processes

MDPI

Article

Low-Carbon and Economic-Oriented Dispatch Method for Multi-Microgrid Considering Green Certificate: Carbon Trading Mechanism Driven by AI Reinforcement Learning-Enhanced Genetic Algorithm

Yiqiao Cheng [1,*], Hongbo Zou [2,*] and Fei Wang [3]

[1] School of Computer Science and Engineering, University of New South Wales, Sydney, NSW 2052, Australia
[2] School of Electricity and New Energy, Three Gorges University, Yichang 443002, China
[3] School of Electrical and Automation, Wuhan University, Wuhan 430072, China; wrjiaofei008@163.com
* Correspondence: chengyiqiao8085@163.com (Y.C.); zhbhorace@ctgu.edu.cn (H.Z.)

Abstract

Aiming at the problem that the existing research mostly focuses on a single microgrid or an independent optimization goal and lacks the cooperative scheduling of multi-microgrids and the deep integration with the green certificate (GC) and carbon trading (CT) mechanisms, this paper proposes a low-carbon and economic-oriented dispatch method for multi-microgrids considering a GC-CT mechanism driven by an artificial intelligence (AI) reinforcement learning-enhanced genetic algorithm (GA). First of all, under the constructed architecture model of the GC-CT mechanism and multi-microgrid, this method constructs an optimal objective model that incorporates economic revenue and GC-CT costs. Secondly, regarding the two key parameters, crossover rate and mutation rate, which seriously influence the performance of the GA, this paper utilizes an AI reinforcement learning algorithm to adaptively adjust them and solves the constructed model based on the AI reinforcement learning-enhanced GA. Finally, based on a regional multi-microgrid system, the simulation results show that the proposed method can significantly improve the operating efficiency of the microgrid system after integrating the GC-CT mechanism into the microgrid system, which provides a theoretical framework and technical path for low-carbon and economic-oriented dispatch of multi-microgrids and helps the power system to evolve into a zero-carbon smart energy system.

Keywords: artificial intelligence; reinforcement learning algorithm; genetic algorithm; green certificate; low-carbon and economic optimal dispatch

check for updates

Academic Editor: Michael C. Georgiadis

Received: 19 July 2025
Revised: 7 August 2025
Accepted: 8 August 2025
Published: 11 August 2025

Citation: Cheng, Y.; Zou, H.; Wang, F. Low-Carbon and Economic-Oriented Dispatch Method for Multi-Microgrid Considering Green Certificate: Carbon Trading Mechanism Driven by AI Reinforcement Learning-Enhanced Genetic Algorithm. *Processes* **2025**, *13*, 2531. https://doi.org/10.3390/pr13082531

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Driven by the dual forces of the global climate crisis and energy transition, building a new-type power system dominated by new energy (NE) has become the core objective of national energy strategies [1–3]. As an important carrier of distributed energy systems, microgrids have emerged as a key unit for achieving low carbonization of the energy structure, owing to their flexible energy coupling characteristics and efficient localized consumption capabilities [4]. However, with the proliferation of microgrids and the increasing penetration rate of distributed energy, the independent operation mode of single microgrids has gradually exposed issues such as low resource utilization efficiency, poor system economy, and limited carbon emission (CE) reduction benefits [5–7]. Against this backdrop,

collaborative scheduling of multi-microgrids has emerged, which can significantly enhance NE consumption capacity and reduce system operation costs through cross-regional energy sharing and optimal allocation. Nevertheless, how to further integrate low-carbon economic incentive policies has become the current research focus [8–11].

The purpose of multi-microgrid collaborative scheduling is to achieve energy complementarity and optimal allocation among various microgrids, reduce power interaction with the upper-level grid, and increase the penetration rate of NE [12–14]. Currently, multi-microgrid systems are mainly classified into centralized and distributed control modes based on their control methodologies. In centralized control, the system management center uniformly schedules the interaction power of each microgrid, aiming to maximize overall benefits [15]. Distributed control, on the other hand, is based on consensus protocols, where each microgrid autonomously decides its power interaction plan to maximize its individual benefits. Depending on the optimization objectives, multi-microgrid collaborative scheduling includes types such as energy optimization, economic optimization, and environmental optimization. Energy optimization relies on predictive and real-time data to allocate energy through optimization algorithms. Economic optimization focuses on reducing electricity purchase costs and enhancing the operational revenue of microgrids, while environmental optimization considers reducing emissions, noise, and pollution to achieve green operation [16–18].

Global climate change and energy security crises are compelling countries to accelerate their energy transitions. GC and CT mechanisms, as market-based emission reduction tools, are increasingly gaining prominence [19–21]. GC employs market-based approaches to quantify the environmental worth of NE, allowing microgrids to obtain additional revenue through trading. CT, through cap-and-trade systems and emission rights pricing, compels microgrids to reduce their reliance on fossil fuels. As core tools of low-carbon economic policies, GCs and CT mechanisms provide a new optimization dimension for multi-microgrid collaborative scheduling. However, existing research primarily focuses on scheduling optimization under a single mechanism and has not fully explored the synergistic effects of GC and CT. For example, GC trading may influence microgrids' generation decisions, such as prioritizing the dispatch of NE, while CT costs must be integrated into the charging/discharging strategies of energy storage (ES) devices. The coupling relationships between the two in terms of time scales, spatial dimensions, and economic attributes require scheduling models to possess stronger dynamic adaptability and multi-objective balancing capabilities. Additionally, GC-CT mechanisms rely on high-precision predictive data, such as NE output and load demand, but actual scheduling must strictly adhere to physical constraints, such as power balance and energy storage lifespan. How to achieve the unity of data-driven decision-making and physical system security remains a challenge. The literature [21] points out that the volatility of GC prices and the differences in carbon quota allocation rules will lead to the need for scheduling models to handle multiple uncertainty factors simultaneously, making traditional deterministic optimization methods inadequate.

GA, as an intelligent algorithm that simulates natural selection and genetic mechanisms, searches for optimal or near-optimal solutions within the solution space through simulating the evolutionary process of biological populations [22–24]. The GA gradually optimizes population individuals through simulated selection, crossover, and mutation operations in biological evolution. However, the GA is sensitive to parameters such as crossover rate and mutation rate, faces slow convergence speed when dealing with high-dimensional complex problems, and has weak local search capabilities. The literature [22] proposes a power-dispatching decision-making assistance model based on an improved GA to address the issue that existing hybrid distribution network scheduling methods do not si-

multaneously consider carbon price fluctuations and wind power uncertainty. This method considers the impact of wind power uncertainty and power generation carbon emission on power balance in wind power integration systems, establishes power-dispatching cost and CE objective functions, and applies strategies such as decision space dimensionality reduction, population initialization, and gene operations to improve algorithm execution efficiency and model robustness. The literature [23] proposes a reactive power-forecasting technology for power systems based on an improved GA. The literature [24] presents a novel black-start weight determination method based on GA and score consistency. This method first compares commonly used similarity methods such as cosine similarity and mean square deviation similarity to select the optimal similarity method. It then assumes the index weights as a set of n-dimensional variables and calculates the scoring vector of the power black-start restoration plan. Representative black-start evaluation methods are selected, and the consistency between the scoring vector of the representative methods and the scoring vector of the proposed evaluation method is calculated using the selected optimal similarity method. A fitness function for maximizing the total similarity of black-start evaluation methods is constructed, and the GA is used to determine the optimal solution for index weights.

Reinforcement learning (RL) algorithm is a pivotal approach within the realm of machine learning in the field of artificial intelligence [25,26]. The fundamental concept underlying RL is to acquire an optimal policy via the iterative interaction between an intelligent agent and its surrounding environment, with the aim of maximizing cumulative rewards [27,28]. Unlike supervised learning, which relies on labeled data, and unsupervised learning, which relies on data structure, RL achieves autonomous learning through a "trial-and-error" mechanism and is widely applied in fields such as gaming, robot control, and autonomous driving [29,30]. The literature [30] proposes an online emergency control strategy based on hybrid distributed deep RL to maintain the dynamic stability performance of systems after a large disturbance. This method first models the transient stability emergency control problem as a Markov decision process. Then, to address issues such as the "curse of dimensionality" and decreased accuracy caused by discretizing the mixed action space in conventional deep RL algorithms, a discrete–continuous hybrid policy architecture is proposed, and the proximal policy optimization-based computational algorithm is adopted as the policy update method to directly handle the mixed action space in emergency control problems. Next, to overcome the drawbacks of conventional deep RL algorithms, such as long training times and insufficient robustness, a distributed parallel training architecture is introduced, and an illegal action-masking mechanism that incorporates prior physical knowledge of emergency control is designed, significantly improving the algorithm's training speed and robustness.

This paper proposes a low-carbon and economic-oriented dispatch method for multi-microgrids considering GC-CT mechanisms driven by an AI reinforcement learning-enhanced GA. Its innovations are reflected in the following aspects:

(1) A GC-CT mechanism and multi-microgrid integrated architecture model is constructed;
(2) The GC-CT costs are explicitly incorporated into the multi-microgrid dispatch objective function to establish an optimal dispatch model for microgrid operators that includes economic revenue, GC costs, and CT costs;
(3) For the two key parameters of crossover rate and mutation rate in the GA, an AI reinforcement learning algorithm is employed for their adaptive adjustment. The RL-enhanced GA is then used to solve the constructed optimal dispatch objective model for microgrids, and simulation verification is conducted.

2. GC-CT Mechanism and Multi-Microgrid Integrated Architecture Model

2.1. GC Trading Model

In the GC trading market, power generators specify the quantity of GC to buy/sell in the day-ahead market to offset their own CE quotas or obtain corresponding benefits. This mechanism can effectively promote the consumption of NE and reduce CE. The market trading price model for GC constructed in this paper is as follows [19]:

$$\begin{cases} C_{i,green} = \rho_{green}(G_{i,green} - Q_{i,green}) \\ G_{i,green} = \beta_{green} \sum_{t=1}^{24} P_{i,t} \\ Q_{i,green} = \beta_{green} \sum_{t=1}^{24} P_{i,t,real} \end{cases} \tag{1}$$

where $C_{i,green}$, $G_{i,green}$, and $Q_{i,green}$ are the transaction cost, GC quota, and GC obtained by consuming NE of power producer i, respectively; ρ_{green} is the benchmark transaction price of GC; β_{green} is the quota coefficient of GC; $P_{i,t}$ is the load at time t of power producer i; and $P_{i,t,real}$ is the actual output of the NE at time t.

2.2. CT Model

In this paper, the initial free CE quota and CE of power producers are as follows [20]:

$$M_{i,green} = \sigma_p Q_{i,green} \tag{2}$$

$$M_{i,trade} = a_{ci}(Q_{i,green})^2 + b_{ci}Q_{i,green} \tag{3}$$

where $M_{i,green}$ is the CE quota of power producers i; $M_{i,trade}$ is the total amount of CT of power producer i; $Q_{i,green}$ is the bid-winning electricity quantity of power producer i; σ_p is the CE quota per unit power generation of power producers; and a_{ci} and b_{ci} are the CE coefficients of power producer i.

According to the allocated CE quota, this paper adopts a stepped CE transaction cost model, which has obvious advantages in the design of green card transaction mechanisms and can avoid the extrusion of "one size fits all" pricing on emerging projects. When the CE of power producers do not exceed the free CE, they can sell the remaining CE quotas and get subsidies. On the other hand, you need to pay an additional penalty for the ladder carbon quota. The more quotas are exceeded, the higher the CT price will be, as shown below:

$$C_{i,cet} = \begin{cases} -\tau(1+2w)(M_{i,green} - M_{i,trade} - v),\ M_{i,trade} \le M_{i,green} - v \\ -\tau vw - \tau(1+w)(M_{i,green} - M_{i,trade}),\ M_{i,green} - v < M_{i,trade} \le M_{i,green} \\ \tau(M_{i,trade} - M_{i,green})w,\ M_{i,green} < M_{i,trade} \le M_{i,green} + v \\ \tau[k + (k-1)w]v + \tau(1+kw)(M_{i,trade} - M_{i,green} - kv),\ M_{i,green} + kv \le M_{i,trade} \le M_{i,green} + (k+1)v \end{cases} \tag{4}$$

where $C_{i,cet}$ is the carbon transaction cost of producer i; τ is the price of CT; w is the CT reward coefficient; v is the CE interval; and k is a step series.

2.3. The Architecture Model of GC-CT Mechanism and Multi-Microgrid

Under the GC-CT mechanism, the strategies adopted by various power generators in the electricity market become more intricate. For high-carbon-emission units, solely relying on increasing power generation and bidding prices to gain more profits may lead to substantial CE costs and GC trading costs. For low-carbon-emission units, the key lies in how to reasonably allocate resources in the GC-CT market to further enhance

competitiveness. For NE power generators, participating in both the electricity market and the GC trading market not only directly raises their revenue ceiling but also indirectly increases their cleared electricity volume in the electricity market. Additionally, due to the presence of "carbon" as a common factor, the prices in the GC market and the CT market influence each other. It can thus be seen that the GC market, the CT market, and the electricity market are mutually coupled, with their coupling relationship illustrated in Figure 1.

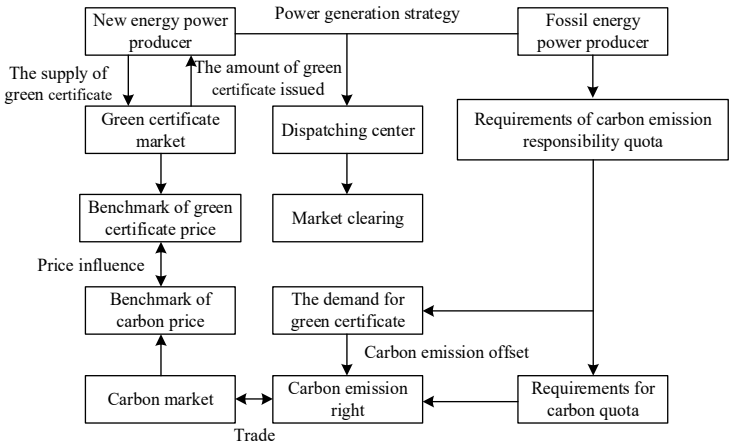


Figure 1. The coupling structure of GC-CT market.

Microgrids utilize NE sources for power generation, and they can also purchase gas from natural gas companies to supply electricity and heat through gas-fired equipment. During peak electricity demand periods, they buy electricity to meet load requirements. Since the actual CEs from purchasing electricity from the main grid (MG) are relatively high, while CT between microgrids is not counted towards the total actual CE, collaboration among multiple microgrids can further unlock their low-carbon potential. When microgrids operate cooperatively, they can engage in negotiated electricity trading with each other. Microgrids with surplus electricity can sell their excess power to other microgrids. Internal electricity transactions between microgrids are transmitted via power interconnection lines, and corresponding grid passage fees are paid to the MG based on the transaction volume. The GC-CT mechanism and multi-microgrid integrated architecture model is shown in Figure 2.

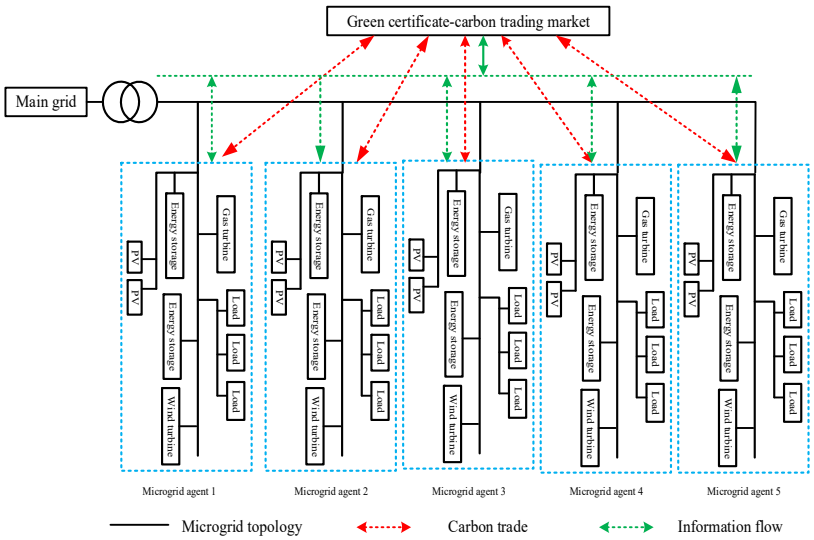


Figure 2. The GC-CT mechanism and multi-microgrid integrated architecture model.

3. The Multi-Microgrid Dispatching Model Considering GC-CT Mechanism

3.1. The Objective Function

In this paper, the goal of multi-microgrid cooperative operation is to maximize the microgrid revenue on the basis of considering its own operating cost and energy sales revenue; the objective function (OF) is as follows:

$$\max F = \sum_{i=1}^{M} F_i \tag{5}$$

$$F_i = F_{i,sell} - C_{i,energy} - C_{i,grid} - C_{i,trade} - C_{i,co_2} \tag{6}$$

where F is the total revenue; F_i refers to the sales revenue of microgrid operator i to its own users; M represents the number of microgrids; $C_{i,energy}$ is the energy production cost of gas turbine and new energy; $C_{i,grid}$ is the energy interaction cost with the MG; $C_{i,trade}$ is the network crossing fee generated by energy interaction among microgrids; and C_{i,co_2} is the cost of participating in the GC-CT market.

$$\begin{cases} F_{i,sell} = \sum\limits_{t=1}^{24} P_{i,t,sale}\rho_{sale} \\ C_{i,energy} = \sum\limits_{t=1}^{24} \left[\begin{array}{l} w_{1,gas}(P_{i,t,gas})^2 + w_{2,gas}P_{i,t,gas} + w_{3,gas} + \\ w_{1,new}(P_{i,t,new})^2 + w_{2,new}P_{i,t,new} + w_{3,new} \end{array} \right] \\ C_{i,grid} = \sum\limits_{t=1}^{24} P_{i,t,g2m}\rho_{g2m} - \sum\limits_{t=1}^{24} P_{i,t,m2g}\rho_{m2g} \\ C_{i,trade} = \frac{1}{2}\sum\limits_{t=1}^{24}\sum\limits_{j=1,i\neq j}^{M}\left[\gamma_1(P_{ij,t})^2 + \gamma_2 P_{ij,t}\right] \\ C_{i,co_2} = C_{i,green} + C_{i,cet} \end{cases} \tag{7}$$

where $P_{i,t,sale}$ represents the user's electric load supplied by microgrid operator i at time t; ρ_{sale} represents the electricity price of the multi-microgrid system at time t, which is determined by the microgrid operators through consultation; $w_{1,gas}$, $w_{2,gas}$, $w_{3,gas}$ and $w_{1,new}$, $w_{2,new}$, $w_{3,new}$ are the cost coefficients of gas turbine and new energy power generation, respectively; $P_{i,t,gas}$ and $P_{i,t,new}$ are the power of gas turbine and new energy at time t, respectively; $P_{i,t,g2m}$ is the electric power purchased by the microgrid operator i from the MG at time t; $P_{i,t,m2g}$ is the electric power that the microgrid operator i sends back to the MG at time t; ρ_{g2m} represents the electricity price of the MG at time t; ρ_{m2g} represents the repurchase price of the MG at time t; $P_{ij,t}$ represents the electric power supplied by the microgrid operator i to the microgrid operator j at time t, with a positive value indicating sale and a negative value indicating purchase; and γ_1 and γ_2 are the conversion coefficients of network fees.

3.2. The Constraints

The power balance constraint of the microgrid system is as follows:

$$P_{i,t,gas} + P_{i,t,new} + P_{i,t,grid} = P_{i,t,sale} + P_{ij,t} \tag{8}$$

Equipment operation constraints are as follows:

$$\begin{cases} P_{i,gas}^{\min} \le P_{i,t,gas} \le P_{i,gas}^{\max} \\ R_{i,gas}^{down} \le P_{i,t,gas} - P_{i,t-1,gas} \le R_{i,gas}^{up} \\ 0 \le P_{i,t,new} \le P_{i,t,new}^{pre} \end{cases} \tag{9}$$

where $P_{i,gas}^{\max}$ and $P_{i,gas}^{\min}$ are the superior and inferior bounds of the gas turbine output power; $R_{i,gas}^{up}$ and $R_{i,gas}^{down}$ are the superior and inferior bounds of climbing power of gas turbine; $P_{i,t-1,gas}$ represents the power of gas turbine at time $t-1$; and $P_{i,t,new}^{pre}$ is the predicted value of new energy power generation.

The constraint of interconnection lines between microgrids is as follows:

$$\begin{cases} 0 \le \left|P_{ij,t}\right| \le P_{ij,\max} \\ \sum_{i=1}^{M} \sum_{j=1,j\neq i}^{M} P_{ij,t} = 0 \end{cases} \tag{10}$$

where $P_{ij,max}$ is the upper transmission limit of the internal electric energy transaction between microgrid i and j. Since all microgrids cooperate to realize the electric energy sharing, the overall electric energy transaction volume within the system is zero.

External connection line constraint with MG is as follows:

$$\begin{cases} 0 \le P_{i,t,g2m} \le P_{grid,\max} \\ 0 \le P_{i,t,m2g} \le P_{grid,\max} \end{cases} \tag{11}$$

where $P_{grid,\max}$ is the superior bound of the transmission power of the external connection line with the MG.

4. Low-Carbon and Economic-Oriented Dispatch Method Based on AI Reinforcement Learning-Enhanced GA

4.1. The GA

The core idea of GA is based on Darwin's natural selection and genetic mechanism. By simulating population evolution (selection, crossover, and mutation), it is suitable for complex nonlinear, multimodal, discrete, or continuous space optimization problems.

The main parameters of the GA solution include population size N_p, crossover probability P_c, mutation probability P_m, and genetic termination evolution algebra $K_{\max}$ (let K represent evolutionary algebra). Population size N_p has a major impact on the final result of genetic optimization and how efficiently the GA runs. P_c mainly controls how often crossover operations are applied. P_m is an auxiliary search way that helps keep the population diverse. Genetic termination evolutionary generation $K_{\max}$ is a parameter for the end condition. The algorithm stops when it reaches the specified generation and outputs the best individual in the current population as the problem's optimal solution.

The procedural steps for the execution of the GA are outlined as follows:

(1) Population Initialization: A set of valid candidate solutions (individuals) are randomly generated. Since the GA represents each individual by using a chromosome, the initial population is essentially a group of chromosomes.

(2) Fitness Calculation: The fitness value (OF value) of each individual is calculated. For the initial population, this operation is performed once. Following the application of genetic operators, namely selection, crossover, and mutation, this process is iteratively executed for each successive new generation.

(3) Selection, Crossover, and Mutation: By applying the genetic operators of selection, crossover, and mutation to the population, a new generation is produced. The selection procedure serves the crucial function of identifying and picking out individuals with favorable traits from the existing population. The crossover mechanism generates progeny from the chosen individuals. Typically, it achieves this by facilitating the exchange of a segment of chromosomes between two pre-selected individuals, thereby producing two novel chromosomes that symbolize the offspring. The mu-

tation process, on the other hand, randomly modifies one or multiple chromosome values (genes) within each newly formed individual.

(4) Algorithm Termination Condition: The algorithm stops when either the upper-bound on the count of iterative executions is reached or the fitness values converge. The optimal solution is then output.

The flow chart of GA is shown in Figure 3.

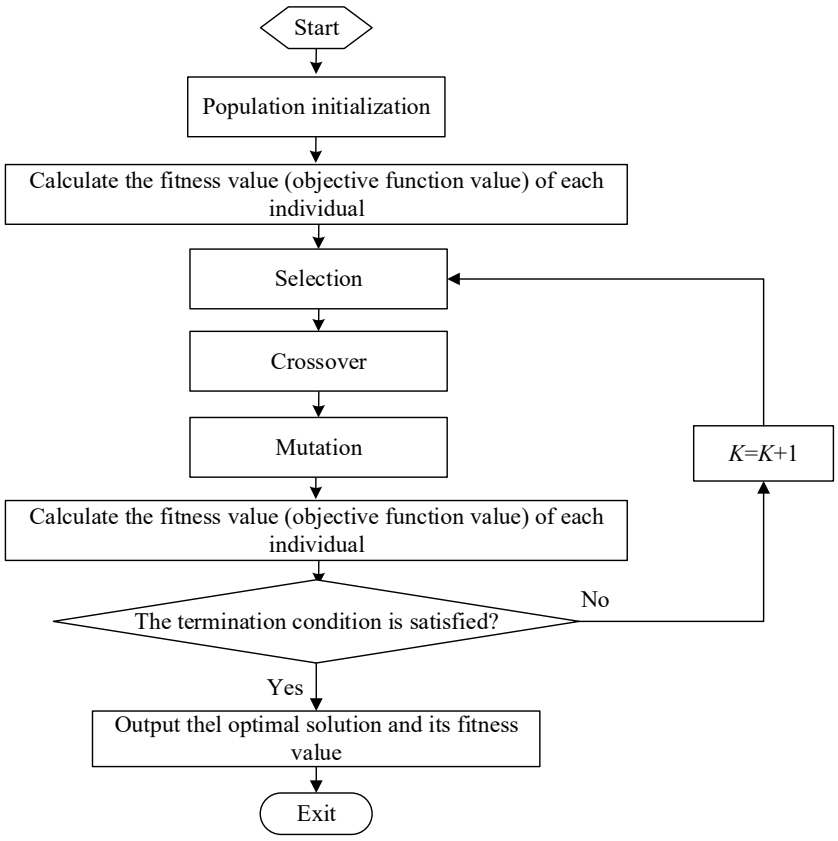


Figure 3. The flow chart of GA.

4.2. AI Reinforcement Learning-Enhanced GA

In GA, P_c and P_m are key parameters that affect performance. If the P_c is too large, the population diversity will decrease rapidly, the algorithm may fall into local optimization, and frequent crossover may destroy the existing high-quality gene combination in the parent and reduce the search efficiency. However, if the P_c is too small, the population cannot effectively combine excellent genes, the convergence speed is too slow, and the optimization efficiency is low, the algorithm will degenerate into relying on random mutation and lose the advantages of the GA. If the P_m is too large, the algorithm will degenerate into random search, lose the ability to use high-quality genes, and frequent mutation will make it difficult to retain excellent genes and have poor convergence. If the P_m is too small, the population genes will tend to be homogeneous and cannot jump out of the local optimum, and the algorithm may fall into sub-optimal solutions in the early stage and cannot be further optimized. Therefore, the P_c and P_m should not be too large or too small, and an appropriate value should be selected. Therefore, based on the AI reinforcement learning algorithm, this paper adaptively adjusts the two parameters.

The Markov property signifies that the state at the present instant is solely contingent upon the state and action at the immediately preceding moment. When $h_t = \{s_1, s_2, \ldots, s_t\}$(all states before time t are included in h_t), it can be expressed as follows:

$$p(s_{t+1}|s_t, a_t) = p(s_{t+1}|h_t, a_t) \tag{12}$$

where h_t represents the set of states before time t; s_t and s_{t+1} represent the state at time t and $t + 1$; a_t represents the action at time t; and p is the Markov property.

In the event that the state transition within a multi-stage decision-making problem adheres to Equation (12), it can be regarded as a Markov decision-making problem, as shown in Equation (13):

$$\psi = \{S, A(s), R, p, \pi(a|s)\} \tag{13}$$

where ψ represents the Markov decision-making problem; S represents state space; s represents state; $A(s)$ is the optional action space when the state is s; R represents agent-related reward variable, which specifically represents the feedback in the form of a reward that the environment provides subsequent to the agent executing an action; and $\pi(a|s)$ is the probability distribution of the output action a when the state is s.

In this paper, the actions mainly include P_c value and P_m value. The setting of reward function is related to the selection of P_c and P_m. When the new generation has higher fitness than the previous generation, it can be rewarded. The reward calculation formulas of optimal and average individual fitness are, respectively, as follows:

$$r_c = \frac{\max f(x_i^k) - \max f(x_i^{k-1})}{\max f(x_i^{k-1})} \tag{14}$$

$$r_m = \frac{\sum_{i=1}^{N} f(x_i^k) - \sum_{i=1}^{N} f(x_i^{k-1})}{\sum_{i=1}^{N} f(x_i^{k-1})} \tag{15}$$

where r_c and r_m represent the reward calculation formulas of optimal and average individual fitness, respectively; $f(x_i^k)$ is the fitness of x_i^k; x_i^k is the i-th individual in the k generation; $f(x_i^{k-1})$ is the fitness of x_i^{k-1}; x_i^{k-1} is the i-th individual in the k-1 generation; and N is the population size.

The RL algorithm uses neural networks to fit the state-action value function Q, which is shown in the following formula:

$$Q_{k+1} = r_k + \gamma Q(s_k, \text{argmax}Q(s_k, a_k, \theta)) \tag{16}$$

where r_k represents the reward in the k generation; γ represents the discount factor; θ represents the parameters of the target network; $\text{argmax}Q(s_t, a_t, \theta)$ represents the maximum action of the Q value; Q_{k+1} represents the Q value in the k+1 generation; s_k represents the state in the k generation; and a_k represents the action in the k generation.

The algorithm training steps are as follows:

(1) Set the population size N of the GA and the maximum iterations $K_{\max}$. Let the initial iteration $k = 1$. Initialize the experience replay buffer D for the RL algorithm and randomly initialize the action-value function Q.
(2) Calculate the state value s_k.
(3) Select an action a_k with a random probability ε.
(4) Observe the next state s_{k+1} and the reward r_k.
(5) Store the tuple (s_k, a_k, r_k, s_{k+1}) in the experience replay buffer D.

(6) Conduct a stochastic sampling operation to extract a batch of data instances from the experience replay buffer.
(7) Calculate the Q_{k+1} value based on Equation (16).
(8) Obtain the optimal parameters P_c and P_m, which are trained by the RL algorithm.
(9) Proceed to the GA computation.
(10) When $k = K_{max}$, the iterative computation ends, and the optimal result is output.

5. Numerical Test and Analysis

5.1. Basic Data and Simulation Conditions

In this paper, a microgrid group system consisting of five microgrids is taken as an example for simulation and verification. The schematic diagram of the system structure is shown in Figure 2. The upper transmission limit of power tie line in microgrid is 1500 kW, the cost coefficient of network access fee is 0.01 RMB/kWh, the purchase price of natural gas is 2.0 RMB/m^3, and the CE quota coefficient is 0.55 kg/kWh. The quota coefficient of NE is 0.2, the basic price of reward and punishment ladder carbon transaction is 0.25 RMB/kg, the growth rate pertaining to the costs associated with rewards and penalties is 25%, and the width of CE ladder interval is 300 kg.

To verify the effectiveness of the proposed method in this paper, the following scenarios are set up for comparative analysis:

Scenario 1: Without considering the GC-CT mechanism, each microgrid operates independently and only exchanges electricity with the MG [16].

Scenario 2: Without considering the GC-CT mechanism, each microgrid operates cooperatively and only exchanges electricity with the MG [18].

Scenario 3: Considering the GC-CT mechanism, each microgrid operates independently and only exchanges electricity with the MG.

Scenario 4: Considering the GC-CT mechanism, each microgrid operates cooperatively and only exchanges electricity with the MG, i.e., the method proposed in this paper.

5.2. Simulation Results and Analysis

The results of microgrid operation under each scenario are shown in Table 1.

As indicated in Table 1, under scenario 1, the total revenue is 39,956.6 RMB, and the energy production cost is 23,678.7 RMB. Under scenario 2, the total revenue is 41,159.4 RMB, and the energy production cost is 22,488.1 RMB. It can thus be seen that, when neither scenario considers the GC-CT mechanism, the mutual cooperation and operation among micro-energy grids yield higher revenues and lower energy production costs. Therefore, cooperative operation is conducive to the healthy development of the entire micro-energy grid system. However, since both scenario 1 and scenario 2 do not take the GC-CT mechanism into account, there are neither carbon gains nor carbon costs. Nevertheless, these two scenarios overlook environmental protection factors. Under strict CE control, these two scenarios will face limitations. In scenario 3, the total revenue is 39,006.9 RMB, the energy production cost is 23,906.8 RMB, and the CT cost is 566 RMB. In scenario 4 (the method proposed in this paper), after the joint operation of the multi-microgrid system, additional revenue is generated through CT, resulting in an increased total revenue for the system, with a total revenue of 42,582.2 RMB, an energy production cost of 22,661.3 RMB, and a CT cost of −1,276 RMB. Therefore, in scenario 4, compared with scenario 1 in the literature [16] and scenario 2 in the literature [18], the total revenue of the proposed method is increased by 6.57% and 3.46%, respectively. It can be inferred that scenario 3, which considers the GC-CT mechanism, incurs additional CE costs, leading to a decrease in the system's overall revenue compared to scenario 1. In contrast, scenario 4 enables the multi-microgrid system to obtain extra revenue through CT after joint operation. Consequently, in a joint

multi-microgrid operation system, by considering the GC-CT mechanism and optimizing the operation of the combined micro-energy grid system, not only are system costs not increased, but additional revenue is generated instead.

Table 1. The results of microgrid operation under each scenario.

Scenario	Microgrid Agent	Total Revenue/RMB	Energy Production Cost/RMB	Carbon Transaction Cost/RMB
1	1	8261.5	4735.5	-
	2	8077.4	4821.8	-
	3	7969.8	4688.9	-
	4	7765.6	4759.3	-
	5	7882.3	4673.2	-
	Sum	39,956.6	23,678.7	-
2	1	8769.4	4574.2	-
	2	8355.7	4628.1	-
	3	8168.3	4419.3	-
	4	7804.8	4527.6	-
	5	8061.2	4338.9	-
	Sum	41,159.4	22,488.1	-
3	1	8081.3	4877.2	367
	2	7726.4	4781.4	284
	3	7625.7	4733.7	−315
	4	7721.2	4825.9	−218
	5	7852.3	4688.6	448
	Sum	39,006.9	23,906.8	566
4	1	9094.2	4611.5	−216
	2	8683.7	4572.3	−353
	3	8411.2	4438.2	−287
	4	8077.4	4653.7	−312
	5	8315.7	4385.6	−108
	Sum	42,582.2	22,661.3	−1276

In scenario 4, under the GC-CT mechanism, the results of electric energy trading between microgrids in cooperative operation are shown in Figure 4, wherein positive numerical values signify the sale of electricity, while negative numerical values denote the procurement of electricity. As shown in Figure 4, in the joint operation of a multi-microgrid system, microgrids exchange and share electricity with each other through tie-lines, thereby reducing the electricity interaction with the MG. This approach helps to decrease the overall energy consumption costs and CE of the system.

Figures 5 and 6 show the total CE of the micro-energy grid system and the consumption of new energy under various scenarios. As illustrated in Figure 5, when neither scenario 1 nor scenario 2 considers the GC-CT mechanism, the overall CEs of the system are significantly high. However, after incorporating the GC-CT mechanism into scenario 3 and scenario 4 by including CT costs in the OF, the CE of the system are notably reduced. Additionally, compared to scenario 1 and scenario 3, the overall CE of scenario 2 and scenario 4 are lower after the joint operation of the multi-microgrid system. Notably, scenario 4, which employs the method proposed in this paper, achieves the lowest overall CE, further demonstrating the superiority of the proposed method.

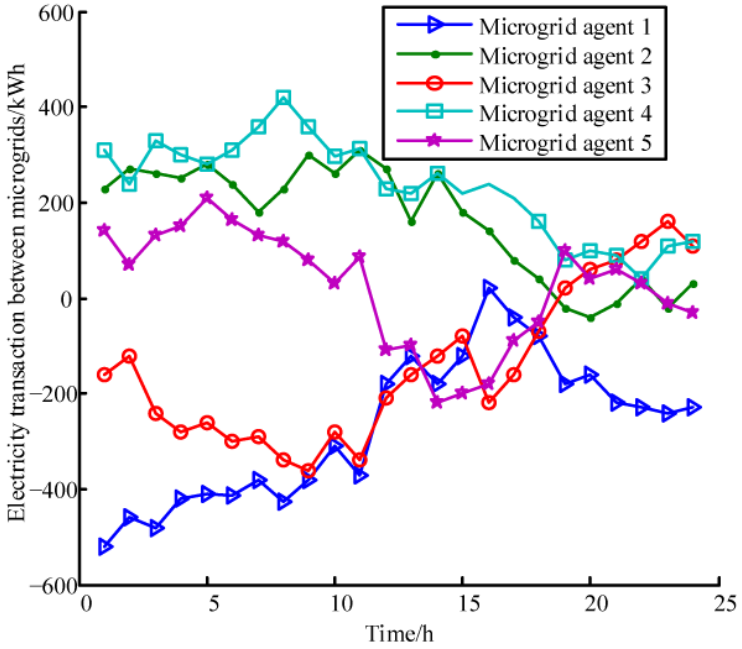


Figure 4. The energy trading results between microgrids.

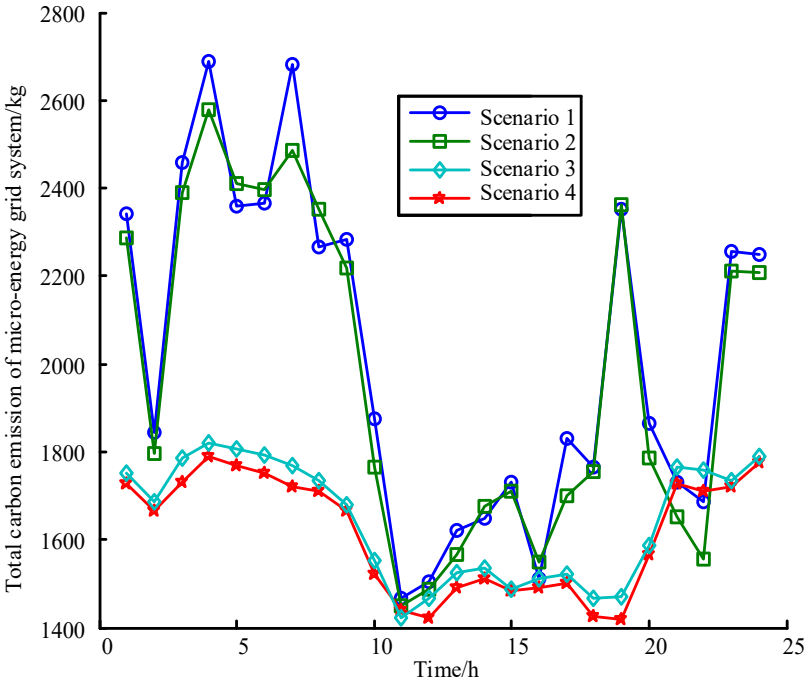


Figure 5. The total CE of micro-grid system under various scenarios.

As can be seen from Figure 6, in scenario 1 and scenario 3, the micro-grid operators operate independently. During certain periods of high load, they are unable to meet the energy supply demands and have to purchase high-carbon-emission energy from the main network. Conversely, during periods of low load, they fail to fully utilize the carbon-free new energy sources, resulting in a low new energy consumption rate and ineffective utilization. In scenario 2, although the operators operate cooperatively, achieving complementary advantages among themselves, the GC-CT mechanism is not taken into account. In contrast, the proposed method in this paper (i.e., scenario 4) involves cooperative operation among micro-grids while considering the GC-CT mechanism. This not only significantly reduces

the amount of energy purchased from the upper-level grid but also enables the effective consumption of new energy sources.

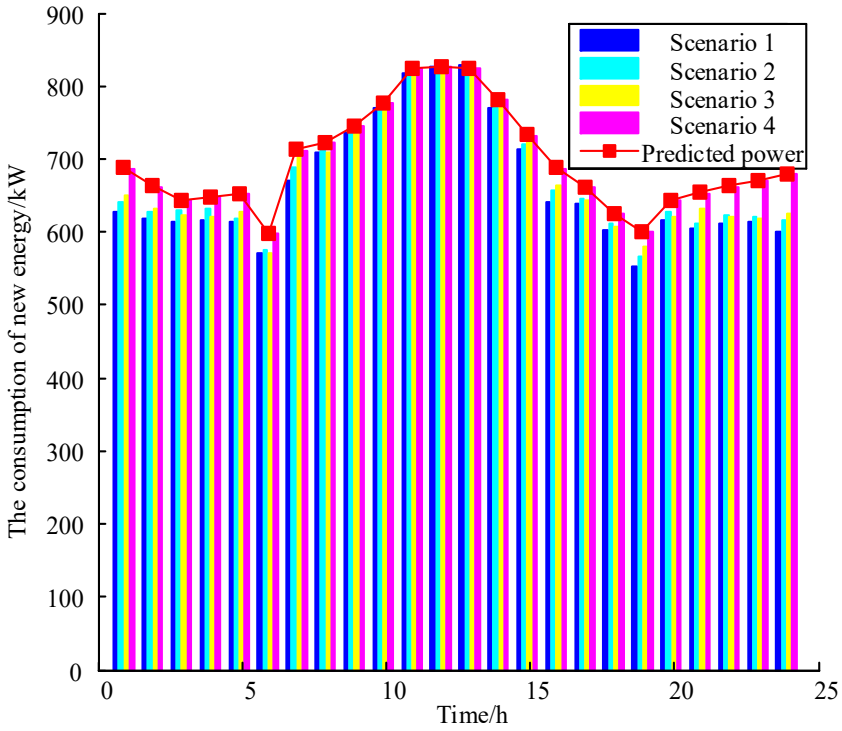


Figure 6. The consumption of new energy under various scenarios.

6. Conclusions

Aiming at the problem that the existing research mostly focuses on a single microgrid or an independent optimization goal and lacks the cooperative scheduling of multi-microgrids and the deep integration with the GC and CT mechanisms, a low-carbon and economic-oriented dispatch method for multi-microgrids considering GC-CT mechanisms driven by AI reinforcement learning-enhanced GA is proposed in this paper. Based on the constructed GC-CT mechanism and multi-microgrid integrated architecture model, a multi-microgrid dispatch objective function that includes economic revenue, GC costs, and CT costs is proposed. The RL-enhanced GA is then used to solve the constructed optimal dispatch objective model. The simulation results of a regional multi-microgrid system show that the proposed method can significantly improve the operating efficiency of the microgrid system and promote the consumption of new energy. The proposed method can provide a theoretical framework and technical path for low-carbon economic dispatch of multi-microgrids and help the power system to evolve into a zero-carbon smart energy system. In the future work, we will deeply discuss the key external factors, such as communication infrastructure, data availability, and regulatory requirements, so as to better promote the practical application of the method proposed in this paper.

Author Contributions: Conceptualization, Y.C., H.Z., F.W.; methodology, Y.C., H.Z., F.W.; software, Y.C., H.Z., F.W.; writing—original draft, Y.C., H.Z., F.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding authors.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Lin, F.J.; Liao, J.C.; Zhang, Y.M.; Huang, Y.C. Optimal Economic Dispatch and Power Generation for Microgrid Using Novel Lagrange Multipliers-Based Method with HIL Verification. *IEEE Syst. J.* **2023**, *17*, 4533–4544. [CrossRef]
2. Salman, M.; Li, Y.; Xiang, J. A Distributed Consensus-Based Optimal Dispatch Control Strategy for Hybrid AC/DC Microgrids. *IEEE Access* **2024**, *12*, 90997–91010. [CrossRef]
3. Sun, L.; Ding, D.; Dong, H.; Yi, X. Distributed Economic Dispatch of Microgrids Based on ADMM Algorithms with Encryption-Decryption Rules. *IEEE Trans. Autom. Sci. Eng.* **2025**, *22*, 8427–8438. [CrossRef]
4. Liu, D.; Jiang, K.; Yan, L.; Ji, X.; Cao, K.; Xiong, P. A Fully Distributed Economic Dispatch Method in DC Microgrid Based on Consensus Algorithm. *IEEE Access* **2022**, *10*, 119345–119356. [CrossRef]
5. Teng, F.; Ban, Z.; Li, T.; Sun, Q.; Li, Y. A Privacy-Preserving Distributed Economic Dispatch Method for Integrated Port Microgrid and Computing Power Network. *IEEE Trans. Ind. Inform.* **2024**, *20*, 10103–10112. [CrossRef]
6. Sun, J.; Hu, C.; Liu, L.; Zhao, B.; Liu, J.; Shi, J. Two-stage Correction Strategy-based Real-time Dispatch for Economic Operation of Microgrids. *Chin. J. Electr. Eng.* **2022**, *8*, 42–51. [CrossRef]
7. Li, Z.; Cheng, Z.; Si, J.; Li, S. Distributed Event-Triggered Hierarchical Control to Improve Economic Operation of Hybrid AC/DC Microgrids. *IEEE Trans. Power Syst.* **2022**, *37*, 3653–3668. [CrossRef]
8. Zheng, X.; Li, Q.; Yuan, J.; Chen, Z. Distributed Weighted Gradient Descent Method with Adaptive Step Sizes for Energy Management of Microgrids. *IEEE Trans. Smart Grid* **2024**, *15*, 4436–4449. [CrossRef]
9. Martinez-Gomez, M.; Orchard, M.E.; Bozhko, S. Dynamic Average Consensus with Anti-Windup Applied to Interlinking Converters in AC/DC Microgrids Under Economic Dispatch and Delays. *IEEE Trans. Smart Grid* **2023**, *14*, 4137–4140. [CrossRef]
10. Dou, Y.; Chi, M.; Liu, Z.-W.; Wen, G.; Sun, Q. Distributed Secondary Control for Voltage Regulation and Optimal Power Sharing in DC Microgrids. *IEEE Trans. Control. Syst. Technol.* **2022**, *30*, 2561–2572. [CrossRef]
11. Sun, B.; Jing, R.; Ge, L.; Zeng, Y.; Dong, S.; Hou, L. Quick Hosting Capacity Evaluation Based on Distributed Dispatching for Smart Distribution Network Planning with Distributed Generation. *J. Mod. Power Syst. Clean Energy* **2024**, *12*, 128–140. [CrossRef]
12. Liu, C.; Zhang, H.; Shahidehpour, M.; Zhou, Q.; Ding, T. A Two-Layer Model for Microgrid Real-Time Scheduling Using Approximate Future Cost Function. *IEEE Trans. Power Syst.* **2022**, *37*, 1264–1273. [CrossRef]
13. Wang, B.; Zhang, C.; Li, C.; Li, P.; Dong, Z.Y.; Lu, J. Hybrid Interval-Robust Adaptive Battery Energy Storage System Dispatch with SoC Interval Management for Unbalanced Microgrids. *IEEE Trans. Sustain. Energy* **2022**, *13*, 44–55. [CrossRef]
14. Zhang, Z.; Wang, C.; Chen, S.; Zhao, Y.; Dong, X.; Han, X. Multitime Scale Co-Optimized Dispatch for Integrated Electricity and Natural Gas System Considering Bidirectional Interactions and Renewable Uncertainties. *IEEE Trans. Ind. Appl.* **2022**, *58*, 5317–5327. [CrossRef]
15. Jiao, F.; Zou, Y.; Zhang, X.; Zhang, B. A Three-Stage Multitimescale Framework for Online Dispatch in a Microgrid with EVs and Renewable Energy. *IEEE Trans. Transp. Electrif.* **2022**, *8*, 442–454. [CrossRef]
16. An, R.; Liu, J.; Liu, Z.; Song, Z. Flexible Transfer Converters Enabling Autonomous Control and Power Dispatch of Microgrids. *IEEE Trans. Power Electron.* **2022**, *37*, 13767–13781. [CrossRef]
17. Ren, Z.; Qu, X.; Wang, M.; Zou, C. Multi-Objective Optimization for DC Microgrid Using Combination of NSGA-II Algorithm and Linear Search Method. *IEEE J. Emerg. Sel. Top. Circuits Syst.* **2023**, *13*, 789–796. [CrossRef]
18. Bai, C.; Li, Q.; Zheng, X.; Yin, X.; Tan, Y. Dynamic Weighted-Gradient Descent Method with Smoothing Momentum for Distributed Energy Management of Multi-Microgrids Systems. *IEEE Trans. Smart Grid* **2023**, *14*, 4152–4168. [CrossRef]
19. Zhang, S.; He, F.; Li, B. Research on the Optimal Scheduling of Multi-Microgrid Double-Layer Game Considering Fair Carbon Trading Strategy in the Green Certificate Trading Market. *IEEE Access* **2024**, *12*, 161620–161636. [CrossRef]
20. Danish, S.M.; Zhang, K.; Amara, F.; Cepeda, J.C.O.; Vasquez, L.F.R.; Marynowski, T. Blockchain for Energy Credits and Certificates: A Comprehensive Review. *IEEE Trans. Sustain. Comput.* **2024**, *9*, 727–739. [CrossRef]
21. Ma, T.; Pei, W.; Xiao, H.; Yang, Y.; Ma, L. A Joint Power and Renewable Energy Certificate Trading Method in the Peer-to-Peer Market. *IEEE Trans. Smart Grid* **2025**, *16*, 1604–1618. [CrossRef]
22. Behbahani, S.; de Silva, C.W. Mechatronic Design Evolution Using Bond Graphs and Hybrid Genetic Algorithm with Genetic Programming. *IEEE/ASME Trans. Mechatron.* **2013**, *18*, 190–199. [CrossRef]
23. Wei, H.; Tang, X.-S. A Genetic-Algorithm-Based Explicit Description of Object Contour and its Ability to Facilitate Recognition. *IEEE Trans. Cybern.* **2015**, *45*, 2558–2571. [CrossRef] [PubMed]
24. Raj, B.; Ahmedy, I.; Idris, M.Y.I.; Noor, R.M. A Hybrid Sperm Swarm Optimization and Genetic Algorithm for Unimodal and Multimodal Optimization Problems. *IEEE Access* **2022**, *10*, 109580–109596. [CrossRef]
25. Singh, A.; Chiu, W.-Y.; Manoharan, S.H.; Romanov, A.M. Energy-Efficient Gait Optimization of Snake-Like Modular Robots by Using Multiobjective Reinforcement Learning and a Fuzzy Inference System. *IEEE Access* **2022**, *10*, 86624–86635. [CrossRef]
26. Abdulazeez, D.H.; Askar, S.K. Offloading Mechanisms Based on Reinforcement Learning and Deep Learning Algorithms in the Fog Computing Environment. *IEEE Access* **2023**, *11*, 12555–12586. [CrossRef]

27. Tan, X.; Qu, C.; Xiong, J.; Zhang, J.; Qiu, X.; Jin, Y. Model-Based Off-Policy Deep Reinforcement Learning with Model-Embedding. *IEEE Trans. Emerg. Top. Comput. Intell.* **2024**, *8*, 2974–2986. [CrossRef]
28. Rapetswa, K.; Cheng, L. Towards a multi-agent reinforcement learning approach for joint sensing and sharing in cognitive radio networks. *Intell. Converg. Netw.* **2023**, *4*, 50–75. [CrossRef]
29. Wang, X.; Wang, S.; Liang, X.; Zhao, D.; Huang, J.; Xu, X.; Dai, B.; Miao, Q. Deep Reinforcement Learning: A Survey. *IEEE Trans. Neural Netw. Learn. Syst.* **2024**, *35*, 5064–5078. [CrossRef]
30. Belyakov, B.; Sizykh, D. Adaptive Algorithm for Selecting the Optimal Trading Strategy Based on Reinforcement Learning for Managing a Hedge Fund. *IEEE Access* **2024**, *12*, 189047–189063. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.