https://doi.org/10.31449/inf.v46i16.9736 Informatica **49** (2025) 429–440 **429**

Integrating DDPG and QPSO for Multi-Objective Optimization in High Proportion Renewable Energy Power Dispatch Systems

Xu'an Qiao[1*], Chaofan Liu[2]
[1]School of Aeronautics, Chongqing City Vocational College, Chongqing 402160, China
[2]School of Physics Sciences, University of Science and Technology of China, Hefei 230026, China
E-mail: qiaoxuan109_2023@126.com, Liuchaofan1980_622@126.com
*Corresponding author

Keywords: Power system, dispatch optimization, renewable energy, DDPG, heuristic algorithm

Received: June 16, 2025

This study proposes a novel dispatch optimization model that integrates deep deterministic policy gradient (DDPG) and quantum particle swarm optimization (QPSO) to address the challenges posed by high proportions of renewable energy in power systems. The proposed multi-objective optimization framework considers system cost reduction, supply-demand balance, and dynamic adaptability to renewable energy fluctuations. The experimental results on the IEEE 30-bus and 118-bus systems demonstrated significant improvements. This method reduced total system costs by 13.6% and 11.4%, respectively. It also increased supply reliability to 97.1% and achieved an energy utilization rate of 94.85%. Additionally, it minimized frequency deviation to 1.25 Hz. The optimization time was also improved, achieving a reduction of 58.3 seconds in efficiency. The research results have important practical application value in improving power system economy, enhancing system reliability, and dynamic adaptability. It can provide efficient and reliable technical support for power dispatch planning, load management, and real-time control under high percentage renewable energy scenarios.

Povzetek:Študija predlaga hibridni model za optimizacijo razporejanja (dispečanja) v omrežjih z visokim deležem obnovljivih virov, ki združuje izboljšani DDPG in QPSO. Model zmanjša stroške, izboljša zanesljivost in stabilnost ter poveča izrabo energije. Preizkusi na IEEE sistemih potrjujejo visoko učinkovitost.

1 Introduction

With the continuous adjustment and optimization of the global energy structure, carbon peaking and carbon neutrality targets have become important strategies for countries around the world to cope with climate change and achieve sustainable development. Because of their clean and low-carbon benefits, renewable energy (RE) sources like solar and wind have been frequently used in this setting [1]. However, the operation and scheduling of the conventional power system (PS) have been severely hampered by the widespread use of wind, solar, and other high percentage (HP) RE sources. To begin with, RE is highly volatile and erratic. Their output is affected by natural conditions, including wind speed, light, etc., and there is a large uncertainty [2]. This uncertainty makes the load balance and stability of the system subject to shocks, and is prone to supply-demand imbalance in the PS in the case of peak power demand or insufficient wind and solar resources [3]. Second, the traditional PS, which relies on precise forecasts of load and generation capacity from the scheduling model, becomes irrelevant when the proportion of RE in the PS increases. Due to the significant impact of unstable factors on RE generation capacity, there is a substantial discrepancy between actual and forecasted values. This discrepancy makes short-term PS scheduling more challenging and complex [4-5]. In addition, current PS scheduling relies mainly on a phased sequential scheduling approach, i.e., unit commitment (UC) is performed first to determine the start/stop status of the units. Then economic dispatch (ED) is performed to optimize the unit output. Finally, real-time regulation is performed through automatic generation control (AGC) [6]. Although this staged dispatch model is simple to operate, there are problems such as response delays between different dispatch modules. In view of this, the study proposes a multilevel cooperative dispatch model for HP of renewable energy power systems (REPS) and introduces heuristic algorithms to accelerate the solution of complex systems. The study aims to solve the limitations of the traditional sequential scheduling method and improve the economy, reliability and feasibility of system operation through refined modeling and cooperative optimization.

This study's novel contribution is its proposal of a joint heuristic algorithm that combines improved deep deterministic policy gradient (DDPG) and quantum particle swarm optimization (QPSO) algorithms to optimize scheduling in high-proportion REPS. This method improves the convergence speed and stability of complex, multi-constraint, multi-timescale problems by introducing dual experience pooling and time-decaying exploration strategies. The proposed mathematical model addresses the inefficiencies and local optima of traditional models. It provides a more efficient and reliable solution for PS scheduling optimization.

2 Related works

The key to ensuring the PS operates steadily is PS schedule optimization. The optimization method directly impacts the stability and adaptability of the PS to RE fluctuations. These are essential to the functioning and advancement of contemporary PS, as well as its economic efficiency. Therefore, many scholars have carried out various researches on PS scheduling optimization. For the optimal reactive power scheduling problem in PS, M. Abd-El Wahab et al. suggested a hybrid method called augmented Jaya and artificial ecosystem-based optimization, which improved system stability, economic viability, and overall efficiency [7]. A nonconvex mixed integer and quadratic restricted planning technique was presented by Cox J L et al. to solve the challenge of optimizing a centralized solar power plant's profitability under changing solar resources. The method improved the solvability of the problem through exact and approximation techniques, thus enabling operational scheduling optimization in real-time decision support [8]. To address the issue of system security and economic cost over time in microgrid scheduling, Zhang et al. suggested a multi-timescale scheduling model that incorporated load voltage and frequency dynamics. To minimize economic cost while maintaining voltage and frequency stability, the study converted it into a multi-objective optimization problem that took into account economic cost, voltage deviation, and frequency stability. This improved the microgrid dispatch's efficiency and dependability. [9]. To address the global issues brought on by the recent explosive increase in the demand for electricity, Hou et al. developed an integrated day-ahead multi-objective microgrid optimization framework. The framework produced more affordable, dependable, and ecologically friendly power supply services by combining demand-side management, forecasting methods, and economic-environmental dispatch [10].

Large-scale access to the PS by RE sources affects the output characteristics of wind and photovoltaic energy sources. These sources exhibit strong intermittency, randomness, and volatility due to weather, climate, and other external natural factors. These challenges have led to the urgent need for innovation and optimization of existing dispatch methods to adapt to the new situation of HP of RE access. A mixed-integer linear programming method was put up by Shirzadi et al. to address the issue of enhancing the efficiency and dependability of RE systems. The study optimized the PS's daily operating expenses and system resilience by combining a unique hybrid model with deep learning and statistical modeling to forecast the load demand and wind power output (PO) for the ensuing three days [11]. Due to the impact of the volatility of wind and solar power generation on PS operation, Guo et al. proposed multi-stage optimization, online optimization, and multi-timescale optimization for RE integration. This study realized the strategic scheduling and control of energy storage units and improved the efficiency of RE integration in the power grid [12]. By proposing a new economic low-carbon clean PS dispatch model that incorporates power-to-gas technology, Cui et al. addressed the issue of increasing the grid's capacity to absorb wind power. This model integrated the effects of multiple price factors, resulting in low-carbon PS operation and cost optimization [13]. An enhanced jellyfish search optimization technique was presented by Gami et al. to solve the optimal reactive power dispatch problem in HP renewable PSs. By improving the algorithm's exploration and development stages, the study successfully optimized the PS's most secure and stable state [14]. This allowed the PS to operate in both deterministic and probabilistic load demands and RE resource states.

In summary, the existing research has made some significant progress in PS scheduling optimization. However, there are still deficiencies in the research for HP of RE access, such as insufficient consideration of the volatility and stochastic characteristics of RE, fewer studies on the collaborative scheduling of multi-timescale modules, and insufficiently perfect uncertainty handling methods. Therefore, the study proposes to construct a mathematical model of HP of REPS scheduling optimization and introduce a heuristic algorithm to solve it. The innovation of the study is to propose a multi-module cooperative optimization framework to accurately deal with uncertainty and extreme scenarios. Meanwhile, the optimization algorithm improves the solution efficiency and provides new ideas for complex PS scheduling.

3 Methods and materials

This section provides a detailed description of the PS scheduling optimization method proposed in the study. The method consists of scheduling optimization mathematical model and scheduling optimization heuristic algorithm. The combination of the two effectively improves the scheduling efficiency and stability of the PS with a HP of RE access.

3.1 Mathematical model construction for power system scheduling optimization

The PS suffers scheduling complexity issues brought on by volatility and uncertainty as a result of the extensive access to HP of RE sources. Moreover, the conventional phased scheduling approach finds it challenging to satisfy the needs of system stability and economy [15-16]. Therefore, the study proposes a PS scheduling optimization method for HP of RE access.

The two main cores of the method are a mathematical model that can achieve multi-module co-optimization by comprehensively considering system costs, constraints, and uncertainties. The heuristic algorithm can efficiently solve complex optimization problems, taking into account the global search and local fine optimization. Figure 1 depicts the method's general framework.

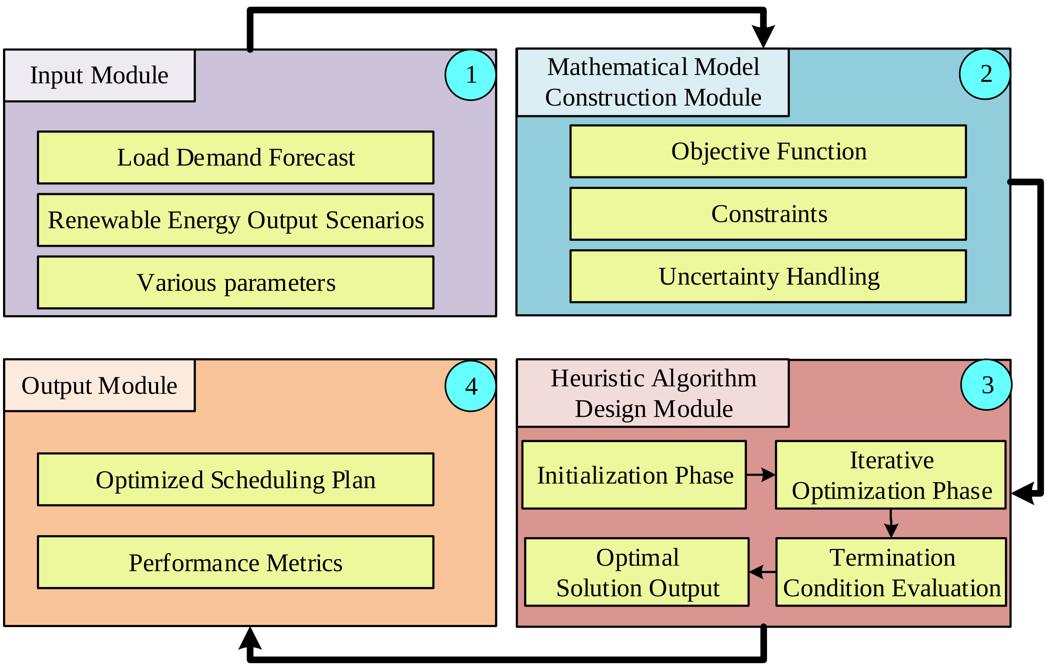


Figure 1: Overall framework of power system scheduling optimization method

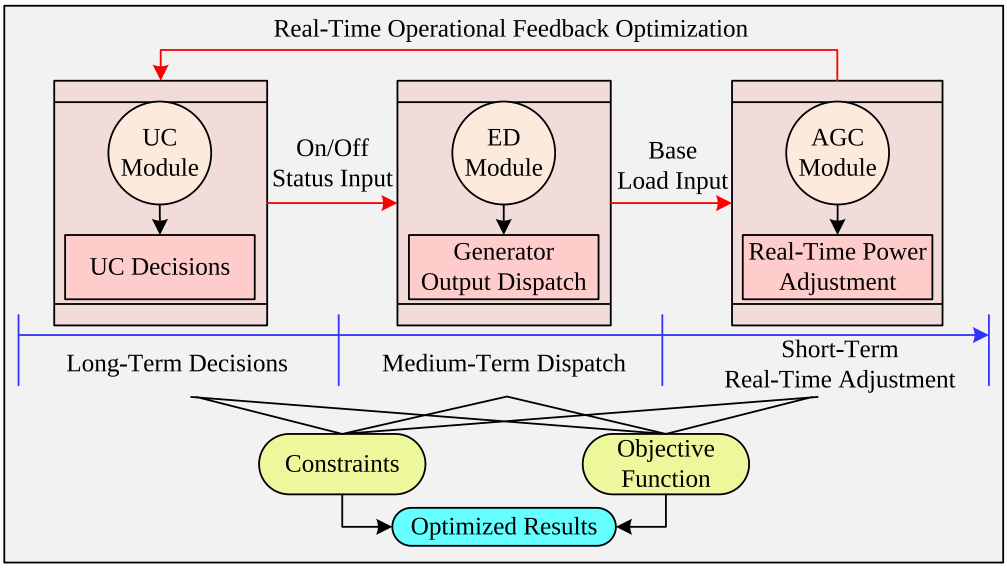


Figure 2: Closed-loop sequential optimization process among UC, ED, and AGC modules

Four components make up the general architecture of the PS scheduling optimization approach suggested in the study, as shown in Figure 1: the input module, the output module, the heuristic algorithm design module, and the mathematical model construction module. The input module contains load demand forecast, RE output scenarios, and various parameters as support. The mathematical model construction module is then responsible for constructing the PS scheduling optimization model based on three main foundations: objective function (OF), constraints, and uncertainty handling [17-18]. The scheduling optimization model based on the four steps is effectively solved using the heuristic algorithm design module. Finally, the output module generates the optimized scheduling plan, including unit start/stop status, power allocation, and standby capacity configuration. It also evaluates the economy and stability of the scheduling scheme through performance indicators.

The mathematical model developed in this study differs from the traditional stage-wise sequential scheduling model in terms of the scheduling optimization approach. The proposed model forms a closed-loop sequential scheduling framework with dynamic feedback coupling among modules by integrating the UC for start-stop decisions, the ED for cost minimization, and the AGC for real-time supply-demand balancing. These modules operate across different time scales and interact through feedback mechanisms to realize coordinated optimization.

The principle of inter-module coordination is illustrated in Figure 2.

As shown in Figure 2, the scheduling optimization process proposed in this study adopts a closed-loop sequential optimization mechanism, consisting of three main modules: UC, ED, and AGC. These modules interact across multiple time scales through real-time feedback to achieve dynamic coordination. During each scheduling cycle, the UC module first optimizes the on-and-off status of units based on current load forecasts, reserve requirements, and other system parameters. Then, it passes the results to the ED module. Then, the ED module performs power allocation and generates a base load profile for the AGC module, which makes real-time, short-term power adjustments. Unlike traditional dispatch models, which operate in isolated stages, the AGC module in this framework generates feedback information, such as load correction values and reserve margin stress levels, continuously during its adjustment process. Rather than discarding this data, it is fed back as correction inputs into the next UC scheduling cycle. Specifically, the system monitors the magnitude and frequency of AGC adjustments in the previous cycle. If frequent or significant real-time corrections are observed, it indicates potential deficiencies in load forecasting or reserve planning. In response, the system increases the reserve capacity settings for the next cycle to improve operational redundancy. At the same time, the load forecast is corrected by incorporating observed deviations into the predicted curve. This enables the UC module to make more accurate start and stop decisions that reflect actual system demand. This adaptive feedback mechanism is repeated in every cycle, progressively refining UC decisions to better match real-world operating conditions and improve overall dispatch responsiveness.

In the mathematical model, schedule optimization aims to reduce the system's overall running costs. The OF is set as shown in Equation (1).

$$\min Z = \sum_{t=1}^{T}\left(\sum_{i=1}^{N}\left(C_{fuel,i,t}+C_{start/stop,i,t}\right)+C_{resreve,t}+C_{EENS,t}\right) \quad (1)$$

In Equation (1), Z denotes the total system cost. T denotes the total quantity of scheduling time segments. N is the total quantity of units. $C_{fuel,i,t}$ is the fuel cost of the i th unit at time t. $C_{start/stop,i,t}$ is the startup and shutdown cost of the i th unit at time t. $C_{resreve,t}$ is the standby cost at time t. $C_{EENS,t}$ is the desired power deficit cost at time t, which mainly measures the supply-demand imbalance caused by RE fluctuations [19]. The UC module is responsible for optimizing the start-stop state (SSS) of the units. The SSS constraint $u_{i,t}$ is shown in Equation (2).

$$u_{i,t} \in \{0,1\}, \forall i,t \quad (2)$$

In Equation (2), a value of 1 for $u_{i,t}$ indicates that the unit is on and a value of 0 indicates that the unit is off. Equation (3) displays the unit start/stop time limitation.

$$\sum_{t=1}^{T} u_{i,t}\Delta t \ge T_{\text{min-on},i} \quad (3)$$

In Equation (3), Δt denotes the time period when the unit is turned on. $T_{\text{min-on},i}$ denotes the minimum continuous operation time of the i th unit. The output power constraint is shown in Equation (4).

$$P_{i,\min} \le P_{i,t} \le P_{i,\max} \quad (4)$$

In Equation (4), $P_{i,\min}$ and $P_{i,\max}$ are the maximum and minimum PO of the i th unit. $P_{i,t}$ denotes the PO of the i th unit at time t. Both Equation (3) and Equation (4) hold only when the value of $u_{i,t}$ is 1. The ED module is responsible for optimizing the power allocation of the turned-on units after the UC determines the SSS of the units [20]. The power balance constraints in the ED module are shown in Equation (5).

$$\sum_{i=1}^{N} P_{i,t} + P_{RES,t} = D_t, \forall t \quad (5)$$

In Equation (5), $P_{RES,t}$ denotes the RE output at time t. D_t denotes the load demand at time t. The climbing capacity constraint is shown in Equation (6).

$$P_{i,t} - P_{i,t-1} \le R_i^{up}, P_{i,t-1} - P_{i,t} \le R_i^{down}, \forall i,t \quad (6)$$

In Equation (6), R_i^{up} and R_i^{down} are the upper and lower climbing limits for unit i, respectively. The reserve capacity constraint is shown in Equation (7).

$$\sum_{i=1}^{N} R_{i,t} \ge R_{required,t}, R_{i,t} = P_{i,\max} - P_{i,t}, \forall t \quad (7)$$

In Equation (7), $R_{required,t}$ is the standby capacity requirement of the system at time t. $R_{i,t}$ denotes the standby capacity that the i th unit can provide at time t. Equation (8) illustrates how the AGC module regulates the fluctuations by modifying the power in real time depending on ED.

$$P_{i,t} = P_{i,t}^{base} + \Delta P_{i,t}, \forall i,t \quad (8)$$

In Equation (8), $P_{i,t}^{base}$ denotes the base point load provided by the ED module [21]. $\Delta P_{i,t}$ is the real-time power adjustment of the AGC module. The real-time balancing constraint is shown in Equation (9).

$$\sum_{i=1}^{N} \Delta P_{i,t} = \Delta D_t, \quad \Delta D_t = D_t - \sum_{i=1}^{N} P_{i,t}^{base} - P_{RES,t} \quad (9)$$

In Equation (9), ΔD_t represents the difference between the actual load demand and the base point load and RE output. The adjustment speed constraint is shown in Equation (10).

$$\Delta P_{i,t} \le R_i^{response} \quad (10)$$

In Equation (10), $R_i^{response}$ denotes the upper limit of real-time regulation speed. As a result, the synergistic relationship among the UC, ED, and AGC modules is realized through the tight coupling of inputs and outputs. The UC provides the SSS for the ED, the ED provides the base point load for the AGC, and the feedback from the AGC optimizes the start-stop strategy of the UC.

The proposed model incorporates uncertainty in wind and solar output directly into the scheduling process to enhance adaptability to RE fluctuations. A limited number of representative renewable output scenarios are generated during each scheduling cycle by applying random deviations to forecasted values based on recent historical variation. These scenarios simulate possible short-term fluctuations in renewable generation. The reserve capacity constraint is adjusted accordingly based on the observed fluctuation range, ensuring sufficient buffer during high-variability periods. In the AGC stage, real-time control targets are fine-tuned using deviation trends derived from these scenarios. The proposed model maintains dispatch feasibility and system stability under uncertain renewable output conditions by dynamically updating reserve settings and AGC parameters.

3.2 Power system scheduling optimization heuristic algorithm design

The proposed mathematical model for optimizing PS scheduling takes into account total system operating costs, the synergistic optimization of multiple modules, and the uncertainty associated with a high proportion of RE sources. This provides a theoretical basis for scheduling. However, the simple model may be inefficient or susceptible to local optimization when dealing with complex, multi-constraint, multi-timescale optimization problems [22]. Therefore, heuristic algorithms are introduced to optimize the mathematical model and solve it. In deep reinforcement learning, the DDPG method effectively optimizes unit SSS and power allocation. However, it may converge slowly and become trapped in local optima when solving complex problems with multiple constraints [23]. QPSO overcomes the limitations of DDPG by improving particle diversity and global search capabilities using quantum behavioral mechanisms. Hence, this study combines improved DDPG and QPSO to propose a joint heuristic algorithm. Figure 3 illustrates the computational flow of the enhanced DDPG in this technique.

In Figure 3, the study introduces a dual experience pooling mechanism in DDPG, which balances exploration and utilization by storing diverse samples and high-value samples separately to improve training efficiency and policy quality. Second, to prevent falling into the local optimum, a time-decaying exploration noise technique is used to boost exploration at the beginning and improve stability at the end. Finally, the target network update strategy is optimized to dynamically adjust the target network parameters through the soft update method to enhance the training stability and convergence speed.

The improved DDPG workflow consists of four main stages. First, in the initialization phase, the Critic network, Actor network, and their target networks are randomly initialized. Additionally, two experience pools, B1 and B2, are established. B1 stores the initial experience samples. B2 stores the high-value samples that are selected using a filtering mechanism. The dual-experience pool design maintains diversity in the training data, which improves sample selection efficiency. Next, within the training iterations and time step loop, the agent generates actions via the policy network. This enhances the exploration of unknown strategies by adding exploration noise. After interacting with the environment, experience samples are generated and stored in B1. High-value samples are then selected based on reward values and stored in B2. In this way, the experience pool contains both common experience samples and high value samples. This ensures the samples are diverse and valuable for training, thereby improving the algorithm's learning efficiency.

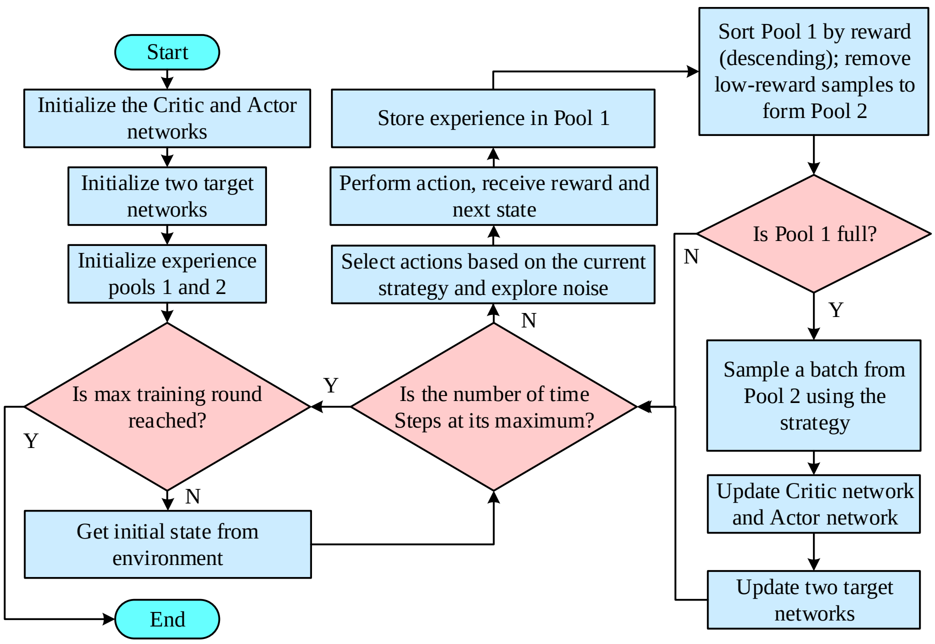


Figure 3: The framework of DDPG-QPSO joint heuristic algorithm

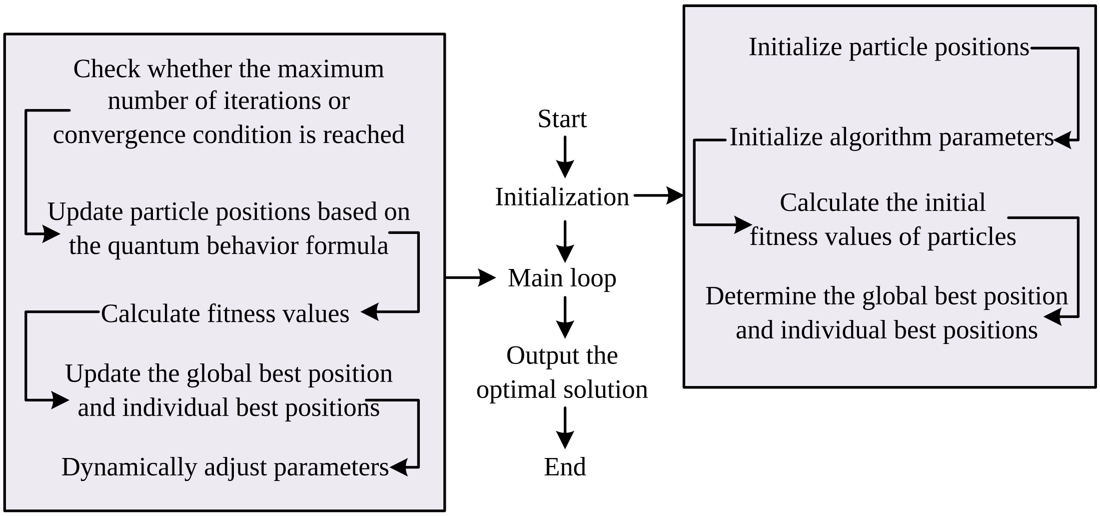


Figure 4: Schematic diagram of the algorithm flow of QPSO

Then, a small batch of high-value samples is sampled from B2 to update the Critic and Actor networks. The Critic network is updated using the error calculated from the target value. Meanwhile, the Actor network is updated using the policy gradient method to maximize the long-term cumulative reward. This optimization allows the model to continuously improve its policy and value function, thereby enhancing the quality of its decisions. Finally, the research employs a soft update method that dynamically adjusts the target network parameters. This further improves training stability and convergence speed. The soft update strategy smoothly adjusts the target network parameters. This prevents excessive fluctuations in the network during training and avoids instability caused by dramatic parameter updates. Figure 4 depicts the QPSO algorithm's flow.

In Figure 4, the overall process of QPSO algorithm is not much different from the traditional PSO algorithm. The steps are initializing particle positions and parameters, calculating fitness values, updating global optimal position (GOP) and individual optimal position (IOP), dynamically adjusting parameters, and iterative judgment to achieve the output optimal solution (OS). However, the core difference between the two is the way of updating the particle position. Traditional PSO is based on the iterative formula of velocity and position, while QPSO adopts the quantum behavioral formula, which constructs the quantum distribution of the particle position through the GOP and IOP. In the QPSO algorithm, the quantum modulation factor controls the randomness of the particle update process and regulates the particles' ability to explore the search space. This enhances search diversity and prevents the algorithm from getting trapped in local OSs. Quantum distribution describes the probabilistic characteristics of particle position updates. New particle positions are generated through formulas based on quantum behavior by combining global and IOPs. This reflects the non-deterministic update mode inspired by quantum mechanics. The quantum behavioral formulation updates the particle positions as shown in Equation (11).

$$x_{i,j}^{(t+1)} = P_{i,j} \pm \beta \cdot \left| x_{i,j}^{(t)} - P_{i,j} \right| \cdot \ln\left(\frac{1}{u}\right) \tag{11}$$

In Equation (11), $x_{i,j}^{(t)}$ and $x_{i,j}^{(t+1)}$ denote the updated position of particle i after t and $t+1$ iterations in the j th dimension. $P_{i,j}$ denotes the reference point of particle i in the j th dimension. β denotes the quantum modulation factor. u denotes the random number, which is used to introduce randomness to give the particle a non-deterministic update property. $P_{i,j}$ is obtained from the GOP and the IOP with certain weights, as shown in Equation (12).

$$P_{i,j} = \phi \cdot p_{best,i,j} + (1-\phi) \cdot g_{best,j} \tag{12}$$

In Equation (12), $p_{best,i,j}$ and $g_{best,j}$ denote the IOP and GOP, respectively. ϕ denotes the inertia factor, which is used to control whether the particle prefers the individual OS or the global OS. In the integration of DDPG and QPSO, the improved DDPG algorithm first generates an initial dispatch strategy based on the current environmental state and load demand information. This strategy includes the start-stop decisions and power allocation for each generation unit over all time periods. The output of DDPG is a deterministic decision vector, representing an executable scheduling solution. This comprehensive scheduling solution serves as a key reference for initializing the population in the QPSO algorithm. More specifically, the DDPG output is encoded as a particle position within the QPSO search space, which is then assigned as the initial position of at least one particle within the swarm. The remaining particles are initialized in the vicinity of this solution through random perturbations, ensuring that the initial population has both guidance and diversity. Based on this initialization, QPSO performs global search optimization. Its quantum-behavior mechanism further refines and adjusts the scheduling strategy, improving the solution's overall stability and adaptability.

Initially, the global search is favored and the local optimization is favored in the later stage. Combining the above, the final PS scheduling optimization flow designed by the study is shown in Figure 5.

In Figure 5, the final PS scheduling optimization process consists of inputting load demand forecasts, RE output scenarios, and related parameters to provide basic data for optimization. With the aim of reducing the overall system cost, a multi-module cooperative scheduling model comprising UC, ED, and AGC modules is built. The improved DDPG is utilized to generate the initial optimization strategy, which is further optimized by QPSO. The optimized scheduling plan covers unit start/stop status, power allocation, and standby capacity configuration. It ultimately achieves efficient and stable scheduling optimization.

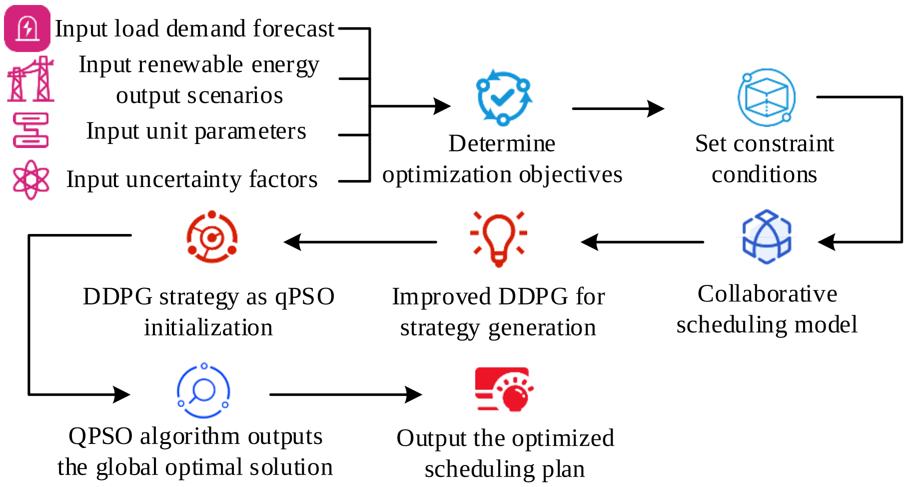


Figure 5: Final power system scheduling optimization process

Table 1: Experimental environment configuration

Hardware configuration		Software configuration	
CPU	Intel Core i9-12900K (16 cores, 3.2 GHz)	Operating system	Ubuntu 22.04 LTS
GPU	NVIDIA GeForce RTX 3090 (24GB VRAM)	Programming language	Python 3.10
Memory	32GB DDR4	Deep learning framework	TensorFlow 2.10
Storage	1TB SSD	Optimization algorithm library	NumPy, SciPy, Pyomo
Power supply	850W High-Efficiency Power Supply	Power system simulation tools	MATPOWER 7.1 (MATLAB Toolbox)
/	/	Data processing tools	Pandas

Table 2: Results of model scheduling performance differences

Indicator	IEEE 30-Bus Test System		IEEE 118-Bus Test System	
	Traditional sequential scheduling model	The proposed model	Traditional sequential scheduling model	The proposed model
Total cost	$12,500	$10,800	$48,200	$42,700
Fuel cost	$7,200	$6,500	$28,500	$26,000
Startup/shutdown cost	$3,000	$2,500	$12,000	$10,200
Reserve cost	$2,000	$1,500	$6,500	$5,500
Demand-supply Imbalance cost	$300	$300	$1,200	$1,000
Supply-demand deviation (MW)	5.5	4.2	25.0	18.5
Response time (s)	10.1	8.3	18.7	13.5

4 Results

The efficiency and superiority of the PS scheduling optimization methods suggested in the study are confirmed in this section using both heuristic algorithms and mathematical models. The focus is on verifying the effectiveness of multi-module collaboration and uncertainty handling, as well as multi-timescale optimization and the performance enhancement and comprehensive optimization capabilities of the improved DDPG, QPSO, and joint heuristic algorithms.

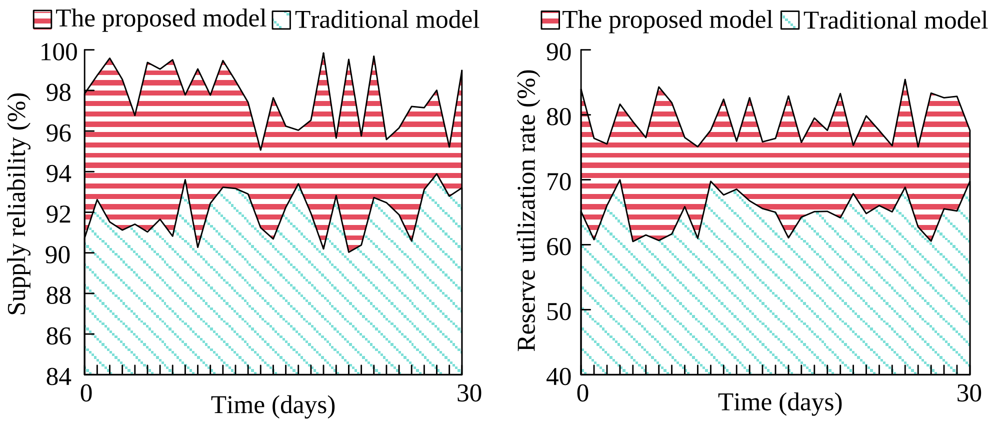
4.1 Validation of mathematical model for power system scheduling optimization

A multi-module cooperative scheduling model including UC, ED, and AGC modules is constructed with the goal of lowering the overall system cost.

Based on Table 1, the study selects the 30-node test system and the 118-node test system from the IEEE standard examples. The former includes 30 nodes, 41 transmission lines, 6 generators, and 20 load nodes, which are suitable for preliminary verification and experimentation. The latter includes 118 nodes, 186 transmission lines, 54 generators, and 99 load nodes, which can be used for in-depth research on the optimization capabilities of multi module collaborative scheduling and heuristic algorithms. First, the performance of dispatches under the proposed closed-loop sequential scheduling model based on UC-ED-AGC feedback is compared with that under a traditional stage-wise sequential scheduling model. The traditional model is a dispatch process in which the UC, ED, and AGC modules run independently in a fixed order. This process does not consider RE uncertainty or provide feedback or coordination. To ensure a fair comparison, both models are solved using the same optimization algorithm (QPSO) under identical system configurations and forecast conditions. This setup ensures that performance differences are attributed to model structure rather than solver differences. Table 2 displays the findings.

In Table 2, the mathematical model proposed in the study demonstrates advantages in both IEEE 30 node and 118 node testing systems. In terms of economy, the total cost of the 30 node system has been reduced by $1700, and the 118 node system has been reduced by $5500. It optimizes fuel, start stop, and reserve capacity costs. In terms of supply-demand balance capability, the supply-demand deviation has been reduced by 1.3 MW and 6.5 MW respectively, effectively addressing the uncertainty of load demand and RE fluctuations. Meanwhile, the real-time adjustment response time is shortened by 1.8s and 5.2s respectively, improving the dynamic response capability. Overall, the mathematical model proposed in the study has achieved more efficient resource utilization in small-scale systems and demonstrated superior adaptability to complex problems in large-scale systems. The suggested model is compared to conventional PS scheduling models that do not take uncertainty handling into account since it takes into account the uncertainty of RE. The result is shown in Figure 6.

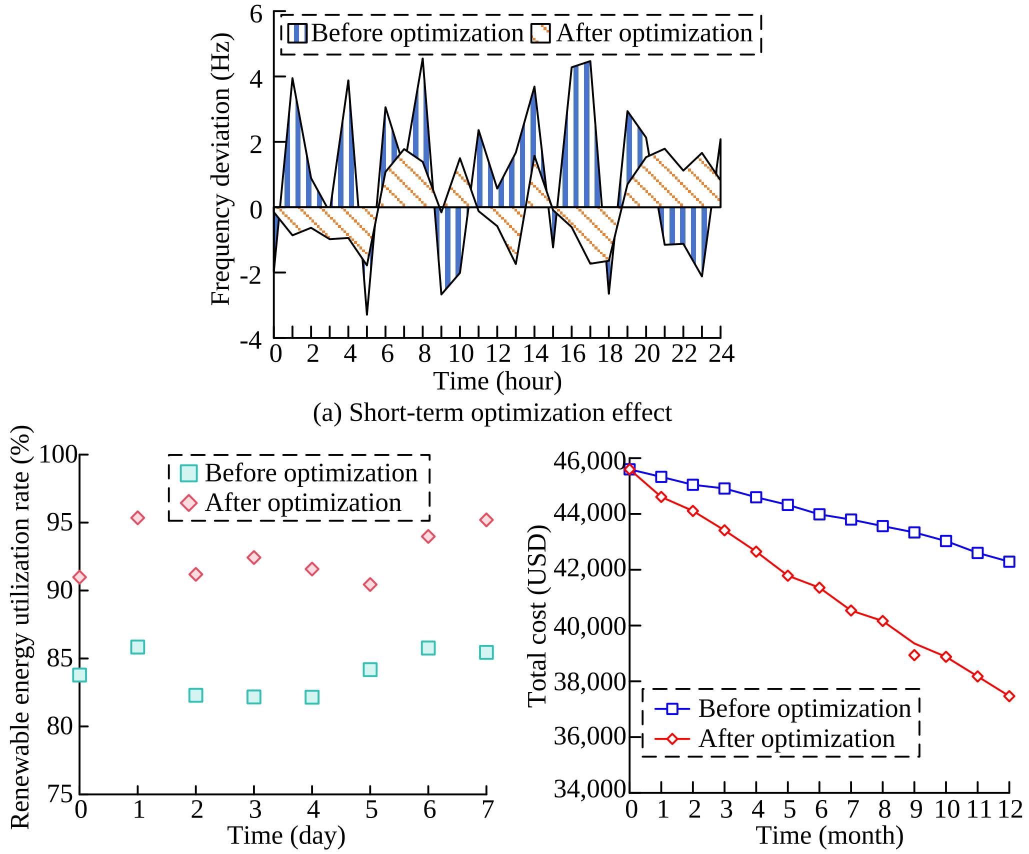
In Figure 6 (a), within 30 days of PS scheduling optimization, the proposed model achieves a power supply reliability of over 94%, with an average of 96.58%. However, traditional models that do not consider uncertainty processing have the highest power supply reliability of only 93.73%, with an average of only 92.16%. In Figure 6 (b), for the utilization rate of backup capacity, after 30 days of model operation, the proposed model increases the utilization rate to between 75% and 87%, while the utilization rate of the traditional model only fluctuated between 60% and 70%. The outcomes displays that the proposed model improves the adaptability of RE fluctuations in PS scheduling optimization. Finally, on various time scales, the impact of scheduling optimization of the study's suggested model is confirmed. The result is shown in Figure 7.



(a) Comparison of supply reliability (b) Comparison of reserve utilization rate

Figure 6: Results compared with traditional models that do not consider uncertainty processing

In Figure 7(a), in the short-term time (24 hours), the frequency deviation (FD) of the PS dispatch before optimization is much larger, up to more than 4Hz. Whereas, after optimization using the research model, the FD of the PS dispatch is effectively controlled and remains between -2Hz and 2Hz. In Figure 7(b), the energy consumption rate of the PS dispatch before optimization averages 84.15% during the interim time i.e., one week, whereas after optimization the energy consumption rate improves to 93.84%. In Figure 7(c), in the long-term time i.e., one year, the optimized PS dispatch significantly reduces the dispatch cost from $45,600 to $37,860 while the pre-optimization dispatch cost is $43,080. The outcomes reveal that the suggested model performs better in terms of long-term economics, medium-term efficiency, and short-term stability.

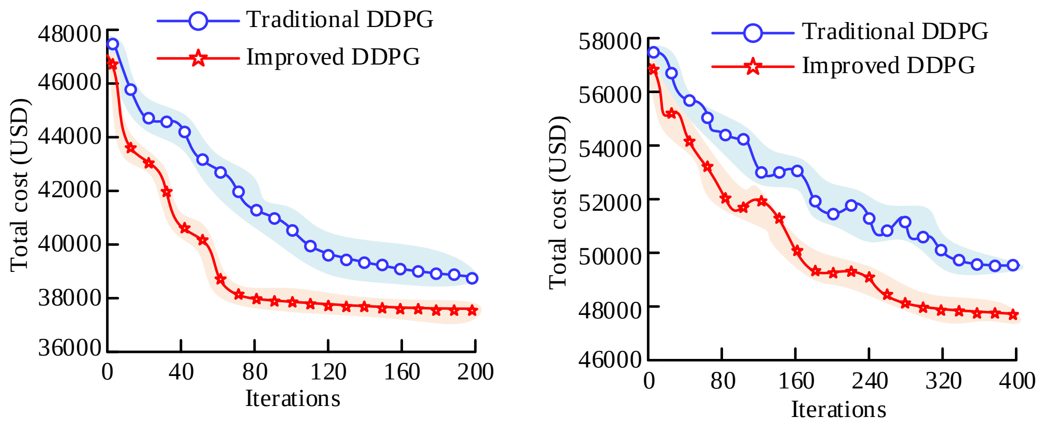


(a) Short-term optimization effect

(b) Medium-term optimization effect

(c) Long-term optimization effect

Figure 7: Scheduling optimization effect on different time scales



(a) IEEE 30-Bus Test System

(b) IEEE 118-Bus Test System

Figure 8: Comparison of DDPG algorithm before and after improvement

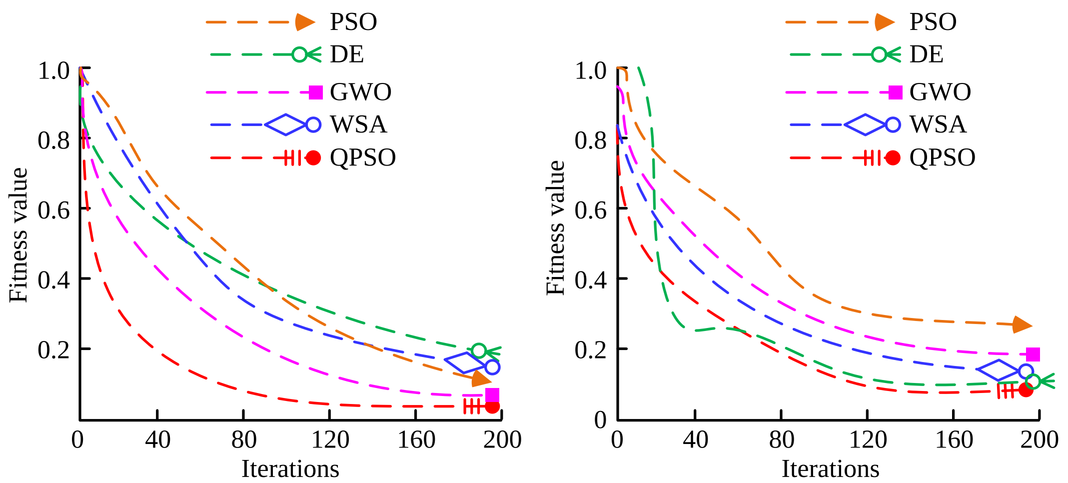
4.2 Validation of heuristic algorithms for power system scheduling optimization

After the validity and superiority of the mathematical model proposed by the study is verified, the study further validates the involved heuristic algorithms. Experiments are first conducted for the improvement of the DDPG algorithm. The DDPG before and after the improvement is applied to solve the mathematical model proposed by the study at the IEEE 30-node and 118-node test systems. When conducting the experiment, in the DDPG algorithm, the learning rates of the Critic network and the Actor network are set to 0.0001. The discount factor is 0.99, the batch size is 64, and the experience pool size is 1,000,000. The noise is examined through the use of an Ornstein-Uhlenbeck process, which has an initial standard deviation of 0.2 and undergoes attenuation during the training process. The target network adopts a soft update with an update parameter of 0.001 to ensure stability. The results are shown in Figure 8.

Figure 8(a) shows that the traditional DDPG algorithm converges after 160 iterations when using the IEEE 30-node system, resulting in a total cost of \$39,560.

The improved DDPG converges after 80 iterations, and the total cost is reduced to $37,960. This improvement benefits from dual experience pooling. Storing low-quality samples and high-value samples separately optimizes the efficiency of sample utilization and improves the training speed. This accelerates the convergence process and reduces the cost. Meanwhile, time-decay exploration improves initial exploration capabilities and stabilizes strategy optimization in later stages, accelerating convergence and reducing costs. In As shown in Figure 8(b), both the traditional and improved DDPG require more iterations to converge when using the more complex IEEE 118-node system. However, the improved DDPG still performs better and has a lower convergence cost. The dual experience pool and time decay exploration strategy effectively improves the algorithm's adaptability and convergence efficiency in large-scale systems, demonstrating its superiority. Furthermore, the optimization effect of QPSO is validated through research, and differential evolution (DE), grey wolf optimizer (GWO), and wolf search algorithm (WSA) are selected for comparison. In the QPSO algorithm, the number of particles is 50, the maximum number of iterations is 1000, and the inertia factor is 0.9. The learning factors are set to 1.5 and 2.0, respectively, to control the global and local optimal attractive forces. Setting the quantum modulation factor to 0.5 enhances the flexibility of the particle position update. Figure 9 displays the findings.

In Figure 9(a), in the IEEE 30-node test system, QPSO has the fastest convergence speed among the five algorithms and the lowest final fitness value. In Figure 9(b), in the IEEE 118-node test system, again QPSO has the fastest convergence speed among the five algorithms and the lowest final fitness value. It can be concluded that QPSO effectively enhances the global search capability of particles through the quantum behavior mechanism. Both in the smaller-scale IEEE 30-node system and in the more complex IEEE 118-node system, QPSO shows superior performance, proving its adaptability to problems of different scales and complexities. Finally, the study applies the proposed mathematical model in combination with the joint DDPG-QPSO heuristic algorithm in HP RE scheduling optimization. The more advanced methods in reference [11], [12], 13], and [14] are selected as comparison methods. The algorithms from references [11] to [14] are re-implemented by the research team based on the original descriptions in the respective papers. Each method is tuned within a reasonable range of parameters based on the recommended settings. Then, it is validated to ensure optimal performance in the current test environment. All methods are evaluated under the same experimental conditions, which includes load forecast profiles, RE output scenarios, system topology, and a unified evaluation period. All performance metrics are kept consistent across experiments. The results are presented in Table 3.



(a) IEEE 30-Bus Test System (b) IEEE 118-Bus Test System

Figure 9: Optimization effect verification of QPSO

Table 3: Comprehensive comparison between research methods and reference methods

Method	Total cost ($)	Energy utilization rate (%)	Supply reliability (%)	Frequency deviation (Hz)	Optimization time (s)
Proposed method	37960	94.85	97.10	1.25	58.3
Reference [11]	40230	90.76	93.85	2.64	125.4
Reference [12]	39760	92.42	94.50	2.18	98.7
Reference [13]	38450	93.57	95.87	1.89	75.6
Reference [14]	38930	93.25	95.30	2.01	88.9

In Table 3, the proposed mathematical model and the joint DDPG-QPSO heuristic algorithm of the study show more obvious advantages in HP of REPS scheduling optimization. The proposed method outperforms other methods with the lowest total cost of $37960 and 94.85% energy consumption rate. Meanwhile, the reliability of power supply reaches 97.10%, the FD is only 1.25 Hz, and the optimization time is 58.3 s. The proposed algorithm demonstrates excellent economy, system stability, and solution efficiency, and provides a highly efficient and reliable solution for PS scheduling optimization in high-percentage RE scenarios.

5 Discussion and conclusion

Targeting the scheduling issue brought on by the HP of RE access in the PS, the study put out a mathematical model for multi-module cooperative scheduling that brought together three main modules and used the enhanced DDPG and QPSO algorithms to address the issue. The efficacy of the study's suggested model and algorithm was confirmed by experimental findings. In the IEEE 30-node and 118-node test systems, the proposed model reduced the total scheduling cost by $1,700 and $5,500, respectively, compared with the traditional sequential scheduling model. It enhanced the energy consumption rate and power supply reliability. The improved DDPG algorithm increased the convergence speed by 50% from $39,560 to $37,960 in the 30-node system by introducing a dual experience pool and a time decay exploration strategy. QPSO exhibited a stronger global search capability, with the fastest convergence speed and the lowest final adaptation value in systems of different sizes compared to other algorithms. In addition, the study's optimization experiments on short-, medium-, and long-term time scales revealed that the FD was effectively reduced, the energy consumption rate was improved by 9.69%, and the total dispatch cost was reduced by 17.6%. The adaptability and superiority of the model in cooperative optimization over multiple time scales were demonstrated.

The DDPG and QPSO algorithms perform well in the test system. However, they may face challenges regarding scalability and adaptability in an actual power grid. As the power grid grows, its computational complexity will increase significantly. This is particularly relevant when working with large volumes of data and real-time dispatching. These factors can lead to a shortage of computing resources and excessively long training times. In addition, the diversity of power grid topologies and operating conditions may affect the algorithm's adaptability. The power grid itself contains complex generators, energy storage systems, and distributed energy resources. Corresponding adjustments to the algorithm are required to effectively address these. In terms of real-time performance, the algorithm functions well in a simulated environment. However, in a highly dynamic actual power grid, it may not respond promptly to load fluctuations and changes in RE. This affects the stability of the system. Therefore, although the algorithm performs well in the test system, it still needs further verification and optimization for practical applications to improve stability and response speed.

References

[1] Lei Gan, Tianyu Yang, Xingying Chen, Gengyin Li, and Kun Yu. Purchased power dispatching potential evaluation of steel plant with joint multienergy system and production process optimization. IEEE Transactions on Industry Applications, 58(2):1581-1591, 2022. https://doi.org/10.1109/TIA.2022.3144652

[2] A. A. Lebedev, A. A. Voloshin, and A. N. Lednev. Conceptual framework for developing highly-automated power distribution networks and micropower systems. Power Technology and Engineering, 58(1):163-168, 2024. https://doi.org/10.1007/s10749-024-01790-2

[3] Ehsan Naderi, Lida Mirzaei, Mahdi Pourakbari-Kasmaei, Fernando V. Cerna, and Matti Lehtonen. Optimization of active power dispatch considering unified power flow controller: application of evolutionary algorithms in a fuzzy framework. Evolutionary Intelligence, 17(3):1357-1387, 2024. https://doi.org/10.1007/s12065-023-00826-2

[4] Lilin Cheng, Haixiang Zang, Anupam Trivedi, Dipti Srinivasan, Zhinong Wei, and Guoqiang Sun. Mitigating the impact of photovoltaic power ramps on intraday economic dispatch using reinforcement forecasting. IEEE Transactions on Sustainable Energy, 15(1):3-12, 2023. https://doi.org/10.1109/TSTE.2023.3261444

[5] Yu Dong, Xin Shan, Yaqin Yan, Xiwu Leng, and Yi Wang. Architecture, key technologies and applications of load dispatching in China power grid. Journal of Modern Power Systems and Clean Energy, 10(2):316-327, 2022. https://doi.org/10.35833/MPCE.2021.000685

[6] Huating Xu, Bin Feng, Chutong Wang, Chuangxin Guo, Jian Qiu, and Mingyang Sun. Exact box-constrained economic operating region for power grids considering renewable energy sources. Journal of Modern Power Systems and Clean Energy, 12(2):514-523, 2023. https://doi.org/10.35833/MPCE.2023.000312

[7] Ahmed M. Abd-El Wahab, Salah Kamel, Mohamed H. Hassan, José Luis Domínguez-García, and Loai Nasrat. Jaya-AEO: an innovative hybrid optimizer for reactive power dispatch optimization in power systems. Electric Power Components and Systems, 52(4):509-531, 2024. https://doi.org/10.1080/15325008.2023.2227176

[8] John L. Cox, William T. Hamilton, Alexandra M. Newman, Michael J. Wagner, and Alex J. Zolan. Real-time dispatch optimization for concentrating solar power with thermal energy storage. Optimization and Engineering, 24(2):847-884, 2023. https://doi.org/10.1007/s11081-022-09711-w

[9] Huifeng Zhang, Dong Yue, Chunxia Dou, and Gerhard P. Hancke. PBI based multi-objective optimization via deep reinforcement elite learning strategy for micro-grid dispatch with frequency dynamics. IEEE Transactions on Power Systems, 38(1):488-498, 2022. https://doi.org/10.1109/TPWRS.2022.3155750

[10] Sicheng Hou, and Shigeru Fujimura. Day-ahead multi-objective microgrid dispatch optimization based on demand side management via particle swarm optimization. IEEJ Transactions on Electrical

and Electronic Engineering, 18(1):25-37, 2023. https://doi.org/10.1002/tee.23711

[11] Navid Shirzadi, Fuzhan Nasiri, Claude El-Bayeh, and Ursula Eicker. Optimal dispatching of renewable energy-based urban microgrids using a deep learning approach for electrical load and wind power forecasting. International Journal of Energy Research, 46(3):3173-3188, 2022. https://doi.org/10.1002/er.7374

[12] Zhongjie Guo, Wei Wei, Mohammad Shahidehpour, Zhaojian Wang, and Shengwei Mei. Optimisation methods for dispatch and control of energy storage with renewable integration. IET Smart Grid, 5(3):137-160, 2022. https://doi.org/10.1049/stg2.12063

[13] Dai Cui, Weichun Ge, Wenguang Zhao, Feng Jiang, and Yushi Zhang. Economic low-carbon clean dispatching of power system containing P2G considering the comprehensive influence of multi-price factor. Journal of Electrical Engineering & Technology, 17(1):155-166, 2022. https://doi.org/10.1007/s42835-021-00877-4

[14] Fatma Gami, Ziyad A. Alrowaili, Mohammed Ezzeldien, Mohamed Ebeed, Salah kamel, Eyad S. Oda, and Shazly A. Mohamed. Stochastic optimal reactive power dispatch at varying time of load demand and renewable energsy resources using an efficient modified jellyfish optimizer. Neural Computing and Applications, 34(22):20395-20410, 2022. https://doi.org/10.1007/s00521-022-07526-5

[15] Jatin Soni, and Kuntal Bhattacharjee. Multi-objective dynamic economic emission dispatch integration with renewable energy sources and plug-in electrical vehicle using equilibrium optimizer. Environment, Development and Sustainability, 26(4):8555-8586, 2024. https://doi.org/10.1007/s10668-023-03058-7

[16] Yukang Shen, Wenchuan Wu, Bin Wang, and Shumin Sun. Optimal allocation of virtual inertia and droop control for renewable energy in stochastic look-ahead power dispatch. IEEE Transactions on Sustainable Energy, 14(3):1881-1894, 2023. https://doi.org/10.1109/TSTE.2023.3254149

[17] Xiaojing Wang, Li Han, Mengjie Li, and Panpan Lu. A time-scale adaptive a forecasting and dispatching integration strategy of the combined heat and power system considering thermal inertia. IET Renewable Power Generation, 17(8):1966-1977, 2023. https://doi.org/10.1049/rpg2.12743

[18] Bing Sun, Ruipeng Jing, Leijiao Ge, Yuan Zeng, Shimeng Dong, and Luyang Hou. Quick hosting capacity evaluation based on distributed dispatching for smart distribution network planning with distributed generation. Journal of Modern Power Systems and Clean Energy, 12(1):128-140, 2023. https://doi.org/10.35833/MPCE.2022.000604

[19] Wei Liu, Tianhao Wang, Shuo Wang, Zhijun E, and Ruiqing Fan. Day-ahead robust optimal dispatching method for urban power grids containing high proportion of renewable energy. Process Safety and Environmental Protection, 178(1):715-727, 2023. https://doi.org/10.1016/j.psep.2023.08.025

[20] Maolin Li, Youwen Tian, Haonan Zhang, and Nannan Zhang. The source-load-storage coordination and optimal dispatch from the high proportion of distributed photovoltaic connected to power grids. Journal of Engineering Research, 12(3):421-432, 2024. https://doi.org/10.1016/j.jer.2023.10.042

[21] Junjie Rong, Ming Zhou, Zhi Zhang, and Gengyin Li. Coordination of preventive and emergency dispatch in renewable energy integrated power systems under extreme weather. IET Renewable Power Generation, 18(7):1164-1176, 2024. https://doi.org/10.1049/rpg2.12893

[22] Zhoujun Ma, Yizhou Zhou, Yuping Zheng, Li Yang, and Zhinong Wei Distributed robust optimal dispatch of regional integrated energy systems based on ADMM algorithm with adaptive step size. Journal of Modern Power Systems and Clean Energy, 12(3):852-862, 2023. https://doi.org/10.35833/MPCE.2023.000204

[23] Jian Hu, Yingjun He, Wenqian Xu, Yixin Jiang, Zhihong Liang, and Yiwei Yang. Anomaly detection in network access-using LSTM and encoder-enhanced generative adversarial networks. Informatica, 49(7):175-186, 2025. https://doi.org/10.31449/inf.v49i7.7246