb temperature	°C
T^{set}	Indoor dry bulb temperature setpoint	°C
δ^{out}	Binary indicator for power outage (1=outage, 0=no outage)	–

aggregates and evaluates these building-level profiles and computes optimal battery charging/discharging schedules to achieve community-wide goals such as peak shaving and carbon emission reduction. This iterative and hierarchical coordination ensures that local control objectives are harmonized with global targets, enabling integrated and efficient operation across the entire community.

3.2. Data-driven MPC for heat pump

3.2.1. Mathematical formulation

The optimization problem is designed to track the indoor air temperature setpoints:

$$\min_{u_t} \sum_{t=1}^{24} (\widetilde{T}_t - T_t^{set})^2 \tag{3a}$$

$$\mathcal{T} = \{1, \dots, 24\}, \tag{3b}$$

$$\text{s.t. } \widetilde{T}_0 = T_{init}, \tag{3c}$$

$$\widetilde{T}_t^{in} = f_{lstm}\left(\widetilde{T}_{t-1}, u_t, d_{t-1}\right), \quad \forall t \in \mathcal{T}, \tag{3d}$$

$$\underline{u}_t \le u_t \le \bar{u}_t, \quad \forall t \in \mathcal{T}, \tag{3e}$$

where T_t^{set} denotes the predicted indoor temperature setpoint, which is an exogenous variable predicted based on past setting habits of occupants; $\widetilde{T}_t$ is the controlled indoor temperature, considered as an endogenous variable; u_t represents the cooling demand serves as the control variable. The descriptions for each equality and inequality constraint, along with the necessary explanation of parameters, are provided as follows:

- Eq. (3c) sets the current indoor temperature as T_{init}.
- Eq. (3d) describes the evolution of the indoor temperature ($\widetilde{T}_t$) using an LSTM-based function f_{lstm}, which takes the controlled temperature ($\widetilde{T}_{t-1}$), the control variable (u_t), and disturbances (d_{t-1}) as inputs, based on the thermal dynamics of the building envelope introduced in Appendix A
- Eq. (3e) imposes constraints on the control inputs.

It is worth noting that $\underline{u}_t = 0$, and the maximum cooling demand $\bar{u}_t$ is not a fixed constant. It varies with the outdoor temperature at each time step. The maximum cooling demand of the heat pump can be calculated using the following equations:

$$COP_t = \eta^{\text{tech}} \cdot \frac{T_{\text{target}}}{T_t^{\text{out}} - T_{\text{target}}} \tag{4}$$

$$\bar{u}_t = COP_t \cdot P_{\text{nom}}^{\text{hp}} \tag{5}$$

First, we calculate the cooling performance coefficient COP_t, where η^{tech} is the technical efficiency of the heat pump; T_t^{out} indicates the outdoor temperature; T_{target} is the target cooling temperature. The maximum cooling demand at time step t is then determined as the product of the COP and the nominal power P_{nominal} of the heat pump. The first action of the optimal control sequence is then applied to control the heat pump. The optimized cooling demand is converted into the actual control command a_t for the heat pump as follows:

$$a_1 = \frac{u_1}{COP_1 \cdot P_{\text{nom}}^{\text{hp}}} \tag{6}$$

To provide the microgrid-level battery dispatch MPC with an accurate estimate of the heat pump's electricity consumption, we calculate the cooling electricity consumption E_t^{hp} as follows:

$$E_t^{\text{hp}} = \frac{u_t}{COP_t}, \quad \forall t \in \mathcal{T}, \tag{7}$$

This yields a 24-hour electricity consumption profile for the heat pump, which is then passed to the microgrid-level control as a disturbance input. To optimize our objective function, we employ the Projected Stochastic Gradient Descent (PSGD) algorithm, a variant of the standard Stochastic Gradient Descent (SGD) optimizer, based on the non-convex characteristics of the LSTM model. By incorporating Projected Gradient Descent (PGD), we ensure that the control inputs satisfy the constraints. After each gradient update, the control vector $\boldsymbol{u}_k$ is projected onto the feasible region $\mathcal{Q}$ using the projection operation shown in Eq. (8), which minimizes the Euclidean distance between the control vector $\boldsymbol{u}_k$ and its projection $\boldsymbol{u}_k^p$.

$$\boldsymbol{u}_k^p = \underset{\boldsymbol{u} \in \mathcal{Q}}{\operatorname{argmin}} \frac{1}{2} \left\| \boldsymbol{u} - \boldsymbol{u}_k \right\|_2^2 \tag{8}$$

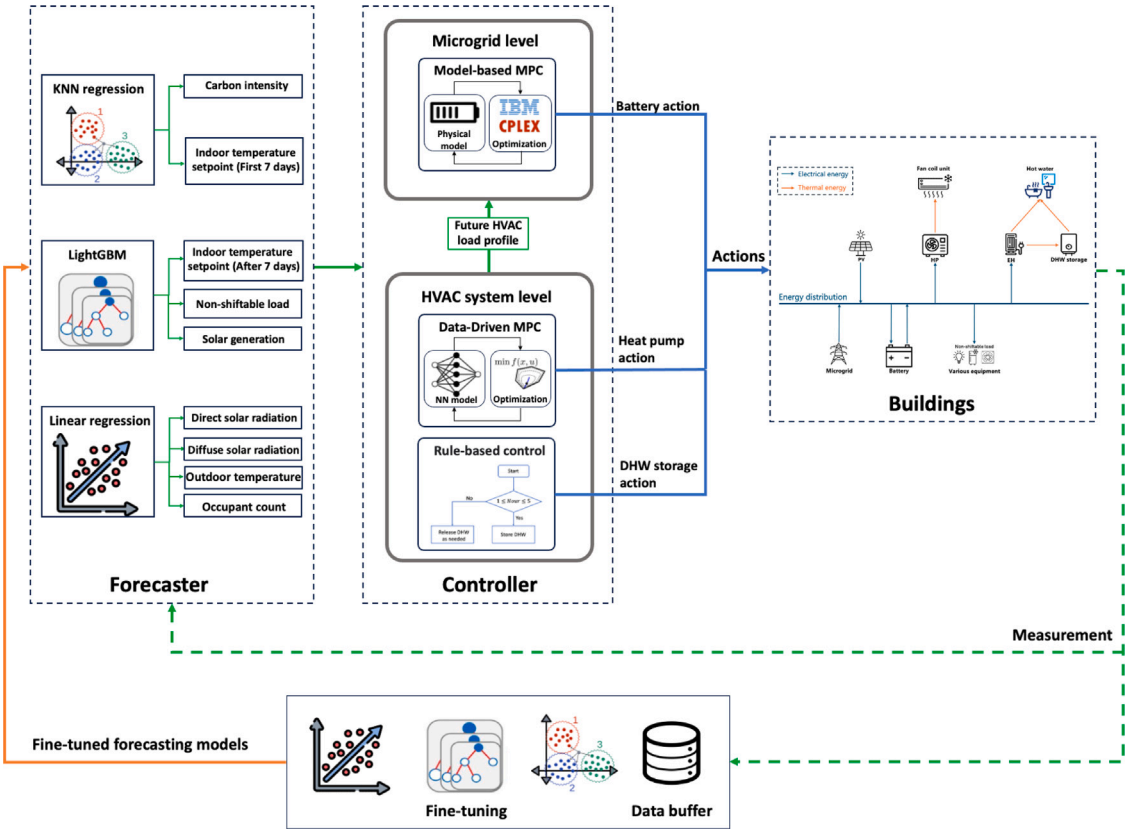

Fig. 3. Optimal control framework for energy management in community buildings.

A transfer learning strategy via fine-tuning is employed for the thermal dynamics of three similar private buildings whose underlying models are unknown. A base model is initially developed using data from one publicly available building, capturing generic thermal dynamics with lower LSTM layers. During the testing phase, real-time data collected from the private buildings adapt the model to their specific thermal behaviors. As illustrated in Fig. 4, the fine-tuning process involves freezing the lower LSTM layers, which learn the fundamental thermal dynamics, to retain the generic knowledge, while the fully connected layer is updated with real-time data. This approach enables the model to maintain robust generalization capabilities while efficiently accommodating the unknown thermal characteristics of private buildings.

3.2.2. Prediction models

For the MPC scheme delineated in Eqs. (3), the prediction of diverse exogenous condition variables is required across the 24-hour control horizon. Table 2 provides details on these variables, the corresponding prediction models, and input features, including *Weekday*, which is represented as a binary variable (1 for weekdays, 0 for non-weekdays), and *Office hour*, defined as a binary indicator (1 for the period from 9:00 to 19:00, and 0 otherwise). Linear regression is selected for most external variables due to its relatively high accuracy and low computational cost, which make it suitable for real-time control scenarios. Predicting the indoor temperature setpoint is crucial, as occupant preferences directly determine the temperature the controller must achieve. The control system cannot optimally regulate cooling without knowing how occupants will adjust their setpoints. Consequently, a two-stage prediction strategy is adopted: KNN regression is used in early operation when only limited historical data are available, followed by a transition to LightGBM after sufficient data accumulation, taking advantage of its ability to handle larger feature sets and deliver higher predictive accuracy. To accommodate the 24-hour predictive horizon, a direct multi-step forecasting approach is employed, as illustrated in Fig. 5. Separate models are trained for each step (e.g., $T_{t+1}, T_{t+2}, \ldots, T_{t+24}$), allowing the algorithm to directly forecast long-range outcomes without relying on recursive predictions.

3.3. RBC for DHW storage

A rule-based control strategy is developed for DHW storage in buildings based on analyzing non-shiftable load data from three public buildings. Due to the difficulty in accurately predicting hot water usage, a simple time-based rule set was adopted. The control objective for the DHW storage is to minimize peak load and smooth grid demand. As shown in Fig. 6, the hourly average electric power demand (depicted by the blue points) and its variance (represented by the gray shading) illustrate the fluctuations in demand throughout the day. Load analysis showed that electricity demand was consistently low between 1:00 AM and 5:00 AM, making this period ideal for charging the DHW storage. To utilize this period, the SOC of the DHW storage is increased by 0.2 units per hour. The increment value of 0.2 was chosen to evenly distribute the charging over the five-hour low grid demand window, avoiding sudden spikes while gradually increasing the thermal storage. During peak hours (6:00 AM to 12:00 AM), the system prioritizes discharging stored hot water to cover demand, reducing grid reliance. If the stored hot water is insufficient, additional power is drawn. Mathematically, the rule is defined as:

$$a^{\mathrm{dhw}} = \begin{cases} 0.2, & \text{if } 1\text{ AM} \le t \le 5\text{ AM} \\ -1, & \text{otherwise} \end{cases} \tag{9}$$

where $a^{\mathrm{dhw}} = -1$ indicates that DHW should be released as needed to meet the demand, ensuring that the stored hot water is used efficiently. Although the rule is simple in design, it can effectively shift a portion of domestic hot water heating to off-peak hours, thereby helping to alleviate peak loads while balancing the overall load curve.

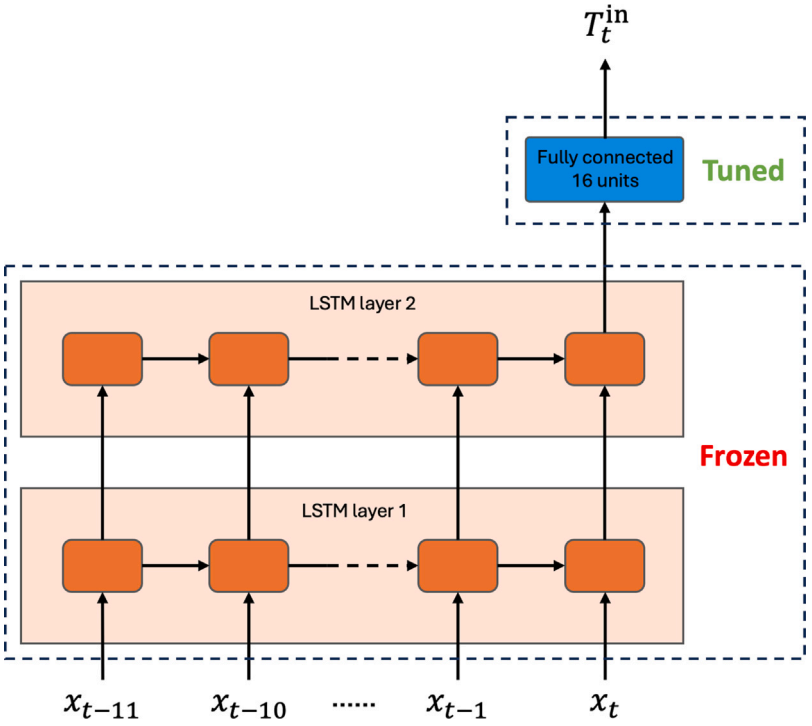

Fig. 4. A transfer learning technique for thermal dynamics modeling.

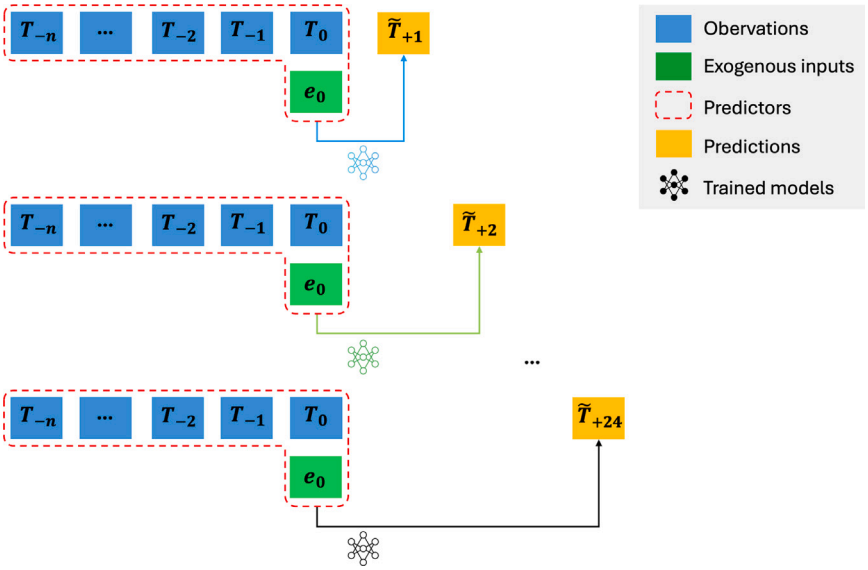

Fig. 5. Schematic representation of the direct multi-step forecasting approach. Separate models are trained to predict each step within the prediction horizon where T represents the target variable. Each model utilizes lagged observations, the current observation (T_0), and current exogenous inputs (e_0) for predictions.

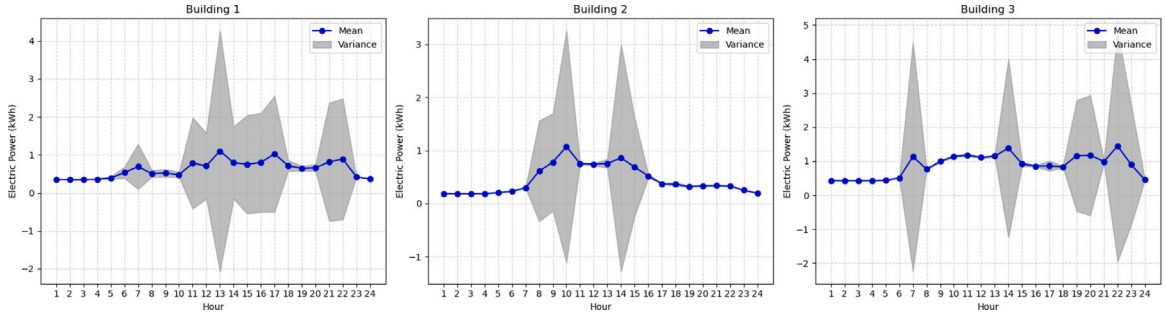

Fig. 6. Hourly average non-shiftable load (blue points) with variance range (gray shading) for three public buildings (Buildings 1, 2, and 3).

Table 2
Forecasting models and input features for exogenous variables used in the MPC-based control of heat pumps.

Target variables (24h prediction)	Model	Input features
I^{diff}	Linear regression	H, D T^{out} (current, 6h, 12h, 24h leads) I^{diff} (current, 6h, 12h, 24h leads, 1–24h lags) I^{dir} (current, 6h, 12h, 24h leads)
I^{dir}	Linear regression	H, D T^{out} (current, 6h, 12h, 24h leads) I^{diff} (current, 6h, 12h, 24h leads) I^{dir} (current, 6h, 12h, 24h leads, 1–24h lags)
T^{out}	Linear regression	H, D T^{out} (current, 6h, 12h, 24h leads, 1–24h lags) I^{diff} (current, 6h, 12h, 24h leads) I^{dir} (current, 6h, 12h, 24h leads)
T^{set}	KNN regression (First 7 days) LightGBM (After 7 days)	KNN regression: H, D, Weekday, Office hour, T^{set} (1-24h lags) LightGBM: H, D, Weekday, Office hour, T^{set} (1–168h lags)
N^{occ}	Linear regression	H, D, Weekday, Office hour, N^{occ} (1–24h lags)

3.4. MPC for battery dispatch

An MPC framework is formulated to govern the charging and discharging of batteries in each building over a 24-hour horizon. Let $\mathcal{I} = \{1, \dots, 6\}$ be the set of buildings and $\mathcal{T} = \{1, \dots, 24\}$ be the set of hourly time steps. The decision variables are the charging and discharging actions, $x_{i,t}$ and $y_{i,t}$. Endogenous state variables include the battery SOC, while exogenous parameters such as carbon intensity (CI_t) remain fixed inputs to the MPC. The cost function J sums four terms over the control horizon: carbon emissions, grid ramping cost, the penalty for exceeding a predefined threshold, and average energy usage. Mathematically:

$$J = \frac{1}{B^{\text{em}}} \cdot \sum_{i=1}^{6} \sum_{t=1}^{24} \left(E_{i,t}^{\text{net}} \cdot CI_t \cdot w_t^{\text{factor}} \right) + \frac{1}{B^{\text{grid}}} \cdot \left(W_1 + \sum_{t=2}^{24} W_t \cdot w_t^{\text{factor}} \right) + \left(E^{\text{over}} \right)^2 + E^{\text{avg}}, \tag{10}$$

where $E_{i,t}^{\text{net}}$ represents the net energy consumption of building i at time step t, CI_t is the predicted carbon intensity, W_t characterizes the power ramp from the grid, E^{over} quantifies the extra consumption beyond a defined threshold, and E^{avg} represents the average total consumption over the control horizon. The time-weighting factor w_t^{factor} prioritizes near-term decisions by assigning higher weights to earlier time steps, calculated as: $w_t^{\text{factor}} = 1.0 - \frac{1.0-0.95}{24} \cdot (t-1), \quad t \in \mathcal{T}$. The overall problem is formulated as follows:

$$\min_{x_{i,t}, y_{i,t}} \ J \tag{11a}$$

$$\mathcal{I} = \{1, \dots, 6\}, \quad \mathcal{T} = \{1, \dots, 24\}, \tag{11b}$$

s.t.
$$L_{i,t} = E_{i,t}^{\text{ns}} - E_{i,t}^{\text{pv}} + E_{i,t}^{\text{hp}} + E_{i,t}^{\text{dhw}}, \quad \forall i \in \mathcal{I}, t \in \mathcal{T}, \tag{11c}$$

$$E_{i,t}^{\text{net}} \geq x_{i,t} \cdot C_i^{\text{bat}} + y_{i,t} \cdot C_i^{\text{bat}} + L_{i,t}, \quad \forall i \in \mathcal{I}, t \in \mathcal{T}, \tag{11d}$$

$$E_t^{\text{agg}} \geq \sum_{i=1}^{6} \left(x_{i,t} \cdot C_i^{\text{bat}} + y_{i,t} \cdot C_i^{\text{bat}} + L_{i,t} \right), \quad \forall t \in \mathcal{T}, \tag{11e}$$

$$E^{\max} \geq E_t^{\text{agg}}, \quad \forall t \in \mathcal{T}, \tag{11f}$$

$$E^{\text{over}} \geq E^{\max} - E^{\text{thres}}, \tag{11g}$$

$$W_1 \geq E_1^{\text{agg}} - E^{\text{init\_agg}}, \tag{11h}$$

$$W_1 \geq E^{\text{init\_agg}} - E_1^{\text{agg}}, \tag{11i}$$

$$W_{t+1} \geq E_{t+1}^{\text{agg}} - E_t^{\text{agg}}, \quad \forall t \in \{1, \dots, 23\}, \tag{11j}$$

$$W_{t+1} \geq E_t^{\text{agg}} - E_{t+1}^{\text{agg}}, \quad \forall t \in \{1, \dots, 23\}, \tag{11k}$$

$$S_{i,t}^{\text{bat}} = S_{i,t}^{\text{bat, 1}} + S_{i,t}^{\text{bat, 2}}, \quad \forall i \in \mathcal{I}, t \in \mathcal{T}, \tag{11l}$$

$$0 \leq S_{i,t}^{\text{bat, 1}} \leq 0.8 \quad \forall i \in \mathcal{I}, t \in \mathcal{T}, \tag{11m}$$

$$0 \leq S_{i,t}^{\text{bat, 2}} \leq 0.2 \quad \forall i \in \mathcal{I}, t \in \mathcal{T}, \tag{11n}$$

$$0 \leq x_{i,t} \leq \frac{P_{\text{nom},i}^{\text{bat}} \cdot \left(1 - 4 \cdot S_{i,t}^{\text{bat, 2}}\right)}{C_i^{\text{bat}}}, \quad \forall i \in \mathcal{I}, t \in \mathcal{T}, \tag{11o}$$

$$-\frac{P_{\text{nom},i}^{\text{bat}} \cdot \left(1 - 4 \cdot S_{i,t}^{\text{bat, 2}}\right)}{C_i^{\text{bat}}} \leq y_{i,t} \leq 0, \quad \forall i \in \mathcal{I}, t \in \mathcal{T}, \tag{11p}$$

$$S_{i,1}^{\text{bat}} = S_{\text{init},i}^{\text{bat}} + x_{i,1} \cdot \eta^{\text{bat, ch}} + y_{i,1} \cdot \eta^{\text{bat, dis}}, \quad \forall i \in \mathcal{I}, \tag{11q}$$

$$S_{i,t+1}^{\text{bat}} = S_{i,t}^{\text{bat}} + x_{i,t} \cdot \eta^{\text{bat, ch}} + y_{i,t} \cdot \eta^{\text{bat, dis}}, \quad \forall i \in \mathcal{I}, t \in \{1, \dots, 23\}, \tag{11r}$$

$$y_{i,1} \geq -\left(S_{\text{init},i}^{\text{bat}} - S_i^{\text{dod}}\right) \cdot \eta^{\text{bat, rt}}, \quad \forall i \in \mathcal{I}, \tag{11s}$$

$$y_{i,t+1} \geq -\left(S_{i,t}^{\text{bat}} - S_i^{\text{dod}}\right) \cdot \eta^{\text{bat, rt}}, \quad \forall i \in \mathcal{I}, t \in \{1, \dots, 23\}. \tag{11t}$$

Each equality and inequality constraint is described as follows.

- Eq. (11c) defines the net load $L_{i,t}$ of the building excluding the battery energy. It aggregates various energy components: non-shiftable consumption ($E_{i,t}^{\text{ns}}$), PV generation ($E_{i,t}^{\text{pv}}$), cooling electricity consumption ($E_{i,t}^{\text{hp}}$), and domestic hot water consumption ($E_{i,t}^{\text{dhw}}$).
- Eq. (11d) ensures that the net grid power consumption $E_{i,t}^{\text{net}}$ covers both the load $L_{i,t}$ and battery power $\left(x_{i,t} + y_{i,t}\right) C_i^{\text{bat}}$.
- Eq. (11e) and (11f) aggregate individual loads into E_t^{agg} and constrain it by $E^{\max}$ to limit peak consumption.
- Eq. (11g) introduces E^{over}, representing the excess of $E^{\max}$ over a certain threshold E^{thres} to penalize large peaks during optimization.
- Eq. (11h)–(11k) define W_t as an upper bound on rapid changes in aggregated consumption between consecutive time steps, thereby smoothing the net load profile.
- Eq. (11l)–Eq. (11n) decompose the SOC $S_{i,t}^{\text{bat}}$ into two zones: a standard range $\left(S_{i,t}^{\text{bat,1}}\right)$ from 0 to 0.8, and a high SOC zone $\left(S_{i,t}^{\text{bat,2}}\right)$ from 0 to 0.2. This formulation regulates battery operation more conservatively in the upper SOC region.
- Eqs. (11o) and (11p) set dynamic bounds on the battery's charging and discharging action, respectively. The bounds scale the nominal power $P_{\text{nom},i}^{\text{bat}}$ by the coefficient $(1 - 4 \cdot S_{i,t}^{\text{bat,2}})$. When $S_{i,t}^{\text{bat,2}} = 0$, charging and discharging occur at nominal power. As $S_{i,t}^{\text{bat,2}}$ increases, the coefficient decreases, gradually restricting power to prevent unsafe operations. $S_{i,t}^{\text{bat,2}}$ becomes active only when $S_{i,t}^{\text{bat, 1}} = 0.8$, requiring the lower SOC zone to fill first. In the

Table 3
Forecasting models and input features for exogenous variables utilized in the proposed MPC framework for battery dispatch.

Target variables (24h prediction)	Model	Input features
E^{ns}	LightGBM	H, D, T, E^{pv} T^{out} (current, 6h, 12h, 24h leads) I^{dif} (current, 6h, 12h, 24h leads) I^{dir} (current, 6h, 12h, 24h leads) E^{ns}(current, 1–24h lags)
E^{pv}	LightGBM	H, D T^{out} (current, 6h, 12h, 24h leads) I^{dif} (current, 6h, 12h, 24h leads) I^{dir} (current, 6h, 12h, 24h leads) E^{pv} (current, 1–24h lags)
CI	KNN regression	CI (current, 1–24h lags)

higher SOC zone (0.8-1), the coefficient more aggressively limits power, reducing stress and extending battery life. See Appendix A for details.

- Eqs. (11q) and (11r) describe the SOC transitions for the initial and subsequent time steps, respectively.
- Eq. (11s) and Eq. (11t) enforce a lower SOC bound through the depth-of-discharge threshold S_i^{dod} and account for round-trip efficiency $\eta^{bat,rt}$.

The above battery dispatch formulation constitutes a linear programming problem and is solved using the CPLEX [33] solver to ensure computational efficiency and global optimality within the 24-hour horizon. By explicitly incorporating the heat pump consumption E_t^{hp} and DHW electricity consumption $E_{i,t}^{dhw}$ into the battery dispatch optimization (Eq. (11c)), the framework can effectively synchronize heat pump operations, domestic hot water usage, and battery charging/discharging. This coordination enables lower peak power requirements and smoother net load profiles across the multi-building system.

3.4.1. Prediction models

In the proposed MPC framework, accurate forecasting of the non-shiftable load (E^{ns}), PV generation (E^{pv}), and carbon intensity (CI) is critical. Table 3 summarizes the models and input features for predicting these variables over a 24-hour horizon. The non-shiftable load and PV generation are predicted using LightGBM models, leveraging features such as time indicators (H, D), weather variables (T^{out}, I^{dif}, I^{dir}), and historical data of the respective target variables. Carbon intensity is forecasted using a KNN regression model based solely on its historical lag values. These forecasts provide the exogenous inputs necessary for solving the optimization problem.

3.5. Power outage

To facilitate the energy distribution during periods of grid unavailability, we propose a dedicated control scheme that is immediately triggered once $\delta^{out} = 1$. During these events, the principal energy supply primarily stems from solar generation, DHW storage, and battery storage. In the following, we describe how to manage each crucial step to ensure an efficient operation under constrained conditions.

Ensuring higher SOC before outages . Before an outage occurs, we keep a higher minimum SOC threshold for the battery (e.g., 0.5). Maintaining this elevated lower bound, the battery retains more energy to power critical loads once the outage begins. This approach helps reduce the impact of sudden supply shortages and enhances the overall resilience of the energy system.

Temperature setpoint adjustment. Recognizing that maintaining occupant comfort must be balanced against energy constraints, we implement a moderate increase (e.g., 0.5 °C) in the indoor temperature setpoint. The adjustment keeps the indoor temperature within the comfort band while reducing energy consumption for cooling. Once the grid supply is restored, the system can revert to standard setpoints, ensuring the balance between comfort and energy savings remains optimal throughout normal and outage conditions.

Precise power allocation. During an outage, we allocate power resources based on the expected cooling demand, available solar generation, and the battery's maximum discharge rate. Specifically, if the forecasted cooling demand is entirely met by solar generation, the battery remains idle to conserve its reserves. Should the solar output alone be insufficient but still within the sum of solar and maximum discharge capacity, we draw the needed difference from the battery while preserving the original cooling demand. In scenarios where the projected demand surpasses solar generation and the battery's maximum discharge, the system simultaneously deploys its full discharge capability and reduces the cooling load to align with the limited power availability. By proportionally adjusting battery discharge and cooling load, we aim to minimize unmet cooling requirements while extending the battery's operational window under outage conditions.

DHW storage. In light of the limited energy resources available during an outage, the hot water storage system is not recharged. Consequently, the hot water supply is limited to the volume already stored before the outage. This measure ensures that electrical energy is conserved for higher-priority demands, thereby optimizing resource utilization under constrained conditions.

4. Results and discussion

4.1. Heat pump control

This section presents and analyzes experimental results on heat pump control for thermal comfort over a three-month testing period. The testing results for the six buildings in a representative week are shown in Fig. 7. Overall, the controlled indoor temperatures remain within acceptable comfort levels for most of the operational time, suggesting that the data-driven MPC controller generally succeeds in regulating the heat pump to align with occupant preferences. In public buildings, the temperature profiles closely follow occupant setpoints, staying well within the comfort bounds. However, the private buildings exhibit larger temperature deviations which is unsurprising given that their thermal dynamics are inferred through the fine-tuning approach rather than derived from detailed building models. Building 6, in particular, consistently maintains higher indoor temperatures than occupants desire, indicating that the current fine-tuning strategy may be insufficient in capturing the building's thermal characteristics. Such limitations underscore the necessity of exploring more robust transfer strategies to enhance control performance when thermal dynamics remain only partially known.

4.2. DHW storage control

Fig. 8 presents the evolution of ramping scores over a three-month testing period for the public and three private buildings, comparing scenarios with and without DHW storage and DR. Buildings equipped with DHW storage and managed under our RBC strategy consistently exhibit lower ramping values, indicating smoother load profiles and reduced stress on the grid. Fig. 9 further illustrates the effects of DHW control on energy usage by showing the distribution of DHW storage SOC and mean net electricity consumption over a 24-hour period. In public buildings, an elevated SOC is generally maintained during off-peak hours (1:00—5:00), allowing daytime hot water demand to be met from stored reserves rather than real-time electricity. Consequently, the net electricity consumption profile remains flatter with lower variance than scenarios lacking DHW storage and DR. Private buildings follow a similar pattern and demonstrate that a simple, time-based DHW storage control policy can substantially enhance grid efficiency.

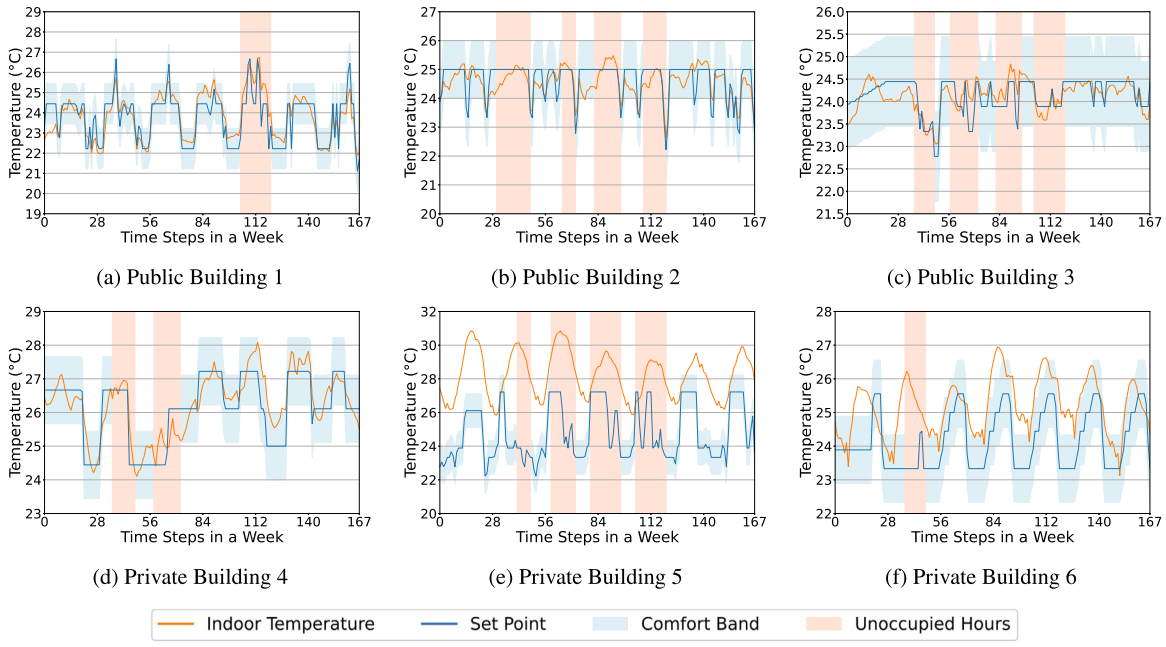

Fig. 7. The comparison of the indoor temperature with the set points of six buildings in a week.

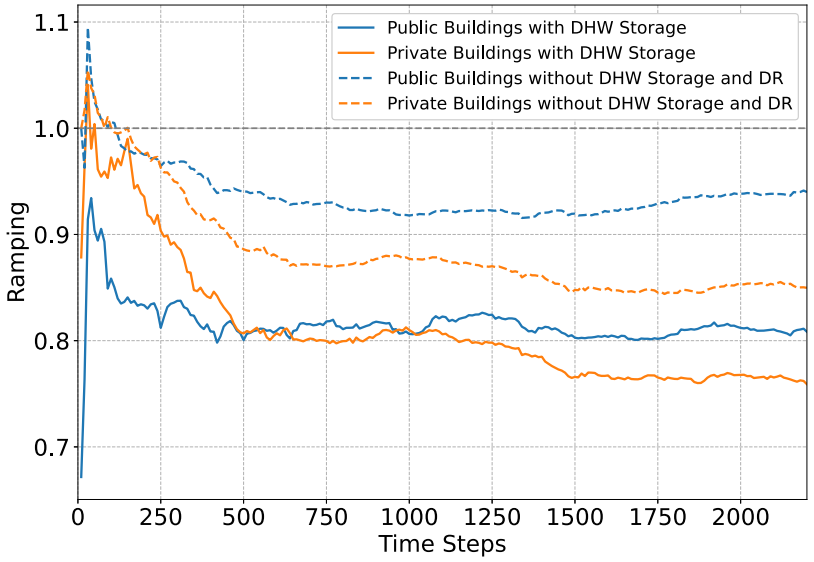

Fig. 8. Ramping score of three public and three private buildings over the three-month testing period, comparing cases with DHW storage and DR to those without them.

4.3. Battery control

Battery storage serves as an essential device for smoothing load profiles and mitigating peak demand, thereby enhancing overall grid efficiency. The MPC-based dispatch strategy coordinates charging and discharging, allowing the battery to store surplus energy when demand is low and supply it back to the grid during peak hours. Fig. 10 shows the mean net electricity consumption along with the corresponding variance ranges and blue bars represent battery SOC trends. The battery typically charges around midday when solar generation is at its peak—and discharges later in the evening to help meet rising demand. The MPC-based dispatch strategy displays a smoother profile with less variance during periods of high demand, which indicates that stored energy is deployed effectively to shift load and mitigate demand fluctuations, thereby enhancing grid stability. Fig. 11 presents a summary of four grid efficiency KPIs after the three-month testing period, where a lower score is more desirable for all metrics. In each case, the battery-enabled scenario shows clear improvements over the baseline, reflecting reduced load volatility and minimized peak demand. These results underscore the efficacy of the MPC dispatch strategy in delivering smoother load profiles and alleviating stress on the grid.

4.4. Challenge results and discussion

Table 4 presents a comprehensive comparison between the top three competing teams on the leaderboard of the CityLearn Challenge.[1] Our solution achieves the best performance in the public buildings category with a score of 0.464, which corresponds to a 57% improvement over

[1] Control track leaderboard of NeurIPS 2023 CityLearn Challenge: https://www.aicrowd.com/challenges/neurips-2023-citylearn-challenge/problems/control-track-citylearn-challenge/leaderboards?challenge_round_id=1303

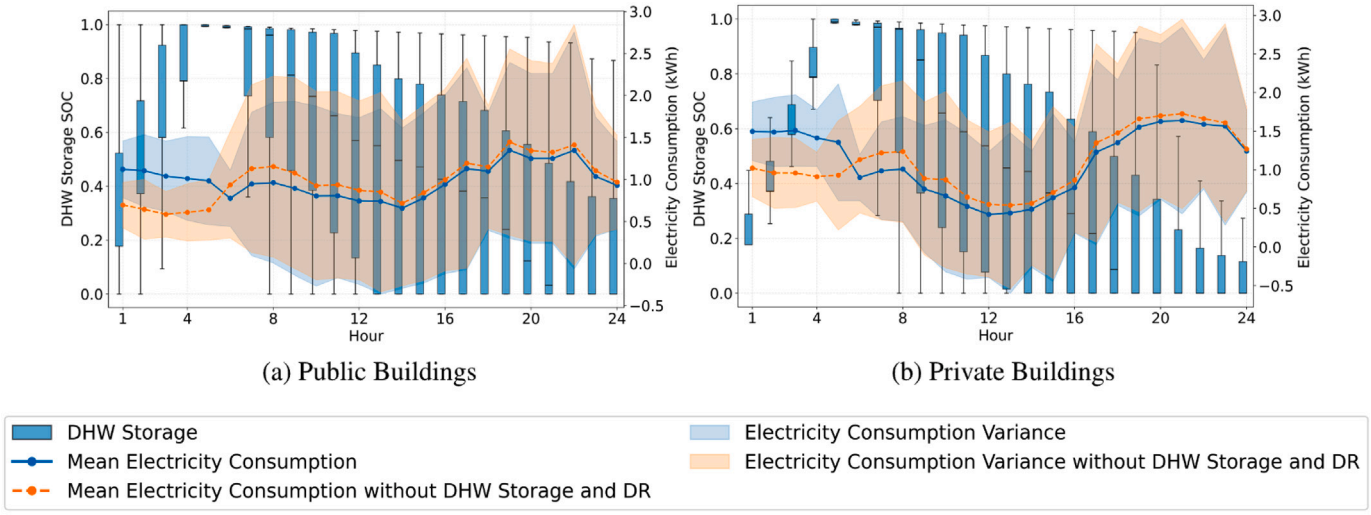

Fig. 9. Hourly distribution of DHW storage SOC and net electricity consumption over a 24-hour period. The data is aggregated from the results of the three-month testing period, comparing cases with DHW storage and DR to those without them.

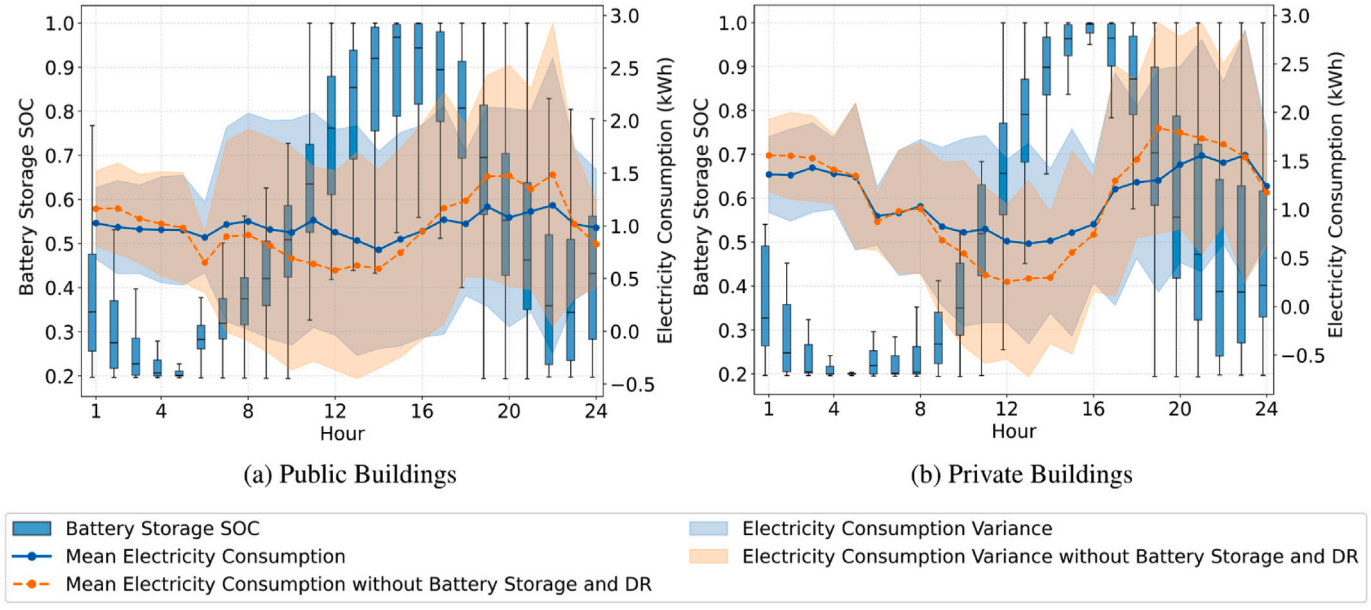

Fig. 10. Hourly distribution of net electricity consumption and battery SOC over a 24-hour period, comparing cases with battery storage and DR to those without them.

the competition baseline, demonstrating the strength of our multi-level MPC-based framework in environments with known dynamics and rich data. Compared with the first-place solution (Team 1), both frameworks adopted a hierarchical control structure but employed fundamentally different control strategies. Our framework combined a data-driven MPC for heat pump cooling, a rule-based controller for DHW scheduling, and a physics-based MPC for battery dispatch. In contrast, Team 1 relied on Proportional–integral–derivative (PID) and heuristic methods, focusing on robustness and real-time responsiveness [34]. In battery management, we utilized a physics-based MPC approach that optimizes across multiple objectives. Team 1, on the other hand, used a heuristic tree search strategy with decentralized charge/discharge suggestions, followed by centralized corrections. Our method resulted in higher improvements in grid cost (28%), resilience (33%), and carbon emissions (60%) compared to Team 1's respective improvements of 23%, 29%, and 57%. This advantage stems from the receding horizon nature of our MPC controller, which continuously re-optimizes battery scheduling based on updated forecasts and system states. For temperature regulation, we introduced a transfer learning mechanism to adapt the thermal model, enabling it to generalize from public to private buildings and addressing the challenge of unknown thermal dynamics in private buildings. Team 1, however, adopted a PID controller independent of thermal modeling and used Bayesian optimization for parameter tuning, contributing to higher robustness and generalizability. Their temperature regulation approach proved more effective in private buildings, achieving an 87% improvement in thermal comfort ($S_c = 0.13$) versus our 69% ($S_c = 0.30$). This advantage in thermal comfort, which holds a relatively high weight in the final score as defined in Eq. (2), ultimately offset our gains in other metrics and led to Team 1 achieving a slightly better overall score.

Given the resource constraints of embedded controllers in real-world deployments, we also evaluated the computational efficiency of the proposed MPC-based framework on a CentOS system equipped with an AMD EPYC 7K62 processor (2.6 GHz, 32 cores) and 62 GB of RAM. The results indicate that the framework achieved excellent computational performance, with an average computational time of 0.4 s per

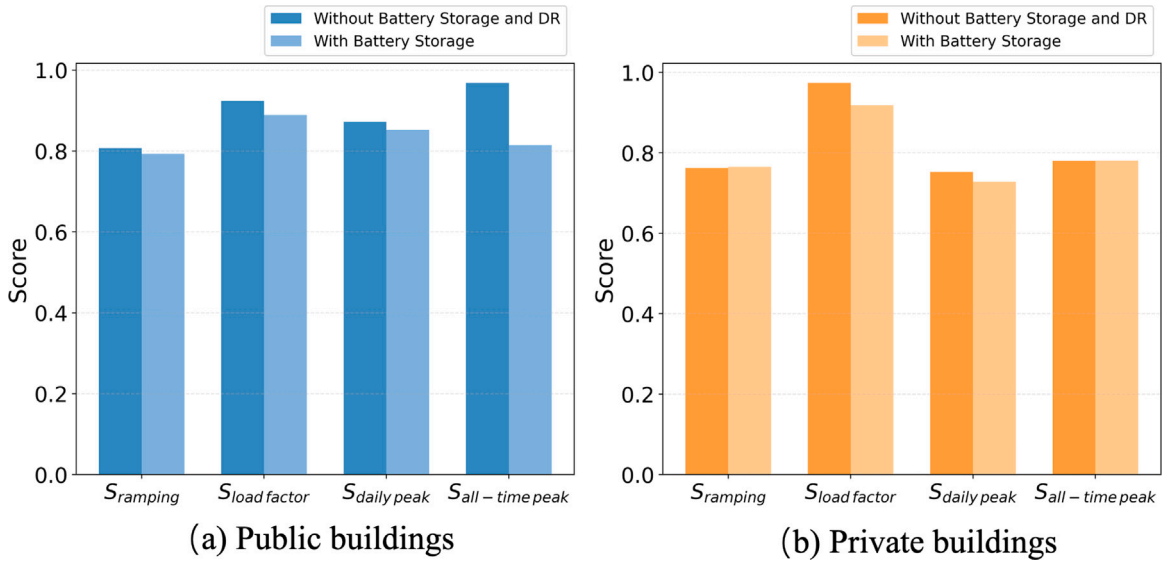

Fig. 11. Comparison of four grid efficiency KPIs ($S_{\text{ramping}}, S_{\text{load factor}}, S_{\text{daily peak}}, S_{\text{all-time peak}}$) for public and private buildings after a three-month testing period. Darker bars represent the baseline scenario without battery storage, while lighter bars indicate battery dispatch by the MPC controller.

Table 4
The performance of the top three teams on the CityLearn Challenge leaderboard. Values in parentheses indicate percentage improvements from the RBC baseline. The best-achieved performance for each metric is highlighted, with lower values indicating better performance.

Team	Overall score	Public score	S_c	S_e	S_g	S_r
Baseline	1.12	1.09	0.99	2.19	1.147	0.775
Team 1	**0.565** (50%)	0.562 (48%)	**0.13** (87%)	0.94 (57%)	0.882 (23%)	0.552 (29%)
Ours	0.575 (49%)	**0.464** (57%)	0.30 (69%)	**0.88** (60%)	**0.821** (28%)	**0.520** (33%)
Team 3	0.582 (48%)	0.508 (53%)	0.20 (80%)	0.93 (58%)	0.847 (26%)	0.590 (24%)

control interval (one hour). This fast computation time highlights the feasibility of deploying the framework in real-time applications with limited computational resources.

5. Conclusion and future work

This paper presents an integrated control framework for optimizing energy management in community buildings. By incorporating data-driven MPC for heat pump control, a rule-based strategy for DHW storage and a model-based MPC for battery dispatch, our approach achieves significant improvements in grid efficiency, peak shaving and resilience. Resilience measures were also incorporated to manage unexpected grid outages, ensuring continued functionality under constrained conditions. The proposed methodology effectively coordinates different energy resources while ensuring thermal comfort for occupants. The experimental results demonstrate that our controller successfully balances real-time energy demand and supply, mitigating peak electricity usage and improving overall grid stability. The findings further highlight the importance of hierarchical coordination between HVAC and microgrid levels, allowing the controllers to adapt dynamically to energy consumption and external constraints.

Moreover, the transfer learning-based approach for fine-tuning the heat pump control model shows promising results, enabling adaptation to buildings with unknown thermal dynamics. However, deviations in temperature regulation observed in some private buildings indicate that further refinement of the adaptation process is necessary. The larger temperature deviations observed in private buildings can be attributed to two primary factors. First, the private buildings rely on transfer learning from source domains instead of directly leveraging detailed thermal models, which inevitably introduces performance gaps, as the transferred models may not fully capture the unique thermal responses of the target buildings. Second, in the current implementation, only the final fully connected layer of the LSTM-based thermal model is fine-tuned. While this design reduces the computational cost and helps mitigate overfitting, it may be insufficient when the thermal characteristics of the target building diverge significantly from those of the source building. This limitation could explain the persistent temperature overshoot observed in Building 6. Despite the limitations in thermal comfort control for some private buildings, the proposed framework performed competitively overall. Our solution secured second place in the 2023 NeurIPS CityLearn Challenge, validating the effectiveness of our control strategy. The framework achieving the highest performance in public buildings and outperforming competitors in key categories such as carbon emissions reduction, grid stability, and resilience. While the proposed framework is developed in the context of community buildings, its core design principles, such as hierarchical coordination between HVAC and microgrid layers and coal control at the device level, can be effectively extended to commercial or mixed-use buildings. These buildings often feature more complex energy systems. However, the control philosophy remains applicable. Beyond the current scope of community buildings, the proposed framework also demonstrates strong potential for broader applicability. Although it was initially developed for community buildings, core design principles such as hierarchical coordination between the HVAC and microgrid layers and local control at the device level can be extended to commercial or mixed-use buildings. These buildings typically involve more complex energy systems, such as centralized chillers, ventilation units, and heterogeneous load profiles. Nevertheless, the control philosophy remains relevant because the primary objectives of ensuring occupant comfort, minimizing energy costs, enhancing load balancing, and supporting grid stability are generally consistent across building types. Due to its modular structure, the framework can be flexibly adapted to accommodate the specific equipment configurations and operational schedules of commercial settings.

For future work, several key directions can be explored to improve the adaptability and performance of the proposed framework. First, refining the fine-tuning strategy of data-driven MPC through advanced transfer learning techniques could improve adaptability to buildings with unknown thermal dynamics. One potential mitigation is to extend the fine-tuning beyond the final output layer and adapt intermediate layers of the LSTM thermal model. This approach may allow the model to retain generalized knowledge from source domains while better capturing the specific thermal characteristics of target buildings, thereby reducing performance gaps introduced by domain shifts. Second, as real-time pricing data is not provided in the current simulation environment, our framework does not yet incorporate price-based demand response. Future work could explore the integration of demand-side management strategies considering real-time pricing. To further enhance the framework's value, synthetic price signals could be incorporated into the control optimization. This would enable flexible loads to respond to cost variations and enhance peak shaving and overall system efficiency. The current approach predicts temperature setpoints based on historical patterns but does not explicitly account for occupant preferences. RL presents a promising approach for developing occupant behavior models [35], allowing agents to learn from interaction and adapt to changing comfort needs. However, integrating RL into the MPC framework poses challenges due to its requirement for multi-step setpoint forecasts. Future work could investigate combining RL with sequence modeling or trajectory planning techniques to generate long-horizon behavioral predictions compatible with MPC. Finally, potential cybersecurity threats and mitigation strategies for smart energy management systems warrant further investigation. Given the framework's dependence on real-time data exchange and control actions, it remains vulnerable to data breaches and system tampering risks. Implementing security measures such as encrypted communication, user authentication, and anomaly detection would strengthen system resilience and protect against cyberattacks.

CRediT authorship contribution statement

Wanfu Zheng: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Ziqi Hu:** Writing – review & editing, Writing – original draft, Visualization, Validation, Formal analysis. **Dan Wang:** Visualization, Validation. **Zhe Wang:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

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.

Acknowledgments

This study is supported by the National Science Fund for Excellent Young Scholars, China (Grant number 52322813); by 2023 HKUST Shenzhen-Hong Kong Collaborative Innovation Institute Green Sustainability Special Fund, China, from Shui On Xintiandi and the InnoSpace GBA, China; and by the Hong Kong University Grants Committee (UGC) Grant 26209323.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.enconman.2025.119927.

Data availability

I am providing the URL link where the research data and code can be accessed: https://github.com/wfzheng/2nd-place-solution-neurips-citylearn2023-control.

References

[1] United Nations Environment Programme. Global status report for buildings and construction: Towards a zero-emission, efficient and resilient buildings and construction sector. 2022, URL https://www.unep.org/resources/publication/2022-global-status-report-buildings-and-construction. [Accessed 12 August 2023].
[2] Shin Minjae, Kim Sungsoo, Kim Youngjin, Song Ahhyun, Kim Yeeun, Kim Ha Young. Development of an HVAC system control method using weather forecasting data with deep reinforcement learning algorithms. Build Environ 2024;248:111069.
[3] Bayasgalan Anujin, Park Yoo Shin, Koh Seak Bai, Son Sung-Yong. Comprehensive review of building energy management models: Grid-interactive efficient building perspective. Energies 2024;17(19):4794.
[4] Zhong Weifeng, Yu Rong, Xie Shengli, Zhang Yan, Yau David KY. On stability and robustness of demand response in V2G mobile energy networks. IEEE Trans Smart Grid 2016;9(4):3203–12.
[5] Nawaz Arshad, Zhou Min, Wu Jing, Long Chengnian. A comprehensive review on energy management, demand response, and coordination schemes utilization in multi-microgrids network. Appl Energy 2022;323:119596.
[6] Wamalwa Fhazhil, Ishimwe Ariane. Optimal energy management in a grid-tied solar PV-battery microgrid for a public building under demand response. Energy Rep 2024;12:3718–31.
[7] Inal Omer Berkehan, Charpentier Jean-Frédéric, Deniz Cengiz. Hybrid power and propulsion systems for ships: Current status and future challenges. Renew Sustain Energy Rev 2022;156:111965.
[8] Yang Shiyu, Wan Man Pun, Chen Wanyu, Ng Bing Feng, Dubey Swapnil. Model predictive control with adaptive machine-learning-based model for building energy efficiency and comfort optimization. Appl Energy 2020;271:115147.
[9] Wang Shufeng, Zhang Baokang, Yuan Yadong, Liu Zhe. A dynamic weight multi-objective model predictive controller for adaptive cruise control system. Automatika 2023;64(4):919–32.
[10] Aldaouab Ibrahim, Daniels Malcolm, Ordóñez Raúl. Model predictive control energy dispatch to optimize renewable penetration for a microgrid with battery and thermal storage. In: 2018 IEEE texas power and energy conference. IEEE; 2018, p. 1–6.
[11] Ginzburg-Ganz Elinor, Segev Itay, Balabanov Alexander, Segev Elior, Kaully Naveh Sivan, Machlev Ram, et al. Reinforcement learning model-based and model-free paradigms for optimal control problems in power systems: Comprehensive review and future directions. Energies 2024;17(21):5307.
[12] Zhang Kun, Prakash Anand, Paul Lazlo, Blum David, Alstone Peter, Zoellick James, et al. Model predictive control for demand flexibility: Real-world operation of a commercial building with photovoltaic and battery systems. Adv Appl Energy 2022;7:100099.
[13] Wamalwa Fhazhil, Williams Nathaniel J. Integration of an off-grid solar-PV-battery system after grid connection using model predictive control: A case study in Kenya. In: 2021 IEEE PES/iAS powerAfrica. IEEE; 2021, p. 1–5.
[14] Zhu Bing, Tazvinga Henerica, Xia Xiaohua. Switched model predictive control for energy dispatching of a photovoltaic-diesel-battery hybrid power system. IEEE Trans Control Syst Technol 2014;23(3):1229–36.
[15] Aruta Giuseppe, Ascione Fabrizio, Bianco Nicola, Mauro Gerardo Maria, Vanoli Giuseppe Peter. Optimizing heating operation via GA-and ANN-based model predictive control: Concept for a real nearly-zero energy building. Energy Build 2023;292:113139.
[16] Stoffel Phillip, Berktold Max, Müller Dirk. Real-life data-driven model predictive control for building energy systems comparing different machine learning models. Energy Build 2024;305:113895.
[17] Zheng Wanfu, Wang Dan, Wang Zhe. Economic model predictive control for building HVAC system: A comparative analysis of model-based and data-driven approaches using the BOPTEST framework. Appl Energy 2024;374:123969.
[18] Jiang Wei, Yi Zhongkai, Wang Li, Zhang Hanwei, Zhang Jihai, Lin Fangquan, et al. A stochastic online forecast-and-optimize framework for real-time energy dispatch in virtual power plants under uncertainty. In: Proceedings of the 32nd ACM international conference on information and knowledge management. 2023, p. 4646–52.
[19] Vanthournout Koen, D'hulst Reinhilde, Geysen Davy, Jacobs Geert. A smart domestic hot water buffer. IEEE Trans Smart Grid 2012;3(4):2121–7.
[20] Pepe Crescenzo, Farooq Ali Mir, Zanoli Silvia Maria. Heat pump management for building's indoor temperature control. In: 2024 25th international carpathian control conference. IEEE; 2024, p. 1–6.
[21] Grant Peter, Gehbauer Christoph. Evaluating the impact of weather forecast inaccuracy on performance of model predictive control for dynamic facades. In: Proceedings of the 2022 building performance modeling conference and simBuild. 2022.

[22] Li Lunlong, Ju Yi, Wang Zhe. Quantifying the impact of building load forecasts on optimizing energy storage systems. Energy Build 2024;307:113913.
[23] Ma Zhenjun, Awan Muhammad Bilal, Lu Menglong, Li Shengteng, Aziz Muhammad Shahbaz, Zhou Xinlei, et al. An overview of emerging and sustainable technologies for increased energy efficiency and carbon emission mitigation in buildings. Buildings 2023;13(10):2658.
[24] Hu Kai, Yan Chengchu, Xu Chao, Li Wenjun, Ye Jing, Gong Yanfeng, et al. Strategies for grid-friendly and uncertainty-adaptive design in zero energy buildings. Energy Build 2024;307:113967.
[25] Joe Jaewan, Karava Panagiota. A model predictive control strategy to optimize the performance of radiant floor heating and cooling systems in office buildings. Appl Energy 2019;245:65–77.
[26] Deng Kun, Sun Yu, Li Sisi, Lu Yan, Brouwer Jack, Mehta Prashant G, et al. Model predictive control of central chiller plant with thermal energy storage via dynamic programming and mixed-integer linear programming. IEEE Trans Autom Sci Eng 2014;12(2):565–79.
[27] Ra Seon Jung, Kim Jin-Hong, Park Cheol Soo. Real-time model predictive cooling control for an HVAC system in a factory building. Energy Build 2023;285:112860.
[28] Dongol Deepranjan, Feldmann Thomas, Bollin Elmar. A model predictive control based peak shaving application for a grid connected household with photovoltaic and battery storage. In: SMARTGREENS. 2018, p. 54–63.
[29] Abraham Sunil, Mishra Yateendra, Cholette Michael E. Mpc based community battery system to minimize the energy cost of a residential community. In: 2023 IEEE power & energy society general meeting. IEEE; 2023, p. 1–5.
[30] Joe Jaewan, Dong Jin, Munk Jeffrey, Kuruganti Teja, Cui Borui. Virtual storage capability of residential buildings for sustainable smart city via model-based predictive control. Sustain Cities Soc 2021;64:102491.
[31] Vazquez-Canteli Jose R, Dey Sourav, Henze Gregor, Nagy Zoltan. Citylearn: Standardizing research in multi-agent reinforcement learning for demand response and urban energy management. 2020, arXiv preprint arXiv:2012.10504.
[32] Ke Guolin, Meng Qi, Finley Thomas, Wang Taifeng, Chen Wei, Ma Weidong, et al. Lightgbm: A highly efficient gradient boosting decision tree. Adv Neural Inf Process Syst 2017;30.
[33] Cplex IIBM ILOG, et al. V12. 1: User's manual for CPLEX. Int Bus Mach Corp 2009;46(53):157.
[34] Garmendia Andoni, Morri Francesco, Cappart Quentin, Cadre Hélène Le. Winning the 2023 CityLearn challenge: a community-based hierarchical energy systems coordination algorithm. In: Proceedings of the 27th European conference on artificial intelligence. IOS Press; 2024, p. 2330–7.
[35] Deng Zhipeng, Chen Qingyan. Reinforcement learning of occupant behavior model for cross-building transfer learning to various HVAC control systems. Energy Build 2021;238:110860.