Energy Proceedings
Vol 32, 2023
ISSN 2004-2965

Spot Electricity Market Bidding Strategy Optimization Based on Risk Analysis with Reinforcement Learning

Yuan Yu[1], Kengjie Li[2], Boon Giin Lee[1,3*]

1 Nottingham Ningbo China Beacons of Excellence Research and Innovation Institute, School of Computer Science, University of Nottingham Ningbo China, Ningbo 315100 China

2 The Chinese University of Hong Kong, Shenzhen 518116 China

3 Key Laboratory of Carbonaceous Waste Processing and Process Intensification of Zhejiang Province
(*Corresponding Author: boon-giin.lee@nottingham.edu.cn)

ABSTRACT

Great effort has been made to restructure the traditional monopolization of power industries to introduce fair competition. The deregulation of the electricity market allows the price of electricity to be formulated based on the bidding price. Nevertheless, it is still challenging to derive an optimal bidding strategy with many factors that need to be considered. This paper proposes a reinforcement learning (RL) based method to devise an optimal bidding strategy for maximizing the profit, taking the risk preferences in the spot electricity market into consideration. The problem is formulated based on Markov decision process (MDP), which is a discrete stochastic optimization method. The objective function is to optimize the cumulative profit over the span. This method also employs temporal difference technique and actor-critic learning algorithm for strategy optimization. In addition, the study introduces smart-market market-clearing method and a Gaussian distribution to formulate the strategy. Two different environmental conditions of the spot electricity market, static and dynamic, are applied in the simulation for analysis completeness. Only the target plant can adjust the bidding strategy in the static environment while all plants can adjust the bidding strategy in the dynamic environment. Simulation cases of nine participants are considered and the obtained results are analyzed.

Keywords: clean energy, reinforcement learning, risk analysis, strategy optimization

1. INTRODUCTION

Deregulation of the electricity industry has become an established practice in many countries. In the day-ahead market, the power-exchange bidding mechanism requires each bidder to submit the bid as a block once per day (24 hours). However, in the spot electricity market, the supply of electricity is distributed by the power stations and subsequently allocated to households and industries. Electricity in the spot market is bought and sold at the spot price.

An optimal strategy aims to maximize a profit with lowest risk involved. Several optimal bidding strategy models focusing on the market clearing-price forecast have been proposed, while others concentrate on the bidding behaviors prediction of competitors. As such, market bidding strategies are implemented to evaluate the accuracy of the electricity consumption based on various market simulations and load forecasting. A basic price-based auction mechanism is firstly proposed by Christie *et al.* [1] where Mielczarski *et al.* [2] proposed an auctioneer method to match the bids of buyers and sellers to find the market clearing price (MCP). Game theory has been used [3] and [4] for optimal bidding based on hourly auctions. David *et al.* [5] proposed a genetic algorithm to competition strategy in the spot market. However, these approaches are only effective if the market is not volatile. Optimization-based bidding strategies are proposed by [6] to address the volatile market issue. The optimal bidding is divided into two optimization problems: a bidder (deterministic) and an independent system operator (ISO) (stochastic). Zaid *et al.* [7] modeled the problem of optimal bidding as a

This is a paper for Applied Energy Symposium 2023: Clean Energy towards Carbon Neutrality (CEN2023), April 23-25, 2023, Ningbo, China.

Markov decision process (MDP) where load on a weekly basis with peak and off-peak loads is considered.

The study on bidding strategy optimization is limited in conventional simulation methods, mainly focusing on the day-ahead market. Another problem is the deficiency analysis of market environment where the simulation under static and dynamic conditions belongs to different agent models. Meanwhile, sampling process of load is not convincing with distinction only in weekly peak and off-peak. What's more, risk preference analysis is not considered in most of the existing market simulations, which is a key component in studying the behavior of power generation company. Risk preference analysis is an important indicator for evaluating the degree of acceptance to risks of power plants during bidding.

In this paper, a reinforcement learning (RL) method was proposed to devise an optimal bidding strategy, with the aims to maximize the profit with consideration of risk preference in spot electricity market. The problem is formulated in the framework of Markov decision process (MDP) under two different environment conditions of the spot electricity market (static and dynamic conditions of bidding in the spot electricity market). The temporal difference technique and actor-critic learning algorithm are employed for strategy optimization. The study utilizes smart-market market-clearing system and Gaussian distribution sampling methods to formulate the forecasting of loads in a simulated market. The major contributions of this paper include:

a. The optimal strategy considers loads, competitors offer, and historical bidding as inputs to a RL model and the performance is evaluated based on a market simulation under static and dynamic environment respectively.
b. This study also research on generating dataset with risk preference data based on MDP and Gaussian distribution method, to be used in the simulated market for analysis.

2. MATERIAL AND METHODS

2.1 Problem simulation

The ISO conducts an energy auction for the spot market. The electricity spot market is a wholesale market and operates everyday day from 7 am to 1 pm. Then, based on the bidding results, the electricity is traded and distributed at midnight. The bids are presented in the form of points in a piecewise linear curve on energy and price coordinates. The seller bids the amount of the energy that he or she is willing to sell at a given price or above, and the buyer bids the amount of the energy he or she is willing to buy for a given price or a price lower than it. The unconstrained MCP is determined by the point of intersection of the aggregate demand and supply bid curves.

It is presumed that the external operation conditions that bidders are aware include the total generation capacity of each other, the load forecast for the next hour, the past cost curves of competitors. The internal operation conditions include the start-up cost of each unit, and the cost curve is assumed to be known, in the form of:

$$CC^i = a^i + b^i U + c^i U^2$$

where the start-up cost for an hour t is su^{it}.

The bidder first decides how many parts it wants to bid, i.e., q parts. It divides its maximum generation capacity into q parts and using the cost curve method to find the marginal generated cost. This forms its middle element M of bid set. The higher H and lower L elements of the bid set are obtained from the middle element as:

$$H_j^1 = 1.1 M_j^1$$

$$L_j^1 = 0.9 M_j^1$$

where $J = 1,2,3, \ldots q$.

The bid set consists of a Cartesian product of the three bid elements for each part. For a two-part bid $(q = 2)$ and three-level (H, M, L) bids, the bid set consists of nine (3^q) bid-set:

$$Bs = [B_1 B_2 \ldots B_{3^q}]$$

where

$$B = [0, P_1, P_2, \ldots P_q, b, pr_1, pr_2, \ldots, pr_q]$$

2.2 Static and Dynamic Environment

Based on MDP method, a problem can be defined where the factors influencing clearing price is set as state and the bid-set of target power plant is defined as action. This study considers both the static and dynamic environment conditions. In a static environment, only the target plant can adjust strategy while others bid with marginal cost in the static environment. Meanwhile, in a dynamic environment, all plants can adjust their strategy accordingly. The state of the system is defined as bids offered by each bidder in a particular hour. The bids offered by each bidder are selected from their respective bid sets.

Bidders are then deciding their bidding actions and their offers in the next clearing round in an hour. Here, the bidding action is transferred from a continuous state space into a discrete state space since the continuous price value could be segmented into n classes according to the n bidding volume shaping as price-volume pair pri_i and vol_i where $i \in 1, \ldots, n$. The reward function is the summation of two parts, the net profit (NP), and the potential risk R with coefficient α.

2.3 Application of the Actor Critic Learning Algorithm

This section considers the application of the deep deterministic policy gradient (DDPG) method to produce an optimal bidding strategy based on inputs of loads, and their corresponding risks to address the problem formulated. Under the static conditions, the simulations are carried out by an agent, operated based on the target bidder. Under the dynamic conditions, the simulations are carried out by the agents operated by all bidders. The proposed DDPG algorithm methods under different environment are shown in Fig. 1.

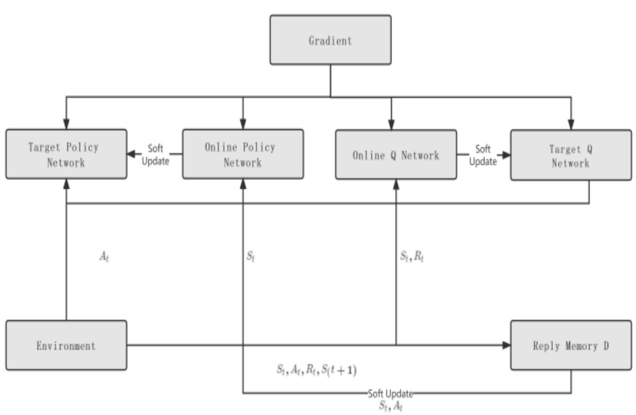

Fig. 1. Framework of DDPG algorithm under different environment

3. RESULTS AND DISCUSSION

3.1 Design of Simulation System

The design of the system simulation considers several elements as follows:

- Three dispatchable loads, bidding three fixed blocks each as shown in Table 1.
- Each of the six electricity generators has three blocks of capacity, and the initial offering is shown in Table 2.
- Load sampling from Gaussian distribution.

TABLE I
THREE DISPATCHABLE LOADS

Generator	Block1	Block2	Block3
	MW*$/MWh	MW*$/MWh	MW*$/MWh
1	10*$100	10*$100	10*$100
2	10*$100	10*$100	10*$100
3	10*$100	10*$100	10*$100

TABLE II
SIX DISPATCHABLE GENERATORS

Generator	Block1	Block2	Block3
	MW*$/MWh	MW*$/MWh	MW*$/MWh
1	12*$20	24*$50	24*$60
2	12*$20	24*$40	24*$70
3	12*$20	24*$42	24*$80
4	12*$20	24*$44	24*$90
5	12*$20	24*$46	24*$75
6	12*$20	24*$48	24*$60

The market simulations under both environments are carried out based on the sample systems. The discount factor γ is 0.8. The parameter of step size α is varied from 5 to 50 according to risk parameter. The result converges at early stage with greater risk parameter. The step size parameter for preferences is 0.02. A total of 10 000 iterations were carried out for each experiment.

3.2 Static Environment

<u>3.2.1 Results without Risk Preference</u>

The self-clear price in 24 hours for the spot market where target generator is 1 to 6 respectively is shown in Fig. 2. The original trading price where all generators bid by marginal cost is also shown in Fig. 2 with points. From the results, it is obvious that self-clear price will increase substantially when target generator take strategy during bidding.

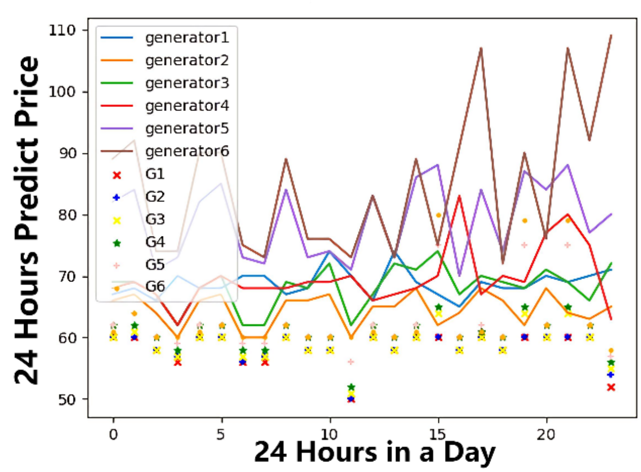

Fig. 2. The self-clear price in 24 hours for the spot market

Table 3 shows the average increasing percentage for each target generator over 24 hours trading. The strategy effect works differently for generators. Generator 1, 5, and 6 show highest significant improvement in performance. It is observed that the strategy effect depends on the generator's attributes including generator physical conditions and risk preference decision.

TABLE III
PROMOTION PERCENTAGE OF PRICE ON 6 GENERATORS

Generator	1	2	3	4	5	6
Percentage	16.6	15.0	15.6	11.3	29.3	39.2

Fig. 3 shows the self-clear volume in 24 hours for the spot market where target generator is 1 to 6 respectively. Fig. 2 together with Fig. 3 forms the bid clear pair of trading. The original trading volume where all generators bid by marginal cost is also shown in Fig. 2, presented as points. From the results, it is seen that self-clear volume will not absolutely increase when target generator take strategy during bidding. The strategy considers the incremental of overall revenue instead of single factor.

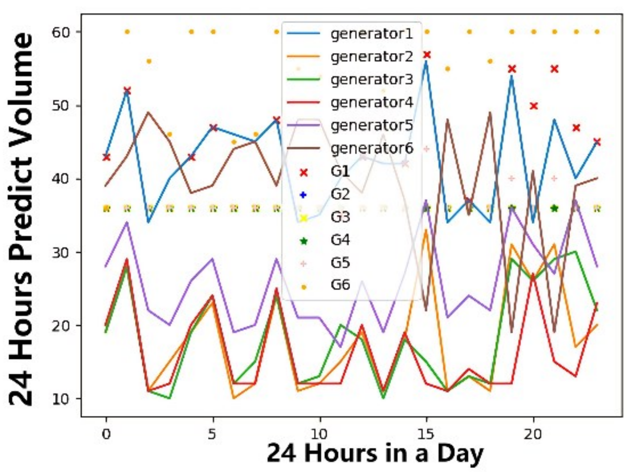

Fig. 3. The self-clear volume in 24 hours for the spot market

Table 4 shows the average increasing percentage for each target generator over 24 hours trading. Compared with the promotion percentage of price, direct proportion can be observed.

TABLE IV
PROMOTION PERCENTAGE OF PROFIT ON 6 GENERATORS

Generator	1	2	3	4	5	6
Percentage	25.0	3.0	1.0	1.5	50.0	62.0

3.2.2 Results with 0.5 to 0.8 Risk Index

Taken one of the generators as target example, the self-clear price, volume pairs in 24 hours for the spot market for target generator 1 to 6 is shown in Fig. 4 where the generators are under low-risk index transaction.

Table 5 shows the average increasing percentage of profit for each target generator over 24 hours trading. It could be seen that the increase in percentage is caused by the decline in strategy bidding thoroughly compared with the static environment without risk preference, regardless of the risk adventure or averse. However, the extent of decline varies depending on its risk preference.

TABLE V
PROMOTION PERCENTAGE OF PRICE ON 6 GENERATORS WITH RISK

Generator	1	2	3	4	5	6
Percentage	15.0	0.5	0.8	1.2	42.0	57.0

3.3 Dynamic Environment

The same sample system is applied in the market simulation under dynamic conditions. Experiment starts with condition without including the risk preference followed by considering the risk factor in the simulation. Iteration of 10 000 were carried out for the above problem.

3.3.1 Results without Risk Preference

The self-clear price and volume pairs in 24 hours in the dynamic spot market bidding without risk for generator 1 to 6 is shown in Fig. 5. Generators 1, 5 and 6 achieve a higher price and larger volume leading to higher net profit. Apart from they have better machine set supporting power generation, their risk tolerance is higher which is shown in Fig. 6.

3.3.2 Results with Risk Preference

The self-clear price and volume pairs in 24 hours in the dynamic spot market bidding with risk for generator 1 to 6 respectively is shown in Fig. 7. Generally, the transaction results almost keeps in same proportion among generators as situation without risk. The numerical value of either clear price or volume decreases respectively for all generators. It is reasonable to achieve such market clear results since every participant will bidding vigilantly trying not to against self-risk tolerance. The relearned actual risk preference for each generator is shown in Fig. 8. It shows that the profit variation for

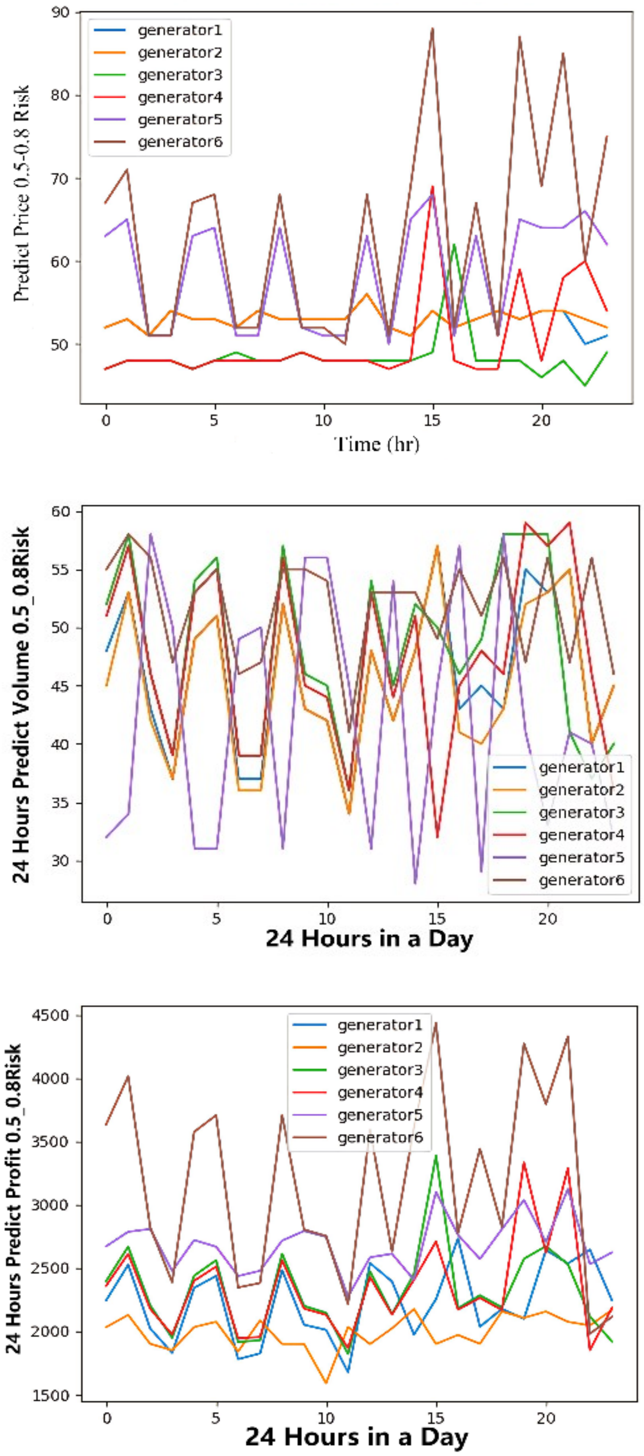

Fig. 4. The self-clear results value in 24 hours for the spot market with 0.5 to 0.8 risk index

participants grows resulting in higher risk compared with original situation.

It is assumed that other agents bid with marginal cost as fixed policy. The experiments results shows that self-clear profit will increase substantially when target generator take strategy during bidding, the promotion from strategy effect depends on the generators private

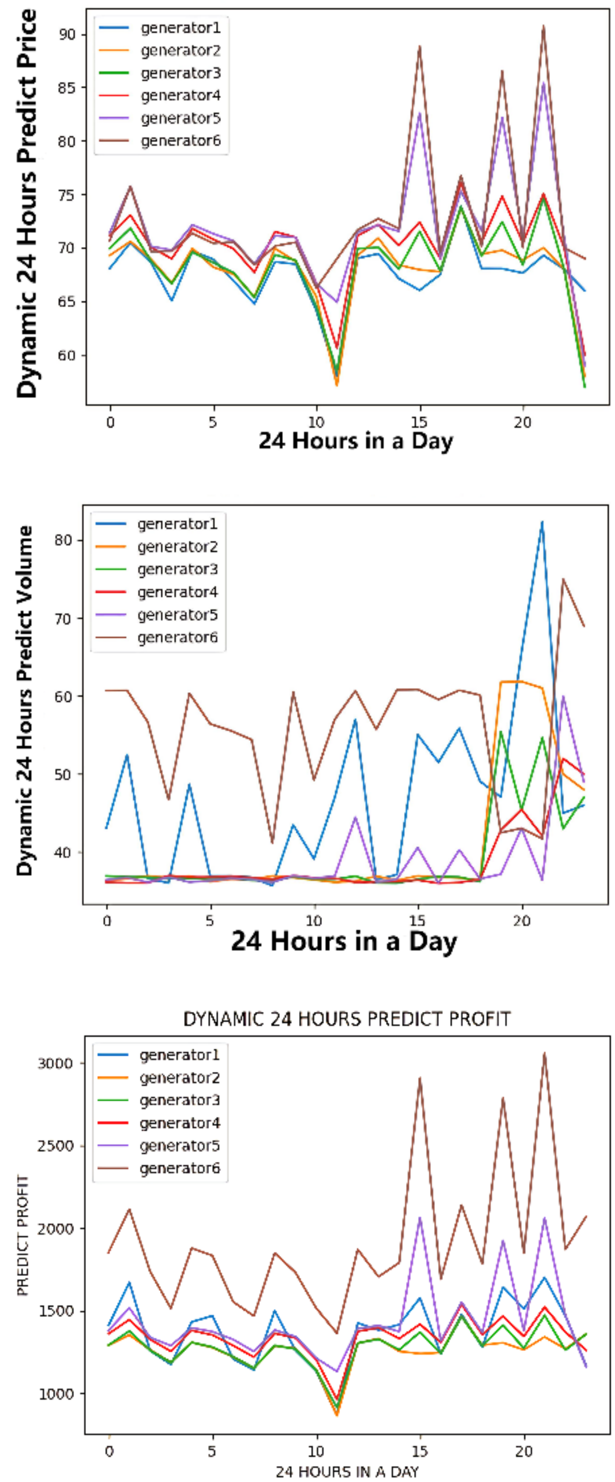

Fig. 5. The self-clear results value in 24 hours for the spot market under dynamic conditions

attribute including generator physical conditions and risk preference decision, risk preference will limit the profit promotion and enlarge the variance of profits, and larger risk index will have larger effects.

The agents bid based on the assumption that applying a policy maximizing the profit earned by the bidders, while minimizing the others' profit constrained

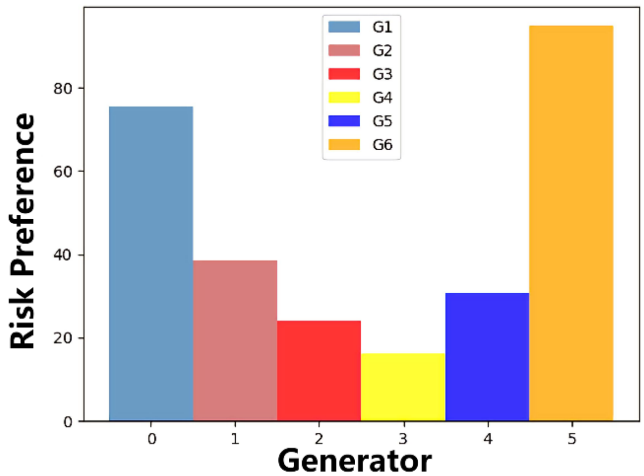

Fig. 6. The risk preference value in 24 hours for the spot market under dynamic conditions without risk

by risk preference. The experiments results show that for all generators, profit increases with results showing better performance under static conditions. In addition, the promotion has less deviation comparing to static conditions. Lastly, the risk preference limits the profit promotion and enlarge the variance of profits whose effect is larger than risk consideration under static situation.

4. CONCLUSIONS

This study models the spot market bidding problem based on MDP method which aims to solve the bidding strategy optimization problems. This work considers the smart-market market-clearing system and Gaussian distribution to formulate the optimization strategy. In addition, this study integrates a reinforcement learning, temporal difference technique and actor-critic learning algorithm to formulate the bidding strategy. The results show that the optimal bidding strategy could maximize the profit by including risk preference using the proposed method, experimented under both the static and dynamic conditions.

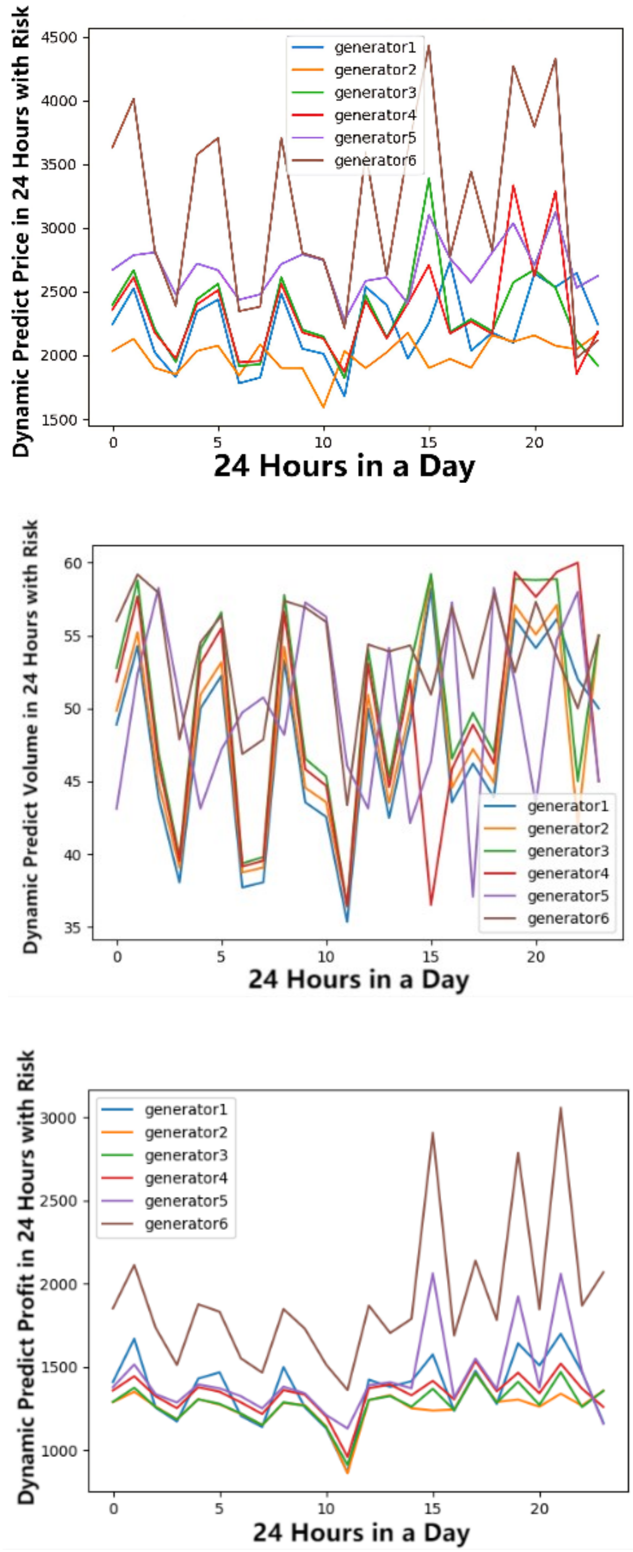

Fig. 7. The self-clear results value in 24 hours for the spot market under dynamic conditions with risk

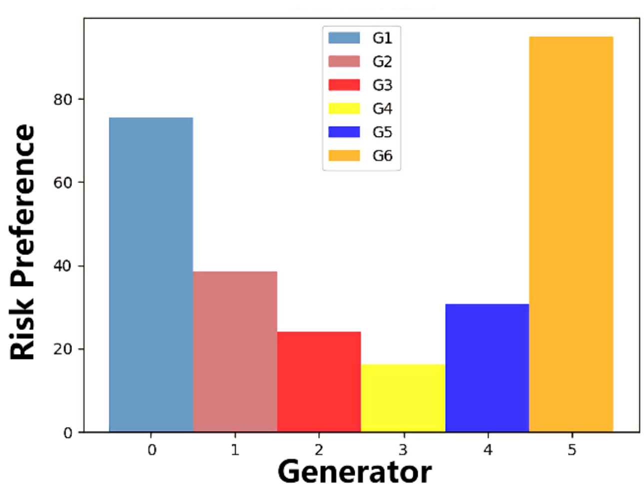

Fig. 8. The risk preference value in 24 hours for the spot market under dynamic conditions with risk

ACKNOWLEDGEMENT

The Zhejiang Provincial Department of Science and Technology is acknowledged for this research under its Provincial Key Laboratory Programme (2020E10018).

DECLARATION OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. All authors read and approved the final manuscript.

REFERENCE

[1] R. D. Christie, B. F. Wollenberg and L. Wangensteen, "Transmission management in the deregulated environment," *Proceedings of the IEEE*, vol. 88, no. 2, February 2000, pp. 170-195.

[2] W. Mielczarski, G. Michalik and M. Widjaja, "Bidding strategies in electricity markets," *Proceedings of the 1999 IEEE PES Power Industry Computer Applications Conference (PICA'99)*, 1999, pp. 7I-76.

[3] M. J. Exelby and N. J. D. Lucas, "Competition in the UK market for electricity generating capacity - a game-theory analysis," *Energy Policy*, vol. 21, no. 4, 1993, pp. 348-354.

[4] R. W. Ferrero, S. M. Shahidehpour and V. C. Ramesh, "Transaction analysis in deregulated power systems using game theory," *IEEE Transactions on Power Systems*, vol. 12, no. 3, 1997, pp. 1340-1347.

[5] D. M. Newbery, "Capacity-constrained supply function equilibria: competition and entry in the electricity spot market," *DAE Working Paper No. 9208*, University of Cambridge, 1992.

[6] A. Rudkevich, M. Duckworth, and R. Rosen, "Modelling electricity pricing in a deregulated generation industry: the potential for oligopoly pricing in a poolco," *Energy Journal*, vol. 19, no. 3, 1998, pp. 19-48.

[7] Z. Younes and M. IIic, "Generation strategies for gaming transmission constraints: will the deregulated electric power market be an oligopoly?" *Decision Support System*, vol. 24, no. 3, 1999, pp. 207-222.