Renewable Energy and Power Quality Journal
https://repqj.com/index.php/repqj/index
RE&PQJ, Vol. 23, No. 4, 2025

Real-time Control Method for Charging and Discharging of Large-capacity Batteries in Intelligent Networks

Ming Lu[1], Yupeng Cai[2,*], Zhiyu Liu[1], Qingyang Tian[1], Meng Wu[2], Xiangdong Gong[2]

[1]State Grid Liaoning Electric Power Co. Ltd, Shenyang, 110006, China
[2]State Grid Liaoning Electric Power Research Institute, Shenyang, 110006, China
[*]Corresponding author's email: lm20250403@163.com

Abstract. This paper proposes a real-time control method for optimizing the charging and discharging of large-capacity batteries, using intelligent algorithms to improve efficiency, scheduling accuracy and response speed. The method improves battery utilization and extends battery life by real-time monitoring of battery status, load demand and grid fluctuations. An improved multi-layer feature fusion long short-term memory (LSTM) model is used to predict the battery state of health (SOH), and an adaptive time window weighting strategy is used to enhance the model's response to short-term grid load changes. The hierarchical reinforcement learning (HRL) framework optimizes the scheduling strategy through high-level task decomposition and low-level dynamic adjustment. In order to accelerate strategy search, an adaptive particle swarm optimization (PSO) algorithm is used. Experimental results show that this method is superior to the comparative method. When the peak-to-valley difference is 30% and the temperature is 15 degrees, the charging and discharging efficiency is improved by up to 9.7%, and the energy consumption is optimized by up to 20%, and it can still maintain good adaptability and stability under extreme conditions.

Key words. Large-capacity battery, Charging and discharging control, Battery health management, Time-varying load prediction, Multi-objective optimization

1. Introduction

As large-capacity batteries are widely applied in power systems, electric vehicles, and renewable energy storage [1,2], battery charging and discharging control has become a key technology to ensure efficient battery operation, extend battery life, and improve grid stability [3,4]. In wind farm energy storage systems, battery control is utilized to balance the fluctuations in wind power output and improve system stability through precise charging and discharging management. In microgrid frequency modulation applications, battery charging and discharging strategy optimization can achieve frequency modulation to ensure the power grid's supply and demand balance and stable operation. Traditional rule-based control methods cannot respond to changes in the battery's state of health and grid load in real time, which often leads to inefficient battery use and makes it difficult to adapt to the needs of complex grid environments.

Existing research has made some progress in optimizing battery charging and discharging control. Many studies have focused on battery health management, using modeling and data-driven methods to evaluate and predict battery state to guide charging and discharging decisions [5,6]. In terms of grid load prediction, there are also studies that improve the precision of load fluctuation predictions through techniques based on machine learning [7,8] and time series analysis [9,10]. Although existing research has made progress in battery charging and discharging control, there are still many shortcomings. On the one hand, the accuracy of battery health prediction and grid load prediction needs to be improved. On the other hand, the dynamic adjustment strategy of each objective weight in multi-objective optimization is not perfect, which can easily lead to uneven optimization effects. Existing charging and discharging control strategies have challenges in dynamic response capabilities, local optimal problems, and multi-objective optimization balance.

To address the above problems, this paper combines HRL with PSO algorithm and improved LSTM model to realize real-time control of charging and discharging of large-capacity batteries in smart grids. For battery health assessment, this paper improves the prediction accuracy through multi-layer feature fusion and evaluates the battery state based on real-time data. In power grid load scheduling, an adaptive time window weighting strategy is adopted in combination with battery health information to improve the responsiveness of power grid load fluctuation prediction. Additionally, for the charging and discharging scheduling strategy, this paper applies the HRL framework, which integrates global scheduling and local control to optimize the decision-making efficiency and precision during the charging and discharging process. This paper makes improvements based on the PSO algorithm to further improve the scheduling efficiency. Through the adaptive particle update strategy, the search behavior of particles is dynamically adjusted to avoid local optimal solutions and achieve the global optimization of battery charging and discharging scheduling. Through these improvements, this paper realizes real-time optimization and control of battery charging and discharging strategies, significantly improves battery utilization efficiency and life, and enhances the system's adaptability to grid load fluctuations.The structure of this paper is shown in Figure 1.

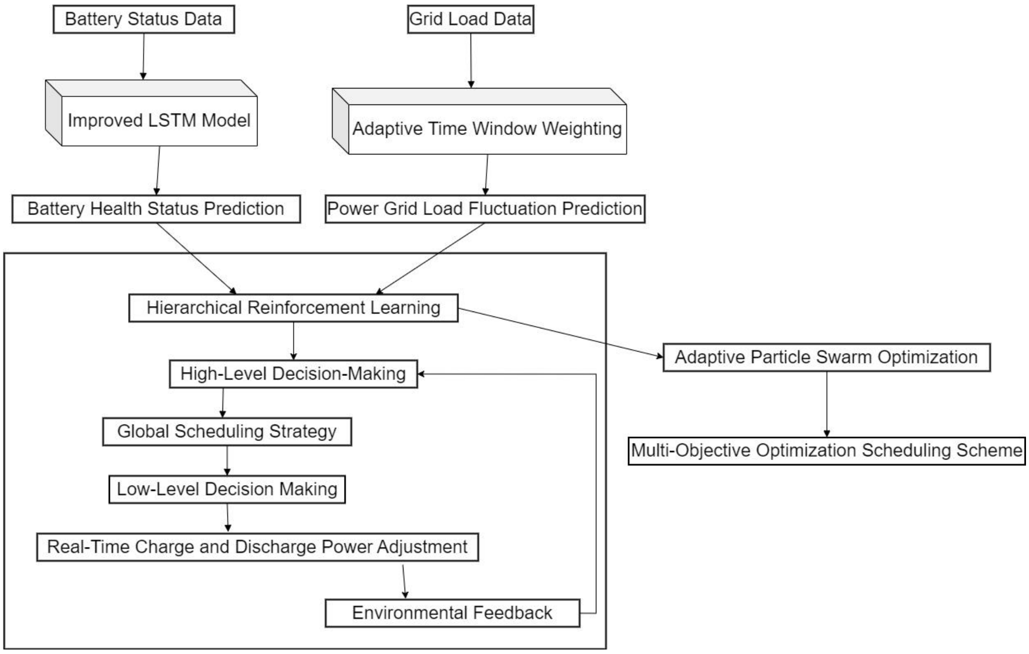

Figure 1. Structure of the method in this article.

2. Related Work

In large-capacity battery charging and discharging control, many studies are devoted to optimizing the battery charging and discharging process [11,12]. Traditional battery management methods do not consider the dynamic characteristics of batteries and the complexity of grid load fluctuations. In response to this, some studies have begun to attempt to apply intelligent algorithms such as deep learning [13,14] and reinforcement learning [15,16] to battery management.

In terms of battery health prediction, researchers have widely applied the LSTM model in deep learning. Some studies use LSTM networks to predict the battery's state of health (SOH), remaining useful life (RUL), and charging and discharging efficiency, providing data support for scheduling decisions. Zhang L constructed a lithium-ion SOH prediction model based on LSTM and achieved an accurate prediction of battery available capacity and RUL [17]. Zhao J used LSTM and Gaussian process regression (GPR) models to achieve the prediction of SOH and RUL, ensuring the prediction accuracy and reliability while reducing the computational complexity [18]. In addition, there are also studies that attempt to combine LSTM with convolutional neural networks to further improve the prediction precision [19,20].

For the prediction of power grid load, many studies have modeled and predicted power grid load fluctuations based on time series analysis and machine learning methods. Luo J used a new support vector regression model for power load prediction. In particular, in the case of data integrity attacks, the weight function was used to reduce the impact of malicious data and achieve more accurate prediction results [21]. Yanmei J proposed the EnGAT-BiLSTM model, which combines graph neural networks and bidirectional LSTM (BiLSTM) technology and achieves short-term power load prediction through methods such as Box-Cox transformation, dynamic load knowledge graph, and graph attention mechanism, significantly improving the prediction accuracy and model robustness [22]. The LSTM model has become the mainstream method for load prediction due to its advantages in dealing with nonlinear and time-varying characteristics [23,24]. However, there are still some limitations in the responsiveness of traditional LSTM models in the face of short-term fluctuations. Some studies have proposed a weighted time window strategy to enhance the model's sensitivity to recent load fluctuations, thereby improving prediction precision.

In charging and discharging scheduling, deep reinforcement learning has been widely used [25,26], especially deep Q-learning [27,28] and HRL [29,30]. These technologies guide models to learn the optimal charging and discharging decision through reward mechanism and adapt to complex dynamic environment. However, they face problems such as slow convergence speed, which limit their application efficiency in large-scale battery systems.

To optimize the multi-objective problem of battery charging and discharging scheduling, the PSO algorithm is widely used because of its strong global search capability [31,32]. However, when faced with complex and multi-dimensional optimization problems, PSO is prone to local optimal solutions. To this end, some studies have proposed adaptive particle update strategies [33,34] and hybrid optimization algorithms [35,36] to improve search precision and global optimality.

In summary, although current research has made significant progress in battery health prediction, power grid load prediction, and charging and discharging scheduling optimization in current research, existing methods still face some challenges, such as insufficient dynamic response capability, local optimal problems, and balancing multi-objective optimization. Therefore, integrating intelligent algorithms, combined with the time-varying characteristics of battery health and grid load fluctuations, to propose a more efficient and intelligent charging and discharging control strategy remains an important direction for future research.

3. Implementation of Battery Charging and Discharging Control Strategies

A. *Battery Health Assessment and Prediction*

In battery health assessment and prediction, this paper uses improved LSTM to evaluate the battery's state of health in real time to ensure its optimal charging and discharging strategy under different working conditions. This process includes several key links, such as input feature selection, LSTM model design and training, and real-time prediction. Each link is critical in solving the dynamic characteristics and battery aging problems in battery health prediction.

Figure 2. Large-capacity batteries.

In the selection of input features, the battery current, voltage, temperature, discharging cycle, battery internal resistance, and other parameters are selected based on the impact of the battery operating state on health assessment [37]. Figure 2 presents several common large-capacity batteries.

The current data is recorded in amperes (A); the voltage is expressed in volts (V); the temperature is in degrees Celsius (°C); the internal resistance is in ohms (Ω). The battery's charging cycle, discharging cycle, and number of charging and discharging times are also included to fully reflect the battery's historical operation. These feature data are all standardized so that each feature has a consistent scale during the model training process to prevent some features from having too much or too little impact on the training process. The mean-variance standardization method is used in the standardization process to convert each feature to zero mean and unit variance:

$$\tilde{x}_t = \frac{x_t - \mu_x}{\sigma_x} \quad (1)$$

x_t represents the raw feature data at time step t. After standardization, the mean μ_x of the feature is 0, and the standard deviation σ_x is 1.

In the model design phase, this paper uses a BiLSTM structure. The BiLSTM network can simultaneously capture the positive and negative dependencies of sequence data, especially in the battery state prediction process, it can more comprehensively consider the long-term and short-term dependencies of battery operation data. In battery health assessment, this structure works especially well for modeling time series data. In the LSTM network, to effectively capture the dependency of the battery's state of health at different time scales, a multi-scale feature fusion strategy is adopted, and the current and historical state data are taken as input. The attention mechanism is applied between the input layer and each LSTM layer to further enhance the model's learning ability. To solve the problems of gradient vanishing and gradient exploding in training of long sequence data, the residual connection mechanism is integrated to ensure the efficient transmission of information in the deep structure and improve the training efficiency and stability. A dropout layer is set between the network layers to avoid overfitting, and its ratio is set to 0.2 to improve the generalization ability. In addition, the ReLU (Rectified Linear Unit) activation function is utilized, and its strong nonlinear expression ability improves the model's ability to handle complex battery states.

In the training phase, the battery data from the Xi'an Jiaotong University battery data set is used to train the LSTM model. 70% of the training data is used for training, 15% for validation, and 15% for testing. The network is optimized using the Adam optimizer during training. To enhance the training stability and convergence speed, the batch size is set to 64, and the learning rate is set to 0.001. The mean-square error (MSE) loss function is adopted, and its goal is to reduce the difference between the battery's state of health predicted by the model and the actual battery's state of health. The parameters of the network are optimized through the backpropagation algorithm during the training process. 200 epochs are set in the training, and the early stopping mechanism is adopted. When the loss value of the validation set does not decrease significantly for 20 consecutive epochs, the training stops to avoid overfitting. During training, the network parameters are optimized by the back propagation algorithm. 200 training epochs are set, and the early stopping mechanism is adopted. When the loss value of the validation set does not decrease significantly within 20 consecutive epochs, the training is stopped to prevent overfitting. Hyperparameter adjustment and optimization during the training process enable the model to effectively predict under different battery's states of health, with high precision.

In the real-time prediction process, the trained LSTM model accepts the real-time collected battery data and dynamically predicts the battery's state of health. The input data includes real-time parameters, such as battery current, voltage, and temperature. These data are input into the LSTM network for prediction after being standardized. A sliding window technique is utilized to process real-time data since the battery condition varies over time. The window size is set to 30 minutes each time. By continuously updating the data input, the LSTM network can precisely predict its state of health based on the latest battery data. Each time window includes historical data on the battery. After the BiLSTM network processes these data, a prediction of the battery's current state of health is obtained.

This study applies an incremental learning strategy to improve real-time prediction precision. During battery operation, in the face of continuously changing data, incremental learning allows the model to dynamically update weights to maintain prediction accuracy. The specific method is to regularly input newly collected battery data into the pre-trained LSTM model and fine-tune the model parameters to ensure that it can adapt to the effects of battery aging and environmental changes.

B. Power Grid Load Prediction and Scheduling Response

In the power grid load prediction and scheduling response, an adaptive time window weighting strategy is applied for time series prediction. This strategy dynamically adjusts the data window size according to the power grid load fluctuation characteristics to more accurately capture short-term load changes. By assigning weighting coefficients to data points in each time window, data points close to the current moment are given higher weights, enhancing the model's ability to respond to short-term load fluctuations. Formula (2) is a dynamic adjustment formula for window length:

$$w_t = \beta \cdot \frac{\sigma_t}{\overline{\sigma}} \quad (2)$$

w_t represents the adaptive window length at time step t. $\overline{\sigma}$ represents the average standard deviation of the historical load data. β represents the window adjustment coefficient.

In the preprocessing phase of power grid load data, outliers are removed, and missing values are filled to ensure the input data's integrity and accuracy. In the adaptive time window weighting strategy, the window size dynamically changes according to the load fluctuation situation. When the load fluctuates greatly, the window length is shortened, and the weight coefficients are adjusted based on the load's temporal characteristics. The data points within the window are weighted in chronological order, with higher weights assigned to data points closer to the current time.

The BiLSTM structure is used in load prediction, which enables the model to learn the load data dependencies from both forward and backward directions, thereby capturing the potential trends in load fluctuations better.

During the model training phase, MSE is adopted as the loss function. The Adam optimizer is utilized during training. The learning rate is set to 0.001. The batch size is 64. The training cycle is set to 200 epochs. The early stopping mechanism is adopted to avoid overfitting. If the validation set error no longer decreases during training, the model automatically stops training.

In the load prediction phase, the adaptive time window weighting strategy is used to dynamically adjust the data window to ensure that the data in the window can precisely reflect the characteristics of the current grid load fluctuations during each prediction. The load prediction results at each time step are input into the scheduling system as the basis for charging and discharging scheduling. The scheduling system adjusts battery charging and discharging strategies based on the battery health prediction results and load prediction information. The objective function comprehensively considers load stability, battery life, and charging and discharging efficiency. Formula (3) shows the optimization objective function form:

$$J = \sum_{t=1}^{n} \left(\alpha_1 \cdot \left| P_t - \widehat{P}_t \right| + \alpha_2 \cdot \mathrm{SOH}_t + \alpha_3 \cdot \mathrm{SOC}_t \right) \quad (3)$$

α_1, α_2, and α_3 are the weight coefficients of load prediction error, battery's state of health, and state of charge, respectively. SOH_t represents the state of remaining battery life. SOC_t represents the battery's state of charge. By analyzing the importance of grid load stability, battery health status and charge status, the weights are initially set to 0.5, 0.3 and 0.2 respectively. These weights will be dynamically adjusted according to factors such as battery charging and discharging status, health status and grid load fluctuations to adapt to different operating conditions and ensure balance during the optimization process.

C. HRL to Optimize Charging and Discharging Scheduling

In optimizing the battery charging and discharging scheduling, the HRL method is used to achieve a balance between battery health management and grid load fluctuations. To this end, the scheduling task is divided into high-level and low-level decision-making. The high-level decision-making starts from a global perspective and formulates strategies based on the battery's state of health and load prediction data, aiming to maximize battery life and keep the grid load stable. Its output is the global scheduling strategy. The low-level decision-making adjusts the charging and discharging strategy based on the real-time battery state and short-term fluctuations in the grid load. The input includes real-time battery state and grid load change data, and the output is a specific charging and discharging power adjustment scheme. The global direction provided by the high-level decision-making guides the low-level decision-making that performs specific scheduling tasks based on real-time data. Figure 3 presents the HRL battery charging and discharging scheduling optimization framework. The high-level decision-making generates a global scheduling strategy every hour and passes the strategy to the low-level decision-making module through the parameter sharing mechanism. The low-level decision-making collects battery SOC, SOH, temperature and grid load data in real time with a cycle of 5 minutes, and dynamically adjusts the charging and discharging power according to the current state. The execution results of the low-level decision are transmitted back to the high-level decision-making module through the feedback channel for the next round of strategy update.

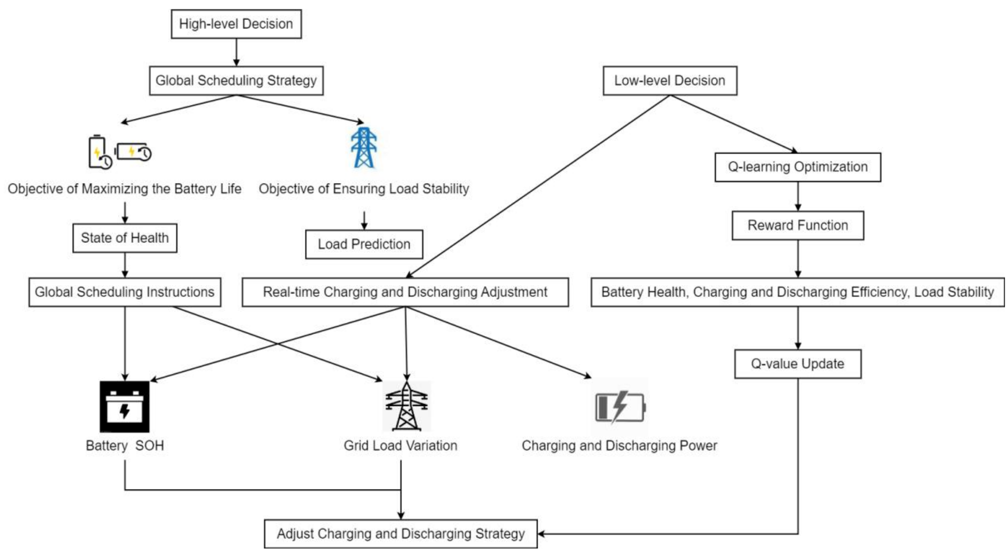

Figure 3. HRL battery charging and discharging scheduling optimization framework.

A global objective function is designed for high-level decision making. The objective is to maximize the battery life and ensure the stability of the load during battery operation. Formula (4) is the objective function:

$$\mathcal{J}_{\text{global}}(s_t, a_t) = \alpha_1 \cdot \mathcal{L}_{\text{battery}}(s_t, a_t) - \alpha_2 \cdot \mathcal{V}_{\text{load}}(s_t, a_t) \quad (4)$$

$\mathcal{L}_{\text{battery}}(s_t, a_t)$ represents the optimization function of battery life. $\mathcal{V}_{\text{load}}(s_t, a_t)$ is the variance of grid load fluctuation. α_1 and α_2 are weight coefficients for adjusting battery life and load stability.

High-level decision-making converts the global objective into a series of instructions that need to be executed by the lower level. The input of the high-level decision-making system includes the battery's state of health and the load prediction data. The output is the global scheduling strategy. The high-level decision-making process includes selecting the most suitable charging and discharging scheduling plan based on the battery's long-term state of health and the prediction information of load fluctuations.

Low-level decision-making focuses on the battery's real-time state and the power grid's short-term fluctuations. When the real-time load fluctuates or the battery state changes, the low-level decision-making is responsible for adjusting the charging and discharging power to adapt to the current load demand of the power grid and the actual charging and discharging capacity of the battery. The input of low-level decision-making is the state variables of the battery, such as state of charge (SOC), SOH, current, voltage, temperature, and data on grid load changes. The output of low-level decision-making is the charging and discharging power adjustment scheme for the current state.

The Q-learning algorithm is utilized to optimize the system's decision-making process. Q-learning evaluates the long-term rewards of taking different actions in each state by constructing a Q-value table. In this problem, the state space of Q-learning includes information such as the battery's SOC, SOH, and grid load fluctuations. The action space is the adjustment of the charging and discharging power. The reward function comprehensively considers multiple factors such as battery health, battery charging and discharging efficiency, and the stability of grid load fluctuations. Formula (5) shows the reward function:

$$R\left(s_t, a_t\right) = -\left(\beta_1 \cdot \Delta\mathrm{SOC}_t^2 + \beta_2 \cdot \Delta P_t^2 + \beta_3 \cdot \Delta\mathrm{SOH}_t\right) \quad (5)$$

$\Delta\mathrm{SOC}_t$ is the deviation of the state of charge. ΔP_t is the deviation of the power. $\Delta\mathrm{SOH}_t$ is the deviation between the current state of health and the target state of health. β_1, β_2, and β_3 are weight coefficients used to balance different factors' impacts on returns.

The training process of the Q-learning algorithm is based on multiple real-time parameters, such as the battery's state of health and load fluctuations. It is updated through a reward mechanism. Each feedback obtained from the environment (battery state and grid load fluctuations) updates the Q-value function to guide the agent in selecting the optimal charging and discharging strategy. During the Q-learning process, the system makes decisions based on the current battery SOC and SOH and future load changes and gradually learns an optimal charging and discharging scheduling strategy.

The Q-learning algorithm adopts the ε-greedy strategy to balance exploration and exploitation. This strategy adjusts the value of ε during the training process and allows a certain proportion of random exploration in the early stage to discover the best possible strategy. As the training progresses, ε gradually decreases, and the system gradually tends to choose the action with the highest Q-value, thereby accelerating the convergence of the strategy and avoiding falling into the local optimal solution.

During training, multiple epochs of training based on a simulation environment are used. The battery and load fluctuations of the power grid are used as input to simulate the charging and discharging process under different battery conditions. Through repeated training, the system can continuously adjust the Q-value so that low-level decisions can make more precise charging and discharging adjustments based on real-time battery state and load fluctuations.

In the scheduling system, high-level decisions are responsible for global optimization, with the main objective of maximizing battery life and grid stability. Low-level decisions are adjusted in real time based on the actual battery state and load fluctuations. Low-level adjustments ensure that the battery is not over-charged or over-discharged and reduce battery aging by controlling the charging and discharging power. The system also combines the battery health management model to evaluate the remaining battery life in real time and dynamically adjust charging and discharging strategies based on this information.

In addition, when training the Q-learning algorithm, the battery health data and the historical data of grid load fluctuations are used, including the historical battery charging and discharging data, load fluctuation data, etc. After each epoch of training, the system evaluates the performance of the current strategy and feeds the evaluation results back to the Q-learning algorithm as a reward signal. Q-learning optimizes charging and discharging decisions through reward signals, allowing the system to gradually learn the optimal charging and discharging strategy.

During the reinforcement learning process, low-level decisions can also make corresponding adjustments to the battery's short-term state of health. For instance, when the battery enters the aging state, the low-level decision adjusts the charging and discharging power according to the battery's health information to avoid over-charging or over-discharging and protect the long-term health of the battery. When the load fluctuates violently, the low-level decision responds to the changes in the power grid in real time and adjusts the charging and discharging power to maintain the stability of the power grid and avoid instability of the power grid due to excessive load fluctuations.

D. Multi-objective Optimization Scheduling Scheme

The adaptive PSO method is adopted in the multi-objective optimization scheduling scheme to optimize the battery charging and discharging scheduling. The objectives are to improve the battery charging and discharging efficiency, extend the battery life, and stabilize the grid load. The grid stability is improved, and the battery system operates efficiently by thoroughly optimizing these three objectives.

The implementation of the PSO algorithm starts with the initialization of particles. Each particle represents a potential charging and discharging scheduling strategy. The initial position and velocity of the particle are set according to the battery's state of health, load demand, and grid fluctuations. The initial particle position $\mathbf{x}_i(0)$ and velocity $\mathbf{v}_i(0)$ are defined as shown in Formulas (6) and (7), respectively:

$$\mathbf{x}_i(0) = \left\{P_{\text{charge},i}, P_{\text{discharge},i}, T_i, \mathrm{SOC}_i\right\} \quad (6)$$

$$\mathbf{v}_i(0) = \left\{\Delta P_{\text{charge},i}, \Delta P_{\text{discharge},i}, \Delta T_i, \Delta\mathrm{SOC}_i\right\} \quad (7)$$

$P_{\text{charge},i}$ and $P_{\text{discharge},i}$ are the initial charging and discharging powers of particle i, respectively. T_i is the battery temperature. SOC_i is the battery's state of charge. $\Delta P_{\text{charge},i}$, $\Delta P_{\text{discharge},i}$, ΔT_i, and ΔSOC_i are the changes in the charging power, discharging power, temperature, and state of charge at that moment.

The particle swarm can search in the entire scheduling space, and the solution space covers various parameters, such as battery charging and discharging power, battery temperature, charging and discharging cycle, etc. The perturbation method is used when initializing particles to prevent the particle swarm from falling into the local optimal solution, thereby increasing the diversity and coverage of the search.

The fitness function is utilized to evaluate each particle's quality during optimization. The energy loss is calculated to evaluate the battery's charging and discharging efficiency. The energy conversion efficiency during charging and discharging directly affects the battery's performance. The objective of extending the battery life depends on the depth of charging and discharging, the charging frequency, and the charging and discharging rate. The objective of grid load stability is to evaluate the smoothness of grid load fluctuations, avoiding excessive load fluctuations on the grid during battery charging and discharging. Considering these three objectives, the fitness function uses a weighted strategy to perform a weighted summation of the objectives:

$$f_i(\mathbf{x}_i) = w_1 \cdot f_{\text{efficiency}}(\mathbf{x}_i) + w_2 \cdot f_{\text{lifetime}}(\mathbf{x}_i) + w_3 \cdot f_{\text{stability}}(\mathbf{x}_i) \quad (8)$$

$f_{\text{efficiency}}(\mathbf{x}_i)$ evaluates the battery charging and discharging efficiency. $w_2 \cdot f_{\text{lifetime}}(\mathbf{x}_i)$ evaluates the battery life extension. $f_{\text{stability}}(\mathbf{x}_i)$ evaluates the grid load stability. w_1, w_2, and w_3 control the importance of each objective in the fitness function, and the initial values are 0.5, 0.3, and 0.2, respectively.

The particle update mechanism is based on the adjustment of velocity and position. After each iteration, the particle velocity and position are updated by comparing the historical best position with the global best position. Inertia weights can give particles a strong local search capability. The adjustment of the acceleration factor enhances the particle's ability to explore the global optimal solution. As the iteration proceeds, the inertia weight gradually decreases, ensuring that the particles can perform more refined local searches when approaching the global optimal solution. The particle velocity update formula adjusts the particle search direction according to the gap between the current particle position and the target optimal solution so that the particle swarm gradually converges to the optimal solution.

In multiple iterations, the particle swarm gradually updates pbest and gbest by continuously comparing each particle's fitness with the historical best fitness. When the fitness of particles meets the predetermined convergence criteria or reaches the maximum number of iterations, the optimization process ends, and the optimal charging and discharging scheduling method is the final optimal solution.

To solve the problem of classic PSO algorithm easily getting stuck in local optimal solutions, this paper adopts adaptive particle update technology, dynamically modifies the particle update method, and improves the update amplitude of particle velocity and position. In the case where the fitness value of particle swarm does not change significantly and no better solution can be found after multiple repetitions, the search ability of particles is improved, and they are guided to leave the local optimal solution they are currently in and enter different solution spaces for further search.

The objective weights in the PSO algorithm are also dynamically adjusted. Under different operating conditions, the weight ratio of the optimization objective is adjusted based on factors such as battery charging and discharging conditions, battery health conditions, and grid load fluctuations. When the battery health is poor, the priority is to optimize the extension of battery life. When the grid load fluctuates significantly, the weight of load stability is increased. When the battery charging and discharging efficiency is low, the priority is to optimize the charging and discharging efficiency. The dynamic adjustment of objective weights is expressed as Formulas (9) to (11):

$$w_1 = \frac{\eta_{\text{efficiency}}}{\eta_{\text{efficiency}} + \eta_{\text{lifetime}} + \eta_{\text{stability}}} \quad (9)$$

$$w_2 = \frac{\eta_{\text{lifetime}}}{\eta_{\text{efficiency}} + \eta_{\text{lifetime}} + \eta_{\text{stability}}} \quad (10)$$

$$w_3 = \frac{\eta_{\text{stability}}}{\eta_{\text{efficiency}} + \eta_{\text{lifetime}} + \eta_{\text{stability}}} \quad (11)$$

$\eta_{\text{efficiency}}$, η_{lifetime}, and $\eta_{\text{stability}}$ represent the importance of charging and discharging efficiency, battery life, and load stability, respectively.

During the entire optimization process, the balance optimization of multiple objectives is achieved through the design of a weighted fitness function. Each particle's fitness not only considers the charging and discharging efficiency, battery life, and grid load stability but also ensures that these three objectives can achieve optimal balance during the scheduling process. The battery efficiency and battery life objectives are usually negatively correlated, so these two objectives need to be carefully weighed during the optimization process. The load stability objective needs to be considered together with the battery efficiency and life objectives to avoid excessively reducing battery efficiency while pursuing load smoothing.

4. Real-time Control Strategy Effectiveness Evaluation

A. Experimental Setup

This experiment uses a simulation environment based on Python and OpenDSS (Open Distributed System Simulator) platform, combined with a battery model, grid load fluctuation data, and charging and discharging control strategy module to simulate the battery's charging and discharging process under various operating conditions and load fluctuation conditions. The battery system consists of 5 groups of lithium iron phosphate batteries with a rated capacity of 50kWh (nominal voltage 48V, maximum charge and discharge current 100A), which are connected to the grid through a bidirectional DC-AC converter (model: ABB ACS800-67). The load parameters are generated based on the typical daily load data of the IEEE 123 node distribution system, with the peak load set to 150kW and the valley load set to 60kW. The experiment lasts for 6 months, covering data collection and load change records at different battery operation phases. The battery's long-term charging and discharging process under real working conditions is simulated, and the control system's adaptability under different grid load fluctuation conditions is evaluated. The experiment records key feature information, such as the state of health, battery internal resistance, temperature changes, and number of charging and discharging cycles, of 10 groups of large-capacity lithium-ion batteries, combined with the grid load fluctuation data with a 15-minute sampling frequency to perform dynamic load prediction and charging and discharging control. During the experiment, the battery's charging and discharging strategies are tested under different operating temperatures (15°C, 25°C, and 35°C) and different load fluctuation intensities (peak-to-valley differences of 10%, 20%, and 30%) to fully simulate the battery's diverse operating states under actual working conditions. In addition, two special scenarios are set up to verify the adaptability of the control system under extreme working conditions: high-load impact (load fluctuation range>40%) and battery aging state (SOH<70%), to examine the robustness of the system under conditions of battery performance degradation and large load fluctuations.

The comparative experiment comprehensively compares the method in this paper (abbreviated as HRL) and the improved deep deterministic policy gradient (DDPG) algorithm and LSTM with adaptive time window weighting strategy to evaluate the differences in the multi-dimensional performance of different methods. During the experiment, 5 rounds of experiments are conducted under the same load conditions using different charging and discharging strategies. Each group of experiments runs 100 cycles, and the charging and discharging effects and energy consumption characteristics of each group of methods at different phases are recorded.

B. Evaluation Indicators

Experimental results are tested for significance through variance analysis, and the adaptability and efficiency of the method in this paper in a complex power grid environment are fully verified through multi-dimensional statistical analysis of the charging and discharging results of different strategies. The specific indicators include:

1) Charging and Discharging Efficiency

This indicator is used to measure the energy conversion efficiency during charging and discharging. The ratio of the energy a battery receives from a charging device to the energy actually stored in it is called the charging efficiency. The ratio of the energy actually released by a battery to its total energy output is called the discharging efficiency. High charging and discharging efficiency can improve the system's sustainability and economy while reducing the energy loss of the battery.

2) Charging and Discharging Energy Consumption Optimization rate

The charging and discharging energy consumption optimization rate is used to measure the degree of energy consumption reduction of different strategies during charging and discharging. By comparing the energy loss of the battery during charging and discharging with different methods, the proportion of optimized energy consumption to total energy consumption is calculated. This indicator reflects the effectiveness of the control system in optimizing battery energy consumption.

3) Battery Health Index

The battery's state of health is a significant indicator for measuring battery health. During the experiment, the battery health index is used to evaluate the impact of different control methods on battery life and to analyze whether the control strategy can effectively delay the battery aging process. By real-time monitoring of factors such as battery internal resistance and voltage changes, the protective effect of the charging and discharging strategy on battery health is evaluated.

4) Response Time

Response time refers to the time delay of the control measures taken by the system according to the fluctuation of grid load or the change of battery state. A fast response can ensure that the battery charging and discharging strategy can be adjusted quickly in the real-time grid fluctuation, thereby avoiding over-charging and over-discharging, ensuring the battery safety and the grid stability.

C. Application Effect

1) Battery Charging and Discharging Efficiency

The battery charging and discharging efficiency is evaluated by comparing the charging and discharging efficiency before and after optimization. The specific method is to first record the battery charging and discharging efficiency when no control strategy is applied and then measure the battery charging and discharging efficiency again after implementing the optimization strategy. By calculating the percentage increase in the optimized charging and discharging efficiency compared to the original efficiency, the optimization effect of the control strategy on the battery charging and discharging process is evaluated. Figure 4 displays the improvement in charging and discharging efficiency.

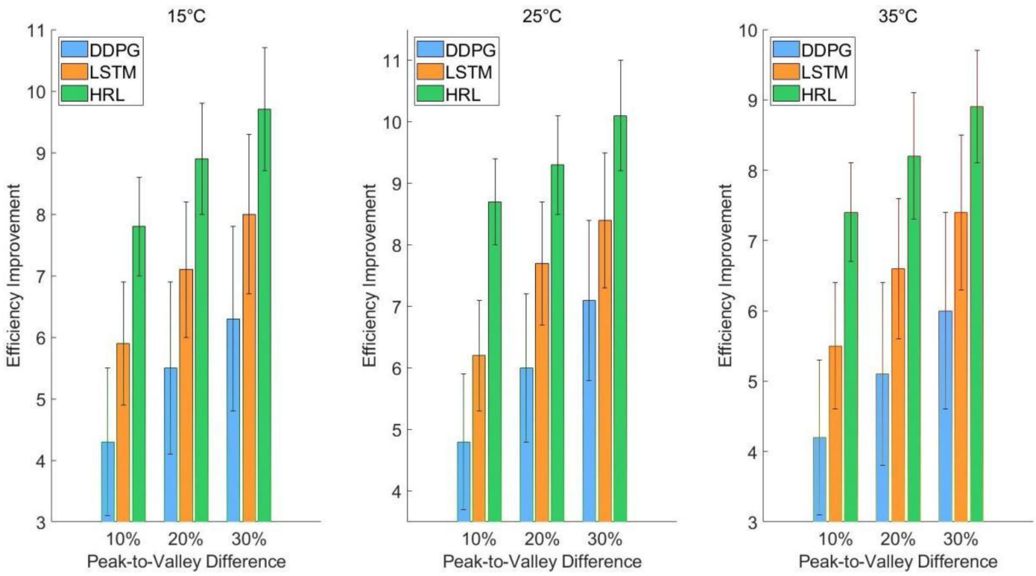

Figure 4. Improvement in charging and discharging efficiency of each method in different environments.

The HRL method achieves significantly higher charging and discharging efficiency than DDPG and LSTM in various experimental environments, and the improvement of HRL's charging and discharging efficiency increases accordingly as the load fluctuation intensity increases. This shows that the HRL method is adaptable to dealing with different load fluctuation intensities. When the load fluctuation is large, it can quickly respond to short-term load changes through hierarchical strategy optimization while ensuring the stability of long-term control.

When the peak-to-valley load difference is 30%, and the temperature is 15 degrees, the improvement in charging and discharging efficiency of HRL reaches 9.7%, which is significantly higher than that of other methods. This indicates that HRL can effectively cope with large fluctuations in load changes and reduce the impact of load fluctuations. In addition, in different temperature environments, HRL also demonstrates its strong stability and adaptability, with an improvement of up to 8.9% in the high-temperature environment. This phenomenon demonstrates that HRL can ensure the stability of battery performance when the ambient temperature changes. Compared with traditional LSTM and DDPG methods, HRL can avoid the negative impact of temperature changes on battery performance and continuously optimize the charging and discharging process in high-temperature environments through a more refined hierarchical learning strategy.

2) Battery State

Figure 5 presents the battery's SOH. The change in SOH shows which method can better protect battery health and delay battery aging.

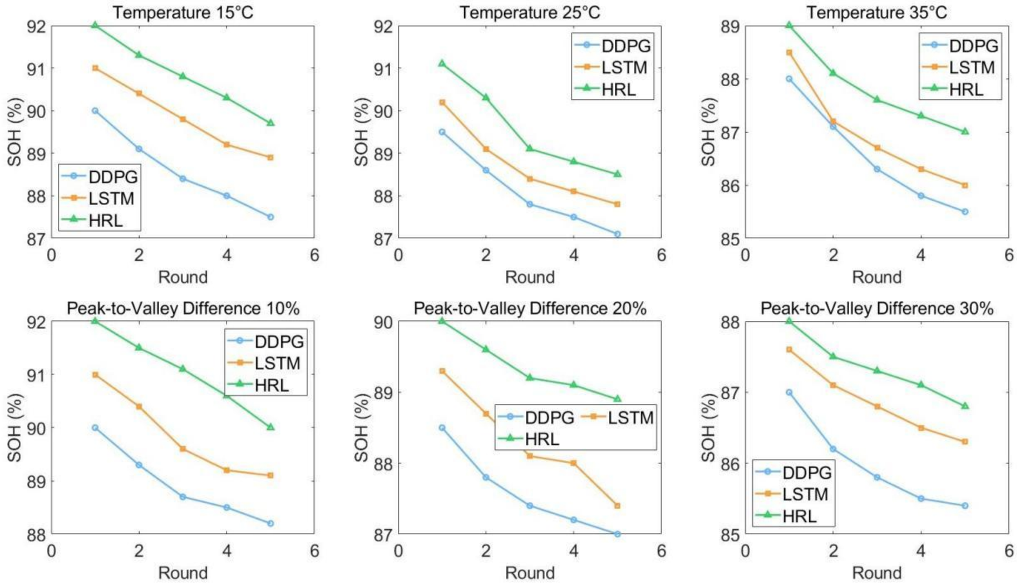

Figure 5. SOH changes of each method.

The HRL method can effectively maintain the SOH of the battery under various temperature conditions. At 25°C, the SOH in the fifth round is 88.50%, which is more stable than that of DDPG and LSTM. In high-temperature environments, the chemical reactions of batteries are accelerated, causing the battery to decay faster. HRL can reduce the negative impact of temperature fluctuations on battery performance by dynamically adjusting the charging and discharging rates and optimizing strategies.

The impact of load fluctuation intensity on battery health is also verified under different load fluctuation conditions. As the peak-to-valley load difference increases from 10% to 30%, HRL can maintain a high SOH in all rounds. When the peak-to-valley load difference is 30%, its SOH in the fifth round is 86.80%, which is still higher than that of DDPG and LSTM. The superior performance of HRL under high load and high temperature proves the adaptability of this method in dealing with extreme situations. Through continuous environmental feedback learning, HRL can dynamically optimize the working state of the battery, avoid over-charging, over-discharging, and over-heating of the battery, protect the battery health, and extend its service life.

3) Battery Charging and Discharging Energy Consumption Optimization Rate

The charging and discharging energy consumption optimization rate evaluates the effectiveness of the control method by comparing the change in battery system energy consumption before and after optimization. The energy consumption optimization percentage is calculated by comparing the battery's energy consumption during the charging and discharging process with its startup energy consumption. The optimization rate indirectly represents the control strategy's performance in improving system efficiency and reducing energy loss because it shows the extent to which battery energy consumption is reduced when various control techniques are applied. Figure 6 presents the results.

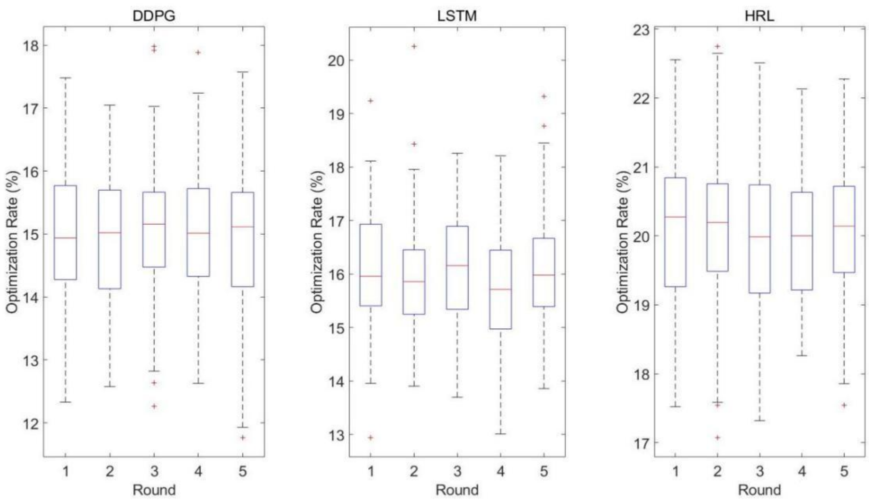

Figure 6. Optimization rate of charging and discharging energy consumption of each method.

The average optimization rate of the DDPG method is around 15%, with large fluctuations, indicating that DDPG has significant instability during multiple rounds of control, and it is difficult to continuously and stably improve the battery charging and discharging efficiency. The optimization rate of the LSTM method is slightly higher than that of DDPG, with an average of about 16% and smaller volatility, indicating that LSTM has a certain degree of stability in control compared with DDPG. The average optimization rate of the HRL method is about 20%. From the box range and the number of outliers in the box plot, the results of HRL show fewer outliers, indicating that the optimization strategy of HRL is more stable between different experimental rounds. In terms of stability, HRL is actually more volatile than LSTM, as evidenced by the wider box of HRL. This is related to the multi-level decision-making structure of HRL. Although it can optimize decisions at a higher level, when faced with certain complex or extreme situations, the synergy between levels may not be stable enough, resulting in significant optimization fluctuations. LSTM performs relatively stably in this regard, with a small fluctuation range, indicating that its time series-based prediction capability can provide more consistent performance during battery optimization.

4) Real-time Control Response Time

Real-time control response time is used as an indicator to evaluate the battery management system's response speed and control ability when facing load fluctuations, revealing the adaptability of different methods in dynamic environments. The data in Figure 7 shows that the HRL method has a significantly better response speed than the other two methods.

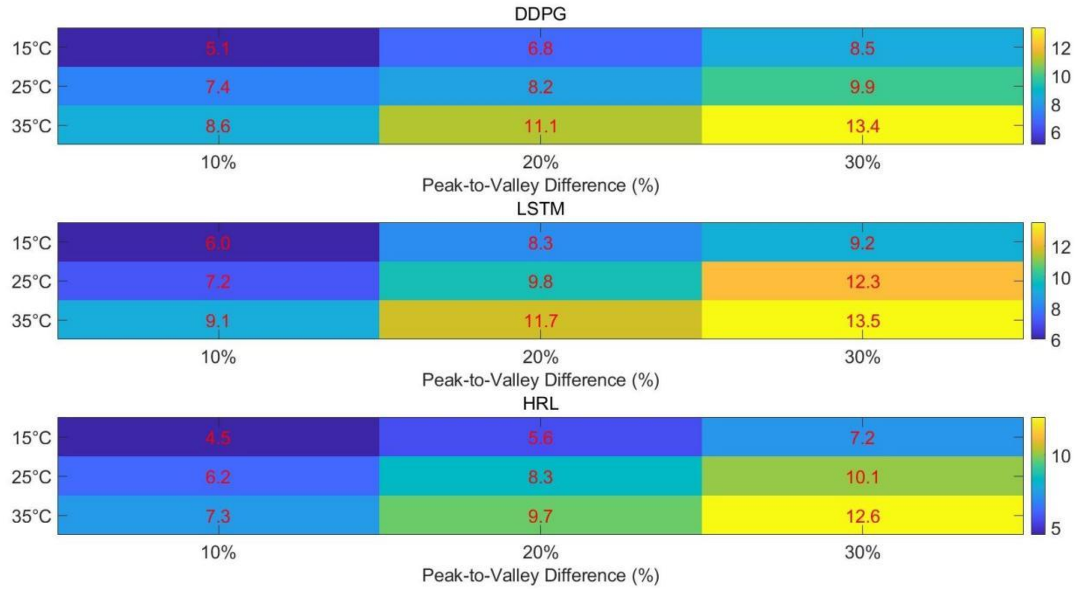

Figure 7. Response time of each method to implement control.

As the temperature increases and the load fluctuation increases, the response time of all methods increases, but HRL still maintains a relatively short response time. At 15°C and a 10% peak-to-valley load difference, HRL's response time is 4.5 seconds, which is lower than 5.1 seconds of DDPG and 6 seconds of LSTM. At 35°C and a 30% peak-to-valley load difference, HRL's response time is still within 12.6 seconds. In contrast, DDPG and LSTM's response times are 13.4 seconds and 13.5 seconds, respectively. This phenomenon demonstrates that the HRL method can maintain a faster system response capability through more efficient strategy optimization when facing temperature fluctuations and load changes.

This performance is attributed to the reinforcement learning mechanism of the HRL method, especially in the multi-level and hierarchical strategy optimization, which can quickly adapt to the system's real-time state changes. Compared with DDPG and LSTM, HRL effectively reduces the dependence on global information in the adjustment process through a hierarchical decision-making structure, thereby having a faster response time in real-time control. This advantage also indicates that when facing more complex battery management tasks, the HRL method can handle more environmental variables while ensuring a low response time, thereby achieving efficient control of the battery system.

5) Adaptability in Extreme Situations

Under the two extreme conditions of high-load impact and battery aging, the key indicators of different strategies, such as the improvement rate of charging and discharging efficiency and the optimization rate of charging and discharging energy consumption, are compared and analyzed. Table 1 lists the results.

Table 1. Adaptability analysis for extreme situations.

Scenario	Method	Battery charging and discharging efficiency improvement rate	SOH	Charging and discharging energy consumption optimization rate	Real-time control response time
High-load impact	DDPG	5.20%	84.30%	4.50%	10.3s
	LSTM	6.10%	85.70%	6.20%	9.8s
	HRL	7.80%	87.10%	7.50%	8.2s
Battery aging	DDPG	3.40%	61.60%	3.00%	12.1s
	LSTM	4.00%	63.20%	4.40%	11.6s
	HRL	5.30%	67.50%	5.70%	10.2s

Under high-load impact, HRL's charging and discharging efficiency improvement rate reaches 7.80%, and the energy consumption optimization rate is 7.50%, which are better than LSTM and DDPG. Regarding real-time control response time, HRL only needs 8.2s. In contrast, LSTM needs 9.8s, and DDPG needs 10.3s. Under battery aging, HRL's SOH drops to 67.5%; the charging and discharging efficiency improvement rate reaches 5.3%; the energy consumption optimization rate reaches 5.70%. At the same time, the real-time control response time of HRL is 10.2s. In summary, the excellent performance of the HRL method under extreme conditions is due to its multi-level strategy optimization, dynamic model update, and efficient modeling of complex battery states. HRL can significantly improve the charging and discharging efficiency, optimize energy consumption, and shorten the control response time through precise state perception and adaptive control mechanisms.

6) Significance Analysis

The p-values in Table 2 are used to evaluate the significant differences between the HRL method in this paper and the two comparative methods under different scenarios and conditions and to verify whether the optimization effect of HRL on charging and discharging efficiency, SOH, energy consumption optimization rate, and real-time control response time is statistically significant.

Table 2. Significance analysis results.

Scenario	Method	p-value			
		Battery charging and discharging efficiency improvement rate	SOH	Charging and discharging energy consumption optimization rate	Real-time control response time
High-load impact	DDPG vs HRL	0.022	0.031	0.017	0.024
	LSTM vs HRL	0.027	0.035	0.019	0.018
Battery aging	DDPG vs HRL	0.024	0.021	0.013	0.022
	LSTM vs HRL	0.029	0.039	0.016	0.014
15°C	DDPG vs HRL	0.026	0.034	0.022	0.027
	LSTM vs HRL	0.023	0.037	0.019	0.021
25°C	DDPG vs HRL	0.029	0.031	0.024	0.019
	LSTM vs HRL	0.028	0.032	0.026	0.022
35°C	DDPG vs HRL	0.023	0.028	0.015	0.026
	LSTM vs HRL	0.025	0.036	0.022	0.019
Peak-to-valley load difference 10%	DDPG vs HRL	0.019	0.029	0.021	0.023
	LSTM vs HRL	0.021	0.027	0.018	0.02
Peak-to-valley load difference 20%	DDPG vs HRL	0.027	0.031	0.024	0.022
	LSTM vs HRL	0.022	0.025	0.017	0.021
Peak-to-valley load difference 30%	DDPG vs HRL	0.021	0.026	0.022	0.02
	LSTM vs HRL	0.023	0.031	0.019	0.022

Under high-load impact conditions, the HRL method shows significant advantages in improving charging and discharging efficiency, maintaining SOH, and optimizing energy consumption, with p-values all below 0.05, proving that the HRL method is more adaptable under extreme load conditions. Under battery aging conditions, the p-values of the HRL method also remain below 0.05, indicating that it can still maintain good charging and discharging control effects when the SOH drops below 70%. In addition, under different temperature conditions, the charging and discharging efficiency, energy consumption optimization rate, and response time of the HRL method and comparative methods show significant differences. The p-values at 25°C are low, indicating that the performance advantage of HRL under this temperature condition is more significant. Under different peak-to-valley load fluctuation conditions (10%, 20%, and 30%), the p-values of the HRL method are all lower than 0.05, reflecting that the changes in load fluctuation amplitude have little impact on the optimization effect of the HRL method, further verifying the robustness and control stability of HRL under large load fluctuations.

7) Load Forecasting and Calculation Time

This table systematically compares the comprehensive performance of the three methods of HRL, DDPG and LSTM in terms of grid load forecasting accuracy and computational efficiency, intuitively reflecting the technical advantages of the real-time control method for large-capacity batteries in smart grids proposed in this paper. From the prediction results, the predicted load value of HRL is 157.8kW, which is closest to the actual load, indicating that the prediction method based on the improved LSTM model and adaptive time window weighting strategy has a stronger response ability to short-term load fluctuations. In terms of computational efficiency, the average time consumption of HRL is 2.7s, which verifies that the hierarchical reinforcement learning framework effectively balances the model complexity and computational overhead through the parameter sharing mechanism.

Table 3. Comparison of load forecasting and computational time.

Method	Forecast Load (kW)	True load (kW)	Average calculation time (s)
HRL	157.8	160	2.7
DDPG	144.2	160	5.8
LSTM	143.9	160	4.2

5. Conclusion

This paper studies an optimization method based on intelligent algorithms to solve the real-time control problem of large-capacity battery charging and discharging in intelligent networks. The improved LSTM model is used to predict battery health, the adaptive time window weighting strategy is used to improve the grid load prediction precision, and the HRL, combined with the improved PSO algorithm, is applied to optimize charging and discharging scheduling. Experiments show that this method can significantly improve battery charging and discharging efficiency, extend battery life, and achieve fast real-time scheduling response. The method in this paper achieves a 9.7% improvement in charging and discharging efficiency and a 20% energy consumption optimization rate under the conditions of a 30% peak-to-valley load difference and a temperature of 15°C. It also maintains stable performance under extreme working conditions and shortens the response time to less than 8.2 seconds. Although this method has achieved certain results, there are still some shortcomings. For example, in some extreme cases, the stability of HRL still must be enhanced. The dynamic adjustment strategy of each objective weight in multi-objective optimization can be further optimized. In terms of battery health prediction, we can try to introduce a multi-physics field coupling model, combining the electrochemical reaction mechanism with an intelligent algorithm to improve the understanding of the internal aging mechanism of the battery and the prediction accuracy. For grid load prediction, we explore deep learning models that integrate multi-source heterogeneous data such as meteorological data and user behavior patterns to further improve the accuracy and generalization of predictions. In terms of charging and discharging scheduling optimization, we study the dynamic adjustment strategy of the weights of each objective in multi-objective optimization, develop more efficient optimization algorithms to enhance the real-time and adaptability of scheduling, and consider the integration and coordinated control of distributed energy resources to improve the stability of the entire power grid and energy efficiency.

Acknowledgment

None

Consent to Publish

The manuscript has neither been previously published nor is under consideration by any other journal. The authors have all approved the content of the paper.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Funding

None

Author Contribution

Ming Lu, Yupeng Cai: Developed and planned the study, performed experiments, and interpreted results. Edited and refined the manuscript with a focus on critical intellectual contributions.

Zhiyu Liu, Qingyang Tian: Participated in collecting, assessing, and interpreting the date.

Meng Wu, Xiangdong Gong: Made significant contributions to date interpretation and manuscript preparation.

Ming Lu, Yupeng Cai: Provided substantial intellectual input during the drafting and revision of the manuscript.

Conflicts of Interest

The authors declare that they have no financial conflicts of interest.

References

[1] P. Parvizi, M. Jalilian, A.M. Amidi, M.R. Zangeneh, J.R. Riba. From Present Innovations to Future Potential: The Promising Journey of Lithium-Ion Batteries. Micromachines, 2025, 16(2), 194-270. DOI: 10.3390/mi16020194

[2] A. Mousaei. Analyzing locational inequalities in the placement of electric vehicle charging stations using machine learning: A case study in Glasgow. Next Research, 2025, 2(1), 100123-100137. DOI: 10.1016/j.nexres.2024.100123

[3] R.R. Kumar, C. Bharatiraja, K. Udhayakumar, S. Devakirubakaran, K.S. Sekar, L. Mihet-Popa. Advances in batteries, battery modeling, battery management system, battery thermal management, SOC, SOH, and charge/discharge characteristics in EV applications. Ieee Access, 2023, 11, 105761-105809. DOI: 10.1109/ACCESS.2023.3318121

[4] N. Ghaeminezhad, M. Monfared. Charging control strategies for lithium-ion battery packs: Review and recent developments. IET Power Electronics, 2022, 15(5), 349-367. DOI: 10.1049/pel2.12219

[5] S. Aghajan-Eshkevari, S. Azad, M. Nazari-Heris, M.T. Ameli, S. Asadi. Charging and discharging of electric vehicles in power systems: An updated and detailed review of methods, control structures, objectives, and optimization methodologies. Sustainability, 2022, 14(4), 2137-2170. DOI: 10.3390/su14042137

[6] M. Adaikkappan, N. Sathiyamoorthy. Modeling, state of charge estimation, and charging of lithium-ion battery in electric vehicle: a review. International Journal of Energy Research, 2022, 46(3), 2141-2165. DOI: 10.1002/er.7339

[7] W. Strielkowski, A. Vlasov, K. Selivanov, K. Muraviev, V. Shakhnov. Prospects and challenges of the machine learning and data-driven methods for the predictive analysis of power systems: A review. Energies, 2023, 16(10), 4025-4056. DOI: 10.3390/en16104025

[8] T. Mazhar, H.M. Irfan, I. Haq, I. Ullah, M. Ashraf, T.A. Shloul, et al. Analysis of challenges and solutions of IoT in smart grids using AI and machine learning techniques: A review. Electronics, 2023, 12(1), 242-267. DOI: 10.3390/electronics12010242

[9] H. Habbak, M. Mahmoud, K. Metwally, M.M. Fouda, M.I. Ibrahem. Load forecasting techniques and their applications in smart grids. Energies, 2023, 16(3), 1480-1513. DOI: 10.3390/en16031480

[10] Z. Masood, R. Gantassi, Ardiansyah, Y. Choi. A multi-step time-series clustering-based Seq2Seq LSTM learning for a single household electricity load forecasting. Energies, 2022, 15(7), 2623-2634. DOI: 10.3390/en15072623

[11] M.U. Tahir, A. Sangwongwanich, D.I. Stroe, F. Blaabjerg. Overview of multi-stage charging strategies for Li-ion batteries. Journal of Energy Chemistry, 2023, 84, 228-241. DOI: 10.1016/j.jechem.2023.05.023

[12] M. Ahwiadi, W. Wang. Battery Health Monitoring and Remaining Useful Life Prediction Techniques: A Review of Technologies. Batteries, 2025, 11(1), 31-60. DOI: 10.3390/batteries11010031

[13] A. Mousaei, Y. Naderi, I.S. Bayram. Advancing state of charge management in electric vehicles with machine learning: A technological review. IEEE Access, 2024, 12, 43255-43283. DOI: 10.1109/ACCESS.2024.3378527

[14] H.P. Bhupathi. Deep Learning and EV Charging: Battery Life and Performance. ESP International Journal of Advancements in Science & Technology (ESP-IJAST), 2023, 1(1), 29-46. DOI: 10.56472/25839322/IJAST-V1I1P106

[15] S. Yalçın, M.S. Herdem. Optimizing EV Battery Management: Advanced Hybrid Reinforcement Learning Models for Efficient Charging and Discharging. Energies, 2024, 17(12), 2883-2904. DOI: 10.3390/en17122883

[16] N. Kaewdornhan, C. Srithapon, R. Liemthong, R. Chatthaworn. Real-time multi-home energy management with EV charging scheduling using multi-agent deep reinforcement learning optimization. Energies, 2023, 16(5), 2357-2382. DOI: 10.3390/en16052357

[17] L.J. Zhang, T. Ji, S.H. Yu, G.C. Liu. Accurate prediction approach of SOH for lithium-ion batteries based on LSTM method. Batteries, 2023, 9(3), 177-195. DOI: 10.3390/batteries9030177

[18] J.H. Zhao, Y. Zhu, B. Zhang, M.Y. Liu, J.X. Wang, C.H. Liu, et al. Method of Predicting SOH and RUL of Lithium-Ion Battery Based on the Combination of LSTM and GPR. Sustainability, 2022, 14(19), 11865-11881. DOI: 10.3390/su141911865

[19] D.X. Gao, X. Liu, Z.Y. Zhu, Q. Yang. A hybrid CNN-BiLSTM approach for remaining useful life prediction of EVs lithium-Ion battery. Measurement and Control, 2023, 56(1-2), 371-383. DOI: 10.1177/00202940221103622

[20] D.W. Chen, X.Y. Zheng, C.Y. Chen, W.D. Zhao. Remaining useful life prediction of the lithium-ion battery based on CNN-LSTM fusion model and grey relational analysis. Electronic Research Archive, 2023, 31(2), 633-655. DOI: 10.3934/era.2023031

[21] J. Luo, T. Hong, Z.M. Gao, S.C. Fang. A robust support vector regression model for electric load forecasting. International Journal of Forecasting, 2023, 39(2), 1005-1020. DOI: 10.1016/j.ijforecast.2022.04.001

[22] Y.M. Jiang, M.S. Liu, Y.Y. Li, Y.P. Liu, J.Y. Zhang, Y.F. Liu, et al. Enhanced neighborhood node graph neural networks for load forecasting in smart grid. International Journal of Machine Learning and Cybernetics, 2024, 15(1), 129-148. DOI: 10.1007/s13042-023-01796-8

[23] H.H. Chen, M.Y. Zhu, X. Hu, J.R. Wang, Y. Sun, J.D. Yang. Research on short-term load forecasting of new-type power system based on GCN-LSTM considering multiple influencing factors. Energy Reports, 2023, 9, 1022-1031. DOI: 10.1016/j.egyr.2023.05.048

[24] T. Bashir, H.Y. Chen, M.F. Tahir, L.Q. Zhu. Short term electricity load forecasting using hybrid prophet-LSTM model optimized by BPNN. Energy Reports, 2022, 8, 1678-1686. DOI: 10.1016/j.egyr.2021.12.067

[25] B. Lami, M. Alsolami, A. Alferidi, S.B. Slama. A Smart Microgrid Platform Integrating AI and Deep Reinforcement Learning for Sustainable Energy Management. Energies, 2025, 18(5), 1157-1187. DOI: 10.3390/en18051157

[26] H. Pourvaziri, H. Sarhadi, N. Azad, H. Afshari, M. Taghavi. Planning of electric vehicle charging stations: An integrated deep learning and queueing theory approach. Transportation Research Part E: Logistics and Transportation Review, 2024, 186, 103568-103591. DOI: 10.1016/j.tre.2024.103568

[27] Y.X. Wen, P.X. Fan, J. Hu, S. Ke, F.Z. Wu, X. Zhu. An optimal scheduling strategy of a microgrid with V2G based on deep Q-learning. Sustainability, 2022, 14(16),

10351-10369. DOI: 10.3390/su141610351
[28] V.R. Chifu, T. Cioara, C.B. Pop, H.G. Rusu, I. Anghel. Deep q-learning-based smart scheduling of evs for demand response in smart grids. Applied Sciences, 2024, 14(4), 1421-1442. DOI: 10.3390/app14041421
[29] S. Mohanty, S. Panda, S.M. Parida, P.K. Rout, B.K. Sahu, M. Bajaj, et al. Demand side management of electric vehicles in smart grids: A survey on strategies, challenges, modeling, and optimization. Energy Reports, 2022, 8, 12466-12490. DOI: 10.1016/j.egyr.2022.09.023
[30] H.L. Zhang, J.K. Peng, H.X. Dong, F. Ding, H.C. Tan. Integrated velocity optimization and energy management strategy for hybrid electric vehicle platoon: A multiagent reinforcement learning approach. IEEE Transactions on Transportation Electrification, 2023, 10(2), 2547-2561. DOI: 10.1109/TTE.2023.3298365
[31] H.H. Bian, Z.Y. Guo, C.G. Zhou, X.M. Wang, S. Peng, X.A. Zhang. Research on orderly charge and discharge strategy of EV based on QPSO algorithm. IEEE Access, 2022, 10, 66430-66448. DOI: 10.1109/ACCESS.2022.3185236
[32] H.I. Shaheen, G.I. Rashed, B. Yang, J. Yang. Optimal electric vehicle charging and discharging scheduling using metaheuristic algorithms: V2G approach for cost reduction and grid support. Journal of Energy Storage, 2024, 90, 111816-111826. DOI: 10.1016/j.est.2024.111816
[33] I. Ahmed, M. Rehan, A. Basit, M. Tufail, K.S. Hong. A dynamic optimal scheduling strategy for multi-charging scenarios of plug-in-electric vehicles over a smart grid. Ieee Access, 2023, 11, 28992-29008. DOI: 10.1109/ACCESS.2023.3258859
[34] S.A. Khou, J. Olamaei, M.H. Hosseini. Strategic scheduling of the electric vehicle-based microgrids under the enhanced particle swarm optimization algorithm. Scientific Reports, 2024, 14(1), 30795-30819. DOI: 10.1038/s41598-024-81049-y
[35] V.M. Kumar, C. Bharatiraja, A. ELrashidi, K.M. AboRas. Chaotic harris hawks optimization algorithm for electric vehicles charge scheduling. Energy Reports, 2024, 11, 4379-4396. DOI: 10.1016/j.egyr.2024.04.006
[36] Q.W. Yang, Y.T. Huang, Q.Q. Zhang, J.J. Zhang. A bi-level optimization and scheduling strategy for charging stations considering battery degradation. Energies, 2023, 16(13), 5070-5085. DOI: 10.3390/en16135070
[37] T.H. Dar, S. Singh. Optimized parameter estimation of lithium-ion batteries using an improved cuckoo search algorithm under variable temperature profile. e-Prime-Advances in Electrical Engineering, Electronics and Energy, 2025, 11, 100902-100914. DOI: 10.1016/j.prime.2025.100902