ISSN 2616-5902 Vol. 3, Issue 8: 72-83, DOI: 10.25236/AJBM.2021.030813

Market trading model before the distribution network under the participation of multi-interest parties at the early stage of marketization

Wenxu Tian

CHN Energy Longyuan Power Technology & Engineering Co., Ltd. Haidian Beijing, 100039, China
183758841@qq.com

Abstract: *Along with the promotion of the electricity sales side market reform, load-side resources, including new energy sources, gradually participate in the distribution network electricity sales market. At the early stage of marketization, in order to give full participation rights to all parties involved in the market and to take into account the responsibility of distribution network operators to ensure the safety and quality of power supply, we propose a distribution network day-ahead market transaction model with the participation of multiple interests at the early stage of marketization. Firstly, a tripartite interest model including distribution network operators, distributed power supply operators and load aggregators is established in the distribution network, and all three parties take time-of-use tariff as the game strategy; secondly, a Pay as Bid (PAB) model is combined with the tripartite interest model to establish a market settlement model, and the three parties can modify their own offers according to the available information in the market settlement process. Finally, the Nash-Q method is used to solve the model. The results show that, compared with the traditional constant/time-sharing tariff, the model can ensure the safe and reliable power supply to customers while stimulating the active participation of new energy and other market players in the market, and can also increase the contribution of new energy to the power balance and reduce the risk of new energy consumption in the distribution network.*

Keywords: *Nash-Q method, distribution network, day-ahead market, game theory*

1. Introduction

The focus of power system reform is to control the middle and liberalize the two ends [1]. In the context of the increasingly mature reform of the power generation market, the opening of the electricity sales market will become an important task to deepen the reform of the power system. The distribution network power market will be transformed from the original model of unified purchase and sale to a market model combining medium- and long-term contract market, day-ahead market and real-time market [2]. The market counterparties will expand from traditional Distribution Network Operator (DNO) to Distributed Generation Operator (DGO), Load Aggregator (LA), and even integrated Energy service providers, etc. [3-4].

There have been many studies on the trading strategies and trading models of the electricity sales market in China. In the literature [5], a day-ahead electricity market-based mechanism based on optimization theory was proposed. This mechanism allows market members to change their bidding strategies and effectively solves the problem of cost information asymmetry among market members. The literature [6] constructs a multi-subject game model of the market containing power generation enterprises, emerging power sales entities and users, but the paper only gives a static trading method for a certain time period, and has not yet covered the dynamic trading strategy for consecutive time periods. Based on the background of direct purchase of electricity by large customers, the paper [7] designs a market model to motivate customers to participate in wind power consumption. The literature [8] constructs a master-slave game model for bilateral contract trading containing multiple generators and users, and gives a general solution to the Nash equilibrium solution. The literature [9] reviews the trading mechanism of distributed subjects participating in the market at the distribution network side. Its results show that distributed subjects participate in trading through real-time tariffs, breaking the traditional distribution network operator-centered trading model, and can achieve a win-win situation for distributed subjects and the system as a whole. The literature [10] outlines some experiences of foreign electricity sales side market opening, including basic trading objects, market access mechanism, and risk control.

ISSN 2616-5902 Vol. 3, Issue 8: 72-83, DOI: 10.25236/AJBM.2021.030813

The studies in the aforementioned literature are all based on the background of relatively mature electricity market, where each market subject is free to participate in market transactions. However, at the early stage of marketization, the electricity sales side should be open to some of the power purchase customers [10], while the traditional grid operators also need to assume the traditional basic responsibility for the safety and quality of power supply. While ensuring safe and reliable power supply to customers, how to mobilize all parties through benefit distribution is a problem that needs to be solved urgently. Based on this, this paper proposes a distribution grid day-ahead market transaction model with the participation of multiple interest parties at the early stage of marketization. The model not only takes into account the different interests of DNO, DGO and LA, but also takes into account the responsibility of DNO to ensure the safety and quality of power supply. The model is solved by Nash-Q method. The results show that, compared with the traditional constant/time-sharing tariff, the model can ensure the safe and reliable power supply to customers while incentivizing the active participation of new energy and other market players in the market, and also improve the contribution of new energy to the power balance and reduce the risk of new energy consumption in the distribution network.

2. Multi-entity game model of electricity market before the distribution network day

2.1 Multi-entity game framework

The traditional distribution network contains numerous resources, such as flexible loads and distributed power sources. In this paper, the resources are clustered into three types of entities: distribution network operators (DNOs), distributed power supply operators (DGOs), and load aggregators (LAs). All three types of subjects participate in the competition for the supply of tradable loads. Among them, the subjects are classified as shown in Figure 1.

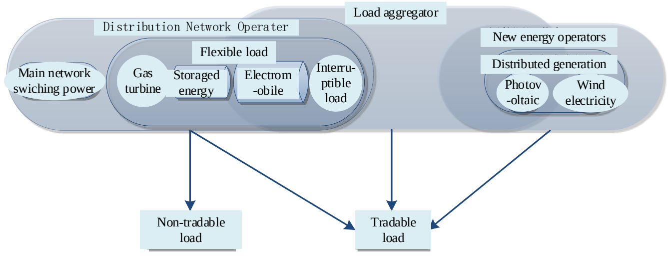


Fig.1 Schematic diagram of party classification

As shown in Figure 1 above, the load aggregator (LA) has more controllable resources, including distributed power and flexible load. In this paper, Load Aggregators (LAs) are selected to operate mainly photovoltaic systems and Interruptible Load (IL). Meanwhile, DGO mainly operates wind power systems, which must be coupled with energy storage to maintain smooth power output due to the volatility of wind power output. The Distribution Network Operator (DNO), as the main supplier of electricity to the distribution network, is responsible for operating the exchange of power with the main grid and a part of the flexible loads, such as gas turbines. Both DNO-controlled gas turbines and LA-controlled interruptible loads can compete for the responsibility of securing power supply. The framework of the tripartite game in the electricity market before the distribution network day is shown in Figure 2.

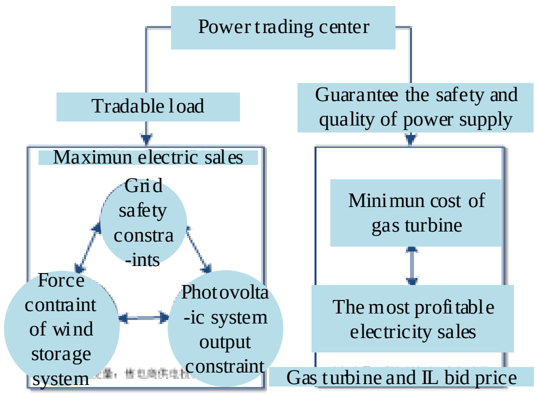


Fig.2 Schematic diagram of the tripartite game framework

ISSN 2616-5902 Vol. 3, Issue 8: 72-83, DOI: 10.25236/AJBM.2021.030813

2.2 Multi-entity quoted revenue model

1) DNO Revenue Model

On the one hand, the DNO competes with the DGO and LA for the maximum profit, and on the other hand, the DNO must assume the responsibility of guaranteeing the power supply. the profit X of the DNO in the whole dispatching cycle can be expressed as equations (1) to (5)

$$F^{DNO} = \max \Delta T \sum_{t=1}^{N_T} (I_t^{DNO} - B_t^{DNO}) \tag{1}$$

$$I_t^{DNO} = P_t^{DNO} w_t^{DNO} + P_{loss,t} w_t^{pd} \tag{2}$$

$$\begin{aligned} B_t^{DNO} = {} & P_t^{DNO} w_t^{pd} + P_t^{IL} w_t^{IL} + \mu_t \overline{\mu_{t-1}} C_t^{M} \\ & + \mu_t P_t^{M} w_t^{M} + C_t^{sd} \end{aligned} \tag{3}$$

In the formula: ΔT is the length of the scheduling period, N_T is the total number of scheduling periods, I_t^{DNO} is the benefit of DNO in D period, and B_t^{DNO} is the cost of DNO in t period.

I_t^{DNO} includes the benefit of electricity sales and the benefit of network loss improvement. In equation (2), P_t^{DNO} is the amount of electricity supply obtained by DNO participating in the game in time t, w_t^{DNO} is the price of electricity sold by DNO, $P_{loss,t}$ is the reduction of active network loss by DNO participating in the market transaction, and w_t^{pd} is the unit price of electricity purchased by DNO from the main grid.

B_t^{DNO} includes the cost of purchasing electricity from the main grid, the cost of purchasing IL, the cost of activating the gas turbine, and the cost of undertaking to guarantee the safety and quality of electricity supply. P_t^{IL} is the amount of IL electricity purchased by DNO from LA, and w_t^{IL} is the price of electricity sold for IL. μ_t is a Boolean variable representing the operating status of the gas turbine, and μ_t =1 if the gas turbine is operating in t time, and 0 if the opposite. C_t^{M} is the cost of starting and stopping the gas turbine once, P_t^{M} is the electricity supplied by the gas turbine, and w_t^{M} is the unit price of electricity for operating the gas turbine. C_t^{sd} is the cost of undertaking to ensure the safety and quality of the electricity supplied.

In this paper, the nodal voltage and branch currents of the grid are taken as the criteria for assessing the responsibility of guaranteeing power supply.

a) The voltage at the nodes of the distribution network must be maintained within a reasonable range to ensure the quality of power supply.

$$C_t^{sd} = \begin{cases} 0 & \forall\ j \in N \quad V_j \in [V_{\min}\ V_{\max}] \\ \lambda_1 & \exists\ j \in N \quad V_j \notin [V_{\min}\ V_{\max}] \end{cases} \tag{4}$$

b) The tide of each branch of the distribution network must be limited to the maximum allowable transmission power of each branch to ensure the safety of the network operation.

$$C_t^{sd} = \begin{cases} 0 & \forall\ i, j \in N \quad |P_{ij}| \le P_{ij,\max} \\ \lambda_2 & \exists\ i, j \in N \quad |P_{ij}| > P_{ij,\max} \end{cases} \tag{5}$$

In the formula: λ_1 and λ_2 are the penalty cost constants for violation of the mechanism.

2) DGO Revenue Model

The revenue function F^{DGO} of a wind storage system in a DGO-operated distribution network can be expressed as equations (6) to (8).

ISSN 2616-5902 Vol. 3, Issue 8: 72-83, DOI: 10.25236/AJBM.2021.030813

$$F^{DGO} = \max \Delta T \sum_{t=1}^{N_T} (I_t^{DGO} - B_t^{DGO}) \tag{6}$$

$$I_t^{DGO} = P_t^{WG} w_t^{DGO} \tag{7}$$

$$B_t^{DGO} = P_t^{WG} w_t^{WG} + P_{ES,t}^{WG} w_{ES,t} \tag{8}$$

In the formula: I_t^{DGO} is the electricity sales benefit of DGO in time period t. P_t^{WG} is the amount of electricity supplied by DGO after participating in the game, w_t^{DGO} is the electricity sales price of DGO, B_t^{DGO} is the cost of DGO in time period t, w_t^{WG} is the unit wind power operation and maintenance cost, $P_{ES,t}^{WG}$ is the amount of wind storage system charge and discharge, and $w_{ES,t}$ is the operation and maintenance cost of energy storage system.

3) LA Revenue Model

LA operates the optical storage system in the distribution network, while controlling IL as a demand-side resource to participate in market transactions, the revenue function F^{LA} can be expressed as equations (9) to (11)

$$F^{LA} = \max \Delta T \sum_{t=1}^{N_T} (I_t^{LA} - B_t^{LA}) \tag{9}$$

$$I_t^{LA} = P_t^{PV} w_t^{LA} + P_t^{IL} w_t^{IL} \tag{10}$$

$$B_t^{LA} = P_t^{PV} w_t^{PV} + P_t^{IL} w_t^{us} + P_{ES,t}^{PV} w_{ES,t} \tag{11}$$

In the formula: I_t^{LA} is the electricity sales benefit of LA in time period t. P_t^{PV} is the amount of electricity supply obtained after LA participates in the game, w_t^{LA} is the electricity sales price of LA. B_t^{LA} is the cost of LA in time period t, w_t^{PV} is the unit photovoltaic operation and maintenance cost, w_t^{us} is the unit subsidy cost that LA needs to give to customers for buying IL, and $P_{ES,t}^{PV}$ is the charge and discharge volume of the optical energy storage system.

The three-party game constraint is as follows.

a) Power balance constraint

The tripartite participation of DNO, DGO and LA in the electricity market game before the distribution day must be premised on the power balance constraint.

$$P_t^{DNO} + P_t^{WG} + P_t^{PV} = P_t^{Load} \tag{12}$$

In the formula: P_t^{Load} is the amount of electricity required by large customers in time period t. At any time, the combined power supply of the three parties must meet the power required by the load.

b) Energy storage equipment operating constraints

To ensure the lifetime of the energy storage system and the continuity of the control cycle, the energy storage system must satisfy the charging and discharging constraints and the constraint of zero energy change within a dispatch cycle, as shown in equations (13) to (15).

$$E_{\min} \le E_{t-1} + [P_{st,t}\delta_{st} - P_{ex,t}/\delta_{ex}]\Delta T \le E_{\max} \tag{13}$$

$$\begin{cases} P_{st,t} \le P_{ES,\max} \\ P_{ex,t} \le P_{ES,\max} \end{cases} \tag{14}$$

$$\sum_{t=1}^{N_T} [P_{st,t}\delta_{st} - P_{ex,t}/\delta_{ex}]\Delta T = 0 \tag{15}$$

In the formula: $E_{\min}$, $E_{\max}$, E_{t-1} are the minimum, maximum capacity of the energy storage device

ISSN 2616-5902 Vol. 3, Issue 8: 72-83, DOI: 10.25236/AJBM.2021.030813

and the stored power at the end of time period $t-1$. $P_{st,t}$, $P_{ex,t}$ and δ_{st}, δ_{ex} are the charging and discharging power and charging and discharging efficiency at time period t. $P_{ES,\max}$ is the maximum allowable charging and discharging power of energy storage in t time period.

c) IL Characteristic Constraints

Interruptible load participates in demand response as a load-side resource. The key factors for enabling IL depend on the transaction duration and the transaction power of IL, as shown in equations (16) to (17)

$$P_{IL,t} \le P_{IL,\max} \tag{16}$$

$$T_{IL} \le T_{IL,\max} \tag{17}$$

In the formula: $P_{IL,t}$, $P_{IL,\max}$ are the trading power and the maximum allowed trading power in IL in t time period; T_{IL} , $T_{IL,\max}$ are the trading hours and the maximum allowed trading hours in IL.

2.3 Multi-subject game equilibrium model

The PAB bidding mechanism refers to the settlement of proceeds in the market clearing process with each seller's own offer. Its basic schematic is shown in Figure 3.

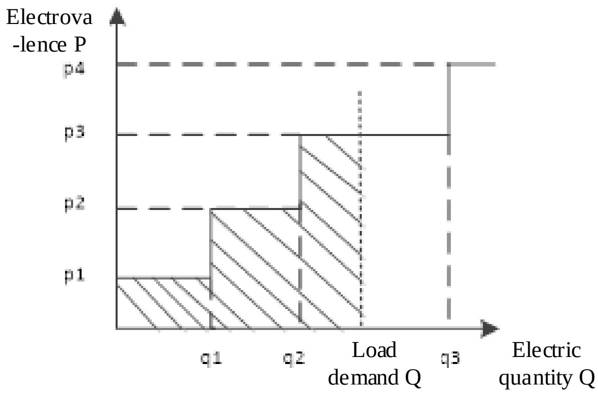


Fig.3 Pay As Bid Price

At the beginning of the bidding process, DNO, DGO, and LA use the information released by ISO a few days ago as a reference to report the t time step $[(q_t^{DNO} q_t^{M}) q_t^{DGO} (q_t^{PV} q_t^{IL})]$ and the corresponding step $[(w_t^{DNO} w_t^{M}) w_t^{DGO} (w_t^{PV} w_t^{IL})]$. In the market clearing process, ISO uses the queuing method to calculate the amount of electricity supplied by each party $[(P_t^{DNO} P_t^{M}) P_t^{WG} (P_t^{PV} P_t^{IL})]$, while the revenue of each party is determined by equations (1)-(11) according to the PAB settlement mechanism.

According to the definition of Nash equilibrium [7], when the tariff of one party of the game changes by itself, neither of the other two parties will be willing to change the tariff voluntarily in order to increase the revenue. Its equilibrium solution is

$$[w_t^{*DNO} w_t^{*M}] = \arg^{-1}\max F^{DNO}(w_t^{*DGO},[w_t^{*PV} w_t^{*IL}]) \tag{18}$$

$$w_t^{*DGO} = \arg^{-1}\max F^{DGO}([w_t^{*DNO} w_t^{*M}],[w_t^{*PV} w_t^{*IL}]) \tag{19}$$

$$[w_t^{*PV} w_t^{*IL}] = \arg^{-1}\max F^{LA}([w_t^{*DNO} w_t^{*M}],w_t^{*DGO}) \tag{20}$$

$$w_t^{*bd} \in K_{bd} w_{t,1}^{bd} \tag{21}$$

In the formula: $(w_t^{*DNO} w_t^{*M})$, w_t^{*DGO} and $(w_t^{*PV} w_t^{*IL})$ are the optimal strategy sets for DNO, DGO and LA, respectively. w_t^{*bd} is the optional offer strategy of each party, K_{bd} is the allowed price fluctuation factor of each party, and $w_{t,1}^{bd}$ is the first offer strategy of each party in time slot t. To ensure the fairness of the three-party game, the equilibrium solution must be within the allowed price fluctuation factor. To ensure the fairness of the three-party game, the equilibrium solution must be limited to the range of the allowable tariff fluctuation factor K_{bd}.

2.4 Model evaluation indicators

In order to compare the difference between the time-stepped tariff-based electric energy trading model established in this paper and the traditional constant/time-sharing tariff-based trading model, the following three indicators are defined.

1) Equilibrium degree of benefit distribution. The equilibrium degree of benefit distribution among each electricity seller in time period t during the market transaction process is evaluated using the equilibrium degree $e(x_{1,t}\, x_{2,t}\, x_{3,t})$. $x_{1,t}\, x_{2,t}\, x_{3,t}$ is the ratio of the respective benefits of DNO, DGO, and LA in time period t to the sum of the benefits of the three parties, respectively. As shown in Equations (22) to (25). It can be seen that the closer the value of the equilibrium degree is to 1, the more balanced the benefits of each party are.

$$e(x_{1,t}\, x_{2,t}\, x_{3,t}) = 1 - \sum_{j=1}^{3} \left(x_{j,t} - \frac{1}{3} \sum_{i=1}^{3} x_{i,t} \right)^2 \tag{22}$$

$$x_{1,t} = F_t^{DNO} / (F_t^{DNO} + F_t^{DGO} + F_t^{LA}) \tag{23}$$

$$x_{2,t} = F_t^{DGO} / (F_t^{DNO} + F_t^{DGO} + F_t^{LA}) \tag{24}$$

$$x_{3,t} = F_t^{LA} / (F_t^{DNO} + F_t^{DGO} + F_t^{LA}) \tag{25}$$

2) Power balance contribution. The power balance contribution degree is used to measure the effect of new energy in the market trading process to relieve the pressure of power supply and participate in peak regulation during the peak load. Its calculation formula is

$$b = \min \left\{ \frac{1}{N_1} \frac{\sum_{t \in T_1} (P_t^{WG} + P_t^{PV})}{P_{\text{Load, 午高峰}}} \quad \frac{1}{N_2} \frac{\sum_{t \in T_2} (P_t^{WG} + P_t^{PV})}{P_{\text{Load, 晚高峰}}} \right\} \times 100\% \tag{26}$$

In the formula: T_1 , T_2 are the sets of hours for the afternoon and evening peaks, respectively; N_1 , N_2 are the number of hours for the afternoon and evening peaks, and $P_{\text{Load, 午高峰}}$ $P_{\text{Load, 晚高峰}}$ is the load of users participating in transactions during the afternoon and evening peaks, respectively.

3) Average penetration volatility of new energy. This indicator is used to measure the value of the risk to be taken by the distribution network to consume new energy.

$$v = \sum_{t=1}^{N_T} \left(\frac{P_t^{WG} + P_t^{PV}}{P_{SumLoad,t}} - \frac{1}{N_T} \sum_{t=1}^{N_T} \frac{P_t^{WG} + P_t^{PV}}{P_{SumLoad,t}} \right)^2 \tag{27}$$

In the formula: $P_{SumLoad,t}$ is the total load of the distribution network in time period t. The smaller the indicator is, the more stable the new energy supply capacity is, the more stable the distribution network needs to cope with the risk of fluctuations in new energy output and provide backup capacity, and the less risk it bears to consume new energy.

3. Game model solving based on Nash-Q method

As described in Section 2.2, the electric energy trading model developed in this paper is a nonlinear, multivariate game problem. The Q-learning algorithm has high convergence reliability and the complexity of the model has less influence on the algorithm, which has some advantages in solving the nonlinear problem. Meanwhile, combining game theory with Q-learning is beneficial to better solve the multi-object game problem.

3.1 Nash-Q Method

The Q-learning algorithm, one of the most frequently used reinforcement learning algorithms, is suitable for optimal policy selection in discrete Markov states [12-13]. The principle of the algorithm is to use the current empirical Q value as the initial value for subsequent Q calculations. The iterative equation is expressed as.

$$Q_{n+1}(a,s) = Q_n(a,s) + \alpha[R(a,s) + \gamma \max_{a' \in A\ s' \in S} Q_n(a',s') - Q_n(a,s)] \tag{28}$$

$$P(a_i) = \frac{e^{Q(s,a_i)}}{\sum_{a \in A} e^{Q(s,a)}} \tag{29}$$

In the formula: s , s' are the current state and future state, respectively, and S is the set of state spaces; $R(a,s)$ denotes the reward value of the intelligence after adopting action a in state s , and G is the set of action strategies; $Q_{n+1}(a,s)$, $Q_n(a,s)$ are the target values of the $n+1$th and n th steps; $\max_{a' \in A\ s' \in S} Q_n(a',s')$ denotes the maximum possible reward value of the intelligence in state s' . α is the learning parameter and γ is the discount factor. $P(a_i)$ is the probability of selecting action a_i in state s .

The literature [14] combines Q-learning and game theory to propose the Nash-Q method. It uses Nash equilibrium solutions to define Q-value functions, thus solving multi-party non-zero and non-cooperative game problems.

Intelligent body i , in state s' , forms a game parity $\pi(s') = [Q_n^1(s'), Q_n^2(s')...Q_n^m(s')]$ with other intelligent bodies and has an equilibrium solution $[\pi^1(s')\,\pi^2(s')...\pi^m(s')]$, where m is the number of intelligent bodies. Thus the Q-value iteration can be expressed as

$$\begin{aligned} Q_{n+1}^i(s,a^1,a^2...,a^m) &= (1-\alpha)Q_n^i(s,a^1,a^2...,a^m) \\ &\quad + \alpha[R^i + \gamma NashQ_n^i(s')] \\ NashQ_n^i(s') &= \pi^1(s')\pi^2(s')...\pi^m(s') \cdot Q_n^i(s') \end{aligned} \tag{30}$$

In the formula: $NashQ_n^i(s')$ denotes the payoff function for the choice of the Nash equilibrium solution of the intelligent body i .

3.2 Game model solving based on Nash-Q method

Based on the previous analysis, the three-party game model for the distribution network is solved as follows.

1) Initialize the Q-value table. the initial value of each element (a,s) *in the Q-value table is taken as 0.*

2) Create the required (action-state) pairs for Q-learning.

The action strategy selection takes the charging and discharging behavior of the energy storage equipment, IL and whether the gas turbine participates in the market trading behavior as the strategy set, i.e., $A = \{A_E\ A_{IL}\ A_M\}$.

For the state space selection, the moment, the predicted output of the new energy source and the stored power value of the energy storage are taken as the state elements. To match the need of Q-learning algorithm, the variables are discretized into the form of equal-length intervals, i.e.

$$E_x = [E_{\min} + x\Delta E \;\; E_{\min} + (x+1)\Delta E] \\ \Delta E = (E_{\max} - E_{\min})/m \;\; 0 \le x \le m-1 \;\; x \in N \tag{31}$$

$$P_y = [y\Delta P \;\; (y+1)\Delta P] \\ \Delta P = P/n \quad 0 \le y \le n-1 \;\; y \in N \tag{32}$$

In the formula: $E_{\max}$, $E_{\min}$ are the maximum and minimum storage capacity of energy storage respectively, and P is the installed capacity of new energy sources (wind power, PV). Therefore, for any moment, given the new energy output value and the current storage capacity, we can determine the unique state $S = \{S_t \; S_P \; S_E\}$.

3) According to the description in Section 2.3, the three parties report the benchmark step (tariff-electricity) group for the first round of time slot t according to the information released by ISO before the day. To ensure fairness and impartiality, after the start of the current round of gaming, the step electricity is kept constant and the step tariff is varied within a limited range of tariff fluctuation factor K_{bd}. The i th round of game tariff will take the $i-1$ th round of game tariff as input.

$$[w_{t,i}^{DNO} \; w_{t,i}^{M}] = \arg^{-1}\max F^{DNO}(w_{t,i-1}^{DGO}, [w_{t,i-1}^{PV} w_{t,i-1}^{IL}]) \tag{33}$$

$$w_{t,i}^{DGO} = \arg^{-1}\max F^{DGO}([w_{t,i-1}^{DNO} w_{t,i-1}^{M}], [w_{t,i-1}^{PV} w_{t,i-1}^{IL}]) \tag{34}$$

$$[w_{t,i}^{PV} \; w_{t,i}^{IL}] = \arg^{-1}\max F^{LA}([w_{t,i-1}^{DNO} w_{t,i-1}^{M}], w_{t,i-1}^{DGO}) \tag{35}$$

4) Determine whether the three-way game reaches Nash equilibrium, i.e., the optimization result of round i is consistent with the optimization result of round $i-1$.

$$[w_t^{*DNO} \; w_t^{*M}] = [w_{t,i}^{DNO} \; w_{t,i}^{M}] = [w_{t,i-1}^{DNO} \; w_{t,i-1}^{M}] \tag{36}$$

$$w_t^{*DGO} = w_{t,i}^{DGO} = w_{t,i-1}^{DGO} \tag{37}$$

$$[w_t^{*PV} \; w_t^{*IL}] = [w_{t,i}^{PV} \; w_{t,i}^{IL}] = [w_{t,i-1}^{PV} \; w_{t,i-1}^{IL}] \tag{38}$$

If the game reaches equilibrium, then go to step (5), otherwise go to step (3).

5) From equations (1) to (11), we solve for the tripartite benefit, i.e., the Q value of the time period to which it belongs, and forecast the new energy output for the next time period.

6) Iterate the Nash-Q algorithm by equation (30), and at the same time calculate the stored power of the energy storage system according to the actual output of the new energy in the next period combined with the corresponding model and get the new state s', so that $s \to s'$.

7) Determine whether Q-learning converges or reaches the predetermined time limit, if it does not converge, return to (2) and continue the calculation.

4. Analysis of algorithms

4.1 Algorithm parameters

To verify the effectiveness of the method in this paper, the IEEE 33-node power distribution system is

ISSN 2616-5902 Vol. 3, Issue 8: 72-83, DOI: 10.25236/AJBM.2021.030813

used as an arithmetic example [21], and its system wiring diagram is shown in Fig 4.

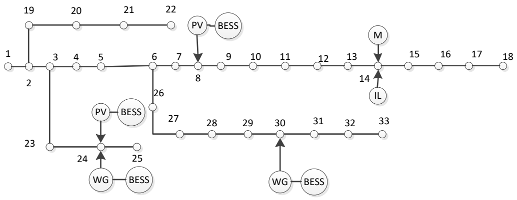


Fig.4 IEEE 33-bus distribution network system

As shown in the figure, DGO operates WG-BESS at Nodes 24 and 30, and LA operates PV-BESS at Nodes 8 and 24. the wind and optical storage system parameters are detailed in Exhibit 1.

Table 1 Profits and partial cost in different scenarios

	DNO				LA		DGO
	Profit/RMB	Cost of calling gas wheel/RMB	Purchase IL cost/RMB	Undertake the cost of guaranteeing power supply mechanism	Profit/RMB	IL electric sales benefit/RMB	Profit/RMB
Scenario1	2089.08	484.61	0	0	2005.22	0	2845.24
Scenario2	2232.43	457.72	0	0	2233.47	0	3008.45
Scenario3	2659.78	163.26	242.36	0	2524.92	242.36	3324.61

DNO puts in gas turbine M at node 14 with a maximum available power of 150 kW. LA has controllable IL at node 14 with a maximum controllable power of 80 kw and a maximum continuous control duration of 5 hours. Both the interruptible load and the gas turbine participate in the market transaction are used to assume the responsibility of securing power supply. The game relationship is shown in Table 2 below.

Table 2 Tripartite game relationship table

	Game party(schedulable device/capacity)		
Market transaction load number	DNO	DGO	LA
8	✓	×	✓
24	✓	✓	✓
30	✓	✓	×

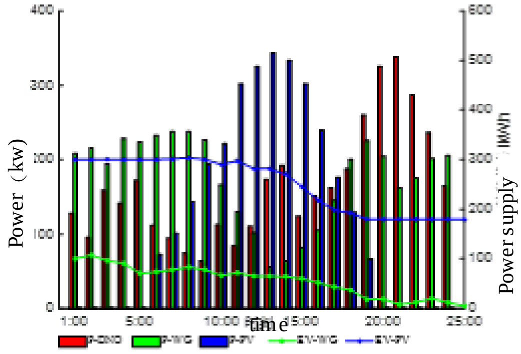
4.2 Scenario Comparison

To illustrate the validity of the model in this paper, the current mainstream electricity sales market trading mechanism-fixed tariff/time share tariff-is introduced for comparison: scenario 1, where new energy in the distribution network is sold at a fixed tariff. In scenario two, a time-sharing tariff is used, and the new energy sources have the right to set their own electricity sales strategy [22]. In scenarios one and two interruptible loads do not participate in the market, the distribution network accommodates new energy sources to the maximum extent, and there is no market gaming behavior among the three parties. In scenario three, the three-party game is conducted by the method in this paper. Where scenario two is solved by Q-learning algorithm and scenario three is solved by Nash-Q method.

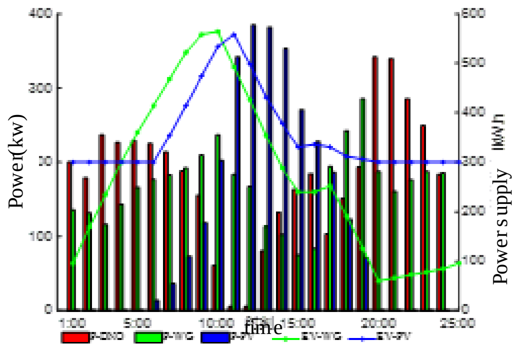
In terms of economic benefits, the results of the tripartite game of load power supply at node 24 under the three scenarios are shown in Figure 5. the results of the game of load power supply at node 8 and node 30 are shown in Figure 1 and Figure 2. in the figure, P-DNO, P-WG and P-PV are the electricity sales of DNO, DGO and LA respectively. EV-WG and EV-PV are the values of electricity stored in wind and light storage. Scenario 1, because the fixed tariff has no guiding effect on the new energy output, the new energy output is completely determined by its own characteristics, and its own profit is the lowest level among the three scenarios, while the gas turbine cost called by the DNO to undertake the guaranteed power supply mechanism is the highest among the three scenarios. The profit of LA and DGO is increased by 228.25 Yuan (11.4%) and 163.21 Yuan (5.7%) respectively compared with Scenario 1. The profit of DNO is increased by 143.35 Yuan (6.9%) compared with Scenario 1 because the amount of power supplied by DNO is lower during the peak hours. 143.35 yuan (6.9%). The time-sharing tariff can guide the new energy to participate in load peaking, so the cost of DNO to move the gas turbine is reduced by $26.89 (5.5%) compared with scenario one. Meanwhile, it can be seen from Figure 5(b) that the light storage system and

ISSN 2616-5902 Vol. 3, Issue 8: 72-83, DOI: 10.25236/AJBM.2021.030813

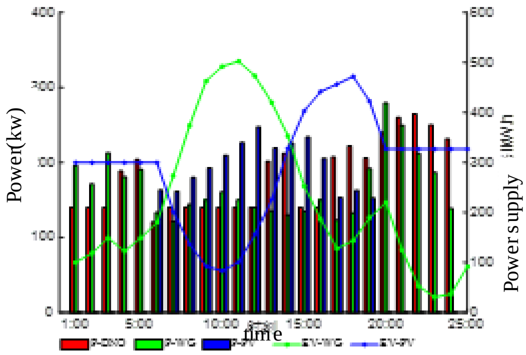
wind storage system participate in power supply with maximum power in the midday period under the premise of satisfying the formula (13)~(15), which leads to the most DNO power supply power resulting in the lowest DNO power supply power of 5kw, while the new energy reduces its own power supply power in the evening period to satisfy the energy storage equipment constraint, so DNO has the maximum power supply power of 355kw. power supply peak-valley difference of 350kw. The difference between peak and valley power supply is 350kw. It can be seen that the time-sharing tariff mechanism can improve the profit of new energy feed-in, but it will sacrifice the profit of DNO and increase the fluctuation of DNO power supply.



(a)Scenario 1



(b) Scenario 2



(c) Scenario 3

Fig.5 Tripartite outputs of node 24 in different scenarios

Therefore, in contrast with scenario 1, the new mechanism can enhance the interests of all parties, and also motivate new energy and other subjects to actively participate in the market, which is conducive to the promotion of market reform. From Fig. 5(c) compared with Fig. 5(a) and Fig. 5(b), the power output of DNO increases significantly during peak load hours, and DNO competes with new energy for load power supply by regulating the offer, which improves the position of DNO in the market game and also inhibits to a certain extent the tendency of new energy and other subjects to supply power arbitrarily in pursuit of benefits. At the same time, because DNO has the bargaining power, DNO's power supply is more moderate in other periods, and the difference between peak and valley power supply is only 110kw, so the pressure of DNO's power supply becomes smaller.

Figure 8 represents the gas turbine and IL power output under three scenario comparisons. Scenario one and scenario two lead to higher cost for DNO to invoke gas turbine power to undertake the mechanism of securing power supply at peak load due to IL not participating in the market. The LA-controlled IL under scenario three participates in the market transaction in time and its profit is $242.36; the total cost borne by the DNO in calling the gas turbine while purchasing the IL is $405.62. Comparing with scenarios one and two, the cost is reduced by about 15%. This shows that mobilizing load-side resources to actively participate in the market has a certain positive effect on reducing DNO's operating costs, while IL itself can also gain a part of the benefits, which can achieve a win-win situation in the interests of all parties involved in the market while incentivizing LA.

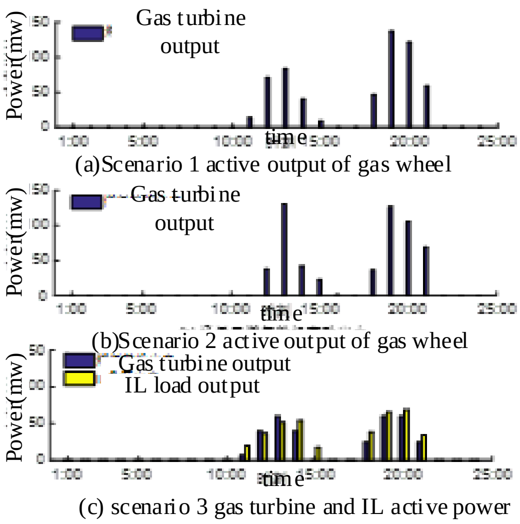


Fig.6 The outputs of IL and gas turbine in different scenarios

5. Conclusion

In this paper, we establish the market transaction model of distribution grid day before the market under the participation of multiple interests at the early stage of marketization in the context of electricity reform on the sales side, and solve it by Nash-Q method. The results show that.

1) Game bargaining can improve the profit of each interest subject participating in market transactions and increase their participation enthusiasm. At the same time, it can mobilize demand-side resources to participate in the market and reduce the operation cost of distribution network.

2) Game bargaining gives equal bargaining power to all parties. It can reduce the profit difference of each main body and facilitate the initial market reform.

3) Game bargaining can improve the effect of new energy participation in peaking and reduce the risk of new energy consumption in the distribution network.

In the next step, we can consider adding load In the next step, we can consider adding load to the game model and consider the influence of load demand response on the whole electric energy trading model, so as to make the whole model more close to the reality.

References

[1] The State Council of People,s Republic of China. Some opinions on further deepening reform of the electric power system[R].2015.
[2] ZOU Peng, CHEN Qixin, XIA Qing, et al. Logical Analysis of Electricity Spot Market Design in Foreign Countries and Enlightenment and Policy Suggestions for China[J].Automation of Electric Power Systems,2014,38(13):18-27. DOI: 10.7500/AEPS20140219003.
[3] TIAN Shiming, YAN Wenpeng, ZHANG Dongxia,et al.Technical forms and key technologies on energy internet[J].Proceedings of the CSEE,2015,35(14):3482-3494
[4] HU Junjie, WANG Kunyu, AI Xin,et al.Transactive energy:an effective mechanism for balancing electric energy system. Proceedings of the CSEE,2019,39(4):953-966.

ISSN 2616-5902 Vol. 3, Issue 8: 72-83, DOI: 10.25236/AJBM.2021.030813

[5] SHU Chang, ZHONG Haiwang, XIA Qing. Day-ahead electricity market design based on market interpretation of optimization theory[J]. Automation of Electric Power Systems,2016,40(2):55-62. DOI:10.7500/AEPS20150922003.

[6] ZHANG Zhonghui, LAI Feiyi, Xie Yimiao. Analysis of trilateral game in electricity market based on Nash equilibrium theory[J]. Power System Technology, 2016,40(12): 3671-3679.

[7] XIA Ye, KANG Chongqing. CHEN Tianen, et al. Day-ahead market mode with power consumers participation in wind power accommodation[J]. Automation of Electric Power Systems, 2015, 39(17): 120-126.DOI:10.7500/AEPS20150317009.

[8] WU Cheng, GAO Bingtuan, TANG Yi, et al. Master-slave Game Based Bilateral Contract Transaction Model for Generation Companies and Large Consumers[J]. Automation of Electric Power Systems,2016,40(22):56-62. DOI: 10.7500/AEPS20151110003.

[9] CHEN Qixin, WANG Kedao, CHEN Sijie, et al. Transactive energy system for distributed agents: architecture, mechanism design and key technologies[J]. Automation of Electric power Systems, 2018,42(3):1-7. DOI:10.7500/AEPS20171031002.

[10] ZHANG Xiaoxuan, XUE Song, YANG Su, et al. International experience and lessons in power sales side market liberalization[J]. Automation of Electric Power Systems, 2016, 40(9):1-8. DOI: 10.7500/AEPS20151128001.

[11] CHEN Yuguo, ZHANG Xuan, LUO Gang, et al. Demand Response Mechanism and Approach of Electricity Spot Market in Bidding Mode without Price on User Side[J].Automation of Electric Power Systems,2019,43(9):179-186. DOI: 10.7500/AEPS20180409009.

[12] Junling Hu, Michael P Wellman. Nash Q-learning for General Sum Stochastic Game [J]. Journal of Machine Learning Research(S1532-4435),2003,4:1039-1069.

[13] WEN Junqiang, ZENG Bo, ZHANG Jianhua. Bi-level Programming Method for Distributed Generator Considering Stakeholders Game Relationship in an Electricity Market Environment[J]. Automation of Electric Power Systems, 2015, 39(15):61-67. DOI: 10.7500/AEPS20141203001.

[14] LIU Guojing, HAN Xueshan, WANG Shang, et al. Optimal decision-making in the cooperation of wind power and energy storage based on reinforcement learning algorithm[J]. Power System Technology, 2016, 40(9): 2729-2736.