



OPEN

# Enhancing renewable energy certificate transactions through reinforcement learning and smart contracts integration

Qingsu He<sup>1,4</sup>✉, Jingsong Wang<sup>2</sup>✉, Ruijie Shi<sup>1</sup>, Yifan He<sup>3</sup>✉ & Muqing Wu<sup>4</sup>

Given the complexity of issuing, verifying, and trading green power certificates in China, along with the challenges posed by policy changes, ensuring that China's green certificate market trading system receives proper mechanisms and technical support is crucial. This study presents a green power certificate trading (GC-TS) architecture based on an equilibrium strategy, which enhances the quoting efficiency and multi-party collaboration capability of green certificate trading by introducing Q-learning, smart contracts, and effectively integrating a multi-agent trading Nash strategy. Firstly, we integrate green certificate trading with electricity and carbon asset trading, constructing pricing strategies for the green certificate, carbon, and electricity trading markets; secondly, we design a certificate-electricity-carbon efficiency model based on ensuring the consistency of green certificates, green electricity, and carbon markets; then, to achieve diversified green certificate trading, we establish a multi-agent reinforcement learning game equilibrium model. Additionally, we propose an integrated Nash Q-learning offer with a smart contract dynamic trading joint clearing mechanism. Experiments show that trading prices have increased by 20%, and the transaction success rate by 30 times, with an analysis of trading performance from groups of 3, 5, 7, and 9 trading agents exhibiting high consistency and redundancy. Compared with models integrating smart contracts, it possesses a higher convergence efficiency of trading quotes.

**Keywords** Green certificate, Nash Q-learning, Smartcontract, Renewable energy, Carbon asset

Amid rising global attention to climate change, the development and use of renewable energy has become crucial for reducing greenhouse gas emissions and achieving sustainable development goals. In China, the world's largest energy consumer and carbon emitter, developing zero-carbon energy and promoting renewable energy consumption through green certificates (GCs) are essential for meeting increasing energy needs and striving to achieve carbon neutrality by 2060. Consequently, China has introduced measures such as the Renewable Energy Portfolio Standards (RPS)<sup>1,2</sup> and Tradable Green Electricity Certificates (GEC) systems to encourage a shift towards sustainable energy production and consumption through market incentives.

The experiences of the United States, the United Kingdom, Italy, Norway, and Sweden provide valuable lessons for China, demonstrating the effectiveness of market mechanisms in promoting renewable energy development.

Due to varying policies and green certificate initiation times across countries, international green certificate systems also face limitations, especially in terms of policy impact and market adaptability, such as price volatility and regulatory challenges that could undermine investor confidence and the economic viability of renewable projects. To enhance the efficiency and adaptability of the green certificate market, innovative methods are being explored, including the use of blockchain technology and the development of green financial products, as well as international cooperation.

Moreover, China's green certificate market still confronts numerous challenges<sup>3</sup>, including the complexities of issuance, verification, and trading, and the impact of policy changes. The current lack of a market-based trading system hampers the interconnectivity of electricity trading, carbon emission rights trading, and the

<sup>1</sup>State Grid Digital Technology Holdings Co., Ltd. (State Grid Xiongan Financial Technology Group Co., Ltd.), Beijing 100010, China. <sup>2</sup>Case Western Reserve University Cleveland, Cleveland 44106, USA. <sup>3</sup>University of California, Santa Cruz 95064, USA. <sup>4</sup>College of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China. ✉email: heqingsu@bupt.edu.cn; hqs1909@gmail.com; jxw741@case.edu; yhe115@ucsc.edu; heyifan\_2001@outlook.com

realization of the environmental attributes of green certificates. The absence of an effective pricing mechanism and sufficient market incentives has led to low transaction volumes, highlighting the urgent need to improve the green certificate trading framework, increase transaction efficiency, and transparency to promote renewable energy development<sup>4</sup>.

In light of these issues, this study proposes a zero-carbon energy green certificate trading system (GC-TS) architecture based on a balanced strategy, integrating Q-learning, smart contracts, and multi-agent Nash strategies. This aims to address current market problems, enhance transaction efficiency, and foster collaboration, making significant contributions to sustainable development goals in China and globally. Through this comprehensive strategy, the goal is to push the green certificate market towards higher efficiency, transparency, and intelligence, ensuring the harmonious coexistence of economic growth and environmental sustainability. The main contributions are:

1. To establish a multi-intelligence reinforcement learning game equilibrium model with the goal of diversified GC trading;
2. Introducing smart contracts to optimize the game equilibrium efficiency of multi-intelligents participating in trading, and establishing a trading system framework.

### Literature review

The global transition to sustainable energy sources necessitates the development of mechanisms like green certificates (GCs) to incentivize renewable energy production. Scholars from China, Europe, America, and other regions have extensively researched and explored issues related to the market mechanisms and models of GCs<sup>5–19</sup>, technological innovations including blockchain and artificial intelligence platform technologies<sup>20–37</sup>, policies and economic strategies and market changes<sup>10,19,38–45</sup>.

Researchers worldwide, especially in China, Europe, and the United States, have thoroughly explored the rules of the green certificate market<sup>5–8</sup>, circulation methods<sup>9</sup>, price simulations<sup>10</sup>, and market models<sup>11,12</sup>. Through trading games<sup>13–15</sup>, equilibrium models<sup>16</sup>, the integration of electricity markets with green certificate markets<sup>17–19</sup>, the coupling of green certificates with carbon emissions<sup>46,47</sup>, market combinations<sup>48</sup>, the impact of carbon pricing<sup>49</sup>, and the effect of electricity load<sup>50</sup>, scholars aim to understand and optimize the dynamics and efficiency of the GC market.

With the evolution and development of information processing technology, utilizing blockchain and artificial intelligence in the pricing system of GCs, trading strategies, and optimizing transaction efficiency has become an important means in the design of new generation trading technology frameworks. The application of blockchain technology<sup>20–23,51</sup>, smart contracts<sup>25,26</sup>, transaction information processing<sup>27–30</sup>, and platform design<sup>31,32</sup> has offered new possibilities for GC trading. These technological innovations not only enhance the transparency and efficiency of transactions but also promote the security and reliability of the market.

Research also involves the policies and economic strategies of GCs<sup>38,46</sup>, especially analyzing how renewable energy tariff surcharge subsidies, subsidy settlement cycles, delay cycles, and consumer preferences affect the sales price of GCs. These strategies aim to improve the market price advantage of GCs<sup>52</sup>, thereby promoting the consumption of renewable energy.

In the GC market, the strength of the herding effect determines the average compliance cost level, and the strategic behavior of market participants mainly affects the stability of the GC price, with greater price volatility leading to slower convergence to the equilibrium price<sup>39</sup>. To address the regional imbalance of energy supply and demand in China, literature<sup>53</sup> has established a cross-regional GC futures model. Additionally, research has used game models<sup>40–44</sup>, self-conclusive and variational particle swarm optimization algorithms<sup>45</sup>, and other methods to assess the synergy and incentives of electricity<sup>54</sup>, GPMs, and the GC market on the decision-making behaviors of non-renewable energy generation companies<sup>55,56</sup>, GPMs, and electricity purchasers.

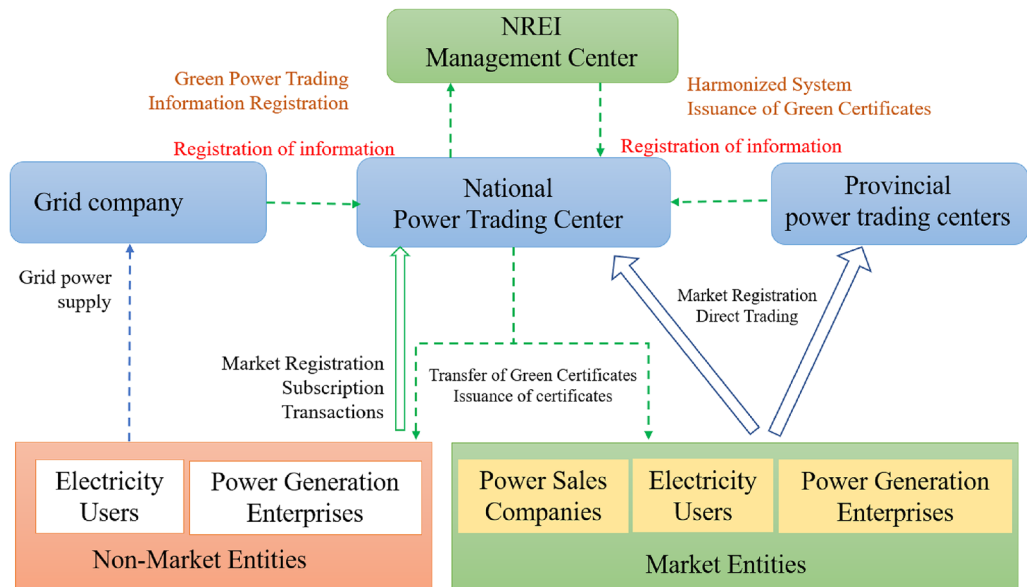
The application of reinforcement learning and Q-learning in financial market forecasting<sup>33</sup>, learning trading rules for specific financial assets<sup>34</sup>, and improving financial trading decisions<sup>35</sup> offers a new perspective for GC trading strategies. Particularly, deep Q-learning in the algorithmic trading system for the commodity futures market<sup>36</sup> and the design of a supply chain carbon allowance allocation auction based on multi-agent modeling and Q-learning<sup>37,57,58</sup> demonstrate the potential of AI technology in energy management and GC trading.

Despite the challenges posed by the complexity of issuing, verifying, and trading GCs, as well as policy changes, the development of the GC market can be effectively enhanced by integrating technological innovations and advanced trading strategies, such as reinforcement learning and smart contracts. Future research needs to focus on further integrating these technologies and strategies, as well as addressing challenges specific to the Chinese market, to contribute to the healthy development of the global GC market.

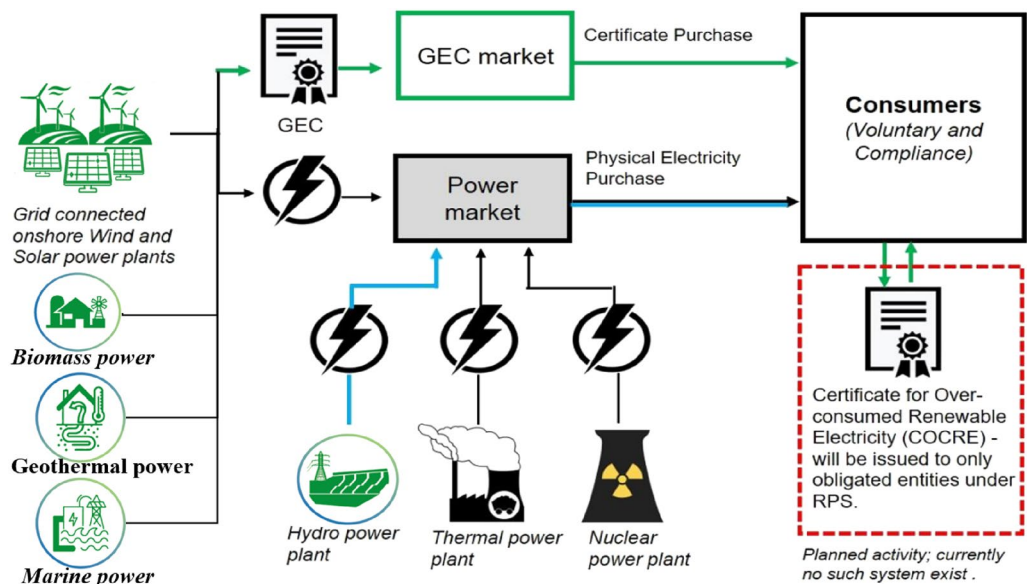
### Green certificate framework: Methods and lifecycle

The trading model for GCs primarily comprises two components: GCs bundled with green power (as Fig. 1) and standalone GCs (as Fig. 2).

The trading of Green Certificates (GCs) alongside green power represents a dual model that integrates both “certificates and power.” This model enhances the traceability of green power throughout its entire lifecycle and aligns the value of green power with environmental benefits. In the process of green power transaction settlements, the power trading center allocates GCs based on various data points that have been mutually agreed upon by the involved parties. The established framework, as depicted in Fig. 1, ensures a harmonized flow of green power and the corresponding GCs, thereby simplifying trading processes and strengthening market participation.



**Figure 1.** Framework for green certificates to accompany green power trading in China.



**Figure 2.** Schematic diagram of the green certificate and green electricity separation trading method.

Independent trading certificates allow renewable energy producers to trade on a voluntary subscription platform, where buyers select GCs through the interface and complete their subscription via negotiation, listing, or auction. Although this method overcomes geographical limitations of electricity transmission, it faces challenges: the platform’s inability to integrate electricity market data makes it difficult to bridge the gap between “certificates and electricity,” and there’s a lack of a quantified relationship between GCs and carbon credits for achieving net-zero carbon. NREI will issue a certificate (One Ps./MWh) based on a certain proportion of green power traded, forming a tradable GC. Generally speaking, grid companies, power sales companies, power users, and carbon emission enterprises (such as fossil energy power plants) constitute the main body of the quota. QEs fulfill their social responsibilities of green environmental protection, carbon reduction, and low carbon by purchasing GCs. GC life cycle is divided into three stages: issuance, Transactions, and verification.

**Revenue model: based on green power manufacturers and quota enterprises**

GPMs sell GCs in the trading market to subsidize their initial investment costs. The price of the GC will be affected by time, region, supply and demand tension, and other factors. The total revenue of the GPM settlement cycle  $u_{gpm}$  includes electricity sales revenue  $I_{rep}$ , government subsidy income  $I_{(rep,subs)}$ , and GC transaction revenue  $I_{gc}$ .

$$u_{gpm} = I_{ep} + I_{(ep,subs)} + I_{gc} - C_{ep} = I_{(ep,subs)} + \sum_{t=1}^T (\lambda_t^{ep} * Q_t^{ep} + \lambda_t^{gc} * Q_t^{gc}) - C^{ep}. \quad (1)$$

where  $\lambda^{gc}$  is the price of GCs,  $Q^{gc}$  is the quantity of GCs,  $\lambda^{ep}$  is the on-grid price of renewable energy power generation, and  $Q^{ep}$  is the electricity generation of RE,  $C^{ep}$  is the cost of generation of the RE. It is assumed that this work does not consider the impact of government subsidy income  $I_{(ep,subs)}$  on total income. Equation (1) can be expressed as:

$$u_{gpm} = \sum_{p \in P} \sum_{t \in T} (\lambda_{p,t}^{ep} * Q_{p,t}^{ep} + \lambda_{p,t}^{gc} * Q_{p,t}^{gc}) - C^{ep} \quad (2)$$

where  $\lambda_{p,t}^{ep}$  and  $Q_{p,t}^{ep}$  are the online trading electricity price and trading electricity quantity of GPMs in times( $t$ ),  $\lambda_{p,t}^{gc}$ ,  $Q_{p,t}^{gc}$  are the transaction price and transaction quantity of GCs of GPMs in times( $t$ ).

$$C^{ep} = \sum_{p \in P} \sum_{t \in T} (m_{ic,p} * h_{p,t} * \lambda_{p,t}^{cost}) \quad (3)$$

where  $\lambda_{p,t}^{cost}$  is the unit cost of electricity generated by renewable energy,  $m_{ic,p}$  is the installed capacity of renewable energy, and  $h_{p,t}$  is the monthly utilization hours of renewable energy generation. Assume that all renewable energy traded electricity will receive GC, satisfying  $m_{ic,p} * h_{p,t} = Q_{p,t}^{ep}$ . Equation (2) can be simplified as follows:

$$\begin{cases} u_{gpm} = \sum_{p \in P} \sum_{t \in T} ((\lambda_{p,t}^{ep} - \lambda_{p,t}^{cost}) * Q_{p,t}^{ep} + \lambda_{p,t}^{gc} * Q_{p,t}^{gc}) \\ s.t. Q_p^{ep} \geq Q_p^c, \lambda_p^{ep} \geq \lambda_p^{cost} \end{cases} \quad (4)$$

without considering other commercial values of GC in Eq. (4), we consider that its market transaction price can be determined by green power cost, revenue balance and opportunity cost.

The utility function of the QE corresponds to its cost function, and the model is as follows:

$$\begin{cases} u_{qe} = \sum_{p \in P} \sum_{t \in T} (\gamma_{p,t} * Q_{p,t}^{ep} + \lambda_{p,t}^{co2} \varphi(Q_{p,t}^{gc}) - \lambda_{p,t}^{gc} * Q_{p,t}^{gc}) \\ \varphi(Q_{p,t}^{gc}) = \omega Q_{p,t}^{gc} + \Pi_{p,t}^{(co2,qe)} \end{cases} \quad (5)$$

where  $u_{qe}$  is the total cost of the QE,  $\lambda^{ep}$  is the on-grid electricity price (the fixed electricity price model will be used in this work),  $\lambda^{gc}$  is the GC transaction price of the QE,  $Q^{gc}$  is the GC transaction volume of the QE.  $\lambda^{co2}$  is the price of carbon.  $\gamma_{p,t}$  is the ratio of income generated from the production and operation of enterprises using renewable energy (We take the reciprocal of carbon emission intensity of QE as the calculated value of  $\gamma$  in this work).

To reflect the characteristics of the seller's market in the function, this work supplements the function  $\varphi(Q_{p,t}^{gc})$  as a component in the cost function of the QE, which represents the ability of a QE to independently complete a certain amount of quota. It is also a functional relationship between the actual GC transaction volume and a certain percentage of the quota. In this work, the quota completion capability function is a constraint rather than a real cost for the QEs. QE needs to consider whether it can meet the quota requirements when bargaining with GPMs and adjusting the transaction volume strategy.  $\omega$  is the carbon asset conversion rate corresponding to the GC (and the quota rate of the QE). Where  $\Pi_{p,t}^{(co2,qe)}$  is the actual GC trading volume of the enterprise that offsets the carbon quota quantity. We ignore the impact of income change trend. In addition, the price of GCs and the price of carbon assets meet a certain linear relationship (the coefficient is  $\eta$ ):  $\Pi_{p,t}^{(co2,qe)} = \gamma_{p,t} * \eta * Q_{p,t}^{gc}$ . So transformed Eq. (5) as follows:

$$\begin{cases} u_{qe} = \sum_{p \in P} \sum_{t \in T} (\gamma_{p,t} * Q_{p,t}^{ep} + \lambda_{p,t}^{co2} \varphi(Q_{p,t}^{gc}) - \lambda_{p,t}^{gc} * Q_{p,t}^{gc}) \\ \varphi(Q_{p,t}^{gc}) = \omega Q_{p,t}^{gc} + \gamma_{p,t} * \eta * Q_{p,t}^{gc} \end{cases} \quad (6)$$

Equation (6) establishes the correlation between carbon allowances, green certificates, and carbon intensity. Equations (4), (6) provide quantitative relationships for game equilibrium, trading offer optimization, and reinforcement learning convergence.

## Algorithms combining game strategies: Q-learning and smart contracts Game and green certificates

We define the multi-agent collaborative trading game,  $G = \langle P, S, U \rangle$ . Where  $P$  is the set of QEs, GPMs of GCs,  $P = 1, 2, \dots, n$ ;  $S$  is the set of possible strategies to be executed by the transaction;  $S_i$  is the action strategy of each buyer(seller) in the team, and each buyer(seller) makes the corresponding action according to the current surroundings and the environment of other participants. The strategy of each Agent can be formalized as  $(A_1^i, A_2^i, \dots, A_n^i)$ , and  $U$  is the payoff function, which denotes the gain or loss after executing the strategy.

In dynamic and complex environments, the information acquired by agent in a multi-Agent system may be complete or incomplete. Let the environment in which a multi-Agent system is located be  $X$ , and  $X_t$  denotes the environment in which the multi-Agent system is located at moment  $t$ . Let the set of agent observable environmental states be  $S_t$ ,  $S_t^i = f(E_t)$ , and  $S_t = (S_t^1, S_t^2, \dots, S_t^n)$  be the joint observation in the system at moment  $t$ , denoted by  $S = \prod_{i \in N} S_i$ . Let the set of agent actions be  $A$ ,  $A^i$  is used to denote the set of actions of agent  $i$ , the

set of actions of agent is denoted by  $A$  ( $A = \prod_{i \in N} A_i$ ), and from the impact on the environment of the actions taken by each agent at the moment  $t$  of the observed environment  $A_t^i \in A_i$ , the joint action between multiple participants ( $A_t^1, A_t^2, \dots, A_t^n$ ) will also affect the state of the environment in which it is currently located.

Let the state transfer function be  $T, T_t : S \times A \rightarrow S$ , denoting the possible impact on the environment through collaboration between a trader and other transactions in a particular environment.

Let the Agent payoff function be  $U, U_i = S \times A \rightarrow U$ , denoting the payoffs after the actions taken by agent  $i$  to accomplish a task in a multi-agent system. Agent's goal set  $G = G_1, G_2, \dots, G_n$ ,  $G_i$  denotes the goal of each agent in the multi-agent system, which can usually be expressed by using the payment function  $U$ , and there may be multiple relationships between the goals of each agent  $i$ : when the goals are consistent, the completion of the goals between the Agents is mutually reinforcing; when the goals conflict, there is a conflict of interest resources. If there exists a joint action  $a^* \in S$  that satisfies the conditions:  $\forall I \in P, \forall a_i \in S, U(a_i^*, a_{i-1}^*) \geq U(a_i, a_{i-1})$ , then  $a^*$  is said to be a Nash equilibrium of the game  $G$ .

According to Eqs. (4), (6), the GMP revenue function is a quadratic function of the green power transaction and GC market price, with a coefficient of not less than zero and a "convex" property. Therefore, the game model considered has Nash equilibrium. Through the process of establishing the above mathematical model, the GC trading price ( $\lambda^{ep}, \lambda^{gc}$ ) is established based on the transformation of the interests of the participating subjects. The trading subjects derive the optimal strategy based on the combination of variables ( $Q_{i,t}^{ep}, Q_{i,t}^{gc}, \Pi_{i,t}^{co2, qe}$ ) and adjustment coefficients ( $\gamma_{i,t}, \eta, \omega$ ), and thus determine the Nash equilibrium point. Considering the introduction of non-convex characteristics of decision variables in equation. The traditional solution algorithm has certain difficulties, so this paper adopts the Therefore, this paper adopts a multi-agent Nash-Q reinforcement learning algorithm and Smart matching of bidding under smart contracts.

### Q-learning-based strategy analysis and green certificate trading process

Q-learning have been heavily researched in finance, especially in stock trading, and in the energy field focusing on energy management. Given the attributes of GCs and the market participants in the transaction subject, in addition to the different liquidity, the mode of its transaction can be referred to. So, we combine multi-agent reinforcement learning with game theory and construct a framework for applying multi-participant bidding games in GC markets using multi-agent Nash-Q reinforcement learning. Agents try out rewards or punishments given by the environment during the learning process and gradually develop expectations about incentives to formulate reward-maximizing strategies. We adopt the Q-learning algorithm in reinforcement learning, which involves four kinds of parts: (1) Q-table,  $Q(s, a^1, \dots, a^n)$  is the cumulative value of executing  $a$  action in state  $s$ , (2) selecting action  $a$ , (3) making action and environment feedback, (4) environment update. During its process  $Agent_i$  observes the surrounding environment and executes the actions in the action strategy set. At time  $t$ ,  $Agent_i$  acts  $a_i$ , while feedback gain  $R(S_t, a_i)$ , updates the Q-value table, and repeats the above process until the end of the task. The value of  $Q(S_t, a_i)$  can be expressed as:  $Q(S_t, a_i) = R(S_t, a_i) + \beta \max Q(S_t + 1, a_i + 1)$ , where  $a$  is an action in the action strategy set;  $\beta$  ( $0 \leq \beta \leq 1$ ) is influence factor.

The GC seller GPMs make a joint action offer based on the green power trading price, generation efficiency, and environmental status such as cost recovery cycle and carbon asset. Therefore, a reinforcement learning algorithm is added to the set of trading strategies to improve the set of strategies selected by the participants' actions. the set of strategies at moment  $t$  is denoted as  $S_t: S_t \subseteq R$ . The purchaser and seller of GCs choose their respective offer actions at a certain moment  $t$ .

The purchaser can separately compute the payment matrix  $U_t$  of the purchaser for both parties under different strategy choices based on the defined payment function. It is difficult for buyers (QEs) to be informed of the state-action values  $Q$  of sellers (GPMs) to accurately find suitable strategies to cope with them, so we incorporate Q-learning methods to learn the action-state values of GPMs and adjust them to a targeted set of trading strategies. The game strategies of GPMs are learned through reinforcement learning to develop suitable bidding strategies for buyers.

The learning task of Step-T cumulative reward is added to the algorithm, starting from the initial state of the offerer, so that the offerer obtains an offer trajectory of GPMs with Step-T after learning:  $\langle x_0, a_0, r_1, x_1, a_1, r_2, \dots, x_{t-1}, a_{t-1}, r_t, x_t \rangle$ . In order to obtain the optimal strategy, the  $\epsilon$ -greedy algorithm is introduced, which selects one action uniformly at random from all actions with probability  $\epsilon$ . The current optimal action is selected with probability  $1 - \epsilon$ , and the identified strategy is labeled as the "original strategy". The policy that uses the  $\epsilon$ -greedy algorithm in the original policy is denoted as Eq. (7):

$$\pi(a|s) = \begin{cases} 1 - \epsilon + \frac{\epsilon}{|\mathcal{A}(s)|}, & \text{if } a = \arg \max_a Q^\pi(s, a) \\ \frac{\epsilon}{|\mathcal{A}(s)|}, & \text{if } a \neq \arg \max_a Q^\pi(s, a) \end{cases} \quad (7)$$

In GC trading process, the sum of cumulative rewards for each pair of state-action  $Q$  in the trajectory is recorded as a one-time sampling value of cumulative rewards with respect to the GC seller. When multiple GC seller trajectories are obtained by sampling the GC seller multiple times, the cumulative reward sampling values obtained multiple times will be averaged using Eq. (8) to obtain an estimate of  $Q$ .

$$Q_n(k) = \frac{1}{n}((n-1) \times Q_{n-1}(k) + u_n) = Q_{n-1}(k) + \frac{1}{n}(u_n - Q_{n-1}(k)) \quad (8)$$

During the Nash-Q intensive learning process for GC trading based on GPMs, QEs, and other participants, the corresponding iterative refinement of the Q-table is illustrated in Algorithm 1. To validate the effectiveness of Nash Q-Learning in optimizing GC trading and enhancing sustainability by promoting the increased use of renewable energy sources, a comprehensive experimental setup and an evaluation strategy are essential. We

employ the trade-off between exploration and exploitation, known as the Exploration-Exploitation Tradeoff ( $\epsilon$ -greedy).

---

**Require:** Initialization to generate GPMs and QEs:  
 $G = \langle P, S, U \rangle \leftarrow$  Build agents Pool{GPMs, QEs}  $\leftarrow \{\forall I \in P, \forall a_i \in S, U(a_i^*, a_{i-1}^*) \geq U(a_i, a_{i-1})\}$ ;  
 Set  $I_{gpm} \in I_{gpm,1}, \dots, I_{gpm,i}$  and the price  $\Delta\lambda \in \Delta(\lambda_{gpm}, \lambda_{qe}), \Delta(\lambda_{gpm}^*, \lambda_{qe}^*) \leftarrow \Delta Q \in \Delta(Q_{gpm}, Q_{qe}), \Delta(Q_{gpm}^*, Q_{qe}^*)$  ;  
 Get the initial state  $s_0, t = 0, s \in S, Q_i^j(s, a_1, \dots, a_n) = 0$ ;  
 $\forall a_i \in A_i, i = 1, 2, \dots, n$ ;  
 Initialization matrix:  $U_i \leftarrow$  EQ. 4 and EQ.6;  
**Ensure:** optimal electricity price  $\lambda_{gpm,i}^*$  and the optimal quantity  $Q_{qe,i}^*$  || Calculate the maximum revenue.  
**for** iteration  $i = 0$  to  $N$  **do**  
   Use agent  $i$  state  $s_i$ , with  $\epsilon$ -greedy in EQ.7  $\rightarrow a_i$  ;  
   Execute  $\pi$ , and Step-T:  $(x_0, a_0, r_1, x_1, a_1, r_2, \dots, x_{t-1}, a_{t-1}, r_t, x_t)$ ;  
   Choose action  $a_t^j \rightarrow$  Cumulative rewards with EQ.8;  
   **if** staes  $s_t$  and action  $a_1 \rightarrow NashQ_i^j(s, a^1, \dots, a^n)$  **then**  
     Goto End for;  
   **else**  
     Get Next State  $s'$  and agent  $i$  || (Observe  $r_t^1, r_t^2, \dots, r_t^n$ ) || Joint action  $(a_1, \dots, a_n) \leftarrow s$  for all agents ||  $s_{t+1} = s'$ ;  
   **end if**  
**end for**  
 Update  $Q_i^j$  for  $i = 1, 2, \dots, n \leftarrow$  EQ.8 ||  $Q_{t+1}^j(s, a^1, \dots, a^n) \leftarrow (1 - \alpha)Q_t^j(s, a^1, \dots, a^n) + \alpha[r_t^j + \beta NashQ_i^j(s')]$ .  
 $count(x_i, a_i) = count(x_i, a_i) + 1$ ,  
 Let  $t \leftarrow t + 1$ .  
**return** result the optimal electricity price  $\lambda_{gpm,i}^*$  and the optimal quantity  $Q_{qe,i}^*$   
 Calculate the maximum revenue

---

**Algorithm 1.** Collaborative algorithms based on game theory and Q-learning.

### Smart contracts and transaction execution

In this work, we integrate game models with smart contracts to solve the bargaining, consensus cooperation, and trading equilibrium problems in GC trading. Based on the Nash equilibrium strategy, a trading contract script under smart contract is established to dynamically implement the trading strategy (NES-SC).

The GC trading game model, coupled with the operational mechanics of smart contracts, guarantees that all transaction and management parties can dynamically and intelligently complete transactions within given constraints. To ensure a systematic approach to equilibrium and market clearing, the execution flow of this process is outlined in Algorithm 2.

---

**Require:** Establish a self-interest model for GPMs and QEs:  
 $\{\lambda_{i,t}^{ep}$  in Eq. 3 and  $\eta_{i,t}, \lambda_{i,t}^{co2}$  and  $\pi_{i,t}^{(co2,em)} \in$  Eq.2.}  $\leftarrow$  Dataset;  
 $G = \langle P, S, U \rangle$  Build agents Pool {GPMs, QEs}  $\leftarrow \{\forall I \in P, \forall a_i \in S, U(a_i^*, a_{i-1}^*) \geq U(a_i, a_{i-1})\}$ ;  
**Ensure:** The maximum revenue in the equilibrium state;  
 initial game scheme  $\lambda_{gpm,i}^*(0)$  and  $Q_{qe,i}^*(0)$   
**for** iteration  $k = 0$  to  $N$  **do**  
    $P: S \rightarrow R, \forall i \in N, \forall s^{-i} \in S^{-i}$  and  $\forall \{x, y\} \in S \forall i \in N, \forall s^{-i} \in S^{-i}$  and  $\forall \{x, y\} \in S$   
   **if**  $P(x, s^{-i}) - P(y, s^{-i}) > 0 \Leftrightarrow u_i(x, s^{-i}) - u_i(y, s^{-i}) > 0$  **then**  
     Game equilibrium strategy and smart contract script generation.  
     Set price fluctuation interval  $\Delta\lambda \in \Delta(\lambda_{gpm}, \lambda_{qe}), \Delta(\lambda_{gpm}^*, \lambda_{qe}^*)$  by  $\Delta Q \in \Delta(Q_{gpm}, Q_{qe}), \Delta(Q_{gpm}^*, Q_{qe}^*)$   
     adjust the distance  $\Delta\lambda$  of the fastest intersection. batch participants of the gradient value with the offer reference value contract set, meet the Eqs. 3, 6;  
     Goto end for  
   **else**  
      $k \leftarrow k + 1$   
      $\lambda_{gpm,i}^*(k + 1)$  and  $Q_{qe,i}^*(k + 1)$   
     Iterative solution;  
     Continue  
   **end if**  
**end for**  
 the optimal electricity price  $\lambda_{gpm,i}^*$  and the optimal quantity  $Q_{qe,i}^*$

---

**Algorithm 2.** NES-GC Process Improvement for Smart Contracts.

## Results and analysis

### Basis of data

In this paper, we focus on the demand for Green Certificates (GCs) from enterprises in thermal power, chemical industry, iron and steel, and cement sectors within China's eight major carbon emission categories. Three GC supply enterprises are selected to partake in the simulation transaction. The relationship between GC price, green

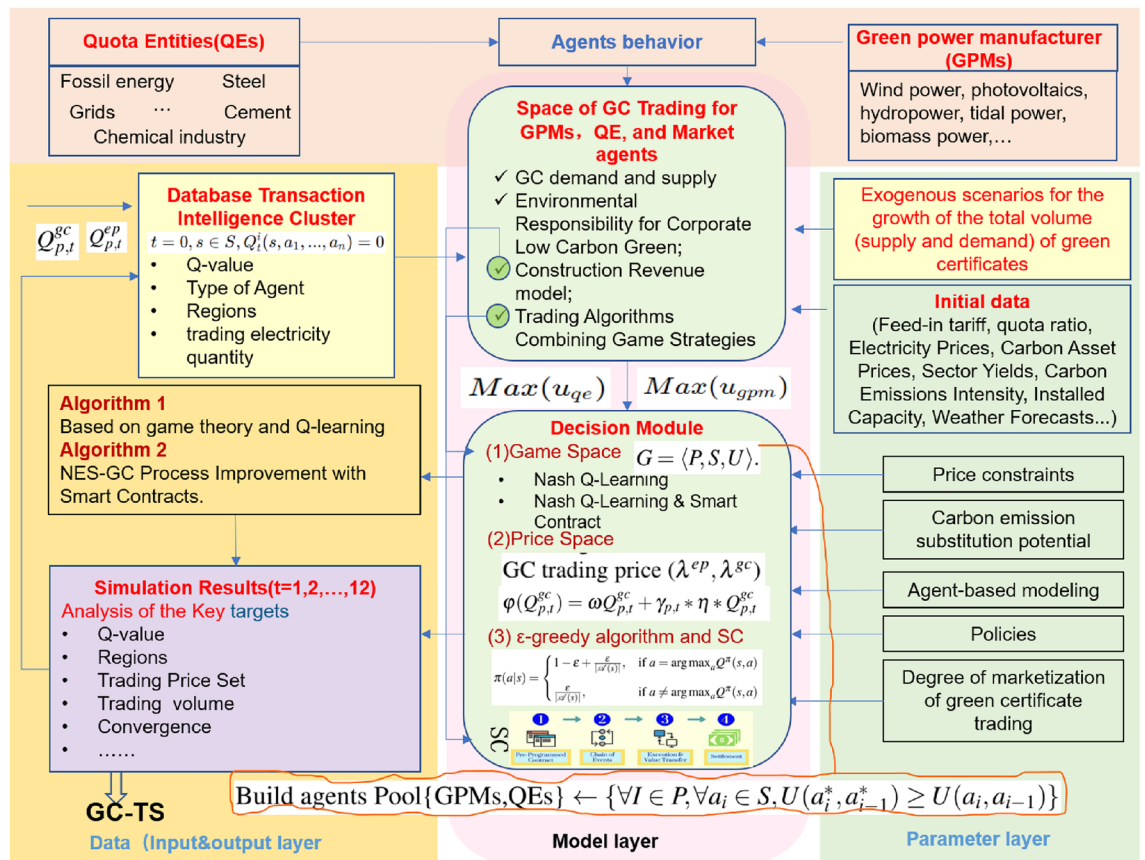


Figure 3. Methodological framework for logical relationships between models.

QEs	Enterprise attributes	Revenue (10 kt)	Carbon emissions (100 million CNY)	Carbon emission intensity (tons/10k CNY)	Data source
QE-1	Non ferrous	1860	8904	4.79	
QE-2	Steel	370	1758	4.75	*
QE-3	cement	71	1430	20.06	*
QE-4	Comprehensive	4645	4337	0.93	
QE-5	Power plant	142	3694	25.98	*
QE-6	Coal	338	3026	8.97	*
QE-7	Petrifaction	21,060	17,094	0.81	
QE-8	Shipping	1713	2084	1.22	
QE-9	Communication	3038	1423	0.47	
QE-10	Chemical industry	167	831	4.98	

Table 1. Carbon emissions of CQ companies (2020, extracts from Chinese listed companies, CNY).

power trading price, and carbon asset price is delineated as Eq. (5). The convergence and trading return behavior are governed by Eqs. (4), (6), with carbon emissions intensity data referenced in Table 1, and the methodological framework for logical relationships between models (Fig. 3).

This study simulates the Nash bargaining scheme for bilateral GC transactions, utilizing historical data of Wind-GC. CQ-fulfilling enterprises across different industries in 2020 are chosen as Qualified Entities (QEs) (see Table 1). In the case study, four QEs and three Green Power Markets (GPMs) (Wind) are selected to engage in the gaming process to substantiate the analysis. The algorithms 1, 2 are coded and executed in python3.9.

In addition, we take the base price as 0.43 (CNY)/kWh, the operating cost as 0.2 (CNY)/kWh,  $\omega=0.96$  (t./MWh) and combine the data in Table 1 (\* is the estimated data of Zhong Chuang Carbon Investment & Finance magazine. Enterprises marked with \* have their own disclosure, and enterprises marked with blank use their own public disclosure of data.) to verify the analysis. In this paper, we set  $\lambda_{gpm,1}^*(0) = 210$ ,  $\lambda_{gpm,2}^*(0) = 200$ ,  $\lambda_{gpm,3}^*(0) = 196$

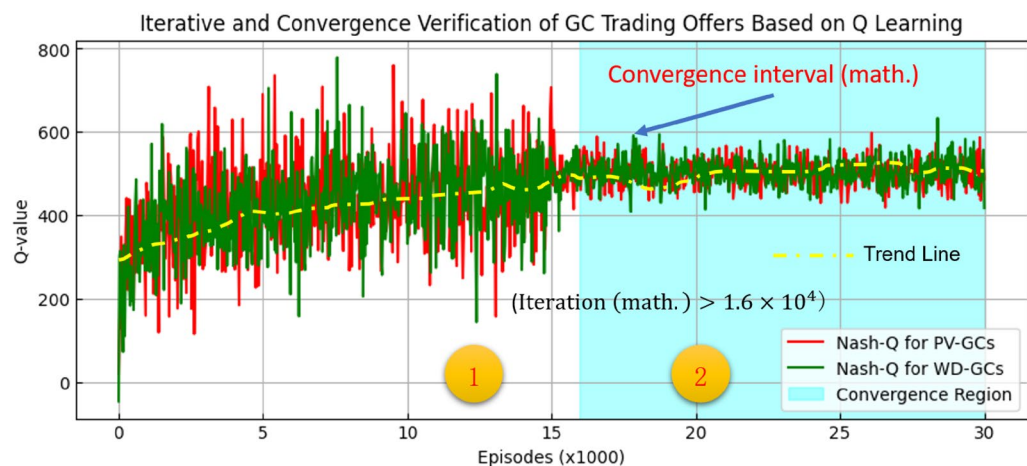
## Analysis of results

In the results analysis section, we present the performance of the GC-TS architecture and trading mechanism through simulations and experimental validation under varying market conditions and policy environments. By establishing a multi-agent reinforcement learning game equilibrium model and proposing a Nash Q-learning offer clearing mechanism, as well as integrating smart contracts to facilitate smart trading, we aim to validate the analysis of the convergence efficiency of the trading offer game, trading efficiency, and optimized trading benefits.

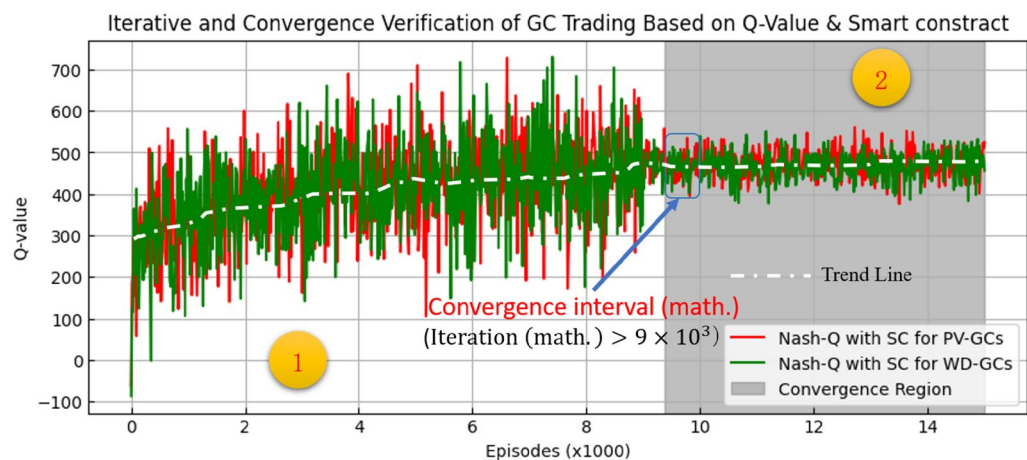
### Convergence analysis

In Fig. 4, we observe the following:

1. On the Q-value Analysis aspect, Fig. 4a shows the Q-values of the Nash Q-learning algorithm for both photovoltaic (PV-GCs) and wind (WD-GCs) green certificates without the integration of smart contracts. The Q-values for photovoltaics tend to be slightly higher than those for wind, suggesting that photovoltaic projects might expect slightly higher returns in the simulated trading environment. In Fig. 4b, when smart contracts (SC) are introduced, the Nash Q-learning with SC demonstrates Q-values that stabilize at higher iterations with less fluctuation, indicating that smart contracts might offer additional stability to the trading process.
2. On the Convergence speed, Fig. 4a indicates that the system begins to show convergence behavior after  $1.6 \times 10^4$  iterations. In contrast, Fig. 4b shows the onset of convergence earlier, after just  $9 \times 10^3$  iterations, suggesting that the integration of smart contracts may accelerate the convergence speed of the Nash Q-learning algorithm.
3. On consistency of GC trade clearing, both Fig. 4a and b, the stability of Q-values within the convergence interval signifies a consistent clearing process. This means that as learning progresses, trading strategies for



(a) Convergence of GC trading under Q-Learning



(b) Convergence of GC trading under integrated smart contracts and Q-Learning

**Figure 4.** Convergence analysis of green certificate trading.



both types of green certificates begin to align, with agents finding more optimal strategies for maximizing returns.

In summary, the analysis of these two Fig. 4a,b reveals the potential advantages of the Nash Q-learning model integrated with smart contracts in terms of convergence speed, stability, and consistency of trading strategies. These advantages may bring significant benefits to actual green certificate trading, thereby promoting the use of renewable energy and the overall development of the market.

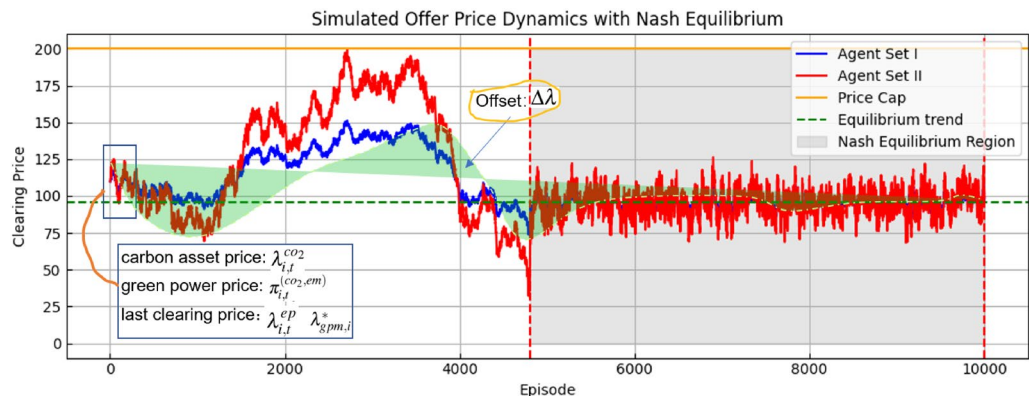
*Price analysis*

Figure 5 shows a simulation of how the bid prices of the two players change over the episode. The Nash equilibrium is represented by the green solid line, with the Episode outside the 5000 range and bids converging to the value of the Nash equilibrium. The chart displays the evolving dynamics of offer prices from two groups of agents involved in the green certificate market, categorized as Agent Set I and Agent Set II, within a Nash Q-learning framework. Their pricing strategies are informed by the current green power price ( $\pi_{(co_2,em)}$ ), carbon asset trading price ( $\lambda_{i,t}^{co_2}$ ), and the last round's clearing price ( $\lambda_{i,t}^{ep}, \lambda_{gpm,i}^*$ ). Regarding the offer trend and adjustment (Offset  $\delta\lambda$ ), the green power price and carbon asset price provide the base price parameters for the trading parties. As seen in Fig. 5, these prices generally fall within the range of offers from both sides; the offer adjustment (Offset  $\delta\lambda$ ) indicates that both sides continually adjust their offers in search of market equilibrium. Before reaching consistency within the Nash Equilibrium Region, there might be a significant difference in the offers from the buyer and seller. The equilibrium trend analysis shows that (1) after initial fluctuations, the offers from both trading parties begin to converge toward the equilibrium trend line (dashed green line), indicating the market is moving towards stability and consistency; (2) in the chart, around the 2000 to 4000 episode range, offers begin to converge within the Nash Equilibrium Region, suggesting that after a series of dynamic adjustments, the market has found an equilibrium price acceptable to both sides. Additionally, the analysis of convergence and market clearing in the chart found that as the number of episodes increases, the offers from Agent Set I and Agent Set II tend to stabilize within the Nash Equilibrium Region. This convergence demonstrates the effectiveness of the proposed Nash Q-learning offer clearing mechanism, capable of implementing smart trading by integrating smart contracts, thus enhancing trading efficiency and optimizing trading benefits.

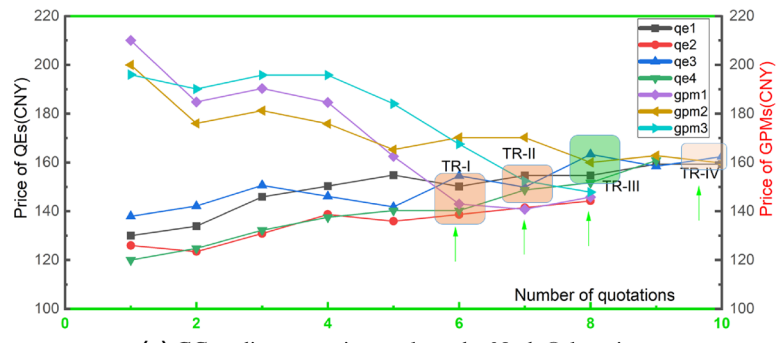
*Comparison of two algorithms*

Using the algorithm 1,2, we compare the speed of convergence after 5000 iterations of 100 3-agent simulations, 20 5-agent simulations, and 20 7-agent simulations. The game convergence scenario becomes more difficult to converge as the number of intelligences increases. This is illustrated by the example runs with 3, 5, 7, and 9 trading intelligence in Figs.6, 7.

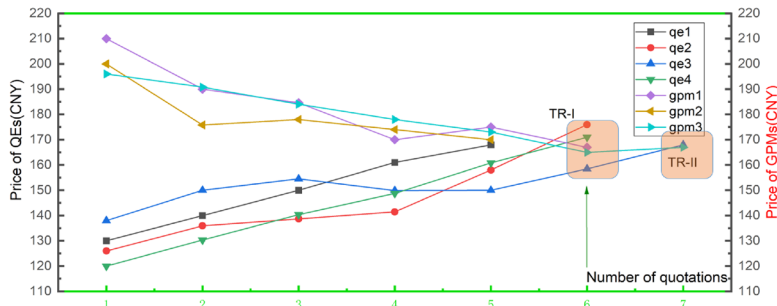
In the Nash Q-learning (NES) and Nash Q-learning & Smart Contract (NES-SC) model model validation cases, a total of seven intelligences, QEs, and GPMs, were used to participate in the GC quotation trading. Four QEs and three GPMs achieve 10 rounds of quotes in this case. Figure 6a,c are the quotation and strategy spending time in NES mode, respectively. The supply and demand sides continuously adjust the price, trading volume, and other game strategies, and multiple quotes form a trading trend when  $k \geq 5$ . There are TR-I, TR-II, TR-III, and TR-IV transaction clusters that satisfy the price and volume demands in the offer cycle in Fig.6a, and the model cannot guarantee that all demands are satisfied when  $k = 10$  and the transaction is not yet completed. The number of offers increases, the amount of information obtained increases, and the time spent by the strategy subsequently decreases as in Fig.6c. In the same data case, we performed a case study of NES-SC, and Fig. 6b,d are the offer and strategy spend time, respectively. Figure 6b shows that both sides basically complete the transaction (two transaction clusters TR-I and TR-II) when  $k = 7$ . Compared with Fig. 6a,c, the transaction success rate of both sides is improved and the strategy spending time is shortened. In conclusion, the two methods are consistent



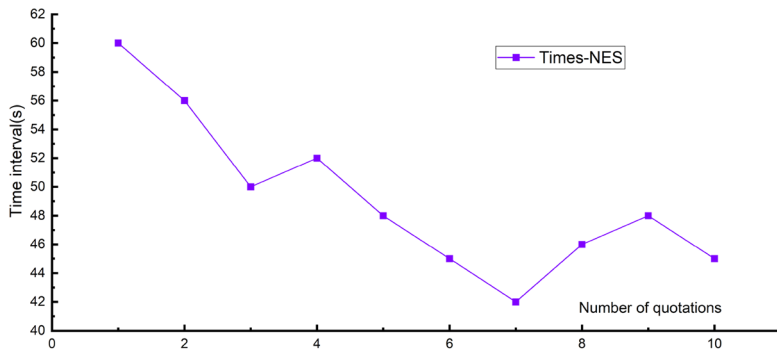
**Figure 5.** Trend analysis of GC price under the model learning process.



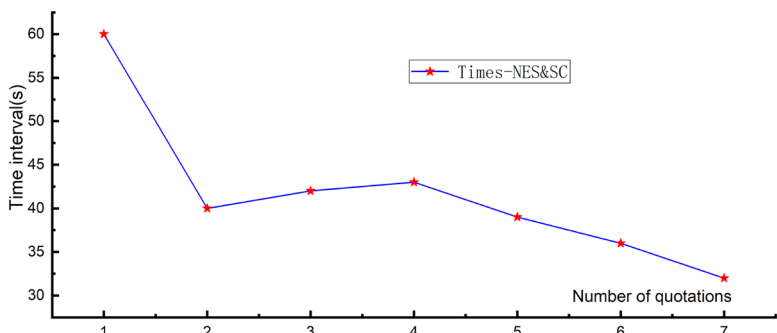
(a) GC trading quotation cycle under Nash Q-learning



(b) GC trading quotation cycle under integrated smart contracts and Q-Learning



(c) Quotation time period under Nash Q-learning



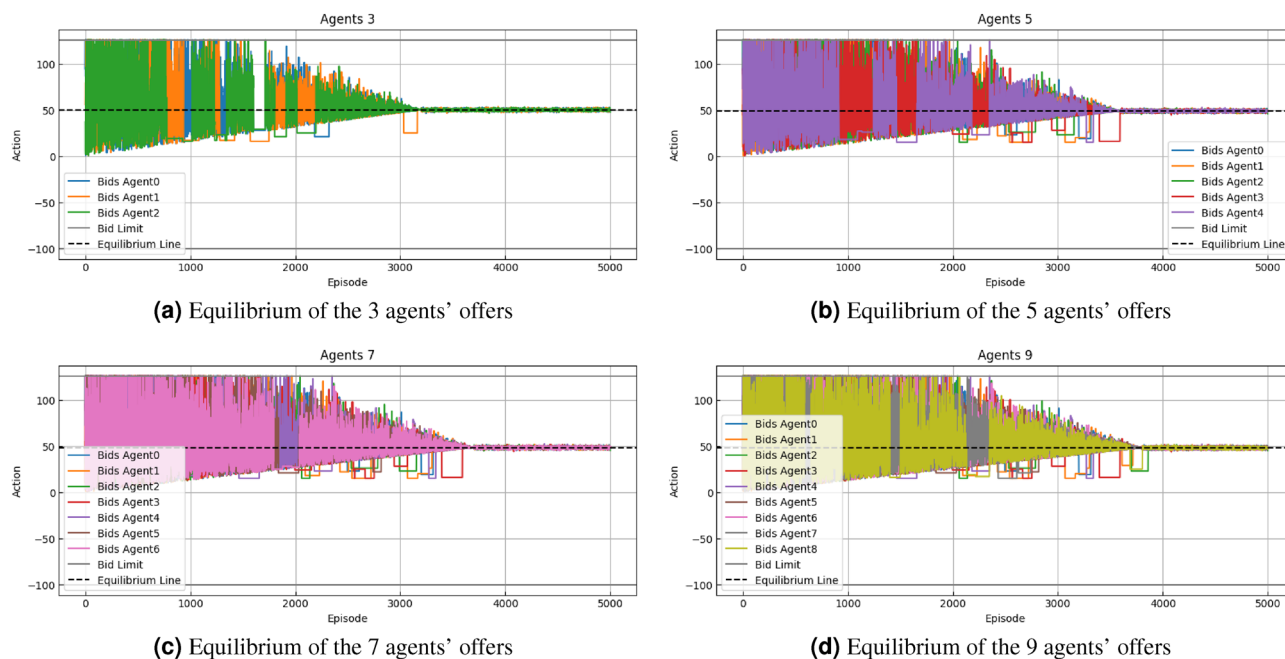
(d) Quotation time period under integrated smart contracts and Q-Learning

**Figure 6.** Quotation and time period in two models (Nash Q Learning: algorithm 1, Nash Q Learning & SC: algorithm 2) in case.

in the trend of quotes, and the NES-SC method is based on the NES model, so both methods can effectively support the trading platform, and NES-SC has better advantages in transaction efficiency.

*Robustness analysis*

To validate the robustness of the green certificate trading model, we increased the complexity of transactions by expanding the number of agents involved. The study employs cluster agents representing different industries,



**Figure 7.** Robustness analysis of green certificate trading models.

such as power grid, steel, chemical, and thermal power, with each group consisting of several individual entities participating in collective quotation trading.

In Fig. 7, we observe the bidding behavior of agent groups with 3, 5, 7, and 9 agents participating in Nash Q-learning and integrated smart contract trading strategies. These charts reflect how agents' bidding actions change over time and attempt to approach the equilibrium line. In Fig. 7a, the bidding behavior of the agents exhibits significant fluctuations, which may reflect the uncertainty of decision-making in a smaller group of agents. Despite the fluctuations, bids tend to converge toward the equilibrium line over time. As the number of agents increases, the volatility of the bidding behavior seems to increase, indicating that reaching a consensus becomes more difficult in a larger group. However, over time, bids tend to stabilize and move closer to the equilibrium line (in Fig. 7b–d). Nevertheless, from the Fig. 7c,d, the largest group of agents, despite significant volatility, the bidding trends gradually approach the equilibrium line over the long term. So this indicates that even in more complex multi-agent environments, the combined strategy of Q-learning and smart contracts can still guide groups of agents toward collaboration and equilibrium.

From these simulations, we can conclude that as the number of agents increases, the complexity of the trading strategy and the difficulty of reaching equilibrium also increase correspondingly. The integration of Nash Q-learning and smart contracts is crucial for enhancing the coordination of group behaviors and driving toward an equilibrium in trades, especially in complex trading environments involving more agents. These simulations emphasize the importance of considering agent diversity and collective behavior patterns when designing green certificate trading systems.

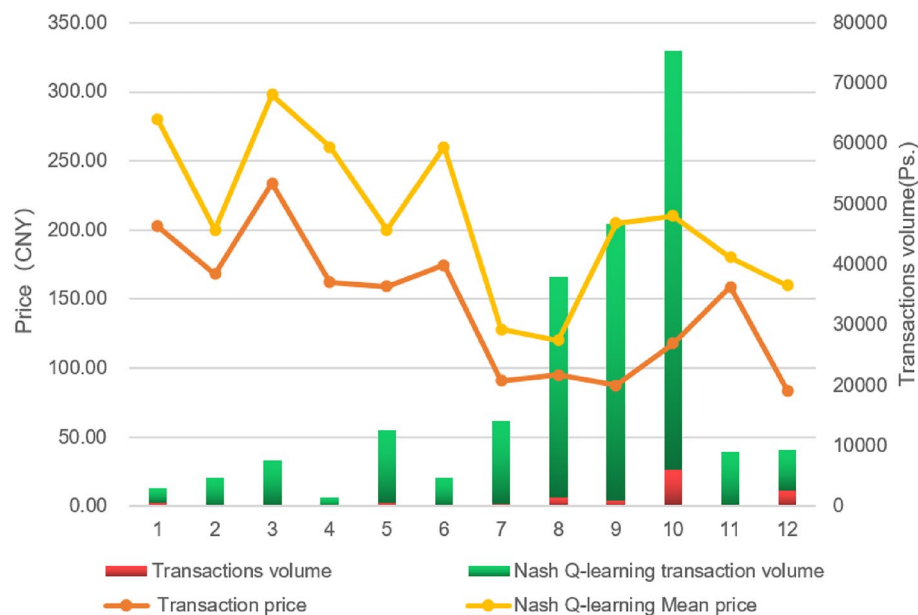
### Comparison of green certificate trading

As the actual value of GCs in the future carbon neutrality is more and more emphasized by the Chinese government, and the value-added benefits of carbon quota enterprises through the carbon offset mechanism of GCs are taken into account.

Figure 8 is the analysis of the monthly mean values of the quoted price and trading volume under the gaming strategy of the participating counterparties. The trading model designed in this paper underpins the overall transaction price and volume. Using an equilibrium strategy, it swiftly locates the trading point for each participant, and compared with the trading data from 2021, the transaction volume has increased by up to 30 times. The average transaction price induced by the Nash Q-learning strategy (yellow line) is generally higher than the actual transaction price (orange line), thanks to the game-theoretic strategy negotiation and bid adjustments of the Nash Q-learning model.

In terms of transaction volume, the graph shows a significant peak in Nash Q-learning transaction volume (green bar) in the ninth month, which occurs concurrently with the peak of the Nash Q-learning mean price (yellow line). The data reveals that the forecasted transaction volume by Nash Q-learning reached nearly 70,000 units, whereas the actual volume was only about 2000 units. As for pricing, the predicted price by Nash Q-learning for the ninth month reached close to 300 CNY, while the actual price was around 200 CNY. The actual transaction volume (red bar) and price (orange line) exhibit some volatility with noticeable peaks and valleys. The transaction volume prediction by Nash Q-learning (green bar) is more stable.

The Fig. 8 illustrates the potential value-added benefits of applying the Nash Q-learning strategy in the GC market, reflecting its capability to quickly locate optimal trading points for participants, potentially leading to



**Figure 8.** Comparison of price and volume based on Nash Q-Learning simulation with actual transactions.

an increase in price and volume compared to historical data. This aligns with the objective of enhancing the effectiveness of the GC trading model and the carbon offset mechanism in the context of carbon neutrality goals.

## Conclusion

Given the complexity of issuing, verifying, and trading GCs in China, coupled with the challenges brought forth by policy shifts, it is imperative to ensure that the market-based GC trading system is bolstered by adequate mechanisms and technological support. This study proposes an architecture for a zero-carbon energy Green Certificate Trading System (GC-TS) that leverages an equilibrium strategy, enhancing the efficiency of GC trading quotes and facilitating multi-party collaboration through the incorporation of Q-learning, smart contracts, and an effectively integrated multi-agent Nash strategy. Initially, the system integrates GC trading with electricity and carbon asset markets, formulating pricing strategies across these domains. Subsequently, a certificate-electricity-carbon efficiency model is designed to maintain consistency across the GC, green electricity, and carbon markets. Aiming for diversified GC trading, a multi-agent reinforcement learning game equilibrium model is established, alongside a Nash Q-learning offer clearing mechanism that employs smart contracts for intelligent trading, thereby increasing the convergence efficiency of the trading offer game and enhancing trading efficacy while optimizing benefits for all parties involved.

In summary, this research plays a constructive role in advancing the development of renewable energy in China. It highlights the scalability of the GC-TS and its potential as a reference model for future GC trading platforms in China. The system offers a swift avenue for green power producers to subsidize the costs of renewable energy generation through GC proceeds, facilitating the involvement of carbon quota enterprises in carbon emission responsibility compliance, and promoting the growth of renewable energy sources.

## Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Received: 6 January 2024; Accepted: 24 April 2024

Published online: 12 May 2024

## References

1. National Development and Reform Commission, National Energy. Notice on establishing and improving renewable energy power consumption guarantee mechanism[EB/OL]. <http://zfxgk.nea.gov.cn/auto87/201905/t20190515-3661.htm>
2. National Energy Administration. Promote the development of hydrogen energy and introduce a renewable energy quota system as soon as possible [EB/OL]. <http://www.nea.gov.cn/2019-%2003/08/c-137877887.htm>
3. Xin, C. *et al.* The coupling effect of carbon emission trading and tradable green certificates under electricity marketization in China. *Renew. Sustain. Energy Rev.* **187**, 113750 (2023).
4. Verhaegen, Karolien, Meeus, Leonardo & Belmans, Ronnie. Towards an international tradable green certificate system-The challenging example of Belgium[J]. *Renew. Sustain. Energy Rev.* **13**(1), 208–215 (2009).
5. Ganhammar, Kajsa. The effect of regulatory uncertainty in green certificate markets: Evidence from the Swedish-Norwegian market. *Energy Policy* **158**, 112583 (2021).
6. Verbruggen, Aviel & Laes, Erik. Early European experience with tradable green certificates neglected by EU ETS architects. *Environ. Sci. Policy* **119**, 66–71 (2021).

7. Amundsen, Eirik S. & Bergman, Lars. Green certificates and market power on the nordic power market. *Energy J.* **33**(2), 101–117 (2012).
8. Nielsen, L. & Jeppesen, T. Tradable green certificates in selected European countries - overview and assessment. *Energy Policy* **31**(1), 3–14 (2003).
9. Coria, J. & Jaraite, J. Ownership structure and prices: A case study of the Swedish tradable green certificate market. *Energy Policy* **183**, 113822 (2023).
10. Ford, Andrew, Vogstad, Klaus & Flynn, Hilary. Simulating price patterns for tradable green certificates to promote electricity generation from wind. *Energy Policy* **35**(1), 91–111 (2007).
11. Marchenko, O. V. Modeling of a green certificate market. *Renew. Energy* **33**(8), 1953–1958 (2008).
12. Wang, Hui *et al.* An agent-based modeling approach for analyzing the influence of market participants' strategic behavior on green certificate trading. *Energy* **218**, 119463 (2021).
13. Tao, Yuechuan *et al.* Renewable energy certificates and electricity trading models: Bi-level game approach. *Int. J. Electr. Power Energy Syst.* **130**, 106940 (2021).
14. Yuechuan, Tao *et al.* Bi-level optimization model of renewable energy certificates trading based on game theory. *J. Glob. Energy Interconnect.* **4**(1), 64–76 (2021).
15. Na, Chunming *et al.* Green power trade behavior in China's renewable portfolio standard: An evolutionary game-based system dynamics approach. *Sustainability* **14**(9), 5155 (2022).
16. Xuena, An. *et al.* Two-stage joint equilibrium model of electricity market with tradable green certificates. *Trans. Inst. Meas. Control* **41**, 1615–1626 (2019).
17. Hasani-Marzooni, M. & Hosseini, S. H. Trading strategies for wind capacity investment in a dynamic model of combined tradable green certificate and electricity markets. *IET Gen. Transm. Distrib.* **6**(4), 320–330 (2012).
18. Hindsberger, M. *et al.* Co-existence of electricity, TEP, and TGC markets in the Baltic Sea Region. *Energy Policy* **31**(1), 85–96 (2003).
19. Zhang, Y. *et al.* Impacts of renewable portfolio standard on carbon emission peaking and tradable green certificate market: A system dynamics analysis method. *Front. Energy Res.* **10**, 963177 (2022).
20. Yamaguchi, J. A. R., Rachael, S. T. & Carvalho, A. P. Blockchain technology in renewable energy certificates in Brazil. *Bar - Braz. Adm. Rev.* **18**, e200069 (2021).
21. Zeyu, Shen *et al.* Distributed energy trading technology based on Blockchain. *Proc. Chin. Soc. Electr. Eng.* **41**(11), 3841–3850 (2021).
22. Karumba, Samuel *et al.* HARB: A hypergraph-based adaptive consortium blockchain for decentralized energy trading. *IEEE Internet Things J.* **9**(16), 14216–14227 (2022).
23. Zuo, Yanjun. Tokenizing renewable energy certificates (RECs)-A blockchain approach for REC issuance and trading. *IEEE Access* **10**, 134477–134490 (2022).
24. Qingsu He, YuXu. *et al.* A consensus and incentive program for charging piles based on consortium blockchain. *CSEE J. Power Energy Syst.* **4**(4), 452–458 (2018).
25. Han, Dong *et al.* Smart contract architecture for decentralized energy trading and management based on blockchains. *Energy* **199**, 117417 (2020).
26. Changsen, Feng *et al.* Design and implementation of joint trading market for green power certificate and carbon based on smart contract. *Autom. Electr. Power Syst.* **45**(23), 1–11 (2021).
27. Helgesen, Per Ivar & Tomasgard, Asgeir. An equilibrium market power model for power markets and tradable green certificates, including Kirchhoff's Laws and Nash-Cournot competition. *Energy Econ.* **70**, 270–288 (2018).
28. Hulshof, Daan, Jepma, Catrinus & Mulder, Machiel. Performance of markets for European renewable energy certificates. *Energy Policy* **128**, 697–710 (2019).
29. Guo, Hongye *et al.* Modeling strategic behaviors of renewable energy with joint consideration on energy and tradable green certificate markets. *IEEE Trans. Power Syst.* **35**(3), 1898–1910 (2020).
30. Yi, Zuo *et al.* Research on tradable green certificate benchmark price and technical conversion coefficient: Bargaining-based cooperative trading. *Energy* **208**, 118376 (2020).
31. Umit, Cali *et al.* Cybersecure and scalable, token-based renewable energy certificate framework using blockchain-enabled trading platform. *Electr. Eng.* **106**, 1–12 (2022).
32. Xian, Zhang *et al.* Design and application of green power trading system based on blockchain technology. *Autom. Electr. Power Syst.* **46**(9), 1–10 (2022).
33. Cheng, L. C. *et al.* A novel trading strategy framework based on reinforcement deep learning for financial market predictions. *Mathematics* **9**(23), 3094 (2021).
34. Taghian, M. *et al.* Learning financial asset-specific trading rules via deep reinforcement learning. *Expert Syst. Appl.* **195**, 116523 (2022).
35. Jeong, G. *et al.* Improving financial trading decisions using deep Q-learning: Predicting the number of shares, action strategies, and transfer learning. *Expert Syst. Appl.* **117**, 125–138 (2019).
36. Massahi, M. *et al.* A deep Q-learning based algorithmic trading system for commodity futures markets. *Expert Syst. Appl.* **237**, 121711 (2024).
37. Avval, A. E. *et al.* Auction design for the allocation of carbon emission allowances to supply chains via multi-agent-based model and Q-learning. *Comput. Appl. Math.* **41**(4), 170 (2022).
38. Jensen, S. G. & Skytte, K. Interactions between the power and green certificate markets. *Energy Policy* **30**(5), 425–435 (2002).
39. Zhang, Qi. *et al.* Substitution effect of renewable portfolio standards and renewable energy certificate trading for feed-in tariff. *Appl. Energy* **227**, 426–435 (2018).
40. Zhang, X. Y. *et al.* Assessing the policy synergy among power, carbon emissions trading and tradable green certificate market mechanisms on strategic GENCOs in China. *Energy* **278**, 127833 (2023).
41. Song, M. *et al.* Investment and production strategies of renewable energy power under the quota and green power certificate system. *Energies* **15**(11), 4110 (2022).
42. Zheng, J. J. *et al.* How to improve the effectiveness of Chinese green certificate market? A complex network and social influence analysis. *J. Clean. Prod.* **380**, 134943 (2022).
43. Wu, Y. X. *et al.* Multi-oligarch dynamic game model for regional power market with renewable portfolio standard policies. *Appl. Math. Model.* **107**, 591–620 (2022).
44. Tan, Y. X. *et al.* A two-phase hybrid trading of green certificate under renewables portfolio standards in community of active energy agents. *Energies* **15**(19), 6915 (2022).
45. Zhang, L. *et al.* An optimal dispatch model for virtual power plant that incorporates carbon trading and green certificate trading. *Int. J. Electr. Power Energy Syst.* **144**, 108558 (2023).
46. Feng, Tian-tian *et al.* Induction mechanism and optimization of tradable green certificates and carbon emission trading acting on electricity market in China. *Resour. Conserv. Recycl.* **169**, 105487 (2021).
47. Jiang, C. *et al.* Aggregated impact of allowance allocation and power dispatching on emission reduction. *J. Mod. Power Syst. Clean Energy* **5**(6), 936–946 (2017).

48. Xue, Yang *et al.* Blockchain-based joint incentive mechanism for tradable green certificate and carbon trading market. *Electr. Power Constr.* **43**(6), 24–33 (2022).
49. Schusser, Sandra & Jaraitė, J. ūratė. Explaining the interplay of three markets: Green certificates, carbon emissions and electricity. *Energy Econ.* **71**, 1–13 (2018).
50. Jiang, Chao & Yue, Yunliang. Sensitivity analysis of key factors influencing carbon prices under the EU ETS. *Polish J. Environ. Stud.* **30**(4), 3645–3658 (2021).
51. Junxiang, Li. *et al.* Blockchain-based dynamic game of electricity price and power for microgrid electricity market. *Autom. Electr. Power Syst.* **45**(17), 11–19 (2021).
52. Xu, S. Q. *et al.* Optimal pricing decision of tradable green certificate for renewable energy power based on carbon-electricity coupling. *J. Clean. Prod.* **410**, 137111 (2023).
53. Zeng, L. J. *et al.* An inter-provincial tradable green certificate futures trading model under renewable portfolio standard policy. *Energy* **257**, 124772 (2022).
54. Xie, H. P. *et al.* Incremental green certificate towards flexibility incentive for renewable dominated power systems. *J. Clean. Prod.* **377**, 134345 (2022).
55. Zhang, F. *et al.* Decision-making behavior of power suppliers in the green certificate market: A system dynamics analysis. *Energy Policy* **171**, 113296 (2022).
56. Xu, Y. B. *et al.* Assessing the effects of tradable green certificates and renewable portfolio standards through demand-side decision-making simulation: A case of a system containing photovoltaic power. *Energies* **16**(8), 3517 (2023).
57. Park, H. *et al.* An intelligent financial portfolio trading strategy using deep Q-learning. *Expert Syst. Appl.* **158**, 113573 (2020).
58. Yang, L. X. *et al.* Nash Q-learning based equilibrium transfer for integrated energy management game with We-Energy. *Neurocomputing* **396**, 216–223 (2020).

## Acknowledgements

The authors express their gratitude to the anonymous reviewers for their valuable comments. The authors also thank the editors for their rigorous scientific work and assistance.

## Author contributions

Q.H., J.W., and F.H. designed the study, analyzed data, and wrote the manuscript. R.S. and M.W. revised the manuscript. All authors have read and approved the version of the manuscript to be published.

## Funding

This study is mainly based on the science and technology project of State Grid Corporation of China (SGCC) (Grant Number: 1400-202272230A-1-1-ZN).

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to Q.H., J.W. or Y.H.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024