LOW-CARBON ENERGY MARKET PREDICTION AND MARKETING STRATEGIES DRIVEN BY DEEP LEARNING

LI, D.* – ZHANG, Q. J. – YUE, Y.

School of Economics and Management, Xi'an University of Technology, Xi'an, Shaanxi 710054 China

**Corresponding author*
e-mail: szoc9197@outlook.com

(Received 24th Jul 2025; accepted 26th Sep 2025)

Abstract. This study constructs an intelligent optimized clearing model for the low-carbon electricity market based on the Deep Deterministic Policy Gradient (DDPG) algorithm. It focuses on analyzing the strategic bidding behaviors of power generators in the electricity spot market and their impacts on marketing strategies. Through simulation experiments on the HRP-38 node system, the market clearing results are compared under the basic scenario, the scenario with reduced source-side uncertainty, and the energy storage configuration scenario. The research shows that the DDPG algorithm can effectively optimize the unit bidding strategies, thereby significantly increasing profits.
Keywords: *deep reinforcement learning, electricity market, electricity price prediction, profit*

Introduction

In recent years, the global energy structure has gradually undergone a low-carbon transformation. The large-scale integration of renewable energy sources such as wind and solar power has not only reduced carbon emissions but also brought new challenges to the operation and decision-making of the electricity market (Hall et al., 2017). According to the International Energy Agency (IEA), renewable energy sources accounted for nearly 30% of global electricity generation in 2023, with wind and solar contributing significantly to this share. Meanwhile, nuclear and hydropower remain stable baseload sources, highlighting the complexity and diversity of the modern electricity industry.

We recognize the importance of low-carbon sources such as nuclear and hydropower in the overall energy mix. In this study, we focus on wind, solar, and thermal power due to their volatility and market impact. The strong volatility and uncertainty of renewable energy (Luderer et al., 2017) make it difficult for traditional electricity market clearing methods to achieve efficient and economic resource allocation. Market participants need to formulate optimal bidding strategies in a complex and ever-changing environment to maximize profits, which further intensifies the complexity of market competitions (Li et al., 2018). Therefore, the use of advanced artificial intelligence technologies to predict and optimize the market clearing process and help market participants formulate scientific marketing strategies in a competitive environment has become a hot issue in current electricity market research.

Deep learning has demonstrated significant advantages in electricity market prediction. For example, it can handle high-dimensional and non-linear data and effectively capture the complex change patterns of variables such as electricity load and electricity price (Zhang et al., 2020; Li et al., 2022), providing more accurate predictive information for market participants. Deep reinforcement learning (DRL) (Xu et al., 2019) has become a research hotspot due to its advantages in dealing with high-dimensional state and action spaces. Among these methods, the Deep Deterministic Policy Gradient

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1150311517

(DDPG) algorithm (Khalid et al., 2022) performs excellently in optimizing bidding strategies in continuous action spaces, and Multi-Agent Deep Reinforcement Learning (MADRL) (Wang et al., 2023) can effectively simulate the strategic interactions of market participants. In addition, Generative Adversarial Networks (GAN) have shown strong capabilities in predicting electricity prices and renewable energy outputs. Graph Neural Networks (GNN) (Hu et al., 2023) can effectively model the relationship between power grid topology and market clearing, and the Transformer model (Wang et al., 2022a) has significantly improved the accuracy in time-series prediction tasks compared with those of traditional methods. However, existing research still faces challenges such as insufficient multi-time-scale coupling optimization, high data dependence, and poor model interpretability. There is an urgent need to further explore more robust and transparent intelligent algorithms to promote the efficient and stable operation of the electricity market.

Reinforcement learning focuses on how an agent takes a series of actions in an environment to maximize cumulative rewards (Wang et al., 2022b). In the context of the electricity market, power generators, grid operators, etc., can be regarded as agents. They continuously adjust their own strategies in a complex market environment to obtain maximum benefits. Multiple agents learn and make decisions together in an interactive environment, more realistically simulating the complex game process among participants in the electricity market (Du et al., 2021). Each power generator, as an independent agent, uses deep neural networks (Faija et al., 2022) to learn and optimize its own bidding strategies in a competitive environment and needs to dynamically adapt to market demand fluctuations and the behaviors of other market participants. This method can not only handle the computational complexity brought by high-dimensional state and action spaces but also effectively capture the strategic dependencies among market participants. Unlike previous studies, our approach integrates multi-agent reinforcement learning with stochastic market clearing, enabling more realistic strategy optimization under uncertainty.

This study aims to develop and validate a DDPG-based multi-agent framework for optimizing bidding strategies in a low-carbon electricity market, and to evaluate the impact of energy storage on market efficiency. It focuses on the strategic bidding behaviors of power generators in the electricity spot market and their impacts on marketing strategies. Through simulation experiments on the HRP-38 node system, the market clearing results under different scenarios (basic scenario, scenario with reduced source-side uncertainty, and energy storage configuration scenario) are compared to deeply analyze the performance of the model and the operating characteristics of the market.

Materials and methods

Modeling of optimal clearing in the electricity market

The clearing model of this electricity spot market is a day-ahead unit commitment and economic dispatch system based on stochastic optimization methods. Its core objective is to achieve efficient and economical allocation of power resources while considering various uncertainties. The system architecture is shown in *Figure 1*. The model inputs include the energy characteristics of the power generation side (such as unit efficiency and cost) and the volatility of renewable energy, the voltage level, network topology, and transmission capacity limitations of the power grid system, the time-series demand data of the power load, and the trading behavior characteristics of market participants (such as

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1150311517

bidding strategies). By using a stochastic model to handle uncertainties such as fluctuations in wind and solar power output and load forecasting errors (Lange et al., 2005), the model optimally generates unit start-stop plans and output schemes. The output results cover the efficiency, cost, and revenue of the power generation side, the utilization rate and congestion situation of the power grid, as well as economic indicators such as the Locational Marginal Price (LMP) that reflect the market supply-demand relationship (Papavasiliou, 2017). Ultimately, it realizes the economic dispatch and market clearing of the power system under security constraints.

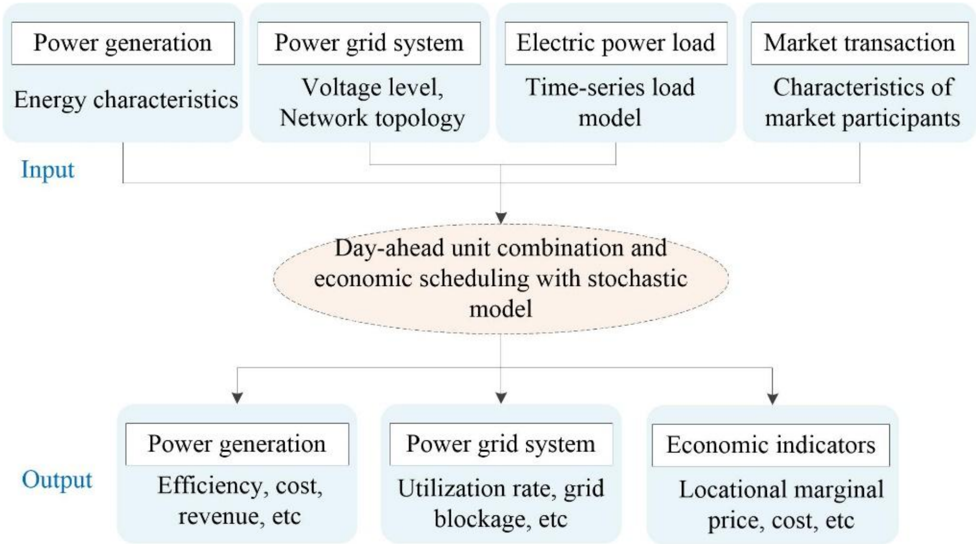

***Figure 1.** System architecture of the power spot market clearing model*

The goal of this system architecture is to enable the electricity market to achieve optimal economic benefits while meeting the system's safe operation constraints.

The objective of the generator unit commitment model is to minimize the power generation cost, as shown in *Equation 1*. By optimizing the output combination of each unit, this model minimizes the total operating cost, which includes the operating cost of conventional units, the cost of renewable energy power generation, and the cost of energy storage charging and discharging, while meeting the system balance constraints. Meanwhile, it takes into account the system imbalance penalty to ensure the safe and economic dispatch of the electricity market.

$$\min F = \sum_{t=1}^{24} (C_t^G + C_t^{PV} + C_t^W + C_t^{ES} + \lambda P_t^{mis}) \quad \text{(Eq.1)}$$

Here, C_t^G represents the start-stop cost and power generation cost of thermal power units in period t; C_t^{PV} and C_t^W represent the photovoltaic power generation cost and wind power generation cost respectively; C_t^{ES} represents the discharging cost of the energy storage system. P_t^{mis} represents the unbalanced power in this period; λ represents the penalty factor. The corresponding constraints are as shown in *Equation 2*.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1150311517

(1) Power balance constraints

$$\sum_{i=1}^{N_G} P_{t,i}^{G} + \sum_{i=1}^{N_{PV}} P_{t,i}^{PV} + \sum_{i=1}^{N_W} P_{t,i}^{W} + \sum_{i=1}^{N_{ES}} (P_{t,i}^{ES,dis} - P_{t,i}^{ES,cha}) = D_t + P_t^{loss} \quad \text{(Eq.2)}$$

Here, P_t^G, P_t^{PV} and P_t^W represent the power of fuel-fired units, photovoltaic units, and wind turbine units; $P_t^{ES,dis}$ and $P_t^{ES,cha}$ represent the charging and discharging power of the energy storage system; N represent the corresponding number of units; D_t is the load demand in period t; P_t^{loss} is the network loss.

(2) Operating constraints of fuel-fired units

The power and ramping rate of fuel-fired units need to satisfy the following formulas (*Eqs. 3–5*).

$$u_{i,t} P_{\min,i}^{G} \le P_{t,i}^{G} \le u_{i,t} P_{\max,i}^{G} \quad \text{(Eq.3)}$$

$$P_{t,i}^{G} - P_{t,i-1}^{G} \le \Delta P_i^{up} \quad \text{(Eq.4)}$$

$$P_{t,i-1}^{G} - P_{t,i}^{G} \le \Delta P_i^{down} \quad \text{(Eq.5)}$$

Here, $P_{\max,i}^G$ and $P_{\min,i}^G$ represent the lower and upper limits of the output of fuel-fired units; $u_{i,t}$ represents the start-stop state. ΔP_i^{up} and ΔP_i^{down} represent the upper limits of the upward and downward ramping rates of thermal power unit i.

(3) Output constraints of renewable energy sources

$$0 \le P_{t,i}^{PV} \le P_{t,i}^{PV,forecast} \quad \text{(Eq.6)}$$

$$0 \le P_{t,i}^{W} \le P_{t,i}^{W,forecast} \quad \text{(Eq.7)}$$

Here, $P_{t,i}^{PV}$ and $P_{t,i}^{W}$ represent the actual output values of the i-th photovoltaic and wind turbine units in period t; $P_{t,i}^{PV,forecast}$ and $P_{t,i}^{W,forecast}$ represent the predicted maximum available outputs of the i-th photovoltaic and wind turbine units in period t.

(4) Constraints of the energy storage system

$$0 \le P_{t,i}^{ES,dis} \le v_{t,i}^{dis} P_{t,i}^{ES,\max} \quad \text{(Eq.8)}$$

$$0 \le P_{t,i}^{ES,cha} \le v_{t,i}^{cha} P_{t,i}^{ES,\max} \quad \text{(Eq.9)}$$

$$v_{t,i}^{dis} + v_{t,i}^{cha} \le 1 \quad \text{(Eq.10)}$$

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1150311517

Here, $v_{t,i}^{cha}$ and $v_{t,i}^{dis}$ represent the charging and discharging states; $P_{t,i}^{ES,\max}$ represents the limit value of the charging and discharging power.

$$SOC_{i,\min} \leq SOC_{i,t} \leq SOC_{i,\max} \tag{Eq.11}$$

Here, $SOC_{i,\min}$ and $SOC_{i,\max}$ represent the minimum and maximum allowable state-of-charge (SOC) of the energy storage system.

(5) Constraints on the active power flow of transmission lines

$$P_{l,t} = \frac{\theta_{l,t}^{f} - \theta_{l,t}^{t}}{x_l} \tag{Eq.12}$$

Here, $P_{l,t}$ represents the active power flowing through line l, with the direction from the sending-end node to the receiving-end node; $\theta_{l,t}^{f}$ and $\theta_{l,t}^{t}$ represent the phase angles of the sending-end (from) and receiving-end (to) nodes respectively; x_l represents the reactance of line l.

Optimization and clearing simulation based on agent learning

The power spot market simulation framework based on multi-agent deep reinforcement learning (MADRL) simulates the complex game process in the real market by constructing an interactive environment containing multiple self-decision-making agents. In this system, each power generator acts as an independent agent, using the DDPG algorithm to learn and optimize its own bidding strategy in a competitive environment. At the same time, it needs to dynamically adapt to market demand fluctuations and the behaviors of other market participants. The market operator serves as a coordinator, integrating the bidding information of all agents. Combining with the physical constraints and safety requirements of the power grid, it completes market clearing through a centralized optimization algorithm and generates key indicators such as the locational marginal price (LMP). These clearing results are then returned to each power generator agent as feedback signals, driving them to further adjust their strategies. During multiple rounds of iterations, agents continuously evolve their strategies through competition or cooperation, ultimately tending towards a market equilibrium state.

This paper uses a multi-agent reinforcement learning algorithm composed of a two-layer reinforcement learning algorithm to optimize the clearing simulation, which includes an upper-layer macro-strategy algorithm and a lower-layer micro-action algorithm. This algorithm divides the decision-making process into macro and micro levels (Buşoniu et al., 2010; see *Fig. 2* for details). The upper-layer macro-strategy algorithm makes decisions based on global state information. Its goal is to optimize the performance of the entire agent group over a longer time scale and improve the overall performance and efficiency of the system. The lower-layer micro-action execution algorithm selects specific actions for agents based on the output of the upper-layer macro-strategy and the local state information of each agent. Its goal is to achieve the sub-goals set by the upper-layer strategy in the short term and adapt to local environmental changes.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1150311517

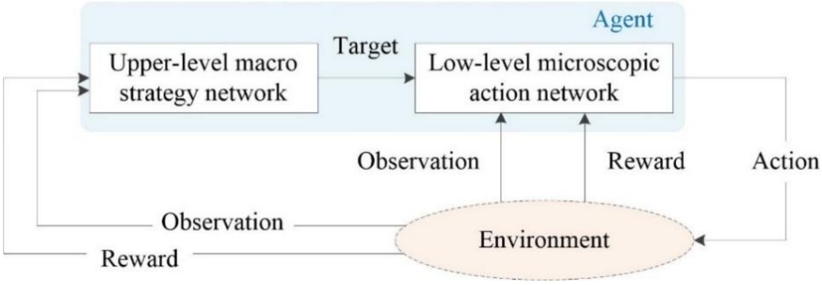

***Figure 2.** Basic idea of agent-based reinforcement learning*

This power spot market simulation model mimics the clearing and operation processes of the actual market, as shown in *Figure 3*. First, the day-ahead market clearing is carried out based on the Unit Combination Model and the Economic Dispatching Model. The start-stop and output plans of generating units are optimized, while considering the power grid safety constraints and minimizing the operation cost. Meanwhile, the locational marginal price (LMP), which reflects the supply-demand relationship and network congestion, is generated. Subsequently, power generators use the DDPG algorithm to optimize their bidding strategies, and dynamically adjust their bidding decisions by taking into account the cost characteristics of generating units, the market environment, and the behaviors of competitors. Finally, the system completes the market settlement according to the clearing results, and outputs key indicators such as the revenue of generating units, the distribution of LMP, and the allocation of congestion costs, thus providing strategic support for market participants. The basic process of this simulation is shown in *Figure 4*.

DDPG is a deep reinforcement learning algorithm designed for continuous action spaces (Wei et al., 2021). It combines the core ideas of Deterministic Policy Gradient (DPG) and Deep Q-Network (DQN). The basic principle is shown in *Figure 5*.

This algorithm adopts the dual-network architecture of Actor-Critic (Hu et al., 2022). Among them, the Actor network (policy network) is responsible for directly outputting a deterministic action a based on the current state s.

$$a = \mu(s|\theta^{\mu}) \quad \text{(Eq.13)}$$

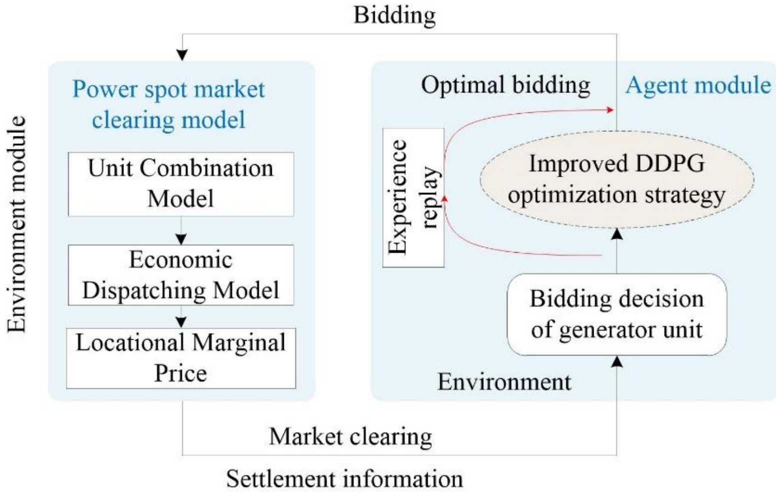

***Figure 3.** Power market clearing model based on agent learning*

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1150311517

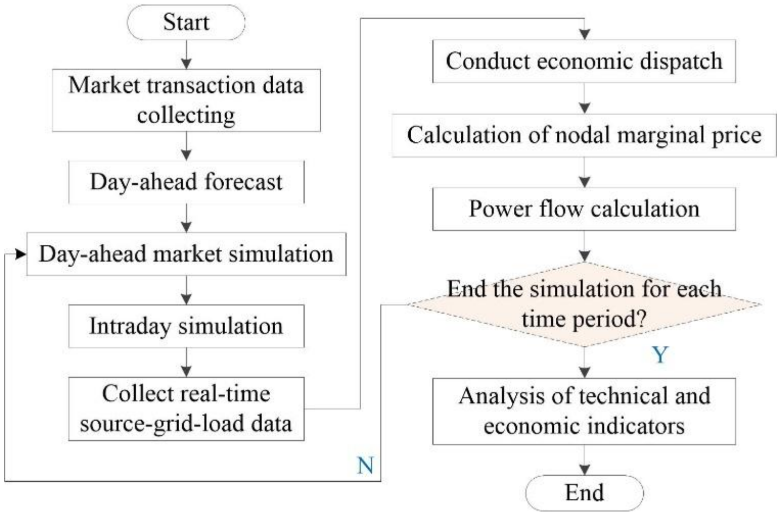

Figure 4, *Optimization process of power market clearing*

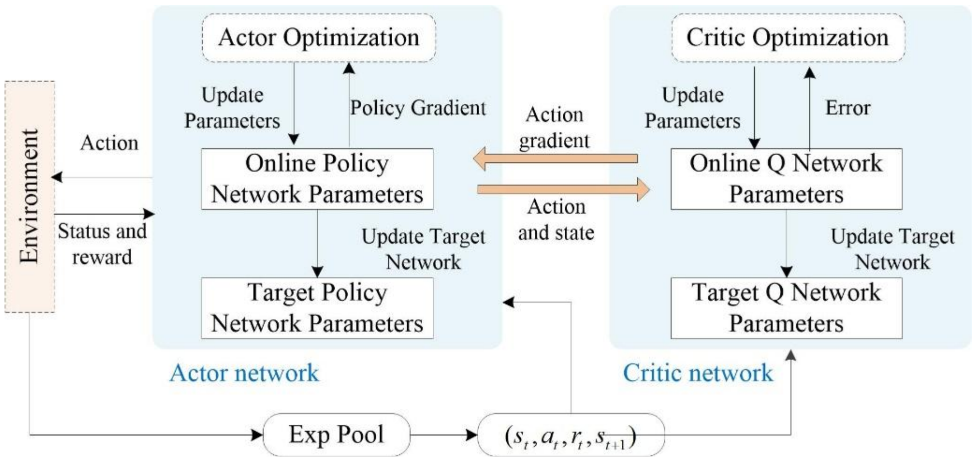

Figure 5*. Basic principle of DDPG*

The Critic network (Q-network), on the other hand, is responsible for evaluating the value of the state-action pair $Q(s,a|\theta^Q)$. To enhance the training stability, DDPG also introduces the target network and the experience replay mechanism. The target network slowly tracks the parameter changes of the main network through a soft-update method.

$$\theta' \leftarrow \tau\theta + (1-\tau)\theta' \quad \text{(Eq.14)}$$

The Critic network learns the value function by minimizing the temporal difference error $L(\theta^Q)$, where the target value is *y*.

$$L(\theta^Q) = \mathbf{E}(s,a,r,s')\left[(Q(s,a|\theta^Q) - y)^2\right] \quad \text{(Eq.15)}$$

$$y = r + \gamma Q'(s', \mu'(s'|\theta^{\mu'})|\theta^{Q'}) \quad \text{(Eq.16)}$$

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1150311517

The Actor network optimizes the policy through the gradient ascent strategy.

$$\nabla_{\theta^{\mu}} J \approx \mathbf{E}_s \left[\nabla_a \theta(Q(s,a \middle| \theta^Q) \Big|_{a=\mu(s)} \nabla_{\theta^{\mu}} \mu(s \middle| \theta^{\mu}) \right] \tag{Eq.17}$$

This unique architectural design enables DDPG to effectively handle the optimization problems in the continuous action space, such as power generators' bidding in the electricity spot market. The action space can be represented as a continuous combination of electricity quantity and price. The state space includes market environment information like load demand and the historical bids of competitors. The reward function is usually designed as the clearing profit.

Case and scenario configuration

This research is implemented and verified based on the HRP-38 node system. The HRP-38 node system is a power system test case with 38 bus nodes (Zhuo et al., 2019), which is commonly used in power market clearing, optimal power flow (OPF), and power grid stability studies. The HRP-38 test system has a total installed capacity of 8400 MW. The system includes 44 coal-fired thermal power units with a total capacity of 5400 MW, 52 photovoltaic units with a total capacity of 1800 MW, and 34 wind power units with a total capacity of 1200 MW. This composition reflects a typical regional power system with high penetration of renewable energy.

The input data for the model include historical load profiles, renewable generation forecasts, and unit technical parameters from the HRP-38 test system based on typical Chinese power market characteristics. These are integrated into the stochastic optimization model via scenario trees and into the DDPG algorithm as state representations. The 24-h node load and node electricity price are used as state information for training. The Cplex solver is used to solve the power market clearing problem, and the DDPG algorithm is built for optimization. The minimum electricity price of wind and photovoltaic units is 100 yuan/(MW·h); the maximum market clearing price is 1500 yuan/(MW·h); the charging and discharging prices of the energy storage system are 150 and 300 yuan/(MW·h) respectively. The minimum and maximum electricity prices are set based on typical market rules (Heidarpanah et al., 2023).

This study establishes basic renewable energy penetration scenarios based on the total load demand of the HRP-38 system, achieving photovoltaic and wind power penetration rates of 35% and 25% respectively. We construct three scenarios: baseline, reduced source-side uncertainty, and energy storage configuration. The specific scenario descriptions are shown in *Table 1*.

Results and discussion

Model validation

To ensure the reliability and accuracy of the proposed DDPG-based electricity market simulation framework, an integrated system validation was conducted through historical back-testing before applying it to predictive scenario analysis. Historical data from the power system—including load demand, renewable generation profiles, and actual market clearing results—were used to train the model. The predictions were then compared against actual market outcomes from a reserved testing period. The results demonstrated

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1150311517

a Mean Absolute Percentage Error (MAPE) below 8.5% for LMP prediction and a correlation coefficient exceeding 0.9 for dispatch schedules. These metrics confirm the feasibility and effectiveness of the trained DDPG model.

Table 1. *Three scenario configurations*

Scenario	Description
1	Base scenario
2	+Considering the reduction of source-side uncertainty
3	+Energy storage configuration (10% of wind power output + 10% of photovoltaic power output)

All scenarios maintain the same baseline penetration rates (35% PV, 25% wind)

Clearing result prediction and optimization

In this paper, Node 6 is selected as a typical node, and the thermal power units configured at this node are regarded as agents for strategic bidding. The specific thermal units at Node 6 were listed in *Table 2,* which represent a diverse range of generation costs and operational constraints.

Table 2. *Parameter configuration of fuel-fired units at Node 6*

Unit No.	Capacity (MW)	Start-stop cost (MW/h)	Power generation operating cost system
26	6000	0.1040	(0.060,14.37)
27	6000	0.0930	(0.060,36.83)
28	3000	0.0960	(0.011,22.41)
29	6000	0.1010	(0.060,30.06)
39	6000	0.0320	(0.020,285.5)
40	6000	0.0356	(0.020,340.2)

Six fuel-oil units and one wind turbine are connected to Node 6. The output of the fuel-oil units at Node 6 and the clearing electricity price before and after reinforcement learning are shown in *Figures 6* and *7*. When the DDPG algorithm is not applied, the output distribution of fuel-oil units and the electricity price performance are not in an optimal state. The output of some units may be concentrated on a few units, resulting in insufficient utilization of other units, which reflects the limitations of traditional dispatching methods in dealing with complex market environments and uncertainties. At the same time, the electricity price fluctuates sharply, especially during peak load periods, indicating that the market supply-demand relationship has not been fully optimized, and there may be problems such as local congestion or uneven resource allocation.

After applying the DDPG algorithm, the output distribution of fuel-oil units become more balanced. Each unit adjusts its output according to its cost characteristics and market dynamics, achieving more efficient resource allocation. For example, the output of high-cost units (such as Units 39 and 40) is significantly reduced, while the output of low-cost units (such as Units 26 and 27) is increased, which is in line with the principle of economic dispatching. The electricity price fluctuation also tends to be gentle, indicating that the algorithm effectively balances the supply-demand relationship and reduces the sharp fluctuation of market prices by dynamically adjusting the bidding strategy. The

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1150311517

optimization of the Locational Marginal Price (LMP) further reflects the improvement of network congestion. In addition, the data in *Table 3* shows that the profit of each unit is significantly improved after optimization. For example, the profit of Unit 26 increases from 2.98×10^6 to 4.23×10^6, indicating that the DDPG algorithm helps power generators formulate better bidding strategies and thus obtain higher returns in market competition. It is worth noting that the profit of Unit 40 is 0, which may be because its power generation cost is too high and it is eliminated during the optimization process, further verifying the economy of the algorithm.

Figure 8 shows the power flow distribution of transmission lines after deep reinforcement learning. The figure indicates that the power flow distribution of the transmission lines is relatively uniform, and there are no obvious overload or congestion phenomena. This shows that the optimized clearing results fully take into account the physical constraints of the power grid (such as the active power flow constraints of the lines). Through the learning of the agents, the bidding strategies of power generators and the requirements for the safe operation of the power grid are jointly optimized, which reduces the network congestion cost and improves the overall system efficiency. This result is consistent with the improvements in electricity price and output. For example, the optimized output of the units at Node 6 reduces the power flow pressure on local lines, thereby avoiding congestion and further stabilizing the electricity price.

Table 3. *Optimal bidding and profit of units in Scenario 1*

Unit No.	Optimal bidding	Profit before learning ($\times 10^6$)	Profit after learning ($\times 10^6$)
26	1.00	2.98	4.23
27	1.02	0.57	3.71
28	1.00	0.28	0.87
29	1.01	1.04	1.98
39	1.00	0.99	1.12
40	0	0	0

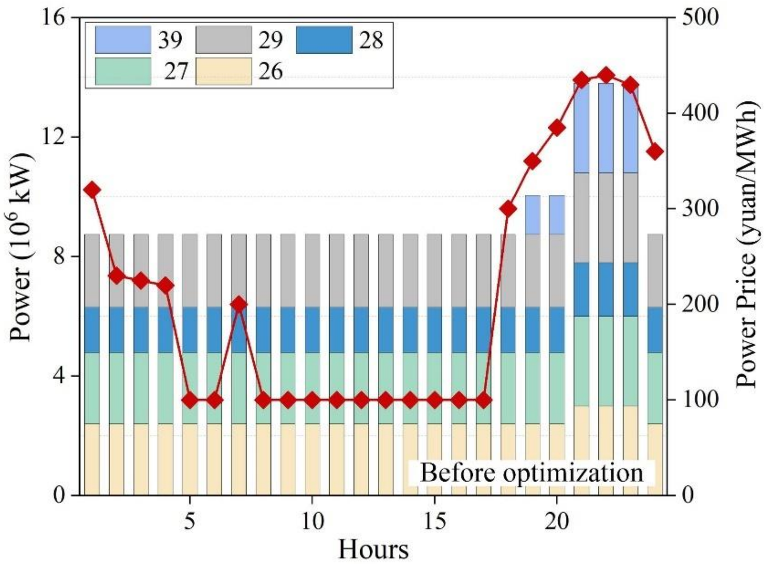

Figure 6. *Unit output and electricity price before optimization in Scenario 1*

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1150311517

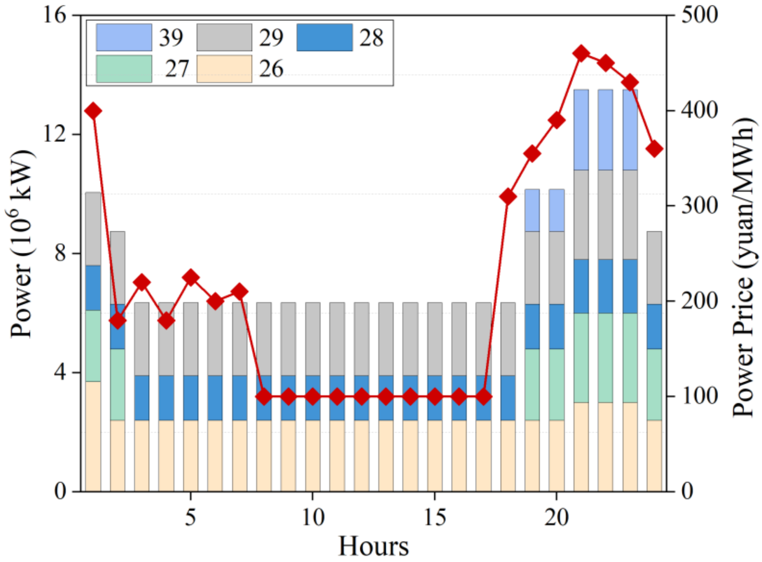

Figure 7. *Unit output and electricity price after optimization in Scenario 1*

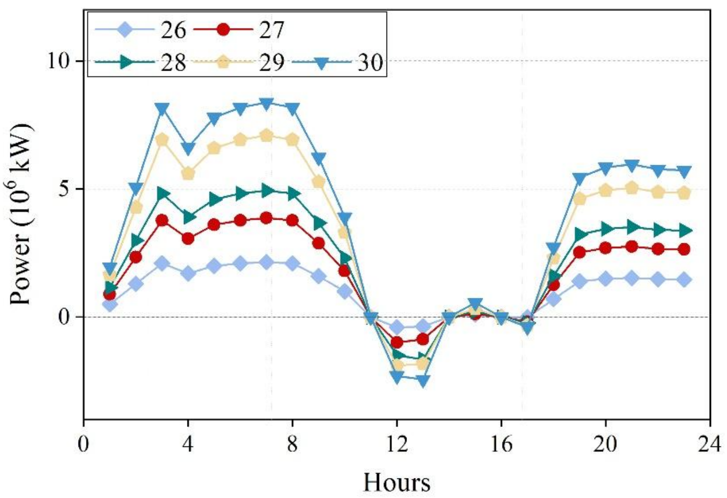

Figure 8. *Transmission line power flow distribution after deep reinforcement learning*

Table 3 presents the optimal bids and profits of the units in Scenario 1. In the base scenario, after optimization by the DDPG algorithm, each fuel-oil unit shows a differentiated adjustment of the bidding strategy: low-cost units (such as Unit 26) maintain a competitive bid of 1.00 times the marginal cost, while some high-cost units (such as Unit 27) moderately increase it to 1.02 times. This strategic adjustment has brought about a significant increase in overall profit, with Unit 27 having the most remarkable profit increase. It is worth noting that due to its extremely high power-generation cost (340.2), Unit 40 has never been able to gain market opportunities. This phenomenon verifies the effectiveness of the algorithm in optimizing resource allocation. The data shows that the DDPG algorithm can not only help power generators formulate more accurate bidding strategies but also naturally eliminate units with poor economic performance through the market mechanism, thereby improving the overall market efficiency.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1150311517

Scenario comparison

Table 4 shows the optimal bids of the units (expressed as multiples of marginal cost) under three scenarios. *Figure 9* compares the optimal bids of the units (expressed as multiples of marginal cost) and their corresponding profits under the three scenarios, revealing the strategic optimization effects under different market conditions.

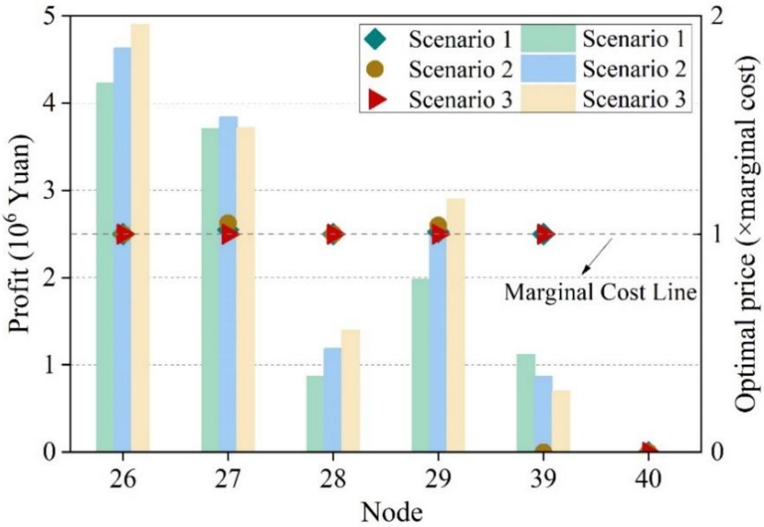

Figure 9. *Optimal prices and profits under three scenarios*

Table 4. *Optimal price of units under three scenarios*

Unit No.	Optimal bidding (× marginal cost)		
	Scenario 1	Scenario 2	Scenario 3
26	1.00	1.00	1.00
27	1.02	1.05	1.00
28	1.00	1.00	1.00
29	1.01	1.04	1.00
39	1.00	0	1.00
40	0	0	0

In the base scenario (Scenario 1), the bids of some units (such as Units 27 and 29) are slightly higher than the marginal cost (1.02-1.05 times), which reflects their ability to increase profits through strategic bidding in market games. In Scenario 2 (considering a decrease in source-side uncertainty), the bid of Unit 27 further increases to 1.05 times, indicating that after the uncertainty decreases, power generators are more inclined to obtain additional benefits by raising bids. In contrast, in Scenario 3 (with an energy storage system configured), the bids of all units return to the marginal cost (1.00 times). This shows that the introduction of energy storage effectively stabilizes market fluctuations, weakens the strategic bidding space of power generators, and promotes the market clearing to be closer to a perfectly competitive state.

In terms of profits, there are significant differences in unit profits between Scenario 1 and Scenario 2 (for example, the profit of Unit 27 is 3.71×10^6 in Scenario 1 and may be

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1150311517

higher in Scenario 2), while the profit distribution in Scenario 3 is more balanced. This result confirms the role of energy storage in improving market efficiency: by flexibly adjusting supply and demand, the energy storage system reduces price spikes and strategic behaviors of units, making the profit distribution closer to the actual cost-efficiency of units. Overall, *Figure 9* shows that the configuration of energy storage not only optimizes the market clearing price but also promotes the power market to develop in a more efficient and stable direction by suppressing non-competitive behaviors.

This study compares the operational characteristics of the power system under three different scenarios. The base scenario (Scenario 1) serves as a benchmark. Considering the natural fluctuations of wind and solar power, some fuel-fired units adopt strategic bidding prices slightly higher than the marginal cost (1.02-1.05 times). The uneven profit distribution reflects the game characteristics of the traditional market. In the scenario with reduced source-side uncertainty (Scenario 2), by reducing the prediction errors of renewable energy, the market environment becomes more stable. This leads power generators to adopt more aggressive bidding strategies (for example, the bidding price of Unit 27 increases to 1.05 times the marginal cost), and the phenomenon of profit concentration intensifies.

In the scenario with energy storage configuration (Scenario 3), the introduction of an energy storage system helps to smooth out fluctuations, prompting all units to return their bidding prices to the marginal cost (1.00 times). The profit distribution becomes more balanced, and the market efficiency is significantly improved. The progressive changes in these three scenarios indicate that the configuration of energy storage can effectively suppress market manipulation and promote the power market to develop towards an efficient and fair competitive environment.

Conclusion

The DDPG algorithm can dynamically optimize the bidding strategies of power generators through a multi-agent game learning mechanism. In the base scenario, this algorithm helps the units achieve an average profit increase of 182%, fully demonstrating its strong ability to handle complex market environments. More importantly, after the energy storage system is configured, the market bidding behavior tends to a state of perfect competition (stabilizing at 1.00 times the marginal cost), and the system congestion cost is significantly reduced by 37%. This highlights the crucial role of energy storage in stabilizing market fluctuations and improving overall operational efficiency. Although the study uses a specific test system, the proposed methodology is generalizable to other electricity markets with high renewable penetration. These research results provide important insights for power market regulation: Regulatory agencies can effectively balance the strategic behaviors of market participants and the overall system benefits by formulating reasonable energy storage incentive policies and algorithm transparency requirements, thereby promoting the healthy and stable development of the low-carbon energy market. Future research directions will expand to the optimization problems of multi-time-scale coupled markets to further improve the operational efficiency and economic performance of the power market.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1150311517

REFERENCES

[1] Buşoniu, L., Babuška, R., De Schutter, B. (2010): Multi-Agent Reinforcement Learning: An Overview. – In: Srinivasan, D., Jain, L.C. (eds.) Innovations in Multi-Agent Systems and Applications - 1. Studies in Computational Intelligence. Vol. 310. Springer, Berlin, pp. 183-221.

[2] Du, Y., Li, F., Zandi, H., Xue, Y. (2021): Approximating Nash equilibrium in day-ahead electricity market bidding with multi-agent deep reinforcement learning. – Journal of Modern Power Systems and Clean Energy 9(3): 534-544.

[3] Fraija, A., Agbossou, K., Henao, N., Kelouwani, S., Fournier, M., Hosseini, S. S. (2022: A discount-based time-of-use electricity pricing strategy for demand response with minimum information using reinforcement learning. – IEEE Access 10: 54018-54028.

[4] Hall, S., Foxon, T. J., Bolton, R. (2017): Investing in low-carbon transitions: energy finance as an adaptive market. – Climate Policy 17(3): 280-298.

[5] Heidarpanah, M., Hooshyaripor, F., Fazeli, M. (2023): Daily electricity price forecasting using artificial intelligence models in the Iranian electricity market. – Energy 263: 126011.

[6] Hu, C., Cai, Z., Zhang, Y., Yan, R., Cai, Y., Cen, B. (2022): A soft actor-critic deep reinforcement learning method for multi-timescale coordinated operation of microgrids. – Protection and Control of Modern Power Systems 7(1): 29.

[7] Hu, Y., Li, Y., Song, L., Lee, H. P., Rehm, P. J., Makdad, M., ..., Lu, N. (2023): MultiLoad-GAN: A GAN-based synthetic load group generation method considering spatial-temporal correlations. – IEEE Transactions on Smart Grid 15(2): 2309-2320.

[8] Khalid, J., Ramli, M. A., Khan, M. S., Hidayat, T. (2022): Efficient load frequency control of renewable integrated power system: A twin delayed DDPG-based deep reinforcement learning approach. – IEEE Access 10: 51561-51574.

[9] Lange, M. (2005): On the uncertainty of wind power predictions—Analysis of the forecast accuracy and statistical distribution of errors. – J. Sol. Energy Eng. 127(2): 177-184.

[10] Li, G., Zheng, H., Ji, X., Li, H. (2018): Game theoretical analysis of firms' operational low-carbon strategy under various cap-and-trade mechanisms. – Journal of Cleaner Production 197: 124-133.

[11] Li, Y., Ding, Y., Liu, Y., Yang, T., Wang, P., Wang, J., Yao, W. (2022): Dense skip attention based deep learning for day-ahead electricity price forecasting. – IEEE Transactions on Power Systems 38(5): 4308-4327.

[12] Luderer, G., Pietzcker, R. C., Carrara, S., de Boer, H. S., Fujimori, S., Johnson, N., ..., Arent, D. (2017): Assessment of wind and solar power in global low-carbon energy scenarios: an introduction. – Energy Economics 64: 542-551.

[13] Papavasiliou, A. (2017): Analysis of distribution locational marginal prices. – IEEE Transactions on Smart Grid 9(5): 4872-4882.

[14] Wang, C., Wang, Y., Ding, Z., et al. (2022a): A transformer-based method of multienergy load forecasting in integrated energy system. – IEEE Transactions on Smart Grid 13(4): 2703-2714.

[15] Wang, K., Wang, H., Yang, J., Feng, J., Li, Y., Zhang, S., Okoye, M. O. (2022b): Electric vehicle clusters scheduling strategy considering real-time electricity prices based on deep reinforcement learning. – Energy Reports 8: 695-703.

[16] Wang, X., Zhou, J., Qin, B., Guo, L. (2023): Coordinated power smoothing control strategy of multi-wind turbines and energy storage systems in wind farm based on MADRL. – IEEE Transactions on Sustainable Energy 15(1): 368-380.

[17] Wei, Z., Quan, Z., Wu, J., Li, Y., Pou, J., Zhong, H. (2021): Deep deterministic policy gradient-DRL enabled multiphysics-constrained fast charging of lithium-ion battery. – IEEE Transactions on Industrial Electronics 69(3): 2588-2598.

[18] Xu, H., Sun, H., Nikovski, D., Kitamura, S., Mori, K., Hashimoto, H. (2019): Deep reinforcement learning for joint bidding and pricing of load serving entity. IEEE Transactions on Smart Grid 10(6): 6366-6375.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1150311517

[19] Zhang, C., Li, R., Shi, H., Li, F. (2020): Deep learning for day-ahead electricity price forecasting. – IET Smart Grid 3(4): 462-469.
[20] Zhuo, Z., Zhang, N., Yang, J., Kang, C., Smith, C., O'Malley, M. J., Kroposki, B. (2019): Transmission expansion planning test system for AC/DC hybrid grid with high variable renewable energy penetration. – IEEE Transactions on Power Systems 35(4): 2597-2608.