



OPEN ACCESS

EDITED BY

Yue Xiang,
Sichuan University, China

REVIEWED BY

Ke Qing,
Sichuan Normal University, China
Xianpeng Xi,
State grid nanchong power supply company,
China
Bin Zhang,
Aalborg University, Denmark
Mingxuan Li,
SPIC Southwest Energy Research Institute,
China

*CORRESPONDENCE

Wentao Feng,
✉ 1179623991@qq.com

RECEIVED 06 May 2024

ACCEPTED 11 June 2024

PUBLISHED 19 July 2024

CITATION

Feng W, Deng B, Zhang Z, Jiang H, Zheng Y, Peng X, Zhang L and Jing Z (2024), Low-carbon economic dispatch strategy for integrated electrical and gas system with GCCP based on multi-agent deep reinforcement learning. *Front. Energy Res.* 12:1428624. doi: 10.3389/fenrg.2024.1428624

COPYRIGHT

© 2024 Feng, Deng, Zhang, Jiang, Zheng, Peng, Zhang and Jing. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Low-carbon economic dispatch strategy for integrated electrical and gas system with GCCP based on multi-agent deep reinforcement learning

Wentao Feng^{1*}, Bingyan Deng¹, Ziwen Zhang¹, He Jiang¹, Yanxi Zheng¹, Xinran Peng¹, Le Zhang¹ and Zhiyuan Jing²

¹State Grid Sichuan Information and Telecommunication Company, Chengdu, Sichuan, China, ²School of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China, Chengdu, Sichuan, China

With the growing concern for the environment, sustainable development centred on a low-carbon economy has become a unifying pursuit for the energy industry. Integrated energy systems (IES) that combine multiple energy sources such as electricity, heat and gas are essential to facilitate the consumption of renewable energy and the reduction of carbon emission. In this paper, gas turbine (GT), carbon capture and storage (CCS) and power-to-gas (P2G) device are introduced to construct a new carbon capture coupling device model, GT-CCS-P2G (GCCP), which is applied to the integrated electrical and gas system (IEGS). Multi-agent soft actor critic (MASAC) applies historical trajectory representations, parameter spatial techniques and deep densification frameworks to reinforcement learning for reducing the detrimental effects of time-series data on the decisional procedure. The energy scheduling problem of IEGS is redefined as a Markov game, which is addressed by adopting a low carbon economic control framework based on MASAC with minimum operating cost and minimum carbon emission as the optimization objectives. To validate the rationality and effectiveness of the proposed low-carbon economy scheduling model of IEGS based on MASAC, this paper simulates and analyses in integrated PJM-5 node system and seven nodes natural gas system.

KEYWORDS

low-carbon, integrated electrical and gas system, carbon emission, GCCP, multi-agent deep reinforcement learning algorithm

1 Introduction

With population growth and accelerated industrialization, energy consumption is increasing, along with significant greenhouse gas emission. These emissions have a huge impact on climate change, with issues such as extreme weather events, sea-level rise and ecosystem collapse (Ma et al., 2024; Wojtaszek et al., 2024). It is therefore crucial to adopt energy-saving and emission-reduction measures (Li et al., 2019). By promoting the use of renewable energy, improving energy efficiency and adopting cleaner production technologies, we can reduce dependence on fossil fuels and lower greenhouse gas emissions (Liu et al., 2024; Okedu et al., 2024).

The United Nations Framework Convention on Climate Change, adopted by the United Nations in 1994, became the world's first international treaty to reduce emissions of greenhouse gases such as carbon dioxide and methane, requiring all countries in the world to take some responsibility for carbon emission reduction (Werksman, 1994). In 2005, countries continued to sign the Kyoto Protocol, which provided programs to reduce carbon emissions and established corresponding cooperation mechanisms for countries to deal with carbon emission reduction, thus advancing the development of global carbon emission reduction (Gallo et al., 2018). The Paris Agreement, signed in Paris in 2015, requires developed countries, while accomplishing their own carbon emission reduction tasks, to provide certain financial support and related technologies to developing countries, contributing to the fulfillment of the global carbon emission reduction tasks (Elsayed et al., 2024). In 2021, China proposes to vigorously develop renewable energy sources such as wind power and photovoltaic, laying a solid foundation for achieving carbon peak by 2030 and carbon neutrality by 2060 (Zhong et al., 2023). However, renewable energy sources such as wind turbine (WT) and photovoltaic (PV) have problems such as chronology, stochasticity and volatility. When a high percentage of renewable energy is injected into the grid it can bring about problems such as sudden voltage changes (Cao et al., 2020a; Cao et al., 2021) and system collapse (Barker and Mello, 2000; Dulăua et al., 2013), which has also become the focus of carbon emission reduction (Gao and Zhang, 2024).

In traditional energy dispatching, various energy sources are dispatched by different sub-networks (e.g., electrical network, gas network, and heat/cooling network), in which the sub-networks are controlled by different departments. In actual operation, power grids are coupled with each other in terms of production, transmission, distribution and utilization. However, a singular scheduling makes it impossible to form an effective information sharing and energy complementarity between sub-networks, so as to fail to guarantee the system stability. Multi-energy flow deployment can improve the efficiency of energy utilization, reduce the total cost of system operation, and realize the stability and safety of system operation (Liu et al., 2018; Wang et al., 2023a). Consequently, breaking the limitations of traditional energy architecture and constructing multi-energy network architecture, such as electricity-heat-gas-cooling integrated energy system (IES) to tap the potential of energy transmission between different systems, is one of the core contents of research by scholars in related fields in various countries (Liu et al., 2023).

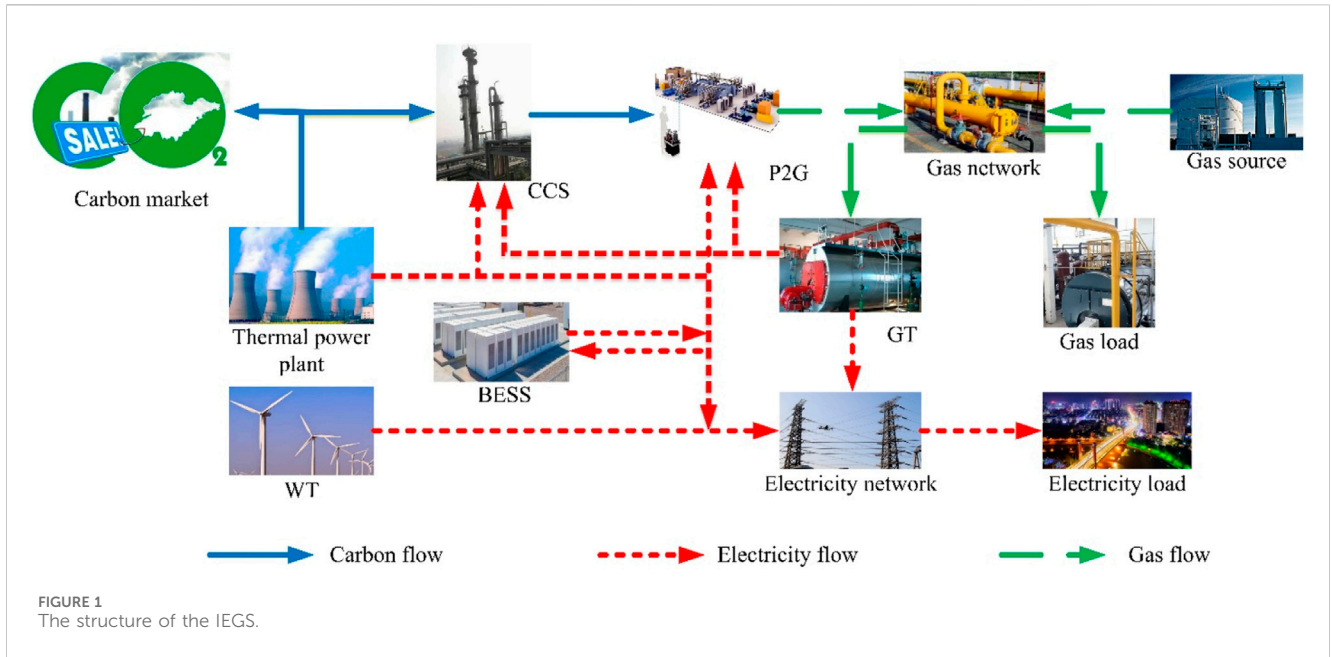
Substituting conventional thermal power generation with renewable energy sources such as wind and photovoltaic, reducing the utilization rate of conventional thermal power plants by applying power-to-gas (P2G) and gas turbines (GT), and achieving carbon dioxide (CO₂) absorption and utilization are significant in reducing carbon emissions. Carbon capture and storage (CCS) provides an alternative and effective technology for dealing with CO₂ emissions, with 92 per cent of the CO₂ produced by coal-fired units being captured and stored in the IES. It can be seen that the combination of P2G and CCS technologies can effectively reduce and improve carbon emissions, and the carbon capture technology also improves the carbon feedstock for methane

generation by P2G, thereby reducing the amount of CO₂ generated by coal-fired units in the IES (Gu et al., 2017; Yang et al., 2019).

He et al. (2022) constructs a near-zero emission park-level IES considering uncertainty with P2G and CCS. Zhang et al. (2020) constructed an integrated electricity-gas energy system (IEGS) optimization model considering P2G and wind power uncertainty based on distributed robust optimization. The superiority of the low carbon emission is effectively verified by the results of three different IEGSS. In order to determine the optimal capacity of the gas turbine and P2G technology for different IEGS, a Monte Carlo based optimization framework has been proposed in Tabebordbar et al. (2023). The experimental results demonstrate the superiority and sophistication of the reliability-oriented optimization framework. However the algorithms adopted in the above mentioned literature struggle to attain the satisfactory results in the presence of the complexity and diversity of the system, the mutual constraints of the coupled energy components and the large dimension of the optimization objective.

With the advancement of science and technology and the rapid progress of artificial intelligence technology, many scholars have proposed a variety of optimization and control strategies based on machine learning in the literature. Reinforcement learning (RL) is currently the most popular method to solve the control optimization problem, which includes both agent and environment (Cao et al., 2020b; Zhang et al., 2023a; Cao et al., 2023; Li et al., 2024a, 2024c). Zhang et al. (2023b) proposed a two-timescale energy management strategy based on multi-agent deep RL (MADRL) for residential multicarrier energy system, where the optimal solution of each coupling element in the system is obtained to achieve the optimal control effect. Aiming at the joint operation of multiple microgrids, a MADRL-based energy management method is proposed in Li et al. (2023). Each microgrid as an agent performs a game with each other, and the continuous training ensures that each agent chooses the local optimal strategy under the global optimal situation. Taking into account the different characteristics between the electricity and the heat network, Monfaredi et al. (2023) achieves an hourly optimal scheduling strategy by scheduling multiple renewable energy sources. During the optimization process, MADRL is applied to achieve the information interaction between the energy storage system, new energy sources, heat and power conversion system and the grid, which results in the desirable control strategy of the system to improve the energy utilization. In order to address the distributed energy management problem of multi-area IES, a MADRL-based energy management strategy is proposed, which effectively decreases the influences of renewable energy uncertainty on the decision-making of the optimization model by exploiting the generalization capability of RL (Ding et al., 2024). A MADRL-based building energy management model has been proposed which achieves excellent dynamic decision making through centralized training and distributed execution (Wang et al., 2024).

This paper proposes a low carbon and economic IEGS scheduling method based on multi-agent soft actor critic (MASAC), which achieves a bidirectional coupling between the electrical network and gas network by utilizing the P2G and GT. the CCS captures the CO₂ produced by the power plant as feedstock for the CH₄ produced by the P2G, which will reduce the carbon emissions of the system, and the excess CO₂ will be purchased



through the carbon trading market to achieve a zero-carbon system. The main contributions of this paper can be summarized as.

- 1) A GT-CCS-P2G (GCCP) model is presented to achieve a two-way coupling between the electric and gas grids in IECS.
- 2) The electrical network and gas network are separately modelled as an agent to enhance the generalization capability of the energy dispatch model through reciprocal gaming.
- 3) A novel energy scheduling strategy model based on MASAC exploiting historical data is proposed.

The reminder structure of this article is as follows. Section 2 focuses on the theory of electrical and gas networks and related coupling elements. Section 3 describes the algorithmic solution process of the method proposed in this paper. Section 4 verifies the superiority of the proposed method through a detailed analysis of examples. Section 5 indicates the summary of the whole paper.

2 Problem formulation

The detailed structure of IECS is shown in Figure 1, including the electrical and gas networks. Electricity in the power network is supplied by thermal power plants, GT and WT, where the CO₂ from the thermal power plants is captured by CCS and used as feedstock for methane generation by P2G. CO₂ from thermal power plants is converted to CH₄ by P2G treatment which is transmitted to the natural gas network. CO₂ that cannot be captured by CCS will be purchased for carbon emission trading rights to achieve the zero carbon target. Perfectly realize the bi-directional coupling between electrical and gas networks through GT, CCS and P2G, enabling a bi-directional flow of energy. Battery energy storage system (BESS) as a rechargeable and dischargeable energy device enhances the proportion of renewable energy consumed by the electrical network.

2.1 Natural gas system modelling

In a natural gas system, natural gas is transmitted to the consumer through a pipeline, which consists of a gas source, a gas load, a transmission pipeline, and a compressor (Zhang et al., 2024a). Natural gas system modelling is mainly gas source, load, nodal pressure and pipeline flow modelling.

2.1.1 Gas source and load

The main components of a typical natural gas network are gas wells and gas storage stations. In practice, the supply of natural gas from gas wells is not unlimited with certain constraints that should exist in its supply. It can be expressed as Eq. 1.

$$S_{i,\min} \leq S_{i,t} \leq S_{i,\max} \quad (1)$$

where $S_{i,t}$ is the Natural gas supply; $S_{i,\min}$ and $S_{i,\max}$ represent the minimum and maximum of Natural gas supply, separately.

2.1.2 Pipeline flow modelling

During the transport of natural gas, its flow rate does not decrease. Analogous to voltage losses in power system, nodal pressure losses will exist at the beginning and end of a natural gas pipeline. Flow always moves from the high pressure node to the low pressure node in the pipeline, with the magnitude depending on the length of the pipeline, diameter, operating temperature and pressure. The relationship between pipe flow rate magnitude and pipe pressure can be expressed as Eqs 2, 3 (Dai et al., 2020).

$$\begin{cases} B_{ij} = \text{sng}(\pi_i, \pi_j) * C_{ij} \sqrt{|\pi_i^2 - \pi_j^2|} \\ \text{sng}(\pi_i, \pi_j) = \begin{cases} 1 & \pi_i \geq \pi_j \\ -1 & \pi_i < \pi_j \end{cases} \end{cases} \quad (2)$$

$$\pi_{i,\min} \leq \pi_i \leq \pi_{i,\max} \quad (3)$$

where B_{ij} is pipe flow rate; π is the node pressure; $\text{sng}(\pi_i, \pi_j)$ is the function whose value is 1 When the pressure at node i is higher than

that at node j and *vice versa* its value is -1 ; C_{ij} is a coefficient; $\pi_{i,\min}$ and $\pi_{i,\max}$ are the minimum and maximum of the node pressure.

The correlation matrix between the injected flow at each node and the pipeline flow can be established by using the forward backward generation method for the acyclic natural gas network. The correlation matrix between the injection flow and the pipeline flow at each node can be established for the non-cyclic natural gas network by using the forward back generation method, which is similar to the concept of the generation shift factor (GSF) in the direct current flow method of the power system. The relationship between the natural gas supply and load at each node with the pipeline flow is represented by.

$$B_{ij} = \sum_{n=1}^N \text{GSF}_{\text{gas},m,j} * (S_n - GL_n) \quad (4)$$

where n denotes the nature gas injection node; GL_n indicates the consumption gas load of node n .

A link between each pipe node is established based on Eq. (4). Therefore, the pressure in each node can be obtained from the acquired pipeline flow rate based on Eq. (2).

2.1.3 Compressor station

As the distance of gas transmission increases, pressure losses between nodes can lead to low pressure at the end of the pipeline which limits the transmission capacity of the network. The most important part of the compressor station is the compressor which consumes electrical energy that increases the pressure of the natural gas. The pressurization station considered in this paper is of fixed variable ratio and the energy consumed is from electrical energy contained in the load of the grid node (Bai et al., 2016). It can be presented as Eqs 5, 6.

$$H_{\text{com}} = EB_{ij} \left[\left(\frac{\pi_i}{\pi_j} \right)^G - 1 \right] \quad (5)$$

$$P_{\text{com}} = H_{\text{com}} * \chi \quad (6)$$

where H_{com} represents the power required by the compressor; E , G and χ represent coefficient; P_{com} represents the electrical load required by the compressor.

2.2 Electricity system modelling

The results of power flow calculations are the basis for analyzing the feasibility, safety, reliability and economics of grid planning and supply options. Power flow analyses perform a vital role in grid operation modelling and designing which can be calculated as Eqs 7, 8.

$$P_{i,t} = V_{i,t} \sum_{j=1}^N V_{j,t} (G_{ij,t} \cos \theta_{ij,t} + B_{ij,t} \sin \theta_{ij,t}) \quad (7)$$

$$Q_{i,t} = V_{i,t} \sum_{j=1}^N V_{j,t} (G_{ij,t} \sin \theta_{ij,t} - B_{ij,t} \cos \theta_{ij,t}) \quad (8)$$

where $P_{i,t}$ denotes the active power injected by node i at time t ; $G_{ij,t}$ and $B_{ij,t}$ represent the correlation values of the node i and node j

conductivity matrices, respectively; N indicates the total amount of nodes; $\theta_{ij,t}$ denotes the phase difference from nodes i with j ; $Q_{i,t}$ denotes the reactive power injected by node i at time t .

2.3 Coupling elements modelling

The GT and the P2G technology enable the deep coupling of the power system with the natural gas system, thereby enabling a bi-directional conversion of the electricity-gas system.

2.3.1 GT

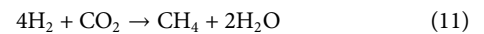
The GT can be viewed as a power source in a power system, however in the natural gas system as the load. The relationship between the power generated and the natural gas consumed can be expressed as Eq. 9 (Ji et al., 2013).

$$P_{\text{GT},t} = \zeta_{\text{GT}} * GL_{\text{GT},t} \quad (9)$$

where $P_{\text{GT},t}$ represents the generation of electricity from the GT at node i at time t ; ζ_{GT} represents conversion efficiency of GT; $GL_{\text{GT},t}$ represents the gas load of the GT at node i at time t .

2.3.2 P2G

P2G technology consists of two main steps: the electrolysis of water and the synthesis of methane. The chemical equations for the two reactions are expressed as Eqs 10, 11 (Clegg and Mancarella, 2015).



The hydrogen produced in the first step can be stored in a hydrogen storage facility or injected into the network in a mixture of natural gas, but the concentration is limited due to safety factors. The methane produced in the second step can be stored in large quantities or transported to other places where it is necessary, which absorbs a large amount of CO_2 and reduces carbon emissions. P2G not only strengthens the coupling of the electricity-gas network, but also consumes the electricity generated by the new energy to increase the proportion of new energy consumption. The conversion relationship between P2G consumption of electrical energy and the generation of natural gas can be demonstrated as Eqs 12, 13.

$$E_{\text{P2G},t} = \zeta_{\text{P2G}} * P_{\text{P2G},t} \quad (12)$$

$$S_{\text{P2G},t} = \frac{E_{\text{P2G},t}}{\text{HHV}} \quad (13)$$

where $E_{\text{P2G},t}$ and $S_{\text{P2G},t}$ denote the power and volume values obtained by the P2G technology at time t , respectively; ζ_{P2G} represents the conversion efficiency of P2G; HHV indicates the conversion relationship between energy and volume of natural gas of P2G; $P_{\text{P2G},t}$ denotes the P2G electrical load at time t .

2.3.3 CCS

CCS has been identified as a key and promising technology for future power generation (Zhang et al., 2024). Capture and storage are the two main phases of current carbon capture technology. The capture process is complex, and the main commercially available CO_2 capture methods fall into three categories: oxygenated fuel

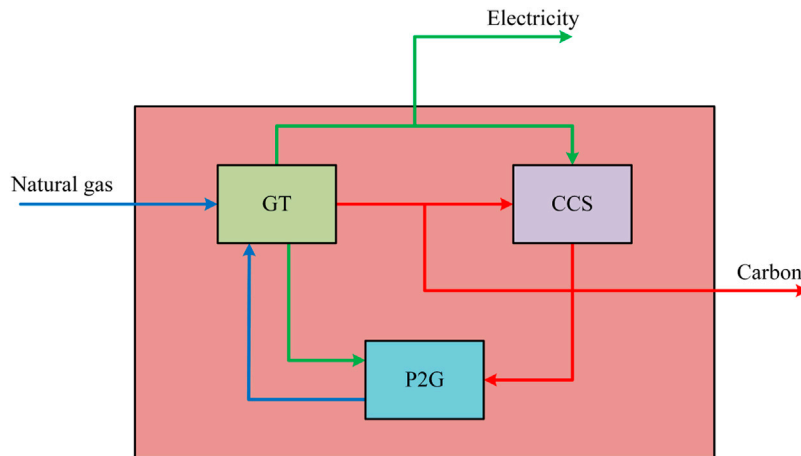


FIGURE 2
GCCP energy flow.

combustion, pre-combustion and post-combustion technologies. Post-combustion technology deals with the CO₂ containing gases produced by conventional fossil fuel plants and achieves the process of separating CO₂ from other gases. Pre-combustion technology, on the other hand, is the process of pre-treating the fuel to sort the carbon in it from other substances. Unlike the first two technologies, oxy-fuel combustion technology changes the environment in which the fuel is burned, allowing it to be burned in an environment containing only oxygen to obtain carbon dioxide and water. Of these, post-combustion is currently the more widely used method, and it is also the most cost-effective of the three CO₂ capture technologies. Sequestration technology begins with the creation of pipelines to transport carbon dioxide, which is then sequestered by compressors.

The electrical energy consumed by the CCS during operation is expressed as Eq. 14.

$$P_{CCS,t} = \zeta_{CCS} * C_{CCS,t}^{CO_2} \quad (14)$$

where $P_{CCS,t}$ represents the power consumed by the CCS to capture CO₂ at time t ; ζ_{CCS} indicates CCS capture efficiency; $C_{CCS,t}^{CO_2}$ denotes the amount of carbon dioxide captured at time t .

2.4 GT-CCS-P2G

Conventionally, carbon capture power plants have operated CCS in combination with thermal power plants or CHP. However, to better reduce carbon emissions, a conventional power plant is replaced by GT coupled with CCS and P2G in this paper, which enhances the coupling of electrical energy flow. In terms of carbon emission, the CCS captures the carbon dioxide emitted by the GT and supplies the P2G with CO₂ to generate methane. In terms of energy supply, the GT unit supplies electricity to P2G and CCS, meanwhile P2G can also supply a small amount of natural gas to the GT. In terms of economic cost, P2G avoids the cost of purchasing CO₂ and CCS reduces purchase electricity from the

main grid. The energy flow route of GT-CCS-P2G (GCCP) is shown in Figure 2.

2.4.1 GCCP operational power

The power consumed by P2G and CCS in the GCCP combined operation model is supplied by GT and the excess power will participate in the power network dispatch which can be calculated as Eq. 15.

$$P_{GCCP,t} = P_{GT,t} - P_{CCS,t} - P_{P2G,t} \quad (15)$$

where $P_{GCCP,t}$ indicates that the GCCP participates in grid dispatch power at time t ; $P_{GT,t}$ denotes the power produced by GT at time t .

The GT, CCS and P2G power constraint can be formulated as Eqs 16–18.

$$P_{GT,min} \leq P_{GT,t} \leq P_{GT,max} \quad (16)$$

$$P_{CCS,min} \leq P_{CCS,t} \leq P_{CCS,max} \quad (17)$$

$$P_{P2G,min} \leq P_{P2G,t} \leq P_{P2G,max} \quad (18)$$

where $P_{GT,min}$ and $P_{GT,max}$ indicates GT operating minimum and maximum power, separately; $P_{CCS,min}$ and $P_{CCS,max}$ represents the minimum and maximum power for capturing CO₂ by CCS, respectively; $P_{P2G,min}$ and $P_{P2G,max}$ denote the minimum and maximum power for P2G operation, separately.

2.4.2 GCCP carbon emission calculation

During operation of the GCCP, GT burning of natural gas releases CO₂, which can be expressed as Eq. 19.

$$C_{GT,t}^{CO_2} = \zeta_{GT}^{CO_2} * P_{GT,t} \quad (19)$$

where $C_{GT,t}^{CO_2}$ denotes the CO₂ produced by the GT operation at time t ; $\zeta_{GT}^{CO_2}$ indicates the carbon emission factor for GT operation.

During GCCP operation, P2G reduces the carbon emissions of the system by capturing CO₂. P2G The synthesis of CH₄ is divided into two main steps: the first step is the electrolysis of water, in which the principle of electrolysis of water is used to convert electrical energy into hydrogen energy in an electrolytic

tank. The second step is methanation, where the hydrogen energy generated is converted to CH₄ and heat energy in a Sabatier reaction with carbon dioxide in a methane reactor which can be calculated as Eq. 20.

$$C_{P2G,t}^{CO_2} = \zeta_{H_2-CO_2} \zeta_{P2G}^{H_2} S_{P2G,t}^{CO_2} \quad (20)$$

where $C_{P2G,t}^{CO_2}$ indicates the amount of CO₂ captured during P2G operation at time t ; $\zeta_{H_2-CO_2}$ denotes the coefficient of conversion between H₂ and CO₂; $\zeta_{P2G}^{H_2}$ represents the efficiency of hydrogen generation by P2G.

In the GCCP coupling model, the CCS captures CO₂ from GT operation simultaneously providing P2G with the CO₂ required for CH₄ production. CH₄. The carbon emissions from the GCCP can be expressed as follow (Eq. 21).

$$C_{GCCP,t}^{CO_2} = C_{GT,t}^{CO_2} - C_{CCP,t}^{CO_2} \quad (21)$$

where $C_{GCCP,t}^{CO_2}$ is the carbon emissions of GCCP at time t .

2.5 Objective function

In consideration of the above-mentioned model, the energy scheduling of the IEGS system is viewed as an optimization problem which involves the minimization of the following objective function (Eq. 22).

$$\min \left\{ \sum_{t=1}^T (C_{CCS,t} + C_{P2G,t} + C_{P,t} + C_{E,t} + C_{Wind,t} + C_{Gas,t}) \right\} \quad (22)$$

where T represents the length of the operating hours; $C_{CCS,t}$ is the cost of CCS at time t ; $C_{P2G,t}$ indicates the cost of P2G at time t ; $C_{P,t}$ denotes the cost of purchasing carbon emissions trading at time t ; $C_{E,t}$ is the cost of purchasing coal at time t ; $C_{Wind,t}$ represents the cost of abandoned wind at time t ; $C_{Gas,t}$ is the cost of acquiring gas at time t .

The electricity consumed by the CCS to collect CO₂ from the gas produced by the thermal power unit through compression and separation mainly consists of fixed energy consumption and operation energy consumption. The detailed calculation is expressed as follows (Eqs 23–26).

$$P_{CCS,t} = P_{CCS}^e + \psi C_{CCS,t}^{CO_2} \quad (23)$$

$$C_{CCS}^e = \sum_{t=1}^T \eta_t^e P_{CCS,t} \Delta t = \sum_{t=1}^T \eta_t^e (P_{CCS}^e + \psi C_{CCS,t}^{CO_2}) \Delta t \quad (24)$$

$$C_{CCS}^r = \frac{C_a (1 - \omega_a)}{365 N_a} \quad (25)$$

$$C_{CCS} = C_{CCS}^e + C_{CCS}^r \quad (26)$$

where P_{CCS}^e represents CCS fixed energy consumption; ψ is the energy coefficient for capturing CO₂; $C_{CCS,t}^{CO_2}$ indicates the capacity to capture CO₂ at time t ; C_{CCS}^e is the cost consumed in the operation of CCS; η_t^e denotes the price of electricity at time t ; C_{CCS}^r represents the depreciated cost of CCS; C_a is the total investment cost of CCS; ω_a represents the depreciation factor of CCS; N_a is the depreciable year of CCS.

Similar to CCS, the cost of P2G can be expressed as (Eqs 27–29).

$$C_{P2G}^e = \sum_{t=1}^T (\eta_t^e P_{P2G,t} \Delta t + \eta^{CO_2} Z_{CO_2} - \eta^{CH_4} E_{CH_4,t}) \quad (27)$$

$$C_{P2G}^r = \frac{C_b (1 - \omega_b)}{365 N_b} \quad (28)$$

$$C_{P2G} = C_{P2G}^e + C_{P2G}^r \quad (29)$$

where C_{P2G}^e is the cost consumed in the operation of P2G; η^{CO_2} denotes the price of CO₂; Z_{CO_2} represents the volume of CO₂ absorbed by P2G; η^{CH_4} is the proceeds from the generation of CH₄; $E_{CH_4,t}$ indicates the total volume of CH₄ produced at time t ; C_{P2G}^r represents the depreciated cost of P2G; C_b is the total investment cost of P2G; ω_b represents the depreciation factor of P2G; N_b is the depreciable year of P2G.

The remaining portion of the costing is shown below (Eqs 30–33).

$$C_P = \sum_{t=1}^T \tau (\zeta P_{electricity,t} - Z_{CO_2,t}) \quad (30)$$

$$C_E = \sum_{t=1}^T (a_E P_{electricity,t}^2 + b_E P_{electricity,t} + c_E) \quad (31)$$

$$C_{Wind} = \sum_{t=1}^T \eta^{Wind} \Delta P_{Wind,t} \quad (32)$$

$$C_{Gas} = \sum_{t=1}^T \eta^{Gas} E_{Gas,t} \quad (33)$$

where τ denotes the coefficient for purchasing carbon emissions; ζ is the CO₂ emission factor for thermal power units; $P_{electricity,t}$ indicates the power purchased from the grid at time t ; $Z_{CO_2,t}$ represents the volume of CO₂ captured by CCS at time t ; a_E , b_E and c_E , is the coefficient of operating costs of thermal power units; η^{Wind} represents the wind discard cost factor; $\Delta P_{Wind,t}$ is the power of the discarded wind; η^{Gas} indicates the price of natural gas; $E_{Gas,t}$ represents the volume of gas consumed by gas network at time t .

2.6 Constraints

The constraints that need to be satisfied during power system operation include power balance, nodal voltage limits and thermal generator output constraints (Eqs 34–37).

$$P_{Load,t} = P_{electricity,t} + P_{Wind,t} - \Delta P_{Wind,t} + P_{GCCP,t} \quad (34)$$

$$V_{min} \leq V_{i,t} \leq V_{max} \quad (35)$$

$$P_{min,t} \leq P_{electricity,t} \leq P_{max,t} \quad (36)$$

$$d_{min,t} \leq \frac{P_{electricity,t} - P_{electricity,t-1}}{t} \leq d_{max,t} \quad (37)$$

where $P_{Load,t}$ denotes the load power at time t ; $V_{i,t}$ represents the voltage at node i at time t ; V_{min} and V_{max} are the upper and lower voltage limits for safe grid operation; $P_{min,t}$ and $P_{max,t}$ denote the upper and lower thresholds for thermal generators, respectively; $d_{min,t}$ and $d_{max,t}$ indicate the upper and lower thresholds of climbing power for thermal generators, separately.

The gas network system consists of three main components: the gas supply source, the gas network and the gas load (Eqs 38, 39).

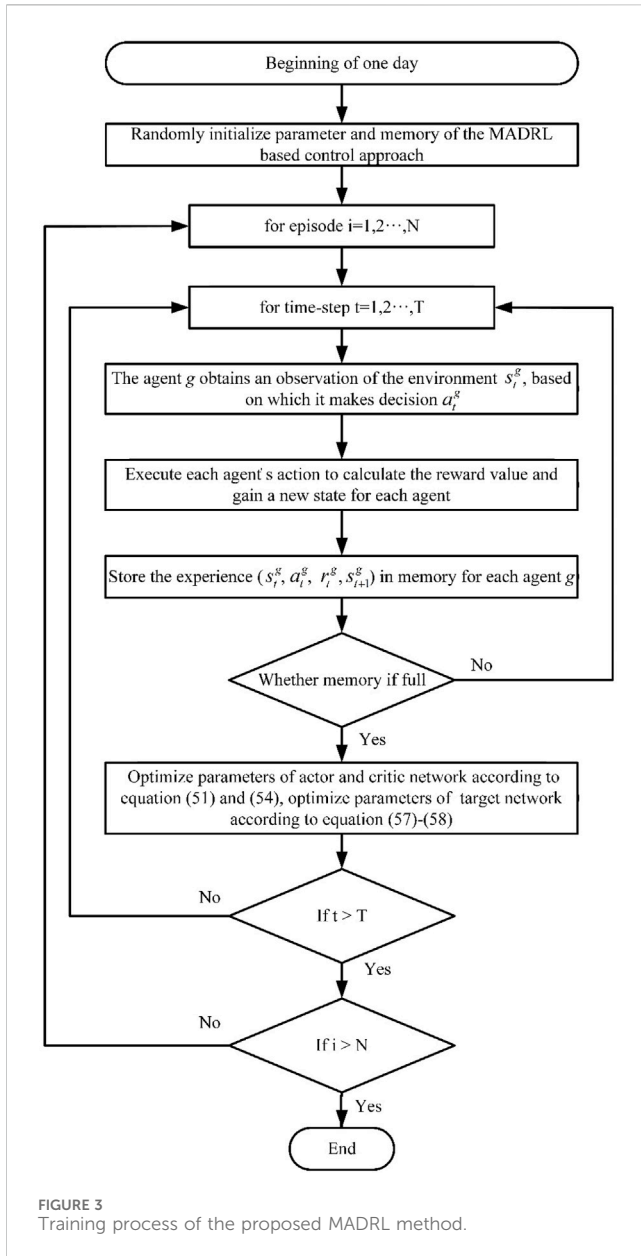


FIGURE 3 Training process of the proposed MADRL method.

$$E_{Gas,a} + E_{P2G,a} = E_{Load,a} + E_{GT,a} \quad (38)$$

$$F_{min} \leq F_t \leq F_{max} \quad (39)$$

where $E_{Gas,a}$ is the injection at node a gas source point; $E_{P2G,a}$ denotes the amount of gas produced by P2G at node a ; $E_{Load,a}$ represents the gas load required by the gas network at node a ; $E_{GT,a}$ is the gas load consumed by GT at node a ; F_{min} and F_{max} denote the upper and lower thresholds for pipeline delivery of natural gas flow, respectively; F_t represents the flow rate conveyed by the pipe at time t .

The constraints for other auxiliary equipment are expressed as follows Eqs 40–41.

$$SOC_{ESS,min} \leq SOC_{ESS,t} \leq SOC_{ESS,max} \quad (40)$$

$$Q_{ESS,min} \leq Q_{ESS,t} \leq Q_{ESS,max} \quad (41)$$

where $SOC_{ESS,min}$ and $Q_{ESS,min}$ indicate the ESS minimum active power and reactive power, respectively; $SOC_{ESS,max}$ and $Q_{ESS,max}$

represent the ESS maximum active power and reactive power, separately; $SOC_{ESS,t}$ and $Q_{ESS,t}$ denote the active and reactive power of ESS at time t , respectively.

3 Proposed control methodology

3.1 Markov game modeling

The energy management optimization problem can be modelled as a Markov game which is solved by the presented MADRL algorithm. The Markov game contains several components (Li et al., 2023).

- **Agent:** In the Markov game, the power grid and the gas grid are respectively modeled as an agent.
- **Environment:** Before each decision, the agents collect information from the nodes in their corresponding region. Each agent makes a decision based on the local observation information which calculates the reward value for each agent based on the decision.
- **State:** The state ensemble $S_t = \{s_t^{grid}, s_t^{gas}\}$ at time t includes all relevant information about the grid and the gas network. The state of the grid agent s_t^{grid} and the state of the gas agent s_t^{gas} are defined as Eqs 42–43:

$$s_t^1 = s_t^{grid} = \{P_{grid,i,t}, \phi_{grid,i,t}, P_{CCS,i,t}, P_{Wind,i,t}, P_{GT,i,t}, SOC_{ESS,i,t}, V_{grid,i,t}, M_{grid,t}\} \quad (42)$$

$$s_t^2 = s_t^{gas} = \{E_{Gas,a,t}, E_{P2G,a,t}, M_{gas,t}\} \quad (43)$$

where $P_{grid,i,t}$ and $\phi_{grid,i,t}$ indicate the active and reactive power demanded by the load at grid node i at time t , respectively; $P_{Wind,i,t}$ denotes the active power injected into node i at time t by WT; $SOC_{ESS,i,t}$ indicates the capacity ratio of the ESS at node i at time t ; $V_{grid,i,t}$ denotes the voltage value of grid at node i at time t ; $M_{grid,t}$ and $M_{gas,t}$ are the price of grid and gas, respectively.

- **Action:** The action ensemble $A_t = \{a_t^{grid}, a_t^{gas}\}$ Each agent performs certain actions based on the observed environmental data, which are denoted respectively as Eqs 44–45.

$$a_t^1 = a_t^{grid} = \{P_{ESS,i,t}, P_{CCS,i,t}, P_{ESS,i,t}\} \quad (44)$$

$$a_t^2 = a_t^{gas} = \{P_{GT,i,t}\} \quad (45)$$

where a_t^{grid} and a_t^{gas} indicate the action of the grid and gas network, respectively.

- **Reward:** The reward value obtained by the system is the value returned when each agent performs an action based on the current state. Each agent shares the same reward function, which is expressed as Eq. 46.

$$r_t = -(l_1 * C_{P,t} + l_2 * (C_{CCS,t} + C_{P2G,t} + C_{E,t} + C_{Wind,t} + C_{Gas,t})) / \hbar \quad (46)$$

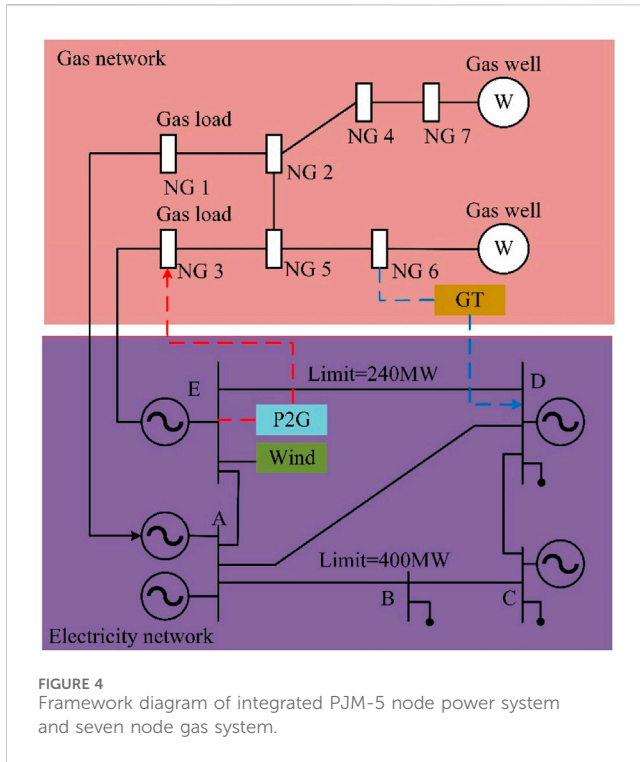


TABLE 1 Electricity price.

Time period	Price (\$/MWh)
0:00–8:00, 22:00–24:00	105.06
8:00–12:00, 18:00–22:00	130.36
12:00–18:00	177.24

TABLE 2 Parameters of the proposed approach.

Parameters	Value
Temperature parameter	0.1
Reward discount factor	0.95
Memory capacity	1e6
Learning rate of actor	1e-3
Learning rate of critic	1e-3
Soft replacement	1e-2
Batch size for updating	256

where l_1 and l_2 denote Weighting coefficients; h indicates a constant.

- **State transition function:** It represents the probability of state transfer to the next step after the agent executes the action. $s_{t+1}^g = F(s_t^g, a_t^g)$, $g = 1, 2$ where $F(\cdot)$ denotes the state transition function.

The energy supply optimization problem of IEGS is transformed into a Markov game, where the grid agent and the gas agent search

for optimal actions by continuously learning the game to attain the best control. In the specific training process, each agent provides the corresponding action by observing part of the state, with the corresponding reward value passed to the agent. While obtaining the reward value, the agent can observe the environment state at the next time. As the number of iterations increases, each agent continuously adjusts its action value through mutual games to maximize the reward value.

3.2 Proposed approach based on MADRL

Each agent has the action-critic network framework which the action network is responsible for strategy program evaluation and the critic network is responsible for strategy parameter updating. Through the interaction and iteration between the two networks, the parameters of the network are continuously updated with the reward value gradually moving towards maximization. The proposed method adopts MASAC (Li X. Y. et al., 2024; Hu et al., 2024) as a kernel, which effectively mitigates the influence of environmental data fluctuations on energy scheduling decisions by sharing environmental and historical information between agents. Each intelligence in MASAC has four deep neural networks, namely, actor network and critic network and target actor network and target critic network. During the training process, only the parameters of the actor network and the critic network are updated, whereas the target actor network and the target critic network are employed to stabilize the learning effect of the actor network and the critic network.

3.2.1 Critic network

The target critic network is mainly employed to mitigate the rate of parameter updates to balance the stability and speed of the training process. It is presented as Eq. 47.

$$Q'(s_t^g, \bar{a}_t^g | \theta') = Q'(s_t^g, \pi_{g'}^g(a_t^g | s_t^g) | \theta') \quad (47)$$

where $\pi_g(\cdot | s_t^g)$ is the value function in the actor network of agent g ; $\pi_{g'}^g(\cdot | s_t^g)$ denotes the function of target actor network; θ' is the parameters of the target actor network; θ' is the parameters of the target critic network; \bar{a}_t^g represents the value passed from the action network of agent g . The computed Q-value is applied to compute the loss function of the criticized network which can be calculated as Eqs 48, 49.

$$h_t = r(s_t^g, a_t^g) + v(Q'(s_t^g, \bar{a}_{t+1}^g | \theta')) \quad (48)$$

$$L(\theta) = \mathbb{E}[(h_t - Q(s_t^g, a_t^g | \theta)) | a_t^g = \pi_g(\cdot | a_t^g)^2] \quad (49)$$

where h_t is the value of Q for the specific situation; $r(s_t^g, a_t^g)$ denotes the total reward obtained by multiple agents performing action a_t^g in global state s_t^g ; v represents the discount factor; θ is the parameters of the critic network; $\mathbb{E}(\cdot)$ indicates the mathematical expectation function.

The gradient of the criticized network parameters $\nabla_{\theta} L(\theta)$ is obtained using the gradient descent method which can be presented as Eqs 50, 51.

$$\nabla_{\theta} L(\theta) = \mathbb{E}[(h_t - Q(s_t^g, a_t^g | \theta)) \nabla_{\theta} Q(s_t^g, a_t^g | \theta)] \quad (50)$$

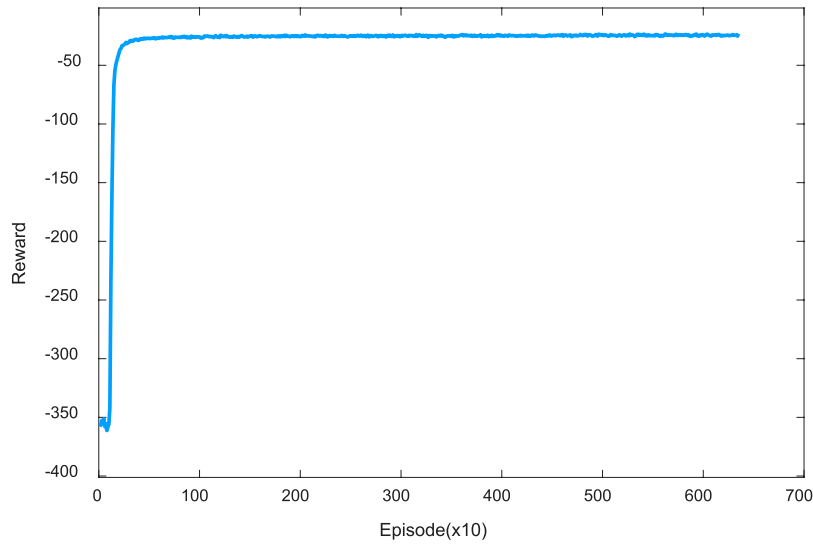


FIGURE 5
Convergence process of proposed method on the train set.

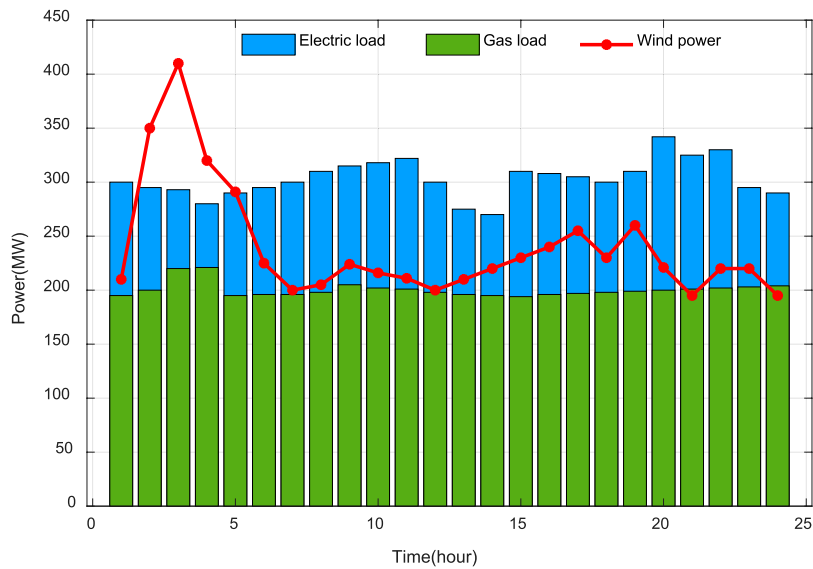


FIGURE 6
Load for a particular day on the test set.

$$\theta_{t+1} \leftarrow \theta_t - \beta_c \nabla_{\theta} L(\theta) \tag{51}$$

where β_c is the learning rate of the critic network; θ_{t+1} denotes the parameters of the critic network at time $t+1$.

3.2.2 Actor network

The expression for the value function in the actor network is Eq. 52:

$$a_t^g = \pi_{\vartheta}(\cdot | s_t^g) \tag{52}$$

where ϑ denotes the parameters of the actor network.

This leads to the gradient function of the actor network value function $\nabla_{\vartheta} L(\vartheta)$ is Eqs 53, 54:

$$\nabla_{\vartheta} L(\vartheta) = \mathbb{E}[\nabla_A Q(s_t^g, a_t^g | \theta) \cdot \nabla_{\vartheta}(\bar{a}_t^g)] \tag{53}$$

$$\vartheta_{t+1} \leftarrow \vartheta_t - \beta_a \nabla_{\vartheta} L(\vartheta) \tag{54}$$

where β_a is the learning rate of the actor network.

In order to prevent the value function in the critic network from overly agreeing with the Q-value calculated by the target value function, a corresponding noise function ζ_t based on a normal distribution is attached to the value passed from the target value

TABLE 3 Comparison results of various approaches on the test day.

Approaches	PSO	MADDPG	MATD3	Proposed approach
Total cost (\$)	457,802.6	434,237.86	429,974.52	422,131.94
The cost of CCS (\$)	9,064.32	9,125.36	9,432.11	10,584.25
The cost of P2G (\$)	30,825.14	31,004.31	31,231.25	32,214.16
The cost of carbon trading (\$)	15,979.9	15,384.45	15,198.9	14,777.1
The cost of thermal power (\$)	220,042.46	206,324.54	204,865.14	202,772.15
The cost of abandoned wind (\$)	6,632.14	6,241.33	6,032.58	5,442.03
The cost of gas source (\$)	175,258.64	166,157.87	163,214.54	156,342.25
Carbon emission (tons)	1,158.66	1,025.63	1,013.26	985.14

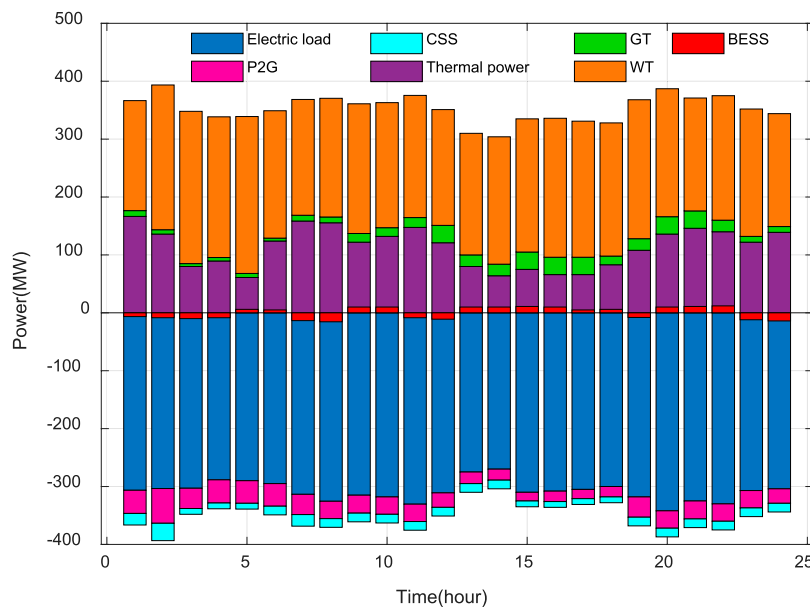


FIGURE 7 Optimization results for electrical system.

function to the critic network in the actor network which can be shown as Eqs 55, 56.

$$\bar{a}_t^g = \pi_{g'}(\cdot | s_t^g) + \zeta_t \tag{55}$$

$$\zeta \sim clip(N(0, \sigma), -z, z) \tag{56}$$

In the training process, the evaluation network is mainly to provide guidance for the actor network to select the optimal action, if the difference between the Q value solved by the critic network and the target value function is huge, the action learned by the actor network will be dispersed and the critic network will be unstable in the learning of the value function. Therefore, in the parameter updating process, the parameters of both the target actor network and the target critic network are updated after a period of training, which

updates their relevant parameters through soft updating as Eqs 57–58 (Li et al., 2023).

$$\theta'_{t+1} \leftarrow \epsilon \theta'_{t+1} + (1 - \epsilon) \theta'_t \tag{57}$$

$$\vartheta'_{t+1} \leftarrow \epsilon \vartheta'_{t+1} + (1 - \epsilon) \vartheta'_t \tag{58}$$

where ϵ is the soft update factor, which has a value much less than 1.

The detailed flowchart of the MADRL algorithm proposed in this paper is shown in Figure 3.

4 Case study

In this session, the parameters of the IEGS and the proposed algorithm are firstly described, followed by an

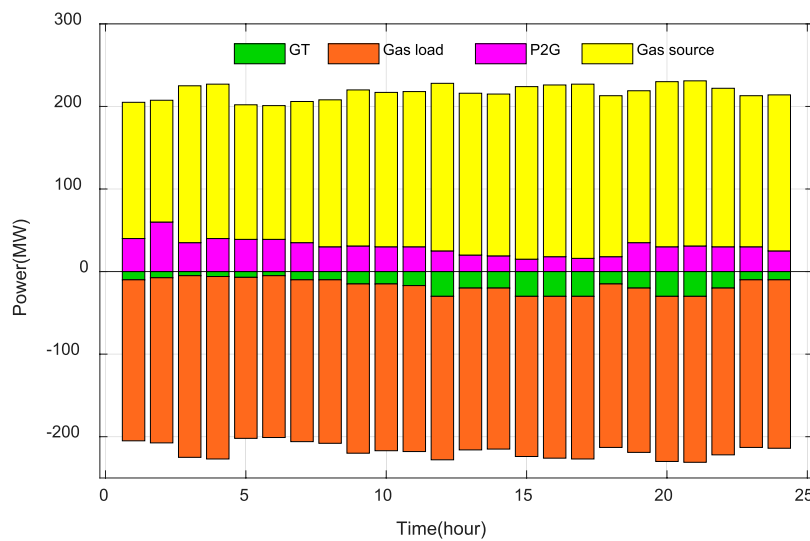


FIGURE 8 Optimization results for gas system.

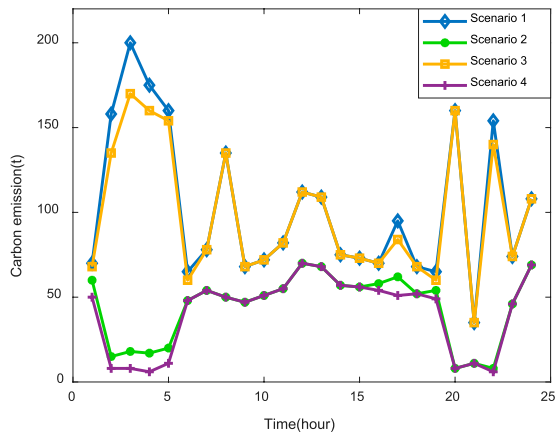


FIGURE 9 Comparison of carbon emission.

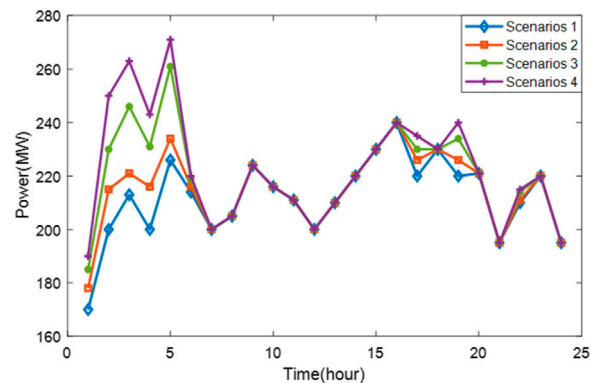


FIGURE 10 Comparison of wind power consumption.

example to assess the effectiveness and superiority of the proposed approach.

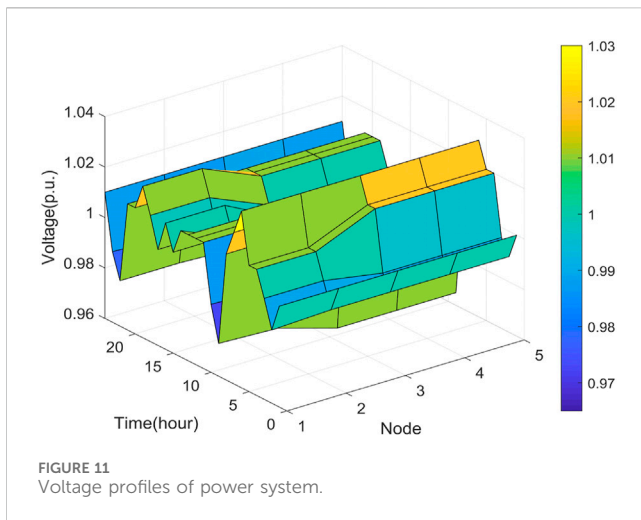
4.1 Case study setup

In order to effectively evaluate the performance of the proposed scheme, the integrated PJM-5 node system (Li et al., 2017) and seven nodes natural gas system (Li et al., 2008) are selected for experimental analysis, where P2G is connected to the WT at PJM-5 node E, which decreases the wind abandonment rate of the WT and the P2G simultaneously achieves the CH₄ transmission through node three of the gas system. GT achieves the conversion between gas and electricity by connecting to the PJM-5 node D and the gas system node 6. The specific system architecture is shown in

Figure 4. The electricity price is divided into three different prices as shown in Table 1, where the electricity price for 0:00–8:00 and 22:00–24:00 is \$105.06/MW, for 8:00–12:00 and 18:00–22:00 is \$130.36/MW and for the remaining hours is \$177.24/MW. The price of gas sold from the two wells is 78.39\$/MW. The price for purchasing carbon credits in the carbon market is 15\$/ton. Detailed parametric data of IEGS can be found in (Li et al., 2023). Parameters of the proposed approach are shown in Table 2.

4.2 Evaluation of the proposed control model

In order to verify the effectiveness and advancement of the proposed model, the following methods are selected for comparative analysis. The reward variation of the proposed method during the training process is shown in Figure 5, where the performance of the



proposed method is evaluated by the variation of the reward value. Since the parameters of the action neural network are randomly initialized at the beginning of the training process, the agent is unaware of how to make decisions to reduce the total operating cost. Therefore, the agent chooses to explore the environment to gain more experience. The experience gained from the pre-training is stored in the experience pool to optimize the control strategy with the experience replay mechanism. Each iteration step samples a certain amount of historical training data from the experience pool for updated parameters of the action and critic neural network. As can be seen from the figure, the cumulative reward earned by the agent gradually increases during the process. At the end of 100 stochastic optimizations, the reward rises rapidly, and the curve starts to converge when the training reaches about 1,000 times.

After several training sessions, the proposed model has acquired the ability to cogitate for optimal decision making in new environments. One particular day of data is chosen for the validation analysis, with specific information displayed in Figure 6. As can be seen, the power of the WT is higher in the early hours of the morning, while the electrical loads are smaller, leaving the system struggling to completely dissipate all the wind power. During the period 5:00–24:00, the power of WT has been less than the electrical load.

In this paper, three algorithms, particle swarm optimization (PSO) (Du et al., 2023), multi-agent deep deterministic policy gradient (MADDPG) (Abid et al., 2024), and multi-agent twin delayed deep deterministic policy gradients (MATD3) (Wang et al., 2023b) are selected for comparative validation as a way to verify the reliability and robustness of the proposed approach. The specific results of the proposed method with other comparison schemes on the test set are shown in Table 3. PSO performs the worst in the face of complex IEGS such that the optimal scheduling scheme is not obtained. Compared to the PSO, MADDPG selects multiple agents for optimal scheduling. Coordinated management between energy sources is achieved using communication between agents to obtain better performance. MATD3 adds two sub-networks agents for Q-value estimation to solve the problem due to Q-value overestimation, which is a further improvement compared to MADDPG. The proposed scheme adopts MASAC as the kernel, which expands the stochasticity of the scheduling process by increasing the entropy function, to obtain the optimal scheduling strategy. The proposed approach represents a total cost

reduction of 35,670.66\$ and a reduction in carbon emission of 173.52 tons compared to PSO.

The detailed outputs of the electrical system and the gas system in the proposed method on a particular day of the test set are presented in Figures 7, 8 respectively. From Figure 7, it can be observed that during the period 0:00–8:00, due to the lower electricity price and higher wind power output, P2G consumes more electricity for CH₄ production, which reduces the CO₂ release from the system while consuming wind power, and the GT is almost inactive at this time. During this period the BESS starts charging and the CCS is capturing carbon to consume electricity. During 8:00–12:00, as the price of electricity increases, the power consumed by the P2G and CCS starts to decrease, the GT gradually starts to work, and the BESS releases the stored power. During the period 12:00–18:00, when the tariff reaches its maximum value, the GT reaches its maximum power to achieve gas to power conversion, which reduces the cost of purchasing electricity for the system. It can also be seen from Figure 8 that the proposed model chooses to increase the power from electricity to gas when the electricity price is low. In contrast, when the electricity price is high, the proposed model chooses to increase the power of gas-to-electricity conversion.

4.3 Evaluation of the proposed GCCP model

In order to verify the validity of the GCCP model, this paper constructs four scenarios for simulation analysis. The economic scheduling strategy in scenario one does not consider CCS and P2G. The economic scheduling strategy in scenario two considers only CCS. The economic scheduling strategy in scenario three considers mainly P2G. The economic scheduling strategy in scenario four introduces the GCCP proposed in this paper.

The comparison of CO₂ emissions under different scenarios is displayed in Figure 9. Comparing the carbon emissions of scenario one and scenario 2, it can be observed that the carbon capture device

can significantly reduce the CO₂ emissions of the IEGS, which is about 53.4% of the total emissions, with a total of 1,210 tons of CO₂ reduced in scenario two compared with scenario 1. By comparing scenario one and scenario 3, the CO₂ emissions of the system are almost unchanged because the CO₂ demanded for P2G is purchased from an external source. The higher carbon emissions in scenario one than in scenario three in the 0:00–6:00 interval are due to the conversion of excess wind energy achieved through P2G, which reduces the amount of gas purchased online to reduce carbon emissions. When the GCCP model was introduced in scenario 4, the carbon emissions at each hour are significantly reduced compared to the other scenarios, with a reduction of 1,476 tons compared to scenario 1.

The wind power output under different scenarios is shown in Figure 10. Comparing scenario 2 with scenario 1, which is the peak period of wind power output from 1:00 to 5:00, it can be observed that there is a partial improvement in wind power consumption after the use of the CCS device. Comparing scenario one and scenario 3, it is clear that the P2G device can significantly increase the wind power output, and the wind power generated in scenario three does not achieve the maximum value due to the maximum input power of the device. In scenario 4, the GCCP coupling device is adopted to significantly enhance the wind power consumption, which reaches 92.81%.

The voltage values of the proposed approach at each moment of each node on a certain day of the test set are shown in Figure 11, from

which it can be seen that although the voltage values of all nodes before and after the moment have a large volatility, the voltage has been located in the range of [0.96, 1.04]. It conforms to the requirement of stability and security operation of the system, which again proves the effectiveness of the proposed approach in voltage control.

5 Conclusion

In this study, a MADRL-based IEGS scheduling approach considering GCCP with simultaneous consideration of system security and economy is proposed. MADRL replicates historical data to address the negative impacts caused by time series data with efficient exploratory techniques for seeking optimization. Agents seek optimal control strategies by continuously interacting with each other with information. The detailed conclusions of the study are summarized as follows: (1) Compared to several other methods, the proposed controlling framework and approach provide the best performance. (2) The explainability of the properties in detail in the IEGS provides additional evidence of validation for the proposed controlling framework. (3) The effectiveness of the proposed GCCP model is verified through four different scenarios, reducing carbon emissions by 1,476 tons and increasing the proportion of wind power consumption by 4.41% compared to Scenario 1.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

WF: Conceptualization, Data curation, Writing—original draft, Writing—review and editing. BD: Conceptualization, Data curation, Investigation, Writing—original draft. ZZ: Data curation, Investigation, Methodology, Writing—original draft. HJ: Project

administration, Supervision, Validation, Writing—review and editing. YZ: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Writing—review and editing. XP: Data curation, Formal Analysis, Investigation, Methodology, Software, Writing—review and editing. LZ: Conceptualization, Methodology, Software, Writing—original draft. ZJ: Methodology, Project administration, Software, Validation, Writing—original draft, Writing—review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Science and Technology Project of State Grid Sichuan Electric Power Company (B7194723R001).

Conflict of interest

Authors WF, BD, ZZ, HJ, YZ, XP, and LZ were employed by State Grid Sichuan Information and Telecommunication Company.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The authors declare that this study received funding from State Grid Sichuan Electric Power Company. The funder had the following involvement in the study: study design, data collection and analysis, decision to publish.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Abid, M. S., Apon, H. J., Hossain, S., Ahmed, A., Ahshan, R., and Lipu, M. S. H. (2024). A novel multi-objective optimization based multi-agent deep reinforcement learning approach for microgrid resources planning. *Appl. Energy* 353, 122029. doi:10.1016/j.apenergy.2023.122029
- Bai, L. Q., Li, F. X., Cui, H. T., Jiang, T., Sun, H. B., and Zhu, J. X. (2016). Interval optimization based operating strategy for gas-electricity integrated energy systems considering demand response and wind uncertainty. *Appl. Energy* 167, 270–279. doi:10.1016/j.apenergy.2015.10.119
- Barker, P. P., and Mello, R. W. (2000). Determining the impact of distributed generation on power systems. I. Radial distribution systems. *2000 Power Eng. Soc. Summer Meet.*, 1645–1656.
- Cao, D., Hu, W. H., Zhao, J. B., Huang, Q., Chen, Z., and Blaabjerg, F. (2020a). A multi-agent deep reinforcement learning based voltage regulation using coordinated PV inverters. *IEEE Trans. Power Syst.* 35, 4120–4123. doi:10.1109/tpwrs.2020.3000652
- Cao, D., Hu, W. H., Zhao, J. B., Zhang, G. Z., Zhang, B., Liu, Z., et al. (2020b). Reinforcement learning and its applications in modern power and energy systems: a review. *J. Mod. Power Syst. Clean Energy* 8 (6), 1029–1042. doi:10.35833/mpce.2020.000552
- Cao, D., Zhao, J. B., Hu, J. X., Pei, Y. S., Huang, Q., Chen, Z., et al. (2023). Physics-informed graphical representation-enabled deep reinforcement learning for robust distribution system voltage control. *IEEE Trans. Smart Grid* 15 (1), 233–246. doi:10.1109/tsg.2023.3267069
- Cao, D., Zhao, J. B., Hu, W. H., Huang, Q., Chen, Z., Blaabjerg, F., et al. (2021). Data-driven multi-agent deep reinforcement learning for distribution system decentralized voltage control with high penetration of PVs. *IEEE Trans. Smart Grid* 12, 4137–4150. doi:10.1109/tsg.2021.3072251
- Clegg, S., and Mancarella, P. (2015). Integrated modeling and assessment of the operational impact of Power-to-Gas (P2G) on electrical and gas transmission networks. *IEEE Trans. Sustain. Energy* 6, 1234–1244. doi:10.1109/tste.2015.2424885
- Dai, W., Yu, J., Yang, Z. F., Huang, H. Y., Lin, W., and Li, W. Y. (2020). A static equivalent model of natural gas network for electricity-gas co-optimization. *IEEE Trans. Sustain. Energy* 11, 1473–1482. doi:10.1109/tste.2019.2927837
- Ding, L. F., Cui, Y. K., Yan, G. F., Huang, Y. J., and Fan, Z. (2024). Distributed energy management of multi-area integrated energy system based on multi-agent deep reinforcement learning. *Int. J. Electr. Power and Energy Syst.* 157, 109867. doi:10.1016/j.ijepes.2024.109867
- Du, W. Y., Ma, J., and Yin, W. J. (2023). Orderly charging strategy of electric vehicle based on improved PSO algorithm. *Energy* 271, 127088. doi:10.1016/j.energy.2023.127088

- Dulua, L., Abrudeanb, M., and Bicac, D. (2013). Effects of distributed generation on electric power systems. *7th Int. Conf. Interdiscip. Eng.*, 1–5.
- Elsayed, A. A. H., Nashwan, M. S., Eltahan, A. M. H., and Shahid, S. (2024). Millions more Egyptians will be exposed to drought by 2100 under the goals of the Paris climate agreement. *Int. J. Disaster Risk Reduct.* 101, 104257. doi:10.1016/j.ijdrr.2024.104257
- Gallo, C., Faccilongo, N., and Sala, P. L. (2018). Clustering analysis of environmental emissions: a study on Kyoto Protocol's impact on member countries. *J. Clean. Prod.* 172, 3685–3703. doi:10.1016/j.jclepro.2017.07.194
- Gao, X., and Zhang, X. (2024). Robust collaborative scheduling strategy for multi-microgrids of renewable energy based on a non-cooperative game and profit allocation mechanism. *Energies* 17, 519. doi:10.3390/en17020519
- Gu, W., Wang, J., Lu, S., Luo, Z., and Wu, C. Y. (2017). Optimal operation for integrated energy system considering thermal inertia of district heating network and buildings. *Appl. Energy* 199, 234–246. doi:10.1016/j.apenergy.2017.05.004
- He, J. M., Wu, Y. N., Yong, X. K., Tan, Q. L., and Liu, F. T. (2022). Bi-level optimization of a near-zero-emission integrated energy system considering electricity-hydrogen-gas nexus: a two-stage framework aiming at economic and environmental benefits. *Energy Convers. Manag.* 274, 116434. doi:10.1016/j.enconman.2022.116434
- Hu, D. E., Li, Z. C., Ye, Z. H., Peng, Y. G., Xi, W., and Cai, T. T. (2024). Multi-agent graph reinforcement learning for decentralized Volt-VAR control in power distribution systems. *Int. J. Electr. Power and Energy Syst.* 155, 109531. doi:10.1016/j.ijepes.2023.109531
- Ji, Z., Kang, C. Q., Chen, Q. X., Xia, Q., Jiang, C. M., Chen, Z. X., et al. (2013). Low-carbon power system dispatch incorporating carbon capture power plants. *IEEE Trans. Power Syst.* 28, 4615–4623. doi:10.1109/tpwrs.2013.2274176
- Li, G. Q., Zhang, R. F., Jiang, T., Chen, H. H., Bai, L. Q., Cui, H. T., et al. (2017). Optimal dispatch strategy for integrated energy systems with CCHP and wind power. *Appl. Energy* 192, 408–419. doi:10.1016/j.apenergy.2016.08.139
- Li, S. C., Hu, W. H., Cao, D., Chen, Z., Huang, Q., Blaabjerg, F., et al. (2023). Physics-model-free heat-electricity energy management of multiple microgrids based on surrogate model-enabled multi-agent deep reinforcement learning. *Appl. Energy* 346, 121359. doi:10.1016/j.apenergy.2023.121359
- Li, S. C., Hu, W. H., Cao, D., Dragičević, T., Huang, Q., Chen, Z., et al. (2024c). Electric vehicle charging management based on deep reinforcement learning. *J. Mod. Power Syst. Clean Energy* 10 (2), 719–730. doi:10.35833/mpce.2020.000460
- Li, S. C., Hu, W. H., Cao, D., Hu, J. X., Huang, Q., Chen, Z., et al. (2024a). A novel MADRL with spatial-temporal pattern capturing ability for robust decentralized control of multiple microgrids under anomalous measurements. *IEEE Trans. Sustain. Energy*, 1–12. doi:10.1109/tste.2024.3386180
- Li, T., Eremia, M., and Shahidehpour, M. (2008). Interdependency of natural gas network and power system security. *IEEE Trans. Power Syst.* 23 (4), 1817–1824. doi:10.1109/tpwrs.2008.2004739
- Li, X. Y., Luo, F. G., and Li, C. J. (2024b). Multi-agent deep reinforcement learning-based autonomous decision-making framework for community virtual power plants. *Appl. Energy* 360, 122813. doi:10.1016/j.apenergy.2024.122813
- Li, Z. M., Liu, J. Y., Xiang, Y., Zhang, X., and Chai, Y. X. (2019). Agricultural load modeling based on crop evapotranspiration and light integration for economic operation of greenhouse power systems. *CSEE J. Power Energy Syst.* 7 (5), 1113–1121.
- Liu, J. B., Su, Y. C., Xiang, Y., Liu, J. Y., Wang, L. F., and Xu, W. T. (2018). Operational reliability assessment for gas-electric integrated distribution feeders. *IEEE Trans. Smart Grid* 10 (1), 1091–1100. doi:10.1109/tsg.2018.2844309
- Liu, T. H., Tian, J., Zhu, H. Y., Goh, H. H., Liu, H., Wu, T., et al. (2023). Key technologies and developments of multi-energy system: three-layer framework, modelling and optimisation. *Energy* 277, 127697. doi:10.1016/j.energy.2023.127697
- Liu, W. J., Chi, W. Y., and Hua, W. Q. (2024). Blockchain-enabled renewable energy certificate trading: a secure and privacy-preserving approach. *Energy* 290, 130110. doi:10.1016/j.energy.2023.130110
- Ma, Z., Jørgensen, B. N., and Ma, Z. (2024). A scoping review of energy-efficient driving behaviors and applied state-of-the-art AI methods. *Energies* 17, 500. doi:10.3390/en17020500
- Monfaredi, F., Shayeghi, H., and Siano, P. (2023). Multi-agent deep reinforcement learning-based optimal energy management for grid-connected multiple energy carrier microgrids. *Int. J. Electr. Power and Energy Syst.* 153, 109292. doi:10.1016/j.ijepes.2023.109292
- Okedu, K. E., Oyinnu, B., Colak, I., and Kalam, A. (2024). Geographical information system based assessment of various renewable energy potentials in Nigeria. *Energy Rep.* 11, 1147–1160. doi:10.1016/j.egyr.2023.12.065
- Tabebordbar, A., Rastegar, M., and Ebrahimi, M. (2023). Reliability-oriented optimal sizing of power-to-gas and combined heat and power technologies in integrated electricity and natural gas transmission systems. *Sustain. Cities Soc.* 95, 104593. doi:10.1016/j.scs.2023.104593
- Wang, X., Wang, S. X., Zhao, Q. Y., and Lin, Z. R. (2023a). Low-carbon coordinated operation of electric-heat-gas-hydrogen interconnected system and benchmark design considering multi-energy spatial and dynamic coupling. *Energy* 279, 128042. doi:10.1016/j.energy.2023.128042
- Wang, X., Zhou, J. S., Qin, B., and Guo, L. Z. (2023b). Coordinated control of wind turbine and hybrid energy storage system based on multi-agent deep reinforcement learning for wind power smoothing. *J. Energy Storage* 57, 106297. doi:10.1016/j.est.2022.106297
- Wang, Z. X., Xiao, F., Ran, Y., and Xu, Y. (2024). Scalable energy management approach of residential hybrid energy system using multi-agent deep reinforcement learning. *Appl. Energy* 367, 123414. doi:10.1016/j.apenergy.2024.123414
- Werksman, J. (1994). The United Nations framework convention on climate change: the first conference of the parties opening in Berlin. *Glob. Environ. Change* 4 (4), 339–340. doi:10.1016/0959-3780(94)90034-5
- Wojtaszek, H., Miciuła, I., Modrzejewska, D., Stecyk, A., Sikora, M., Wójcik-Czerniawska, A., et al. (2024). Energy policy until 2050—comparative analysis between Poland and Germany. *Energies* 17, 421. doi:10.3390/en17020421
- Yang, J. W., Zhang, N., Cheng, Y. H., Kang, C. Q., and Xia, Q. (2019). Modeling the operation mechanism of combined P2G and gas-fired plant with CO2 recycling. *IEEE Trans. Smart Grid* 10, 1111–1121. doi:10.1109/tsg.2018.2849619
- Zhang, B., Hu, W. H., Ghias, A. M. Y. M., Xu, X., and Chen, Z. (2023b). Two-timescale autonomous energy management strategy based on multi-agent deep reinforcement learning approach for residential multicarrier energy system. *Appl. Energy* 351, 121777. doi:10.1016/j.apenergy.2023.121777
- Zhang, B., Hu, W. H., Xu, X., Zhang, Z. Y., and Chen, Z. (2023a). Hybrid data-driven method for low-carbon economic energy management strategy in electricity-gas coupled energy systems based on transformer network and deep reinforcement learning. *Energy* 273, 127183. doi:10.1016/j.energy.2023.127183
- Zhang, B., Xu, N., Zhang, H. R., Qiu, R., Wei, X. M., Wang, Z., et al. (2024a). Influence of hydrogen blending on the operation of natural gas pipeline network considering the compressor power optimization. *Appl. Energy* 358, 122594. doi:10.1016/j.apenergy.2023.122594
- Zhang, W. W., Chen, X. M., and Tian, J. (2024b). The evolutionary analysis of investment in CCS-EOR under dual carbon target—from the perspective of multi-agent involvement. *Int. J. Greenh. Gas Control* 135, 104107. doi:10.1016/j.ijggc.2024.104107
- Zhang, Y. C., Zheng, F., Shu, S. W., Le, J., and Zhu, S. (2020). Distributionally robust optimization scheduling of electricity and natural gas integrated energy system considering confidence bands for probability density functions. *Int. J. Electr. Power and Energy Syst.* 123, 106321. doi:10.1016/j.ijepes.2020.106321
- Zhong, Z. Q., Chen, Y. Q., Fu, M. Y., Li, M. Z., Yang, K. S., Zeng, L. P., et al. (2023). Role of CO2 geological storage in China's pledge to carbon peak by 2030 and carbon neutrality by 2060. *Energy* 272, 127165. doi:10.1016/j.energy.2023.127165