Electrical Power and Energy Systems 178 (2026) 111918

Contents lists available at ScienceDirect

International Journal of Electrical Power and Energy Systems

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

Dynamic characterization of power user profiles for low-carbon regulation potential assessment

Jinfeng Li [a], Run Tang [a], Siyu Jiang [b,c,*], Yu Yao [a], Jiahe Li [b,c], Yishun Zhu [a], Ziyan Pan [a], Guanbin Huang [a], Hongxun Hui [b,c,*]

[a] Guangzhou Power Supply Bureau of Guangdong Power Grid Co., Ltd., Guangzhou, 510620, China
[b] State Key Laboratory of Internet of Things for Smart City, University of Macau, 999078, Macao Special Administrative Region of China
[c] Department of Electrical and Computer Engineering, University of Macau, 999078, Macao Special Administrative Region of China

ARTICLE INFO

Keywords:
Power user profiling
Deep learning
Low-carbon regulation
Potential assessment

ABSTRACT

Effective demand-side management is pivotal for maintaining the supply–demand balance in renewable-dominated power systems. Low-carbon regulation potential reflects a user's ability to adjust consumption in response to renewable energy availability. Accurately assessing this potential is a prerequisite for efficient regulation strategies. Consequently, precise user profiling serves as the critical foundation for identifying and quantifying this potential. However, existing profiling approaches are predominantly static, failing to capture the temporal volatility and diverse patterns of user flexibility. To bridge this gap, this paper proposes a novel dynamic profiling framework dedicated to assessing low-carbon regulation potential. We construct a multi-dimensional characteristic tag system that is continuously updated via a deep learning model. In this architecture, the Convolutional Neural Network extracts intricate local features from load patterns, while the Long Short-Term Memory network captures the temporal evolution of user regulation capabilities. A decision tree-based aggregation mechanism then synthesizes historical behavior and real-time potential scores to form a comprehensive evaluation. Experiments on the Building Data Genome Project-2 dataset demonstrate that the proposed method effectively captures dynamic variations in user behavior. Ultimately, our framework provides a quantifiable and iteratively updatable user profiling method, serving as a dynamic foundation for carbon-oriented energy management.

1. Introduction

In response to global climate change, carbon peaking and carbon neutrality targets have been widely adopted, accelerating the transition toward low-carbon power system operation [1,2]. As a major contributor to global energy consumption and carbon emissions, the building sector is increasingly recognized as a critical lever for sustainable development through low-carbon operation. Meanwhile, renewable energy sources (RES), such as photovoltaics and wind, are being deployed at scale and progressively displacing conventional fossil-fuel generation, thereby reshaping system operation. However, the inherent volatility and intermittency of RES intensify net-load variability and impose stringent requirements on real-time balancing [3]. Different from dispatchable conventional generation, RES outputs are strongly conditioned on meteorological factors, which can readily induce supply–demand mismatches and compromise grid security and stability. To preserve secure operation, power systems must maintain adequate regulation resources (i.e., balancing/ancillary services) that can respond rapidly to net-load deviations. Historically, such services have been largely provided by costly and carbon-intensive fossil-fuel generators (e.g., gas turbines). Consequently, with increasing RES penetration, it becomes imperative to identify cleaner and more economical regulation mechanisms. In this context, demand-side flexibility has emerged as a promising pathway, and building loads are being increasingly exploited for low-carbon regulation and the power grid balancing [4,5].

Public buildings equipped with heating, ventilation, and air conditioning (HVAC) systems exhibit considerable thermal inertia and controllable loads, making them promising resources for providing low-carbon regulation potential [6]. Nevertheless, before such resources can be dispatched safely and efficiently, the power grid must be able to quantify, in real time, how much regulation potential a building can reliably deliver at a given moment. Meanwhile, the deployment of smart-metering-driven sensing and management frameworks enables the acquisition of fine-grained electricity data and supports real-time decision making [7]. This data foundation facilitates the analysis of user

* Corresponding authors.
E-mail addresses: sy.jiang@connect.um.edu.mo (S. Jiang), hongxunhui@um.edu.mo (H. Hui).

https://doi.org/10.1016/j.ijepes.2026.111918
Received 10 March 2026; Received in revised form 27 April 2026; Accepted 5 May 2026
Available online 15 May 2026

behavior and the quantification of its characteristics, thereby making power user profiling a promising way for assessing building low-carbon regulation potential [8,9].

Existing studies on regulation-potential assessment can be broadly categorized into three technical streams: load forecasting, flexibility estimation, and profiling-based approaches. Load forecasting methods mainly aim to predict future electricity demand trajectories based on historical load, weather, and calendar information. These methods are valuable for anticipating short-term operating conditions, but their outputs are typically limited to future consumption levels and do not directly describe the user's adjustable capability or historical response characteristics [10,11]. Flexibility estimation approaches focus more explicitly on quantifying the amount of load that can be shifted, curtailed, or sustained under specific operational constraints [12,13]. Such methods are closer to the objective of regulation-potential assessment, yet they are often task-specific and tend to emphasize instantaneous or scenario-dependent capacity quantification rather than a persistent and updateable characterization of the user. Profiling-based approaches provide a third perspective by organizing user characteristics into a structured representation, such as behavior patterns, load features, or response labels, thereby offering a potentially more interpretable basis for differentiated regulation strategies [14,15].

Among these three streams, profiling-based approaches are particularly relevant to this study, as they provide a structured way to characterize user-level regulation features. A common line of research applies unsupervised learning methods to represent power-consumption behavior. For example, Sandels et al. [16] and Yang et al. [17] employ clustering on smart-meter time series to characterize electricity-consumption patterns. While Ayon et al. [18] use feature extraction and selection to improve the interpretability of clustering results. In addition, some studies expand the feature space by incorporating user information obtained from surveys and questionnaires. For instance, Bampoulas et al. [19] and Chen et al. [20] combine load data with survey-based user information. Ramírez-Mendiola et al. [21] investigate the associations between electricity usage and demographic or activity-related factors. Markovic et al. [22] further leverage machine learning to identify household user profiles.

However, existing profiling studies still face several limitations when applied to regulation-potential assessment in public buildings. Trotta et al. [23] summarize load curves over extended periods, while Flygare et al. [24] employ generic features and cluster-centroid representations to characterize user patterns. Although these approaches are effective for capturing common behavioral characteristics, they mainly emphasize representative or average usage patterns rather than the time-varying operational states of an individual building. As a result, they may be less suitable for public buildings, whose operations are inherently diverse and complex. In particular, generic or centroid-based representations may fail to faithfully describe the temporal variability of a specific building, thereby limiting the precision of individualized profiles. In addition, survey-based data collection is often difficult to implement in multi-tenant public buildings, which restricts the availability of detailed user-level information. More importantly, many existing power-user profiling methods lack dynamic updating capability and are therefore less suited to the strong temporal variability of public-building operations.

Therefore, although previous studies have provided useful foundations from the perspectives of forecasting, flexibility quantification, and consumption-pattern profiling, an important gap remains. Existing forecasting methods can describe what the load is likely to be, and flexibility estimation methods can approximate how much adjustable capacity may exist under certain conditions, but they do not by themselves provide a dynamic, interpretable, and continuously updateable representation of the regulation potential of an individual public building. Meanwhile, existing profiling studies often emphasize static segmentation or long-term consumption patterns, rather than real-time or near-real-time characterization of regulation capability. This limitation constrains the practical deployment of low-carbon demand-side regulation in public buildings, where the usable regulation potential is strongly time-varying and closely related to both current operating conditions and historical response behavior.

Based on the aforementioned analysis, the main contributions of this paper are as follows:

- A dynamic power user profiling framework is proposed to continuously update user characteristics over time, overcoming the limitations of static profiling and enabling timely and adaptive profiling-based assessment of relative regulation potential.
- Key user tags are developed to support the interpretable assessment of low-carbon regulation potential. The proposed indicators include historical peak-load reduction behavior (HPLRB) and current peak-load reduction potential (CPLRP), which are obtained by calculating key scores from multi-dimensional building data and converting them into a standardized profiling format.
- A high-accuracy load forecasting model is introduced as the foundation for the profile. The paper utilizes a Convolutional Neural Network-Long Short-Term Memory (CNN–LSTM) time series forecasting model to evaluate future load changes for each public building, thereby supporting dynamic profile generation and time-varying regulation-potential assessment.

The remainder of this paper is organized as follows: Section 2 describes the power user characteristic tag framework. Section 3 proposes the core methodology, detailing the process for constructing the dynamic user profile. Section 4 introduces the CNN–LSTM time series forecasting algorithm used for modeling consumption patterns. Section 5 utilizes a dataset to validate the proposed method's effectiveness. Finally, Section 6 presents the conclusions and discusses the limitations and directions for future work.

2. Power user characteristic tag framework

Power user profiling functions fundamentally as a data modeling technique grounded in multi-dimensional attribute tagging. Beyond the mere identification of users, power user profiling serves as a critical instrument for quantifying the low-carbon regulation potential of building loads. By systematically aggregating the physical attributes and behavioral characteristics of buildings, this technique identifies the heterogeneity in low-carbon regulation capabilities across different power users, thereby providing a quantitative basis for the formulation of precise decarbonization strategies.

Existing literatures construct power user characteristics by static foundational attributes. Static user profile primarily encompass long-term stable parameters such as building morphology, envelope thermal performance, HVAC system type, and rated power consumption. These parameters determine the energy consumption baseline and the physical boundaries for low-carbon regulation. Dynamic user profile needs to take load fluctuation patterns and response behaviors extracted through time-series analysis into consideration. To address the lack of comprehensive feature dimension construction in assessing the low-carbon regulation potential of public buildings, this paper proposes a multi-dimensional power user characteristic tag framework.

The proposed framework comprises three core categories of tags: basic building information tags, key power characteristic tags, and power user sensitivity tags. Basic building information tags are derived from building archives and ledgers, whereas the latter two categories consist of quantitative features extracted through time-series mining of operational data and coupling analysis with foundational attributes.

The basic building information tags encompass four sub-dimensions designed to establish the physical and geographical boundaries of the assessment targets:

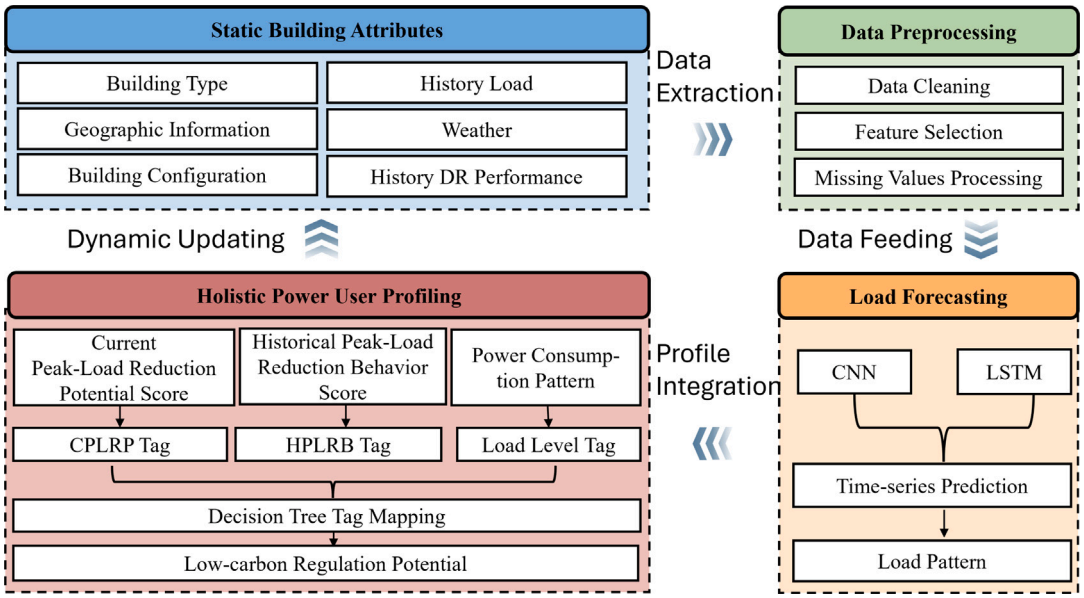


Fig. 1. Power user profiling process for public buildings.

- Customer Profile: Includes the unique building identifier, grid connection date, and historical preference for participating in low-carbon regulation programs, with demand response (DR) as a representative mechanism.
- Building Type: Classifies buildings as commercial/entertainment, office, residential, healthcare, cultural/educational, or government to distinguish baseline load variations driven by functional use.
- Geographic Information: Integrates longitude/latitude, elevation, and climate zone data to correct for the impact of meteorological factors on low-carbon regulation potential.
- Building Configuration: Captures the technical parameters determining the hard capability for low-carbon regulation, including heating system type, photovoltaic installed capacity, and energy storage specifications.

Key power characteristic tags focus on quantifying low-carbon regulation potential, incorporating power consumption pattern clustering models, historical peak-load reduction behavior (HPLRB) indices, and current peak-load reduction potential (CPLRP) predictive values.

User sensitivity tags include electricity price sensitivity and temperature sensitivity. These tags serve to reflect the probability of response willingness of public buildings under specific incentives or environmental changes.

The practical significance of the selected tags lies in constructing a physical capability–behavioral willingness dual-assessment dimension. The building configuration and power characteristic tags define the physical upper limit of low-carbon regulation, while user sensitivity tags calibrate the practically executable regulation potential. Furthermore, the framework adopts a modular design principle, supporting the continuous fusion of multi-source heterogeneous data. With the development of advanced metering infrastructure and Internet of Things technologies, this framework can directly integrate sub-metering, high-frequency data, thereby continuously improving the interpretability and predictive accuracy of low-carbon regulation potential assessment models through feature engineering.

2.1. Power user profiling construction process

The process for constructing power user profiling for public buildings is summarized in Fig. 1.

Initially, the independent system operator (ISO) establishes a basic user profile when a public building is connected to the grid. This profile provides basic tags and static building attributes (e.g., building type, location, and configuration), which are used to construct features and support subsequent characterization

Subsequently, the ISO aggregates time-varying data streams that are directly related to regulation capability, including historical load, weather conditions (e.g., temperature and dew-point temperature), and historical demand response (DR) performance (e.g., participation timestamps, response duration and frequency, adjusted-load ratio, and response evaluation scores). These multi-source data are then processed through a data preprocessing pipeline, including data cleaning, missing-value processing, and feature selection, to ensure reliability for downstream modeling.

Next, a power consumption pattern (load forecasting) module is employed to capture the temporal characteristics of building electricity usage. Specifically, CNN/LSTM-based time-series prediction is used to derive the load pattern representation, which further enables the extraction of load-level tags. This step provides a dynamic description of when and how the building demand aligns with grid regulation windows.

Based on the preprocessed data and the identified load pattern, the ISO evaluates user-side regulation capability from two complementary perspectives: (i) the historical peak-load reduction behavior (HPLRB) score, which quantifies the building's historical regulation contribution in terms of response duration, load-adjustment intensity (regulated-load participation) and (ii) the current peak-load reduction potential (CPLRP) score, which reflects the building's momentary regulation capability under prevailing operating conditions. The specific methodological approaches for these evaluations will be detailed in the subsequent section.

The calculated scores are then converted into the corresponding HPLRB tag and CPLRP tag via a standardized score-to-tag mapping procedure, and integrated with the load-level tag through tag mapping. Subsequently, the ISO performs a comprehensive evaluation by aggregating all derived tags to assess the building's low-carbon regulation potential and to stratify users for dispatch decisions. And a decision-tree-based rule can be constructed to identify high-value users.

Finally, the ISO checks whether updated data are available and, if so, triggers the framework to re-aggregate the data and update the building power user profile; otherwise, the latest profile is output without modification. In this way, the proposed framework enables dynamic

Table 1
Typical update schedule of major tags in the proposed framework.

Tag category	Example tags	Update rule	Rationale
Basic building tags	Building type, configuration, geographic information	Event-driven/infrequent	These attributes are relatively stable and only change when building records are updated.
HPLRB tags	Historical DR behavior	Weekly or after new DR events	Updated based on newly accumulated DR participation records and response evaluation cycles.
Load-level tags	High/middle/low load level	Daily forecasting cycle	Derived from the predicted 24-hour load profile.
CPLRP tags	Current regulation potential	Daily/data refresh	Updated according to current end-use load conditions and operating states.

updates of building power user profiles and supports time-varying, data-driven assessment of low-carbon regulation potential.

To make the dynamic updating mechanism more explicit, Table 1 summarizes the typical update frequency or trigger for each major tag category in the proposed framework.

It should be noted that dynamic profile updating does not imply that the forecasting model must be retrained whenever new data arrive. In the proposed framework, profile tags are refreshed at the corresponding data-update frequency, while the forecasting model may follow a slower life-cycle and be retrained only under periodic maintenance schedules or when sustained distributional changes are detected, such as seasonal shifts, occupancy-pattern changes, or persistent prediction-performance degradation.

2.2. Historical peak-load reduction behavior score

Historical peak-load reduction behavior (HPLRB) scores are derived from each power user's historical DR performance records. As current DR practice is predominantly incentive-based and price-based DR remains limited, this work focuses on incentive-based HPLRB scoring. The proposed HPLRB score provides a comprehensive assessment of past DR performance by jointly accounting for participation frequency, response quality, and behavioral consistency. Specifically, it integrates response type diversity, response duration, load-adjustment contribution, response effectiveness evaluation, recency weighting (temporal decay), and response frequency. In this paper, low-carbon regulation potential refers to the broader ability of public buildings to provide flexible demand-side support under renewable-dominated system operation, whereas peak-load reduction is used as a specific operational manifestation of this potential in the current framework.

In incentive-based DR, peak-load reduction events can be categorized by the advance notification horizon, including day-ahead response (approximately 24 h), short-notice response (approximately 4 h), and real-time response. Shorter notification horizons typically reflect higher system urgency and impose more stringent requirements on users; therefore, the HPLRB scoring assigns differentiated weights/adjustments across response types. The resulting incentive-based HPLRB formulation is given as follows.

$$S_{\mathrm{HPLRB}} = \sum_{i=1}^{n} w_i \cdot d_i \cdot l_i \cdot f_i \cdot \delta_h(t_i) \tag{1}$$

where N denotes the number of HPLRB events. For the ith event, w_i is the weight associated with its response type; d_i is the normalized duration score obtained by mapping the actual duration τ_i to $[0,1]$ (e.g., $d_i = \min(\tau_i/\tau_{\mathrm{ref}}, 1)$ with $\tau_{\mathrm{ref}} = 15$ min); l_i is the load-adjustment score derived from the regulated-load ratio (i.e., the ratio of the load involved in regulation to the total load); and f_i is the response feedback score assigned by the grid operator or aggregator. The term $\delta_h(t_i)$ is a temporal decay function that discounts the contribution of event i according to its occurrence time t_i; its explicit form is given below.

$$\delta_h(t_i) = \exp(-\lambda t_i), \qquad \lambda = 0.3, \tag{2}$$

where t_i denotes the elapsed time between the ith historical DR response event and the current evaluation time, measured in weeks. In particular, $t_i = 0$ if the ith event occurs within the same calendar week as the current evaluation. The exponential form in (2) implements a recency-weighting mechanism, so that recent responses contribute more to the HPLRB score while outdated events are progressively discounted.

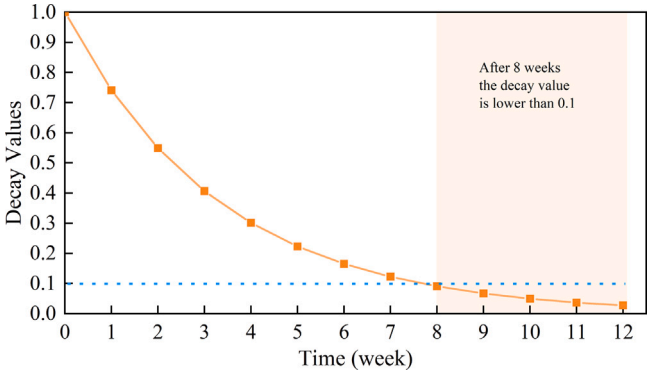


Fig. 2. Time decay function in 12 weeks.

The decay parameter λ is specified based on the operational characteristics of incentive-based DR programs, in which performance is typically evaluated on a monthly rolling basis and settled on a quarterly basis. To align the scoring mechanism with this practice, we set $\lambda = 0.3$, resulting in an effective memory horizon of approximately one quarter. As illustrated in Fig. 2, the decay factor drops below 0.1 after about 8 weeks, indicating that responses older than two months already have a substantially reduced influence. When $t_i = 12$ weeks (i.e., approximately three months), $\delta_h(t_i) = \exp(-0.3\times12) \approx 0.027$, implying that such events contribute less than 3% of their original weight and therefore have a negligible impact on the aggregated HPLRB score.

It should be noted that $\lambda = 0.3$ is not intended to be a universally fixed constant. In real-world deployment, the decay parameter may be adjusted according to the response evaluation cycle, settlement horizon, and operational preference of the grid operator or aggregator.

2.3. Current peak-load reduction potential score

The electrical equipment that can participate in DR usually includes HVAC and electric vehicle (EV) charging stations. EV charging stations are important movable loads. For a building with a certain number of charging piles, the load of the charging piles occupies a considerable proportion of its total load and instantaneous load. Therefore, the calculation formula for the CPLRP score is shown below.

$$S_{\mathrm{CPLRP}} = k_{\mathrm{EV}}\, r_{\mathrm{EV}} + k_{\mathrm{HVAC}}\, r_{\mathrm{HVAC}} = k_{\mathrm{EV}} \frac{L_{\mathrm{EV}}}{L_{\mathrm{Total}}} + k_{\mathrm{HVAC}} \frac{L_{\mathrm{HVAC}}}{L_{\mathrm{Total}}}, \tag{3}$$

where k_{EV} and k_{HVAC} denote the weighting coefficients associated with EV charging stations and HVAC, respectively. The terms r_{EV} and r_{HVAC} represent the corresponding load ratios, defined as the proportion of EV and HVAC loads relative to the total building load, respectively. Here,

Table 2
Power consumption pattern conversion to load level tag.

Normalized consumption range	Load-level tag
$\tilde{p} \geq 0.7$	High
$0.3 < \tilde{p} < 0.7$	Middle
$0 \leq \tilde{p} \leq 0.3$	Low

L_{EV}, L_{HVAC}, and L_{Total} denote the EV charging load, the HVAC load, and the total building load, respectively.

The calculation of L_{EV} and L_{HVAC} assumes that end-use load information can be obtained either from direct sub-metering/building energy management systems or from reliable load disaggregation methods. In practice, this assumption is increasingly reasonable in public buildings with categorized electricity metering for major systems such as HVAC equipment and EV charging facilities. When direct sub-metering is unavailable, Non-Intrusive Load Monitoring or similar disaggregation techniques may be used to estimate these load components. Accordingly, the proposed CPLRP formulation is most readily applicable in data-rich public-building scenarios.

3. Score-to-tag conversion

3.1. Power consumption pattern model-to-tag conversion

The continuous CPLRP and HPLRB scores preserve the quantitative basis of the evaluation, whereas the corresponding discrete tags are introduced as a standardized and interpretable representation layer. This design improves comparability across buildings and supports the subsequent decision-tree-based CLCRP aggregation, in which categorical inputs are more suitable for transparent rule-based user stratification.

According to the power consumption of the building at different times in the next 24 h, its relative position within the daily power-consumption range (i.e., from the minimum to the maximum power value over the next 24 h) is converted into a load-level tag, as shown in Table 2. To transform the continuous power-consumption profile into discrete and interpretable load-level categories, the power consumption is first normalized into a unit-free indicator, denoted by $\tilde{p} \in [0, 1]$, and then partitioned using rule-based thresholds. Specifically, samples with $\tilde{p} \geq 0.7$ are regarded as belonging to the peak period and are therefore assigned the load-level tag of high; samples with $0.3 < \tilde{p} < 0.7$ are considered to fall within the flat period and are assigned the tag of middle; and samples with $\tilde{p} \leq 0.3$ are treated as belonging to the valley period and are assigned the tag of low. This threshold setting is intended to provide a simple and operationally interpretable peak–flat–valley categorization of the normalized load profile, consistent with the common practice of identifying relatively high-load periods using the upper portion of the daily load distribution [25]. It should be noted that such thresholds are adopted here to balance interpretability, operational practicality, and comparability across buildings; therefore, while the tagging logic is transferable, the specific numerical cutoffs may need to be adjusted for other building groups or application scenarios.

3.2. Current peak-load reduction potential score-to-tag conversion

According to the value of CPLRP score, it is converted into CPLRP tag, as shown in Table 3. Accordingly, buildings with $X \geq Q_{0.70}^{CPLRP}$ (top 30%) are tagged as High, those with $Q_{0.30}^{CPLRP} < X < Q_{0.70}^{CPLRP}$ (middle 40%) as Middle, and those with $X \leq Q_{0.30}^{CPLRP}$ (bottom 30%) as Low. In this study, the computed 30th and 70th percentiles are $Q_{0.70}^{CPLRP} = 0.21$ and $Q_{0.30}^{CPLRP} = 0.14$, respectively.

These thresholds are adopted to provide an interpretable categorical partition of the continuous scores while preserving comparability across buildings in the present dataset. Accordingly, the threshold-setting logic is transferable, whereas the exact numerical cutoffs may need to be recalibrated under different data distributions, building portfolios, climate conditions, and operational requirements.

Table 3
CPLRP score-to-tag conversion.

CPLRP score	CPLRP tag
$X \geq Q_{0.70}^{CPLRP}$	High
$Q_{0.70}^{CPLRP} < X < Q_{0.30}^{CPLRP}$	Middle
$X \leq Q_{0.30}^{CPLRP}$	Low

Table 4
HPLRB score-to-tag conversion.

HPLRB score range	HPLRB tag
$X \geq Q_{0.70}^{HPLRB}$	High
$Q_{0.30}^{HPLRB} < X < Q_{0.70}^{HPLRB}$	Middle
$X \leq Q_{0.30}^{HPLRB}$	Low

3.3. Historical peak-load reduction behavior score-to-tag conversion

According to the value of HPLRB score, it is converted into HPLRB tag, as shown in Table 4.

The thresholds in Table 4 are determined by quantiles of the HPLRB score distribution. Accordingly, buildings with $X \geq Q_{0.70}^{HPLRB}$ (top 30%) are tagged as High, those with $Q_{0.30}^{HPLRB} < X < Q_{0.70}^{HPLRB}$ (middle 40%) as Middle, and those with $X \leq Q_{0.30}^{HPLRB}$ (bottom 30%) as Low. In this study, the computed 30th and 70th percentiles are $Q_{0.70}^{HPLRB} = 4$ and $Q_{0.30}^{HPLRB} = 2$, respectively.

3.4. Decision tree to comprehensive low carbon regulation potential tag

We establish a five-level classification scheme for the comprehensive low-carbon regulation potential (CLCRP) tag, namely Highest, High, Middle, Low, and Lowest, as illustrated in Fig. 3.

The CLCRP tag is determined using a rule-based decision tree that hierarchically integrates three key power characteristic tags: the load level tag derived from the power consumption pattern model, the CPLRP tag, and the HPLRB tag. The decision process follows a top–down structure, in which the load level tag serves as the root node, reflecting the structural importance of the user's load magnitude during peak periods.

In the proposed framework, the load-level tag is selected as the root node because it provides the most direct indication of the current scale of available load that may participate in regulation. Before considering historical response behavior or the relative share of flexible end uses, it is first necessary to determine whether the building is operating at a sufficiently high load level during the relevant period. This design is also consistent with previous studies showing that peak-period load characteristics are critical for identifying buildings with strong demand response potential [26,27].

For users with a load level High tag, the CLCRP tag is jointly determined by the CPLRP tag and the HPLRB tag. A High CPLRP tag indicates stronger instantaneous regulation capability, while a High HPLRB tag reflects more reliable historical response behavior. Their combination leads to CLCRP levels ranging from Highest to Lowest, as shown in the upper branch of the decision tree.

For users with a Middle load level tag, the same CPLRP–HPLRB combination logic is applied; however, the resulting CLCRP level is generally shifted downward, reflecting the reduced absolute contribution to peak-load reduction compared with high-load users.

For users with a Low load level tag, the decision tree applies a pruning rule: regardless of the CPLRP or HPLRB tags, the CLCRP tag is directly assigned as Lowest, since limited load magnitude implies insufficient demand-side resources for effective peak-load reduction. In addition, if the CPLRP tag is classified as Low, the user is considered to have insufficient current regulation potential, and the CLCRP tag is also directly assigned as Lowest. This rule reflects a practical

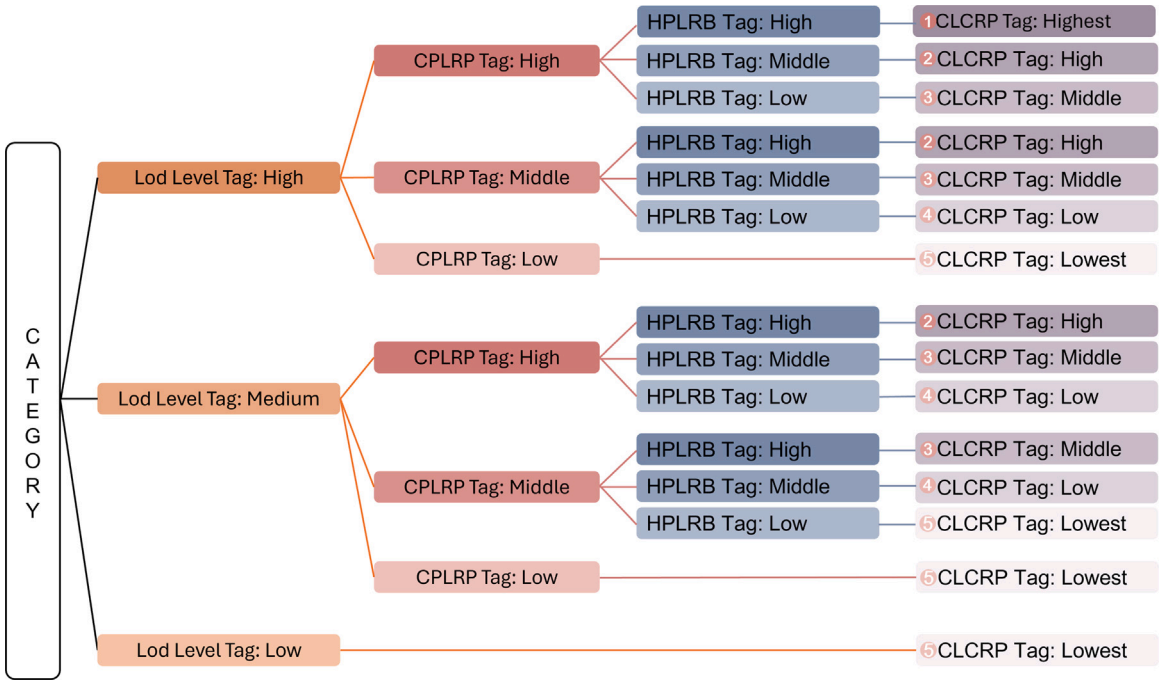


Fig. 3. Decision tree to CLCRP tag.

screening logic: if either the overall load scale or the currently available flexible-load share is insufficient, the building's near-term comprehensive regulation potential should be considered limited regardless of its historical response record.

Through this hierarchical and pruned decision structure, the CLCRP tag provides an interpretable and operational assessment of power users' low-carbon regulation potential.

For users with a High load level tag, the CLCRP tag is jointly determined by the CPLRP tag and the HPLRB tag. A High CPLRP tag indicates stronger instantaneous regulation capability, while a High HPLRB tag reflects more reliable historical response behavior. Their combination leads to CLCRP levels ranging from Highest to Lowest, as shown in the upper branch of the decision tree.

For users with a middle load level tag, the same CPLRP–HPLRB combination logic is applied; however, the resulting CLCRP level is generally shifted downward, reflecting the reduced absolute contribution to peak-load reduction compared with high-load users.

For users with a low load level tag, the decision tree applies a pruning rule: regardless of the CPLRP or HPLRB tags, the CLCRP tag is directly assigned as lowest, since limited load magnitude implies insufficient demand-side resources for effective peak-load reduction.

In addition, a pruning rule is applied across all branches of the decision tree: if the CPLRP tag is classified as low, the user is considered to have insufficient current regulation potential, and the CLCRP tag is directly assigned as lowest. Through this hierarchical and pruned decision structure, the CLCRP tag provides an interpretable and operational assessment of power users' CLCRP.

4. Power consumption pattern model based on CNN-LSTM

4.1. CNN-LSTM model structure

To accurately characterize the power consumption pattern of public buildings, a multi-feature multi-step time-series forecasting model is employed to generate 24-hour ahead load profiles. Such forecasts provide a dynamic representation of daily load evolution and enable continuous updates of power user profiling. Meanwhile, by selecting an appropriate forecasting architecture, a balance can be achieved between prediction accuracy and computational overhead, which is particularly important for deployment in Internet of Things -enabled building environments with limited computing resources.

Convolutional neural network (CNN) are effective at extracting local and short-term patterns from structured time-series data, while long short-term memory (LSTM) networks are well suited for capturing temporal dependencies. By integrating CNN-based feature extraction with LSTM-based temporal modeling, the CNN–LSTM architecture forms a powerful hybrid framework for multi-step load forecasting.

In this work, the CNN–LSTM model is designed to forecast the electricity consumption of a public building over a 24-hour horizon, serving as an input to the dynamic power user profiling process. The CNN component extracts representative features from multi-dimensional input sequences, and the LSTM component models their temporal evolution to generate multi-step predictions. The overall structure of the proposed CNN–LSTM forecasting model is illustrated in Fig. 4.

CNN is widely used for feature extraction from structured sequential data due to their ability to capture local temporal patterns through convolutional operations. In this work, the CNN module employs 1D convolution to process multivariate load-related time-series inputs. Specifically, the input sequence is arranged such that different features are treated as channels, while the convolution is performed along the temporal dimension. The 1D convolution operation can be expressed as

$$s(t) = (x * w)(t) = \sum_{m=0}^{M-1} x(t-m)\, w(m), \tag{4}$$

where $x(\cdot)$ denotes the input sequence, $w(\cdot)$ is the 1D convolution kernel, M is the kernel size, and $s(t)$ is the resulting feature at time step t. For multivariate inputs, each feature dimension is treated as an input channel, and the convolution kernel moves along the time axis to extract local temporal features, which are then passed to subsequent layers for higher-level representation learning.

To introduce nonlinearity, CNN employ activation functions following convolutional operations. A commonly used activation function is the Rectified Linear Unit (ReLU), defined as

$$\mathrm{ReLU}(x) = \max(0, x), \tag{5}$$

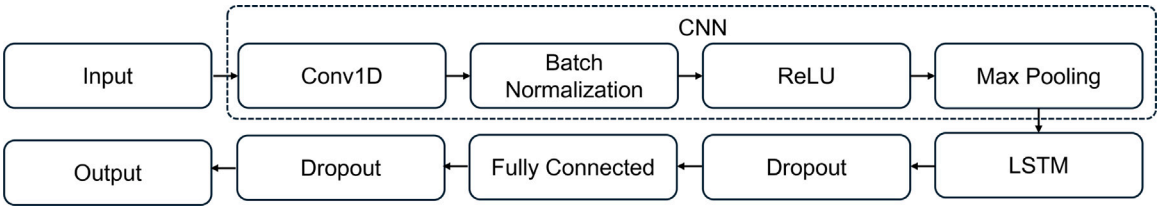


Fig. 4. Structure of CNN–LSTM model.

where x denotes the pre-activation input of a neuron. ReLU improves learning efficiency and alleviates the vanishing gradient problem, thereby facilitating deeper network architectures.

Pooling layers are typically inserted after convolutional layers to downsample feature maps, reduce computational complexity, and enhance robustness to small input variations. Among various pooling strategies, max pooling is most commonly adopted, as it preserves the most salient local features. The extracted features are then passed to subsequent layers for downstream tasks.

An LSTM cell consists of a memory cell and three gating mechanisms: the forget gate, the input gate, and the output gate. These gates regulate the information flow within the network. The forget gate controls how much past information is retained and is given by

$$f_t = \sigma\left(W_f[h_{t-1}, x_t] + b_f\right), \tag{6}$$

where x_t is the current input, h_{t-1} is the previous hidden state, and $\sigma(\cdot)$ denotes the sigmoid activation function.

The input gate determines how much new information is written into the memory cell and is computed as

$$\tilde{C}_t = \tanh\left(W_C[h_{t-1}, x_t] + b_C\right), \tag{7}$$

$$i_t = \sigma\left(W_i[h_{t-1}, x_t] + b_i\right). \tag{8}$$

The memory cell state is then updated according to

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t, \tag{9}$$

where C_t and C_{t-1} denote the current and previous cell states, respectively.

Finally, the output gate controls the information exposed to the next layer:

$$o_t = \sigma\left(W_o[h_{t-1}, x_t] + b_o\right), \tag{10}$$

$$h_t = o_t \cdot \tanh(C_t), \tag{11}$$

where h_t is the current hidden state. Through this gated mechanism, LSTMs can effectively model both short-term dynamics and long-term temporal dependencies.

By combining CNN-based feature extraction with LSTM-based temporal modeling, the CNN–LSTM architecture leverages the complementary strengths of both networks, making it well suited for multi-step load forecasting in power consumption analysis.

4.2. Evaluation metrics

Regression models are typically assessed based on how closely the predicted values approximate the ground truth, with smaller prediction errors indicating better performance. Accordingly, several error-based and goodness-of-fit metrics are adopted to evaluate forecasting accuracy and reliability over multiple prediction horizons.

The mean absolute error (MAE) [19] measures the average magnitude of prediction errors without considering their direction and is defined as:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|, \tag{12}$$

where y_i and $\hat{y}_i$ denote the ground-truth and predicted values at the ith time step, respectively, and n is the total number of forecasted points across all samples.

The root mean squared error (RMSE) is defined as the square root of MSE [28]:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}, \tag{13}$$

which restores the error metric to the same physical unit as the original data. RMSE therefore provides an interpretable measure of average error magnitude while retaining sensitivity to large prediction errors.

The coefficient of determination R^2 [29] measures the proportion of variance in the target variable that is explained by the model and is given by:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}, \tag{14}$$

where $\bar{y}$ denotes the mean of the observed values. An R^2 value of 1 indicates perfect prediction, whereas negative values imply that the model performs worse than a mean-based baseline. In multi-step forecasting tasks, R^2 provides a normalized measure of model fit by jointly accounting for prediction error magnitude and variability in the target series.

5. Case study

5.1. Experimental environment

All experiments are conducted on a laptop equipped with an AMD Ryzen 7 5800H processor (3.20 GHz), 32 GB of DDR4 memory, and an NVIDIA GeForce RTX 3070 Laptop GPU with 8 GB of GDDR6 memory.

For all time-series forecasting models, a sliding-window strategy is adopted, where historical load data from the preceding 14 days (336 time steps) are used as inputs to predict the subsequent 24-hour load profile (24 time steps). The 14-day input window is selected to balance temporal representativeness and computational efficiency. Since the load data are hourly and public-building electricity consumption typically exhibits both short-term continuity and weekly periodicity, a two-week historical window can cover more than one full weekly cycle while still retaining recent operational information. To provide a comprehensive comparison, five representative forecasting models are evaluated, including CNN, LSTM, CNN–LSTM, CNN–LSTM with attention mechanism, and the Transformer model.

Model hyperparameters are specified according to the architectural characteristics of each model, and the detailed configurations are summarized in Table 5. The hyperparameter settings are determined based on preliminary comparative tuning and architecture-specific empirical considerations, with the aim of balancing forecasting accuracy and model complexity for the present dataset, rather than claiming a universally optimal configuration. These configurations include batch size is 64, learning rate is 0.001, dropout rate is 0.3 [30].

To improve the fairness and reproducibility of the model comparison, all five forecasting models were trained under the same optimization budget. Specifically, the maximum number of training epochs was set to 100 for all models, and early stopping was applied by monitoring

Table 5
Hyperparameters of different multi-feature time series forecasting models.

	CNN	LSTM	CNN–LSTM	CNN–LSTM-Attention	Transformer
Batch size	64	64	64	64	64
Learning rate	0.001	0.001	0.001	0.001	0.001
Hidden size	128	128	128	128	–
Drop rate	0.2	0.2	0.3	0.3	0.3
Conv1d layers	2	–	1	1	–
CNN kernel size	3	–	5	5	–
LSTM layers	–	2	1	1	–
Attention heads	–	–	–	4	4
Attention drop	–	–	–	0.1	0.3
Dimensionality	–	–	–	–	64
Encoder layers	–	–	–	–	2

the validation loss with a patience of 10 epochs; the parameter set corresponding to the best validation epoch was retained. Random seeds were fixed at 42 for weight initialization and data shuffling, and each architecture was trained and evaluated over five repeated runs under the same train–test split. We did not perform an exhaustive grid search or random search; instead, a limited preliminary tuning procedure was carried out for each architecture around commonly used settings, and the final configurations in Table 5 were selected under this common training budget. Since the compared models are architecturally heterogeneous, their parameter counts were not forced to be exactly identical; rather, the hidden dimensions and layer depths were kept within a comparable scale so that the comparison reflects architecture-level performance under a consistent computational budget.

5.2. Data preprocessing

The case study is conducted using a subset of the Building Data Genome Project-2 (BDG2) dataset [31], consisting of electricity consumption data from 45 public buildings. The dataset spans from 00:00 on January 1, 2016 to 23:00 on December 31, 2017 with an hourly resolution. The data are chronologically divided into a training set (80%, from January 1, 2016 to July 31, 2017) and a test set (20%, from August 1, 2017 to December 31, 2017).

For reproducibility, the forecasting-model comparison in Section 5.3 is based on these 45 retained public buildings, whereas the downstream profiling demonstration in Section 5.4 further focuses on the 27 office buildings within this retained set in order to keep the final CLCRP illustration within a homogeneous building category. The remaining 18 retained public buildings are therefore not used in the final office-building profiling visualization, although they are included in the forecasting-stage comparison.

Prior to model training, the data are preprocessed to ensure consistency with the research scope and data quality requirements. Only electricity meter readings and their corresponding weather data are retained. Buildings with incomplete records, excessive missing values, or anomalously low consumption levels are excluded to avoid the influence of meter malfunctions or data acquisition errors.

To ensure climatic and behavioral consistency with the target application context, we exclude buildings located in climate zones that are not representative of typical conditions in China. This step mitigates potential biases arising from heterogeneous long-term climatic effects on load behavior, particularly for thermostatically controlled loads.

Feature selection is performed using a Random Forest model to identify the most influential predictors of electricity consumption. Based on the resulting importance rankings, air temperature, sea level pressure, dew temperature, and day of week are selected as input features for subsequent time-series forecasting.

Finally, missing values are handled using an explicit threshold-based interpolation strategy. A building is excluded if more than 5% of its hourly electricity or matched weather records are missing, if the electricity series contains a continuous missing segment longer than 6 h, or if the building shows persistently near-zero consumption indicative of inactive or faulty metering. For the retained buildings, short missing gaps are repaired by linear interpolation along the time axis. When a missing point appears at the beginning or end of the series, forward/backward filling with the nearest valid observation is used after interpolation to avoid boundary nulls. A final check is then performed to confirm that no missing values remain before model training.

5.3. CNN LSTM result analysis

The forecasting performance of five models was comparatively evaluated: CNN–LSTM (M0), CNN (M1), LSTM (M2),CNN–LSTM-Attention (M3), and Transformer (M4). Each model was trained separately for each of the selected public buildings.

The evaluation metrics presented R^2, RMSE, and MAE represent the average values across all public buildings. The comparative experimental results are summarized in Table 6 and illustrated in Fig. 5. Overall, the results show that the standalone LSTM significantly outperforms the standalone CNN, indicating that temporal dependency modeling is the dominant component for the present public-building load forecasting task. Compared with the standalone LSTM, the CNN–LSTM model achieves further improvement, suggesting that the convolutional module contributes useful local temporal feature extraction before recurrent modeling. By contrast, the CNN–LSTM-Attention model does not outperform the standard CNN–LSTM model, implying that the additional attention mechanism does not provide extra benefit under the current dataset and forecasting setting. These results indicate that the final performance gain mainly comes from the combination of temporal dependency modeling and local feature extraction, while additional architectural complexity beyond this combination is not necessarily beneficial.

Fig. 5(a) shows that the CNN–LSTM model achieves the highest average R^2, while Fig. 5(b) and (c) further confirm that it also yields the lowest RMSE and MAE among the evaluated models. This consistent advantage across all three metrics supports the effectiveness of the hybrid CNN–LSTM architecture under the present experimental setting.

However, since the improvement of CNN–LSTM over the standalone LSTM remains relatively modest, the choice between these two models should be understood as a trade-off between forecasting accuracy and computational complexity, particularly for resource-constrained IoT-edge deployment scenarios. In this sense, the CNN–LSTM model can be regarded as a practically effective hybrid solution, while the standalone LSTM remains an attractive alternative when computational simplicity is prioritized.

Fig. 5(a) illustrates the average R^2 for all models on the test set. A higher R^2 score signifies superior model accuracy; consequently, the CNN–LSTM model exhibited the best performance among the five models. Fig. 5(b), and (c) present the average RMSE, and MAE, respectively. For these error metrics, lower values indicate a closer approximation to the actual observed values, and the CNN–LSTM model once again demonstrated optimal performance. Overall, the results suggest that the hybrid CNN–LSTM model provides the best predictive performance among the evaluated models under the present experimental setting. However, since its improvement over the standalone LSTM is relatively modest, the choice between CNN–LSTM and LSTM should be understood as a trade-off between forecasting accuracy and computational complexity, particularly for resource-constrained IoT-edge deployment scenarios. In this sense, the CNN–LSTM model can be regarded as a practically effective hybrid solution, while the standalone LSTM remains an attractive alternative when computational simplicity is prioritized.

Among all evaluated architectures, the hybrid CNN–LSTM model achieved the best overall performance, yielding the highest R^2 score (0.8349) and the lowest values across all error metrics (RMSE = 28.93, MAE = 18.75), as shown in Table 7. This indicates that the combination

Table 6
Comparison of different multi-feature time series forecasting models.

	Model	R^2	RMSE (kW)	MAE (kW)
M0	CNN–LSTM	0.8349	28.93	18.75
M1	CNN	0.5656	45.58	33.14
M2	LSTM	0.8116	31.51	19.33
M3	CNN–LSTM-Attention	0.7463	38.55	26.97
M4	Transformer	0.7621	35.67	23.21

Table 7
Sensitivity analysis results of key hyperparameters for the CNN–LSTM model.

	Batch size	CNN kernel size	LSTM layers	R^2	RMSE (kW)	MAE (kW)
S1	32	5	1	0.8280	29.75	18.90
S2	64	5	1	0.8349	28.93	18.75
S3	128	5	1	0.8297	29.42	19.05
S4	64	4	1	0.8250	29.02	18.96
S5	64	6	1	0.8336	29.21	18.78
S6	64	5	2	0.8317	29.38	19.34
S7	64	5	3	0.8254	30.40	20.24

of convolutional layers for local feature extraction and LSTM layers for temporal dependency modeling is particularly effective for this dataset. The standalone LSTM model also demonstrated strong performance, surpassing the CNN, CNN–LSTM-Attention, and Transformer models on every metric. This observation indicates that temporal dependency modeling is a key factor for the present forecasting task. At the same time, the slightly better performance of CNN–LSTM suggests that the convolutional component provides additional benefits by extracting short-term local patterns and feature interactions before temporal sequence modeling.

For the hyperparameter sensitivity analysis, a controlled-variable strategy was adopted. Specifically, S1–S3 are designed as the comparison group for batch size, where only the batch size is varied while the CNN kernel size and the number of LSTM layers are kept unchanged. Similarly, S2, S4, and S5 formed the comparison group for CNN kernel size, and S2, S6, and S7 formed the comparison group for the number of LSTM layers. Since S2 is selected as the baseline setting, it is included in all three comparison groups to evaluate the effect of each hyperparameter change on model performance. During each sensitivity test, only one hyperparameter is varied while the others are fixed at the baseline setting. The results show that the proposed model is generally robust within the tested hyperparameter ranges, as the performance fluctuations remain relatively limited. For batch size, the setting of 64 achieves the best overall balance, with the highest R^2 (0.8349) and the lowest RMSE (28.93 kW) and MAE (18.75 kW), while both smaller and larger batch sizes lead to slight performance degradation. Among the tested factors, the number of LSTM layers has the most pronounced impact. Increasing the LSTM depth from 1 to 3 gradually reduces forecasting accuracy, with R^2 decreasing from 0.8349 to 0.8254, while RMSE and MAE increase noticeably. This suggests that a relatively shallow recurrent structure is more suitable for the present public-building load forecasting task, whereas deeper LSTM stacks may introduce unnecessary complexity and optimization difficulty. Overall, the sensitivity analysis supports that the adopted parameter setting provides a reasonable trade-off between predictive accuracy and model complexity.

Notably, the Transformer model, despite its recognized capability for modeling long-range dependencies, did not yield superior performance compared to the simpler LSTM-based approaches in this application. This outcome may be jointly attributed to the limited size of the dataset, the relatively short input sequence length (14 days), and the operational characteristics of public-building loads themselves. In particular, public-building electricity consumption often exhibits strong temporal continuity, daily or weekly regularity, and local dependence associated with occupancy schedules and HVAC operation, which may

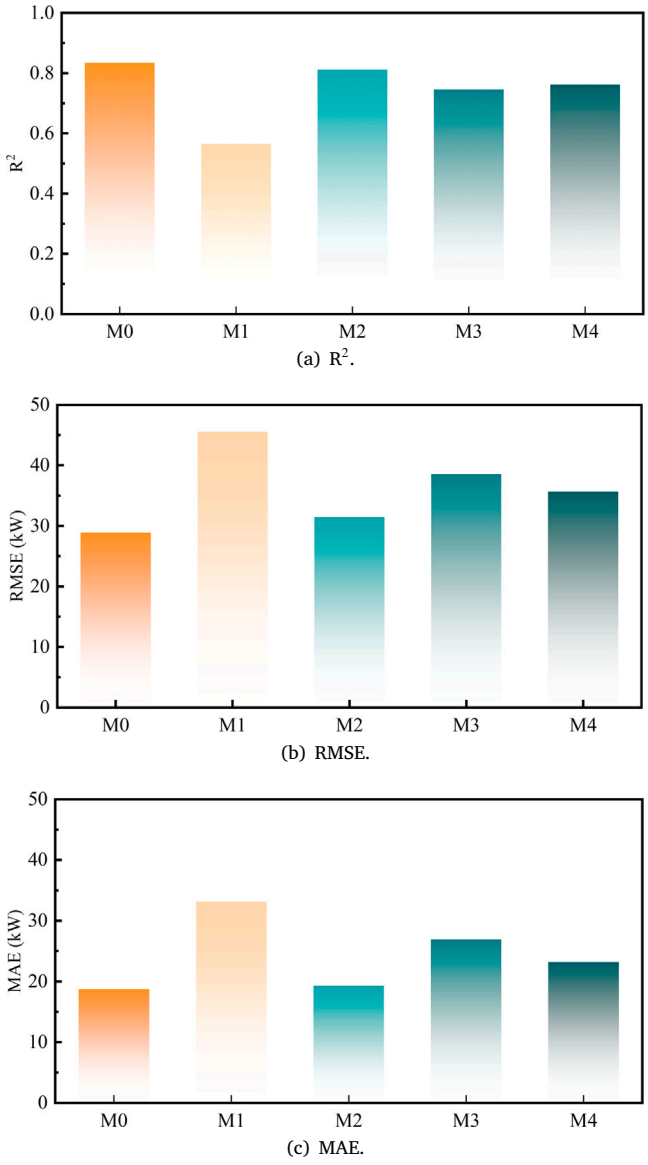


Fig. 5. Comparison of different multi-feature time series forecasting models: (a) Coefficient of Determination (R^2) ; (b) Root Mean Squared Error (RMSE); (c) Mean Absolute Error (MAE).

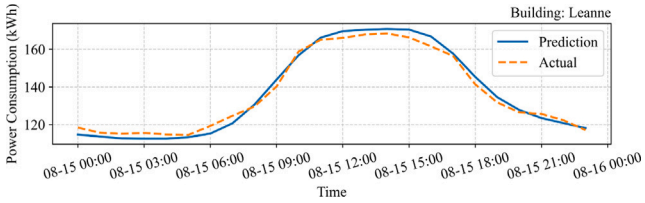
favor LSTM-based or hybrid CNN–LSTM-based architectures under the present forecasting setting.

Overall, these results highlight the effectiveness of the hybrid CNN–LSTM model as the power consumption pattern model, successfully balancing predictive accuracy and computational efficiency.

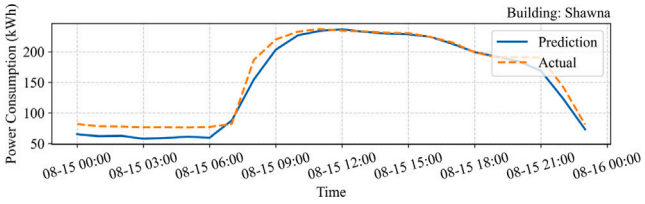
5.4. Power user profiling results

Data from 27 office buildings in the BDG2 dataset, spanning August 1 to August 14, 2017, were selected to construct the user profiles for August 15. First, it was necessary to validate the capability of the power consumption pattern model (CNN–LSTM) to accurately forecast the consumption curves for these buildings on August 15, 2017.

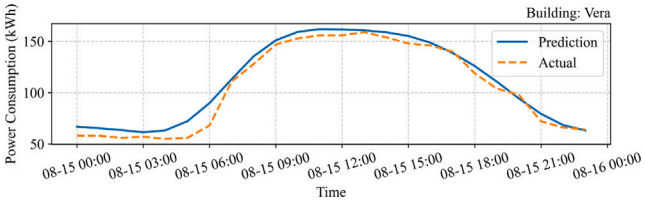
As illustrated in Fig. 6, the predicted consumption curves for these typical buildings demonstrate a strong correlation with the actual observations, effectively capturing the primary load evolution trends throughout the day. Although the actual curves exhibit localized fluctuations, which are likely attributable to stochastic occupant behavior



(a) Building Leanne.



(b) Building Shawna.



(c) Building Vera.

Fig. 6. The forecasting results of buildings: (a) Building Leanne; (b) Building Shawna; (c) Building Vera.

and other subjective factors, the model successfully captures the underlying power-consumption patterns. Furthermore, when aggregating the load across the broader building portfolio, such individual stochastic noise tends to be smoothed out. This compensatory effect is further evidenced by the high degree of fitness between the predicted and actual total consumption curves, confirming the model's robustness in both individual building forecasting. In the current study, the user sensitivity tags are intended to represent regulation willingness rather than regulation capability. Therefore, they are retained as an extensible dimension of the overall profiling framework, but are not yet directly incorporated into the present decision-tree-based CLCRP evaluation.

The decision-tree mapping in this work is a deterministic rule-based classifier, not a trained prediction model. It converts the pre-computed load-level tag, CPLRP tag, and HPLRB tag into the comprehensive CLCRP label, and therefore should not be evaluated by model-oriented metrics such as F1 score.

Instead, we evaluate the downstream tagging reliability of the forecasting module by checking whether the load-level tags derived from the forecasted 24-h load curves agree with those derived from the actual 24-h load curves. In both cases, the same within-day min–max normalization and the same 0.3/0.7 thresholds defined in Section 3 are used to assign High/Middle/Low categories.

This evaluation focuses on the key question of whether forecast errors change the categorical load-level assignment. The results in Table 8 indicate that the forecast-derived load-level tags remain highly consistent with those derived from actual loads, with an average tag-level accuracy of 95.77%.

5.5. Low-carbon regulation result analysis

To demonstrate the power consumption pattern model-to-tag conversion procedure, the public building Hog office Valda is selected as a representative case. As illustrated in Fig. 7, the predicted hourly power consumption on August 15, 2017 is transformed into categorical load-level tags, including High, Middle, and Low. The tagging strategy is based on the relative position of each predicted value within the daily load distribution, with the top 30%, middle 40%, and bottom 30% corresponding to High, Middle, and Low, respectively.

Table 8
Tag-level consistency between forecast-derived and actual-load-derived load-level tags for representative buildings.

Building	Accuracy (%)
Leanne	95.83
Shawna	91.30
Vera	100.00
Average	95.77

To clearly illustrate the CPLRP score-to-tag conversion process, twelve representative public buildings are selected for visualization. Fig. 8 presents this conversion process, where the area between the dotted line and the x-axis denotes the CPLRP score contribution from EV charging stations, while the area between the solid line and the dotted line represents the contribution from HVAC systems. In this respect, Fig. 8 not only illustrates the CPLRP score-to-tag conversion process, but also reveals the end-use composition behind the CPLRP score. These results suggest that the current regulation potential is determined not only by the total load level, but also by the relative contribution of flexible end uses that can actually participate in regulation.

For the HPLRB score-to-tag conversion, all selected public buildings are considered, as shown in Fig. 9. The predefined thresholds enable effective differentiation of buildings as their HPLRB tags evolve over time. In addition, the time decay function $\delta_h(t_i)$ ensures that recent low-carbon regulation participation exerts a stronger influence on the aggregated HPLRB score. This mechanism encourages sustained participation while allowing newly involved users to rapidly accumulate scores and achieve a High tag through high-quality responses within a relatively short period.

Finally, by integrating load-level tags, CPLRP tags, and HPLRB tags through the constructed decision tree, the hourly CLCRP tags of the 27 office buildings on August 15, 2017 are obtained. The results are visualized using the heatmap in Fig. 10 where darker colors indicate Higher CLCRP tag levels. It presents the building-level hourly CLCRP tags of the selected office buildings, thereby illustrating the temporal heterogeneity of regulation potential across individual buildings.

The hourly CLCRP tags facilitate an aggregate analysis of buildings of the same type. Fig. 11 presents the hourly distribution of the 27 office buildings, categorized by their CLCRP tags. Based on the building-level CLCRP results in Fig. 10, Fig. 11 further summarizes the hourly distribution of CLCRP categories across the office-building portfolio, thereby providing an aggregate view of the temporal evolution of regulation potential.

Subsequently, the actual power consumption corresponding to each hourly CLCRP tag level was aggregated, as shown in the stacked area chart in Fig. 12. Fig. 12 further maps the hourly CLCRP distribution in Fig. 11 to the corresponding aggregated actual power consumption, so as to examine how the assessed regulation potential relates to the load composition of the building cluster.

It is conventionally assumed that the period of highest electricity consumption possesses the greatest peak-load reduction potential. However, as evidenced by Fig. 12, the user profiling results obtained in this paper diverge from this assumption. Specifically, the time identified by the CLCRP tags as having the highest reduction potential does not necessarily align with the time of highest total electricity consumption. This discrepancy may be attributed to the fact that DR-participating loads primarily consist of HVAC and EV charging stations; these specific

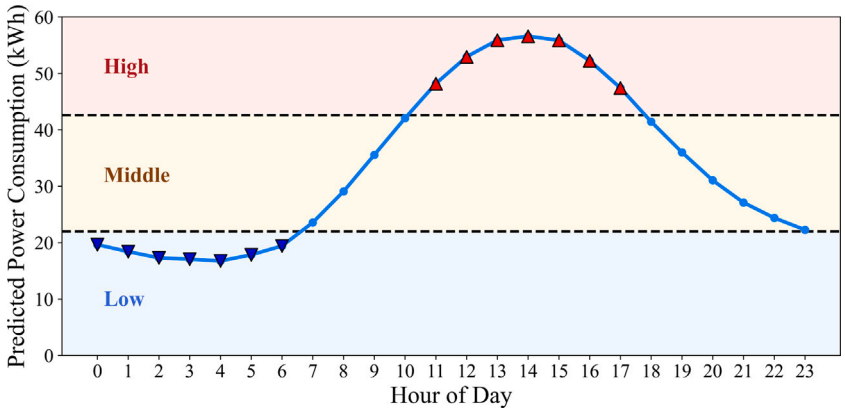


Fig. 7. Power consumption pattern model-to-tag conversion tagging using Hog office Valda as an example.

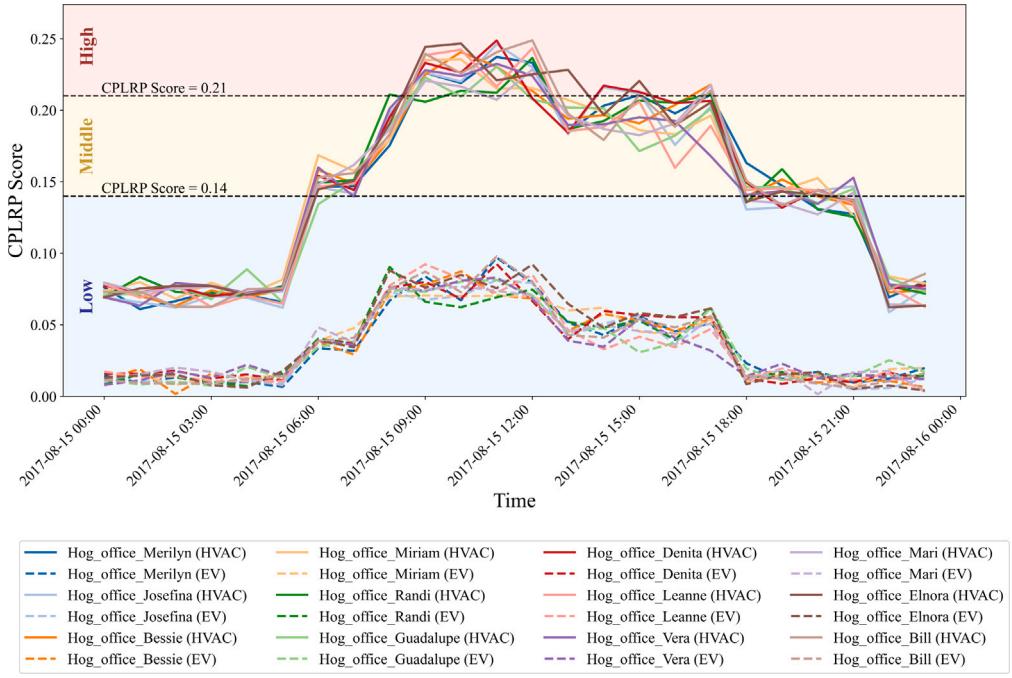


Fig. 8. CPLRP score-to-tag conversion for 12 randomly selected buildings.

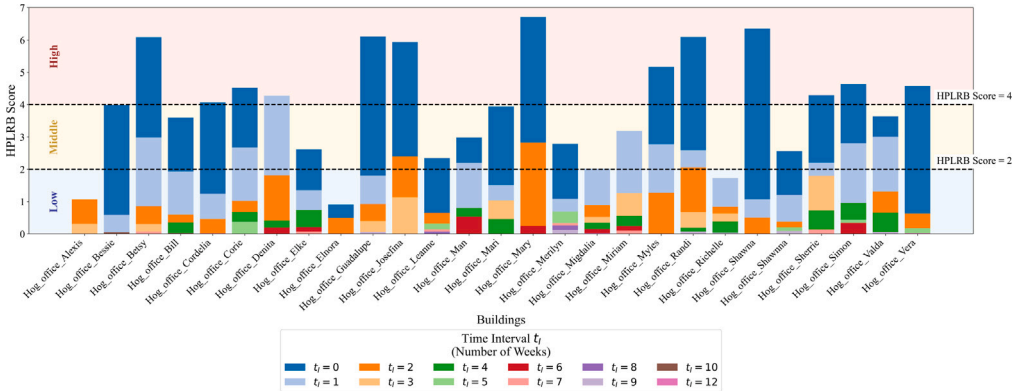


Fig. 9. HPLRB score-to-tag conversion for all 27 selected buildings.

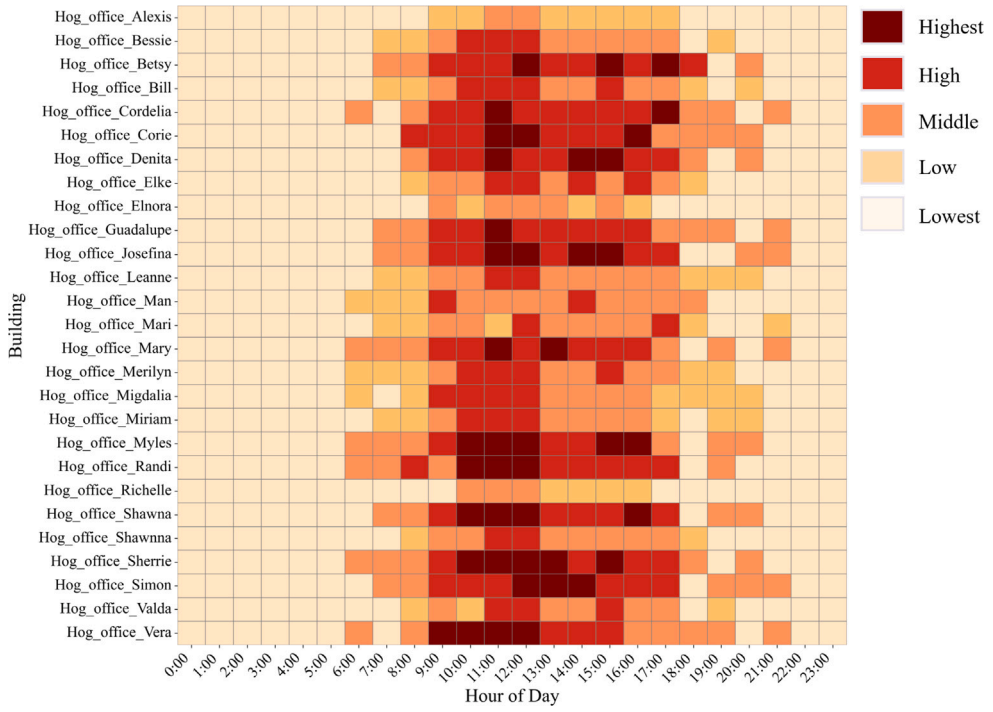


Fig. 10. Hourly CLCRP tag for 27 selected office buildings on August 15, 2017.

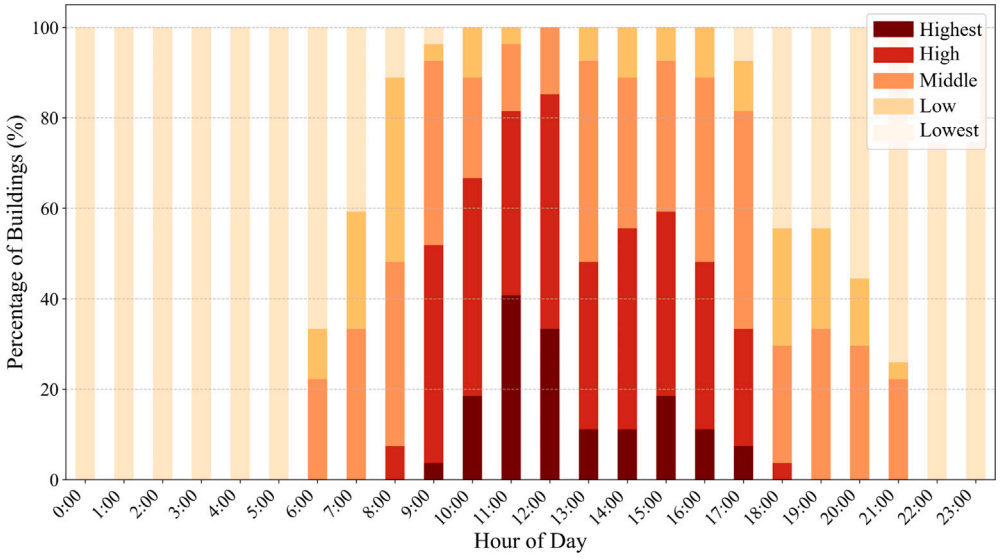


Fig. 11. Hourly distribution of all 27 office buildings by CLCRP tags on August 15, 2017.

loads may not constitute their largest proportion of the total load precisely when the total load is at its peak.

6. Conclusion and future work

6.1. Conclusion

User profiling is a critical enabling technology for evaluating the low-carbon regulation potential of power users, particularly in the context of public buildings with complex and time-varying load characteristics. To address the need for accurate and dynamic assessment of low-carbon regulation potential, this paper proposes a deep-learning-based power user profiling framework for public buildings, integrating time-series forecasting with a structured power user characteristic tag system.

The proposed method organizes user characteristic tags into three categories: basic building information tags, key power characteristic tags, and user sensitivity tags, thereby providing a comprehensive and interpretable representation of user-side low-carbon regulation potential. A power consumption pattern model based on multi-step time-series forecasting is employed to capture the temporal evolution of building loads, enabling daily dynamic updates of power user profiles. In addition, user scoring mechanisms are designed by drawing on experience-based scoring principles, allowing historical regulation behavior and current regulation potential to be jointly quantified.

Case study results demonstrate that the proposed framework can dynamically derive the hourly comprehensive low-carbon regulation potential of public buildings, effectively linking power consumption patterns with regulation capability assessment.

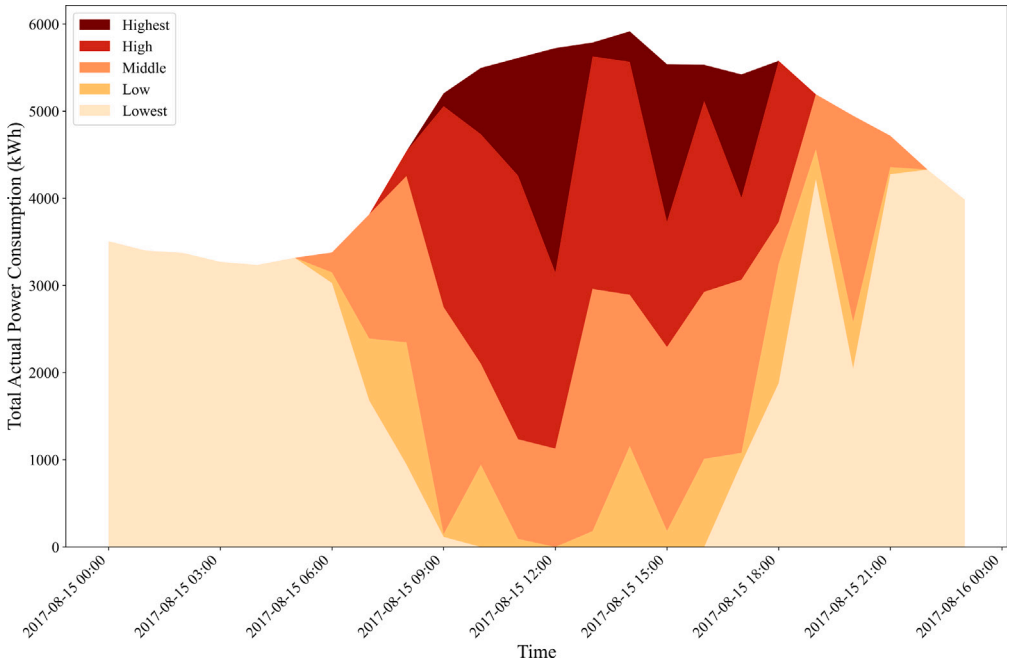


Fig. 12. Actual power consumption corresponding to hourly CLCRP tags of all 27 office buildings on August 15, 2017.

6.2. Future work

Future work will consider three targeted extensions of the proposed framework. From a methodological perspective, the current dynamic power user profiling approach can be further enriched by incorporating uncertainty propagation from load forecasting into current peak-load reduction potential and comprehensive low-carbon regulation potential derivation. By explicitly mapping forecast uncertainty into user scores and tags, the framework can distinguish between expected regulation potential and confidence-aware regulation potential, thereby enhancing robustness without altering the core profiling structure.

From an application perspective, future research will focus on coupling the proposed profiling results with specific low-carbon regulation decision layers, such as user ranking, capacity allocation, or regulation signal design.

Another important direction for future work is to validate the CLCRP tags against actual demand response events or simulation-based regulation-capacity benchmarks, so as to further examine the extent to which the proposed profiling results correspond to realized usable regulation capacity.

CRediT authorship contribution statement

Jinfeng Li: Writing – original draft, Conceptualization. **Run Tang:** Writing – original draft, Validation. **Siyu Jiang:** Writing – review & editing, Writing – original draft, Software. **Yu Yao:** Writing – review & editing, Writing – original draft, Methodology. **Jiahe Li:** Writing – original draft, Data curation, Conceptualization. **Yishun Zhu:** Writing – original draft, Validation. **Ziyan Pan:** Writing – original draft, Data curation. **Guanbin Huang:** Conceptualization, Validation, Writing – review & editing. **Hongxun Hui:** Supervision, Data curation, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Hongxun Hui reports financial support was provided by China Southern Power Grid Co Ltd. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding sources

This paper is supported by the China Southern Power Grid Technology Project (GDKJXM20231411).

Acknowledgments

The authors gratefully acknowledge the Science and Technology Development Fund, Macau SAR with 0050/2025/AIJ.

Data availability

Data will be made available on request.

References

[1] Li J, Wu Q, Tan J, Chen Y, Zhang X. Carbon trading markets: Literature review on mechanisms, accounting, and market models coupled with energy markets. Int J Electr Powe Energy Syst 2026;174:111487. http://dx.doi.org/10.1016/j.ijepes.2025.111487.

[2] Jia X, Xia Y, Yan Z, Gao H, Qiu D, Guerrero JM, et al. Coordinated operation of multi-energy microgrids considering green hydrogen and congestion management via a safe policy learning approach. Appl Energy 2025;401:126611. http://dx.doi.org/10.1016/j.apenergy.2025.126611.

[3] Hwang H, Yoon AY, Lee GS, Moon SI. Evaluating demand flexibility of commercial building HVAC systems in South Korea. Inter J Electr Power Energy Syst 2025;169:110769. http://dx.doi.org/10.1016/j.ijepes.2025.110769.

[4] Alrumayh O, Bhattacharya K. Flexibility of residential loads for demand response provisions in smart grid. IEEE Trans Smart Grid 2019;10(6):6284–97. http://dx.doi.org/10.1109/TSG.2019.2901191.

[5] Wang L, Weng Z, Li R, Zhang Y, Wang S, Li M. Collaborative operational model for shared hydrogen storage and energy prosumers: A quantum game approach. Inter J Hydrog Energy 2026;204:153266. http://dx.doi.org/10.1016/j.ijhydene.2025.153266.

[6] Jiang S, Hui H, Song Y. Credible demand response capacity evaluation for building hvac systems based on grey-box models. Appl Energy 2025;395:126144. http://dx.doi.org/10.1016/j.apenergy.2025.126144.

[7] Tifoura K, Meliani H, Mahrane A. Home energy management system based on applied real-time load scheduling for self-consumption enhancement. Energy Build 2025;345:116107. http://dx.doi.org/10.1016/j.enbuild.2025.116107.
[8] Lu C, Cui J, Wang H, Yi H, Wu C. Privacy preserving user energy consumption profiling: From theory to application. IEEE Trans Smart Grid 2024;15(2):2417–25. http://dx.doi.org/10.1109/TSG.2023.3315690.
[9] Hou Q, Yang Q, Jin X, Dai N. Voltage regulation enhanced volt/var control with discretization-timescale-dependent stability condition in active distribution networks. IEEE Trans Sustain Energy 2025;1–13. http://dx.doi.org/10.1109/TSTE.2025.3646684.
[10] Chen T, Li L, Fan Y, Liu Y, Xu J. A regional short-term load forecasting model based on a cross-attention causal spatio-temporal convolutional network. Inter J Electr Power Energy Syst 2025;172:111275. http://dx.doi.org/10.1016/j.ijepes.2025.111275.
[11] Zhang Y, Xu L, Hong Y, Yang D. Short-term power load forecasting and management based on quadratic decomposition and deep learning. Inter J Electr Power Energy Syst 2026;176:111690. http://dx.doi.org/10.1016/j.ijepes.2026.111690.
[12] Zhang Z, Hui H, Song Y. Response capacity allocation of air conditioners for peak-valley regulation considering interaction with surrounding microclimate. IEEE Trans Smart Grid 2025;16(2):1155–67. http://dx.doi.org/10.1109/TSG.2024.3482361.
[13] Zhang Z, Wang Y, Wang C, Su Y, Wang Y, Dai Y, et al. Distributed chance-constrained optimal dispatch for integrated energy system with electro-thermal couple and wind-storage coordination. IEEE Trans Ind Appl 2025;61(1):833–46. http://dx.doi.org/10.1109/TIA.2024.3472653.
[14] Jin X, Qi F, Wu Q, Mu Y, Jia H, Yu X, et al. Integrated optimal scheduling and predictive control for energy management of an urban complex considering building thermal dynamics. Inter J Electr Power Energy Syst 2020;123:106273. http://dx.doi.org/10.1016/j.ijepes.2020.106273.
[15] Piscitelli MS, Brandi S, Capozzoli A. Recognition and classification of typical load profiles in buildings with non-intrusive learning approach. Appl Energy 2019;255:113727. http://dx.doi.org/10.1016/j.apenergy.2019.113727.
[16] Sandels C, Widén J, Nordström L. Simulating occupancy in office buildings with non-homogeneous markov chains for demand response analysis. In: IEEE power and energy society general meeting. 2015, 7285865. http://dx.doi.org/10.1109/PESGM.2015.7285865.
[17] Yang J, Zhao J, Wen F, Dong Z. A model of customizing electricity retail prices based on load profile clustering analysis. IEEE Trans Smart Grid 2019;10(3):3374–86. http://dx.doi.org/10.1109/TSG.2018.2825335.
[18] Ayón X, Gruber JK, Hayes BP, Usaola J, Prodanovic M. An optimal day-ahead load scheduling approach based on the flexibility of aggregate demands. Appl Energy 2017;198:1–11. http://dx.doi.org/10.1016/j.apenergy.2017.04.038.
[19] Bampoulas A, Pallonetto F, Mangina E, Finn DP. A Bayesian deep-learning framework for assessing the energy flexibility of residential buildings with multicomponent energy systems. Appl Energy 2023;348:121576. http://dx.doi.org/10.1016/j.apenergy.2023.121576.
[20] Chen S, Gan L, Chen C, Yu K, Pi H, Qian Z, et al. Demand-response oriented multi-dimension refined portrait of adjustable resources based on load and survey data fusion. Front Energy Res 2022;Volume 10 - 2022. http://dx.doi.org/10.3389/fenrg.2022.968368.
[21] Ramírez-Mendiola JL, Grünewald P, Eyre N. Linking intra-day variations in residential electricity demand loads to consumers' activities: What's missing? Energy Build 2018;161:63–71. http://dx.doi.org/10.1016/j.enbuild.2017.12.012.
[22] Markovič R, Gosak M, Grubelnik V, Marhl M, Virtič P. Data-driven classification of residential energy consumption patterns by means of functional connectivity networks. Appl Energy 2019;242:506–15. http://dx.doi.org/10.1016/j.apenergy.2019.03.134.
[23] Trotta G. An empirical analysis of domestic electricity load profiles: Who consumes how much and when? Appl Energy 2020;275:115399. http://dx.doi.org/10.1016/j.apenergy.2020.115399.
[24] Flygare C, Nystrand M, Eriksson R, Castellucci V. Load profiling using a two-step clustering framework: Capturing electricity usage variability based on temporal and ambient factors. Int J Electr Powe Energy Syst 2025;171:110924. http://dx.doi.org/10.1016/j.ijepes.2025.110924.
[25] Marmaras C, Javed A, Cipcigan L, Rana O. Predicting the energy demand of buildings during triad peaks in GB. Energy Build 2017;141:262–73. http://dx.doi.org/10.1016/j.enbuild.2017.02.046.
[26] Zhang Y. An assessment and ranking method on demand response potential for urban-scale buildings based on the energy usage portrait. Energy 2025;335:138196. http://dx.doi.org/10.1016/j.energy.2025.138196.
[27] Ma Z. Demand response through ventilation and latent load control for typical medium commercial office buildings in different humid climates. Appl Energy 2024;373:123940. http://dx.doi.org/10.1016/j.apenergy.2024.123940.
[28] Kim H, Hu Y, Ye K, Lu N. A novel vision transformer based load profile analysis using load images as inputs. In: IEEE power and energy society general meeting. 2024, http://dx.doi.org/10.1109/PESGM51994.2024.10688520.
[29] Wang Z, Hong T, Piette MA. Building thermal load prediction through shallow machine learning and deep learning. Appl Energy 2020;263:114683. http://dx.doi.org/10.1016/j.apenergy.2020.114683.
[30] Ruan G, Kirschen DS, Zhong H, Xia Q, Kang C. Estimating demand flexibility using Siamese LSTM neural networks. IEEE Trans Power Syst 2022;37(3):2360–70. http://dx.doi.org/10.1109/TPWRS.2021.3110723.
[31] Miller C, Kathirgamanathan A, Picchetti B, Arjunan P, Park JY, Nagy Z, et al. The building data genome project 2, energy meter data from the ASHRAE great energy predictor III competition. Sci Data 2020;7:368. http://dx.doi.org/10.1038/s41597-020-00712-x.