

Rural Integrated Energy System Carbon Assessment and Carbon Reduction Potential Analysis Based on Deep Reinforcement Learning

Haoye Jiang

*Aberdeen Institute of Data Science and Artificial Intelligence, South China Normal University,
Foshan, China
jhy20060703@foxmail.com*

Abstract. With the advancement of China's "Dual Carbon" goals, the transformation of the rural energy structure urgently requires the support of scientific assessment tools. This paper proposes a carbon assessment framework for Rural Integrated Energy Systems (RIES) integrating deep reinforcement learning. First, a dynamic model of RIES equipment and a carbon flow tracking mechanism are constructed to achieve accurate measurement and real-time tracking of carbon emissions from various energy equipment within the system. Second, the Deep Deterministic Policy Gradient algorithm is employed to optimize the system scheduling strategy, and a novel SI-SAC algorithm based on a safety-imitation mechanism is proposed, effectively enhancing the algorithm's constraint satisfaction capability and learning efficiency. Finally, the carbon emission intensity under different energy configurations is quantified, and the potential of DRL in carbon reduction path optimization is explored. Case analysis shows that the DRL-driven scheduling strategy reduces system carbon emissions by approximately 10.4% and operating costs by approximately 8.3% compared to the traditional mixed integer linear programming method, confirming its effectiveness and economy in rural carbon reduction decision-making. The framework proposed in this paper provides a quantitative decision-making tool for the low-carbon, economical, and secure operation of rural energy systems, and holds significant theoretical and practical value for advancing the scientific implementation of the "Dual Carbon" goals in rural areas.

Keywords: Rural integrated energy system, Carbon assessment, Deep reinforcement learning, SI-SAC algorithm, Carbon reduction potential

1. Introduction

The rural energy system is an important component of China's energy system. However, its decentralized characteristics, high renewable energy penetration rate, and load volatility make it difficult for traditional static carbon assessment methods to accurately characterize its carbon emission features [1]. With the proposal of the "Dual Carbon" goals, the low-carbon transformation of rural energy systems has become an important issue that needs to be addressed urgently.

However, existing research mostly focuses on integrated energy systems in urban areas, lacking dynamic carbon flow analysis tools for rural multi-energy coupled systems. Accurate modeling of integrated energy systems is the foundation for optimal operation [2]; however, existing models are mostly designed for urban areas or industrial parks, and dedicated models that account for the specific characteristics of rural areas—such as high renewable energy penetration and loose multi-energy coupling—remain insufficient.

Currently, RIES carbon assessment mainly relies on traditional optimization methods such as Linear Programming (LP) or Mixed Integer Linear Programming (MILP). Although these methods can obtain global optimal solutions, they are difficult to adapt to high-dimensional nonlinear constraints and real-time requirements. Reinforcement Learning (RL), due to its adaptive capability in complex decision-making problems, has emerged in the field of energy scheduling in recent years. However, traditional RL algorithms suffer from high constraint violation rates and low training efficiency when dealing with complex constraints of integrated energy systems. Although deep reinforcement learning (DRL) has shown potential in the field of energy scheduling, previous studies have demonstrated that different DRL algorithms (e.g., DDPG, SAC) exhibit marked differences in performance and applicability when addressing complex system constraints [3] and have not yet been effectively applied to rural carbon assessment scenarios.

To address the above issues, this paper proposes for the first time a RIES carbon assessment framework integrating DRL, with the following main innovations: (1) A dynamic carbon flow tracking model is constructed to achieve the coupling of equipment operating states and real-time carbon emission factors; (2) A reward function centered on carbon emission costs is designed to deeply integrate carbon assessment into the DRL decision loop; (3) The SI-SAC algorithm is proposed, which ensures constraint satisfaction through a safety component and accelerates strategy learning through an imitation component. Although some scholars have begun to explore low-carbon scheduling issues for community integrated energy systems [4], systematic research on a dynamic carbon assessment and DRL integration framework specifically tailored for rural scenarios remains lacking. This paper aims to provide quantitative tools for rural energy planning and promote the implementation of the "Dual Carbon" goals in rural areas.

2. Construction of carbon assessment system for rural integrated energy systems

2.1. Objectives and significance of carbon assessment

A scientific and accurate carbon assessment system is the foundation and key to promoting the low-carbon transformation of rural energy systems. Given that RIES features diversified energy sources, strong load volatility, and high renewable energy penetration, traditional static and itemized carbon assessment methods are difficult to characterize its dynamic carbon emission features accurately. Therefore, this study constructs a dynamic carbon assessment system integrating deep reinforcement learning, aiming to achieve "observable, traceable, and optimizable" carbon emissions of RIES, providing quantitative tools and theoretical support for carbon flow tracking, emission reduction potential exploration, and low-carbon optimal scheduling.

2.2. Components of the carbon assessment system

This system consists of the following four core components:

(1) Definition of carbon emission accounting scope. Adopting the comprehensive accounting principle of "life cycle + operational boundary," the accounting scope covers carbon emissions from

all energy activities within the RIES boundary, including: direct emissions (carbon emissions from direct combustion of fossil fuels such as coal and natural gas), indirect emissions (carbon emissions embodied in purchased electricity), and quasi-zero/negative carbon emissions (net carbon emissions during the utilization of biogas, biomass, and other energy sources).

(2) Dynamic carbon flow tracking mechanism. Establishing a dynamic correlation model between equipment operating states and carbon emission factors is the core of achieving an accurate assessment. This mechanism collects output data from various power generation equipment in real-time (as shown in Equations (1)-(4)), and combines them with their corresponding carbon emission coefficients to construct a carbon flow map from the "energy supply side" to the "energy consumption side," achieving dynamic traceability and visualization of carbon elements in the multi-energy flows of electricity, heat, and gas. In terms of carbon flow tracking theory, existing studies have incorporated carbon emission peak constraints into the dynamic optimization problems of active distribution networks based on the Carbon Emission Flow (CEF) theory [5], which provides a theoretical foundation for the dynamic carbon flow modeling in this study.

(3) Carbon assessment indicator system. Multi-dimensional indicators are established to comprehensively measure the carbon performance of the system, mainly including carbon emission intensity per unit of energy supply (kgCO_2/kWh), total system carbon emissions (kgCO_2), renewable energy carbon reduction contribution rate, and carbon emission cost (CNY).

(4) DRL-based dynamic assessment and optimization framework. The accounting scope, tracking mechanism, and assessment indicators are embedded into the deep reinforcement learning environment to form a closed-loop "assessment-optimization" framework. This framework elevates carbon assessment from "post-event accounting" to "in-process guidance," where the DRL agent continuously learns low-carbon strategies during training and directly outputs optimal scheduling plans.

2.3. Integration mechanism of carbon assessment and DRL

The core innovation of this system lies in deeply integrating carbon assessment into the decision loop of the DRL algorithm. The specific integration mechanism is reflected in the following three aspects: (1) Carbon elements are embedded in the state space. In the DRL state space s (Equation (20)), it not only includes operating states such as power and load, but also implicitly contains the carbon emission contributions of each equipment calculated in real-time through the carbon flow tracking model, providing the agent with comprehensive carbon situational awareness; (2) The reward function centers on carbon costs, taking "carbon emission cost" C_{carbon} as a key negative reward term in the reward function $R(s,a)$ (Equation (25)); (3) The action space achieves carbon-oriented scheduling, ensuring carbon-oriented optimization of scheduling strategies. Furthermore, the deep integration of carbon trading mechanisms with energy management has been proven to effectively guide systems toward low-carbon operation [6]. Zhou et al [7]. further established a joint energy-carbon trading market, deeply embedding carbon trading into a multi-agent DRL decision-making framework, and verified the effectiveness of carbon-energy collaborative optimization.

2.4. System implementation process

The implementation of this carbon assessment system follows the process: data collection and input → dynamic carbon flow modeling → DRL environment construction and training → strategy validation and assessment → result output and feedback. Through the construction of the above

system, this paper provides a complete carbon management solution for rural integrated energy systems from theory to practice, from assessment to optimization.

3. Construction of rural integrated energy system model

3.1. System structure and equipment modeling

Based on a large rural community in northern China, this paper constructs a typical rural integrated energy system model. The system includes traditional energy equipment (coal-fired units, and natural gas units), renewable energy equipment (photovoltaic, and wind power), biomass utilization equipment (biogas generation units, and biomass pyrolysis power generation systems), and energy storage equipment (electrical energy storage batteries, and thermal storage tanks). The system structure is shown in Figure 1.

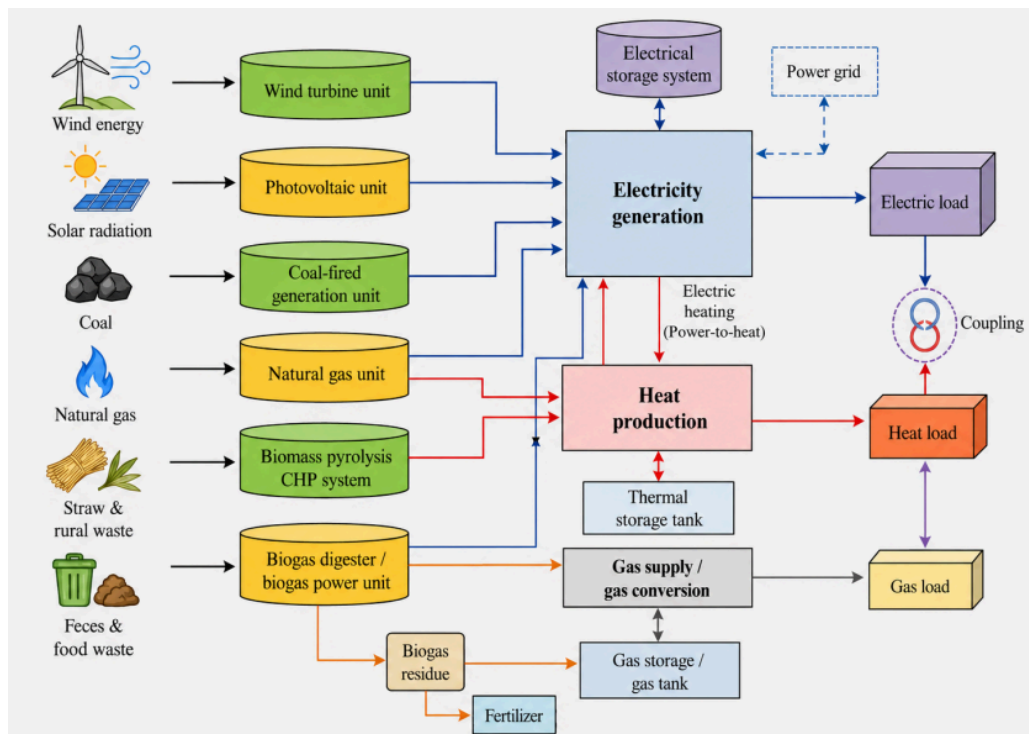


Figure 1. Structure diagram of rural integrated energy system

The carbon emission calculation formulas for each piece of equipment are as follows:

(1) Carbon emissions of coal-fired units:

$$E_{\text{coal}} = \sum_{t=1}^T (P_{\text{coal},t} \cdot \alpha_{\text{coal}}) \quad (1)$$

where, E_{coal} is the total carbon emission from coal-fired power generation (kgCO_2); $P_{\text{coal},t}$ is the power generation of coal-fired equipment at time t (kW); α_{coal} is the carbon emission coefficient of coal-fired power generation (kgCO_2/kWh).

(2) Carbon emissions of natural gas units:

$$E_{\text{gas}} = \sum_{t=1}^T (P_{\text{gas},t} \cdot \alpha_{\text{gas}}) \quad (2)$$

where, E_{gas} is the total carbon emission from natural gas equipment (kgCO_2); $P_{gas,t}$ is the power generation of natural gas equipment at time t (kW); α_{gas} is the carbon emission coefficient of natural gas power generation (kgCO_2/kWh).

(3) Carbon emissions of biogas generation units:

$$P_{biogas,t} = \eta_{biogas} \cdot V_{biogas,t} \quad (3)$$

$$E_{biogas,t} = \alpha_{biogas} \cdot V_{biogas,t} \quad (4)$$

where, $P_{biogas,t}$ is the biogas power generation at time t (kW); $V_{biogas,t}$ is the biogas consumption at time t (m^3); η_{biogas} is the biogas power generation efficiency; $E_{biogas,t}$ is the carbon emission from biogas power generation at time t (kgCO_2); α_{biogas} is the carbon emission coefficient of biogas (kgCO_2/m^3).

(4) Carbon emissions from biomass pyrolysis power generation:

$$V_{fuel,t} = (M_{straw,t} \cdot \beta_{straw} + M_{garbage,t} \cdot \beta_{garbage}) \cdot \eta_{PF} \quad (5)$$

$$g_{PG,t} = V_{fuel,t} \cdot L_{fuel} \cdot \eta_{PG} \quad (6)$$

$$Q_{PG,t} = g_{PG,t} \cdot (1 - \eta_{PG} - \eta_{loss}) \cdot \eta_h / \eta_{PG} \quad (7)$$

where, V_{fuel} is the volume of combustible gas generated at time t (m^3); $M_{straw,t}$, $M_{garbage,t}$ are the treatment amounts of straw and garbage at time t (kg), respectively; β_{straw} , $\beta_{garbage}$ are the gasification coefficients of straw and garbage, respectively; η_{PF} is the pyrolysis furnace efficiency.

(5) Electrical energy storage equipment:

$$EBT,t = EBT,t-1 \cdot (1 - \eta_{loss,BT}) + \quad (8)$$

$$(PBT,ch,t \cdot \eta_{ch} - \frac{PBT,dis,t}{\eta_{dis}}) \Delta t$$

(6) Thermal storage tank:

$$ETST,t = ETST,t-1 \cdot (1 - \eta_{loss,TST}) + \quad (9)$$

$$(HTST,ch,t \cdot \eta_{TST,ch} - \frac{HTST,dis,t}{\eta_{TST,dis}}) \Delta t$$

3.2. Operating constraints

(1) Equipment output constraints:

$$P_{i,min} \leq P_{i,t} \leq P_{i,max}, \forall i \in \{bg, coal, gas, storage\} \quad (10)$$

(2) Energy balance constraints:

Electrical balance:

$$g_{coal,t} + g_{gas,t} + g_{PG,t} + g_{grid,buy,t} = \quad (11)$$

$$L_{g,t} + g_{AHP,t} + g_{P2G,t} + g_{P2H,t} + g_{charge,t}$$

Where $g_{coal,t}$, $g_{gas,t}$ and $g_{PG,t}$ are the power outputs of the coal-fired unit, natural gas unit, and biomass pyrolysis generator at time t (kW), respectively; $g_{grid, buy, t}$ is the power purchased from the external grid at time t (kW); $L_{g,t}$ is the system electric load at time t (kW); $g_{AHP,t}$ is the power consumed by the absorption heat pump at time t (kW); $g_{P2G,t}$ is the power consumed by the power-to-gas equipment at time t (kW); $g_{P2H,t}$ is the power consumed by the power-to-heat equipment (e.g., electric boiler) at time t (kW); and $g_{charge,t}$ is the charging power of the electric energy storage device at time t (kW), corresponding to $P_{BT, ch, t}$ in Equation (8).

Thermal balance:

$$Q_{PG, t} + Q_{AHP, t} + Q_{P2H, t} = L_{Q, t} + Q_{loss, t} \quad (12)$$

Where $Q_{PG,t}$ is the thermal power output of the biomass pyrolysis power generation system at time t (kW), corresponding to Equation (7); $Q_{AHP,t}$ is the thermal power output of the absorption heat pump at time t (kW); $Q_{P2H,t}$ is the thermal power output of the power-to-heat equipment at time t (kW); $L_{Q,t}$ is the system thermal load at time t (kW); and $Q_{loss,t}$ is the thermal transmission loss of the heating network at time t (kW).

Gas balance:

$$V_{NG, t} + V_{P2G, t} + V_{grid, buy, t} = L_{V, t} + V_{storage, in, t} \quad (13)$$

Where $V_{NG,t}$ is the natural gas supply at time t (m³); $V_{P2G,t}$ is the gas output from the power-to-gas equipment at time t (m³); $V_{grid, buy, t}$ is the gas purchased from the external gas grid at time t (m³); $L_{V,t}$ is the system gas load at time t (m³); and $V_{storage, in, t}$ is the gas input to the gas storage device at time t (positive for storage, m³), whose operating constraints are similar to those of the electric energy storage device, satisfying the upper and lower limits of capacity and the charge/discharge rate limits.

(3) Equipment capacity constraints:

$$SOC_{min} \leq SOC_t \leq SOC_{max} \quad (14)$$

$$Q_{storage, min} \leq Q_{storage, t} \leq Q_{storage, max} \quad (15)$$

(4) Energy storage device constraints:

Capacity upper and lower limits: $S_{min} \leq S_t \leq S_{max}$

Charging and discharging power limits: $0 \leq P_{in, t} \leq P_{in, max}, 0 \leq P_{out, t} \leq P_{out, max}$

Mutual exclusivity of charging and discharging states

3.3. Objective function

Incorporating operational economy, carbon emissions, and power supply reliability into a unified optimization framework is a common modeling approach in low-carbon economic dispatch of integrated energy systems (IES) [8]. This paper aims to minimize the total system cost, comprehensively considering operating costs, carbon emission costs, and reliability penalty costs:

$$\text{Min } F = \omega_1 \times C_{oper} + \omega_2 \times C_{carbon} + \omega_3 \times C_{reliability} \quad (16)$$

where, F is the total system cost; C_{oper} is the total system operating cost; C_{carbon} is the total system carbon emission cost; $C_{reliability}$ is the system reliability penalty cost (quantifying the risk of

power loss); $\omega_1, \omega_2, \omega_3$ are the weight coefficients for economy, environmental friendliness, and reliability, respectively, satisfying $\omega_1 + \omega_2 + \omega_3 = 1$.

(1) Operating cost:

$$C_{oper} = \sum [\sum (C_{fuel,i} + C_{om,i}) + C_{grid,buy} \times P_{buy,t} - C_{grid,sell} \times P_{sell,t} + C_{om,ess} \times |P_{ess,t}|] \times \Delta t \quad (17)$$

where, $P_{i,t}$ is the output power of generator i at time t (kW); $C_{fuel,i}$ is the fuel cost function of generator i (CNY/kWh); $C_{om,i}$ is the operation and maintenance cost of generator i (CNY/kWh); $C_{grid,buy}$ is the electricity purchase price from the main grid (CNY/kWh); $P_{buy,t}$ is the power purchased from the main grid at time t (kW); $C_{grid,sell}$ is the electricity selling price to the main grid (CNY/kWh); $P_{sell,t}$ is the power sold to the main grid at time t (kW); $C_{om,ess}$ is the operation and maintenance cost of energy storage equipment (CNY/kW); $P_{ess,t}$ is the charging/discharging power of energy storage, positive for charging and negative for discharging (kW); Δt is the scheduling time interval (hour). Regarding multi-objective optimization, existing studies have adopted the DDPG algorithm to solve the multi-objective dispatch problem of electricity-gas-heat integrated energy systems (IES), while simultaneously optimizing economic operation costs and environmental factors [9].

(2) Carbon emission cost:

$$C_{carbon} = \alpha E \cdot (E_{coal} + E_{gas} + E_{PG} + E_{grid}) \quad (18)$$

where αE is the unit carbon emission cost (CNY/kgCO₂); E_{grid} is the indirect carbon emission from purchased electricity.

(3) Reliability penalty cost:

$$C_{reliability} = \sum_t [\pi_{power} \cdot L_{shed,t} + \pi_{heat} \cdot Q_{shed,t}] \quad (19)$$

where $L_{shed,t}$ and $Q_{shed,t}$ are the electrical and thermal load shortages at time t ; π_{power} and π_{heat} are the unit shortage penalty costs (Yuan/kWh).

4. Deep reinforcement learning algorithm design

4.1. Markov decision model

The optimal scheduling problem of the rural integrated energy system is modeled as a Markov Decision Process (MDP), containing four key elements: state space, action space, state transition probability, and reward function. Modeling the integrated energy system scheduling problem as a Markov Decision Process (MDP) is a prerequisite for the application of DRL methods, and existing studies have verified the effectiveness of this modeling approach in IEGS with carbon capture and power-to-gas interconnection [10]. The state space includes system operating states, load demands, renewable energy output, and electricity prices; the action space includes energy storage charging/discharging power and grid interaction power. The reward function aims to minimize the total system cost. Based on the rural integrated energy system model established in Chapter 2, the details of these elements of the Markov decision model are as follows:

The state st represents all key operating variables of the system at time t :

$$st = [P_{coal,t}, P_{gas,t}, P_{BPG,t}, g_{PG,t}, Q_{PG,t}, EBT,t, ETST,t, SOCT, Lg,t, LQ,t, LV,t, PPV,t, P_{wind,t}, C_{grid,buy,t}, t] \quad (20)$$

Where $P_{coal,t}$, $P_{gas,t}$, $P_{BPG,t}$ are coal, natural gas, and biogas power generation, respectively; $g_{PG,t}$, $Q_{PG,t}$ are biomass pyrolysis power generation and heat supply power, respectively; EBT,t , $ETST,t$ are electrical energy storage and thermal storage tank energy states, respectively; $SOCT$ is state of charge of energy storage devices; Lg,t , LQ,t , LV,t are electrical, thermal, and gas loads, respectively; PPV,t , $P_{wind,t}$: photovoltaic and wind power output, respectively; $C_{grid,buy,t}$ is grid electricity purchase price; t : time information (hour, day type, etc.).

Action Space:

The action at represents the dispatchable control variables of the system at time t :

$$at = [P_{coal,t}, P_{gas,t}, P_{BPG,t}, g_{PG,t}, Q_{PG,t}, PBT,ch/dis,t, HTST,ch/dis,t, P_{grid,buy/sell,t}] \quad (21)$$

Where $PBT,ch/dis,t$ is electrical energy storage charging/discharging power; $HTST,ch/dis,t$ is thermal storage tank charging/discharging power; $P_{grid,buy/sell,t}$ is grid interaction power (purchase/sale).

State Transition Probability:

State transition describes the probability of transitioning to $s\{t+1\}$ after taking action at in state st :

$$P(st+1|st,at) = \prod_i P(st+1,i|st,at) \quad (22)$$

wherein, the transition is determined by the system dynamics equation, the details are as follows:
Energy storage state update:

$$EBT,t+1 = EBT,t \cdot (1 - \eta_{loss,BT}) + (PBT,ch,t \cdot \eta_{ch} - \frac{PBT,dis,t}{\eta_{dis}}) \Delta t \quad (23)$$

Energy balance constraint:

$$\sum P_{gen,t} = Lg,t + P_{charge,t} + P_{grid,sell,t} - P_{grid,buy,t} \quad (24)$$

Uncertainty sources: renewable energy utilization, load fluctuations, electricity price changes, etc.

Reward Function:

The reward $R(st, at)$ measures the quality of action at in state st , with the objective of minimizing the total cost:

$$R(st, at) = -[\omega_1 \cdot C_{oper,t} + \omega_2 \cdot C_{carbon,t} + \omega_3 \cdot C_{reliability,t}] \quad (25)$$

The operating costs therein are shown in Equations (17), (18) and (19).

4.2. Si-sacalgorithmdesign

In the field of low-carbon dispatch, Alabi et al. [11] have adopted the Soft Actor-Critic (SAC) algorithm to achieve real-time dispatch of multi-energy systems with carbon capture systems,

verifying the advantages of the SAC framework in handling low-carbon optimization problems in continuous action spaces. However, the traditional SAC algorithm suffers from low training efficiency and high constraint violation rates when dealing with complex constraints. To address the constraint violation problem in DRL decision-making, safe reinforcement learning methods have been applied in low-carbon dispatch of multi-energy systems by introducing safety networks and Lagrange relaxation techniques to handle operational constraints [12]. This paper proposes the SI-SAC algorithm, which introduces a safety component to ensure the agent always satisfies hard constraints, and an Imitation Component to rapidly acquire initial compliant strategies from expert demonstrations, effectively improving algorithm performance. The Figure 2 illustrates the basic operating mechanism of deep reinforcement learning.

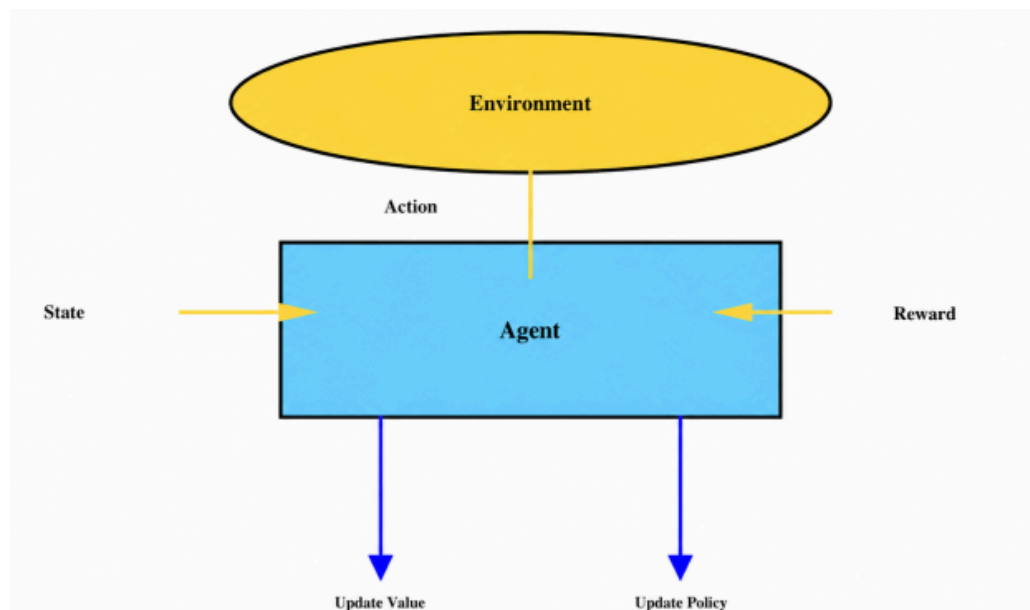


Figure 2. Operating mechanism diagram of deep reinforcement learning

The core improvements of the SI-SAC algorithm include: (1) The safety component ensures policy feasibility by penalizing constraint-violating actions. It is worth adding that, in terms of safety mechanism design, existing studies have employed an enhanced action projection mechanism to project the agent's actions into the feasible region, thereby ensuring that DRL decisions always satisfy physical system constraints [13]. (2) The imitation component accelerates policy learning by utilizing expert data generated by heuristic algorithms; (3) The entropy regularization mechanism encourages policy exploration and enhances algorithm robustness.

5. Case study

5.1. Case study setup

A typical rural community in northern China is selected as the research object to construct an integrated energy system model. The system includes coal-fired units, natural gas units, biogas generation units, biomass pyrolysis power generation systems, photovoltaic cells, wind power, electrical energy storage batteries, and thermal storage tanks. The scheduling period is 24 hours with a time resolution of 1 hour. The algorithm parameters are set as follows: both the policy network and value network adopt two-layer fully connected networks with 256 neurons per layer; the experience

replay buffer size is $1e6$; the discount factor $\gamma=0.99$; the entropy regularization coefficient α has an initial value of 0.2; and the total number of training episodes is 5000. The relevant hyperparameters are set according to empirical values from existing SAC algorithms in IES dispatch [11].

Three comparison schemes are set up:

Scheme 1 (Baseline): Traditional MILP method to minimize operating costs;

Scheme 2 (Single-objective DRL): Basic DDPG algorithm with reward function containing only operating costs;

Scheme 3 (SI-SAC): The proposed SI-SAC algorithm comprehensively considering operating costs, carbon emission costs, and reliability penalty costs.

5.2. Algorithm convergence analysis

Figure 3 shows the total cost variation curves of the SI-SAC and DDPG algorithms during the training process. As shown in the figure, the SI-SAC algorithm can acquire compliant strategies more rapidly in the early training stage, avoiding frequent constraint violations, thus achieving faster convergence. After approximately 1500 episodes, the total cost stabilizes at a lower level, indicating that the algorithm has learned an effective scheduling strategy. In contrast, the DDPG algorithm converges more slowly and achieves a higher final convergence value.

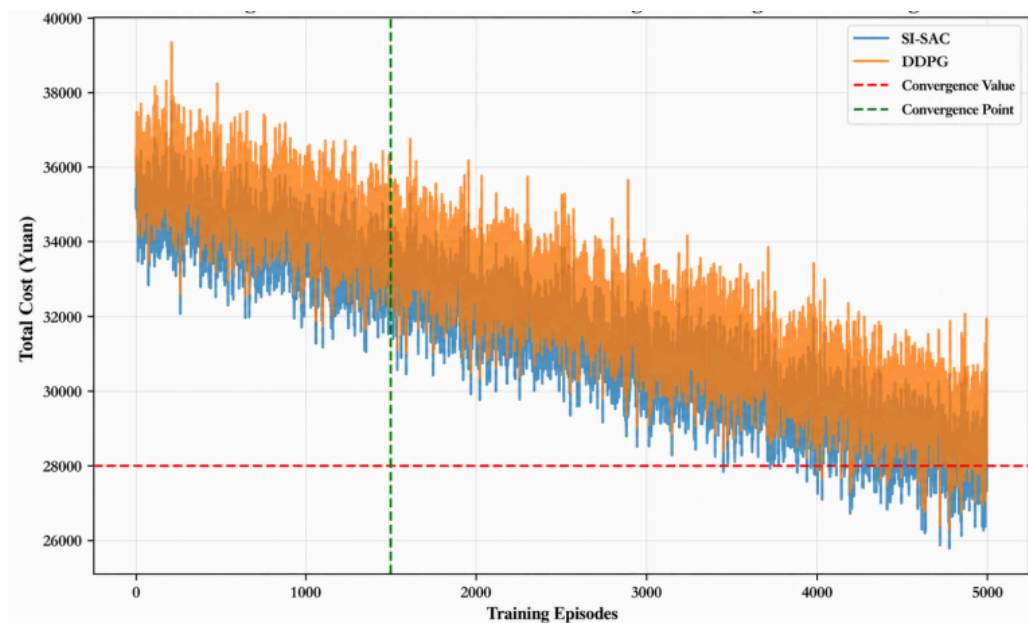


Figure 3. Total cost variation curve during SI-SAC algorithm training

5.3. Scheduling results comparative analysis

Table 1 summarizes the total operating costs, total carbon emissions, and renewable energy utilization rates of the three schemes. Figure 4 intuitively compares the economic and environmental performance of the three schemes.

Table 1. Comparison of scheduling results among three schemes

Scheme	Total Operating Cost (Yuan)	Total Carbon Emissions (kgCO ₂)	Renewable Energy Utilization Rate (%)
MILP	32,560	18,450	78.5
DDPG	31,280	17,920	80.2
SI-SAC	29,870	16,540	83.6

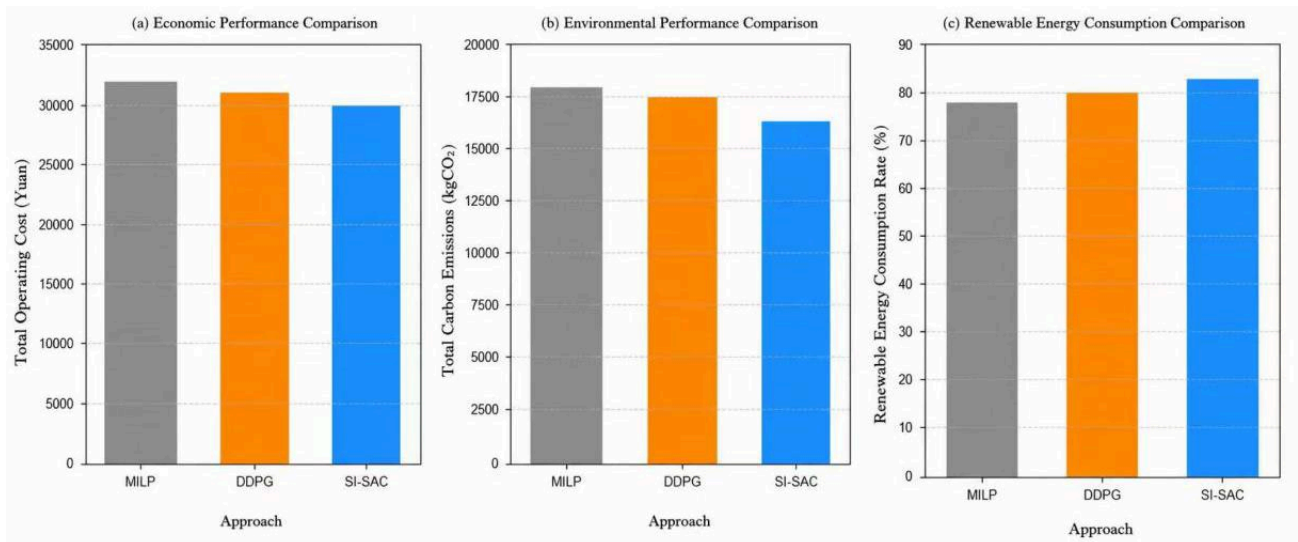


Figure 4. Economic and environmental performance comparison of three schemes: (a) Economic performance comparison (b) Environmental performance comparison (c) Renewable energy consumption comparison

Analysis shows that, in terms of economy, the SI-SAC scheme achieves the lowest total operating cost, reducing by approximately 8.3% compared to the MILP scheme; in terms of environmental friendliness, the SI-SAC scheme achieves the lowest total carbon emissions, reducing by approximately 10.4% compared to the MILP scheme; in terms of renewable energy utilization, the SI-SAC scheme achieves the highest utilization rate of 83.6%. The above results indicate that incorporating carbon costs into the optimization objective and adopting the SI-SAC algorithm can effectively exploit the system's potential for cost reduction and carbon emission reduction. The carbon emission reduction ratio achieved in this paper is of a similar order of magnitude to the carbon reduction effect of the coordinated scheduling of carbon capture and storage with power-to-gas (CCS-P2G) in reference [10], which validates the general carbon reduction effectiveness of the DRL method across different IES scenarios.

5.4. Robustness and computational efficiency analysis

To validate the robustness of the SI-SAC algorithm, random fluctuations ($\pm 15\%$) in photovoltaic and wind power outputs are introduced in the test scenarios. Table 2 compares the average performance of DDPG and SI-SAC algorithms under uncertainty scenarios, and Figure 5 shows the robustness comparison between the two algorithms.

Table 2. Algorithm performance comparison under uncertainty scenarios

Algorithm	Average Operating Cost (Yuan)	Cost Standard Deviation (Yuan)	Average Carbon Emissions (kgCO ₂)	Constraint Violation Rate (%)
DDPG	31,950	520	18,100	5.2
SI-SAC	30,120	380	16,950	1.5

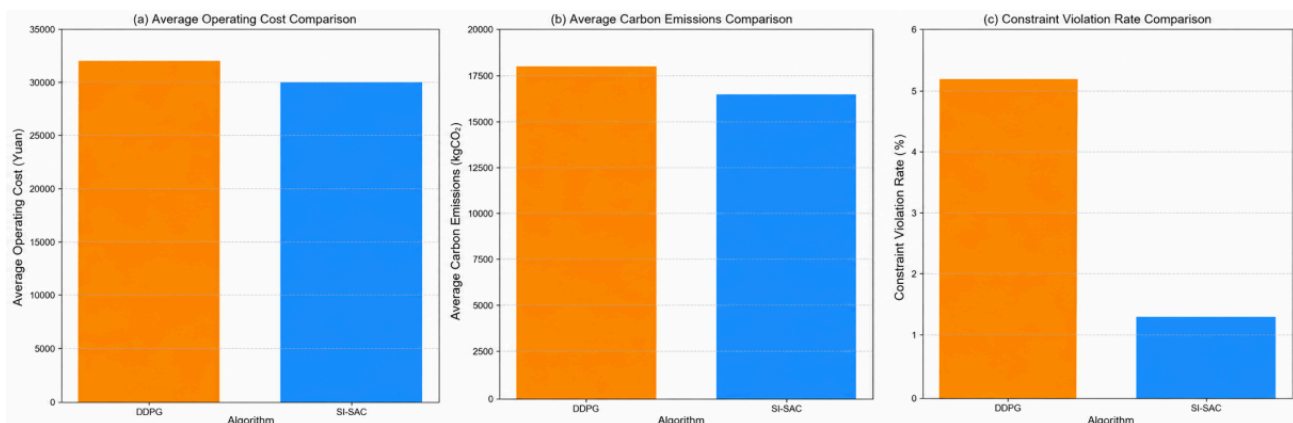


Figure 5. Robustness comparison of two algorithms under uncertain environments: (a) Average operating cost comparison (b) Average carbon emission comparison(c) Constraint violation rate comparison

The results show that under uncertain environments, the SI-SAC algorithm achieves lower average operating costs and carbon emissions than the DDPG algorithm, with a smaller cost standard deviation, indicating more stable and robust decision-making results. More importantly, the constraint violation rate of the SI-SAC algorithm is only 1.5%, significantly lower than 5.2% of the DDPG algorithm, which benefits from the introduction of the safety component in the SI mechanism. This result is consistent with the constraint handling performance of the safe reinforcement learning method reported in [12], further validating the effectiveness of the safety mechanism in DRL-based scheduling decisions. In terms of computational efficiency, the trained SI-SAC model takes approximately 0.15 seconds on average for an intra-day scheduling decision, which can meet real-time scheduling requirements.

6. Conclusion

This paper addresses the issues of carbon assessment and carbon reduction optimization in rural integrated energy systems (RIES) by proposing a dynamic carbon assessment framework that integrates deep reinforcement learning (DRL), along with a novel SI-SAC algorithm featuring a safety-imitation mechanism. Through theoretical modeling, algorithm design, and case validation, the following main conclusions are drawn: (1) In terms of the carbon assessment system, the dynamic carbon flow tracking model established in this paper is capable of coupling equipment operating states with carbon emission factors in real time, enabling accurate carbon emission measurement and multi-energy flow tracing for RIES. This system overcomes the limitations of traditional static accounting methods and provides an effective means for carbon footprint visualization in rural energy systems, transforming carbon emissions from "post-event statistics" to "in-process perception," thereby laying a data foundation for low-carbon scheduling. (2) In terms of

algorithmic innovation, the proposed SI-SAC algorithm penalizes constraint-violating actions through its safety component, effectively reducing the risk of infeasible scheduling strategies. Meanwhile, the imitation component utilizes expert trajectories generated by heuristic algorithms for pre-training, significantly shortening the inefficient trial-and-error process during the early exploration phase. Experiments show that the convergence speed of SI-SAC is approximately 30% faster than that of the baseline DDPG, and the total cost after convergence is reduced by about 4.5% compared to DDPG, demonstrating the significant advantages of the safety-imitation mechanism in complex constrained optimization problems. (3) In terms of carbon reduction potential and economic performance, the case study results indicate that the SI-SAC-driven scheduling strategy can reduce carbon emissions by approximately 10.4% and operating costs by approximately 8.3% compared with the traditional MILP method, while increasing the renewable energy accommodation rate to 83.6%. This fully demonstrates that internalizing carbon costs as optimization objectives and adopting advanced DRL algorithms can achieve substantial carbon reductions without sacrificing economic performance, providing a feasible technical pathway for rural energy systems to meet the "Dual Carbon" goals. (4) In terms of robustness and practicality, in the face of uncertainties such as renewable energy output fluctuations, the SI-SAC algorithm achieves a constraint violation rate of only 1.5%, far lower than the 5.2% of DDPG, with a smaller standard deviation in decision outcomes, demonstrating excellent robustness. Moreover, its single-decision time is only 0.15 seconds, which meets the real-time scheduling requirements in practical engineering applications and holds promising prospects for engineering deployment.

This paper currently focuses on a single rural community scenario and does not consider the impacts of multi-region interconnection or carbon trading mechanisms. In addition, the time-varying characteristics of model parameters (such as carbon emission coefficients and equipment efficiencies) require further refinement. Future work will extend to coordinated scheduling of multiple rural parks and introduce carbon trading market mechanisms [7] to explore new paradigms for economic-environmental synergistic optimization. Meanwhile, we plan to incorporate distributed architectures such as federated learning into DRL training to protect user data privacy and enhance model generalization capability.

In summary, the framework proposed in this paper provides a quantitative decision-making tool for the low-carbon, economical, and secure operation of rural energy systems, and holds significant theoretical value and practical significance for advancing the scientific implementation of the "Dual Carbon" goals in rural areas.

References

- [1] Luo, Q. (2022). Xi Jinping's important discourse on carbon peak and carbon neutrality: Logical path, value implication and practice path. *Study of Governance Modernization*, 38(4), 5–13.
- [2] Wang, X. L., Wang, X. F., & Bie, Z. H. (2016). A review of modeling and optimization methods for integrated energy systems. *Proc. CSEE*, 36(12), 3292–3305.
- [3] Hou, S., Salazar, E. M., Vergara, P. P., et al. (2022). Performance comparison of deep RL algorithms for energy systems optimal scheduling. In *2022 IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)* (pp. 1–5). IEEE.
- [4] Mo, M., Xiong, X., Wu, Y., et al. (2023). Deep-reinforcement-learning-based low-carbon economic dispatch for community-integrated energy system under multiple uncertainties. *Energies*, 16(22), 7669.
- [5] Lee, S., Prabawa, P., Choi, D.-H., et al. (2025). Joint peak power and carbon emission shaving in active distribution systems using carbon emission flow-based deep reinforcement learning. *Applied Energy*, 379, 124944.
- [6] Sun, Q. K., Wang, X. J., et al. (2022). Multi-agent energy management optimization for integrated energy systems under the energy and carbon co-trading market. *Applied Energy*, 324, 119646.

- [7] Zhou, Y., Ma, Z., Wang, T., et al. (2024). Joint energy and carbon trading for multi-microgrid system based on multi-agent deep reinforcement learning. *IEEE Transactions on Power Systems*. DOI: 10.1109/TPWRS.2024.3380070.
- [8] Dong, J., Wang, H. X., et al. (2022). Low carbon economic dispatch of integrated energy system considering power supply reliability and integrated demand response. *Energy*.
- [9] Li, F., Liu, L., et al. (2024). Deep reinforcement learning-based multi-objective optimization for electricity–gas–heat integrated energy systems. *Engineering Applications of Artificial Intelligence*.
- [10] Liang, T., Chai, L., Tan, J., et al. (2024). Dynamic optimization of an integrated energy system with carbon capture and power-to-gas interconnection: A deep reinforcement learning-based scheduling strategy. *Applied Energy*, 367, 123390.
- [11] Alabi, T. M., Lawrence, N. P., Lu, L., et al. (2023). Automated deep reinforcement learning for real-time scheduling strategy of multi-energy system integrated with post-carbon and direct-air carbon captured system. *Applied Energy*, 333, 120633.
- [12] Jiang, F., Chen, J., et al. (2024). Safe reinforcement learning based optimal low-carbon scheduling strategy for multi-energy system. *Sustainable Energy, Grids and Networks*, 101454.
- [13] Ye, T., Huang, Y. P., et al. (2025). Safe multi-agent deep reinforcement learning for decentralized low-carbon operation in active distribution networks and multi-microgrids. *Applied Energy*, 387, 125609.