ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 3 - 2026

The Application of multi-agent game model in the analysis of agent behavior in electricity market

Zhenduo Gao[1,*], *Zheyong Piao*[1], *Tao Meng*[2] *and Le Liu*[1]

[1] Jilin Electric Power Trading Center Co.,Ltd.,Changchun130000, Jilin Province, China
[2] State Grid Jilin Electric Power Co, Ltd., Changchun 130021 Jilin Province, China

SUMMARY: *A multi-agent game model for agent behavior analysis was proposed to solve the problems of strong coupling, fast strategy feedback and obvious price fluctuation among power generators, electricity selling companies, load aggregators and large users in the electricity market. In this method, the market state coding, agent payoff modeling, optimal response mechanism and iterative solution process are integrated into a unified computing framework to describe the evolution law of multi-agent strategies in continuous trading environment. The simulation results based on 120 trading days and 56 market players show that the social welfare of the model reaches 91.8, the market efficiency reaches 0.943, enters the stable range after an average of 735 rounds, and the price volatility coefficient decreases to 0.118. The overall performance of the model is better than that of the single-agent reinforcement learning, non-cooperative game and centralized optimization model. The research shows that the method can better reveal the formation mechanism of power market main body behavior, and provide computational support for market operation analysis and intelligent decision-making.*

KEYWORDS: *Electricity market; Multi-agent game; Agent behavior analysis; Strategy learning*

1 Introduction

The core of power market operation lies in the continuous game between power generation side, power sales side, load side and aggregators around power, price and trading opportunities. The result not only affects the efficiency of resource allocation, but also relates to the market fairness, system security and energy transition process. With the expansion of new energy grid-connected scale, the enhancement of load fluctuation and the continuous improvement of spot trading mechanism, the power market has evolved from a relatively stable single-layer trading structure to a complex decision-making system with multi-period, multi-agent and strong coupling characteristics. Traditional electric power economic analysis methods are mostly based on the assumptions of complete rationality, static equilibrium or single agent optimization. Although they can explain local market phenomena, they are still insufficient to describe the problems such as strategy linkage, incomplete information and dynamic adjustment of behavior, so it is difficult to accurately reveal the evolution law of subject behavior in the real market environment.

In recent years, the development of computational game theory and intelligent computing methods has provided a new research path for the analysis of the behavior of electricity

*15164356699@163.com
https://doi.org/10.65102/is20261047 Published: 10 June, 2026

market players. The game model can describe the competition and synergy relationship between different market participants, and integrate the bidding, electricity purchase, clearing response and risk aversion behaviors into the strategy space. The stability and accessibility of trading results are analyzed by means of Nash equilibrium and Stackelberg equilibrium. At the same time, the multi-agent learning method has the ability of online update, environment interaction and adaptive decision-making, and can deal with dynamic factors such as price fluctuation, load uncertainty and policy feedback in high-dimensional state space. The combination of multi-agent game and computer modeling technology not only helps to improve the level of refinement of electricity market behavior analysis, but also provides computational support for strategy deduction, mechanism evaluation and simulation verification under complex market rules.

Compared with the research path that only discusses static equilibrium results, the multi-agent game model emphasizes the behavior adjustment mechanism of agents in the continuous trading process. Different types of agents can be abstracted as agents with heterogeneous objective functions and response constraints, and a dynamic interaction network is formed through state perception, benefit evaluation and policy update. The computer simulation platform can further realize the reconstruction of market environment, the iterative solution of parameters and the comparison of multiple rounds of experiments, so as to reflect the "strategy-feedback-re-strategy" cycle process in the electricity market more truly. Based on this understanding, this paper constructs a multi-agent game modeling framework around the behavior analysis of electricity market players, and focuses on the expression of agent relationship, strategy update method, model solution process and behavior analysis effect. Combined with comparative experiments, the performance of different models in market efficiency, revenue distribution and strategy stability is investigated. In order to provide a more interpretable calculation method for electricity market operation analysis and intelligent decision-making research.

2 Related Research

With the continuous evolution of power market from planning and dispatching to market configuration, agent behavior analysis has become an important issue in power economic research and intelligent decision modeling. The existing results generally follow two paths: one focuses on the recognition of behavior rules of market participants in price formation, bidding declaration and transaction feedback, and the other focuses on the reconstruction of the dynamic interaction process of the electricity market with multi-agent, reinforcement learning and computational game methods. The former provides the institutional background and economic explanation for the description of the subject's behavior, while the latter enhances the computability and simulability in complex environments. In general, the related research has formed a relatively clear technical context, but there is still room for further expansion in the unified modeling of heterogeneous agents, dynamic strategy iteration, and linkage analysis of market efficiency and stability. In order to facilitate the comparison of the research focus and application boundaries of the existing work, this paper summarizes the representative results as shown in Table 1.

Table 1: Comparison of representative studies related to behavior analysis of electricity market players

Reference	Research Object	Method Characteristic	Main Conclusion	Main Limitation
[1]	Electricity market and renewable energy trading	Review of game theory	Clarified the theoretical value of multi-agent games in electricity price formation analysis	Focuses mainly on review, lacking computable implementation
[3]	Aggregator-based ancillary service provision	Bi-level Stackelberg game	Capable of capturing strategy transmission among hierarchical agents	Strongly scenario-specific
[4]	Prosumer trading in electricity communities	Game-based coordination mechanism	Strengthened the analysis of multi-agent coordination and community trading efficiency	Insufficient consideration of dynamic learning
[6]	Peer-to-peer trading of shared energy storage	Multi-agent trading modeling	Improved the representation of interactions in distributed trading scenarios	Generalizability remains limited
[8]	Day-ahead market bidding agents	Multi-agent learning	Enables adaptive updating of bidding strategies	Insufficient analysis of equilibrium stability
[10]	Strategic bidding of multiple aggregators	Nash–Stackelberg game	Suitable for analyzing competition among heterogeneous agents	Insufficient discussion of overall market efficiency

2.1 Analysis and research status of behaviors of power market players

In the early stage, the research on the behavior of power market players mainly started from the price mechanism, competition relationship and market structure, trying to explain the strategy selection of power generation enterprises, power sellers and power users under different trading systems. Huang and Li systematically sorted out the game research in the electricity market and renewable energy trading, and pointed out that the formation of electricity price was not the result of single cost transmission, but the comprehensive reflection of repeated games between multiple types of subjects under constraints, which laid a theoretical foundation for market behavior analysis [1]. Subsequently, Tsaousoglou et al. reviewed the local electricity market mechanism, solution methods and algorithm technology, and showed that with the continuous development of distributed resources and two-way trading mechanisms, the analysis of agent behavior has shifted from traditional equilibrium deduction to algorithmic driven interactive simulation [2]. This change means that it is difficult to fully characterize the continuous decision-making process of market players by relying solely on static economic models.

In specific market scenarios, the analysis of agent behavior is gradually deepened in the direction of hierarchical bidding, aggregator decision-making and peer-to-peer trading. Li et al. constructed a double-loop Stackelberg game model for aggregators around the problem of multi-type auxiliary service supply, which revealed the coupling relationship between upper decision-making and lower response, and extended the study of agent behavior from single-round bidding to hierarchical linkage analysis [3]. Lee et al. incorporated producers and consumers into the framework of power community trading and studied the multi-agent strategy coordination problem under the guidance of grid connection, indicating that the objective function of market players has been gradually extended from profit maximization to the comprehensive optimization of transaction efficiency, system security and community welfare [4]. Shojaabadi et al. further analyzed the influence of price fluctuation and uncertain renewable output on strategy selection from the perspective of collaborative bidding between electric vehicle aggregators and wind power entities, reflecting the obvious risk-sensitive and random response characteristics of subject behavior research [5].

2.2 Application of multi-agent game method in electricity market

With the growth of data size and the complexity of decision states in the power trading process, multi-agent game method has become an important computational tool for electricity market analysis. In this method, power generators, electricity sellers, load aggregators, energy storage operators and grid operators are abstracted as agents with independent objectives and local information. By constructing state space, action space and profit function, the strategy update process of market players in repeated games is simulated. Zheng et al. introduced multi-agent transaction modeling into the residential shared energy storage embedded peer-to-peer trading market, and proved that market participants could form a more flexible trading structure through local decision-making and information interaction [6]. Liu et al. also pointed out in their evaluation of the Guangdong electricity spot pilot operation that the subject behavior in the real market is jointly affected by rule design, price signals and regional operation characteristics, and the construction of an iterative calculation model is helpful to improve the explanatory power of behavior analysis [7].

On this basis, researchers begin to embed learning mechanisms into game models to enhance the adaptive ability of agent strategy adjustment. Chandrakala and Kiran constructed a day-ahead bidding method based on multi-agent learning, so that each market agent could continuously revise the bidding strategy in historical trading feedback, which strengthened the dynamic learning dimension in behavior analysis [8]. Lei et al. studied the strategic bidding of distributed energy aggregators based on data-driven Stackelberg game and multi-aggregator Nash-Stackelberg game respectively. They showed that when the number of agents increases and the information asymmetry increases, the combination of multi-agent game and computational solution is adopted. It is more suitable for dealing with heterogeneous interactions in the electricity market [9, 10]. From the perspective of methodology evolution, electricity market research is shifting from "equilibrium for a given strategy" to "parallel advancement of strategy learning and equilibrium approximation", which is a direct reflection of the continuous deepening of computer modeling in this field.

2.3 Deficiency of existing research and research focus of this paper

Although the existing results have made much progress in the design of trading mechanism and decision-making analysis of agent, there are still several shortcomings. First, many studies focus on a certain type of agent or specific transaction scenario, and the model has strong pertinency. However, it lacks a unified expression for multiple types of agents such as power generation side, power sales side, load side and aggregator, which makes it difficult to transfer the analysis results to a more general market environment. Second, some game models can obtain equilibrium solutions, but the description of how the agent's strategies evolve, converge and destabilize in continuous trading is still insufficient. In particular, there is a lack of research on the strategy update mechanism, profit feedback mechanism and market efficiency indicators under the same framework. Third, although some works introduce learning algorithms, they pay more attention to revenue improvement, and lack of discussion on the linkage relationship between market stability, fairness and clearing efficiency, so it is difficult to support a more complete explanation of market behavior.

Based on the above problems, this paper intends to construct a unified model for the behavior analysis of electricity market players under the framework of multi-agent game. On the one hand, different market participants are abstracted as heterogeneous agents, and a computable interaction network is established by combining trading rules, clearing feedback and revenue constraints. On the other hand, the strategy update mechanism and the model solving process are introduced to systematically analyze the bidding adjustment, cooperative behavior and game stability of agents in dynamic environments. At the same time, by setting the comparison model and evaluation index, this paper tests the validity of the model from three levels of behavior recognition results, market efficiency and strategy stability, so as to enhance the applicability and explanatory power of the model in the simulation analysis of the electricity market.

3 Method Design

3.1 Electricity market analysis framework and problem description

In the electricity market environment, there is no independent one-way decision-making relationship among power generators, electricity selling companies, large users, load aggregators and market operators, but a continuous coupling strategic interaction is formed around the quotation declaration, electricity purchase and sale arrangement, load response and market clearing results. Due to the simultaneous existence of new energy output fluctuations, load demand changes, and market rules constraints, the analysis of subject behavior is no longer suitable for only using static equilibrium methods to explain, but needs to use computer modeling methods to build an iterative, updable, and simulable dynamic analysis framework. In this paper, the electricity market is abstracted as a multi-agent interactive system composed of environment layer, agent layer and strategy layer. The environment layer is responsible for receiving load, electricity price, unit constraints and external disturbance information, the agent layer describes the profit goals and behavior boundaries of different participants, and the strategy layer completes the game solution, strategy update and result output. To facilitate the description of the market evolution process in a unified state space, let the market state vector at time t be as follows.

$$s_t = [L_t, P_t, R_t, C_t, D_t] \tag{1}$$

Here, L_t represents the system load level, P_t represents the node electricity price, R_t represents the renewable energy output, C_t represents the unit marginal cost, and Dt represents the demand response intensity. This state vector not only preserves the key economic signals in the electricity market operation, but also provides a standardized input for subsequent multi-agent policy learning. On this basis, the comprehensive revenue function of the ith market agent at time t is defined as follows.

$$U_i^t = \lambda_1 \Pi_i^t - \lambda_2 K_i^t + \lambda_3 S_i^t \tag{2}$$

Here, Π_i^t is the trading profit, K_i^t is the strategy adjustment cost, S_i^t is the market rule satisfaction, and $\lambda_1, \lambda_2, \lambda_3$ are the weight coefficients. The expression not only considers the short-term profit of the subject, but also incorporates the strategy change cost and rule constraints into the same evaluation system, so as to be closer to the decision logic of "profit-risk-compliance" coexistence in the real electricity market.

The problem to be solved in this paper can be summarized as: under the environment of incomplete information and continuous trading, how to use the multi-agent game model to describe the strategy formation mechanism of different agents, and then identify the differences in their behavior, the intensity of competition, and the impact on market efficiency. Following this idea, the analytical framework established in this paper is shown in Figure 1, and the main subjects and their corresponding input-output relations are shown in Table 2.

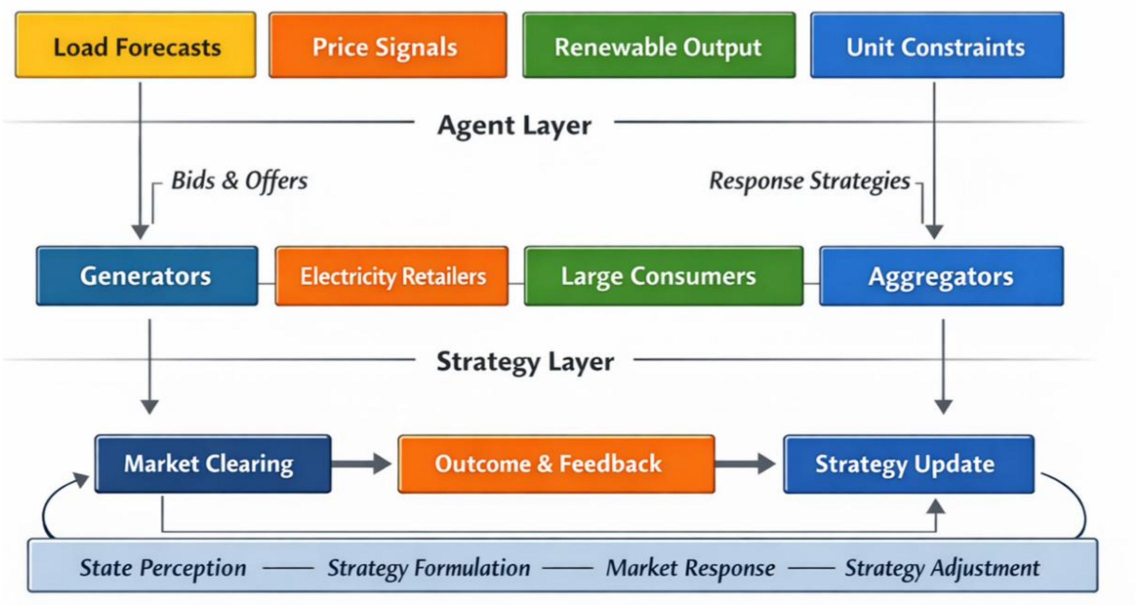

Figure 1: Framework for multi-agent behavior analysis in electricity markets

As shown in Table 2, the observation information, decision variables and optimization objectives of different agents in the market are not consistent, which is an important reason why heterogeneous multi-agent modeling must be used for electricity market behavior analysis. Power suppliers pay more attention to quotation, output and cost recovery, power sellers pay more attention to the purchase and sale price spread and user retention rate, and load aggregators pay more attention to the benefits and response stability of adjustable load calls. Because of the difference between the objective function and the constraint boundary, the evolution of market behavior presents the typical characteristics of multi-agent game.

Table 2: Electricity market main agents and their computing task definitions

Agent Type	Main Input Information	Decision Output	Optimization Objective
Power Generator	Marginal cost, generation unit capacity, forecasted electricity price	Bid price, declared generation quantity	Profit maximization and higher clearing probability
Electricity Retailer	User demand, electricity purchase price, contract constraints	Retail price quotation, electricity purchase portfolio	Revenue stability and risk control
Load Aggregator	Adjustable load capacity, compensation price, response period	Response quantity, participation strategy	Regulation revenue and execution reliability
Large Consumer	Electricity consumption plan, price signals, production constraints	Electricity purchase schedule, peak-shaving arrangement	Minimization of electricity consumption cost
Market Operator	Market-wide bids, network constraints, balancing demand	Clearing results, settlement price	Market efficiency and operational security

3.2 Power market main body modeling and game relationship construction

After completing the definition of the electricity market analysis framework, it is necessary to further transform the key participants in the market into computable, interactive and iterative agent units, so as to provide a structural basis for subsequent strategy solving and behavior analysis. Considering the simultaneous existence of pricing competition, electricity purchase and sale matching, demand response call and market clearing feedback in the spot market, this paper takes the power generation company, electricity sales company, load aggregator and large users into the unified game system, and considers the market operation organization as the node of rule implementation and information feedback. Therefore, the multi-agent game system of the whole electricity market can be expressed as follows.

$$\Gamma = (\mathcal{N}, \mathcal{S}, \mathcal{U}, \mathcal{T}) \tag{3}$$

where N is the set of agents, S is the set of strategies, U is the set of payoff functions, and T is the set of trading and information interaction rules. This expression indicates that the problem studied in this paper is not a single-agent optimization problem, but a multi-agent strategy coupling process constrained by market rules.

At the level of computer modeling, the object-oriented multi-agent representation is used to assign independent state variables, action variables and benefit evaluation functions to each type of agent. For the generator, its decision-making core lies in the coordination between the declared electricity quantity and the quoted price level. Therefore, let the revenue of the g-th generator subject in time period t be:

$$\Pi_g^t = p_t^c q_g^t - C_g(q_g^t) - \eta_g \left| q_g^t - \hat{q}_g^t \right| \tag{4}$$

where p_t^c is the market clearing electricity price, q_g^t is the actual transaction electricity quantity, $C_g(\cdot)$ is the generation cost function, $\hat{q}_g^t$ is the declared planned electricity quantity, η_g is the deviation penalty coefficient. The model not only reflects the transaction profit, but also reflects the constraint effect of deviation assessment on the bidding behavior of the main body. The behavior logic of the electricity selling company in the market is different from that of the power generation company, and its focus is not on production, but on the management of the purchase and sale price difference and the demand matching on the power side. Therefore, the revenue of the RTH electricity selling entity is defined as follows.

$$\Pi_r^t = p_{r,t}^s d_r^t - p_t^c d_r^t - \rho_r \sigma_r^t \tag{5}$$

Here, η_g is the electricity selling price, d_r^t is the electricity sold, σ_r^t is the price volatility risk measure, and ρ_r is the risk penalty weight. This formula shows that the strategy of electricity sales entity is not simply to pursue high-price sales, but to find a relatively stable balance between income and volatility risk.

For the load aggregator, its behavior is mainly manifested as demand response resource invocation, adjustable load integration and auxiliary service participation, which has obvious flexible resource scheduling characteristics. Let the revenue function of the a-th aggregator in time period t be as follows.

$$\Pi_a^t = \omega_t l_a^t + \phi_t e_a^t - \kappa_a (l_a^t + e_a^t) \tag{6}$$

Here, l_a^t represents the adjustable load response, e_a^t represents the amount of energy storage or flexible resource deployment, ω_t and ϕ_t are the response compensation coefficient and the auxiliary service revenue coefficient, respectively, and κ_a is the execution cost coefficient. This payoff design strengthens the intermediary attribute of the aggregator in the market, and also provides a mathematical description for its game coupling with the generation side and the sales side.

Since different main strategies will eventually be transmitted to the market clearing results, it is necessary to set a uniform transaction balance constraint to ensure that the constructed game relationship conforms to the operation logic of the electricity market. In this paper, the following power balance conditions are used:

$$\sum_{g\in G} q_g^t + \sum_{a\in A} e_a^t = \sum_{r\in R} d_r^t + \sum_{u\in U} x_u^t \tag{7}$$

Here, x_u^t denotes the actual electricity purchase demand of the u-th large user in time period t. The constraint integrates the supply of generation side, the adjustment ability of aggregator and the demand of user side into the market settlement plane, so that the behavior changes of various subjects can be transmitted to each other through the clearing results. On this basis, in order to depict the direct competition and indirect influence relationship between the subjects, this paper further constructs the game interaction matrix:

$$M^t = \left[m_{ij}^t\right]_{N\times N}, \qquad m_{ij}^t = \frac{\partial \Pi_i^t}{\partial s_j^t} \tag{8}$$

Here, m_{ij}^t represents the marginal impact strength of agent j's strategy change on agent i's payoff, and s_j^t is the strategy variable of agent j. With the help of this matrix, the intensity of competition among power producers, the sensitivity of electricity sellers to price signals, and the feedback effect on market structure caused by the participation of aggregators can be identified in computer simulation. Therefore, the electricity market is no longer regarded as a simple superposition of several isolated decision making units, but is expressed as a dynamic coupling game network composed of multiple heterogeneous agents.

As shown in Figure 2, the subject relationship constructed in this paper is not a linear conduction structure, but a closed-loop interactive system of "market environment -- subject decision -- clearing feedback -- income correction". The structure not only retains the economic attributes of electricity market operation, but also meets the expression requirements of multi-agent computational modeling for state update, information interaction and strategy coupling, which lays the foundation for subsequent agent strategy updating and model solving.

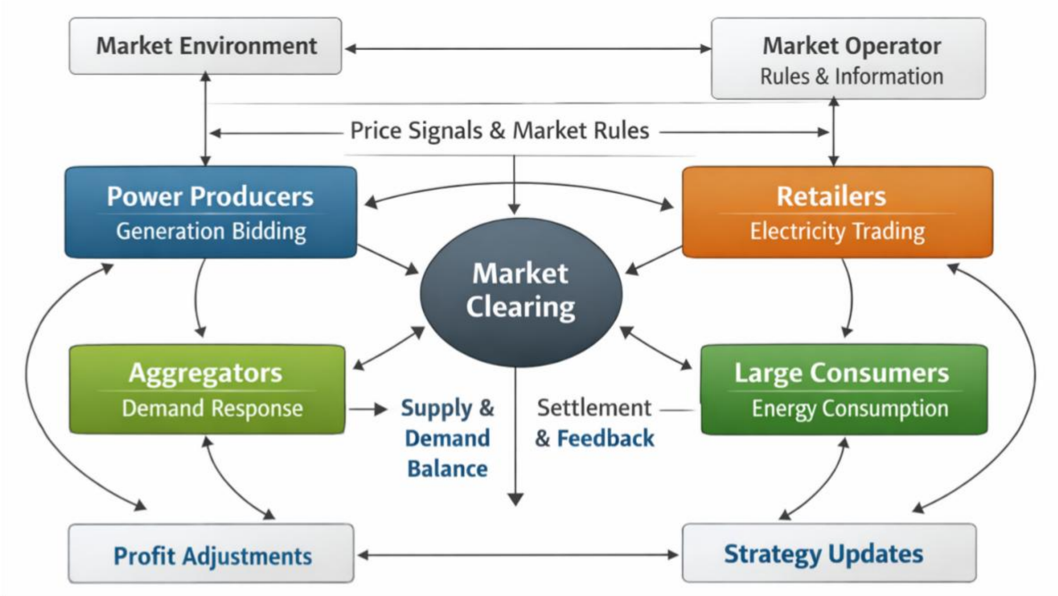

Figure 2: Schematic diagram of power market entity modeling and game relationship construction

3.3 Agent policy update and model solving method

After the subject relationship is determined, the research focus turns to how the strategy is modified, screened and gradually stabilized in successive transactions. The quotation, purchase, response and regulation in the electricity market are not completed in one time, but are constantly adjusted under the joint action of price fluctuation, transaction results and income feedback. Therefore, this paper does not adopt a single round of static solution, but constructs an iterative update mechanism of "state encoder-strategy generation-market clearing-revenue returning-equilibrium correction", so that the model can describe the dynamic evolution trajectory of the agent's behavior in the computer simulation environment. Let the original market state in time period t be s_t, which is mapped by the state-coded network to obtain a compact representation as follows.

$$z_t = \varphi_\omega(s_t) \tag{9}$$

Here, φ_ω is the state representation function with parameter ω. The function of this step is to compress high-dimensional information such as load level, renewable energy output, marginal cost and price volatility, and provide a unified input for subsequent policy learning. For any agent i, its policy output in state z_t is expressed as follows.

$$\pi_i(a_i^t|z_t) = \frac{\exp(h_i(a_i^t, z_t))}{\sum_{a_i \in A_i} \exp(h_i(a_i, z_t))} \tag{10}$$

where hi(·) is the policy scoring function of agent i and A_i is the action space. This equation shows that the model does not directly fix a certain action, but generates the offer or response strategy in a probabilistic form, so as to preserve the behavioral resilience of the electricity market players in an uncertain environment.

When each subject completes the strategy selection, the market clearing module returns the immediate profit and the next period state according to the trading rules. In order to improve the perception of long-term benefits of policy update, this paper uses temporal difference objective to modify state value estimation:

$$\delta_i^t = r_i^t + \gamma V_i(z_{t+1}) - V_i(z_t) \tag{11}$$

Here, r_i^t is the immediate payoff of subject i in time period t, γ is the discount factor, and $V_i(\cdot)$ is the value function. If $\delta_i^t > 0$, it means that the current policy is better than expected, and the model will enhance the selection tendency of this kind of action. If $\delta_i^t < 0$, its probability weight is reduced accordingly. Based on this, the principal policy parameters are updated as follows:

$$\theta_i^{t+1} = \theta_i^t + \alpha\, \delta_i^t \nabla_{\theta_i} \log \pi_i\,(a_i^t|z_t) \tag{12}$$

Here, α is the learning rate and θ_i is the policy network parameters of agent i. Different from the traditional empirical parameter adjustment, this update method can automatically adjust the strategy direction according to the market feedback, which is more in line with the characteristics of adaptive decision-making in multi-agent environment.

Since the agents in the electricity market do not learn in isolation, and the revenue changes of a single agent are often synchronously coupled with the behaviors of other agents, it is necessary to add the equilibrium correction process in addition to the strategy learning. In this paper, the optimal response approximation idea is adopted to calculate the equilibrium deviation of the strategy combination after each round of updating:

$$\varepsilon_t = \sum_{i=1}^{N} \left\| \pi_i^{t+1} - BR_i(\pi_{-i}^{t+1}) \right\|_2 \tag{13}$$

Here, $BR_i(\pi_{-i}^{t+1})$ denotes the optimal response of agent i given that other agent strategies are fixed, and $\|\cdot\|_2$ is the two-norm. When ε_t continues to decrease and is less than the preset threshold, the system can be considered to have entered a relatively stable policy interval. This design avoids the problem that the model only pursues the local revenue rise and ignores the overall strategy coordination, and makes the solution results more suitable for analyzing the market competition intensity, the degree of strategy differentiation and the behavior stability.

According to the above process, the proposed model executes five steps of state sensing, strategy sampling, market clearing, value evaluation and equilibrium correction in each round of simulation, and outputs a stable strategy combination after multiple rounds of repeated interaction. Compared with the method relying solely on the analytical solution of static game, this solution method is more suitable for dealing with the situation of high-frequency feedback, the coexistence of heterogeneous agents and complex rule constraints in the electricity market, and can more truly restore the evolution process of agent behavior from trial to convergence. Figure 3 shows the process of agent policy updating and model solving.

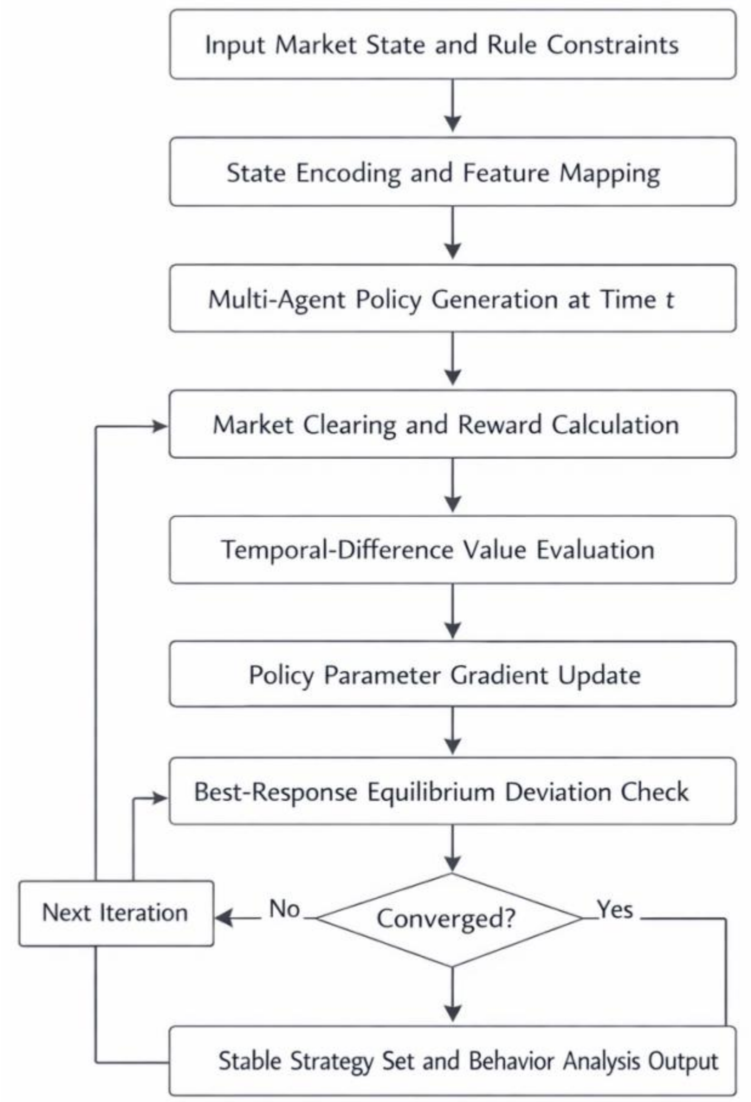

Figure 3: Flow chart of agent policy update and model solution

3.4 Comparison of models and evaluation indicators

In order to judge whether the constructed multi-agent game model can effectively reveal the behavioral characteristics of the main players in the electricity market, this paper sets up comparative experiments under unified data input, the same trading rules and consistent constraints, and makes comprehensive evaluations from four aspects of revenue performance, market efficiency, price volatility and strategy stability. The purpose of this treatment is not to simply compare the value of a certain value, but to investigate the explanatory ability of different modeling ideas to the market competition process, the response mode of the main body and the convergence characteristics of the results. The comparative models set in this paper are shown in Table 3, which include both the single-agent reinforcement learning model without considering multi-agent coupling, the non-cooperative game model emphasizing independent optimal response, and the centralized optimization model guided by the global goal. Table 3 reflects that there are obvious differences between different models in the way of information utilization, decision structure and applicable scenarios, so it has strong comparative significance.

Table 3: compares model Settings and evaluation highlights

Model Name	Decision Characteristic	Computational Method	Applicable Advantages	Limitations
Single-Agent Reinforcement Learning Model	Independent learning by a single agent	Iterative updating based on reward feedback	Simple to implement and low training cost	Difficult to represent game interactions among agents
Non-Cooperative Game Model	Independent competition among multiple agents	Static equilibrium or best-response solving	Capable of analyzing competitive behavior	Insufficient in capturing dynamic strategy adjustment
Centralized Optimization Model	Globally unified decision-making	Mathematical programming solution	High resource allocation efficiency	Weakens agent heterogeneity and independent behavior
Proposed Model	Parallel multi-agent interactive learning and equilibrium correction	Iterative solving based on multi-agent game theory	Can simultaneously reflect behavioral evolution and market feedback	Training and solving process is relatively complex

From the perspective of computational implementation, the single-agent reinforcement learning model regards the market environment as an exogenous feedback system, and trains the strategy network only for a certain agent. Although the solution process is relatively simple, it is difficult to describe the direct game relationship between agents. Non-cooperative game models can describe competitive behavior, but usually rely on static equilibrium solution, and pay insufficient attention to the strategy modification process in continuous trading. The centralized optimization model takes the overall revenue of the market as a unified goal, which can obtain better global allocation results, but it is easy to weaken the independent decision-making and mutual check and balance behavior characteristics of each subject in the actual electricity market. Compared with them, the proposed model takes into account the three links of multi-agent interaction, strategy learning and equilibrium approximation, which is more suitable for the analysis of electricity market behavior. Figure 4 shows the correspondence between the comparative experiments and the evaluation metrics in this paper.

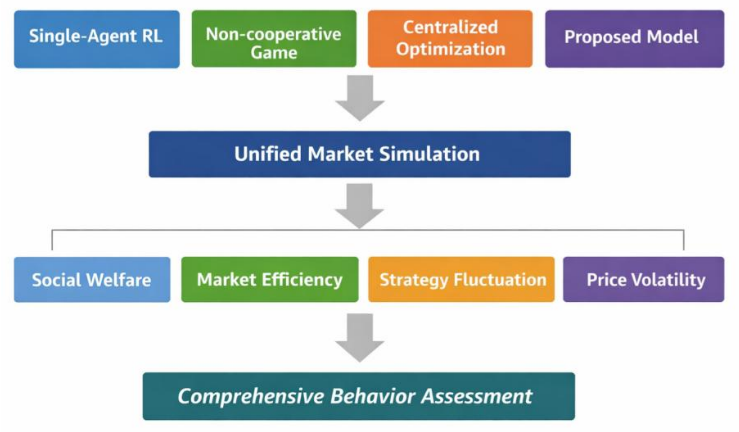

Figure 4: Schematic diagram of the comparison model and the evaluation index system

In order to measure the overall performance of different models in the market, this paper uses the social welfare index to represent the comprehensive income of market operation:

$$W = \sum_{g \in G} \Pi_g + \sum_{r \in R} \Pi_r + \sum_{a \in A} \Pi_a - \Omega \tag{14}$$

Here, Π_g、Π_r and Π_a represent the revenue of the generator, the electricity seller and the load aggregator, respectively, and Ω is the additional cost of system operation. This index can reflect the influence of the model on the overall return level of the market in a unified scale.

Only examining the total revenue is not enough to show the merits of the model, and further evaluation of the market clearing results and the quality of resource allocation is needed. To this end, this paper defines the market efficiency index as follows.

$$E = 1 - \frac{\sum_{t=1}^{T} \left| Q_t^s - Q_t^d \right|}{\sum_{t=1}^{T} Q_t^d} \tag{15}$$

Here, Q_t^s and Q_t^d denote the supply power and demand power at time period t, respectively. This formula describes the matching degree of market supply and demand. The closer E is to 1, the more effective the resource allocation is and the more fully the market is cleared.

Considering that the analysis of power market main body behavior also involves whether the strategy is stable, this paper further constructs the strategy volatility index:

$$F = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=2}^{T} \left\| s_i^t - s_i^{t-1} \right\|_2 \tag{16}$$

Here, s_i^t is the strategy vector of agent i in time period t. The smaller this index is, the more gentle the adjustment range of the agent's strategy is, and the model is easier to converge to a relatively stable game state.

In addition, price stability is also an important dimension to test whether the model is suitable for the simulation of electricity market behavior. In this paper, the standard deviation of market clearing price is used to measure the degree of price volatility:

$$\sigma_p = \sqrt{\frac{1}{T} \sum_{t=1}^{T} (p_t - \bar{p})^2} \tag{17}$$

Here, p_t is the clearing price in period t and $\bar{p}$ is the average price in the sample period. If σ_p is too large, it often means that the agent strategies generated by the model have excessive competition or response imbalance. On the contrary, it shows that the model can maintain the smoothness of market operation to a certain extent. In summary, the evaluation system constructed in this paper does not focus on single prediction accuracy, but integrates revenue, efficiency, stability and price characteristics into the same analysis framework, so that the subsequent experimental results can not only illustrate the computational performance of the model, but also reflect its ability to describe the behavior law of the main body of the electricity market.

4 Experimental design and result analysis

4.1 Experimental environment and data description

In order to verify the applicability of the multi-agent game model constructed in this paper in the behavior analysis of power market agents, the experiment is completed under the unified simulation platform, and the market environment, agent size, sample structure and training parameters are uniformly set. Considering the obvious timing and interactivity of the power market behavior, this paper builds a multi-agent simulation environment based on the operation rules of the regional spot market, and integrates the power generators, electricity sales companies, load aggregators and large users into the experimental scene at the same time. The price feedback, transaction electricity and revenue results are generated through the market clearing module, and then the multi-round iterative learning is completed by the strategy update module. The whole experimental platform is deployed on Ubuntu 22.04 operating system, Python 3.11 is used as the main development language, and PyTorch 2.2 is used as the deep learning framework. The hardware environment is Intel Core i7-12700 processor, 32 GB memory and NVIDIA RTX 4080 16 GB graphics card to ensure the computational efficiency of multi-agent parallel training and repeated simulation.

At the data level, this paper constructs a time-series sample set for electricity market behavior analysis. The sample covers 120 consecutive trading days and is expanded according to the time granularity of 15 min, forming a total of 11520 market session records. The main body size is set to 18 power generators, 6 electricity selling companies, 8 load aggregators and 24 large users, for a total of 56 market players. In order to enhance the authenticity of behavior recognition, the input variables include eight core features such as system load, node electricity price, marginal cost, renewable energy output prediction, subject quotation, electricity declaration, demand response ability, and deviation assessment coefficient. After filling missing values, removing outliers and standardizing Z-score, the original samples were divided into training set, validation set and test set according to the proportion of 70%, 15% and 15%. In the training phase, the batch size was set to 64, the learning rate was 0.001, the discount factor was 0.99, and the maximum number of iterations was 1000. To reduce the perturbation caused by random initialization, all results are averaged over five independent experiments. See Table 4 for the experimental environment and data configuration.

Table 4: Experimental environment and data configuration

Category	Item	Setting Value
Runtime Environment	Operating System	Ubuntu 22.04
Runtime Environment	Programming Language	Python 3.11
Runtime Environment	Deep Learning Framework	PyTorch 2.2
Hardware Environment	CPU	Intel Core i7-12700
Hardware Environment	Memory	32 GB
Hardware Environment	GPU	NVIDIA RTX 4080 16 GB
Market Participants	Number of Power Generators	18
Market Participants	Number of Electricity Retailers	6
Market Participants	Number of Load Aggregators	8
Market Participants	Number of Large Consumers	24
Data Settings	Number of Trading Days	120
Data Settings	Time Granularity	15 min
Data Settings	Number of Sample Time Periods	11520
Data Settings	Feature Dimension	8
Data Settings	Dataset Split	70% / 15% / 15%
Training Parameters	Batch Size	64
Training Parameters	Learning Rate	0.001
Training Parameters	Discount Factor	0.99
Training Parameters	Maximum Number of Iterations	1000
Result Statistics	Number of Repeated Experiments	5

The above Settings make the experiment not only reflect the basic characteristics of the multi-agent parallel game in the electricity market, but also provide a unified data basis for the subsequent behavioral comparison of different models, market efficiency evaluation and strategy stability analysis. Compared with the single static sample test, this experiment organization method based on time series simulation and repeated training is more conducive to observe the evolution process of the agent strategy in the complex market environment and its continuous influence on the clearing results.

4.2 Analysis results of power market main body behavior

After the training is completed, the agent behavior trajectories output by the multi-agent game model are statistically analyzed. The results show that under the joint effect of continuous trading and income feedback, different types of agents show clear strategy convergence characteristics, but there are obvious differences in convergence speed and behavior change range. The average results of five independent experiments show that the model enters the stable interval after about 735 iterations, and the interaction between agents gradually shifts from high-frequency probing in the early stage to relative balance in the middle and late stage. Figure 5 shows the results of the change of the adjustment amplitude of different agent strategies with the iteration rounds.

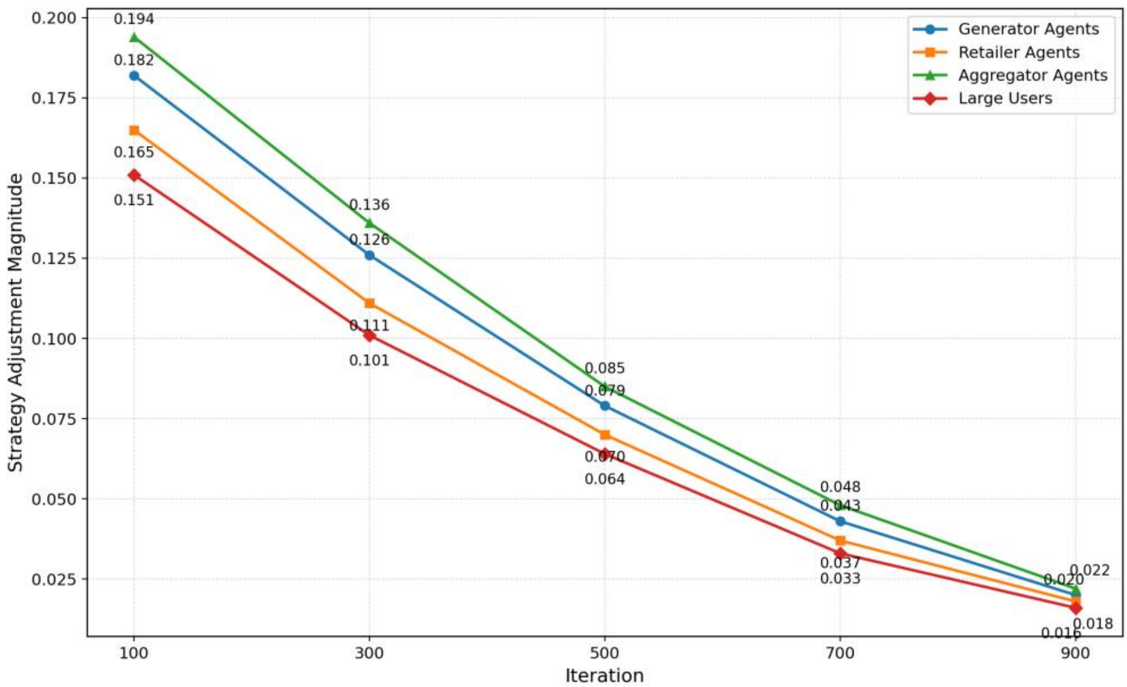

Figure 5: The adjustment amplitude of different subject strategies varies with iteration rounds

From the generation side, the generator has a strong tendency to competitive bidding in the early stage of training, and the proportion of the average bidding deviation from the marginal cost is 13.6%. With the continuous progress of strategy learning, this index drops to 5.4% in the stable stage, indicating that the model can effectively suppress excessive bidding behavior. At the same time, the average clearing success rate of power generators increased from 71.8% to 84.3%, and the unit electricity revenue increased by 9.7%. This shows that in the framework of multi-agent game, power producers no longer rely solely on aggressive bidding to win a deal, but gradually form a rational strategy considering both transaction probability and profit level.

The behavior changes of electricity selling companies and load aggregators are also representative. In the first 200 iterations, the fluctuation of the price difference between purchase and sale of the electricity selling company was obvious, and the standard deviation reached 0.024 yuan /kWh. After entering the stable stage, the average purchase and sale price spread converged to 0.069 yuan /kWh, and the fluctuation standard deviation decreased to 0.009 yuan /kWh, indicating that its price following and risk control capabilities were enhanced. The average demand response amount of load aggregators increased from 41.7 MWh to 56.9 MWh in the initial stage, and the response execution rate reached 92.1%, reflecting that the model can more accurately identify the high-yield response period, and enhance the enthusiasm of flexible resources to participate in the market.

The change of the user-side behavior is mainly reflected in the reallocation of the power purchase timing. The proportion of power purchase of large users in high-price periods decreases from 42.8% to 31.5%, and the proportion of power purchase in low and flat periods increases accordingly, so that the average power purchase cost of large users decreases by 8.4%. From the overall results of the market, the deviation rate of supply and demand is reduced from 4.9% to 2.2%, and the volatility coefficient of clearing price is reduced from 0.186 to 0.118, indicating that the model in this paper not only describes the evolution of the main body's behavior, but also improves the degree of coordination of market operation. In general, the multi-agent game model can not only identify the strategy differentiation paths of different agents, but also promote the market behavior from local confrontation to relative stability through the continuous feedback mechanism, which provides the result basis for the subsequent model comparison analysis and market efficiency evaluation.

4.3 Comparative analysis of different models

In order to test the effectiveness of the proposed model in the analysis of agent behavior in the electricity market, this paper compares it with the single agent reinforcement learning model, the non-cooperative game model and the centralized optimization model. Five independent experiments are completed under the conditions of the same data set, the same agent size and the same number of training rounds, and the results are averaged. The comparison results show that the proposed model performs better in three aspects: income level, market coordination and strategy convergence speed. As shown in Figure 6, the social welfare value of the proposed model reaches 91.8, which is significantly higher than 79.6 of the single-agent reinforcement learning model, 85.4 of the non-cooperative game model, and 89.7 of the centralized optimization model. This shows that in the multi-agent continuous interaction scenario, the combination of game relationship construction and strategy learning mechanism can more fully identify the competition-response-readjust process between agents, and form better market behavior results on this basis.

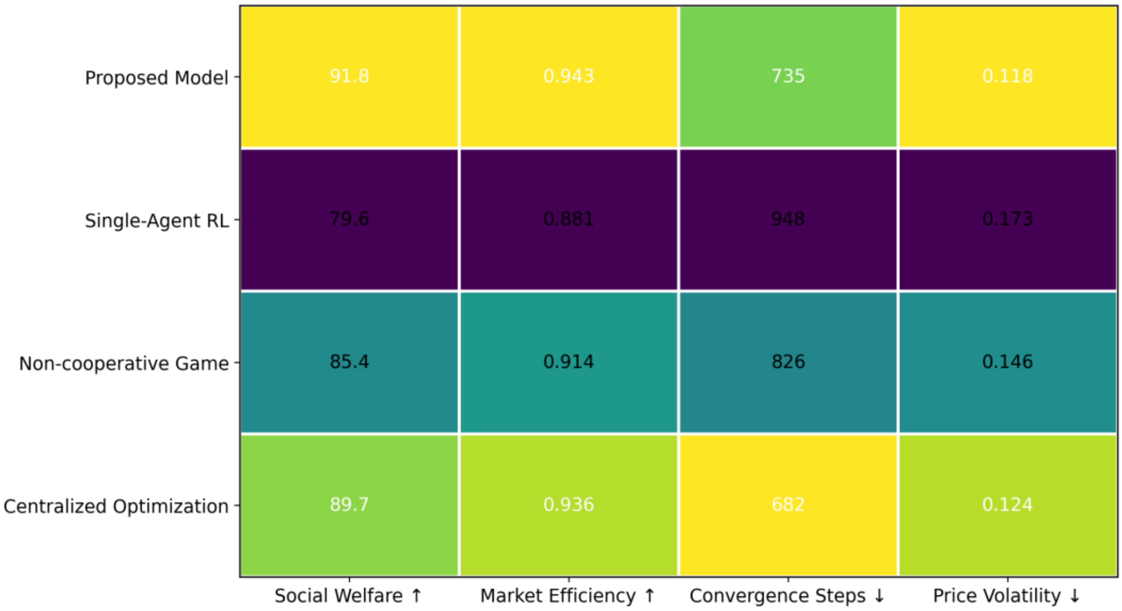

Figure 6: Comparison of main performances of different models

From the perspective of market efficiency index, the supply-demand matching efficiency of the proposed model reaches 0.943, which is higher than 0.881 of the single-agent reinforcement learning model, 0.914 of the non-cooperative game model, and 0.936 of the centralized optimization model. The reason is that the single agent model regards the market environment as an exogenous feedback system, and it is difficult to describe the behavior interaction between different agents. Although the non-cooperative game model can describe the competitive relationship, its adaptability to dynamic strategy modification is still limited under the condition of continuous trading. The centralized optimization model has strong global allocation ability, but weakens the realistic characteristics of the coexistence of decentralized decision-making and local game in the electricity market. In contrast, through the multi-agent iterative update mechanism, the model in this paper can gradually form a relatively stable strategy combination in multiple rounds of the game, so as to achieve a better balance between the market clearing efficiency and the behavior explanation ability.

In terms of convergence performance, the proposed model enters the stable interval after an average of 735 iterations, which is faster than the 948 rounds of the single-agent reinforcement learning model and the 826 rounds of the non-cooperative game model, but slightly slower than the 682 rounds of the centralized optimization model. The centralized optimization model has certain advantages only from the perspective of convergence speed, but its centralized solution method does not fully conform to the operation characteristics of independent decision-making, incomplete information and feedback lag in the real electricity market. Combined with the price volatility comparison results in Figure 6, it can be seen that the price volatility coefficient corresponding to the proposed model is 0.118, which is lower than 0.173 of the single-agent reinforcement learning model, 0.146 of the non-cooperative game model and 0.124 of the centralized optimization model, indicating that the proposed model has stronger adaptability in controlling price shocks and alleviating excessive competition.

4.4 Analysis of market efficiency and strategy stability

Based on the results of agent behavior identification and model comparison, this paper further investigates the reliability of model output from two dimensions of market efficiency and strategy stability. For the electricity market, if the model can only improve the profit of local agents, but can not maintain the matching between supply and demand, price stability and strategy convergence, the analysis results are still difficult to explain the real market behavior. Based on this, this paper selects five indicators of supply and demand deviation rate, market clearing rate, price volatility coefficient, strategy volatility amplitude and equilibrium deviation for joint analysis, and the statistical results are shown in Table 5.

Table 5: Results of market efficiency and strategy stability indicators

Indicator	Initial Stage	Stable Stage	Change Characteristic
Supply–Demand Deviation Rate	4.9%	2.2%	Significant decline
Market Clearing Rate	88.7%	94.3%	Continuous improvement
Price Volatility Coefficient	0.186	0.118	Reduced volatility
Average Strategy Fluctuation Amplitude	0.128	0.019	Rapid convergence
Equilibrium Deviation	0.117	0.012	Approaching a stable equilibrium
Renewable Energy Consumption Rate	89.6%	93.8%	Coordinated improvement

In terms of market efficiency, the average deviation rate between supply and demand of the proposed model on the test set is 2.2%, which is significantly lower than 4.9% at the beginning of training. The market clearing rate reaches 94.3%, which indicates that most offers and power purchase requests can be effectively matched under the uniform clearing rules. The volatility coefficient of the time clearing price was stable at 0.118, and there was no persistent violent shock, indicating that the multi-agent strategy update process did not cause obvious price instability. Table 5 further shows that the consumption rate of new energy increases from 89.6% to 93.8%, indicating that after the coordinated response of load aggregators and large users, the market's adaptability to fluctuating power sources is enhanced, and the efficiency of resource allocation is also improved.

In terms of strategy stability, the model in this paper still has obvious tentative adjustment in the first 200 rounds of iterations, and the average strategy fluctuation range of different subjects is 0.128. When the iteration is advanced to 700 rounds, the index drops to 0.027, and further converges to 0.019 after 900 rounds. As shown in Figure 7, the market efficiency curve generally shows an upward trend, while the strategy fluctuation amplitude and equilibrium deviation continue to decline, and the two change directions have a strong correspondence. This indicates that after continuously receiving market feedback, the subject can gradually correct the non-steady state behaviors such as excessive quotation, conservative power purchase or excessive response, and then form a strategy combination closer to the equilibrium state. The equilibrium deviation index decreases from 0.117 in the initial stage to 0.012 in the stable stage, which also shows that the model in this paper has better strategy coordination ability from another aspect.

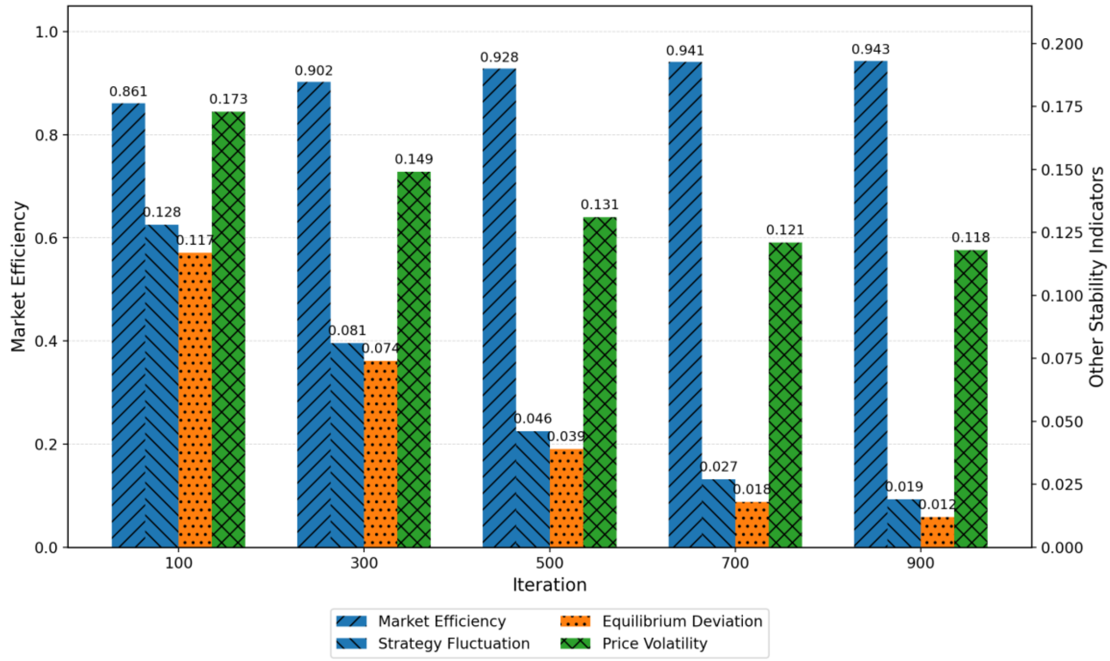

Figure 7: Trends of market efficiency and strategy stability

In general, the model does not sacrifice stability for short-term efficiency improvement, but improves the quality of market allocation and the level of strategy convergence synchronously in the process of continuous game and feedback learning. For the analysis of power market main body behavior, this means that the model can not only identify the difference of main body behavior, but also more truly reflect the dynamic evolution process of the market from competition game to relative stability.

5 Discussion

Based on the above experimental results, it can be seen that the multi-agent game model constructed in this paper shows good adaptability and explanatory ability in the analysis of power market main body behavior. Compared with the single agent reinforcement learning model, the proposed model can simultaneously depict the strategic linkage relationship among power generators, electricity selling companies, load aggregators and large users, thus achieving better results in terms of social welfare, market efficiency and price stability. Compared with the non-cooperative game model, the model does not stop at static equilibrium solution, but introduces a strategy learning mechanism under continuous feedback, so that agents can continuously modify their behavior paths according to the market clearing results, thereby weakening the strategy shock and improving the convergence efficiency. Compared with the centralized optimization model, the proposed model does not rely on the global centralized control, but still maintains a high level of resource allocation, which is closer to the operation characteristics of decentralized decision-making, dynamic competition and continuous adjustment in the real electricity market.

The good effect of the model in this paper is mainly related to the following factors. On the one hand, the modeling of game relationship enables the payoff conflicts, behavioral constraints and mutual influence of different agents to be expressed in a unified computing framework, avoiding the deviation caused by simplifying market behavior into an isolated decision-making process. On the other hand, the strategy update mechanism combines state representation, profit feedback and optimal response modification, so that the agent gradually approaches the stable strategy interval in multiple rounds of interaction, which has a direct effect on suppressing excessive price volatility and improving the matching between supply and demand. In addition, the multi-agent simulation method adopted in this paper enhances the adaptation ability of the model to the complex market environment, so that it can not only output numerical results, but also reveal the evolution logic of agent behavior from competition temptation to relative equilibrium. At the same time, there are still some limitations in this study. The current experiments are mainly based on the simulation environment under unified rules, and the description of cross-regional trading, network blocking constraints and extreme volatility scenarios is still not sufficient. In the future, fine-grained grid topology information and uncertainty disturbance mechanism can be further introduced to improve the fitting ability of the model to the real market operation situation.

6 Conclusions

Focusing on the problem of behavior analysis of power market players, this paper constructs a multi-agent game model, and integrates state perception, strategy update and market clearing feedback into a unified computing framework, and systematically studies the interactive behavior of power generation providers, power sales companies, load aggregators and large users. The experimental results show that the proposed model can better describe the evolution process of agent strategies from competitive exploration to dynamic convergence in a continuous trading environment, and show strong advantages in social welfare, market efficiency, price volatility control and strategy stability. Compared with the single agent reinforcement learning, non-cooperative game and centralized optimization model, the proposed method can better reflect the operation characteristics of the coexistence of heterogeneous agents and the interweaving of local game and global feedback in the electricity market. The research also shows that the combination of game relationship modeling and computer iterative solution mechanism is helpful to improve the interpretability and simulation adaptation ability of market behavior analysis. The work of this paper provides a new analysis idea for the behavior modeling of electricity market players, and also lays a method foundation for the subsequent research on intelligent decision-making in more complex scenarios.

About the Author

Zhenduo Gao, born in July 1989, male, Han ethnicity, from Changchun City, Jilin Province, graduated from Northeast Electric Power University with a Master's degree in Electrical Engineering and holds the title of Senior Engineer. His primary research focuses on functional design, development, operation, and maintenance management of power trading platforms. Email: 15164356699@163.com

Zheyong Piao, born in October 1975, is a male of Korean ethnicity from Tonghua City, Jilin Province. He graduated from Northeast Electric Power University with a Master's degree in Project Management and holds the title of Senior Engineer. His primary research focuses on the development and operation of electricity markets, as well as power system stability analysis.

Tao Meng, born in September 1990, male, Han ethnicity, from Shaoyang City, Hunan Province, graduated from Northeast Electric Power University with a doctoral degree and holds the title of Senior Engineer. His primary research focuses on power system stability

analysis and the development and operation of power markets.

Le Liu, born in August 1980, is a female of Han ethnicity from Jilin City, Jilin Province. She graduated from Dalian University of Technology with a Ph.D. in Engineering and holds the title of Senior Engineer. Her primary research focuses include power grid dispatching and operational control, as well as electricity market trading and settlement.

Acknowledgments

State Grid Jilin Electric Power Co., Ltd. Science and Technology Project(2025ZX-15)

References

[1] Huang W, Li H. Game theory applications in the electricity market and renewable energy trading: A critical survey[J]. Frontiers in Energy Research, 2022, 10: 1009217.

[2] Tsaousoglou G, Giraldo J S, Paterakis N G. Market mechanisms for local electricity markets: A review of models, solution concepts and algorithmic techniques[J]. Renewable and Sustainable Energy Reviews, 2022, 156: 111890.

[3] Li J, Ai Q, Yin S, et al. An aggregator-oriented hierarchical market mechanism for multi-type ancillary service provision based on the two-loop Stackelberg game[J]. Applied Energy, 2022, 323: 119644.

[4] Lee W P, Han D, Won D. Grid-oriented coordination strategy of prosumers using game-theoretic peer-to-peer trading framework in energy community[J]. Applied Energy, 2022, 326: 119980.

[5] Shojaabadi S, Talavat V, Galvani S. A game theory-based price bidding strategy for electric vehicle aggregators in the presence of wind power producers[J]. Renewable energy, 2022, 193: 407-417.

[6] Zheng B, Wei W, Chen Y, et al. A peer-to-peer energy trading market embedded with residential shared energy storage units[J]. Applied Energy, 2022, 308: 118400.

[7] Liu Y, Jiang Z, Guo B. Assessing China's provincial electricity spot market pilot operations: Lessons from Guangdong province[J]. Energy Policy, 2022, 164: 112917.

[8] Chandrakala K R M V, Kiran P. Multi-agent based modeling and learning approach for intelligent day-ahead bidding strategy in wholesale electricity market[J]. Expert Systems with Applications, 2023, 233: 121014.

[9] Lei Z, Liu M, Shen Z, et al. A data-driven Stackelberg game approach applied to analysis of strategic bidding for distributed energy resource aggregator in electricity markets[J]. Renewable Energy, 2023, 215: 118959.

[10] Lei Z, Liu M, Shen Z, et al. A Nash–Stackelberg game approach to analyze strategic bidding for multiple DER aggregators in electricity markets[J]. Sustainable Energy, Grids and Networks, 2023, 35: 101111.

[11] Wu C, Gu W, Yi Z, et al. Non-cooperative differential game and feedback Nash

equilibrium analysis for real-time electricity markets[J]. International Journal of Electrical Power & Energy Systems, 2023, 144: 108561.

[12] Teng M, Lv K, Han C, et al. Trading behavior strategy of power plants and the grid under renewable portfolio standards in China: A tripartite evolutionary game analysis[J]. Energy, 2023, 284: 128398.

[13] Xie J, Guan B, Yao Y, et al. Market power risk prevention mechanism of China's electricity spot market based on stochastic evolutionary game dynamics[J]. Frontiers in Energy Research, 2023, 11: 1270681.

[14] Zhang X, Guo X, Zhang X. Bidding modes for renewable energy considering electricity-carbon integrated market mechanism based on multi-agent hybrid game[J]. Energy, 2023, 263: 125616.

[15] Zheng Y, Wang Y, Yang Q. Bidding strategy design for electric vehicle aggregators in the day-ahead electricity market considering price volatility: A risk-averse approach[J]. Energy, 2023, 283: 129138.

[16] Mohsenzadeh-Yazdi H, Kebriaei H, Aminifar F. Multi-agent reinforcement learning in a new transactive energy mechanism[J]. IET Generation, Transmission & Distribution, 2024, 18(18): 2943-2955.

[17] Li Y, Yang Y, Zhang F, et al. A Stackelberg game-based approach to load aggregator bidding strategies in electricity spot markets[J]. Journal of Energy Storage, 2024, 95: 112509.

[18] Tsao Y C, Ai H T T, Lu J C, et al. Game theory-based electricity pricing decisions incorporating prosumer energy preferences and renewable portfolio standard[J]. Energy, 2024, 306: 132418.

[19] Chen L, Ye Q, Wu X, et al. Stackelberg game-based optimal electricity trading method for distribution networks with small-micro industrial parks[J]. Frontiers in Energy Research, 2024, 12: 1348823.

[20] Wu X, Ye Q, Chen L, et al. Electricity market clearing for multiple stakeholders based on the Stackelberg game[J]. Frontiers in Energy Research, 2024, 12: 1342516.