DOI: 10.22146/jnteti.v15i1.24882

Parameter Optimization of Battery Energy Storage System Considering Degradation Using Reinforcement Learning

Muhamad Dzaky Ashidqi[1], Silviana Windasari[1], Rahmat[2], Probokusumo[3]

[1] Department of Electrical Engineering, Faculty of Engineering, Universitas Sains Indonesia, Bekasi, Jawa Barat 17530, Indonesia
[2] Department of Mechanical Engineering, Faculty of Engineering, Universitas Sains Indonesia, Bekasi, Jawa Barat 17530, Indonesia
[3] Department of Industrial Engineering, Faculty of Engineering, Universitas Sains Indonesia, Bekasi, Jawa Barat 17530, Indonesia

[Received: 13 October 2025, Revised: 15 January 2026, Accepted: 2 February 2026]
Corresponding Author: Muhamad Dzaky Ashidqi (email: muhamad.dzaky@lecturer.sains.ac.id)

ABSTRACT — Accurate and sustainable operation of battery energy storage systems (BESS) is critical for supporting renewable energy integration, ensuring both short-term reliability and long-term asset preservation. This study proposed a reinforcement learning (RL)-based scheduling framework designed to minimize power mismatch while mitigating degradation in lithium-ion batteries. The framework dynamically adapted to fluctuations in photovoltaic generation and residential load, enabling real-time decision-making. The performance was evaluated over a 30-day horizon using three indicators: average power mismatch, cumulative capacity loss, and system stability index (SSI). Results demonstrated that the proposed method achieved near-perfect load balance with an average mismatch of only 0.002 kW, while cumulative degradation remained limited to 0.22% and SSI was maintained at 0.96, reflecting high operational stability. The estimated daily degradation rate of 0.0073% corresponded to an annual capacity loss of approximately 2.7%, significantly lower than the 5–6% typically observed in uncontrolled cycling scenarios. Comparative analysis with simulated annealing (SA) and multi-objective genetic algorithm (MOGA) highlighted the balanced performance of the RL method. While MOGA eliminated mismatch at the expense of excessive degradation (0.60%) and simulated annealing reduced degradation but suffers from high mismatch (0.012 kW), the RL framework delivered the most balanced trade-off across all metrics. These findings confirm the potential of RL as a practical and sustainable strategy for PV–BESS integration, providing both technical resilience and extended battery lifetime.

KEYWORDS — Battery Degradation, Battery Energy Storage System (BESS), Parameter Optimization, Reinforcement Learning.

I. INTRODUCTION

The integration of renewable energy sources, particularly photovoltaic (PV) power plants, has gained increasing attention as part of the global transition toward cleaner and low-carbon electricity systems [1]. However, despite the rapid growth and technological progress in PV deployment, significant challenges remain in ensuring the economic feasibility and operational reliability of these systems. One of the most critical issues involves the high capital and maintenance costs associated with large-scale PV installations. A considerable portion of these expenses arises from the implementation of battery energy storage systems (BESS), which are indispensable for addressing the inherent intermittency and variability of solar generation [2]. Beyond serving as supplementary components, batteries function as integral elements that enable PV systems to deliver stable, predictable, and high-quality power to end users [3]. By providing essential capabilities such as energy balancing, frequency regulation, and backup support, BESS play a decisive role in enhancing the overall resilience and sustainability of solar-based electricity networks, thereby reinforcing their long-term viability in modern energy systems [4], [5].

The growing dependence on BESS is primarily driven by the intermittent and variable nature of solar power generation. As discussed in [6], solar energy, like other renewable resources, exhibits inherent fluctuations that prevent it from providing a constant and predictable power output. PV systems typically generate surplus energy during periods of high solar irradiation; however, their output can decrease significantly under cloudy conditions and completely cease during nighttime hours [7]. This temporal inconsistency poses substantial challenges to maintaining grid stability and ensuring a reliable energy supply, particularly in systems with high penetration levels of solar generation.

In response to these operational challenges, the deployment of BESS has become a crucial strategy for supporting renewable-based electricity networks. As indicated in [8], batteries perform a dual operational role by storing surplus energy during periods of excess generation and discharging it when production falls short or when demand peaks. This balancing mechanism not only mitigates short-term fluctuations in power output but also enhances the overall stability, reliability, and efficiency of the grid. Consequently, BESS serves as a key enabler for the large-scale integration of solar energy, ensuring smoother power delivery and improved system resilience under variable generation conditions.

Despite the numerous technical and environmental advantages of PV–battery integration, this configuration also presents considerable economic challenges. The inclusion of energy storage significantly increases both the capital expenditure and the ongoing maintenance costs of PV systems, making them less competitive than conventional generation units that operate without storage support [9]. The initial investment required for battery procurement, installation, and system integration is often substantial, and the associated maintenance costs can further burden project economics over time. As a result, the financial feasibility of large-scale PV deployment increasingly depends on the ability to manage and reduce the lifecycle costs associated with battery storage [10].

A key factor influencing these costs is the rate of battery degradation, which directly determines the frequency of replacement and the overall maintenance expenditure. Battery degradation is a multifaceted issue influenced by various factors, including usage patterns, environmental conditions, and intrinsic material properties. During storage, batteries degrade even when not in use due to self-discharge mechanisms, which are accelerated at elevated temperatures [11]. In addition, higher storage temperatures and higher state of charge (SoC) levels lead to faster degradation, increasing internal resistance and reducing capacity over time [12]. During operation cycles, frequent cycling at high charge/discharge current rates (C-rates) and high depth of discharge (DoD) level accelerates degradation. Higher C-rates and DoD lead to increased stress on the battery materials, causing faster capacity fade [13].

As highlighted in [14], extending the operational life of batteries through degradation management is therefore essential to achieving economically viable PV-based electricity production. Since degradation naturally progresses with repeated charge–discharge cycles, excessive cycling or suboptimal operational conditions can accelerate capacity loss and performance decline [15]. This accelerated wear not only shortens the service life of the storage system but also leads to higher operational and replacement costs, undermining the long-term sustainability of PV–battery investments. Consequently, optimizing battery usage to minimize degradation has become a central focus in both research and practical implementation of renewable energy systems.

Battery degradation is closely governed by its operational conditions, particularly the parameters that define charging and discharging behavior. As noted in [16], variables such as charging rate, discharging rate, and charge–discharge current play a decisive role in determining the overall lifespan and performance of the storage system. Operating at excessively high values for these parameters tends to accelerate chemical and structural wear within the battery, leading to faster capacity fading and efficiency losses. Conversely, adopting overly conservative operating settings may preserve the battery's condition but often results in underutilization of the available capacity, limiting its capability to supply energy when needed and reducing the overall effectiveness of the system [17].

Given these trade-offs, the optimization of operational parameters becomes an essential strategy for balancing performance and durability. By identifying optimal operating ranges, it is possible to mitigate degradation while ensuring that the battery continues to provide sufficient flexibility to support load variations and stabilize the grid [18]. Achieving this balance is fundamental to maintaining the long-term reliability, sustainability, and cost-effectiveness of renewable energy systems. Thus, developing intelligent control and optimization frameworks represents a critical step toward realizing stable and economically viable renewable-based power networks.

Several studies have examined optimization approaches for the integration of BESS into PV-based power systems, with the objective of improving stability and reducing operational costs. Previous study proposed an optimization framework for determining the optimal capacity of BESS in PV–diesel microgrids by employing a solar variability index to smooth out fluctuations in power output [19]. This work highlighted the critical role of variability indices in capturing solar intermittency and demonstrated that appropriately sized storage can significantly mitigate power fluctuations in hybrid systems. Similarly, another study investigated the optimal sizing of residential PV–battery systems in off-grid configurations. By using linear simulation models, this study emphasized the importance of appropriate system sizing to meet load demand reliably while minimizing installation costs [20]. These contributions confirm that accurate sizing and capacity allocation are central to enhancing the reliability and economic feasibility of PV–BESS systems.

Beyond system sizing, recent research has extended into advanced optimization methods to maximize the operational and economic benefits of PV systems. Recent study employed metaheuristic algorithms, specifically the Grey-Wolf optimization (GWO), to improve parameter extraction of solar PV modules from manufacturer datasheets [21]. The findings of this research demonstrated that GWO provided a high level of accuracy in parameter estimation, thereby improving system modeling and forecasting performance. Building upon the integration perspective, a simultaneous sizing and scheduling optimization for farmhouse PV–battery systems was proposed using a multi-structured power system model [22]. By leveraging simulated annealing, this study achieved significant improvements in cost efficiency and energy loss reduction, showing that optimization at both design and operational levels can enhance overall system performance.

Another study introduced a multi-objective optimization approach based on Pareto frontier analysis for hybrid PV–diesel–battery systems, considering different battery types [23]. The framework simultaneously accounted for technical, economic, and environmental objectives, demonstrating that multi-objective genetic algorithms (MOGA) could identify optimal trade-offs between cost reduction, environmental impact mitigation, and system reliability. However, despite these promising results, most existing studies rely on simulated data and traditional optimization algorithms, which may not fully capture the dynamic and stochastic behavior of real-world PV systems. Furthermore, approaches such as linear programming (LP), GWO, simulated annealing, and MOGA often involve high computational costs and limited adaptability under rapidly changing operating conditions [24]. These limitations open an opportunity for more adaptive and data-driven methods, such as reinforcement learning, which can better handle dynamic environments and real operational data while reducing computational burden.

While these studies provide valuable insights, most rely on linear or simulated data and traditional optimization algorithms, which may not fully capture the dynamic behavior of real-world PV systems. The novelty of the present study lies in two aspects. First, unlike prior works, this research incorporates real and dynamic load scenarios obtained from experimental design of a residential-scale PV system. This approach enables more accurate optimization results that better reflect actual system characteristics and conditions in Indonesia. Second, this study formulates the BESS operation problem within a reinforcement learning (RL) framework that explicitly considers power mismatch minimization under battery operational and degradation-related constraints. RL offers the advantage of adaptive and efficient search for optimal operational parameters, especially under dynamic and uncertain environments, making it highly suitable for enhancing both accuracy and speed in optimization tasks [25].

By addressing these gaps, this study contributes not only to reducing operational costs through optimized BESS operation but also to improving the reliability and scalability of PV systems for real-world implementation.

II. METHODOLOGY

A. EXPERIMENTAL SETUP AND DATA ACQUISITION

To ensure realistic and dynamic input for the optimization process, this study employed a two-stage experimental approach: (a) operational testing of a residential-scale PV–BESS system under household load conditions, and (b) controlled cycling experiments to collect battery degradation data. The experimental setup is as shown in Figure 1.

1) PV–BESS OPERATIONAL DATA

The experiment setup for PV-BESS system was developed consisting of a 1.2 kW PV plant integrated with a lithium-ion battery energy storage system and connected to typical residential household appliances. The purpose of this setup was to replicate realistic operating conditions in a grid-independent environment, where the PV system must dynamically balance energy supply and demand. During operation, the following parameters were continuously measured and logged:

- PV system variables: solar irradiance, ambient temperature, and instantaneous PV power generation.
- Battery variables: SoC, DoD, charging/discharging current, terminal voltage, and temperature.
- Load variables: instantaneous and aggregated household load demand, measured at different times of the day to capture variability.

The system was monitored over extended operational periods, covering both day-to-day and weather-related fluctuations. All parameters were recorded with a 1-second sampling interval, enabling the dataset to capture rapid fluctuation in both solar generation and load consumption. This high-resolution monitoring produced a detailed time-series dataset reflecting the stochastic and nonlinear behavior of PV–BESS systems under real-world conditions. Unlike purely simulated datasets, this experimental dataset closely represents practical operating scenarios, thereby improving the reliability and applicability of the optimization results.

2) BATTERY DEGRRADATION DATA

In addition to the PV operational data, a battery degradation experiment was conducted in the laboratory using the CHROMA 17020 regenerative battery pack test system. The tested battery was a 12 V 100 Ah lithium-ion unit operated under controlled charge and discharge cycles. Throughout the process, key parameters such as SoC, DoD, current, and voltage were continuously recorded to capture the degradation characteristics of the battery.

The experiment was carried out for 30 complete cycles. The limitation in the number of cycles was due to the lengthy duration required for each full charge–discharge process, which made it impractical to extend the test to a larger number of repetitions. To ensure broader applicability, the results obtained from these 30 cycles were generalized to represent longer-term cycling behavior. This generalization is justified by the fact that residential electrical loads, which serve as the case study in this research, typically follow repetitive and relatively consistent daily consumption patterns. Consequently, the degradation trends observed in the laboratory are considered representative of real-world residential battery operation over extended periods.

The degradation dataset captured key indicators of battery ageing, including:

- state of charge dynamics
- capacity fade: reduction of usable capacity over repeated cycles.

As with the operational dataset, degradation parameters were logged every 1 s, ensuring high temporal granularity to capture transient behaviors and subtle ageing effects during cycling

B. OPTIMIZATION OBJECTIVE

The primary objective was to identify the optimal operating parameters of the battery, specifically the DoD and C-rate that minimize the power mismatch between supply (PV and battery) and demand, while indirectly reducing degradation. Here, DoD represents the fraction of the battery's nominal capacity that is utilized during a charge–discharge cycle, while the C-rate defines the charging or discharging current normalized to the battery capacity. High DoD operation and high C-rates are well known to accelerate battery degradation by increasing electrochemical stress, temperature rise, and side reactions within the cell [13].

Power mismatch serves as a quantifiable proxy for stability since persistent mismatch leads to frequency deviations in islanded systems or excessive grid exchange in grid-connected systems. The instantaneous power mismatch is defined as (1).

$$\Delta P(t) = | P_{load}(t) - (P_{PV}(t) + P_{batt}(t)) | \quad (1)$$

The optimization problem is formulated as (2).

$$\min J = \sum_{t=1}^{T} \Delta P(t) \quad (2)$$

This optimization was subject to battery operational constraints (SoC, current, and efficiency) and explored the effects of DoD and C-rate, which were critical determinants of degradation. The optimal pair of DoD and C-rate was obtained by simulation-based search. To solve this optimization problem, RL was employed as a data-driven decision-making framework in which the BESS operated as an agent interacting with the power system environment. At each time step, the agent observed the system state and selected control actions related to battery operation, implicitly determining the effective DoD and C-rate, while receiving a reward based on the resulting power mismatch. Through repeated interactions, the RL agent learnt an optimal operating patterns to achieve stable operation with reduced battery degradation.

C. SYSTEM MODELING

1) POWER BALANCE

The power balance of the PV-BESS is modelled as (3).

$$P_{PV}(t) + P_{batt}(t) = P_{load}(t) + P_{loss}(t) \quad (3)$$

where $P_{batt}(t) > 0$ denotes discharging, $P_{batt}(t) < 0$ denotes charging and P_{loss} represent inverter or conversion loss.

2) BATTERY CAPACITY DYNAMIC

The battery's SoC evolves according to the following discrete-time relation:

$$SoC_{t+1} = SoC_t + \frac{\eta_{ch} P_{ch,t} - \frac{1}{\eta_{dis}} P_{dis,t}}{E_{nom}} \Delta t \quad (4)$$

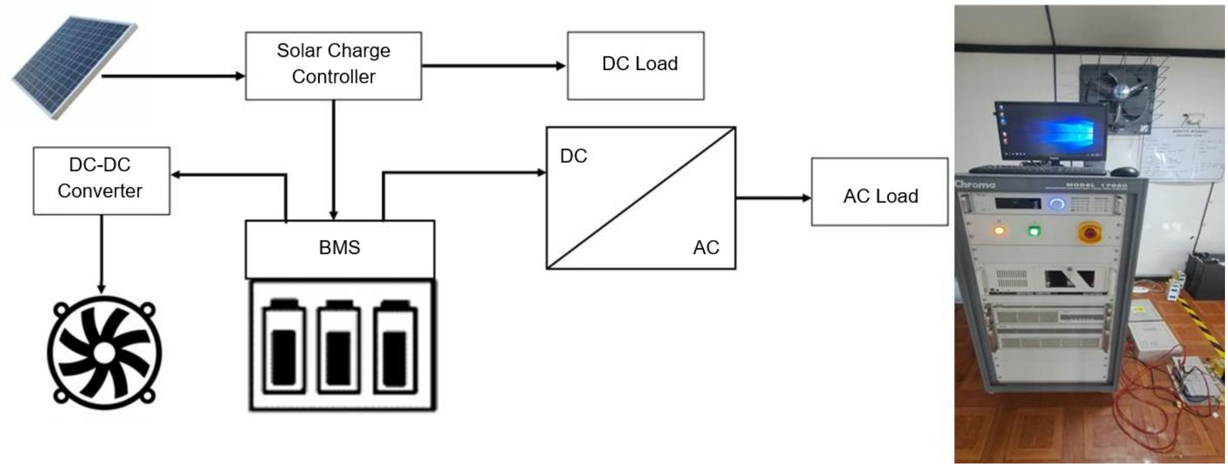

Figure 1. Experiment setup for PV and battery degradation data collection.

subject to the operational bounds:

$$SoC_{min} \leq SoC_t \leq SoC_{max} \tag{5}$$

These constraints ensured safe operation and defined the operating envelope for degradation analysis. Here, DoD and C-rate directly influenced SoC trajectories, current stress, and degradation exposure. DoD and C-rate were explicitly varied in the simulation to quantify their impact on mismatch minimization and long-term degradation. The maximum current constraint is expressed as:

$$|I_t| \leq I_{max} \tag{6}$$

P_{ch} and P_{dis} denote charging and discharging power, η_{ch} and η_{dis} are charging and discharging efficiencies, and E_{nom} is the battery's nominal energy capacity.

In addition to SoC and current limits, battery degradation was explicitly considered as a constraint in the optimization problem. Degradation is primarily driven by two operational stress factors: DoD and C-rate. Empirical studies have shown that capacity fade can be modeled as a nonlinear function of these parameters:

$$\Delta Q \propto (DoD)^{\alpha}.(C\_rate)^{\beta} \tag{7}$$

To ensure practical applicability, the optimization was constrained such that the cumulative degradation over the scheduling horizon remains below a specified threshold:

$$f_{deg}(DoD, C\_rate) \leq f_{max} \tag{8}$$

This constraint guaranteed that the RL agent did not select operational policies that excessively stressed the battery, even if such policies minimized instantaneous power mismatch. Consequently, the optimization resulted in the determination of an optimal DoD and C-rate operating region that jointly minimized power dispatch mismatch while ensuring long-term battery health.

D. REINFORCEMENT LEARNING OPTIMIZATION

The energy management problem is formulated as a Markov decision process (MDP), which provides a structured framework to model sequential decision-making under uncertainty. The MDP is characterized by the state space, action space, state transition dynamics, and reward function, respectively.

1) STATE VECTOR

At each time step, the system state is represented as:

$$s_t = [SoC_t, P_{PV,t}, P_{load,t}, I_{t-1}, T_t, TOD_t] \tag{9}$$

where SoC_t is state of charge, $P_{PV,t}$ is the photovoltaic generation, $P_{load,t}$ is the demand load, I_{t-1} is the previous battery current, TOD_t represents time-of-delay index. This combination captures the essential information for decision-making including the battery internal state, external supply–demand conditions, thermal effects, and temporal context. SoC_t links directly to operational constraints and indirectly to degradation.

2) ACTION VECTOR

The control action was the battery charging or discharging power P_{batt} as explained in (3). Positive values correspond to discharging (supplying the grid or load), whereas negative values correspond to charging. By directly controlling $P_{batt,t}$, the RL agent influenced the SoC trajectory and consequently managed both operational performance and degradation exposure.

3) TRANSITION DYNAMIC

The state transition was governed by the battery dynamics in (4)–(6). These equations ensured that any action taken respected the physical evolution of SoC, the safety bounds on SoC, and the maximum allowable current. Thus, the RL environment integrated electrochemical feasibility into the learning process, preventing infeasible actions from being explored.

E. REWARD FUNCTION

The reward was designed to minimize power mismatch between supply and demand:

$$r_t = \frac{-\Delta P(t)}{P_{base}} \tag{10}$$

where $\Delta P(t)$ is power mismatch which refer to (1). P_{base} is the nominal PV rating used for normalization. The RL agent was used here as a simulation optimizer to explore the trade-off between mismatch minimization and degradation across different DoD and C-rate settings. A smaller mismatch resulted in a higher reward (closer to zero), guiding the agent toward balancing local generation and consumption.

The reward design ensured two desirable outcomes:

- Grid support: By reducing mismatch, the system stabilized power exchange with the grid or minimizes off-grid imbalance. In practice, acceptable power

mismatch levels were usually constrained to within approximately 1–5% of the rated system power to limit frequency and voltage deviations.
- Battery health preservation: Since mismatch minimization inherently discouraged unnecessary cycling and excessive SoC swings, the RL agent indirectly reduced degradation stress factors, such as deep discharges and high DoD cycles.

F. RL ALGORITHM

A soft actor-critic (SAC) algorithm was adopted due to its suitability for continuous action spaces and its entropy-regularized objective, thereby promoting exploration. The SAC's actor and critic networks each consisted of three fully connected layers with 256 neurons per layer (ReLU activation). Training parameters used in this research is as shown in Table I.

G. SIMULATION AND VALIDATION

The environment was implemented in Python using Google Collab platform. Each episode corresponded to 24 hours of operation. Data were split into training (70%), validation (15%), and testing (15%). The evaluation metrics include:

- Average power mismatch:

$$\overline{\Delta P} = \frac{1}{T}\sum_{t=1}^{T} \Delta P(t) \tag{11}$$

- Cumulative capacity loss was obtained from empirical degradation functions that relate capacity fade to DoD and C-rate. This metric reflects the long-term impact of the optimized operating strategy on battery health.
- System stability index (SSI) was derived from the standard deviation of net power mismatch, indicating the degree of power balance stability over the scheduling horizon as expressed in (12):

$$SSI = \max\left(0.1 - \frac{\sigma P}{P_{base}}\right) \tag{12}$$

where σP is the standard deviation of net power mismatch.

These metrics allowed identification of the optimal DoD and C-rate settings that minimized mismatch while indirectly limiting degradation.

III. RESULTS AND DISCUSSION

A. EXPERIMENT RESULTS

The experimental results in Figure 2(a) show the battery capacity degradation over the first 30 charge–discharge cycles. The capacity decreased gradually from nearly 100% to around 99.4%, indicating a small but consistent decline with each cycle. This trend demonstrates the early-stage degradation behavior where the rate of capacity fade is relatively linear and slow. Meanwhile, Figure 2(b) presents the simulated daily power usage over 30 days. The consumption fluctuated irregularly between 0.6–1.0 kWh per day, with an average of approximately 0.78 kWh.

The small deviation around the mean reflects realistic daily variations in residential load while maintaining a stable demand level below the 1 kWh limit. These results confirm the assumption that daily power consumption does not vary significantly, as evidenced by the narrow range of fluctuations in Figure 2(b). Together, the findings highlight the relationship between repetitive cycling and gradual capacity fade, as well as the relatively consistent daily energy usage profile that drives the cycling stress.

TABLE I
TRAINING PARAMETERS

Parameters	Values
Learning rate	3×10^{-4}
Replay buffer size	10^{6}
Batch size	256
Discount factor	0.99

B. LEARNING PERFORMANCE OF RL AGENT

The learning performance of the proposed RL agent using the SAC algorithm is illustrated in Figure 3. The training curve shows a clear improvement in average reward over the course of 200 episodes, indicating that the agent effectively learns to minimize the normalized power mismatch. Initially, the reward was close to −1, reflecting large mismatches between supply and demand. However, as training progresses, the curve rose steadily and converged near zero, corresponding to the condition of nearly perfect power balance ($\Delta P \approx 0$). The smooth convergence trend without significant oscillations suggests that the SAC agent achieves both stable learning and robust policy optimization. This result validates the effectiveness of the proposed method in guiding the agent toward optimal charging and discharging decisions that balance load and generation while respecting system dynamics.

C. OPTIMAL PARAMETERS SETTING

After confirming the stable learning behavior of the SAC-based RL agent in Figure 3, the next step was to evaluate the operational parameters suggested by the trained model. The agent gradually improved its decisions during training by continuously interacting with the environment, adjusting charging and discharging actions, and receiving feedback through the reward function. Once the learning process converged, the final policy consistently produced a set of parameter values reflecting a balanced trade-off between system efficiency and battery longevity. These values are presented in Table II.

The RL results showed that the most suitable depth of discharge was around 65%, balancing usable capacity while reducing stress on the battery. This level prevents deep cycles that accelerate capacity fade while still ensuring sufficient energy to cover load fluctuations. In practice, such a mid-range DoD is widely regarded as a compromise between performance and durability, and in this case, the RL agent identified it autonomously through repeated interactions with the environment.

For the charging and discharging rates, the RL policy converged to 0.5C, a moderate and safe limit. In practical terms, for a 12 V 1000 Ah battery (nominal capacity 12 kWh), this translates to a maximum charge or discharge power of about 50 A or 6 kW. Coupled with a 65% DoD setting, the usable energy is roughly 7.8 kWh per cycle. These constraints reduce risks such as overheating and lithium plating while maintaining adequate flexibility to respond to daily PV and load variations. Taken together, the results show that the RL agent not only optimizes operational performance but also embeds a safeguard for long-term battery health, making the approach practical and sustainable for PV–BESS applications.

D. EVALUATION OF OPTIMIZATION METRICS

To further assess the effectiveness of the proposed RL framework, several optimization metrics were evaluated. These

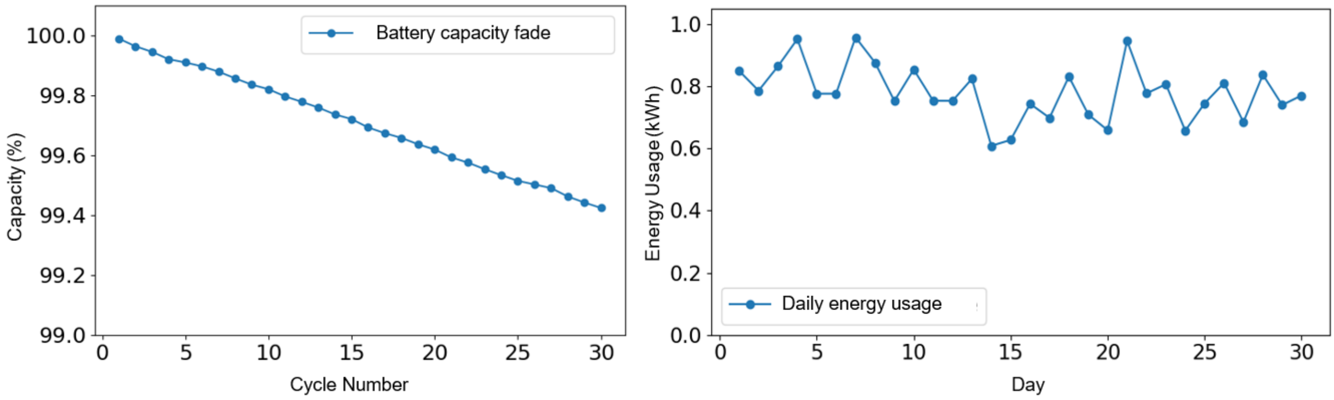

Figure 2. Battery capacity fade recorded from (a) battery test experiment and (b) daily power usage of experimental object.

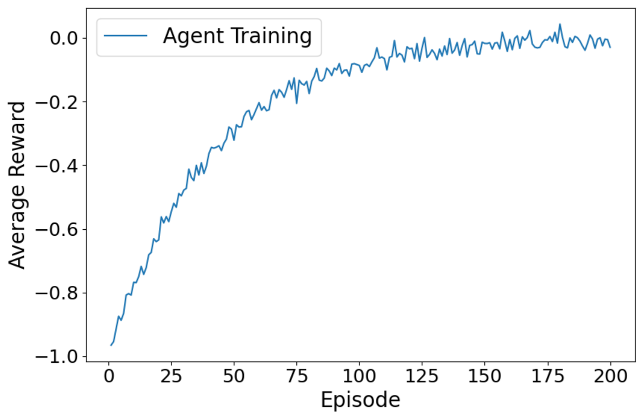

Figure 3. Learning performance of RL agent (SAC).

TABLE II
OPTIMUM BATTERY PARAMETERS SETTING

Parameters	Values	Practical Implementation
DoD	65%	7.8 kWh
C-rate	0.5	50 A

metrics were selected to capture both operational performance and long-term battery sustainability.

The three primary indicators were average power mismatch, cumulative capacity loss, and SSI. The illustration of these evaluations for 30 days based on simulated data is plotted in Figure 4. The first subplot illustrates the variation of average power mismatch over a 30-day horizon. Prior to the implementation of the RL–based optimization, the average power mismatch was maintained at approximately 0.1 kW under conventional operation, indicating a stable but suboptimal balance between supply and demand. After applying the proposed RL-based scheduling framework, the mismatch values were significantly reduced and remained centered close to zero, with fluctuations generally confined within ±0.01 kW. This substantial reduction demonstrates the ability of the RL agent to effectively coordinate the charging and discharging cycles of the battery energy storage system in response to daily demand variations. Although small oscillations were observed, they remained within acceptable operational limits and can be attributed to the inherent variability of residential load and renewable generation. Overall, the comparison confirms that the proposed strategy markedly improves power balance performance beyond the baseline operation.

The second subplot presents the progression of cumulative capacity loss over the same 30-day period. Prior to the implementation of the RL–based optimization, the cumulative capacity loss reached approximately 0.27%, reflecting the impact of repeated cycling with less regulated depth of discharge and C-rate. After applying the proposed RL-based scheduling strategy, the cumulative degradation was reduced and remained below 0.22% throughout the observation window, indicating a measurable improvement in battery health preservation. The gradual and smooth increases in capacity loss was consistent with the expected aging behavior of lithium-ion batteries under moderate cycling conditions. Moreover, the absence of abrupt changes suggests that the optimized scheduling effectively avoids severe degradation mechanisms, such as deep cycling or operation at extreme states of charge. Overall, this comparison demonstrates the capability of the proposed method to minimize battery degradation.

The third subplot illustrates the evolution of the SSI over the 30-day simulation period. Prior to the implementation of the RL-based optimization, the SSI was maintained at approximately 0.95, indicating a generally stable but less optimized operating condition. Following the implementation of the proposed scheduling framework, the SSI consistently improved, remaining within the range of 0.95 to 0.97, with an average value of approximately 0.96. This improvement reflects enhanced power balance stability and reduced volatility in system operation. Although a minor dip was observed around day 16, the system rapidly recovered to its nominal stability range, demonstrating the robustness of the optimized control strategy under fluctuating load and generation conditions. Overall, the results confirm that the proposed framework provides reliable and sustained stability, which is essential for practical deployment in off-grid residential PV–BESS systems.

To further contextualize the results, the average daily rate of capacity loss was approximated using the linear growth observed. With the cumulative capacity loss was 0.22% at day 30, the average daily degradation rate was 0.0073% per day as shown in Table III. Extrapolating this trend over a year yielded an estimated annual degradation of around 2.7%, which was considerably lower than the 2.5–3% typically reported in uncontrolled cycling scenarios.

The quantitative comparison as shown in Table III provides further evidence that the proposed method not only enhances operational efficiency but also preserves the long-term functionality of the battery. The balance achieved between

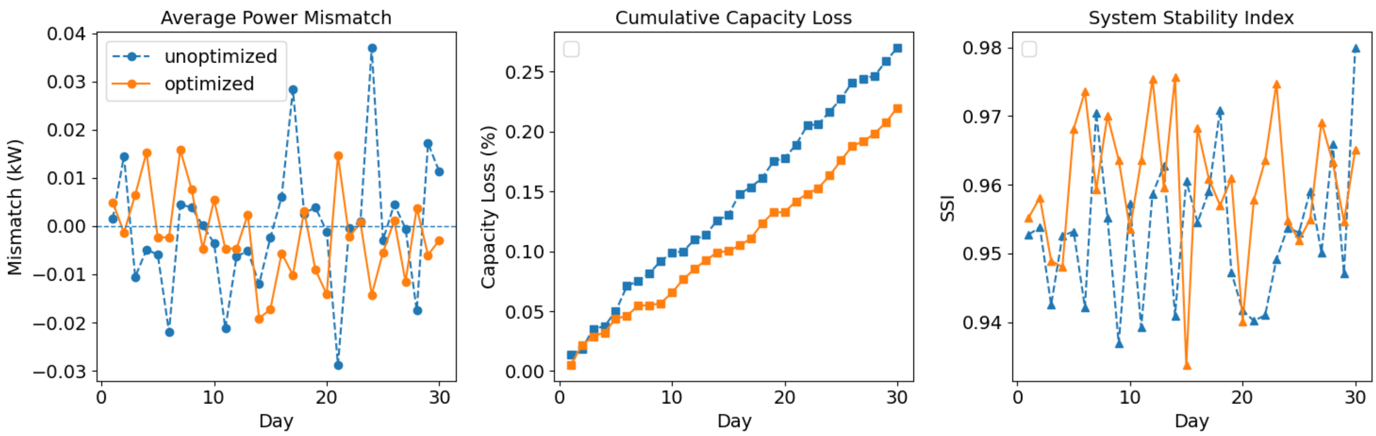

Figure 4. Daily performance of the proposed reinforcement learning-based scheduling framework over a 30-day horizon: (a) average power mismatch, (b) cumulative capacity loss, and (c) system stability index.

TABLE III
KEY METRICS AFTER 30 DAYS

Parameters	Values	Interpretation
Average power Mismatch	0.002 kW	Near-perfect load balance
Cumulative capacity loss	0.22%	Low degradation
SSI	0.96	High stability, minimal fluctuation
Daily degradation rate	0.0073%	Allow long lifetime

low mismatch, controlled degradation, and high stability underscore the practical relevance of the strategy. The integrated assessment of mismatch, degradation, and stability demonstrates the effectiveness of the proposed method. By jointly considering short-term market alignment (mismatch minimization) and long-term asset preservation (capacity loss reduction), the contribution lies in bridging operational efficiency with sustainability. Unlike conventional strategies that optimize only for immediate profit or only for longevity, this approach achieves a balanced trade-off. Thus, the framework provides both economic and technical benefits, supporting higher penetration of BESS in renewable-dominated systems.

The metrics summarized in Table III provide a concise representation of the system's overall performance after the 30-day horizon. The extremely low average power mismatch (~0.002 kW) reflects the precision of the optimization framework in synchronizing charging and discharging with daily demand variations. Similarly, the cumulative capacity loss of only 0.22% demonstrates that degradation remains well within acceptable limits, confirming the ability of the scheduling policy to preserve long-term battery health. The consistently high stability index, with an average of 0.96, further validates the resilience of the control method in maintaining reliable operation. Finally, the derived daily degradation rate of 0.0073%/day offers a useful benchmark for estimating long-term impacts, showing that the system could operates sustainably for years without compromising technical or economic feasibility. Together, these results reinforce the value of the proposed method in delivering balanced, durable, and stable performance for PV–BESS integration.

To further assess the performance, the proposed RL method was compared with two widely optimization techniques from previous study: simulated annealing (SA) algorithm and MOGA. The comparison is illustrated in the Figure 5, which presents three key performance indicators: average power mismatch, cumulative capacity loss, and system stability index. These indicators provide a comprehensive view of both short-term operational performance and long-term sustainability of the battery system under different optimization strategies.

Figure 5 illustrates the trade-offs among three optimization methods across three performance metrics: average power mismatch, cumulative capacity loss, and system stability index. The proposed RL method demonstrated a well-balanced performance, showing very low average power mismatch (0.002 kW), moderate degradation (0.22%), and the highest stability index (0.96). These results indicate that RL can adaptively balance short-term operational accuracy and long-term system sustainability, making it the most promising candidate for real-world deployment.

In contrast, MOGA achieved zero mismatches, indicating optimal load balancing. However, the same figure shows that this comes at the cost of significantly higher cumulative degradation (0.60%). This suggests that MOGA enforces aggressive cycling of the battery to eliminate mismatch, accelerating its wear and reducing its lifespan. Despite a relatively strong stability index (0.92), the high degradation rate raises concerns over long-term economic viability, as frequent battery replacements would offset short-term performance gains.

SA, on the other hand, represents the opposite trade-off. It produced the largest mismatch (0.012 kW), indicating less effective performance in minimizing load imbalances. However, this conservative operation resulted in the lowest degradation among the methods (0.18%), reflecting its bias toward preserving battery health by reducing cycling intensity. Its stability index of 0.90, while acceptable, was the lowest of the three, highlighting greater variability in system operation compared to RL and MOGA. Overall, the figure underscores the importance of considering multiple performance dimensions rather than focusing on a single objective. While MOGA prioritizes mismatch minimization and SA emphasizes battery preservation, only the proposed RL method provides a balanced outcome by achieving low mismatch, reasonable degradation, and high stability simultaneously. This balanced trade-off is critical for practical energy storage applications, where both short-term reliability and long-term sustainability must be optimized together.

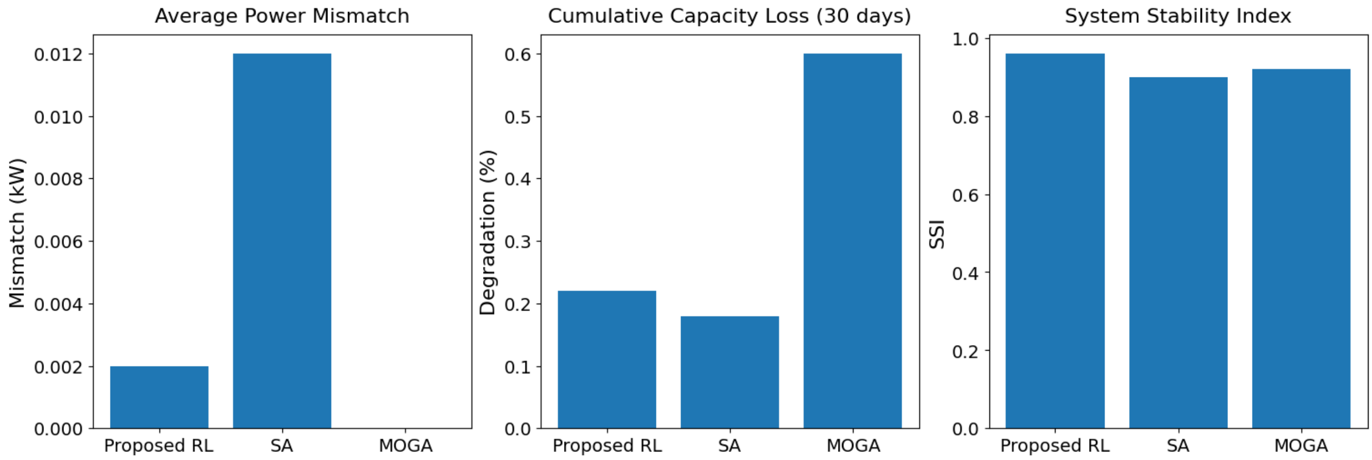

Figure 5. Comparative performance of the proposed reinforcement learning approach, simulated annealing, and MOGA over a 30-day horizon: (a) average power mismatch, (b) cumulative capacity loss, and (c) system stability index.

IV. CONCLUSION

This study introduced an RL-based scheduling framework for lithium-ion battery storage, designed to balance short-term operational performance with long-term degradation management. Over a 30-day horizon, the proposed method achieved an exceptionally low average power mismatch of 0.002 kW, confirming near-perfect load balancing. At the same time, cumulative capacity loss was limited to 0.22%, corresponding to a daily degradation rate of 0.0073%, which, when extrapolated to 2.7% per year, was significantly lower than the 2.5–3% typically reported under uncontrolled cycling. In addition, the SSI consistently averaged 0.96, underscoring the robustness of the framework under varying residential load and renewable generation conditions. A comparative assessment against SA and MOGA further demonstrated the superiority of the proposed RL method. While MOGA minimized mismatch to zero, this was accompanied by a much higher degradation of 0.60%, highlighting the risk of accelerated battery wear. Conversely, SA preserved battery health more effectively, with only 0.18% degradation, but suffered the largest mismatch (0.012 kW) and the lowest stability (0.90). The RL framework achieved the most balanced outcome, combining low mismatch, controlled degradation, and the highest stability, making it more suitable for real-world deployment where both reliability and sustainability are required. Despite these promising results, the study has several limitations. The degradation model used is simplified and does not account for factors such as temperature variation, calendar aging, or heterogeneous cell behavior. In addition, the evaluation was conducted solely on simulation data, which may not fully capture the uncertainties of real-world electricity markets, user demand variability, or hardware constraints. Future work should address these limitations by integrating physics-informed degradation models, incorporating stochastic market and load scenarios, and validating performance through hardware-in-the-loop or field experiments.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

AUTHORS' CONTRIBUTIONS

Conceptualization, Muhamad Dzaky Ashidqi; methodology, Muhamad Dzaky Ashidqi; software, Muhamad Dzaky Ashidqi; validation, Muhamad Dzaky Ashidqi, Silviana Windasari; formal analysis, Muhamad Dzaky Ashidqi; Probokusumo; Silviana Windasari; investigation, Muhamad Dzaky Ashidqi; Rahmat; resources, Muhamad Dzaky Ashidqi; Silviana Windasari; Rahmat; data curation, Muhamad Dzaky Ashidqi, Silviana Windasari, Rahmat; writing—original draft preparation, Muhamad Dzaky Ashidqi; Silviana Windasari; writing—reviewing and editing, Silviana Windasari; Probokusumo; visualization, Probokusumo; supervision, Muhamad Dzaky Ashidqi; project administration, Silviana Windasari; funding acquisition, Muhamad Dzaky Ashidqi.

ACKNOWLEDGMENT

The authors extend their gratitude to the Ministry of Higher Education, Research and Technology (Kemdiktisaintek) as this work was funded through Penelitian Dosen Pemula (PDP) Research Grant, with main contract number 125/C3/DT.05.00/PL/2025 and derived contract number 7971/LL4/PG/2025.

REFERENCES

[1] N. Mlilo, J. Brown, and T. Ahfock, "Impact of intermittent renewable energy generation penetration on the power system networks – A review," *Technology and Economics of Smart Grids and Sustainable Energy*, vol. 6, no. 1, 2021, doi: 10.1007/s40866-021-00123-w.

[2] S. Motta, M. Aro, C. Evens, A. Hentunen, and J. Ikäheimo, "A Cost-Benefit Analysis of Large-Scale Battery Energy Storage Systems For Frequency Markets," in *IET Conference Proceedings*, 2021, pp. 3130–3134. doi: 10.1049/icp.2021.1470.

[3] C. Zhao, P. B. Andersen, C. Træholt, and S. Hashemi, "Grid-connected battery energy storage system: a review on application and integration," *Renewable and Sustainable Energy Reviews*, vol. 182, 2023, doi: 10.1016/j.rser.2023.113400.

[4] H. Wigger, F. Sill Torres, J. Bartels, L. Feller, U. Brand-Daniels, and T. Vogt, "Quantitative Assessment of Sector Coupling Battery Energy Storage Systems and Their Contribution to the Resilience of the Power System," *Int J Energy Res*, vol. 2024, no. 1, 2024, doi: 10.1155/2024/8812115.

[5] M. Sangsoda, A. Prasertnoi, and J. Kluabwang, "A Comprehensive Review of Battery Energy Storage Systems for Grid Stability and Renewable Integration: A Focus on Lithium-Ion's Grid Batteries in Thailand," in *10th IEEE International Smart Cities Conference: Smart Cities: Revolution for Mankind, ISC2 2024 - Proceedings*, 2024. doi: 10.1109/ISC260477.2024.11004256.

[6] Y. Gu and T. C. Green, "Power System Stability With a High Penetration of Inverter-Based Resources," *Proceedings of the IEEE*, vol. 111, no. 7, pp. 832–853, 2023, doi: 10.1109/JPROC.2022.3179826.

[7] C. Ghanjati and S. Tnani, "Optimal sizing and energy management of a stand-alone photovoltaic/pumped storage hydropower/battery hybrid system using Genetic Algorithm for reducing cost and increasing reliability," *Energy and Environment*, vol. 34, no. 6, pp. 2186–2203, 2023, doi: 10.1177/0958305X221110529.

[8] Y. Yang, S. Bremner, C. Menictas, and M. Kay, "Modelling and optimal energy management for battery energy storage systems in renewable energy systems: A review," *Renewable and Sustainable Energy Reviews*, vol. 167, 2022, doi: 10.1016/j.rser.2022.112671.

[9] T. Rahman and T. Alharbi, "Exploring Lithium-Ion Battery Degradation: A Concise Review of Critical Factors, Impacts, Data-Driven Degradation Estimation Techniques, and Sustainable Directions for Energy Storage Systems," *Batteries*, vol. 10, no. 7, 2024, doi: 10.3390/batteries10070220.

[10] G. Zhang, X. Wei, S. Chen, G. Han, J. Zhu, and H. Dai, "Investigation the Degradation Mechanisms of Lithium-Ion Batteries under Low-Temperature High-Rate Cycling," *ACS Appl Energy Mater*, vol. 5, no. 5, pp. 6462–6471, 2022, doi: 10.1021/acsaem.2c00957.

[11] J. Schmitt, A. Maheshwari, M. Heck, S. Lux, and M. Vetter, "Impedance change and capacity fade of lithium nickel manganese cobalt oxide-based batteries during calendar aging," *J Power Sources*, vol. 353, pp. 183–194, 2017, doi: 10.1016/j.jpowsour.2017.03.090.

[12] B. Sood, L. Severn, M. Osterman, M. Pecht, A. Bougaev, and D. McElfresh, "Lithium-ion battery degradation mechanisms and failure analysis methodology," in *Conference Proceedings from the International Symposium for Testing and Failure Analysis*, 2012, pp. 239–249.

[13] G. Aravindhanathan, K. Narayanan, A. S. Rana, A. A. Tellez, T. Senjyu, and A. Sharma, "Analysis of Cyclic Degradation of Li-ion batteries," in *Proceedings of the IEEE Power India International Conference, PIICON*, 2024. doi: 10.1109/PIICON63519.2024.10995090.

[14] S. Nazaralizadeh, P. Banerjee, A. K. Srivastava, and P. Famouri, "Battery Energy Storage Systems: A Review of Energy Management Systems and Health Metrics," *Energies (Basel)*, vol. 17, no. 5, 2024, doi: 10.3390/en17051250.

[15] M. Dubarry and A. Devie, "Battery durability and reliability under electric utility grid operations: Representative usage aging and calendar aging," *J Energy Storage*, vol. 18, pp. 185–195, 2018, doi: 10.1016/j.est.2018.04.004.

[16] M. D. Ashidqi, A. I. Cahyadi, and A. Ataka, "Capacity Loss Modeling of Li-Ion Battery Using Lightweight Neural Network Considering Equivalent Circuit Model," in *ICT-PEP 2023 - 2023 International Conference on Technology and Policy in Energy and Electric Power: Decarbonizing the Power Sector: Opportunities and Challenges for Renewable Energy Integration, Proceedings*, 2023, pp. 133–138. doi: 10.1109/ICT-PEP60152.2023.10351143.

[17] L. A. Wong, V. K. Ramachandaramurthy, S. L. Walker, P. Taylor, and M. J. Sanjari, "Optimal placement and sizing of battery energy storage system for losses reduction using whale optimization algorithm," *J Energy Storage*, vol. 26, 2019, doi: 10.1016/j.est.2019.100892.

[18] B. Cortés-Caicedo, L. F. Grisales-Noreña, O. D. Montoya, and R. I. Bolaños, "Optimization of BESS placement, technology selection, and operation in microgrids for minimizing energy losses and CO2 emissions: A hybrid approach," *J Energy Storage*, vol. 73, 2023, doi: 10.1016/j.est.2023.108975.

[19] P. Parczyk and R. Burduk, "Optimisation of Battery Energy Storage Systems Capacity for Purpose to Reduce Energy Cost," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2024, pp. 414–427. doi: 10.1007/978-3-031-71115-2_29.

[20] V. Saxena, N. Kumar, and U. Nangia, "Computation and Optimization of BESS in the Modeling of Renewable Energy Based Framework," *Archives of Computational Methods in Engineering*, vol. 31, no. 4, pp. 2385–2416, 2024, doi: 10.1007/s11831-023-10046-7.

[21] C. Jiang, S. Qin, K. Lin, S. Liu, F. Deng, and J. Zhang, "Optimal Operation Strategy for Integrated Energy Systems Based on Improved Grey Wolf Optimization Algorithm," in *2024 3rd International Conference on Power Systems and Electrical Technology, PSET 2024*, 2024, pp. 720–724. doi: 10.1109/PSET62496.2024.10808556.

[22] Y. Zhi, D. Gao, G. Wei, and X. Yang, "Simultaneous sizing and scheduling optimization for farmhouse PV-battery systems with a multi-structured power system model," *J Energy Storage*, vol. 104, 2024, doi: 10.1016/j.est.2024.114564.

[23] R. G. Mohamed, A. A. Hassan, and S. H. E. Abdel Aleem, "Techno-economical-environmental approach using multi-objective-based Pareto frontier optimization for hybrid photovoltaic/diesel/battery system considering different battery types," *J Energy Storage*, vol. 115, 2025, doi: 10.1016/j.est.2025.115979.

[24] S. Pires Pimentel *et al.*, "A Systematic Literature Review on Li-Ion BESSs Integrated with Photovoltaic Systems for Power Supply to Auxiliary Services in High-Voltage Power Stations," *Energies (Basel)*, vol. 18, no. 13, 2025, doi: 10.3390/en18133544.

[25] J. He, T. Yang, L. Xie, Y. Yang, C. Chen, and J. Wei, "A Data-Driven Reinforcement Learning Enabled Battery Fast Charging Optimization Using Real-World Experimental Data," *IEEE Transactions on Industrial Electronics*, vol. 72, no. 1, pp. 430–438, 2025, doi: 10.1109/TIE.2024.3398687.