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

Dynamic deduction of market behavior of electricity price formation mechanism in electricity market simulation system

Tao Meng[1,*], *Le Liu*[2], *Mingyu Che*[2] *and Zhenduo Gao*[2]

[1] State Grid Jilin Electric Power Co, Ltd., Changchun 130000 Jilin Province, China
[2] Jilin Electric Power Trading Center Co., Ltd., Changchun 130021, Jilin Province, China

SUMMARY: *Under the background of increasing volatility of new energy and complex trading behavior, the formation of electricity price in power market has shown significant dynamic coupling characteristics. Aiming at the problem of insufficient description of the linkage between price generation process and agent behavior in traditional methods, this paper constructs a dynamic deduction method of market behavior of electricity price formation mechanism for electricity market simulation system, which integrates market data collection, scenario construction, feature representation, state modeling, multi-agent game and intelligent deduction mechanism into a unified computing framework. Based on Python 3.11 and PyTorch 2.2 platform, the experiment was carried out with continuous samples of 180 d and 15 min granularity. The results show that the root mean square error of the electricity price in this paper is 4.87 yuan /MWh, the average absolute percentage error is 3.96%, the price turning point capture rate is 88.4%, the market clearing rate is 94.6%, and the supply-demand deviation rate is 2.3%. The results show that the proposed method can better reveal the internal relationship between electricity price evolution and market behavior adjustment, and provide computational support for electricity market simulation analysis and rule optimization.*

KEYWORDS: *Electricity market simulation; Electricity price formation mechanism; Multi-agent game; Dynamic deduction*

1 Introduction

Under the background of the continuous evolution of the new power system, the fluctuation of new energy output, the enhancement of load response and the increase of market transaction frequency are jointly changing the operation logic of electricity price formation. The price in the electricity market is no longer a static result of supply and demand, but a dynamic variable evolving under the joint action of unit constraints, marginal costs, trading rules, bidding strategies, load elasticity and external disturbances. Especially in the simulation environment, there is a continuous game relationship between power generators, electricity sellers, large users, load aggregators and market operators, and the responses of different agents to historical returns, real-time prices and expected risks will further amplify the linkage effect of market behavior, making the electricity price formation mechanism show significant time-varying, coupling and nonlinear characteristics [1-5]. In this case, if the single-round equilibrium analysis or static rules are still used to infer the market results, it is often difficult to explain the practical problems such as price sudden changes, strategy deviation and local

*13578780564@163.com
https://doi.org/10.65102/is20261048
Published: 10 June, 2026

imbalance.

With the development of computer simulation technology, intelligent optimization method and multi-agent learning framework, electricity market research has begun to shift from "result calculation" to "process deduction". Based on the simulation system, the dynamic behavior of market entities can be described, and the quotation reporting, market clearing, profit feedback and strategy modification can be integrated into the unified computing link, and the price response path and behavior evolution law under different mechanisms can be tested in a repeatable experimental environment. Existing researches have made progress in multi-agent reinforcement learning, bidding modeling, market game analysis and electricity price prediction, etc. However, most methods focus on single bidding optimization or short-term price prediction, and the mechanistic description of "how electricity price is gradually shaped, revised and stabilized in multi-agent interaction" is still insufficient [6-10]. At the same time, although some models have strong predictive ability, they lack the state expression and intelligent deduction design oriented to the simulation system, and it is difficult to give consideration to the requirements of interpretability, transferability and policy evolution analysis [11-14].

The value of electricity market simulation system is not only to reproduce the transaction results, but also to reveal the dynamic transmission process between price and behavior. The combination of market data collection, state coding, multi-agent game and intelligent deduction mechanism is helpful to reconstruct the internal chain formed by electricity price in the computer environment, and identify the influence boundaries of different agent strategy adjustment on market efficiency, price fluctuation and equilibrium stability [15-21]. Based on this understanding, this paper focuses on the problem of electricity price formation mechanism and market behavior dynamic deduction in the power market simulation system.

RQ1: Can the electricity market simulation system more accurately depict the formation process of electricity price under multi-agent interaction conditions and reflect the stage characteristics of price fluctuations?

RQ2: Can the coupling of state modeling, multi-agent game and intelligent inference mechanism improve the stability and interpretability of market behavior evolution analysis?

RQ3: Compared to traditional static analysis methods, can the dynamic deduction framework more effectively reveal the impact of strategy adjustment on market clearing results, price response and behavioral convergence?

Research objectives: Aiming at the shortcomings of the existing research in the connection between dynamic mechanism characterization and simulation deduction, this paper constructs a dynamic deduction method of market behavior for electricity market simulation system, which integrates market scene generation, price state representation, multi-agent game solution and intelligent deduction mechanism into a unified modeling framework, in order to reveal the dynamic evolution law of electricity price formation mechanism from the perspective of computer simulation. It also provides computable and verifiable technical support for market design optimization and subject behavior analysis.

2 Related Research

Focusing on the price formation and subject behavior analysis in the electricity market, the existing research has been carried out from multiple directions, such as bidding optimization, market simulation, price prediction and game modeling. Wu et al. proposed a strategic bidding method based on multi-agent transfer learning and reinforcement learning, which can improve the adaptability of the agent's bidding decision in the competitive market [1]. Namalomba et al. adopt bi-level programming and Q-learning to build a centralized trading market simulation framework, so that trading entities can continuously revise their strategies in feedback [2]. This kind of research shows that with the help of computer simulation and learning algorithms, the behavior adjustment process of market players in repeated games can be better described. Aliabadi et al. further used deep Q network to analyze the market competition under the condition of artificial intelligence participation, and pointed out that after the learning ability of agents is enhanced, the market price may show stronger strategic sensitivity [3]. Xu et al. incorporated the retailer's joint bidding and pricing process into the multi-task deep reinforcement learning framework to improve the collaborative optimization ability under the condition of multiple decision variables [4].

From the perspective of market mechanism and game modeling, Huang et al. systematically sorted out the application path of game theory in electricity market and renewable energy trading, and showed that revenue distribution, strategy interaction and rule design are always important pivots of market analysis [5]. Jain et al. constructed a multi-agent strategic bidding simulator for day-ahead market, so that power generation agents could complete strategy search and profit comparison in the simulation environment [6]. Zhang et al. introduced the electric-carbon coupled trading mechanism into the multi-agent hybrid game model, expanding the constraint dimension of market behavior research [7]. Li et al. focus on the influence of historical returns on subsequent bidding strategies, indicating that the subject's behavior is not only determined by the current market state, but also continuously driven by memory information and expected returns [8]. See Table 1 for a comparison of related studies.

Table 1: Comparison of studies related to price formation and behavior analysis in electricity markets

Reference	Research Objective	Main Method	Key Feature	Main Limotbebox लिमitation
Wu et al. [1]	Improve strategic bidding capability in competitive markets	Multi-agent transfer learning + reinforcement learning	Strong strategy adaptability	Relatively limited explanation of market mechanisms
Namalomba et al. [2]	Construct a simulation framework for centralized electricity trading markets	Bi-level programming + Q-learning	Enables feedback-driven strategy updating	Scenario complexity remains relatively limited
Aliabadi et al. [3]	Analyze the impact of intelligent algorithms on market competition	Deep Q-network	Reveals the sensitivity of prices to strategy learning	Insufficient discussion of multi-agent co-evolution
Xu et al. [4]	Optimize joint bidding and pricing for retailers	Multi-task deep reinforcement learning	Suitable for multivariable decision-making scenarios	Focuses more on agent optimization while weakening system-level simulation
Jain et al. [6]	Simulate the strategic bidding process in the day-ahead market	Multi-agent simulator	Facilitates comparison of returns under different bidding strategies	Insufficient depth in state representation and behavioral evolution
Tang et al. [9]	Predict individual bidding behavior in real electricity markets	Machine learning framework	Strengthens bidding prediction capability	Focuses more on prediction and lacks mechanism-oriented simulation

In terms of electricity price prediction and simulation application, Tang et al. used the machine learning framework to predict individual quotes in the real market, which provided data support for the microscopic behavior recognition of price formation [9]. Liu et al. studied the optimal bidding strategy in real-time multi-party market based on deep reinforcement learning, which enhanced the responsiveness of the model to short-term load disturbances [10]. Harder et al. emphasize that to truly depict the operation of wholesale electricity market, it is still not enough to only rely on simplified price function or single-agent optimal model, and it is necessary to retain agent heterogeneity, rule constraints and dynamic feedback in the simulation system at the same time [11]. Nitsch et al. applied machine learning to electricity price prediction under simulated energy market scenarios, indicating that data-driven methods have advantages in price trend identification [12]. However, simply predicting prices does not fully explain how prices are gradually pushed up, down, or stabilized by agent interactions.

In general, the existing researches have provided an important foundation for the computer modeling of electricity market, but there are still two shortcomings. One is that the research focuses on bidding optimization or price prediction, and the description of the coupling chain between electricity price formation mechanism and market behavior is not complete. Although the other kind of research introduces multi-agent game and simulation technology, it often deals with state representation, market clearing, payoff feedback and intelligent deduction separately, which leads to that the model still has room for improvement in behavior evolution analysis and mechanism explanation. Based on this, this paper intends to build a unified analysis framework of "state modeling-game interaction-intelligent deduction-result feedback" in the electricity market simulation system, so as to reveal the dynamic evolution process of market behavior under the formation mechanism of electricity price in more detail.

3 Electricity price formation mechanism and market behavior dynamic deduction method in electricity market simulation system

3.1 Market data collection and simulation scenario construction

The premise of dynamic deduction of electricity price formation mechanism is to transform the market information scattered in different business links into a unified data stream that can be calculated, interactive and replayed. The power market simulation system is not only a simple reproduction of historical price series, but also needs to simultaneously depict the linkage relationship among load change, unit output, quotation declaration, market clearing, demand response and rule constraints. Based on this consideration, this paper designs market data collection and scenario construction as an integrated process, and forms a pre-link of "raw data access, feature collation, state mapping, and scene generation" in the computer environment, which provides basic input for subsequent electricity price state modeling and multi-agent behavior deduction. The overall process is shown in Figure 1.

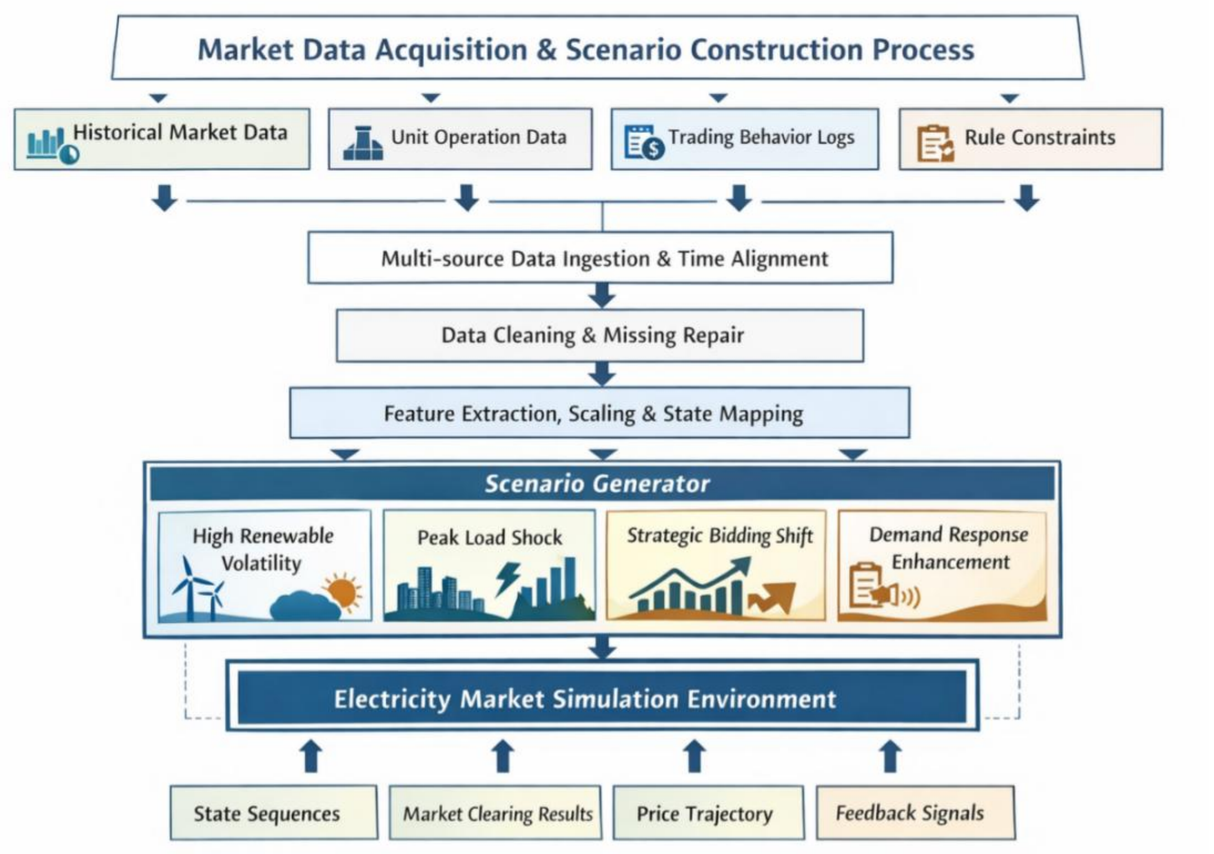

Figure 1: Process of market data collection and simulation scenario construction

The data collected in this study mainly includes four categories. The first is market operation data, covering time-sharing load, node electricity purchase and sale demand, historical clearing price and transaction electricity. The second is power side data, including thermal power unit capacity limit, climbing rate, marginal cost, start and stop state, as well as wind power and photovoltaic forecast output. The third is the transaction behavior data, including the quotation range of power generators, the power purchase declaration of power sales companies, the proportion of large users' transferable load and the response records of load aggregators. The fourth is the rule constraint data, including reserve requirements, clearing time series, price upper and lower limits and deviation assessment parameters. In order to maintain the consistency of the simulation time scale, this paper uniformly maps all the data into a 15-min granularity, forming a total of 180 d continuous samples and obtaining 17 280 time period records. Missing values were imputed by adjacent time periods, outliers were corrected by quantile censoring method, and different dimensional variables were normalized to avoid the dominant effect of large-scale variables on model training. The standardization process of continuous variables is expressed as follows.

$$x_t^* = \frac{x_t - x_{min}}{x_{max} - x_{min}} \tag{1}$$

where, x_t is the original observation value at time period t, x_{min} and x_{max} represent the minimum and maximum value of the variable in the training set, respectively, and x_t^* is the normalized input result.

After data cleaning, this paper further organizes the multi-source variables into power market simulation state vectors, which are used to describe the overall environment of market operation in a given period of time. Let the market state in period t be s_t, then:

$$s_t = [d_t, \hat{p}_t^w, \hat{p}_t^s, g_t, c_t, b_t^g, b_t^r, r_t, m_t] \tag{2}$$

Among them, d_t represents system load demand, $\hat{p}_t^w$ and $\hat{p}_t^s$ represent the forecast output of wind power and PV respectively, g_t represents the available capacity of conventional units, c_t represents the marginal cost vector, b_t^g represents the price set of generation side, b_t^r represents the declaration set of power purchase side, r_t represents the state of demand response resources, and m_t represents the parameters of market rules. This state representation can preserve both physical operation constraints and transaction behavior information, so that the formation of electricity price is no longer regarded as a numerical output at a single time point, but is included in the continuous evolution of the environment state.

In order to make the simulation system cover different market scenarios, this paper constructs four extended environments based on the basic samples: high volatility new energy scenario, load peak scenario, strategic pricing scenario and demand response enhancement scenario. The scene perturbation is generated by parameter injection, that is, the controlled perturbation term is superimposed on the original sequence to obtain the scene input:

$$\tilde{x}_t = x_t(1 + \varepsilon_t), \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma_k^2) \tag{3}$$

where, $\tilde{x}_t$ is the variable value after scene perturbation, ε_t is the disturbance term following normal distribution, and σ_k represents the fluctuation intensity corresponding to the KTH type of scene. By controlling the disturbance amplitude and effect object, the typical situations such as new energy deviation expansion, electricity demand jump, main strategy deviation and flexible load release can be more stably reproduced.

In the scenario operation phase, the system needs to meet the market supply and demand balance and the basic operation security constraints, and its core balance relationship can be written as follows.

$$\sum_{i=1}^{N} p_{i,t}^g + p_t^{imp} + p_t^{dr} = d_t + p_t^{ch} + r_t^{res} \tag{4}$$

Here, $p_{i,t}^g$ are the cleared power of the ith unit at time period t, p_t^{imp} is the external purchased power, p_t^{dr} is the power released in demand response, p_t^{ch} is the energy storage charging power, and r_t^{res} is the reserve demand. Equation (4) ensures that the price formation and transaction results in the simulation environment are always established within the feasible operating boundary, so as to avoid false deduction results that deviate from market constraints.

Table 2 shows the core data composition of the simulation scene. As can be seen from Table 2, this paper does not treat the electricity price only as an outcome variable, but puts the transaction price, quotation behavior, cost structure and system constraints into the unified data framework. This design helps the subsequent model to simultaneously identify the questions "where does the price come from" and "how does the behavior change" during the learning process.

Table 2: Core data composition of simulation scenario

Data Category	Main Variables	Time Granularity	Data Function
Market Operation Data	System load, nodal demand, historical clearing price, traded electricity volume	15 min	Reflects the basic supply–demand status of the market
Power-Side Data	Thermal power capacity, marginal cost, ramp rate, forecasted wind and solar output	15 min	Characterizes generation capability and cost boundaries
Trading Behavior Data	Generation bids, electricity purchase declarations, demand response records, transferable load ratio	15 min	Supports agent behavior modeling and strategy updating
Rule Constraint Data	Price caps and floors, reserve requirements, deviation assessment, clearing rules	Intraday update	Constrains market clearing and electricity price formation
Disturbance Scenario Data	Renewable energy fluctuation parameters, load shock parameters, strategy deviation parameters	Scenario level	Constructs a dynamic simulation environment

3.2 Feature representation and state modeling of electricity price formation mechanism

After the completion of market data collection and simulation scenario construction, the key problem is not to simply stack various variables into the system, but to extract effective features that can characterize the formation process of electricity price from multi-source, heterogeneous and continuously fluctuating data. The price generation in the electricity market is not a direct mapping of the single supply and demand balance, but a comprehensive result jointly shaped by the cost boundary of the generation side, the declaration behavior of purchasing and selling electricity, the deviation of new energy output, the release degree of demand response and the constraints of market rules. Therefore, if the model is only based on the historical price series, it is easy to weaken the strategic driving and constraint transmission behind the price. Based on this, this paper divides the feature representation of the electricity price formation mechanism into four levels: supply feature, demand feature, behavior feature and rule feature. On this basis, the state modeling method for time series deduction is constructed, and its structure is shown in Figure 2.

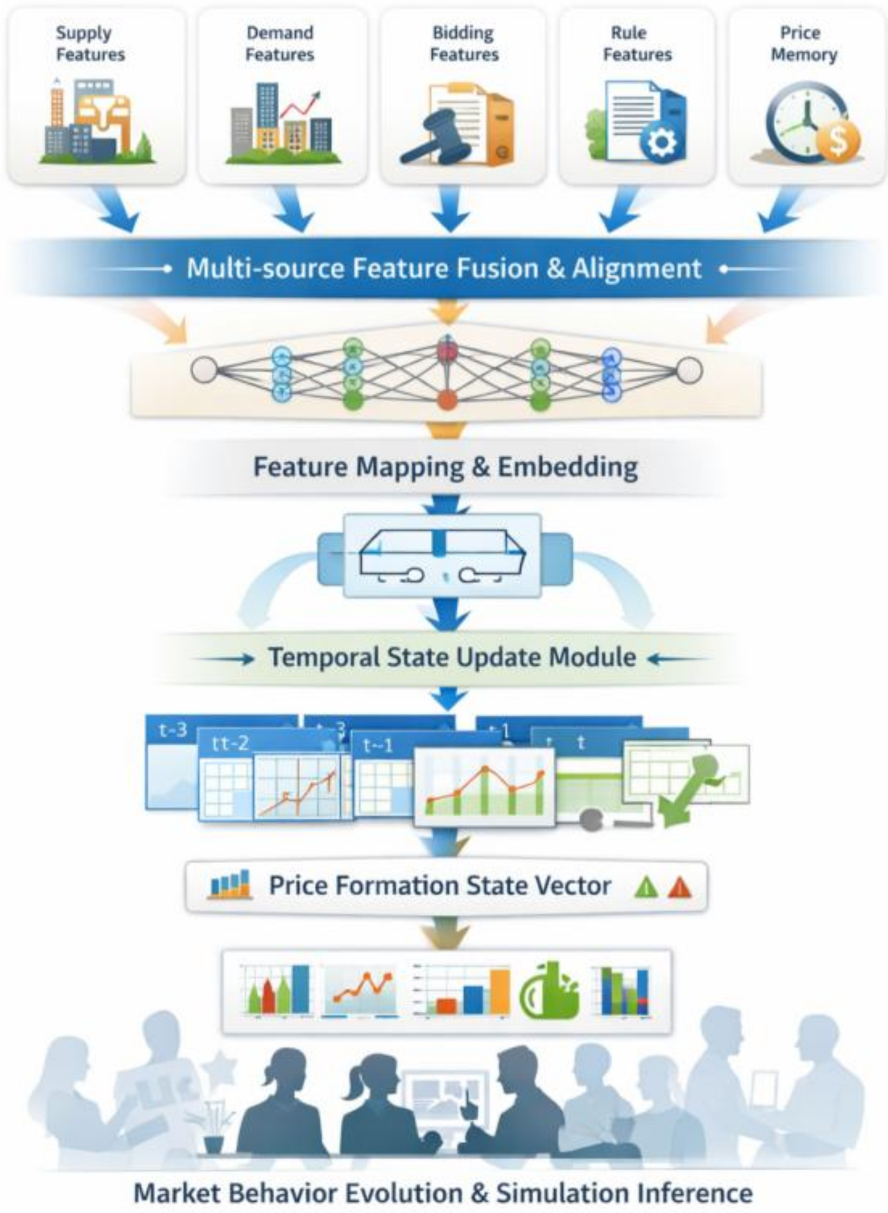

Figure 2: Feature representation and state modeling structure of electricity price formation mechanism

Let the original eigenvector of time period t be x_t, then it can be expressed as follows.

$$x_t = \left[g_t^c, g_t^r, d_t, b_t^s, b_t^b, u_t^{dr}, m_t, p_{t-1}\right] \tag{5}$$

Among them, g_t^c represents the combination characteristics of available output and marginal cost of conventional units, g_t^r represents the forecast output characteristics of new energy such as wind power and photovoltaic, d_t represents the load demand characteristics, b_t^s and b_t^b represent the seller's quotation and buyer's declaration characteristics respectively, u_t^{dr} represents the demand response state, m_t represents the rule constraint characteristics, and p_{t-1} represents the price information of the last period. The expression integrates physical operation, trade declaration and market memory into the unified input space at the same time, so that the model can identify the source of price change from a more complete mechanism chain.

Considering that the economic significance of different variables is not consistent with the time sensitivity, this paper does not directly use the original vector to enter the deduction model, but complete the low-dimensional embedding through the feature mapping layer to obtain a compact representation of the time period t:

$$e_t = \sigma(W_e x_t + b_e) \tag{6}$$

where, W_e is the feature mapping matrix, b_e is the bias term, $\sigma(\cdot)$ is the nonlinear activation function, e_t is the encoded market representation vector. The effect of this process is to compress redundant information, highlight the formation of more sensitive variable combinations to electricity prices, and reduce the impact of high-dimensional inputs on the stability of subsequent time series modeling.

The formation of electricity price has obvious characteristics of time sequence conduction. The price in a certain period not only depends on the current supply and demand and the offer state, but also is continuously affected by the previous transaction results, the profit expectation of the subject and the clearing deviation. In order to characterize this inter-period dependence, a recursive state update mechanism is introduced to jointly input the coding result e_t and the hidden state h_{t-1} of the previous period to form the current hidden state h_t of the market:

$$h_t = \tanh(W_h e_t + U_h h_{t-1} + b_h) \tag{7}$$

Here, W_h and U_h are the state transition parameters, and b_h is the bias term. The obtained h_t is no longer a simple snapshot of a static time point, but a dynamic state representation integrated with the historical market trajectory, which is more suitable for subsequent service subject behavior deduction and price response analysis.

Considering that the electricity price often shows short-time abrupt characteristics under the circumstances of peak load, rapid fluctuation of new energy and strategic declaration, only relying on sequential recurrence may still weaken the impact of critical periods. Based on this, this paper further introduces the time attention weight to reconstruct the historical state in the neighbor window and obtain the enhanced representation z_t:

$$\alpha_{t,k} = \frac{\exp(h_t^{\top} h_k)}{\sum_{j=t-L}^{t} \exp(h_t^{\top} h_j)}, \qquad z_t = \sum_{k=t-L}^{t} \alpha_{t,k} h_k \tag{8}$$

where $\alpha_{t,k}$ is the contribution weight of time period k to the current time period t, and L is the length of the backtracking window. In this way, the model can automatically increase the proportion of the key fluctuation range in the state representation, so as to enhance the recognition ability of the price abrupt change precursor and behavior deviation signal. On this basis, this paper defines the state of electricity price formation as follows.

$$s_t = [h_t, z_t, \Delta d_t, \Delta g_t, \Delta b_t] \tag{9}$$

Here, Δd_t represents the load change rate, Δg_t represents the adjustable output force change rate, and Δb_t represents the quoted offset strength. The state vector preserves the long-term evolution information, local key memory and the characteristics of immediate disturbance, and can better reflect under which supply-demand relationship, strategy environment and rule constraints the market price is gradually pushed up, suppressed or stabilized.

3.3 Dynamic deduction model of market behavior based on multi-agent game

After completing the feature representation and state modeling of the electricity price formation mechanism, a more mechanistic question needs to be further answered, that is, how exactly price changes are generated, amplified and gradually stabilized in the continuous interaction of multiple types of market players. In the electricity market, power producers, electricity selling companies, large users and load aggregators do not make decisions in isolation, but under the constraints of uniform market rules, and continuously adjust their strategies according to price signals, profit feedback and competitor behavior. Therefore, this paper constructs a dynamic deduction model of market behavior based on multi-agent game in the simulation system, and brings the quotation declaration, electricity purchase adjustment, demand response release and market clearing results into the same interactive framework to realize the continuous deduction of market behavior evolution path. Let the dynamic game system of electricity market be expressed as follows.

$$\mathcal{G} = (\mathcal{N}, \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}) \tag{10}$$

where, N is the set of market entities, including power generators, electricity selling companies, large users and load aggregators. S is the market state space, which corresponds to the price formation state vector constructed in Section 3.2. A is the main action space, which is used to describe the quotation, power purchase declaration, load adjustment and response strategy. P is the state transition function; R is the revenue feedback function. The above definitions show that the model constructed in this paper no longer regards the electricity price as a static equilibrium solution, but as a phased result of the joint action of multiple agents in a continuous game. For any time period t, the ith agent outputs the strategic action $\mathrm{a_{n,t}}$ according to the current market state $\mathrm{s_t}$, then the joint action can be written as follows.

$$\mathrm{a_t} = \left[\mathrm{a_{1,t}}, \mathrm{a_{2,t}}, \cdots, \mathrm{a_{n,t}}\right] \tag{11}$$

Among them, the actions of power producers are mainly reflected in the segmented quotation and electricity declaration adjustment, the actions of power sales companies are reflected in the power purchase and power purchase price correction, the large users are reflected in the power time transfer and demand reduction decision, and the load aggregators are reflected in the adjustable load release proportion and response time selection. After the joint action $\mathrm{a_t}$ is input into the market clearing module, the system calculates the transaction result and clearing price of the period according to the matching relationship between supply and demand and the rule constraints. The corresponding price generation relation can be expressed as follows.

$$\mathrm{p_t} = \Phi(\mathrm{s_t}, \mathrm{a_t}, \mathrm{m_t}) \tag{12}$$

where $\Phi(\cdot)$ represents the market clearing mapping function and $\mathrm{m_t}$ is the rule constraint parameter of time period t. The function of the mapping function is to link the subject strategy, system state and market rules, so that the price not only reflects the marginal relationship between supply and demand, but also retains the regulatory effect of institutional constraints on trading results.

In the aspect of revenue characterization, this paper defines the return of the subject in a single period as the comprehensive result of "market revenue, operating cost and deviation penalty". For the generation side agent, its revenue function can be written as follows.

$$r_{i,t} = p_t q_{i,t} - C_i(q_{i,t}) - \lambda_i \left| q_{i,t} - \hat{q}_{i,t} \right| \tag{13}$$

Here, $q_{i,t}$ is the actual transaction quantity of the subject, $C_i(\cdot)$ is the cost function, $\hat{q}_{i,t}$ is the declared quantity, and λ_i is the deviation penalty coefficient. For the power purchase side subject and load aggregator, the power purchase expenditure, response compensation and default cost can be integrated into the same income structure for unified representation. The design can ensure that different types of agents participate in the market interaction with the goal of maximizing profit, and avoid the strategy deduction from the actual trading constraints.

In order to reflect the dynamic evolution characteristics of market behavior, this paper adopts the strategy update mechanism based on state feedback. Let the policy function of the third agent under the parameters θ_i be π_{θ_i}, then its action generation process is as follows.

$$a_{i,t} = \pi_{\theta_i}(s_t) \tag{14}$$

After completing the market clearing and obtaining the return feedback, the agent modifies the strategy parameters according to the immediate return and the expected future return, and the cumulative objective function is as follows.

$$J_i(\theta_i) = \mathbb{E}\left[\sum_{t=0}^{T} \gamma^t r_{i,t}\right] \tag{15}$$

Here, γ is the discount factor. By repeatedly executing the cyclic process of "state observation, strategy generation, market clearing, profit return and parameter update" in the simulation system, the bidding behavior and response decision of different agents will gradually show obvious adaptive adjustment, so as to form a relatively stable market behavior trajectory in the long-term game.

Due to the significant strategy coupling in the multi-agent market, the change of the payoff of a single agent often causes the synchronous response of other agents. Therefore, this paper introduces the joint action influence mechanism in the state transition, and expresses the market state in the next time period as follows.

$$s_{t+1} = f(s_t, a_t, \xi_t) \tag{16}$$

where $f(\cdot)$ is the environmental transfer function, ξ_t is the external disturbance term such as new energy fluctuation and load deviation. Equation (16) shows that market behavior inference is not a local static optimal problem, but a dynamic process influenced by historical states, joint strategies and random disturbances. Through this setting, the simulation system can more completely reproduce the centralized quotation adjustment in the price increase stage, the response expansion under the peak load, and the strategy convergence process after the price drops. Figure 3 illustrates the dynamic deduction structure of market behavior based on the multi-agent game.

Figure 3: Dynamic deduction structure of market behavior based on multi-agent game

3.4 Design of intelligent deduction mechanism for electricity market simulation

After the state modeling of the electricity price formation mechanism and the construction of the multi-agent game deduction framework, there is a further problem to be solved at the implementation level, that is, how to complete the closed-loop operation of market state perception, strategy generation, price clearing, profit feedback and behavior modification in a continuous period of time. Without this closed-loop mechanism, the above state representation and game model can only stay at the theoretical level, and it is difficult to support large-scale, multi-round, and replayable dynamic deduction process. Based on this, this paper designs an intelligent deduction mechanism for power market simulation environment, which connects the state coding module, joint decision module, market clearing module and feedback learning module in series to form an adaptive operation link of "observation-deduction-correction-deduction", so that the evolution of market behavior can continue in the computer system.

Let the market state in time period t be s_t. After receiving this state, the intelligent pusher generates strategic actions for each agent. In order to balance the continuous decision-making accuracy and exploration ability, this paper introduces a disturbance term into the policy output, and the action generation process is expressed as follows.

$$\tilde{a}_{i,t} = \mu_{\theta_i}(s_t) + \epsilon_t \tag{17}$$

Here, μ_{θ_i} is the policy network of the ith agent and ϵ_t is the restricted random disturbance. The role of this design is to avoid premature convergence of the deducing system to a local strategy, so that the model can still retain the necessary search ability in the face of new energy fluctuations, load spikes, and competitive offers.

After the joint action is submitted to the market environment, the system calculates the transaction power, the clearing price and the profit of each subject according to the uniform clearing rules, and writes these results back to the experience buffer. In order to enhance the ability of intelligent inference to identify price volatility and behavioral deviation, the single-period return is designed as a weighted combination of return, stability and deviation penalty:

$$R_t = \omega_1 \Pi_t - \omega_2 V_t - \omega_3 D_t \tag{18}$$

Here, Π_t represents the total return during the period, V_t represents the intensity of price volatility, D_t represents the degree of supply and demand deviation or strategy deviation, and ω_1、ω_2、ω_3 are the weight coefficients. Through this reward function, the intelligent deduction mechanism no longer only pursues the maximum local profit, but maintains a balance between economy, stability and market coordination.

In the parameter update stage, this paper uses the dual-value network to estimate the next period reward, so as to weaken the interference of overestimation phenomenon on policy learning. The target value can be written as follows:

$$y_t = R_t + \gamma \min_{j=1,2} Q_{\bar{\phi}_j}\left(s_{t+1}, \mu_{\bar{\theta}}(s_{t+1}) + \varepsilon_t{}'\right) \tag{19}$$

Here, γ is the discount factor, $Q_{\bar{\phi}_j}$ is the target value network, $\mu_{\bar{\theta}}$ is the target policy network, and $\varepsilon_t{}'$ is the smoothing noise term. The significance of Equation (19) is to use the smaller values in the dual network to construct the training target, so as to improve the robustness of the value estimation in the deduction process and avoid excessive deviation of the strategy due to short-term price anomalies. Accordingly, the value network loss function is defined as follows.

$$L(\phi_j) = \frac{1}{N}\sum_{t=1}^{N}\left(Q_{\phi_j}(s_t, a_t) - y_t\right)^2 \tag{20}$$

where, N is the batch sample number. By minimizing this loss, the system can gradually improve its ability to fit market feedback. The policy network is updated according to the value gradient, and its objective function is expressed as follows.

$$\nabla_\theta J(\theta) = \frac{1}{N}\sum_{t=1}^{N} \nabla_a Q_\phi(s_t, a)|_{a=\mu_\theta(s_t)} \nabla_\theta \mu_\theta(s_t) \tag{21}$$

This update method makes the strategy adjustment always focus on the direction of expected revenue improvement, and can continuously revise the behavior of each agent according to the real feedback after the market clears.

Considering that the electricity market simulation is a continuous sequential system, it is easy to cause learning oscillation if the target network is completely replaced by the latest parameters in each round. Therefore, this paper further adopts the soft update mechanism to keep the training process stable, which is in the form of:

$$\bar{\theta} \leftarrow \tau\theta + (1-\tau)\bar{\theta}, \qquad \bar{\phi} \leftarrow \tau\phi + (1-\tau)\bar{\phi} \tag{22}$$

Here, τ is the soft update coefficient. Through this mechanism, the intelligent deducer can maintain the continuity of parameter evolution while retaining new information, so it is more suitable for market trajectory deduction in long time scales.

Based on the above design, this paper summarizes the intelligent deduction process as follows: the market representation is extracted by the state coding layer, and then the main action is generated by the strategy network. The electricity price and transaction result are calculated by the market clearing module, and then the income feedback and volatility information are transmitted back to the learning module to complete the parameter update, and the revised strategy is used for the next time period. In this way, the simulation system can output the price trajectory, strategy change and market stability indicators on the continuous time axis, and realize the dynamic tracking of the formation process of electricity price. Figure 4 illustrates the structure of the intelligent deduction mechanism for electricity market simulation.

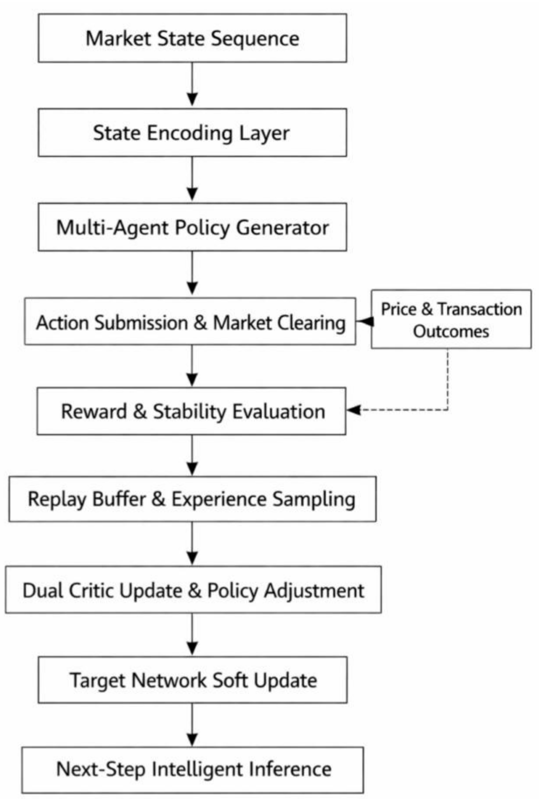

Figure 4: Structure of intelligent deduction mechanism for electricity market simulation

4 Results and discussion

4.1 Experimental environment and parameter setting

In order to test the effectiveness of the electricity market simulation system constructed in this paper in describing the electricity price formation mechanism and the dynamic deduction of market behavior, this paper completed the model training, strategy deduction and result evaluation in a unified computer experiment environment. The experimental platform is implemented by Python 3.11, and the deep learning framework is PyTorch 2.2. The simulation module is composed of a custom electricity market environment, a state coding module, a multi-agent game module and an intelligent inference module, which can complete quotation submission, market clearing, revenue feedback and strategy update on the same timeline. The hardware environment is configured with Ubuntu 22.04 operating system, Intel Core i7-12700 processor, 32 GB RAM and NVIDIA RTX 4080 16 GB GPU. This configuration can meet the needs of multi-round market scenario playback and continuous strategy training, and provide stable support for repeatable verification of subsequent results.

The continuous simulation sample constructed in Section 3.1 is used for the experimental data, which contains 180 d and 15 min granularity market operation records, totaling 17 280 time periods. In order to ensure the independence of training, validation and testing processes, this paper divides the sample sets according to the proportion of 70%, 15% and 15%, which are used for parameter learning, hyperparameter correction and generalization ability test respectively. The input variables cover information such as system load, new energy forecast output, unit available capacity, marginal cost, quotation declaration, demand response status and rule constraints, and the output results correspond to clearing price, transaction electricity, subject income and behavior evolution indicators. See Table 3 for the experimental environment and data Settings.

Table 3: Experimental environment and data Settings

Category	Specific Content
Operating System	Ubuntu 22.04
Programming Language	Python 3.11
Deep Learning Framework	PyTorch 2.2
CPU	Intel Core i7-12700
Memory	32 GB
GPU	NVIDIA RTX 4080 16 GB
Simulation Time Granularity	15 min
Sample Duration	180 d
Total Sample Size	17,280 time-interval records
Data Split	70% training set, 15% validation set, 15% test set
Input Variables	Load, wind and solar forecasts, generation unit capacity, marginal cost, bids, response status, rule parameters
Output Variables	Clearing price, traded electricity volume, revenue feedback, behavioral evolution results

In the process of model training, we use the experience replay and soft update mechanism to maintain the stability of the inference, and use the dual-value network to alleviate the overestimation problem under the condition of price volatility. Considering that the main strategy of the power market has the characteristics of continuous adjustment, the action space is set as a continuous variable, covering the correction coefficient of power quotation, the adjustment proportion of power purchase declaration and the release intensity of demand response. The main optimization objectives of the training are cumulative return, price volatility suppression ability, market clearing rate and strategy stability. At the same time, the model convergence rounds and single step decision time are recorded to evaluate its executability in dynamic market simulation. The core training parameters and evaluation metrics are shown in Table 4.

Table 4: Model training parameters and evaluation index Settings

Category	Parameter or Metric	Value
Training Parameters	Number of Training Epochs	3000
Training Parameters	Batch Size	64
Training Parameters	Replay Buffer	100000
Training Parameters	Actor Learning Rate	0.0001
Training Parameters	Critic Learning Rate	0.0002
Training Parameters	Discount Factor ((\gamma))	0.99
Training Parameters	Soft Update Coefficient ((\tau))	0.005
Training Parameters	Standard Deviation of Exploration Noise	0.10
Training Parameters	Target Network Update Step	Updated once every 2 steps
Scenario Parameters	Renewable Energy Fluctuation Intensity	0.08
Scenario Parameters	Load Shock Intensity	0.10
Evaluation Metrics	Electricity Price Volatility Coefficient	CV
Evaluation Metrics	Market Clearing Rate	%
Evaluation Metrics	Strategy Stability	%
Evaluation Metrics	Supply–Demand Deviation Rate	%
Evaluation Metrics	Single-Step Decision Time	s

4.2 Analysis of deduction results of electricity price fluctuation

After completing the deployment of the experimental environment and parameter setting, this paper further analyzes the electricity price fluctuation deduction results on the test set. The test set contains 2592 time period samples, covering various market situations such as high-tech energy fluctuations, load spikes and demand response enhancement. Since the formation of electricity price is affected by both supply and demand deviation, quotation behavior and rule constraints, simply comparing the mean error is not enough to explain whether the model really captures the price evolution process. Therefore, this paper evaluates the deduction effect from four aspects of price fitting error, peak and valley identification ability, volatility amplitude reduction and turning point capture rate, and compares the proposed model with the static clearing model and the LSTM price prediction model.

The results show that the root mean square error of the electricity price of the proposed model on the test set is 4.87 yuan /MWh, and the mean absolute percentage error is 3.96%, which is significantly lower than the 10.35 yuan /MWh and 8.43% of the static clearing model, and better than the 7.62 yuan /MWh and 6.11% of the LSTM model. In terms of price turning point identification, the capture rate of the proposed model reaches 88.4%, which is 13.5 percentage points higher than that of LSTM and 26.7 percentage points higher than that of the static model. This shows that the joint mechanism based on state modeling and multi-agent behavior inference can better identify the dynamic process of electricity price from the stable range to the rapid rise range, and then from the peak to the equilibrium range, rather than just fitting the historical price curve smoothly.

In order to intuitively show the reduction effects of different models on the price volatility path, representative samples of 24 consecutive periods in the test set are selected for comparison, and the results are shown in Figure 5. It can be seen that the actual electricity price rises from 328.6 yuan /MWh to 492.7 yuan /MWh when the load rises rapidly and the output of new energy decreases from the 8th to the 11th period. If the whole 24-time series is calculated, the actual peak valley difference of the sample is 214.1 yuan /MWh. The peak price deduced by the proposed model in the same interval is 486.3 yuan /MWh, and the peak deviation is only 6.4 yuan /MWh, while the peak price of the LSTM model is 461.5 yuan

/MWh, and the static model is only 438.2 yuan /MWh. There are different degrees of peak compression.

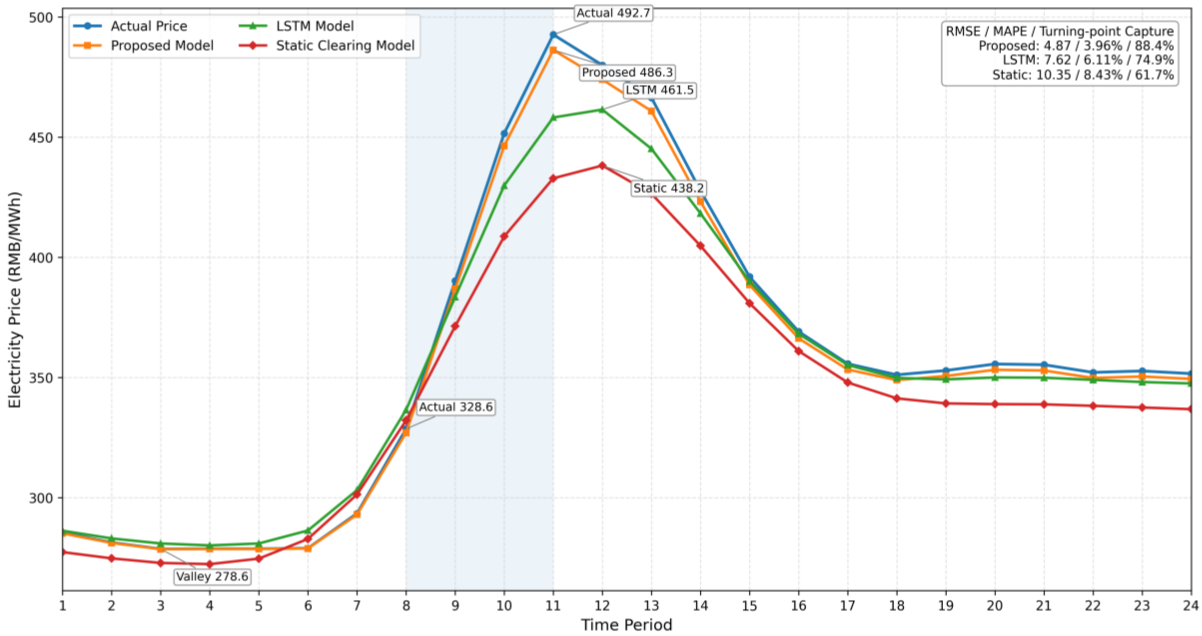

Figure 5: Comparison of electricity price fluctuation deduction results

From the perspective of fluctuation range, the fluctuation coefficient of the actual electricity price series of the test set is 0.184, and the deduced result of the model in this paper is 0.176, and the difference between them is controlled within 0.008. The LSTM model and the static model are 0.152 and 0.137, respectively, indicating that these two types of methods are easier to weaken price spikes and local shocks. Further investigating the peakvalley difference index, the peakvalley difference of the actual sample is 214.1 yuan /MWh, the proposed model is 207.8 yuan /MWh, the LSTM is 181.4 yuan /MWh, and the static model is 165.9 yuan /MWh. It can be seen that the proposed model is closer to the real market trajectory in terms of price volatility intensity recovery.

4.3 Analysis on evolution effect of market behavior

After the results of electricity price fluctuation are verified, it is necessary to further investigate the ability of the proposed model to describe the evolution process of market main body behavior. Electricity market simulation does not take price fitting as the only goal. If the model cannot simultaneously reflect the dynamic changes of price modification by power producers, power purchase adjustment by power selling companies and response release by load aggregators, its explanation of electricity price formation mechanism is still incomplete. Based on this, this paper evaluates the effect of behavior evolution from five aspects of strategy stability, market clearing rate, supply and demand deviation rate, equilibrium deviation and demand response cash rate on the test set, and compares the results with the static game model and the single agent reinforcement learning model.

The experimental results show that the average market clearing rate of the proposed model on 2592 consecutive test periods reaches 94.6%, which is higher than 90.8% of the single agent reinforcement learning model and 87.9% of the static game model. The deviation rate of supply and demand is controlled at 2.3%, which is 1.4 percentage points lower than that of the single agent model and 2.7 percentage points lower than that of the static model. In terms of strategy convergence performance, the main action of the model in this paper began to enter a stable range after the 1680 round, the quotation correction range gradually decreased from 12.4% at the beginning of the training to 4.1%, and the fluctuation range of electricity purchase declaration decreased from 9.8% to 3.7%, indicating that the model could form a relatively stable behavior adjustment path in multiple rounds of interaction. At the same time, the demand response conversion rate reaches 89.7%, indicating that the load aggregator can release the flexible load more accurately in the high-price period, thereby weakening the further amplification effect of the local supply-demand imbalance on the price.

To show the convergence trend of the agent behavior in the iteration process, Figure 6 shows the variation of the policy fluctuation amplitude with the training rounds under different models. It can be seen that the static game model basically maintains a fixed strategy in the whole process, and although there is no significant oscillation, it lacks response ability to market disturbances. The single-agent reinforcement learning model declines rapidly in the early stage, but there is still repeated swing in the middle and late stage. The model in this paper obviously plateaus after about 1600 rounds, and the fluctuation range is the smallest in the later period.

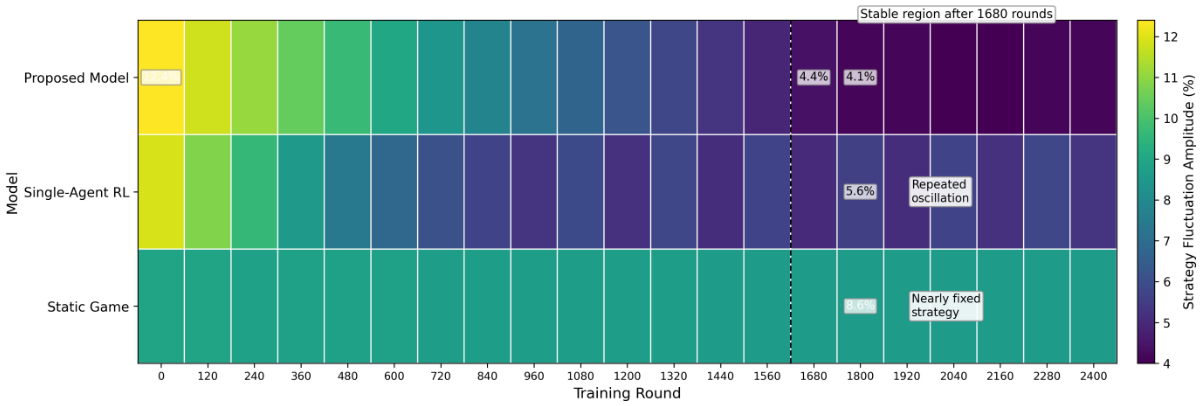

Figure 6: Changes in the volatility range of the strategies of market players under different models

From the perspective of equilibrium deviation, the average equilibrium deviation of the proposed model in the test stage is 0.041, 0.067 for the single-agent reinforcement learning model, and 0.089 for the static game model. This means that the model in this paper can pull the market state back to the relatively coordinated clearing interval faster and avoid persistent deviation when dealing with the multi-agent strategy coupling relationship. Further observing the revenue distribution of the main body, the standard deviation of the revenue of the power generator, the electricity selling company and the load aggregator under the proposed model is decreased by 11.8%, 9.6% and 13.2% respectively compared with the static model, indicating that the proposed model not only improves the market activity, but also improves the stability in the process of behavior evolution.

4.4 Discussion

The existing electricity market analysis methods have made progress in price prediction or strategy optimization, but there are still obvious shortcomings when they are used to explain the electricity price formation mechanism in the simulation system. The time series forecasting model represented by LSTM can well fit the price curve, but it is difficult to simultaneously depict the linkage relationship between quotation correction, electricity purchase adjustment and demand response release, so it is prone to peak compression in high volatility scenarios. In the experiment of this paper, the root mean square error of the electricity price is 7.62 yuan /MWh, and the price turning point capture rate is 74.9%, which is lower than 4.87 yuan /MWh and 88.4% of the model in this paper. Although the static game model can reflect the equilibrium characteristics under the rule constraints, it lacks dynamic expression for continuous feedback and strategy modification. In the test stage, the market clearing rate is only 87.9%, and the supply-demand deviation rate reaches 5.0%, which shows that its adaptability to complex disturbance scenarios is still limited. The single-agent reinforcement learning model has been improved compared with the static method, but it is still prone to policy swing under the condition of multi-agent coupling, and the equilibrium deviation in the middle and late stage is maintained at 0.067, which is less stable than 0.041 of the model in this paper.

The dynamic deduction method constructed in this paper integrates state modeling, multi-agent game and intelligent deduction mechanism into the unified simulation link, so that the electricity price change is no longer regarded as an isolated output, but is reduced to the evolution result of continuous competition, feedback and correction of multi-agent. The experimental results show that the price volatility coefficient of the proposed model is controlled at 0.176, which is closer to 0.184 than the actual series. The market clearing rate increased to 94.6%, the supply-demand deviation rate decreased to 2.3%, and the demand response cash rate reached 89.7%. These results show that the proposed method has a good comprehensive performance at the three levels of price reduction, behavioral convergence and market coordination. The improvement is mainly reflected in three aspects. Firstly, the multi-source feature representation enhances the joint recognition ability of the model for load change, marginal cost and quotation deviation. Second, the multi-agent interaction mechanism retains the strategic traction relationship in the real market. Third, the closed-loop learning structure improves the adaptability of the model to new energy fluctuations and load shocks.

At the same time, this method still has some limitations. First, model training relies on more complete scene parameters and continuous samples. If the market rules are frequently adjusted, parameter transfer still needs further verification. Secondly, the intelligent inference mechanism contains multiple modules such as state encoding, joint decision making and dual-value update, and the computational complexity is higher than that of the static model. The real-time deployment in larger markets still needs to be further optimized. Thirdly, the simulation environment of this paper mainly focuses on the continuous trading process under a uniform time granularity, and the description of cross-market coupling, extreme risk events and long-term policy disturbances can still be further expanded.

5 Conclusion

Under the background of increasing volatility of new energy output, accelerating load change and complex trading behavior, the formation process of electricity price in the power market shows stronger dynamic coupling characteristics. Aiming at this problem, this paper constructs the electricity price formation mechanism and the market behavior dynamic deduction method for the electricity market simulation system. The market data collection, scene construction, feature representation, state modeling, multi-agent game and intelligent deduction mechanism are integrated into the unified computing framework. The model training and simulation verification were completed in the environment of Python 3.11 and PyTorch 2.2. The experimental data covers continuous samples of 180 days and 15 min granularity, and a total of 17 280 time periods are recorded, which are divided into training set, validation set and test set according to the proportion of 70%, 15% and 15%. The results show that the proposed method can better describe the trajectory of electricity price fluctuation and the behavior evolution process behind it. On the test set, the root mean square error of the model electricity price is 4.87 yuan /MWh, the mean absolute percentage error is 3.96%, the price turning point capture rate reaches 88.4%, and the deduced electricity price volatility coefficient is 0.176, which is close to the actual series of 0.184. At the same time, the market clearing rate reaches 94.6%, the supply-demand deviation rate is controlled at 2.3%, the demand response exchange rate is 89.7%, and the equilibrium deviation is 0.041, which shows that the method has good performance in price reduction, behavior coordination and market stability maintenance. This paper shows that the combination of computer state modeling and multi-agent interactive deduction is helpful to explain how the electricity price is gradually formed in continuous bidding, profit feedback and strategy modification from the process level. This method not only improves the dynamic expression ability of power market simulation, but also provides a computable technical path for subsequent market rule optimization, subject behavior recrecognitiomplex scenario deduction.

About the Author

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 research focuses on power system stability analysis and power market development and operation. Contact information: Phone: 430522199009155890; Email: 13578780564@163.com.

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.

Mingyu Che, born in August 1986, is a female of Han ethnicity from Changchun City, Jilin Province. She holds a bachelor's degree from Jilin University of Finance and Economics and is a certified Intermediate Economist. Her primary research focuses on electricity market trading and settlement.

Zhenduo Gao, born in July 1989, is a male of Han ethnicity from Changchun City, Jilin Province. He holds a Master's degree in Electrical Engineering from Northeast Electric Power University and is a Senior Engineer. His primary research focuses on the functional design, development, and operation management of power trading platforms.

Acknowledgments

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

References

[1] Wu J, Wang J, Kong X. Strategic bidding in a competitive electricity market: An intelligent method using Multi-Agent Transfer Learning based on reinforcement learning[J]. Energy, 2022, 256: 124657.

[2] Namalomba E, Feihu H, Shi H. Agent based simulation of centralized electricity transaction market using bi-level and Q-learning algorithm approach[J]. International journal of electrical power & energy systems, 2022, 134: 107415.

[3] Aliabadi D E, Chan K. The emerging threat of artificial intelligence on competition in liberalized electricity markets: A deep Q-network approach[J]. Applied energy, 2022, 325: 119813.

[4] Xu H, Wu Q, Wen J, et al. Joint bidding and pricing for electricity retailers based on multi-task deep reinforcement learning[J]. International journal of electrical power & energy systems, 2022, 138: 107897.

[5] 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.

[6] Jain P, Saxena A. A Multi-Agent based simulator for strategic bidding in day-ahead energy market[J]. Sustainable Energy, Grids and Networks, 2023, 33: 100979.

[7] 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.

[8] Li Q, Yang Z, Yu J, et al. Impacts of previous revenues on bidding strategies in electricity market: A quantitative analysis[J]. Applied Energy, 2023, 345: 121304.

[9] Tang Q, Guo H, Zheng K, et al. Forecasting individual bids in real electricity markets through machine learning framework[J]. Applied Energy, 2024, 363: 123053.

[10] Liu C, Rao X, Zhao B, et al. Deep reinforcement learning-based optimal bidding strategy for real-time multi-participant electricity market with short-term load[J]. Electric Power Systems Research, 2024, 233: 110404.

[11] Harder N, Qussous R, Weidlich A. Fit for purpose: Modeling wholesale electricity markets realistically with multi-agent deep reinforcement learning[J]. Energy and AI, 2023, 14: 100295.

[12] Nitsch F, Schimeczek C, Bertsch V. Applying machine learning to electricity price forecasting in simulated energy market scenarios[J]. Energy Reports, 2024, 12: 5268-5279.

[13] Peng F, Zhang W, Zhou W, et al. Review on bidding strategies for renewable energy power producers participating in electricity spot markets[J]. Sustainable Energy Technologies and Assessments, 2023, 58: 103329.

[14] Bichler M, Knörr J. Getting prices right on electricity spot markets: On the economic impact of advanced power flow models[J]. Energy Economics, 2023, 126: 106968.

[15] Han Z, Fang D, Yang P, et al. Cooperative mechanisms for multi-energy complementarity in the electricity spot market[J]. Energy economics, 2023, 127: 107108.

[16] Zhang C, Wu X, Zhao S, et al. Multi-agent simulation of the effects of Japanese electricity market policies on the low-carbon transition[J]. Energy Strategy Reviews, 2024, 52: 101333.

[17] Turdybek B, Tostado-Véliz M, Mansouri S A, et al. A local electricity market mechanism for flexibility provision in industrial parks involving Heterogenous flexible loads[J]. Applied Energy, 2024, 359: 122748.

[18] Barbosa J, Döllinger F, Steinke F. Game-theoretic analysis of suppliers' pricing power in thermal-electric local energy markets[J]. Energy, 2024, 313: 133591.

[19] Kafshian H, Monfared M A S. A multi-layer–multi-player game model in electricity market[J]. IET Generation, Transmission & Distribution, 2024, 18(7): 1494-1515.

[20] Wu J, Wang J, Kong X. Intelligent strategic bidding in competitive electricity markets using multi-agent simulation and deep reinforcement learning[J]. Applied Soft Computing, 2024, 152: 111235.

[21] Wang Z, Lam J S L, Huo J. The bidding strategy for renewable energy auctions under government subsidies[J]. Applied Energy, 2024, 353: 122148.