Reinforcement Learning- and Option-jointed Modeling for Cross-market and Cross-time Trading of Generators in Electricity and Carbon Markets

Kai Jiang, *Member*, *IEEE*, Kunyu Wang, Lin Yang, and Nian Liu, *Member*, *IEEE*

***Abstract*—With the development of the carbon markets (CMs) and electricity markets (EMs), discrepancies in prices between the two markets and between two time periods offer profit opportunities for generation companies (GenCos). Motivated by the carbon option and Black-Scholes (B-S) model, GenCos are given the right but not the obligation to trade carbon emission allowances (CEAs) and use instruments to hedge against price risks. To model the strategic behaviors of GenCos that capitalize on these cross-market and cross-time opportunities, a multi-market trading strategy that incorporates option-jointed daily trading and reinforcement learning-jointed weekly continuous trading are modeled. The daily trading is built with a bi-level structure, where a profit-oriented bidding model that jointly considers both the optimal CEA holding shares and the best bidding curves is developed at the upper level. At the lower level, in addition to market clearing models of the day-ahead EM and auction-based CM, a B-S model that considers carbon trading asynchronism and option pricing is constructed. Then, by expanding the daily trading, the weekly continuous trading is modeled and solved using reinforcement learning. Binary expansion and strike-to-spot price ratio are utilized to address the nonlinearity. Finally, case studies on an IEEE 30-bus system are conducted to validate the effectiveness of the proposed trading strategy. Results show that the proposed trading strategy can increase GenCo profits by influencing market prices and leveraging carbon options.**

***Index Terms*—Multi-market trading, carbon market, electricity market, Black-Sholes model, carbon option, reinforcement learning.**

NOMENCLATURE

A. Indices and Sets

Ω_g	Generator set in generation company (GenCo) g
$\varphi_n^{\mathrm{S}}, \varphi_n^{\mathrm{N}}$	Generator set and node set at node n
g	Index of GenCos
h	Index of selling block in carbon market (CM)
i	Index of generators
k	Index of buying block in CM
m, n	Indices of power network nodes
t	Index of time periods
v	Index of selling block in electricity market (EM)

B. Parameters and Constraints

$\alpha_{i,h}^{\mathrm{CS,min}}, \alpha_{i,h}^{\mathrm{CS,max}}$	The minimum and maximum carbon emission allowance (CEA) prices of generator i in each selling block h
$\alpha_{i,k}^{\mathrm{CB,min}}, \alpha_{i,k}^{\mathrm{CB,max}}$	The minimum and maximum CEA prices of generator i in each buying block k
$\alpha_{i,v}^{\mathrm{G,min}}, \alpha_{i,v}^{\mathrm{G,max}}$	The minimum and maximum electricity prices of generator i in each selling block v
β	Ratio between strike price of carbon option and CM clearing price (i.e., strike-to-spot price ratio)
ξ	Volatility rate of CEA prices
σ	Variance value of predicted CEA price
μ	Expected value of predicted CEA price
ρ	Soft update coefficient
$\lambda_{\min}^{\mathrm{C}}, \lambda_{\max}^{\mathrm{C}}$	Lower and upper values of CM clearing price
$\lambda_{i,v}^{\mathrm{G}}$	Generation cost of generator i related to bidding block v
$\lambda^{\mathrm{RS}}, \lambda^{\mathrm{RB}}$	Selling and buying prices for CEA reserve
$B_{n,m}$	Admittance between nodes n and m
M	A sufficiently large positive number
$P_{n,t}^{\mathrm{D}}$	Power demand of node n at time t
$P_{i,v}^{\mathrm{G,min}}, P_{i,v}^{\mathrm{G,max}}$	Lower and upper EM bidding limits of generator i in block v
$P_{i,t}^{\min}, P_{i,t}^{\max}$	The minimum and maximum power generations of generator i at time t
$P_{n,m}^{\mathrm{L,max}}$	Line capacity between nodes n and m
Q	Expansion coefficient

Manuscript received: January 25, 2024; revised: April 21, 2024; accepted: September 10, 2024. Date of CrossCheck: September 10, 2024. Date of online publication: October 22, 2024.

This work was supported by the National Science Foundation of Jiangsu Province (No. BK20232026).

K. Jiang, K. Wang, L. Yang, and N. Liu (corresponding author) are with the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Beijing 102206, China (e-mail: kai_jiang@ncepu. edu. cn; wky@ncepu. edu. cn; 13663031116@163. com; nianliu@ncepu.edu.cn).

DOI: 10.35833/MPCE.2024.000013

JOURNAL OF MODERN POWER SYSTEMS AND CLEAN ENERGY

$Q_k^{\mathrm{B,min}}, Q_k^{\mathrm{B,max}}$ Lower and upper limits during CEA buying of block k

$Q_h^{\mathrm{S,min}}, Q_h^{\mathrm{S,max}}$ Lower and upper limits during CEA selling of block h

Q_i^{trans} Surplus or deficit CEA allowed for transfer of generator i

r Continuously compounded risk-free rate

r_d^{pre} Price premium ratio of carbon option

RU, RD Ramping up and down limits

T Number of daily time periods

T_1 Expiration time of carbon option

C. Variables

$\lambda_{n,t}^{\mathrm{E}}$ Locational marginal price at node n at time t

$\lambda^{\mathrm{FS}}, \lambda^{\mathrm{FB}}$ Future selling and buying prices of CM

λ^{C} CM clearing price

$\delta_{m/n,t}$ Voltage angle at each node m or n at time t

$\alpha_{i,h/k}^{\mathrm{CS}}, \alpha_{i,h/k}^{\mathrm{CB}}$ Selling and buying prices of CEA of generator i in block h or k

$\alpha_{i,v}^{\mathrm{G}}$ Bidding price of generator i in block v

θ, ϕ Weight coefficients of critic and actor networks

π_ϕ Actor network with weight ϕ

$\varphi, \varepsilon, \kappa$ Lagrangian multipliers from individual sides of each constraint

$C_g^{\mathrm{SH}}, C_g^{\mathrm{BH}}$ Costs of buying put and call options of GenCo g

$C^{\mathrm{call}}, C^{\mathrm{put}}$ Call and put option prices

d_1, d_2 Intermediate variables

I_i^{base} Carbon emission intensity (CEI) benchmark of generator i

I_i Real CEI of generator i

L Loss function of actor network

$P_{i,v,t}^{\mathrm{G}}$ Power generated by of generator i in block v at time t

p_s Strike price of carbon option

Q_i^{net} Net surplus CEA of generator i

$Q_{i,h/k}^{\mathrm{S}}, Q_{i,h/k}^{\mathrm{B}}$ Selling and buying amounts of CEA of generator i in block h or k

$Q_i^{\mathrm{ST}}, Q_i^{\mathrm{BT}}$ Selling and buying amounts of CEA of generator i

$Q_g^{\mathrm{SH}}, Q_g^{\mathrm{BH}}$ Holding amounts of CEA for seller and buyer of GenCo g

$Q_{i,d}^{\mathrm{SHT}}, Q_{i,d}^{\mathrm{BHT}}$ Transferred holding amount of CEA of generator i for seller and buyer on day d to the next day

$r_d^{\min}$ Forced transaction ratio of CM on day d

$reward_d$ Reward achieved after action on day d

R_g^{E} Income of GenCo g in EM

$R_g^{\mathrm{CS}}, C_g^{\mathrm{CB}}$ Income and expense of selling and buying amounts of CEA of GenCo g in CM at present

$R_g^{\mathrm{FS}}, C_g^{\mathrm{FB}}$ Income and expense of selling and buying amounts of CEA of GenCo g in CM in future

$state_d$ State array of reinforcement learning on day d

u_i Auxiliary binary variable of generator i

$V_{\theta_1}, V_{\theta_2}$ Critic networks with different weights θ_1 and θ_2

Y Estimated value of state

I. Introduction

As the carbon markets (CMs) and electricity markets (EMs) evolve, power generators are confronted with the challenge of integrating power production with carbon mitigation within their EM-CM trading strategies. Fluctuations in equivalent prices across different time and markets create unique profit opportunities. Traditional trading strategies that concentrate exclusively on EMs fail to account for carbon obligations and are ill-equipped to exploit these newly emerging profit landscapes.

In CMs, generators must provide an equivalent quantity of carbon emission allowances (CEAs) to counterbalance their real carbon emissions at the end of each compliance period [1]. These CEAs are first distributed by governments in the primary CM [2], and can then be traded among generation companies (GenCos) in the daily secondary CM [3]. In general, there are two types of allocation mechanisms: auction-based allocation (ABA) and output-based allocation (OBA). The ABA mechanism is used by the European Emissions Trading System [4] and Regional Greenhouse Gas Initiative [5], which requires generators to buy CEAs through auctions. By contrast, the OBA mechanism implemented in China and Canada [6], [7] grants GenCos complementary CEAs equal to the baseline carbon emission intensity (CEI) per unit of power generated.

Regarding the OBA, generators experience either gains or losses based on their actual CEI relative to the baseline value [8]. GenCos that act as sellers with a lower CEI than the baseline can benefit from selling surplus CEAs. By contrast, GenCos with a higher CEI incur losses as buyers. These dynamics subsequently affect the behavior of GenCos in the EM, where bidding prices are adjusted for arbitrage or loss conduction [9]. Consequently, generators must consider the implications of cross-market trading. In addition, given the asynchronism between the annual compliance and daily transactions of CEAs, GenCos may hold their CEA shortages or surpluses for more favorable prices in the future. Therefore, generators must also consider cross-time trading in the CM.

Cross-market trading in CMs and EMs has been widely analyzed and can generally be categorized into three groups. The first group focuses solely on the end-user side. Based on peer-to-peer transactions, a total emission cap decomposition method for prosumers and a suitable carbon-aware pricing scheme are proposed in [10] and [11], respectively. In [12], carbon-electricity integrated optimal bidding strategies for a power plant are studied. In [13] and [14], transactive trading strategies are designed for microgrids in the energy market. Other studies have been conducted based on various demand-side resources [15], [16]. The second group has shed some light on the generation side. In [17] and [18], an optimal price-taker trading strategy and a risk-seeking stochastic offering strategy for wind power are proposed, respectively. Strategic reactions and the corresponding equilibria for power producers in the EM, CM, and natural gas market are modeled in [19]. The last group attempts to link the transmission and distribution sides using a carbon flow tracing technique [20]. In [21], a hierarchical EM-CM framework is developed, where the peer-to-peer transaction results are used for carbon responsibility measures. However, most existing studies have focused on the ABA mechanism for CEAs, wherein market participants can only buy CEAs in CMs. Although some studies such as [22] have indicated the differences between ABA and OBA mechanisms, these studies all assume that the CEA price is predefined and remains static. However, the generators under the OBA can be either CEA buyers or sellers, and the CM clearing price, in practice, is dynamically determined by demand and supply, making the aforementioned strategies unsuitable.

Cross-time arbitrage in EMs for generators has been thoroughly analyzed, which is primarily based on price-forecasting techniques [17] and long-term forward contracts [23]. Notably, via the forward or futures EM contract, generators can settle the future selling price (called the strike price) prior to physical delivery in the spot market. Thus, once the strike price is higher than the spot price (also known as the stock price), cross-time profits can be achieved [24]. Even if the spot price is higher, this type of financial derivative tool can be used to hedge against risks.

A similar derivative market also exists in CMs as the carbon option [25], which is a practical instrument for generators to hedge against the risks of CEA price fluctuations [26]. Carbon options provide buyers with the right but not the obligation to buy or sell CEAs at a predetermined price and date [23]. On the one hand, the carbon option can ensure CEA delivery in the future and prevent economic losses from shortfall penalty or surplus logout. On the other hand, the carbon option provides a cross-time profit space between the strike prices and exercise date stock prices [27]. Take CEA buyers as an example. The generator may prefer to buy CEAs in the future if the future price is expected to be lower. The waiting risks derived from price fluctuations can be hedged through buying carbon options in the opposite price-changing direction. This refers to a call option if the strike price is higher than the spot price; otherwise, it refers to a put option. By contrast, EM generators are normally sellers, whereas those with trading positions in the CM can be bidirectional. Few studies have considered the reality of carbon options in hedging against risks and making profits during CM trading as well as when combining the carbon option with the EM.

Given that carbon options can provide the aforementioned benefits, it is crucial to evaluate their prices for the further use of generators. Specifically, only after the generator weighs profits against the costs of carbon options can the optimal CEA trading-holding portfolio be determined. The commonly used pricing models of options include the Black-Scholes (B-S) model [28], binomial model [29], and fractal Brownian motion model [30], with the B-S model being one of the most mature and widely used [31]. As most carbon options are based on European options [32] and the CEA price follows a geometric Brownian motion with constant volatility and no bankruptcy [30], the carbon option value can be calculated using the B-S model with five variables: the strike price of an option, current stock price, time to expiration, risk-free rate, and volatility [33]. However, in its nonlinear logarithmic functions, the B-S model contains several variables, making it mathematically complex. Therefore, it can be challenging to incorporate the B-S model directly into an optimization problem because of the nonlinearity and complexity involved.

Notably, in our previous work, the strategic behaviors of a GenCo are modeled in EM-CM markets and the effects of dynamic CEIs on the markets are investigated [9]. However, an ideal yet practical assumption was made that any carbon emissions not accounted for in the day-ahead electricity market must be promptly offset with an equivalent number of CEAs in the spot market. This is necessary to unify the multi-market trading period and foreground the effects of dynamic CEI. However, in practice, asynchronism in trading periods exists between two markets, where the physical dispatch schedules of generators must be balanced in the day-ahead EM, whereas their surplus or deficit CEAs can be stocked temporally and must only be annually balanced prior to the end of each compliance period. Under these circumstances, generators have widely adopted a hold-and-see strategy for buying CEAs in a piecemeal manner. Consequently, cross-market trading decisions have been expanded to include cross-time long-term optimization problems. So far, only a few studies have captured these features and built adaptive simulation frameworks.

Based on the aforementioned investigations, three issues need be addressed.

1) A model must be constructed that evaluates hold-and-see actions in the CM while jointly considering expectations, uncertainties, and lasting time from individual price perceptions over the annual compliance period and under different time-waiting windows.

2) A trading strategy must be developed to simulate the practical behaviors of generators while considering the asynchronized interactions of EM and CM during cross-market and cross-time trading.

3) A reformulation method and an explainable algorithm must be devised to linearize and accelerate the multi-market bidding strategy with the derivative carbon option.

This paper thus makes the following contributions to the existing literature.

1) The B-S model is introduced to price the carbon option value and CEA hold-and-see costs under various triangular probability distributions that reflect an individual's perception of future price expectations, price variances, and waiting periods. In addition, temporal asynchronism between daily power balance and annual carbon compliance can be coordinated.

2) A multi-market trading strategy that incorporates the option-jointed daily trading and reinforcement learning (RL)-jointed weekly continuous trading with three core markets of electricity spot, carbon spot, and carbon option is proposed

to model the cross-market and cross-time EM-CM trading of generators. Accordingly, fact-based CM mechanisms including call auctions, stability reserves, and OBA are adaptively supplemented.

3) To facilitate the solution method and better incorporate it into the proposed trading strategy, nonlinear terms during market clearing are partly reformulated using binary expansion, whereas those in the B-S model are modified by the strike-to-spot price ratio. By scanning this parameter, we can achieve the best CEA trading-holding portfolio and holding period in the carbon option contract.

The remainder of this paper is organized as follows. Section II introduces the modeling assumptions and framework. Section III describes how the multi-market trading strategy of GenCos is modeled for cross-time and cross-market arbitrage. Section IV describes the reformulation and solution methods. Section V describes case studies that test and validate the proposed trading strategy. Section VI concludes this paper.

II. Modeling Assumptions and Framework

A. Modeling Assumptions

1) EM

In this paper, EM refers to the day-ahead spot market organized by an independent system operator (ISO). Renewable uncertainties can be accounted for, as in [9], and they are ignored here to streamline the model. Prior to the start of EM, the reserve is determined by the ancillary service markets, and generators that have been informed of the exact granted reserve will reduce their upper power limits in the EM. This aligns with the method used in West Inner Mongolia, China [34].

2) CM

In an OBA-based secondary CM, CEA prices can be determined through over-the-counter or on-floor trading [35]. Research works on CEA exchanges such as those conducted by the European Energy Exchange [36], Intercontinental Exchange [25], and Korea Exchange [37] suggest that on-floor "continuous trading with auctions" is a common form of trading. Thus, the closing auction price, recognized for its highly efficient price discovery [38], is predominantly used in this paper.

Given that the trading actions of generators in the CM are influenced by their carbon emissions from power generation in the EM, and because clearing carbon prices are determined by the bidding curves of both supply and demand-side generators, this paper assumes that the trading strategy of power producers can affect the carbon price.

The objective of this paper is to model the behavior of strategic GenCos. Other entities are modeled using a sequential transaction strategy, where the bidding amounts in the CM are calculated based on the clearing results from the first EM. Normal generators without market information employ fixed bidding prices determined by predetermined costs.

B. Trading Process and Model Framework

Figure 1 shows the trading framework and information flow under the option-jointed daily trading and the RL-jointed weekly continuous trading for generator cross-market and cross-time trading in the EM and CM. T_2, T_3, and T_4 represent different CEA holding time for different carbon price distributions. The details are described as follows.

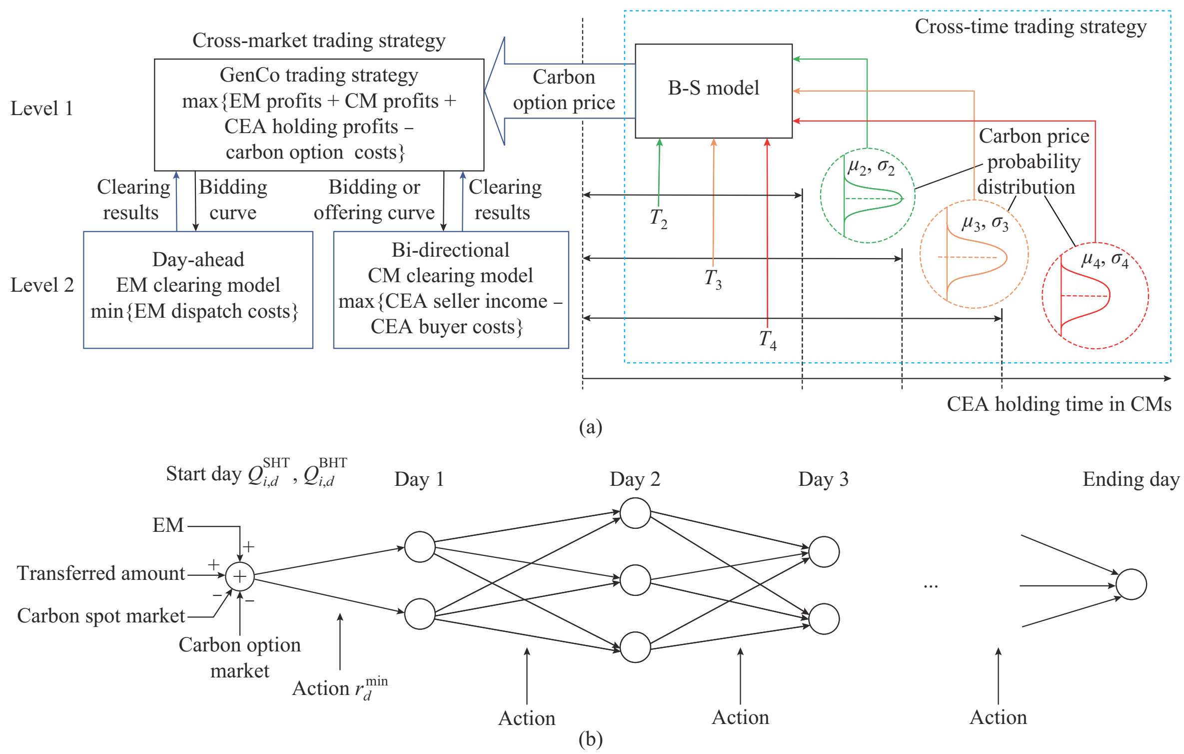

Fig. 1. Trading framework and information flow. (a) Daily trading. (b) Weekly continuous trading.

1) The GenCo submits a price-quantity curve to the EM ISO, which shows the price and quantity of electricity that the GenCo is willing to buy or sell in the market.

2) The ISO checks the bidding ceiling of the GenCo based on results from the reserve ancillary market, and conducts an EM clearing. The allocated power and locational marginal prices are then fed back to the GenCo.

3) The GenCo calculates its net demand for CEAs using the information from the EM ISO and employs the B-S model to price carbon options. The GenCo determines the CEA trading quantity for buying or selling in the market and the amount to hold for future delivery. This is accomplished by obtaining call or put carbon options.

4) Once the bidding order of GenCo in the CM is received, both the CM clearing outcomes and option prices in the B-S model can be updated.

Based on this trading process, the option-jointed daily trading in the EM and CM is modeled. At Level 1, the prices of carbon options with different configurations (i.e., different CEA holding time, predicted expectations, and variances of different future CEA prices) are valued using the B-S model. From this valuation, the strategy of GenCo is determined by optimizing its bidding prices and trading or holding amounts of CEA. At Level 2, the model encompasses an EM clearing model and a bidirectional CM clearing model, generating clearing outcomes in terms of prices and quantities.

In general, two sources of CEA net trading demand exist: net shortage from EM generation, and the amount transferred from previous trading. By contrast, three strategies are used to fulfill these demands: trading in the carbon spot market, holding options immediately through buying, or transferring them to the next day. In terms of daily trading, GenCos must optimize their multi-market strategic behaviors simultaneously. The same conclusion has been obtained in [39]. Thus, the upper level is optimized with feedback from the lower level.

As Fig. 1(b) shows, via the expansion of the aforementioned daily trading to weekly continuous trading, the GenCo can execute continuous trading by determining the exact day to buy carbon options rather than allocating equal amounts of carbon options to net holding CEAs. The daily trading state transition processes and optimal trading action decisions are then determined via RL.

III. Model Formulation

A. Upper-level 1: Daily Trading of GenCo

At this level, the option-jointed daily trading of GenCo for electricity-carbon trading under the OBA mechanism is proposed. The profit-maximization objective function *Obj* of the strategic GenCo is given by (1)-(8).

$$\max Obj = R_g^{\mathrm{E}} + R_g^{\mathrm{CS}} - C_g^{\mathrm{CB}} + R_g^{\mathrm{FS}} - C_g^{\mathrm{FB}} - C_g^{\mathrm{SH}} - C_g^{\mathrm{BH}} \tag{1}$$

$$R_g^{\mathrm{E}} = \sum_t \sum_{i \in \Omega_g} \left(\lambda_{n,t}^{\mathrm{E}} - \lambda_{i,v}^{\mathrm{G}} \right) P_{i,v,t}^{\mathrm{G}} \Delta t \tag{2}$$

$$C_g^{\mathrm{CB}} = \lambda^{\mathrm{C}} \sum_{i \in \Omega_g, k} Q_{i,k}^{\mathrm{B}} + \lambda^{\mathrm{RB}} \sum_{i \in \Omega_g} \left(Q_i^{\mathrm{BT}} - \sum_k Q_{i,k}^{\mathrm{B}} \right) \tag{3}$$

$$R_g^{\mathrm{CS}} = \lambda^{\mathrm{C}} \sum_{i \in \Omega_g, h} Q_{i,h}^{\mathrm{S}} + \lambda^{\mathrm{RS}} \sum_{i \in \Omega_g} \left(Q_i^{\mathrm{ST}} - \sum_h Q_{i,h}^{\mathrm{S}} \right) \tag{4}$$

$$C_g^{\mathrm{FB}} = \lambda^{\mathrm{FB}} Q_g^{\mathrm{BH}} \mathrm{e}^{-rT_1} \tag{5}$$

$$R_g^{\mathrm{FS}} = \lambda^{\mathrm{FS}} Q_g^{\mathrm{SH}} \mathrm{e}^{-rT_1} \tag{6}$$

$$C_g^{\mathrm{BH}} = C^{\mathrm{call}} Q_g^{\mathrm{BH}} \tag{7}$$

$$C_g^{\mathrm{SH}} = C^{\mathrm{put}} Q_g^{\mathrm{SH}} \tag{8}$$

In the EM, the profit depends on the discrepancy between locational marginal prices and generation costs, as expressed in (2). CEA transactions affect the profits of GenCo when its actual CEI diverges from the benchmark. Mandated by CEAs reflecting real carbon emissions, the profits of GenCo fluctuate with CEA trading that aligns with the CEI. For example, a CEA buyer can opt to acquire CEAs through a current-delivery exchange, as shown in (3), where settled CEAs are paid at the CM clearing price and unsettled CEAs are absorbed by the market stability reserve using upper and lower limit prices to stabilize CEA circulation. By contrast, the buyer can hold CEAs for future procurement, as shown in (5). To ensure future delivery, avoid market penalties, and achieve cross-time profit, the GenCo can buy call options via (7). The profit components of the CEA sellers in (6) and (8) are similar to those of CEA buyers. The constraints of the strategic behaviors of GenCo are expressed by (9)-(20).

$$\alpha_{i,h}^{\mathrm{CS,min}} \le \alpha_{i,h}^{\mathrm{CS}} \le \alpha_{i,h}^{\mathrm{CS,max}} \quad \forall i, h \tag{9}$$

$$\alpha_{i,k}^{\mathrm{CB,min}} \le \alpha_{i,k}^{\mathrm{CB}} \le \alpha_{i,k}^{\mathrm{CB,max}} \quad \forall i, k \tag{10}$$

$$\alpha_{i,v}^{\mathrm{G,min}} \le \alpha_{i,v}^{\mathrm{G}} \le \alpha_{i,v}^{\mathrm{G,max}} \quad \forall i, v \tag{11}$$

$$\alpha_{i,h-1}^{\mathrm{CS}} \le \alpha_{i,h}^{\mathrm{CS}} \quad \forall h \ge 2, \forall i \tag{12}$$

$$\alpha_{i,k-1}^{\mathrm{CB}} \ge \alpha_{i,k}^{\mathrm{CB}} \quad \forall k \ge 2, \forall i \tag{13}$$

$$\alpha_{i,v-1}^{\mathrm{G}} \le \alpha_{i,v}^{\mathrm{G}} \quad \forall v \ge 2, \forall i \tag{14}$$

$$Q_i^{\mathrm{net}} = \sum_t \left[\left(\sum_v P_{i,v,t}^{\mathrm{G}} \right) \left(I_i^{\mathrm{base}} - I_i \right) \right] \quad \forall i \tag{15}$$

$$Q_i^{\mathrm{net}} = Q_i^{\mathrm{ST}} + Q_i^{\mathrm{SH}} - Q_i^{\mathrm{BT}} - Q_i^{\mathrm{BH}} \quad \forall i \tag{16}$$

$$0 \le Q_i^{\mathrm{ST}} + Q_i^{\mathrm{SH}} \le M u_i \quad \forall i \tag{17}$$

$$0 \le Q_i^{\mathrm{BT}} + Q_i^{\mathrm{BH}} \le M \left(1 - u_i \right) \quad \forall i \tag{18}$$

$$Q_i^{\mathrm{ST}}, Q_i^{\mathrm{SH}}, Q_i^{\mathrm{BT}}, Q_i^{\mathrm{BH}} \ge 0 \quad \forall i \tag{19}$$

$$Q_g^{(\cdot)} = \sum_{i \in \Omega_g} Q_i^{(\cdot)} \tag{20}$$

The bidding prices of GenCo adhere to the specified price range expressed in (9)-(11). For the GenCo that acts as a seller or buyer in markets, bidding orders follow non-decreasing or non-increasing curves, as constrained in (12)-(14). Equation (15) defines the daily net surplus of the CEA of the GenCo. The transaction status of the GenCo as a seller or buyer in the CM is identified by (16). Constraints (17)-(19) ensure that the holding or trading amounts of CEAs remain positive based on the big-*M* method [8]. GenCo manages the CM trading of the owned generators, as outlined in (20).

In the proposed trading strategy, future CEAs can be exchanged based on the CEA price predictions of the GenCo.

However, in real-life situations, a precise probability distribution of future CEA prices may be unavailable. A triangular distribution can be used to solve this problem, with only the lowest and highest price values being required [40]. Via the probability density function as presented in Fig. 2, the equivalent expectation value of the forecasted price can be obtained as the mean of the two extreme values. The equivalent variance of the forecasted price can be calculated based on the difference between the highest and mean prices.

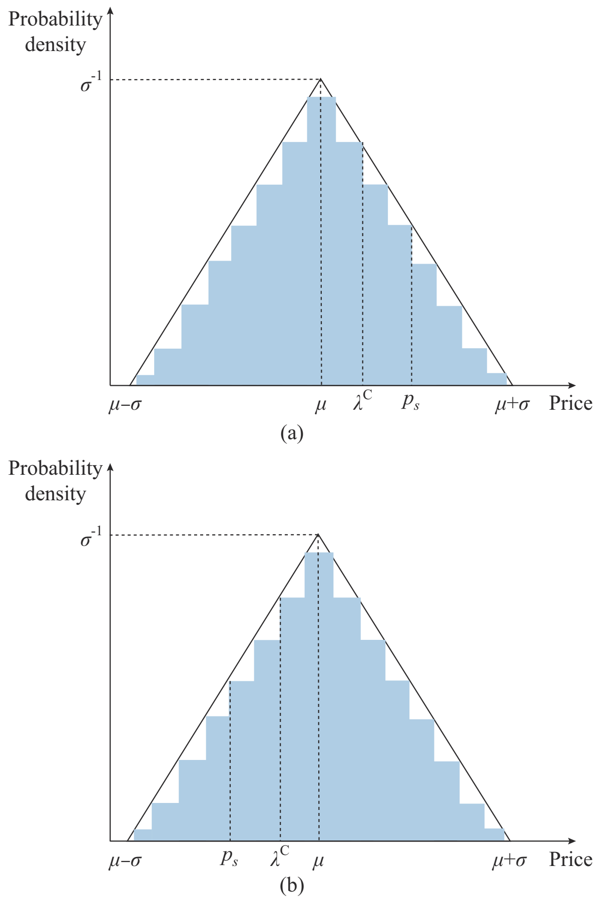

Fig. 2. Triangle probability distribution of future CEA price. (a) Call option for CEA buyer. (b) Put option for CEA seller.

Note that hedging risk is the key objective with the use of carbon options. In the CM, the GenCo is obliged to submit equal amounts of CEAs to real carbon emissions. Thus, if the generator chooses to wait-and-buy CEAs, the call option can avoid future penalties for CEA shortages. By contrast, if the generator chooses to wait-and-sell CEAs, the put option can prevent losses when CEAs are cancelled after their expiration date (and thus cannot be sold).

If we consider the CEA buyer in Fig. 2(a) as an example, the original intention of holding demands and waiting to buy is based on the prediction that the CEA expectation price μ will be less than the current price λ^{C}. To hedge against risks from unexpected extreme values, the GenCo can buy call options with a strike price p_s higher than λ^{C}. With these call options, transactions can be settled using p_s when real prices are higher. Similarly, the cross-time trading of CEA sellers can be modeled by buying put options, as shown in Fig. 2(b). Finally, considering carbon options, the settlement price of the CEA buyer or seller can be calculated by obtaining the present value of the future price expectation in (21) and (22), respectively.

$$\lambda^{\mathrm{FB}}=\left[\int_{\mu-\sigma}^{\mu}\left(\frac{1}{\sigma^2}x+\frac{\sigma-\mu}{\sigma^2}\right)x\mathrm{d}x+\int_{\mu}^{p_s}\left(\frac{\sigma+\mu}{\sigma^2}-\frac{1}{\sigma^2}x\right)x\mathrm{d}x+\int_{p_s}^{\mu+\sigma}\left(\frac{\sigma+\mu}{\sigma^2}-\frac{1}{\sigma^2}x\right)p_s\mathrm{d}x\right]\mathrm{e}^{-rT}=\left[\frac{1}{6\sigma^2}p_s^3-\frac{\sigma+\mu}{2\sigma^2}p_s^2+\frac{1}{2\sigma^2}(\sigma+\mu)^2p_s-\frac{1}{3\sigma^2}\mu^3+\frac{1}{6\sigma^2}(\mu-\sigma)^3\right]\mathrm{e}^{-rT} \tag{21}$$

$$\lambda^{\mathrm{FS}}=\left[\int_{\mu-\sigma}^{p_s}\left(\frac{1}{\sigma^2}x+\frac{\sigma-\mu}{\sigma^2}\right)p_s\mathrm{d}x+\int_{p_s}^{\mu}\left(\frac{1}{\sigma^2}x+\frac{\sigma-\mu}{\sigma^2}\right)x\mathrm{d}x+\int_{\mu}^{\mu+\sigma}\left(\frac{-1}{\sigma^2}x+\frac{\sigma+\mu}{\sigma^2}\right)x\mathrm{d}x\right]\mathrm{e}^{-rT}=\left[\frac{1}{6\sigma^2}p_s^3+\frac{\sigma-\mu}{2\sigma^2}p_s^2+\frac{1}{2\sigma^2}(\sigma-\mu)^2p_s-\frac{1}{3\sigma^2}\mu^3+\frac{1}{6\sigma^2}(\mu+\sigma)^3\right]\mathrm{e}^{-rT} \tag{22}$$

B. Upper-level 2: B-S Option Pricing Model

The B-S option pricing model is presented in (23)-(27). Developed in 1973 by Black, Merton, and Scholes, the B-S model was the first mathematical method to be widely used in calculating the theoretical value of an option contract under current stock prices λ^{C} (also known as the clearing price in CMs), the option's strike price p_s, risk-free rate r, expiration time T_1, and price volatility rate ξ [33].

To hedge against future price increases and decreases and generate cross-time profits, the CEA buyer or seller may buy call and put options valued by (23) and (24), respectively. $N(\cdot)$ is the cumulative probability density function of a normally distributed variable. In (27), β is defined as the quotient of p_s and λ^{C}, which is a pre-determined constant, and via settling, the nonlinearity from $\ln\left(\lambda^{C}/p_s\right)$ in (25) can be eliminated.

$$C^{\mathrm{call}}=N(d_1)\lambda^{C}-N(d_2)p_s\mathrm{e}^{-rT_1} \tag{23}$$

$$C^{\mathrm{put}}=N(-d_2)p_s\mathrm{e}^{-rT_1}-N(-d_1)\lambda^{C} \tag{24}$$

$$d_1=\frac{1}{\xi\sqrt{T_1}}\left[\ln\left(\frac{\lambda^{C}}{p_s}\right)+\left(r+\frac{\xi^2}{2}\right)T_1\right] \tag{25}$$

$$d_2=d_1-\xi\sqrt{T_1} \tag{26}$$

$$p_s=\beta\lambda^{C} \tag{27}$$

C. Lower-level 1: EM Clearing

We next establish an EM clearing model based on [41] for the lowest dispatch costs derived from (28).

$$\min\sum_{t}\sum_{i,v}\alpha_{i,v}^{G}P_{i,v,t}^{G} \tag{28}$$

s.t.

$$-\sum_{m\in\varphi_n^{\mathrm{N}}}B_{n,m}\left(\delta_{n,t}-\delta_{m,t}\right)+\sum_{i\in\varphi_n^{\mathrm{S}},v}P_{i,v,t}^{G}-P_{n,t}^{D}=0\quad\left[\lambda_{n,t}^{E}\right]\quad\forall n,t \tag{29}$$

$$P_i^{\mathrm{G,min}}\le P_{i,v,t}^{G}\le P_i^{\mathrm{G,max}}\quad\left[\varphi_{i,v,t}^{\mathrm{G,min}},\varphi_{i,v,t}^{\mathrm{G,max}}\right]\quad\forall i,v,t \tag{30}$$

$$P_{i,t}^{\min} \le \sum_{v} P_{i,v,t}^{\mathrm{G}} \le P_{i,t}^{\max} \quad \left[\varphi_{i,t}^{\min}, \varphi_{i,t}^{\max}\right] \quad \forall i, v, t \tag{31}$$

$$-RD_i \le \sum_{v} (P_{i,v,t+1}^{\mathrm{G}} - P_{i,v,t}^{\mathrm{G}}) \le RU_i \quad \left[\varepsilon_{i,t}^{\mathrm{D}}, \varepsilon_{i,t}^{\mathrm{U}}\right] \quad \forall i, \forall tT \tag{32}$$

$$B_{n,m}\left(\delta_{n,t} - \delta_{m,t}\right) \le P_{n,m}^{\mathrm{L,max}} \quad \left[\varepsilon_{n,m,t}^{\mathrm{L,max}}\right] \quad \forall n, t, m \in \varphi_n^{\mathrm{N}} \tag{33}$$

$$-\pi \le \delta_{n,t} \le \pi \quad \left[\varepsilon_{n,t}^{\delta,\min}, \varepsilon_{n,t}^{\delta,\max}\right] \quad \forall n, t \tag{34}$$

$$\delta_{1,t} = 0 \quad \left[\varepsilon_t^{\delta_1}\right] \quad \forall t \tag{35}$$

The variables in the square brackets on the right side indicate the Lagrangian multipliers from the individual sides of (29)-(35). The constraints on the energy balance and output limits for each offering block are defined by (29) and (30), respectively. Restrictions on the total output power, generator ramping capability, transmission line capacity, and voltage angle are enforced through (31)-(34), respectively, with specified upper and lower limits. Node 1 is established in (35) as the reference node for the power angle with a fixed value of 0. Importantly, the dual variables for the individual constraints are shown in the square brackets on the right side, with the locational marginal price identified as the dual variable in (29).

D. Lower-level 2: CM Clearing

We next describe a CEA exchange model that utilizes call auctions based on [9] to facilitate bidirectional trading. All GenCos must provide their CM selling and buying curves in a price-quantity format, following a stepwise method, while ensuring compliance with the designated time restriction. The model is characterized by (36)-(41).

$$\max\left(\sum_{i,k} \alpha_{i,k}^{\mathrm{CB}} Q_{i,k}^{\mathrm{B}} - \sum_{i,h} \alpha_{i,h}^{\mathrm{CS}} Q_{i,h}^{\mathrm{S}}\right) \tag{36}$$

s.t.

$$\sum_{i,k} Q_{i,k}^{\mathrm{B}} = \sum_{i,h} Q_{i,h}^{\mathrm{S}} \quad \left[\lambda^{\mathrm{C}}\right] \tag{37}$$

$$Q_h^{\mathrm{S,min}} \le Q_{i,h}^{\mathrm{S}} \le Q_h^{\mathrm{S,max}} \quad \left[\kappa_{i,h}^{\mathrm{S,min}}, \kappa_{i,h}^{\mathrm{S,max}}\right] \quad \forall i, h \tag{38}$$

$$Q_k^{\mathrm{B,min}} \le Q_{i,k}^{\mathrm{B}} \le Q_k^{\mathrm{B,max}} \quad \left[\kappa_{i,k}^{\mathrm{B,min}}, \kappa_{i,k}^{\mathrm{B,max}}\right] \quad \forall i, k \tag{39}$$

$$\sum_{h} Q_{i,h}^{S} \le Q_i^{\mathrm{ST}} \quad \left[\kappa_i^{\mathrm{ST}}\right] \quad \forall i \tag{40}$$

$$\sum_{k} Q_{i,k}^{\mathrm{B}} \le Q_i^{\mathrm{BT}} \quad \left[\kappa_i^{\mathrm{BT}}\right] \quad \forall i \tag{41}$$

The objective function of the CM-clearing model is formulated in (36) to maximize social welfare. The variables in the square brackets on the right side indicate the Lagrangian multipliers from the individual sides of (37)-(41). Equation (37) represents the CEA balance constraint. The clearing price in the CM, which establishes the marginal selling price to maximize turnover, is defined in terms of the dual variable associated with the CEA balance constraint. Constraint (38) sets the lower cap for each bidding block for CEA sellers, whereas (39) sets the upper cap for CEA buyers. Constraints (40) and (41) set limitations on the CEA capacity for sellers and buyers, respectively.

E. Model Modification for Weekly Continuous Trading

Next, the daily trading is expanded to a weekly continuous trading model by introducing a choice to transfer the holding CEA surplus or deficit. For modification, in addition to the subscript of day d, the constraints in (5), (6), (17), and (18) of the aforementioned daily trading are reformulated as:

$$C_{g,d}^{\mathrm{FB}} = \lambda_d^{\mathrm{FB}}\left(Q_{g,d}^{\mathrm{BH}} + Q_{g,d}^{\mathrm{BHT}}\right)\mathrm{e}^{-rT_1} \tag{42}$$

$$R_{g,d}^{\mathrm{FS}} = \lambda_d^{\mathrm{FS}}\left(Q_{g,d}^{\mathrm{SH}} + Q_{g,d}^{\mathrm{SHT}}\right)\mathrm{e}^{-rT_1} \tag{43}$$

$$0 \le Q_{i,d}^{\mathrm{ST}} + Q_{i,d}^{\mathrm{SH}} + Q_{i,d}^{\mathrm{SHT}} \le Mu_{i,d} \quad \forall i \tag{44}$$

$$0 \le Q_{i,d}^{\mathrm{BT}} + Q_{i,d}^{\mathrm{BH}} + Q_{i,d}^{\mathrm{BHT}} \le M\left(1 - u_{i,d}\right) \quad \forall i \tag{45}$$

When the daily separated trading is linked with transferrable $Q_{i,d}^{\mathrm{SHT}}$ or $Q_{i,d}^{\mathrm{BHT}}$, the weekly continuous trading model becomes a Markov decision process, which can be transformed into an RL problem for accelerating the simulation. The key elements of RL (i. e., state, action, reward, and state-transition functions) are defined as $state = \left(d, Q_{i,d}^{\mathrm{ST}}, Q_{i,d}^{\mathrm{SH}}, Q_{i,d}^{\mathrm{SHT}}, Q_{i,d}^{\mathrm{BT}}, Q_{i,d}^{\mathrm{BH}}, Q_{i,d}^{\mathrm{BHT}}, \beta_d, Obj_d\right)$ and $action = \left(r_d^{\min}\right)$.

The state transition between states d and $d+1$ is expressed in (46), which can be activated by the action $r_d^{\min}$ through the constraints (47) and (48). Here, a certain share $r_d^{\min}$ of the overall CEA surplus or deficit should be traded in the spot market or protected through buying pull or call carbon options.

$$Q_{i,d}^{\mathrm{net}} + Q_{i,d-1}^{\mathrm{SHT}} - Q_{i,d-1}^{\mathrm{BHT}} = Q_{i,d}^{\mathrm{ST}} + Q_{i,d}^{\mathrm{SH}} + Q_{i,d}^{\mathrm{SHT}} - Q_{i,d}^{\mathrm{BT}} - Q_{i,d}^{\mathrm{BH}} - Q_{i,d}^{\mathrm{BHT}} \tag{46}$$

$$Q_{i,d}^{\mathrm{ST}} + Q_{i,d}^{\mathrm{SH}} \ge r_d^{\min}\left(Q_{i,d}^{\mathrm{ST}} + Q_{i,d}^{\mathrm{SH}} + Q_{i,d}^{\mathrm{SHT}} + Q_{i,d-1}^{\mathrm{SHT}}\right) \tag{47}$$

$$Q_{i,d}^{\mathrm{BT}} + Q_{i,d}^{\mathrm{BH}} \ge r_d^{\min}\left(Q_{i,d}^{\mathrm{BT}} + Q_{i,d}^{\mathrm{BH}} + Q_{i,d}^{\mathrm{BHT}} + Q_{i,d-1}^{\mathrm{BHT}}\right) \tag{48}$$

The reward value of each action is equal to the objective value Obj_d of the trading state on day d. We assume that only a certain number of CEA demands Q_i^{trans} can be transferred to the next week, whereas the rest of the weekly CEA surplus or deficit of GenCo should be traded in the spot market or protected by carbon options.

IV. Reformulation and Solution Methods

A. Reformulation of Lower-level Problems

The application of the Karush-Kuhn-Tucker conditions [8] enables the initial problem to be restated as a single-level model. The complementary constraints of (36)-(41) are obtained from (S1)-(S5) in Supplementary Material A. As the EM model can be reformulated in the same manner, the detailed process is not presented here.

B. Linearization Methods

Despite the transformation of the model into a single-level one, the presence of compounded nonlinearities poses challenges that hinder the resolution process. The model exhibits two categories of nonlinearity.

1) Nonlinearities of Complementary Slackness

These nonlinearities are mainly expressed in (S1)-(S5) in the form of the product of decision and dual variables, which can be linearized via the big-M method [8]. As an example, (S3) can be reformulated as:

$$0 \le Q_{i,k}^{\mathrm{B}} - Q_k^{\mathrm{B,min}} \le Mu \tag{49}$$

$$0 \le \kappa_{i,k}^{\mathrm{B,min}} \le M(1-u) \tag{50}$$

2) Bilinear Terms from Product of Two Decision Variables

As an example, $\lambda^{\mathrm{C}} Q_g^{\mathrm{SH}}$ in (8) can be discretized and linearized using the binary expansion method [42] expressed in (51)-(56). The same reformulation technique can be used with the bilinear terms $\lambda^{\mathrm{C}} Q_g^{\mathrm{BH}}$, $\lambda^{\mathrm{C}} Q_{i,k}^{\mathrm{B}}$, and $\lambda^{\mathrm{C}} Q_{i,h}^{\mathrm{S}}$.

The carbon price λ^{C} can be ranged in $\left[\lambda_{\min}^{\mathrm{C}}, \lambda_{\max}^{\mathrm{C}}\right]$ according to the price limits set by the government or historical data. Then, when the auxiliary variable x_y^{SH} is introduced, the carbon price can be expanded as in (51) and (52). Note that the discretization resolution of λ^{C} is determined by the expansion number Y, where a higher Y can yield a more precise approximation but higher complexity.

$$\lambda^{\mathrm{C}} = \lambda_{\min}^{\mathrm{C}} + \Delta\lambda \sum_{y=0}^{Y-1} 2^y x_y^{\mathrm{SH}} \tag{51}$$

$$\Delta\lambda = \left(\lambda_{\max}^{\mathrm{C}} - \lambda_{\min}^{\mathrm{C}}\right) / 2^Y \tag{52}$$

Then, after both sides are multiplied by Q_g^{SH}, (51) can be reformulated to (53). To simplify model expression, a new variable $QX_y^{\mathrm{SH}} = Q_g^{\mathrm{SH}} x_y^{\mathrm{SH}}$ is defined and incorporated into linear equation (54).

$$\lambda^{\mathrm{C}} Q_g^{\mathrm{SH}} = \lambda_{\min}^{\mathrm{C}} Q_g^{\mathrm{SH}} + \Delta\lambda \sum_{y=0}^{Y-1} 2^y Q_g^{\mathrm{SH}} x_y^{\mathrm{SH}} \tag{53}$$

$$\lambda^{\mathrm{C}} Q_g^{\mathrm{SH}} = \lambda_{\min}^{\mathrm{C}} Q_g^{\mathrm{SH}} + \Delta\lambda \sum_{y=0}^{Y-1} 2^y \cdot QX_y^{\mathrm{SH}} \tag{54}$$

The nonlinearity in $QX_y^{\mathrm{SH}} = Q_g^{\mathrm{SH}} x_y^{\mathrm{SH}}$ can be solved using the big-M method as:

$$0 \le Q_k^{\mathrm{SH}} - QX_k^{\mathrm{SH}} \le M\left(1 - x_k^{\mathrm{SH}}\right) \tag{55}$$

$$0 \le QX_k^{\mathrm{SH}} \le M x_k^{\mathrm{SH}} \tag{56}$$

C. Twin-delayed Deep Deterministic (TD3) Policy Gradients

Based on the Markov decision described in Section III-E, multiple RL methods can be adopted for simulation acceleration. However, as a continuous variable is used here, a policy-based deep RL method is the most suitable. TD3 policy gradients are regarded as more stable successors of the traditional deep deterministic policy gradients (DDPGs) [43]. Accordingly, they are adopted here with the detailed Algorithm 1 [44].

The relevant equations include (57)-(61). Equation (57) indicates that action $r_d^{\min}$ and noise ε (following a normal distribution $\mathcal{N}$) cannot exceed the pre-defined limits. Equation (58) is the Bellman transition equation for calculating the estimated value. Equation (59) shows the minimum-seeking update for the critic network. Equation (60) shows the DPG-based update for the actor network. The soft-update technique [45] is utilized in (61) to eliminate the target value fluctuation during the estimation.

$$r_d^{\min} = clip\left(\left(\pi_\phi(s) + \varepsilon\right), r_{d,\min}^{\min}, r_{d,\max}^{\min}\right) \quad \varepsilon \sim \mathcal{N}(0, \sigma) \tag{57}$$

Algorithm 1: weekly continuous trading by TD3

Generate a replay buffer B by random generation
Initialize critic networks V_{θ_1} and V_{θ_2}, and actor network π_ϕ
Initialize target network $\theta_i' \leftarrow \theta_i$, $\phi' \leftarrow \phi$
For *episode* in $[1, N_{episodes}]$ **do**
  For d in $[1, 7]$ **do**
    The strategic GenCo selects action $r_d^{\min}$ with exploration noise ε
    Note that the action and noise should satisfy constraint (57)
    Use the minimum trading action $r_d^{\min}$ on the daily trading model
    Calculate the next state and reward by models described in Section III-E
    Store transition $(state_d, r_d^{\min}, reward_d, state_{d+1})$ in buffer B
    Sample N transitions batched from B
    Calculate the target action $r_{d+1}^{\min}$ and estimated value Y by (58)
    Update critic network by (59)
    If $t\%N_{update} = 0$ **then**
      Update ϕ by (60)
      Update critic and actor targets using soft-update method via (61)
    End if
  End iteration d
End iteration *episode*

$$Y = reward_d + \gamma \min_{i=1,2} V_{\theta_i'}\left(state_{d+1}, r_{d+1}^{\min}\right) \tag{58}$$

$$\theta_i \leftarrow \arg\min_{\theta_i} \mathcal{N}^{-1} \sum \left(Y - V_{\theta_i}\left(state_d, r_d^{\min}\right)\right)^2 \tag{59}$$

$$\nabla_\phi L(\phi) = \mathcal{N}^{-1} \sum \nabla_r V_{\theta_1}\left(state_d, r_d^{\min}\right)\Big|_{r_d^{\min} = \pi_\phi(state_d)} \nabla_\phi \pi_\phi\left(state_d\right) \tag{60}$$

$$\begin{cases} \theta_i' = \rho\theta_i + (1-\rho)\theta_i' \\ \phi' = \rho\phi_i + (1-\rho)\phi' \end{cases} \tag{61}$$

V. Case Studies

Two case studies are conducted utilizing the IEEE 30-bus system: case-buyer and case-seller. Each case examines two scenarios: ① the "reference (REF)" scenario, in which the GenCo focuses solely on the EM strategy; and ② the "double-cross" (DC) scenario, in which the GenCo determines its bidding strategy while considering cross-time and cross-market profits. The model is solved using Gurobi 9.5.2 [46] and Python 3.9 [47] running on a computer equipped with a 3.9 GHz CPU and 16 GB RAM.

A. Basic Data

In the power system, 20 generators are used with an aggregate installed capacity of 8000 MW and comprising two wind farms (900 MW), two solar plants (600 MW), and 16 conventional units (6500 MW). The generation costs for the fuel-fired generators are randomly assigned in the range of 20-60 \$/MWh. For renewable energy, the generation costs are fixed in the range of 3.2-6.4 \$/MWh [9]. The rules under the Chinese CM [48] establish the CEI benchmarks for various generators, whereas the CM bidding price is constrained in the range of 21-45 \$/t [49]. The time period is defined as six units, and line congestion is omitted from the model for simplification. The EM and CM feature-bidding blocks are each set to be five. The strategic GenCo operates five conventional generators (three gas units of 240 MW capacity each, and two coal units of 600 and 330 MW capacities, respectively). Table I lists the actual CEI values of these generators in various scenarios.

TABLE I
ACTUAL CEI VALUES OF GENERATORS IN VARIOUS SCENARIOS

Capacity (MW)	Benchmark (g/kWh)	Case-buyer (g/kWh)	Case-seller (g/kWh)
600	850	918	784
330	898	970	826
240	398	430	366

In the B-S model, the CM price volatility rate ξ is set to be 0.3 according to the daily carbon prices in the European emission trading system [49]. The continuously compounded risk-free rate r in the B-S model is 5.83% (equivalent to an annual interest rate of 6%). The best strike price for the carbon options trading of GenCo can be achieved by scanning the predetermined strike-to-spot ratio β under different conditions.

B. Case-buyer Daily Trading: GenCo with Higher CEI

1) Cross-market Trading Analysis

The cross-market trading is exerted by adjusting the bidding prices of GenCo in both the EM and CM, and the equilibrium in the two markets is analyzed in this section. Since cross-market trading analysis may affect the cross-time decisions of GenCo, the expectation and variance of the future CEA price are fixed at 34 \$/t and 3 \$/t, respectively. The expiration time of the carbon option is set to be 200 days, and the strike-to-spot price ratio is set to be 1.1.

Figure 3(a) and (b) illustrates the power supply curves in different scenarios. In the REF and DC scenarios, the gross generation capacity is separated into two parts with differing functions in the REF and DC scenarios. To ensure a minimum output level, the first segment bids at lower prices.

The strategic GenCo manipulates the second half to raise EM clearing prices. For comparison, Supplementary Material A Fig. SA1 shows the system marginal prices (SMPs) in different scenarios, revealing that the SMPs in the DC scenario are occasionally higher than those in the REF scenario. Thus, after the CM is accounted for, the bidding strategy of GenCo aims to increase the EM price by coordinating containing generators and exercising market power, which passes on the CEA buying costs to customers.

Figure 3(c) and (d) shows the cumulative CEA buying curves in different scenarios. The bidding strategy of GenCo in the CM is conducted without plans in the REF scenario, and the uniform clearing price is 35 \$/t. In the DC scenario, part of the demand (66.3 t) is balanced in the CM market at 34.4 \$/t, whereas the rest is held and traded in the future. Fewer CEAs are currently required and the GenCo tries to lower CM prices by exerting market power; therefore, the bidding strategy of GenCo decreases the CM price by 0.6 \$/t.

2) Cross-time Trading Analysis

Supported by the carbon options and B-S model, the GenCo may choose to trade or hold by assessing costs and earnings from the holding action. Different trading scenarios are tested by adjusting β, and the ratio with the highest total profit may be used for call option trading.

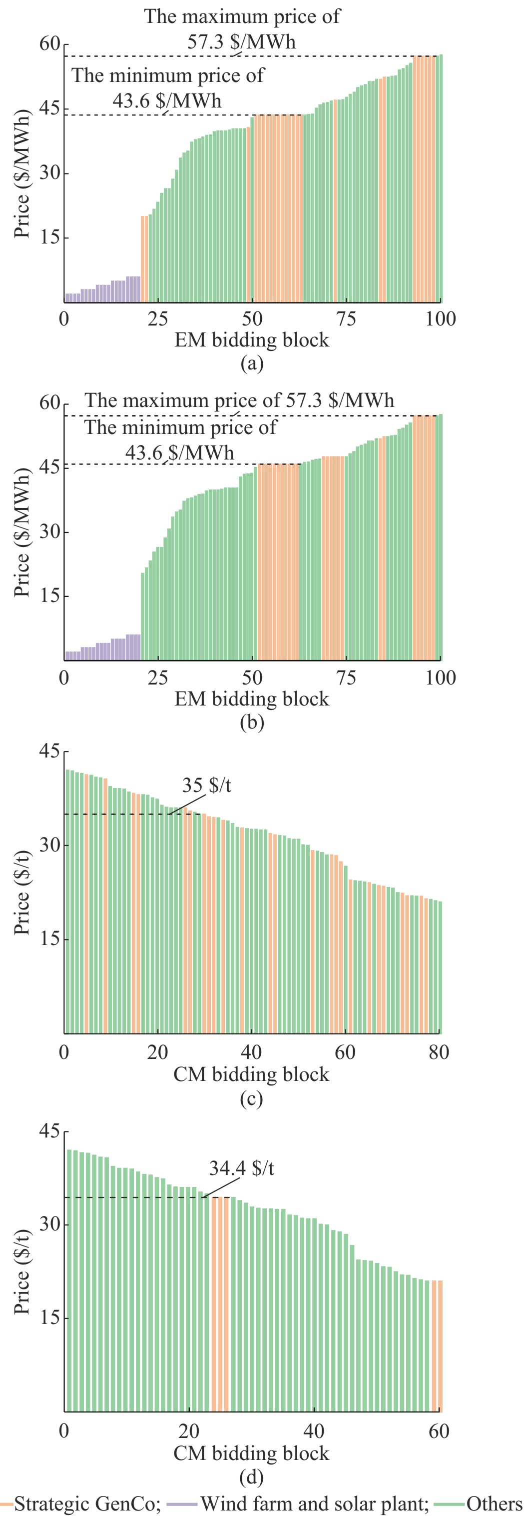

Fig. 3. Multi-market bidding curves in case-buyer. (a) Accumulated power supply curves in EM in REF scenario. (b) Accumulated power supply curves in EM in DC scenario. (c) Accumulated CEA buying curves in CM in REF scenario. (d) Accumulated CEA buying curves in CM in DC scenario.

To hedge against holding risks, the option strike price should be set in the opposite direction of the current stock price (i.e., if the expected future price is less than the present price, the option strike price should be set higher). Therefore, β is set to be larger than 1.

Figure 4 shows both the CEA holding amount and CM prices in case-buyer under different price ratios. The first aspect to note is the increase in holding amounts from 114.9 to 282.2 t, followed by a precipitous drop to zero. When this ratio increases from 1 to 1.2, the option price decreases, making the holding more appealing. However, as the final strike price may be too close to or even surpass the upper limit of the expected range, the depreciation process gradually invalidates the call option, resulting in a dramatic decrease in holding amounts.

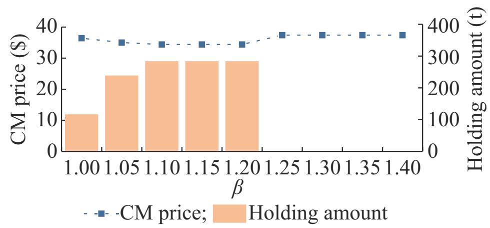

Fig. 4. CEA holding amount and CM prices under different price ratios in case-buyer.

The CM price fluctuates more slowly and is inversely related to holding levels, as the holding actions reduce CEA demand at current and lower CM prices. The effects of the option parameters on the overall earnings of GenCo are shown in Figs. 5-7.

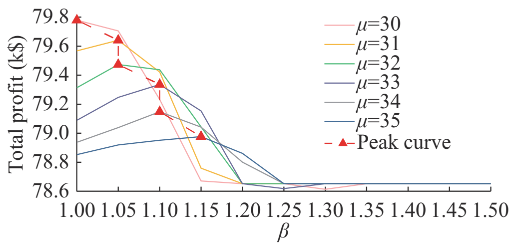

Fig. 5. Total profits under different price ratios and expected future prices μ.

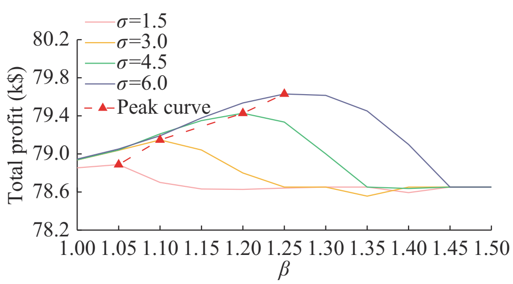

Fig. 6. Total profits under different price ratios and future price variances σ.

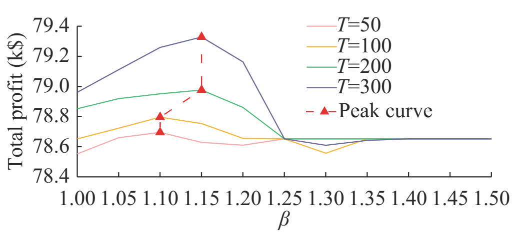

Fig. 7. Total profits under different price ratios and expiration time T.

The key motivation for holding actions is the prediction of a lower future price, as represented by the expected CEA price. Figure 5 shows that the total profit increases following a decrease in the expected CEA price. However, the uncertainties derived from future price variance and the expiration time of options bring risks and in turn a desire for carbon options.

A comparison of the results presented in Figs. 5-7 under the same expectation value reveal that the carbon option can provide larger profits in a more uncertain environment (higher variance and longer expiration time). The findings reveal that the expected value of the future price is critical in determining the best return and that buying call options can effectively hedge the risks of time and variation.

C. Daily Trading in Case-seller: GenCo with Lower CEI

1) Cross-market Trading Analysis

In this case, the actual CEI of GenCo falls below the established benchmarks. The future CEA price is predetermined at 37 \$/t with a variance of 3 \$/t. The carbon option expires in 200 days with a strike price-to-stock price ratio of 0.9.

2) Cross-time Trading Analysis

To mitigate risk, the option strike price should be set lower than the current price when the expected future price exceeds it. Accordingly, the price ratio is adjusted to less than 1.

D. Weekly Trading Analysis

By selecting the forced transaction ratio $r_d^{\min}$ of the CM as the action, we can conduct weekly continuous trading cases. To simulate cross-week CEA holding behavior, the initial CEA amount transferred from the previous week for both the buyer and seller is set to be 500 t, which is also the remaining limit of CEA trading demands at the end of the targeted week. According to the aforementioned daily power load curve, the weekly load curve is designed using the normalized expansion ratio, as shown in Fig. 8.

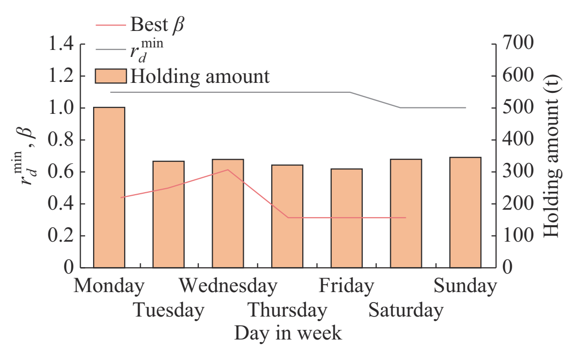

Fig. 8. Weekly CEA holding amount, action $r_d^{\min}$, and best β in case-buyer.

As Fig. 8 illustrates, CEA buyers prefer to buy a substantial amount of CEA at the beginning of the trading cycle to hedge against the invalidation of call options over time. From the buyer's perspective, it is an economically sound strategy to minimize $r_d^{\min}$ to reduce the CEA holding amount at an early stage under a weekly timescale. This resonates with the conclusions drawn from Figs. 4-7, where the total buying costs are reduced from \$0.408 million to \$0.377 million, a decrease of 7.60%.

1) Case-buyer

In the TD3 training optimization process, the buyer sets the action value of $r_d^{\min}$ as a continuous variable between 0 and 1. This allows the buyer to explore different action states within a single day to obtain the β value that maximizes economic benefits in the DC scenario for that state. Through this process, the buyer can determine the optimal trading strategy, which includes both the option trading volume and the remaining CEA transferred to the next trading day.

2) Case-seller

The ratio β for the seller is set to be less than 1. During the TD3 training optimization process, the seller also sets the action value of $r_d^{\min}$ as a continuous variable from 0 to 1.

As given in Supplementary Material A, compared with case-buyer, the initial actions of seller are more conservative, i.e., values of $r_d^{\min}$ are generally greater than 0.85, indicating a preference for holding a larger portion of the CEA demand for additional days. Although additional expenditures are incurred in buying put options, the overall profits of GenCo in the CM are in the range of \$8.607-8.799 million (an increase of 2.23%), which is consistent with the conclusions drawn from Supplementary Material A.

E. Market Power Analysis

A market power analysis is next conducted. As the widely adopted Herfindahl-Hirschman (H-H) index [50] is unsuitable for measuring the market power of a firm (where the H-H index instead measures the concentration of the entire market), the Lerner index (LI) [51] is used in this paper to analyze the market power of strategic GenCo. The LI has a value between 0 and 1 and can be calculated as the percentage markup of price P above the marginal cost MC:

$$LI = (P - MC)/P \tag{61}$$

The LI values in the REF or DC scenarios and the EM and CM are calculated and compared. Figure 9 shows the LI indices in EM in different cases. The CM power values under different conditions are listed in Supplementary Material A.

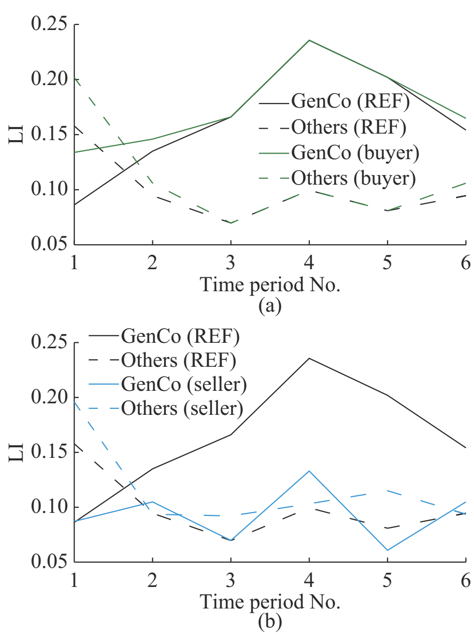

Fig. 9. LI indices in EM in different cases. (a) Case-buyer. (b) Case-seller.

In the case-buyer, as the GenCo exercises EM power, its LI value surpasses those of normal generators. Specifically, the LI value in EM in the DC scenario exceeds that in the REF scenario. Regarding carbon trading, the strategic GenCo observes a significant increase in its carbon LI value, increasing from 2.9% (REF) to 14.4% (DC).

By contrast, the findings diverge for CEA sellers. In the DC scenario, where less EM power is utilized, the strategic GenCo exhibits an LI value that is less than that of normal generators. Nevertheless, the GenCo still maintains a competitive edge in CM, with an LI value of 8.6% which surpasses that of normal generators by 7.3%.

The EM power exertion of strategic GenCo depends on its CM trading position. In the case-buyer, EM power is utilized to pass its carbon costs onto consumers, whereas the strategic GenCo in the case-seller prefers to reduce its EM bidding prices with the aim of accumulating more CEAs for sale. The EM power exertion is not sustained throughout the entire trading period and is more likely to be employed during peak load periods. Regardless of the case under CM power, cross-time and cross-market trading consistently leverages market power in the CM.

F. Calculation Efficiency and Convergence

Binary expansion is an approximation method, so its accuracy and solution speed are strongly related to the expansion coefficient Q. A higher value indicates a more precise solution and a longer solution time. Figure 10 shows the results of a convergence analysis conducted to select the best expansion value. The results show that the models under the case-seller and case-buyer converge with coefficients of 8 and 6, respectively. However, to maintain a certain margin, Q is set to be 10. Supplementary Material A shows the iterative processes and convergence curves of the TD3 technique, which validates its effectiveness in finding the optimal action.

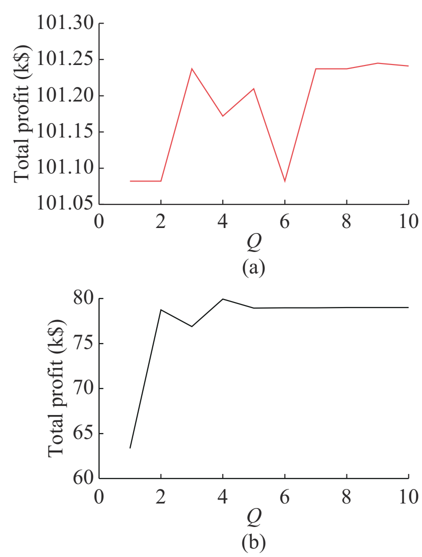

Fig. 10. Total profits under different expansion coefficients. (a) Case-seller. (b) Case-buyer.

VI. Conclusion

This paper proposes a multi-market trading strategy that incorporates an option-jointed daily strategy and a reinforcement learning-jointed weekly strategy for cross-market and cross-time trading of generators in EMs and CMs. Case studies on an IEEE 30-bus system are reconducted, and the following three conclusions are drawn.

1) When the cross-market bidding curves are coordinated and a hold-and-see cross-time trading in CM is adopted, the portions of CEA demand and supply are traded in the future. Although an additional amount is paid in the carbon option to hedge against risks from price uncertainty, the strategic GenCo can still boost its overall profits by approximately 3%.

2) The expected future price is crucial for evaluating the best return from buying a carbon option, which can be more valuable in an uncertain market with longer holding periods and larger prediction variance. With an increase of $1.5 in price forecasting variance, the integrated option strategy is projected to boost the annual profits of strategic GenCo by $73000-182500.

3) The utilization of EM power is contingent on the position of strategic GenCo in CM trading. Typically, this exertion is not sustained across the complete trading duration but is more aptly utilized during peak load intervals. By contrast, the proposed trading strategy ensures consistent leveraging of power in the CM.

Based on the proposed trading strategy, future work will develop a model that exhibits more intelligent trading behavior to simulate market equilibrium with numerous agents. However, the bidding information of other generators is difficult to access; thus, a data-driven method capable of extracting valuable EM-CM data from the limited market disclosure data deserves an in-depth investigation.

References

[1] T. Levin, J. Kwon, and A. Botterud, "The long-term impacts of carbon and variable renewable energy policies on electricity markets," *Energy Policy*, vol. 131, pp. 53-71, Aug. 2019.

[2] L. Tang, J. Shi, and Q. Bao, "Designing an emissions trading scheme for China with a dynamic computable general equilibrium model," *Energy Policy*, vol. 97, pp. 507-520, Jul. 2016.

[3] B. Mizrach, "Integration of the global carbon markets," *Energy Economics*, vol. 34, no. 1, pp. 335-349, Jan. 2012.

[4] B. R. Huber, "How did RGGI do it? Political economy and emissions auctions," *Ecology Law Quarterly*, vol. 40, no. 1, pp. 59-106, Mar. 2013.

[5] Y. J. Zhang and Y. M. Wei, "An overview of current research on EU ETS: evidence from its operating mechanism and economic effect," *Applied Energy*, vol. 87, no. 6, pp. 1804-1814, Jun. 2010.

[6] S. Colenbrander, Y. Cao, and R. Nadin. (2021, Mar.). Five expert views on China's pledge to become carbon neutral by 2060. [Online]. Available: https://odi. org/en/insights/five-expert-views-on-chinas-pledge-to-become-carbon-neutral-by-2060/

[7] Quebec Government. (2022, Aug.). Calculation method used to determine the quantity of emission units allocated free of charge. [Online]. Available: https://www.environnement.gouv.qc.ca/changements/carbone/methode-calcul-en.htm

[8] K. Jiang, P. Wang, J. Wang *et al.*, "Reserve cost allocation mechanism in renewable portfolio standard-constrained spot market," *IEEE Transactions on Sustainable Energy*, vol. 13, no. 1, pp. 56-66, Aug. 2022.

[9] K. Jiang, N. Liu, X. Yan *et al.*, "Modeling strategic behaviors for GenCo with joint consideration on electricity and carbon markets," *IEEE Transactions on Power Systems*, vol. 38, no. 5, pp. 4724-4738, Oct. 2023.

[10] Y. Xia, Q. Xu, S. Tao *et al.*, "Cutting the cake of the emission cap in the peer-to-peer multi-energy sharing market," *CSEE Journal of Power and Energy Systems*, vol. 9, no. 2, pp. 470-482, May 2023.

[11] Z. Lu, L. Bai, J. Wang *et al.*, "Peer-to-peer joint electricity and carbon trading based on carbon-aware distribution locational marginal pricing," *IEEE Transactions on Power Systems*, vol. 38, no. 1, pp. 835-852, Apr. 2023.

[12] X. Liu, "Research on bidding strategy of virtual power plant considering carbon-electricity integrated market mechanism," *International Journal of Electrical Power & Energy Systems*, vol. 137, p. 107891, May 2022.

[13] M. Daneshvar, B. Mohammadi-Ivatloo, K. Zare *et al.*, "A novel operational model for interconnected microgrids participation in transactive energy market: a hybrid IGDT/stochastic approach," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 6, pp. 4025-4035, Jul. 2021.

[14] M. Yan, M. Shahidehpour, A. Alabdulwahab *et al.*, "Blockchain for transacting energy and carbon allowance in networked microgrids," *IEEE Transactions on Smart Grid*, vol. 12, no. 6, pp. 4702-4714, Aug. 2021.

[15] T. Wan, Y. Tao, J. Qiu *et al.*, "Distributed energy and carbon emission right trading in local energy systems considering the emission obligation on demand side," *IEEE Systems Journal*, vol. 17, no. 4, pp. 6292-6301, Aug. 2023.

[16] Z. Yan, C. Li, Y. Yao *et al.*, "Bi-level carbon trading model on demand side for integrated electricity-gas system," *IEEE Transactions on Smart Grid*, vol. 14, no. 4, pp. 2681-2696, Dec. 2023.

[17] D. Xiao, M. K. AlAshery, and W. Qiao, "Optimal price-maker trading strategy of wind power producer using virtual bidding," *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 3, pp. 766-778, Jun. 2022.

[18] D. Xiao, H. Chen, C. Wei *et al.*, "Statistical measure for risk-seeking stochastic wind power offering strategies in electricity markets," *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 5, pp. 1437-1442, Dec. 2022.

[19] A. Coelho, J. Iria, and F. Soares, "Network-secure bidding optimization of aggregators of multi-energy systems in electricity, gas, and carbon markets," *Applied Energy*, vol. 301, p. 117460, Nov. 2021.

[20] C. Kang, T. Zhou, Q. Chen *et al.*, "Carbon emission flow from generation to demand: a network-based model," *IEEE Transactions on Smart Grid*, vol. 6, no. 5, pp. 2386-2394, Jan. 2015.

[21] J. Li, S. Ge, H. Liu *et al.*, "An electricity and carbon trading mechanism integrated with TSO-DSO-prosumer coordination," *Applied Energy*, vol. 356, p. 122328, Feb. 2024.

[22] A. K. Xenophon and D. J. Hill, "Emissions reduction and wholesale electricity price targeting using an output-based mechanism," *Applied Energy*, vol. 242, pp. 1050-1063, May 2019.

[23] G. A. V. Sánchez, J. M. Alzate, A. I. Cadena *et al.*, "Setting up standard power options to hedge price-quantity risk in a competitive electricity market: the colombian case," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1493-1500, Nov. 2011.

[24] J. Zhao, T. Zheng, and E. Litvinov, "A multi-period market design for markets with intertemporal constraints," *IEEE Transactions on Power Systems*, vol. 35, no. 4, pp. 3015-3025, Jul. 2020.

[25] ICE. (2022, Dec.). Products – futures & options – US environmental. [Online]. Available: https://www. theice. com/products/Futures-Options/US-Environmental

[26] Y. Liu, L. Tian, H. Sun *et al.*, "Option pricing of carbon asset and its application in digital decision-making of carbon asset," *Applied Energy*, vol. 310, pp. 118375-118375, Mar. 2022.

[27] S. Viteva, Y. V. Veld-Merkoulova, and K. Campbell, "The forecasting accuracy of implied volatility from ECX carbon options," *Energy Economics*, vol. 45, pp. 475-484, Sept. 2014.

[28] H. R. Sheybani and M. O. Buygi, "Equilibrium-based Black-Scholes option pricing in electricity markets," *IEEE Systems Journal*, vol. 16, no. 4, pp. 5413-5423, Dec. 2021.

[29] X. Yan, C. Gu, H. Zhang *et al*., "Waiting cost based long-run network investment decision-making under uncertainty," *IEEE Transactions on Power Systems*, vol. 36, no. 4, pp. 3340-3348, Dec. 2021.

[30] Z. Liu and S. Huang, "Carbon option price forecasting based on modified fractional Brownian motion optimized by GARCH model in carbon emission trading," *The North American Journal of Economics and Finance*, vol. 55, pp. 101307-101307, Jan. 2021.

[31] K. Zhou and Y. Li, "Carbon finance and carbon market in China: progress and challenges," *Journal of Cleaner Production*, vol. 214, pp. 536-549, Mar. 2019.

[32] S. Prabakaran, I. C. Garcia, and J. U. Mora, "A temperature stochastic model for option pricing and its impacts on the electricity market," *Economic Analysis and Policy*, vol. 68, pp. 58-77, Dec. 2020.

[33] F. Black and M. Scholes, "The pricing of options and corporate liabili-

ties," *The Journal of Political Economy*, vol. 81, no. 3, pp. 637-654, May 1973.

[34] North China Energy Regulatory Bureau. (2020, Dec.). Implementation rules for trading of frequency regulation and reserve in power market of West Inner Mongolia. [Online]. Available: http://hbj.nea.gov.cn/adminContent/initViewContent. do? pk=0000000072a334090172c6aba49-7002c

[35] I. Kalaitzoglou and B. M. Ibrahim, "Trading patterns in the European carbon market: the role of trading intensity and OTC transactions," *Quarterly Review of Economics and Finance*, vol. 53, no. 4, pp. 402-416, Nov. 2013.

[36] EEX. (2022, Dec.). Emissions products overview in European energy exchange. [Online]. Available: https://www. eex. com/en/markets/environmental-markets/eu-ets-spot-futures-options

[37] KRX. (2022, Dec.). Emission trading markets buy/sell. [Online]. Available: https://global. krx. co. kr/contents/GLB/06/0606/0606030101/GLB0606030101.jsp#0f86e223afe34c23526e9f69ac7c5e99=3

[38] G. Ibikunle, "Opening and closing price efficiency: do financial markets need the call auction?" *Journal of International Financial Markets*, *Institutions and Money*, vol. 34, pp. 208-227, Jan. 2015.

[39] H. Guo, Q. Chen, Q. Xia *et al.*, "Modeling strategic behaviors of renewable energy with joint consideration on energy and tradable green certificate markets," *IEEE Transactions on Power Systems*, vol. 35, no. 3, pp. 1898-1910, May 2019.

[40] M. H. Shamsi, U. Ali, E. Mangina *et al*., "A framework for uncertainty quantification in building heat demand simulations using reduced-order grey-box energy models," *Applied Energy*, vol. 275, no. April, pp. 115141-115141, Oct. 2020.

[41] C. Wang, W. Wei, J. Wang *et al*., "Strategic offering and equilibrium in coupled gas and electricity markets," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 290-306, Jan. 2018.

[42] M. V. Pereira, S. Granville, M. H. C. Fampa *et al*., "Strategic bidding under uncertainty: a binary expansion approach," *IEEE Transactions on Power Systems*, vol. 20, no. 1, pp. 180-188, Feb. 2005.

[43] Y. Liu, T. Qie, Y. Yu *et al.*, "A Novel integral reinforcement learning-based control method assisted by twin delayed deep deterministic policy gradient for solid oxide fuel cell in DC microgrid," *IEEE Transactions on Sustainable Energy*, vol. 14, no. 1, pp. 688-703, Jan. 2023.

[44] K. Jiang, K. Wang, C. Wu *et al.*, "Trajectory simulation and optimization for interactive electricity-carbon system evolution," *Applied Energy*, vol. 360, p. 122808, Apr. 2024.

[45] S. Fujimoto, H. Hoof, and D. Meger. (2018, Feb.). Addressing function approximation error in actor-critic methods. [Online]. Available: http://arxiv.org/pdf/1802.09477

[46] Gurobi. (2022, Jan.). Gurobi Optimizer Reference Manual. [Online]. Available: https://www.gurobi.com/ documentation/9.0/refman/index.html

[47] G. V. Rossum and L. Fred, *Python 3 Reference Manual*. Charleston: Createspace Independent Publishing, 2008.

[48] Guangdong Provincial Department of Ecology. (2020, Dec.). Circular of Guangdong provincial department of ecology and environment on printing and distributing the 2020 annual carbon emission quota allocation implementation plan of Guangdong province. [Online]. Available: https://gdee.gd.gov.cn/attachment/0/406/406501/3143292.pdf

[49] EMBER. (2022, Nov.). Daily carbon prices – intercontinental exchange. [Online]. Available: https://ember-climate.org/data/carbon-price-viewer/

[50] S. A. Rhoades. (1993, Feb.). The Herfindahl-Hirschman index. [Online]. Available: https://heinonline. org/HOL/P? h=hein. journals/fedred79&i=376.189

[51] J. A. Aguado, S. Martin, C. A. Pérez-Molina *et al*., "Market power mitigation in transmission expansion planning problems," *IEEE Transactions on Energy Markets*, *Policy and Regulation*, vol. 1, no. 2, pp. 73-84, Jun. 2023.

Kai Jiang received the Ph.D. degree in electrical engineering from North China Electric Power University, Beijing, China, in 2024. From 2022 to 2023, he was a Visiting Research Fellow with The University of New South Wales, Sydney, Australia. He is currently a Lecturer with School of Electrical and Electronic Engineering, North China Electric Power University. His research interests include smart grid operation and planning, extreme weather influence, and carbon and electricity markets.

Kunyu Wang received the B.S. degree in electrical engineering from North China Electric Power University, Beijing, China, in 2024, where he is currently working toward the Ph.D. degree. His research interests include power system scheduling and planning.

Lin Yang received the B.S. degree in new energy science and engineering from Shanghai University of Electric Power, Shanghai, China, in 2023. She is currently working toward the M.S. degree in School of Electrical and Electronic Engineering, North China Electric Power University, Beijing, China. Her research interests include power system scheduling and planning.

Nian Liu received the B.S. and M.S. degrees in electric engineering from Xiangtan University, Xiangtan, China, in 2003 and 2006, respectively, and the Ph.D. degree in electrical engineering from North China Electric Power University, Beijing, China, in 2009. He is currently a Professor and the Vice Dean with School of Electrical and Electronic Engineering, North China Electric Power University. His research interests include multi-energy system integration and microgrid.