IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 23 July 2025, accepted 4 August 2025, date of publication 6 August 2025, date of current version 21 August 2025.
Digital Object Identifier 10.1109/ACCESS.2025.3596545

RESEARCH ARTICLE

Minimum EVCS Aggregation Requirements for Reliable Customer Baseline Load in Electricity Markets

YONGWOO JEE, (Graduate Student Member, IEEE), AND JINHO KIM, (Member, IEEE)
Department of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea
Corresponding author: Jinho Kim (jeikim@gist.ac.kr)

This work was supported by Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Trade, Industry and Energy (MOTIE) of Republic of Korea under Grant 20226210100020 and Grant 20226210100070.

ABSTRACT The electrification of the transportation sector, driven by decarbonization policies, can significantly impact power demand patterns. Strategically integrated electric vehicles (EVs) can serve as flexible grid resources by enabling load shifting, voltage regulation, and frequency stabilization. In particular, to evaluate their potential participation in the wholesale electricity market, it is essential to assess whether aggregated electric vehicle charging stations (EVCSs) meet the eligibility and reliability criteria required for existing demand response (DR) programs. This study analyzes the reliability of aggregated EVCSs based on their aggregation level using an average-based Customer Baseline Load (CBL) method, performance metrics such as RRMSE and RMAE, and a two-stage random sampling approach. The analysis utilizes load data from 873 EVCSs located in Jeju, Korea, as of 2023. The results show that as the level of aggregation increases, both RRMSE and RMAE decrease, significantly reducing CBL estimation errors. Furthermore, it was confirmed that aggregating EVCSs beyond 10 MW of capacity or 500kWh of average hourly load achieves reliability levels comparable to those of traditional DR resources. These findings provide a practical foundation for the integration of EV-based loads into wholesale market participation and grid control strategies, while contributing to the development of CBL methods and evaluation metrics for small-scale resources.

INDEX TERMS EVCS aggregation effect, EVCS CBL calculation, EVCS data analysis, EVCS integration into DR resources, minimum EVCS capacity.

I. INTRODUCTION

Since the Paris Agreement, the power sector has actively pursued various decarbonization strategies, with the electrification of transportation emerging as a key initiative [1], [2], [3]. Globally, the number of electric vehicles (EVs) exceeded 45 million in 2023—more than four times the 2020 level—and is projected to surpass 400 million battery electric light-duty vehicles (BEVs) by 2035 under the Announced Pledges Scenario [4]. Beyond reducing emissions, large-scale EV adoption introduces significant structural changes to power system operations, necessitating strategic grid integration policies. The growing EV demand, if not properly coordinated with the power grid, may lead to increased peak demand and grid reliability concerns [5]. However, when strategically integrated, EVs can serve as flexible resources, mitigating demand fluctuations through optimized charging and discharging strategies [6], [7]. Therefore, it is crucial to establish strategies that enable EVs to function as flexible resources without negatively impacting the power system.

EVs are being explored for various applications beyond transportation, particularly in power system operations, and their utilization can be categorized into three main levels: load-side, distribution-level, and transmission-level applications. On the load side, research has focused on optimizing EV charging schedules to align with pricing schemes [8]. Additionally, studies have investigated the integration of EVs into energy management systems (EMS)

The associate editor coordinating the review of this manuscript and approving it for publication was Furkan Ahmad.

tailored for residential and non-residential environments to mitigate peak electricity demand [9], [10], [11], [12]. At the distribution level, EVs have been utilized to mitigate the variability of renewable energy sources within distribution network and to balance load through valley filling and peak shaving strategies, thereby flattening demand curves [13], [14]. Furthermore, EVs play a significant role in voltage stabilization within distribution networks, ensuring stable grid operation [15]. On the transmission level, EV scheduling strategies aim to balance energy markets by reducing load fluctuations and peak demand [16], [17], [18]. Moreover, EVs have been studied as potential providers of ancillary services, such as reserve power and primary frequency regulation, to enhance grid reliability and stability [19], [20], [21]. EVs are primarily leveraged at all levels of the power system to reduce peak demand and smooth load profiles. Additionally, depending on grid operating conditions, they contribute to voltage regulation and frequency stabilization, highlighting their potential as a multifunctional and flexible resource in modern power systems. To effectively fulfill these roles, it is essential to consider the characteristics of EVs as demand-side resources. Furthermore, a phased approach to EV utilization is necessary—particularly to evaluate whether EVs, which are operated based on optimized charging strategies at the load level under tariff schemes, can be further leveraged at the distribution and transmission levels.

Unlike conventional generation resources, EVs function primarily as electricity consumers, with grid regulation serving as a secondary function. The utilization of demand-side resources in power systems can be broadly categorized into two main approaches: price-based demand response(DR) programs and incentive-based DR programs, such as those implemented in electricity markets or utility programs [22]. Price-based DR programs adjust demand according to tariff structures, such as Time-of-Use (ToU) and Real-Time Pricing (RTP), where EV charging management strategies tailored to tariff schemes fall under price-based DR [8], [9], [10], [11], [12]. Otherwise, incentive-based DR regulates demand based on signals from a central entity or allocated bid quantities, with EVs participating in incentive-based DR programs for various purposes, including distribution grid stabilization and electricity cost minimization [13], [14], [15], [16], [17], [18], [19], [20], [21]. The DR programs for EVs follow a sequential approach. Primarily, each EV demand on the load side is shaped through tariff structures imposed by retailers, establishing baseline consumption patterns at the EV charging station (EVCS) level. EVCSs, due to their fixed geographical locations and capability to aggregate multiple EVs, are inherently more suitable for centralized control compared to individual EVs. Consequently, these aggregated demand patterns present opportunities for further utilization at both the distribution and transmission levels, particularly within incentive-based DR programs such as wholesale DR markets.

Facilitating a seamless transition from price-based DR to incentive-based DR for EVs necessitates a thorough assessment of the viability of EVCSs within existing DR market structures. Participation in such markets typically requires formal registration, which involves evaluation across several critical dimensions, including minimum capacity thresholds, customer baseline load (CBL) calculation methodologies, and mechanisms for ensuring response reliability. The minimum capacity required for DR participation varies across markets. For example, California Independent System Operator (CAISO), and Pennsylvania-New Jersey-Maryland Interconnection (PJM) demand response programs require a minimum capacity of 100 kW, while Midcontinent Independent System Operator (MISO) sets 100 kW as the threshold for demand response programs but enforce a 1 MW minimum capacity for market bidding resources [23], [24], [25], [26]. In contrast, Korea Power Exchange (KPX) operates three separate programs with minimum capacities of 2 MW, 10 MW, or no restrictions [27]. Similarly, CBL (Customer Baseline Load) is typically calculated using the Average Method, but some markets adopt regression analysis and Control Group-based approaches [28]. For EV supply equipment (EVSE)-based CBL determination, CAISO applies 10/10 and 5/10 methods, whereas KPX utilizes Mid 4/6 and Mid 8/10 methods, which select the median from 4 or 8 days out of the past 6 or 10 respectively [24], [27]. Additionally, ongoing research explores control group approaches that more accurately reflect load characteristics [29], [30]. Reliability assessment of demand resources commonly relies on the Relative Root Mean Squared Error (RRMSE) metric. For example, PJM and KPX enforce 20% and 30% RRMSE thresholds, respectively, while MISO allows up to 12% deviation from bid quantities before imposing penalties [25], [27], [31]. Considering these factors, EVSE participation in demand response markets should be structured at a minimum aggregation capacity level that falls within the operational capabilities of market operators. Furthermore, before utilizing EVCS loads, it is crucial to assess whether EVCS load patterns are consistent and predictable to accurately calculate the baseline, as this plays a key role in ensuring system reliability [32]. Table 1 categorizes and summarizes the relevant literature according to key thematic areas.

This paper aims to evaluate the feasibility of registering EVCSs as DR resources by empirically analyzing the relationship between aggregation levels and the accuracy of CBL estimation. While many previous studies have explored the potential of EVs in demand-side management, they have primarily focused on individual EV charging patterns or employed advanced forecasting techniques—such as machine learning or deep learning—to improve prediction accuracy at the device level [9], [10], [32], [33]. However, these approaches often overlook key practical requirements for DR market participation, such as minimum capacity thresholds, baseline estimation feasibility at

TABLE 1. Keyword-based literature grouping and key contributions.

Keyword / Theme	Key Contributions in Literature	References
Load-side EV Utilization (Price-based DR)	Focus on individual EV charging optimization and integration with home/enterprise EMS under price-based DR schemes.	[8–12]
Distribution-level EV Utilization (Incentive-based DR)	Peak shaving, valley filling, voltage stabilization within distribution networks using aggregated EV loads.	[13–15]
Transmission-level EV Utilization (Incentive-based DR)	Provide wholesale ancillary services (e.g., reserve, frequency regulation) and balances market load fluctuations.	[16–21]
DR Market Requirements	Outlines participation thresholds, CBL computation methods, and reliability criteria in various DR markets (e.g., PJM, CAISO, MISO, KPX).	[22–26, 31]
CBL Methods and Evaluation	Discuss average, regression, and control group-based CBL methodologies; importance of accurate baselines.	[24, 27–30, 32]

aggregated levels, and reliability standards. In particular, while previous research suggests that aggregating EVCS loads can reduce forecasting uncertainty, some methodologies indicate that individual resources may yield more accurate demand forecasts [33], [34]. Nevertheless, empirical studies that quantify how different levels of aggregation affect baseline accuracy—and whether aggregated EVCSs can meet industry-standard reliability thresholds (e.g., 20–30% RRMSE)—are still lacking. **To address this research gap, this study analyzes one year of hourly operational data from 873 EVCSs located in Jeju, South Korea, to evaluate baseline estimation performance across various aggregation levels.** By focusing on standardized average-based CBL methods and commonly used performance metrics such as RRMSE and RMAE, this study offers **practical insights and policy-relevant guidelines for integrating EVCSs into existing DR programs**.

This paper offers the following five major contributions:

1) **Large-scale empirical analysis:** Unlike previous studies focusing on one or two EVCSs, this research analyzes real-world data from 873 EVCSs in Jeju, Korea, collected hourly over one year. The data were randomly grouped by capacity to perform a comprehensive analysis across different aggregation levels.
2) **Threshold-based aggregation guideline:** By evaluating CBL accuracy across randomly aggregated groups, this study proposes minimum aggregation capacities required to meet specific target error thresholds, providing practical guidelines for market participation.
3) **Use of standardized average-based CBL methods:** To isolate the effect of aggregation, the study intentionally excludes complex or customized CBL methods and focuses solely on standardized average-based approaches. A comparative analysis with a regression-based method is also conducted to validate the approach.
4) **Interpretability of the RRMSE metric:** The study proposes a method to interpret RRMSE values quantitatively, enhancing the ability of system operators and practitioners to assess the variability of demand resources and apply the metric in practice.
5) **Policy implications for DR market integration:** The findings provide actionable insights for designing DR market entry strategies, particularly for small, high-variability assets such as EVCSs. They support the development of standardized CBL methods and aggregation capacity criteria necessary for market integration.

The rest of the paper is organized as follows: Section II describes the applied CBL methodology, evaluation metrics, and capacity assessment framework. Section III introduces the experimental dataset. Section IV compares the performance of individual and aggregated resources, while Section V discusses aggregation effects and policy implications. Finally, Section VI presents conclusions and future research directions.

II. METHODOLOGY AND METRICS

To utilize EVCS as DR resources in the wholesale market, a certain amount of capacity must be aggregated, a baseline must be established, and then the resource can be deployed for DR. Accordingly, this section describes the methodology used for baseline setting and the corresponding evaluation approach, and presents a method for analyzing performance variations according to different levels of aggregated capacity.

A. APPLIED CBL

To evaluate the eligibility of EVCSs for DR market participation alongside other demand-side resources, the CBL was determined using the average method, a fundamental and widely used approach. The three average-based CBL methods were selected because they are widely adopted in real-world energy markets and provide a practical and neutral benchmark. To maintain objectivity in assessing the effects of aggregation, we intentionally excluded complex or model-specific methods. Specifically, KPX's Mid 4/6, CAISO's 10-in-10, and PJM's Max 4/5(Max 2/3) were applied, as summarized in Table 2 [24], [25], [27]. With 2 day types, days are split into weekdays and weekends/holidays. With 3, they are divided into weekdays, Saturdays, and Sundays/holidays.

B. METRICS FOR CBL EVALUATION

CBL evaluation is performed by comparing the actual load with the calculated baseline load. The evaluation metrics include the RRMSE method adopted by PJM and KPX,

TABLE 2. Summary of applied average CBL methods.

Method	System Operator	Day Types	Description
Mid 4/6	KPX	2	Average of 4 days electricity usage, excluding the highest and lowest from the last 6 same type days
10-in-10	CAISO	2	Average of the last 10 same type days
Max 4/5 (Max 2/3)	PJM	3	Average of the 4 highest usage days from the last 5 same type days

as well as the Relative Mean Absolute Error(RMAE). Since MAE offers a more interpretable measure of average deviation compared to RMSE [35], the RMAE was introduced to represent the error relative to the average power consumption. The detailed formulations can be found in equations (1) and (2). Here, T denotes the total number of time intervals in the evaluation period (e.g., 24 for one day, 720 for one month, or 8,760 for one year). Since the average power consumption varies depending on the value of T, it is important to appropriately determine T when applying this metric.

$$RRMSE(L, \hat{CBL}) = \frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\hat{CBL}(t) - L(t)\right)^2}}{\frac{1}{T}\sum_{t=1}^{T} L(t)} \tag{1}$$

$$RMAE(L, \hat{CBL}) = \frac{\frac{1}{T}\sum_{t=1}^{T}\left|\hat{CBL}(t) - L(t)\right|}{\frac{1}{T}\sum_{t=1}^{T} L(t)} \tag{2}$$

C. RANDOM AGGREGATION

Although the same CBL methodology is applied, the accuracy of the calculated baseline may vary depending on the characteristics of the aggregated resources. While selectively targeting resources with high baseline accuracy can be an effective strategy, this study instead designs a random aggregation approach to examine how CBL accuracy changes with varying levels of aggregated capacity. When all resources are independently selected through repeated random trials and aggregated, the total aggregated capacity tends to converge to a normal distribution according to the central limit theorem. This convergence leads to a concentration of aggregated capacities around the expected mean, making it difficult to obtain samples with capacities that deviate significantly from the average. However, this study aims to generate random aggregations that preserve the randomness of selection while ensuring uniform distribution across different capacity levels. To achieve this, a two-stage random sampling method is proposed to ensure capacity-balanced sampling over the entire range of interest.

- **Stage 1**: A selection probability p is defined. This parameter controls the likelihood that each resource is eligible for inclusion in the aggregation, effectively determining the expected aggregated capacity in this execution.
- **Stage 2**: Each resource is assigned a random value between 0 and 1, and is selected if this value is less than the selection probability p defined in Stage 1. This step ensures that each resource is aggregated in a manner that preserves randomness.

This process is repeated multiple times to ensure uniform random sampling across different capacity levels, as illustrated in Fig. 1. In addition, the number of sampling iterations (N_s) is determined such that the number of resources included within the target capacity C_t — that is, the predefined capacity threshold — meets or exceeds the target count N_t, which is the desired number of resources within that threshold. To satisfy this condition, Equation 3 specifies that N_s must be greater than or equal to the product of the maximum installed capacity (C_m), the maximum allowable sampling probability (P), and the target count (N_t), divided by the target capacity (C_t)

$$\mathbf{N_s} \geq P \times \frac{C_m \times N_t}{C_t} \tag{3}$$

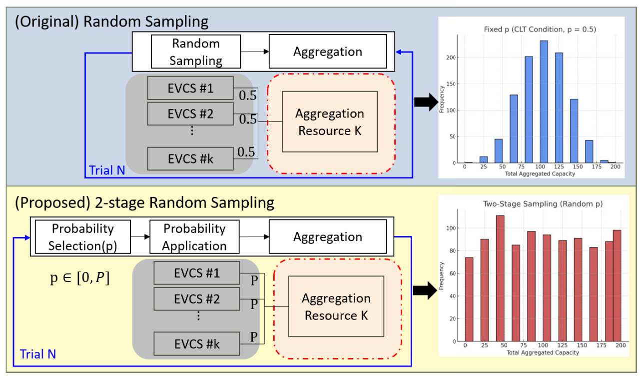

FIGURE 1. Comparison between conventional random sampling and the proposed two-stage random sampling.

III. DATASET AND TEST SYSTEM

A. EVCS DATASET

The EVCS dataset used in this study was provided by Gridwiz for the Korean project titled "Technical Development of Sector Coupling for Surplus Renewable Energy." The dataset includes charging data from 873 EVCSs located in Jeju Island, covering the entire year of 2023. The EVSEs consist of 936 slow chargers and 731 fast chargers, with a total installed capacity of 42,241 kW. Most of the slow chargers have a capacity of 7 to 8 kW, while the fast chargers are primarily 50 kW units. All remaining identifying information was anonymized to protect the privacy of charging service providers and asset owners.

B. DATA PREPROCESSING

The dataset was preprocessed to address outliers, missing values, and to align time resolution. The charging data from EVCSs, recorded in 15-minute intervals, was preprocessed by removing data points that exceed 30% of the EVCS's rated capacity per 15-minute interval (equivalent to 120% of the rated hourly capacity), as outliers [36]. Days with more than 4 hours of missing data were excluded, while those with fewer than 4 hours of missing data were linearly interpolated and forward/backward filled [36], [37]. In addition, since the electricity market operates on an hourly basis, the data was converted to hourly intervals. After filtering out EVCSs with fewer than 100 valid operating days, 830 EVCSs remained. Among these, 528 EVCSs had complete daily data for the full year.

C. TEST SYSTEM

The experimental setup was constructed using the preprocessed EVCS dataset from Jeju. Given the window-based characteristics of the average CBL method, which relies on historical data, this study compared CBL data from February to December 2023.

Across all experiments, the following **three common steps** were applied:

1) CBL Calculation: Based on the preprocessed or 2-stage randomly sampled dataset, CBL values were calculated for the period from February to December using the average CBL methods defined in Section II-A.
2) Error Metric Calculation: The calculated CBL values were compared with actual measured data to compute the RRMSE or RMAE metrics as defined in Section II-B.
3) Performance Evaluation: The resulting metrics served as the basis for evaluating performance under the following criteria.

The experimental framework consisted of two parts:

(a) Evaluation of CBL performance for individual and total resources
(b) Evaluation of CBL performance for randomly aggregated resources

In part (a), the evaluation focused on the following aspects:

1) Performance differences between the average CBL methods presented in Section II-A were assessed using the RRMSE metric from Section II-B.
2) The time window T defined in Equation (1) was varied (daily, monthly, yearly) to determine the most suitable duration for CBL estimation.
3) The RMAE metric was used to evaluate how the variability of individual resource participation differs from that of total resources.

In part (b), the evaluation followed this procedure:

1) A 2-stage random sampling method was applied to aggregate resources with uniformly distributed installed capacity and average power characteristics.
2) RRMSE and RMAE metrics were analyzed according to the aggregated installed capacity and the aggregated hourly power consumption.
3) The relationship and interaction between the two metrics (RRMSE and RMAE) were further investigated.

IV. RESULTS

A. INDIVIDUAL AND TOTAL RESOURCES ANALYSIS

1) RRMSE COMPARISON WITH CBL METHODS

CBL was applied at the individual EVCS level according to the test system, and the RRMSE was calculated. Out of 830 EVCSs, seven units with a zero mean load—resulting in infinite RRMSE values—were excluded from the analysis. A lower RRMSE indicates higher consistency and better suitability for DR participation; however, among the remaining 823 units, only three units showed RRMSE values below 40%, indicating a very limited number of highly consistent resources.

As shown in Fig. 2, a histogram was generated with 20% RRMSE intervals, and EVCSs with RRMSE values exceeding 500% were grouped together for clarity in visualization. In addition, probability density functions were plotted to compare the performance across different CBL methods. According to the analysis of both the histogram and density functions, the 10-in-10 method showed the best overall performance, with its distribution skewed furthest to the left, indicating a larger proportion of EVCSs with relatively low RRMSE values. However, for high-performing resources with RRMSE values below a certain threshold, performance differences between CBL methods were negligible.

These results suggest that the choice of average-based CBL methodology does not significantly affect the performance at the individual EVCS level. This analysis utilized the entire dataset from February to December without segmentation in the time domain, and the RRMSE results may vary if the data are segmented by day or by month.

2) TEMPORAL GRANULARITY APPROPRIATENESS

To explore the effect of temporal granularity evaluation, the RRMSE was further averaged on monthly and daily bases. When calculated daily, the squared error influence was reduced, and the number of EVCSs with RRMSE below 40% under the Mid 4/6 method increased to 13. To examine trends across different time domain settings—daily, monthly, and yearly—probability density functions were plotted for each CBL method. As shown in Fig. 3, daily RRMSE values tended to cluster near zero, but the occurrence of RRMSE exceeding 500 also increased. Meanwhile, monthly and yearly averages exhibited similar levels. Given that DR participation is typically verified on a monthly or quarterly basis in most wholesale programs, yearly RRMSE serves as a practical evaluation metric. Monthly or yearly evaluation is also more suitable for detecting large errors that occur during a few specific time intervals. These findings suggest that the temporal resolution of baseline evaluation significantly influences the reliability assessment of EVCSs, highlighting the need for tailored granularity depending on the application context.

3) RMAE COMPARISON BETWEEN INDIVIDUAL AND AGGREGATED RESOURCES

The deviation was calculated in two ways: treating each EVCS as an individual resource and aggregating

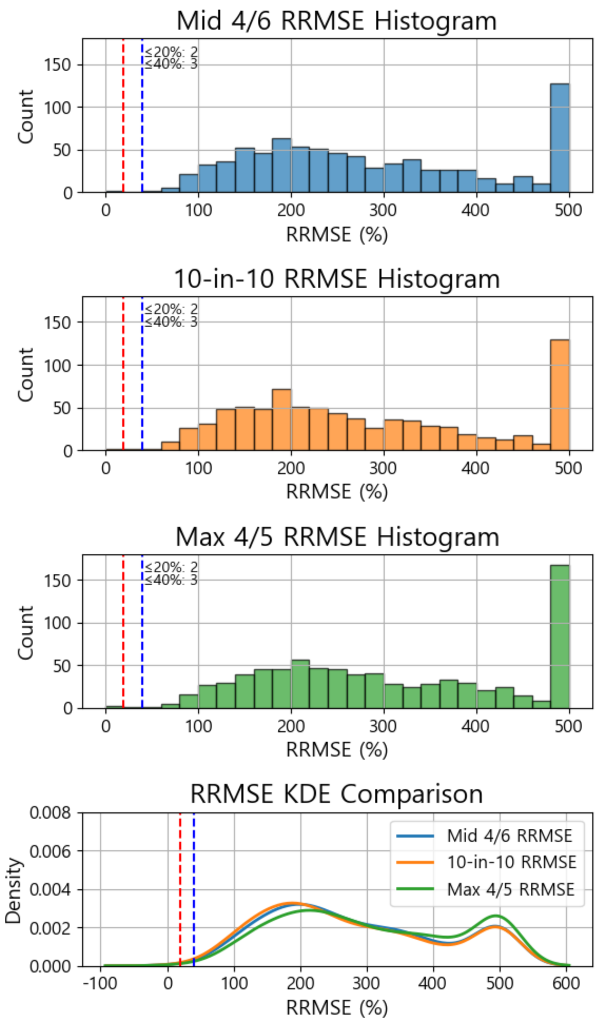

FIGURE 2. **RRMSE comparison across CBL methods.**

all 830 EVCSs into a single virtual resource. The entire set of 830 EVCSs consumed an average of 2,133 kWh per hour over the 11-month period. Based on this consumption, the RMAE was approximately 100% when each EVCS was treated individually, indicating that CBL estimation errors averaged above 2,000 kWh. However, when the EVCSs were treated as a single resource, the RMAE dropped to below 10%, reducing the error magnitude to one-tenth. As shown in Fig. 4, individual-level estimations exhibited high sensitivity to the choice of CBL method, whereas in the aggregated case, method-specific differences were significantly diminished. Based on this result, we conducted additional experiments to quantitatively analyze how the performance of CBL methods varies depending on the level of aggregation.

B. RANDOMLY AGGREGATED RESOURCES ANALYSIS

1) SAMPLING GENERATION

The two-stage random sampling method was applied to evenly aggregate EVCS resources based on capacity. A total of 2,500 random sampling iterations (N_s) were conducted using Equation 3, ensuring that at least 10 units (N_t) were included within each 100kW segment (C_t). The total available capacity (C_m) was preprocessed and set to 41,885kW. In the "probability selection" stage, the maximum selection probability was limited to 0.5, favoring the inclusion of lower-capacity resources.

Fig. 5 demonstrated that the randomly aggregated resources were evenly distributed with respect to both installed capacity and average power consumption. To statistically validate this uniformity, a chi-squared goodness-of-fit test was conducted. To avoid violating the expected frequency assumptions under the 0.5 maximum selection probability condition, the tail-end histogram bins were merged. Specifically, capacity segments above 20 MW and average power segments above 1 MW were grouped into a single bin before testing. As a result, the chi-squared statistics were calculated to be 27.98 for the installed capacity distribution and 13.36 for the average power consumption. The corresponding p-values were 0.110 and 0.204, respectively—both greater than the significance level of 0.05—indicating that the distributions can be considered statistically uniform. These randomly aggregated groups were subsequently used to estimate CBL and compare error metrics such as RRMSE and RMAE.

2) RRMSE/RMAE BY AGGREGATION SIZE

The aggregated resources were grouped by installed capacity in 1MW increments and by average hourly consumption in 100kWh increments, and their performance was analyzed. Fig.6 and Fig.7 presented the evaluation results using RRMSE and RMAE as performance metrics, respectively. In both figures, the first and second subplots averaged the error values within each group to evaluate performance differences across CBL methods. The third and fourth subplots focused on the Mid4/6 method, comparing the mean, maximum, and minimum values within each group. All subplots included error bars based on t score–derived 95% confidence intervals, which indicated the uncertainty of each group's estimate.

As shown in Fig.6, both capacity-based and consumption-based groupings exhibited a clear trend: larger resource groups tended to have lower RRMSE values, regardless of the CBL method. The performance gap among CBL methods was more prominent in smaller-capacity groups, but narrowed as the aggregation size increased. A sharp reduction in average RRMSE was observed between the 1–2MW and 100–200kWh ranges. In addition, the range between the maximum and minimum values, as well as the confidence intervals, became narrower for larger groups, suggesting reduced variability.

When these results were applied to qualification thresholds, it was appropriate to register aggregated resources with at least 4MW of installed capacity or 200kWh of average consumption, based on the 30% RRMSE criterion suggested by KPX. Under the stricter 20% threshold proposed by PJM,

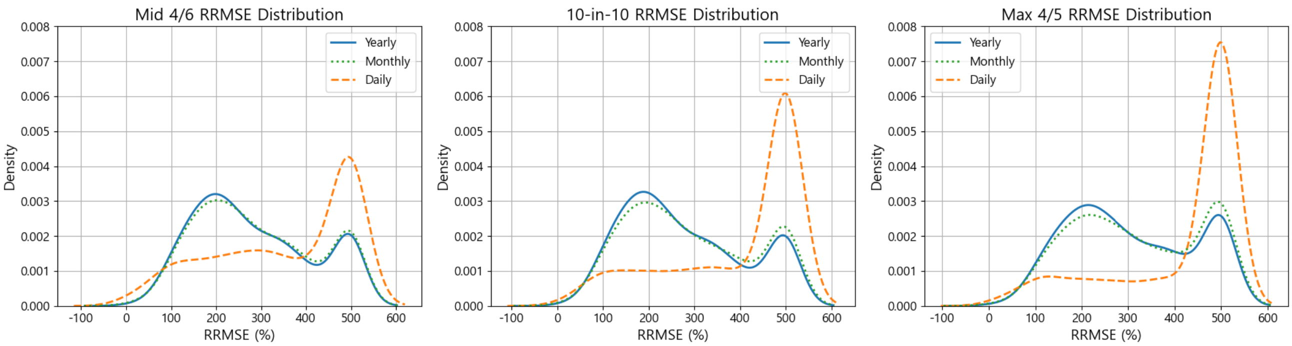

FIGURE 3. RRMSE distribution by temporal granularity.

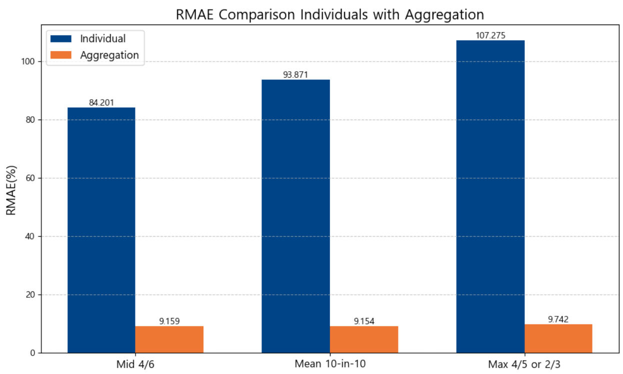

FIGURE 4. RMAE comparison individuals with aggregation across CBL methods.

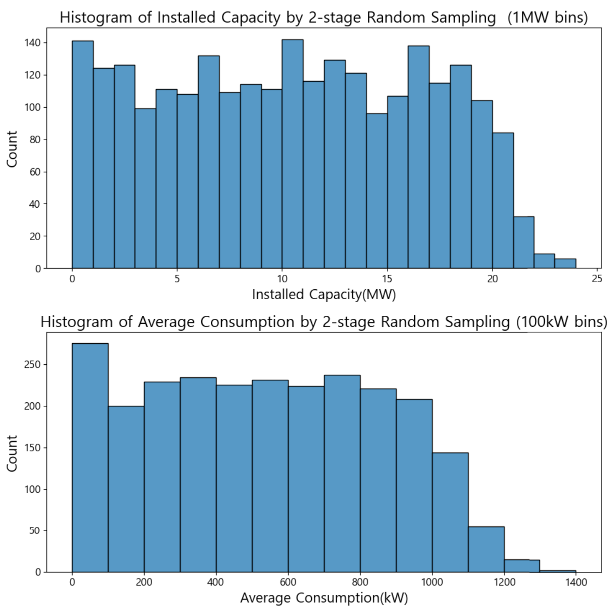

FIGURE 5. Random sampling capacity(average power consumption) histogram.

the aggregation thresholds should be increased to 10MW and 500kWh, respectively. If a more robust approach is required, the maximum RRMSE within each group can be used as a benchmark. Conversely, the presence of low RRMSE values in small-capacity groups indicated that even small-scale aggregations can include reliable resources.

Fig.7 demonstrated a similar trend using RMAE as the error metric, though with less variability compared to RRMSE. This finding suggested that RRMSE was more sensitive in capturing variation of resources.

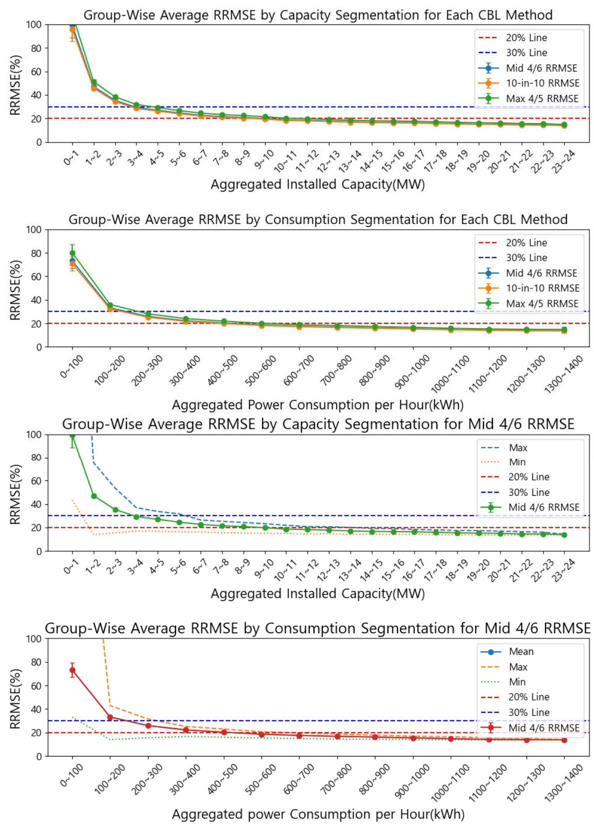

FIGURE 6. Group-wise average RRMSE by installed capacity and average hourly consumption.

3) CORRELATION BETWEEN RRMSE AND RMAE

By comparing Fig.6 and Fig.7, it was confirmed that RRMSE was more sensitive in capturing variability among resources.

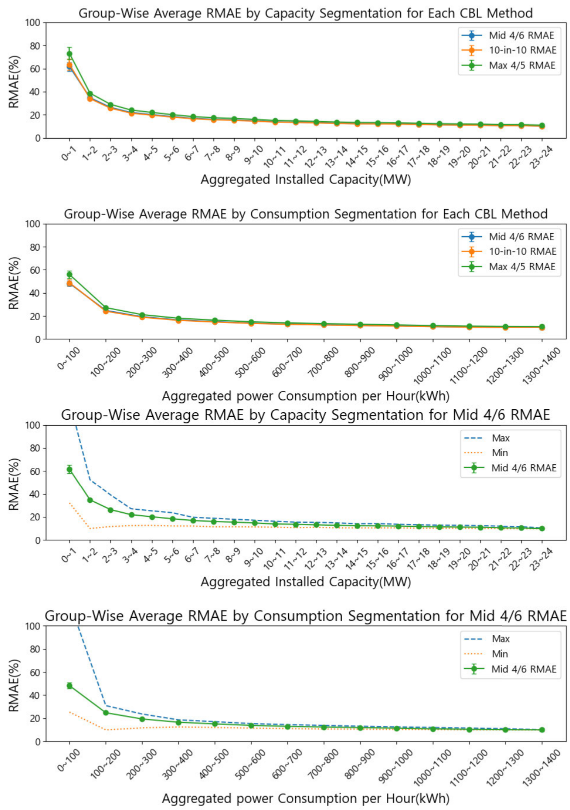

FIGURE 7. **Group-wise Average RMAE by installed capacity and average hourly consumption.**

However, the RRMSE value itself is not easily interpretable in intuitive terms, whereas RMAE directly reflects the deviation relative to the average power consumption. Therefore, a comparative analysis between RRMSE and RMAE was conducted to better understand the implication of the RRMSE scale.

For this purpose, scatter plots were generated for 2,500 sampled resources, where the x-axis represented RRMSE and the y-axis represented RMAE for each CBL method, as shown in Fig.8. Regardless of the CBL method, the plots revealed a clear linear relationship between RRMSE and RMAE in the lower range of RRMSE values (below 60%). A linear regression was applied to this subset of data, resulting in a slope of approximately 0.75. This finding suggests that an RRMSE of 20% corresponds to a deviation of about 0.15kWh per 1 kWh of actual consumption.

V. DISCUSSION

A. PARTICIPATION OF EVs IN THE WHOLESALE ELECTRICITY MARKET

When participating in the wholesale electricity market at the individual EV level, the resource location is typically unfixed, and individual charging behavior tends to be temporally sparse and irregular, making it difficult to ensure reliability. However, by aggregating and registering EVCSs instead of individual EVs, it becomes possible to associate each resource with a fixed location and more consistent charging patterns. As demonstrated in this study, when EVCSs are aggregated to a rated capacity of 4MW (or an average hourly consumption of 200kWh), the resulting aggregate pattern shows temporal consistency comparable to that of conventional DR resources. This improved reliability indicates that aggregated EVCSs can be considered viable resources for DR participation, including upward/downward demand adjustment and contingency reserve scenarios. In particular, Europe is currently making efforts to expand and develop local flexibility markets, and in this context, establishing reliability thresholds for EVCS resources can provide valuable insights for future market design [38].

B. INTERPRETATION OF RRMSE THROUGH COMPARISON WITH RMAE

While the RRMSE is more effective in detecting inconsistent or irregular load behavior, it tends to be more difficult to interpret directly. To address this, the study compared RRMSE with RMAE, providing a clearer quantitative interpretation of deviation magnitudes. Specifically, within the range where RRMSE is less than 60%, a linear relationship was identified, allowing system operators to approximate expected deviations as a function of unit consumption. For example, a 20% RRMSE approximately corresponds to a deviation of 0.15 kWh per 1 kWh of hourly consumption. This enables a more quantitative assessment of variability and supports better planning of operating reserves for demand-side resources. The practical applications of the RRMSE metric can vary depending on the type of market participant, as outlined below:

- **System Operators**: For demand resources with an RRMSE value below 60%, the inherent variability can be quantitatively estimated. This enables system operators to assess the potential impact of resource fluctuations on grid stability and refine reliability DR planning accordingly.
- **Utilities**: The RRMSE metric can be applied not only to demand resources but also to load forecasting. By quantifying the deviation between actual load and day-ahead forecasts, utilities can better formulate demand bidding strategies and secure adequate demand resources in advance for real-time operations.
- **Demand Resource Aggregators (or Virtual Power Plant Operators)**: When participating in electricity markets, aggregators can use the RRMSE metric to estimate the variability of their aggregated demand resources. This supports the development of robust bidding strategies and helps minimize imbalance penalties.

However, it is important to note that this analysis is limited to EVCS resources, and further validation across diverse types of resources is required for broader generalization.

C. IMPLICATIONS FOR CBL METHOD DEVELOPMENT

The performance differences among CBL methods were more pronounced in small aggregation groups, but this gap tended to diminish as the aggregation size increased.

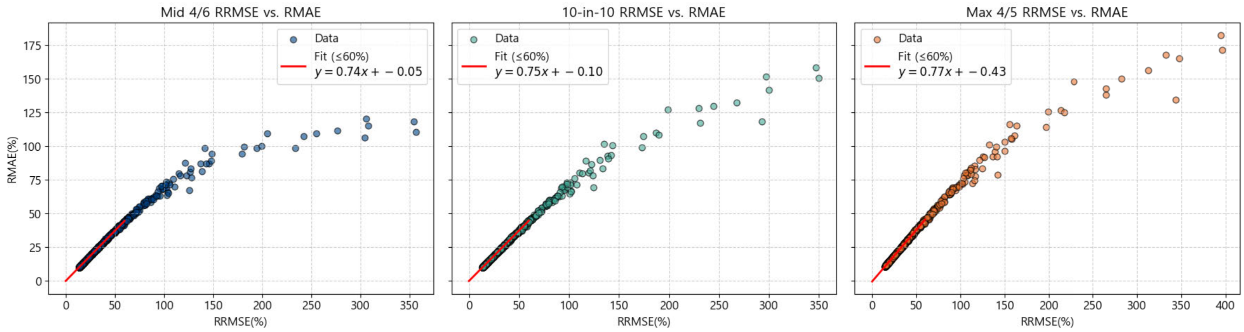

FIGURE 8. RRMSE vs RMAE by 2500 sampling.

This suggests that the sensitivity to CBL method selection decreases with larger aggregation sizes. To validate this, we compared our results with a prior study that employed an autoregressive (AR) model [34]. In that study, when EV resources were aggregated to daily consumption levels of 309kWh (approximately 12.9kWh/hour) and 806kWh (approximately 33.6kWh/hour), the reported RRMSE values were 70% and 46%, respectively. In contrast, our results for average power consumption in the 0–100kWh/hour range showed higher variability, with RRMSE values of 73%, 71%, and 80% for the Mid4/6, Mean10, and Max4/5 methods, respectively. However, when the aggregation size increased to 8,319kWh per day (approximately 346.6kWh/hour), the AR-based study achieved an RRMSE of 21%, which is highly comparable to the results in our study (22%, 22%, and 24% for the same three methods). These findings suggest that, beyond a certain aggregation threshold, standardized and intuitive average-based methods—such as Mid4/6—can deliver performance comparable to more complex regression-based models. While customized CBL approaches may be necessary for enabling individual small-scale EVCSs to participate in the market, average-based methods become sufficiently robust and reliable for settlement and forecasting once resources are aggregated at scale.

Aggregation is essential to enable small-scale and highly variable resources, such as EVs, to meaningfully participate in the wholesale electricity market. While large-scale demand resources—such as water electrolysis systems, electrode boilers, and industrial facilities—may require custom CBL methods to capture their unique operational characteristics [39], [40], [41], this study confirms that aggregated EVCSs can be effectively managed using standardized average-based CBL methods. These findings suggest that while aggregation enables reliable settlement and forecasting for small-scale resources, developing tailored CBL methodologies remains critical for large-scale and operationally heterogeneous demand resources.

VI. CONCLUSION AND FUTURE WORK

In this study, we evaluated the conditions required for EVCSs to reliably participate in existing wholesale DR markets. To this end, we utilized preprocessed load data from 873 EVCSs in the Jeju region as of 2023. Rather than selecting only high-performing resources, we randomly aggregated the EVCSs by capacity groups and compared the average CBL with actual load profiles across each group. The results indicate that, to ensure a level of reliability comparable to conventional DR resources, EVCSs should be aggregated with at least 10MW (4MW) of installed capacity or 500kWh (200kWh) of average hourly power consumption, based on a 20% (30%) RRMSE threshold. These findings suggest that resource consistency and reliability can be effectively ensured by setting minimum registration capacity requirements, without the need for customized CBL development for EVCSs. Additionally, by jointly analyzing both RRMSE and RMAE metrics, we improved the interpretability of the RRMSE metric and proposed practical ways for system operators and market participants to quantitatively assess the variability of demand resources. This approach offers actionable insights for improving the design and operation of DR programs involving aggregated EVCSs. However, this study is based on data from the Jeju region, which may reflect region-specific characteristics, and thus its findings have limitations in terms of generalizability to other regions. Nevertheless, similar results have been reported in a prior study using data from California [34], suggesting that aggregating EVCSs can also be a viable approach to ensuring reliability in different environments.

To advance the current study, we propose two key directions for future research:

1. Enhancing CBL Reliability with Advanced Methods: While this study demonstrated that average-based CBL methods can provide sufficient reliability, future work will explore whether machine learning (ML) or deep learning (DL) approaches can further improve reliability, particularly when applied to large-scale EVCS aggregations in a time-efficient manner. We aim to develop a scalable validation framework that evaluates the performance of these data-driven models under real-world constraints. Additionally, we plan to examine how key hyperparameters in ML/DL-based CBL models can be standardized or shared across different stakeholders—such as system operators, utilities, and aggregators—to ensure settlement transparency and operational consistency.

2. Role-Specific Utilization of Aggregated EV Resources: We will investigate the most appropriate operational roles for aggregated EVCSs that meet baseline reliability criteria. This includes evaluating their suitability for demand reduction, demand increase, reserve provision, and frequency response services. In particular:

- First, we will assess the technical feasibility of enabling EVCS-based DR resources to participate in ancillary services markets.
- Second, we will analyze the potential market impacts of such participation across different seasons and time periods, taking into account system load dynamics.
- Lastly, we will explore control schemes tailored to high-value service types to ensure effective and responsive integration of EVCSs into grid operations.

These future directions aim to address both methodological limitations and practical implementation pathways, enhancing the generalizability and impact of this research.

REFERENCES

[1] V. Masson-Delmotte, P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J. Matthews, Y. Chen, X. Zhou, M. Gomis, E. Lonnoy, T. Maycock, M. Tignor, and T. Waterfield, *Global Warming of 1.5C: An IPCC Special Report on the Impacts of Global Warming of 1.5C Above Pre-industrial Levels and Related Global Greenhouse Gas Emission Pathways, in the Context of Strengthening the Global Response to the Threat of Climate Change, Sustainable Development, and Efforts to Eradicate Poverty*. Cambridge, U.K.: Intergovernmental Panel on Climate Change (IPCC), 2018, doi: 10.1017/9781009157940.

[2] *Net Zero Roadmap: A Global Pathway to Keep the 1.5 °C Goal in Reach*, Licence: CC BY 4.0, International Energy Agency (IEA), Paris, France, 2023. [Online]. Available: https://www.iea.org/reports/net-zero-roadmap-a-global-pathway-to-keep-the-15-0c-goal-in-reach

[3] *World Energy Transitions Outlook 2023: 1.5°C Pathway, Volume 1*, International Renewable Energy Agency (IRENA), Abu Dhabi, United Arab Emirates, 2023.

[4] (2025). *Global EV Outlook 2025*, International Energy Agency (IEA), Paris, France. Accessed: Jul. 2025. [Online]. Available: https://www.iea.org/reports/global-ev-outlook-2025

[5] A. Dubey and S. Santoso, "Electric vehicle charging on residential distribution systems: Impacts and mitigations," *IEEE Access*, vol. 3, pp. 1871–1893, 2015.

[6] S. Powell, G. V. Cezar, L. Min, I. M. L. Azevedo, and R. Rajagopal, "Charging infrastructure access and operation to reduce the grid impacts of deep electric vehicle adoption," *Nature Energy*, vol. 7, no. 10, pp. 932–945, Sep. 2022.

[7] A. A. Munshi and Y. A.-R.-I. Mohamed, "Extracting and defining flexibility of residential electrical vehicle charging loads," *IEEE Trans. Ind. Informat.*, vol. 14, no. 2, pp. 448–461, Feb. 2018.

[8] H. Li, Z. Wan, and H. He, "Constrained EV charging scheduling based on safe deep reinforcement learning," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2427–2439, May 2020.

[9] O. Erdinc, N. G. Paterakis, T. D. P. Mendes, A. G. Bakirtzis, and J. P. S. Catalão, "Smart household operation considering bi-directional EV and ESS utilization by real-time pricing-based DR," *IEEE Trans. Smart Grid*, vol. 6, no. 3, pp. 1281–1291, May 2015.

[10] X. Xu, Y. Jia, Y. Xu, Z. Xu, S. Chai, and C. S. Lai, "A multi-agent reinforcement learning-based data-driven method for home energy management," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 3201–3211, Jul. 2020.

[11] G. Zhang, S. T. Tan, and G. G. Wang, "Real-time smart charging of electric vehicles for demand charge reduction at non-residential sites," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4027–4037, Sep. 2018.

[12] I. Sengör, O. Erdinç, B. Yener, A. Tascikaraoglu, and J. P. S. Catalão, "Optimal energy management of EV parking lots under peak load reduction based DR programs considering uncertainty," *IEEE Trans. Sustain. Energy*, vol. 10, no. 3, pp. 1034–1043, Jul. 2019.

[13] C. Wei, J. Xu, S. Liao, and Y. Sun, "Aggregation and scheduling models for electric vehicles in distribution networks considering power fluctuations and load rebound," *IEEE Trans. Sustain. Energy*, vol. 11, no. 4, pp. 2755–2764, Oct. 2020.

[14] N. I. Nimalsiri, E. L. Ratnam, D. B. Smith, C. P. Mediwaththe, and S. K. Halgamuge, "Coordinated charge and discharge scheduling of electric vehicles for load curve shaping," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 7653–7665, Jul. 2022.

[15] X. Shi, Y. Xu, Q. Guo, H. Sun, and X. Zhang, "Day-ahead distributionally robust optimization-based scheduling for distribution systems with electric vehicles," *IEEE Trans. Smart Grid*, vol. 14, no. 4, pp. 2837–2850, Jul. 2023.

[16] F. Rassaei, W.-S. Soh, and K.-C. Chua, "Demand response for residential electric vehicles with random usage patterns in smart grids," *IEEE Trans. Sustain. Energy*, vol. 6, no. 4, pp. 1367–1376, Oct. 2015.

[17] J. Soares, H. Morais, T. Sousa, Z. Vale, and P. Faria, "Day-ahead resource scheduling including demand response for electric vehicles," *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 596–605, Mar. 2013.

[18] Z. Liu, Q. Wu, S. Huang, L. Wang, M. Shahidehpour, and Y. Xue, "Optimal day-ahead charging scheduling of electric vehicles through an aggregative game model," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 5173–5184, Sep. 2018.

[19] Z. Liu, D. Wang, H. Jia, N. Djilali, and W. Zhang, "Aggregation and bidirectional charging power control of plug-in hybrid electric vehicles: Generation system adequacy analysis," *IEEE Trans. Sustain. Energy*, vol. 6, no. 2, pp. 325–335, Apr. 2015.

[20] D. F. R. Melo, A. Trippe, H. B. Gooi, and T. Massier, "Robust electric vehicle aggregation for ancillary service provision considering battery aging," *IEEE Trans. Smart Grid*, vol. 9, no. 3, pp. 1728–1738, May 2018.

[21] S. Izadkhast, P. Garcia-Gonzalez, P. Frias, L. Ramirez-Elizondo, and P. Bauer, "An aggregate model of plug-in electric vehicles including distribution network characteristics for primary frequency control," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2987–2998, Jul. 2016.

[22] M. Nikzad and B. Mozafari, "Reliability assessment of incentive- and priced-based demand response programs in restructured power systems," *Int. J. Electr. Power Energy Syst.*, vol. 56, pp. 83–96, Mar. 2014.

[23] I. Pavic, M. Beus, H. Pandžic, T. Capuder, and I. Štritof, "Electricity markets overview—Market participation possibilities for renewable and distributed energy resources," in *Proc. 14th Int. Conf. Eur. Energy Market (EEM)*, Jun. 2017, pp. 1–5.

[24] California Independent System Operator (CAISO). (Jun. 22, 2020). *Energy Storage and Distributed Energy Resources (ESDER) Phase 3B*. Training Session: Market Simulation Readiness. [Online]. Available: http://www.caiso.com/Documents/BusinessRequirementsSpecification Clean-EnergyStorageandDistributedEnergyResourcesPhase3.pdf

[25] L. PJM Interconnection. (Dec. 17, 2024). *PJM Manual 11: Energy & Ancillary Services Market Operations*, revision 133 ed., PJM Interconnection. [Online]. Available: https://www.pjm.com/-/media/documents/manuals/m11.ashx

[26] M. I.S. O. (MISO), (Oct. 1, 2024). *Demand Response Bus. Practices Manual (BPM-026-r11)*, Midcontinent Independent System Operator. [Online]. Available: https://www.misoenergy.org

[27] Korea Power Exchange (KPX). (2024). *Electricity Market Operation Rules*. [Online]. Available: https://www.kpx.or.kr

[28] S. Mohajeryami, M. Doostan, A. Asadinejad, and P. Schwarz, "Error analysis of customer baseline load (CBL) calculation methods for residential customers," *IEEE Trans. Ind. Appl.*, vol. 53, no. 1, pp. 5–14, Jan. 2017.

[29] L. Zhang, G. Li, Y. Huang, J. Jiang, Z. Bie, X. Li, Y. Ling, and H. Tian, "Distributed baseline load estimation for load aggregators based on joint FCM clustering," *IEEE Trans. Ind. Appl.*, vol. 59, no. 1, pp. 567–577, Jan. 2023.

[30] F. Wang, K. Li, C. Liu, Z. Mi, M. Shafie-Khah, and J. P. S. Catalão, "Synchronous pattern matching principle-based residential demand response baseline estimation: Mechanism analysis and approach description," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 6972–6985, Nov. 2018.

[31] M. I.S. O. (MISO). (Jul. 11, 2025). *Module C-Energy and Operating Reserve Markets*. fERC Electric Tariff. [Online]. Available: https://www.misoenergy.org/legal/rules-manuals-and-agreements/tariff/

[32] T. Unterluggauer, K. Rauma, P. Järventausta, and C. Rehtanz, "Short-term load forecasting at electric vehicle charging sites using a multivariate multi-step long short-term memory: A case study from Finland," *IET Electr. Syst. Transp.*, vol. 11, no. 4, pp. 405–419, Dec. 2021.

[33] Y. Lu, Y. Li, D. Xie, E. Wei, X. Bao, H. Chen, and X. Zhong, "The application of improved random forest algorithm on the prediction of electric vehicle charging load," *Energies*, vol. 11, no. 11, p. 3207, Nov. 2018.

[34] M. Pertl, F. Carducci, M. Tabone, M. Marinelli, S. Kiliccote, and E. C. Kara, "An equivalent time-variant storage model to harness EV flexibility: Forecast and aggregation," *IEEE Trans. Ind. Informat.*, vol. 15, no. 4, pp. 1899–1910, Apr. 2019.

[35] C. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Climate Res.*, vol. 30, pp. 79–82, Dec. 2005.

[36] K. Baek, S. Kim, E. Lee, Y. Cho, and J. Kim, "Data-driven evaluation for demand flexibility of segmented electric vehicle chargers in the Korean residential sector," *Energies*, vol. 14, no. 4, p. 866, Feb. 2021.

[37] B. Lee, H. Lee, and H. Ahn, "Improving load forecasting of electric vehicle charging stations through missing data imputation," *Energies*, vol. 13, no. 18, p. 4893, Sep. 2020.

[38] S. Chondrogiannis, J. Vasiljevska, A. Marinopoulos, I. Papaioannou, and G. Flego, "Local electricity flexibility markets in Europe," Publications Office Eur. Union, Joint Res. Centre (JRC), Eur. Commission, Brussels, Belgium, Tech. Rep. EUR 31194 EN, jRC130070, 2022.

[39] O. Valentini, N. Andreadou, P. Bertoldi, A. Lucas, I. Saviuc, and E. Kotsakis, "Demand response impact evaluation: A review of methods for estimating the customer baseline load," *Energies*, vol. 15, no. 14, p. 5259, Jul. 2022.

[40] A. Ghasemi, M. Hojjat, J. Saebi, H. R. Neisaz, and M. R. Hosseinzade, "An investigation of the customer baseline load (Cbl) calculation for industrial demand response participants—A regional case study from Iran," *Appl. Energy*, vol. 277, Jan. 2020, Art. no. 115945.

[41] J. Cui, M. Gao, X. Zhou, Y. Li, W. Liu, J. Tian, and X. Zhang, "Demand response method considering multiple types of flexible loads in industrial parks," *Appl. Energy*, vol. 335, Jun. 2023, Art. no. 120753.

YONGWOO JEE (Graduate Student Member, IEEE) received the B.S. degree in physics from Gwangju Institute of Science and Technology, Gwangju, South Korea, in 2022, where he is currently pursuing the integrated M.S./Ph.D. degree in electrical engineering and computer science. His research interests include sector coupling, especially upward demand flexibility.

JINHO KIM (Member, IEEE) received the Ph.D. degree in electrical engineering from Seoul National University, Seoul, South Korea, in 2001, and the M.B.A. degree from the University of Illinois, in 2012. He is currently a Professor with Gwangju Institute of Science and Technology, Gwangju, South Korea. His current research interests include power system economics, energy big data analytics, optimal energy management system in power systems and electricity markets, energy policy and implementations, demand response, electric vehicle grid integration, virtual power plants, smart/micro grid, and management of innovation.