Check for updates

Coordinated Control Strategy of Electricity-Heat-Gas Integrated Energy System Considering Renewable Energy Uncertainty and Multi-Agent Mixed Game

Yan Jing[1]*[†], Duan Zhiqiang[2†], Gao Jiemai[3†], Chen Siyuan[3†], Zhou Bing[1†] and Wang Yajie[2†]

[1]State Grid Hubei Electric Power Co., Ltd., Wuhan, China, [2]State Grid Hubei Electric Power Co., Ltd., Economic and Technological Research Institute, Wuhan, China, [3]School of Electrical and Automation, Wuhan University, Wuhan, China

The Integrated Energy System (IES) can promote social energy transformation and low-carbon development, which is also an effective means to make contributions to energy structure optimization, energy consumption reduction, and new energy consumption. However, the IES has the characteristics of complex energy flow, and strong uncertainty with multi-agents. Therefore, traditional mathematical optimization models are difficult to comprehensively and accurately reflect the interest needs of different entities in the integrated energy microgrid. Aiming at this problem, a two-level collaborative control strategy model of "electricity-heat-gas" IES based on multi-agent deep reinforcement learning is proposed in this paper. The upper layer of this model is a multi-agent hybrid game decision-making model based on the Multi-Agent Deep Deterministic Policy Gradient algorithm (MADDPG), and the lower layer contains the power and gas flow calculation model. The lower model provides the upper model with the energy flow data of the IES and the upper layer rewards the decision-making behavior of the agent based on the energy flow data provided by the lower layer. Effectively solving the high-dimensional nonlinear optimization problem existing in the complex coupling network, this method can improve the convergence and training speed of the model. In this paper, the IEEE 33-node distribution network and 20-node gas network coupling system are provided to verify the model. The simulation results show that the proposed collaborative control strategy method can provide effective decision-making for electric-agent and gas-agent and realize the efficient and economic operation of the integrated energy system.

Keywords: integrated energy system, mixed power flow model, dynamic control, deep reinforcement learning, energy management system

1 INTRODUCTION

With the trend of low-carbon, cleanness and sustainability of electric industry going further, the energy structure of all countries in the world is gradually realizing electric energy substitution and clean energy substitution. China's "carbon peaking and carbon neutrality (dual carbon)" goals are also important to the international community (Ali et al., 2021). Effective response to green and low-carbon development. The Integrated Energy System (IES) runs through the coordinated operation of

electric and thermal energy, which can effectively improve energy utilization, and is of great significance for promoting the consumption of renewable energy and achieving the "dual carbon" goal (Mohamed, 2022). On the energy side, wind power, photovoltaic, natural gas and other clean energy access the grid and will increase the uncertainty of power supply, heating and other aspects (Rezaei et al., 2021; Xia et al., 2021). Meanwhile, on the load side, new types of loads such as electric vehicles, smart homes and other forms of energy loads can deeply participate in energy interactions and also increase new uncertainties in the demand side (Al-Ghussain et al., 2021; Lan et al., 2021). Moreover, the integrated energy system integrates coal, oil, natural gas, electricity, heat and other energy sources in the region (Tan et al., 2022). At the same time, the existing market investment body is increasing. This means that in the existing market environment in China, grid companies have changed from being the dominant player in the construction of energy networks to being an important player in the construction of energy networks (Sun et al., 2015; Ma et al., 2021). Unlike the traditional integrated energy synergy from the overall perspective of decision-making, the integrated energy system is oriented to multiple competing subjects, and there is a significant game relationship between different subjects (Weinstein, 2010). In this background, it is important to study and propose an integrated energy system control strategy that takes into account the game relationship of diverse investment agents.

The problem of integrated energy system regulation has been studied by scholars in China and abroad and certain results have been achieved (Abdelaziz Mohamed and Eltamaly, 2018a). In order to solve the energy station-grid configuration problem, Huang Wei et al. (Huang and Liu, 2020) proposed an integrated energy station-grid double-layer planning optimization model for parks considering the multi-energy complementary characteristics of energy stations. The model effectively alleviates the problem of under-supply or over-supply caused by zonal supply and peak-to-valley demand. The authors in (Yao and Wang, 2020) propose a two-level collaborative optimal allocation method for integrated energy systems considering wind and solar uncertainty. The approach improves both total annualized system cost and average annual equipment utilization. The article verifies the effectiveness and economy of the proposed method. Walter et al. (2020) proposed the integration of thermal energy systems into microgrid energy management systems. He builds models based on consideration of fuel costs, thermal comfort and other factors. Moreover, its economic efficiency is fully considered the optimal solution to achieve the purpose of reducing the cost of electricity and heating. All of the above literature uses a game theory and traditional centralized method approach to regulate aspects such as power system generation, and grid structure. The disadvantage of above approach is that it is difficult to allocate the power system dynamically and also having the problem of solving quickly.

It is worth noting the close coupling of different energy networks such as electricity and natural gas in the integrated energy system (Abdelaziz Mohamed and Eltamaly, 2018b). The problem of solving control strategies for integrated energy

systems is a nonlinear, non-convex optimization problem. Therefore, it is difficult to solve the problem by relying only on traditional mathematical modeling methods (Mohamed et al., 2020).

In recent years, Reinforce Learning (RL) has made significant breakthroughs in data resolution, learning power and computational power. RL has been applied to smart manufacturing, smart medical and other fields. In terms of application results, it has shown good application effects (Sutton and Barto, 1998; Leo Kumar, 2017; Park and Han, 2018). Nowadays, some researchers have applied multi-agent and game theory to energy power systems. These researches make use of the characteristics of multi-agent such as autonomy, interactivity and distributed computing properties and analyze the interesting pursuit of different agents and the possible interest equilibrium relationship through game theory for economic scheduling and energy management (Makhadmeh et al., 2021). In (Yang et al., 2021), a dynamic economic dispatch for integrated energy systems based on the DDPG algorithm was proposed. And this method can achieve dynamic economic scheduling of the system better than traditional methods, but this mechanism uses uniform random sampling to extract empirical data from the experience pool without considering the importance of different experiences. The authors in (Nie et al., 2021) proposed a double-layer reinforcement learning model to achieve real-time economic dispatch of IES with higher solving efficiency. The authors in (Liu et al., 2019) divide the microgrid system into multi-agent, establishes a multi-agent game coordination scheduling model, and proposes an integrated energy microgrid scheduling method based on the Nash game and Q-learning algorithm. However, this model does not take into account the end-load demand. In (Qiao et al., 2021), the authors present a deep reinforcement learning approach based on a soft actor-critic framework for the problem of optimizing the operation of integrated electricity-gas energy systems. The method can realize the continuous action control of a multi-energy flow system, and can flexibly handle the source-load uncertainty of wind power, photovoltaic, multi-energy load, etc. However, the thermal energy in the integrated energy system is not considered. In (Liu et al., 2020), an energy management and optimization method for micro-energy networks based on deep reinforcement learning is proposed. The method uses an experience replay mechanism and fixed network parameters mechanism, which effectively solves the problems of difficult modeling of integrated energy systems, slow convergence of traditional algorithm operation and difficulty in meeting real-time optimization requirements as well as system openness. Reference (Ying et al., 2019) proposes a new energy management approach for real-time dispatch considering load demand, renewable energy and tariff uncertainty. The method adaptively adjusts based on trends in marginal prices and netload, and develops cost-effective dispatch schemes for microgrids in uncertain environments. Reference (Nie et al., 2021) combines deep reinforcement learning with a realistic peer-to-peer (P2P) energy trading model to solve the decision-making problem for microgrids in local energy markets. This strategy improves microgrid utilization and uses virtual

penalties to reduce inefficient power plant generation points and to determine the optimal battery capacity for the microgrid. However, there are few studies on integrated energy systems with electricity, heat and gas as the main game (Fan et al., 2021). In the case that all three energy subjects can provide energy in the Integrated Energy System, further research is needed on how to rationally allocate energy to reduce the cost and energy waste of the Integrated Energy System, improve the energy utilization rate, and meet the energy supply and demand and the interest game behavior of the three subjects while responding to the low-carbon economy.

Aiming at the difficulty solving problem of the traditional modelling methods are difficult to solve the problem, this paper model the multi-agent game process, applies DRL to the multi-agent game process of IES, and simulates the multi-agent systemization based on the MADRL algorithm, which at the same time, may contribute to further reduce the cost of the Integrated Energy System, optimize the energy efficiency and meet the daily load. To be specific, considering the game relationship of diversified investment entities in the IES, this article takes the electric-heat-gas integrated energy system as the object and constructs a multi-agent strategy model based on DRL. The main contributions of this article are as follows:

(1) Considering different kinds of energy equipment and their coupling relations in the integrated energy system, the mathematical model of the integrated energy system is constructed;
(2) Based on the theory of DRL, the dynamic game process and decision model of interests multi-energy players of electricity, heat and gas are constructed according to the decisions of different players and their respective benefits;
(3) A two-layer model is established. The lower layer decides the output of electricity, heat and gas networks, and the upper layer model calculates the equilibrium output of the Integrated Energy System based on the output decision of the lower layer and the constraints of multi-power flow;
(4) The reliability and validity of the proposed model is proved by the analysis of calculation examples.

The rest of this paper is organized as follows. **Section 2** introduced the structure of Integrated Energy System and multi-agent game theory; In **Section 3**, mathematical model of Integrated Energy System is discussed in detail, including the constraints. In **Section 4**, the two-level model is introduced and proposed the setting method to a DRL model; **Section 5** presents the numerical simulation results. Finally, conclusions and suggestions for future work are drawn in **Section 5**

2 INTEGRATED ENERGY SYSTEM STRUCTURE AND MULTI-AGENT GAME

In this paper, the IES is constructed by combining electricity/gas/heat energy sources. It is based on a typical industrial park and includes Renewable Energy (RE) Combined Heat and Power (CHP), Gas Turbine (GT), Gas Boiler (GB), Electric Boiler

FIGURE 1 | Integrated energy system.
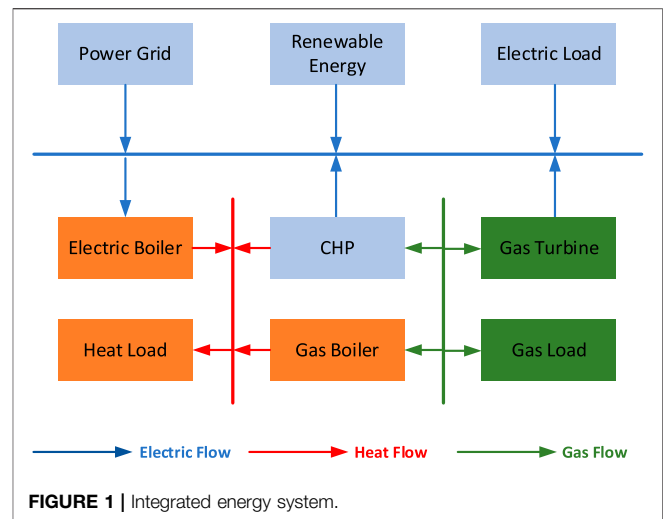
(EB) and other equipment. The system structure is shown in **Figure 1**.

2.1 Integrated Energy System Structure
In this integrated energy park system, electricity and natural gas can be purchased from external sources; Conversion between electricity and gas can be done by combined heat and power units and gas turbines; Conversion between electricity and heat through combined heat and power units and electric boilers; Conversion between heat and gas *via* combined heat and power units and gas boilers.

On this basis, this paper takes the three networks that exist in the system as different subjects of interest. The structure divides IES into three investment entities: electricity (main grid, renewable energy, combined heat and power units), heat (gas boilers, electric boilers) and natural gas.

2.2 Multi-Agent Game Model
This section analyzes the control problem of multi-energy synergy within an integrated energy system and constructs a hybrid game Markov process for electrical and thermal multi-subjects. The structure can fully consider the interests of different subjects and avoids the dimensional catastrophe in the state and action space that arises when deep reinforcement learning is applied.

In this paper, the Nash equilibrium point of solving the game between multiple agents is used as a control strategy for integrated energy systems. Its target expression is

$$G = g\left(N_g, A_g, u_g\right) \tag{1}$$

Where, G is the equilibrium point reached by the game between multiple agents; $g(\cdot)$ is the game function between multiple agents; N_g is the set of behaviors of the agents involved in the game. It is expressed as the operating action of each type of equipment in this article; u_g is the set of utilities of the agents participating in the game. It is represented the economic interests of the respective agents in this paper.

3 MATHEMATICAL MODEL OF INTEGRATED ENERGY SYSTEM CONTROL STRATEGY

The three investment entities included in the actual IES: electricity, natural gas and heat. These three investment entities will consider their interests more in the operation process. Traditional optimization methods are difficult to solve optimally for the game relationship between the three entities and to minimize the operating cost of the system. Therefore, this section considers the return of the investment subject based on the intelligent body division method proposed in **Section 2**. And the control strategy revenue model of electric, natural gas and thermal energy investment entities is constructed.

3.1 Objective Function

The control objective of this system is to maximize the net profit of its system over a period of T by controlling the output of controllable equipment. The system is set to T for 24 h and the net profit of IES is the operating revenue minus the investment cost and operating cost. The objective function form of **Eq. 1** can be written

$$F = \max \sum_{t=1}^{T} I(t) - C(t)$$

$$s.t. \quad I(t) = \sum_{n=1}^{N} i_e(t) + i_h(t) + i_g(t) \tag{2}$$

$$C(t) = \sum_{n=1}^{N} c_{\cos t}(t)$$

Where, $I(t)$ is the profit of the system at time t; $C(t)$ is the operating cost of the system at time t; $c_{\cos t}(t)$ is the cost of purchased energy. The specific expression for the cost of purchased energy is as follows.

3.1.1 Sales of Electricity, Heat and Gas Revenue

$$i_e = c_{se} P_{grid}^{sell}(t) \tag{3}$$

$$i_h = c_{sh} H^{sell}(t) \tag{4}$$

$$i_g = c_{sg} V^{sell}(t) \tag{5}$$

Where, c_{se}, c_{sh} and c_{sg} are the sold prices of electricity, heat and natural gas, respectively; $P_{grid}^{sell}(t)$ is the electrical power delivered to the consumer by subject n at time period t; $H^{sell}(t)$ is the thermal power to supply heat to the customer in time period t; $V^{sell}(t)$ is the volume of natural gas required by the customer in time period t.

3.1.2 Cost of Purchased Electricity, Heat and Gas

$$c_{\cos t}(t) = \left(c_{be} P^{buy}(t) + c_{bg} V^{buy}(t) \right) \tag{6}$$

Where, c_{be} and c_{bg} are the buying prices of electricity and natural gas, respectively; $P^{buy}(t)$ is the electric power delivered from the grid to the IES system at moment t; $V^{buy}(t)$ is the amount of natural gas delivered from the natural gas network to the IES system at time t.

For heat networks, their power generation is delivered internally by the integrated energy system and is not purchased from outside. Heat network costs have been partially converted from gas source power output and electric power output.

3.2 Equipment Model

The input-output relationship of CHP, GT, EB and GB are introduced as follow. These equipment can convert energy from one form to another.

3.2.1 Combined Heat and Power Unit Model

For the CHP unit, there is a coupling between its output electric power and thermal power. Depending on whether the CHP unit electric heat ratio varies or not, there are two types: fixed heat and power ratio and variable heat and power ratio. In this paper, a fixed heat and power ratio is set, denoted by variable η_{CHP}.

$$\eta_{CHP} = \frac{h_{CHP}(t)}{p_{CHP}(t)} \tag{7}$$

Where, $h_{CHP}(t)$ represents the thermal power of the CHP unit at moment t; $p_{CHP}(t)$ is the electrical power of the CHP unit at moment t. In addition, the natural gas volume consumed by CHP unit is:

$$V_g = \frac{p_{CHP}(t) \cdot \Delta t}{Q_{LHV} \cdot \eta_{G\_E}} \tag{8}$$

Where, η_{G_E} represents the gas-to-power efficiency of CHP unit; Natural gas low calorific value Q_{LHV} is 9.7kwh/m³.

3.2.2 Gas Turbine Model

A gas turbine generates electricity by burning natural gas and has the following equation

$$p_{GT}(t) = \eta_{GT} V_{GT}(t) \tag{9}$$

Where, $p_{GT}(t)$ represents the electrical power of the gas turbine unit at time t; η_{GT} is the gas-electric conversion factor of the gas turbine.

3.2.3 Electric Boiler Model

The electric boiler is connected to the grid as a load, and its mathematical model is shown in the following equation

$$\eta_{EB} = \frac{h_{EB}(t)}{p_{EB}(t)} \tag{10}$$

Where, $h_{EB}(t)$ represents the thermal power emitted by the electric boiler unit at time t; $p_{EB}(t)$ is the electrical power consumed by the EB unit at time t.

3.2.4 Gas Boiler Model

The gas boiler is connected to the gas network in the form of a load, and its mathematical model is shown in the following equation:

$$V_{GB}(t) = \frac{h_{GB}(t)}{\eta_{G\_H}\big/Q_{LHV}} \tag{11}$$

Where, $h_{GB}(t)$ represents the thermal power emitted by the gas boiler unit at time t; $V_{GB}(t)$ is the amount of natural gas consumed by the gas boiler at time t; η_{G_H} is the gas-to-heat efficiency of the gas boiler.

3.3 Constraint Condition

The constraints of the integrated energy system control problem include power balance constraints, external energy supply constraints, and equipment operation constraints.

3.3.1 Power Constraints

At moment t, the electric power balance constraint and the thermal power balance constraint can be expressed respectively as:

$$\begin{aligned} p_{\mathrm{grid}}(t) + p_{\mathrm{RES}}(t) + p_{GT}(t) \\ p_{\mathrm{CHP}}(t) - p_{\mathrm{EB}}(t) = p_{\mathrm{load}}(t) \end{aligned} \tag{12}$$

$$h_{\mathrm{CHP}}(t) + h_{\mathrm{GB}}(t) + h_{\mathrm{EB}}(t) = h_{\mathrm{load}}(t) \tag{13}$$

Where, $p_{grid}(t)$ is the power exchanged between the system and the main grid at time t. A positive power indicates that the system purchases power from the main grid, and a negative power indicates that the system carries out surplus power to the grid; $p_{RES}(t)$ is the output power of renewable energy; $p_{load}(t)$ is the electrical load for time period t; $h_{load}(t)$ is the heat load for time slot t.

3.3.2 Equipment Operating Constraints

The devices in the IES system have an upper and lower operating limit range. There are requirements for the electric power output of CHP units, the thermal power output of gas boilers and electric boilers.

$$P_{\mathrm{CHP}}^{\min} \leqslant p_{\mathrm{CHP}}(t) \leqslant P_{\mathrm{CHP}}^{\max} \tag{14}$$

$$P_{\mathrm{GT}}^{\min} \leqslant p_{\mathrm{GT}}(t) \leqslant P_{\mathrm{GT}}^{\max} \tag{15}$$

$$H_{\mathrm{GB}}^{\min} \leqslant h_{\mathrm{GB}}(t) \leqslant H_{\mathrm{GB}}^{\max} \tag{16}$$

$$H_{\mathrm{EB}}^{\min} \leqslant h_{\mathrm{EB}}(t) \leqslant H_{\mathrm{EB}}^{\max} \tag{17}$$

Where, $P_{\mathrm{CHP}}^{\max}$ and $P_{\mathrm{CHP}}^{\min}$ are the upper and lower limits of the output power of the CHP unit; $P_{\mathrm{GT}}^{\max}$ and $P_{\mathrm{GT}}^{\min}$ are the upper and lower limits of the electrical power output of the CHP unit; $H_{\mathrm{GB}}^{\max}$ and $H_{\mathrm{GB}}^{\min}$ are the upper and lower limits of the thermal power output of the gas boiler; $H_{\mathrm{EB}}^{\max}$ and $H_{\mathrm{EB}}^{\min}$ are the upper and lower limits of the thermal power output of electric boiler respectively.

3.3.3 Energy Conversion Equipment Constraint

The energy conversion equipment of IES must meet the constraints shown in **7–11**.

4 TWO-LEVEL DECISION MODEL

IES control problems have complex constraints and are difficult to converge. To address this problem, this paper proposes a two-level model to achieve efficient model learning and effective control of IES. The upper layer DRL agent is responsible for learning strategies to make decisions for the output of the integrated energy system. The lower layer model receives the actions of the upper layer and performs energy flow calculations to verify the security of the model. When the action of the upper RL model has a tide crossing the line, energy flow imbalance, etc., add an appropriate penalty to the reward. The reward design complexity when the upper output action violates the constraint is reduced by a two-level model design. Meanwhile, the model avoids the problem of the difficult convergence of currents.

4.1 Upper Layer DRL Model Design

In RL, there are two interactive objects: agent and environment. The agents sense the external state and obtain reward. After that, the agent learns and makes decisions. The environment will be affected by the action of the agent and change its state, and give feedback to the agent to get the reward in response. DRL has the general intelligence to solve complex problems to a certain extent and has achieved great success in many tasks. A Markov decision process (MDP) is a sequence of random variables with Markov properties, in which the state of the next moment depends only on the current state. If an action is added to the Markov process, the state at the next moment is related to action and the current state.

The upper layer of the two-level model is the DRL agent. In the paper, IES is the environment of the intelligent, which makes optimal control strategy decisions by regulating the device output in the system.

4.1.1 Reward Function Design

When designing the reward function, setting the scaling factor facilitates the agent to learn the gains of different subjects more equally. This helps to balance the weight of the contribution of each agent. Therefore, the reward function for the upper layer DRL agents is designed as follows:

$$r = \left(-c_{\cos t} - \gamma_{pls} c_{pls} - r_{pun}\right) \tag{18}$$

Where, γ_{pls} is the line loss penalty factor, taken as 0.5; r_{pun} is the line crossing penalty, taken as 200.

4.1.2 State-Space Design

The DRL status should contain sufficient information to make decisions, including electric load demand, thermal load demand, gas load demand and renewable generation power, as well as real-time electricity sales prices. It can be expressed as follows:

$$s_t = \left\{ \begin{array}{l} p_{\mathrm{load}}(t), h_{\mathrm{load}}(t), V_{load}(t), \\ p_{\mathrm{RES}}(t), c_{e\_sell}(t), c_{g\_buy}(t) \end{array} \right\} \tag{19}$$

TABLE 1 | Algorithm of MADDPG.

Algorithm 1: Multi-Agent Deep Deterministic Policy Gradient for N agents

for episode = 1 to M **do**

Initialize a random process $\mathcal{N}$ for action exploration

Receive initial state x

for t = 1 to max-episode-length **do**

  for each agent i, select action $a_i = \mu_{\theta_i}(o_i) + \mathcal{N}_t$ w.r.t the current policy and exploration

  Execute actions $a = (a_1, \ldots, a_N)$ and observe reward r and new state x'

  Store $(\mathbf{x}, a, r, \mathbf{x}')$ in replay buffer $\mathcal{D}$

  $\mathbf{x} \leftarrow \mathbf{x}'$

  for agent i = 1 to N do

    Sample a random minibatch of S sample $(\mathbf{x}^j, a^j, r^j, \mathbf{x}'^j)$ from $\mathcal{D}$

    Set $y^j = r_i^j + \gamma Q_i^{\mu'}\left(x'^j, a_1', \ldots, a_N'\right)\Big|_{a_k' = \mu_k'(o_k^j)}$

    Update critic by minimizing the loss $\mathcal{L}(\theta_i) = \dfrac{1}{S}\sum_j \left(y^j - Q_i^{\mu}\left(\mathbf{x}^j, a_1^j, \ldots, a_N^j\right)\right)^2$

    Update actor using the sampled policy gradient:

$$\nabla_{\theta_i} J \approx \frac{1}{S}\sum_j \nabla_{\theta_i} \mu_i\left(o_i^j\right)$$
$$\nabla_{a_i} Q_i^{\mu}\left(\mathbf{x}^j, a_1^j, \ldots, a_i, \ldots, a_N^j\right)\Big|_{a_i = \mu_i(o_i^j)}$$

  end for

  Update target network parameters for each agent i:

$$\theta_i' \leftarrow \tau\theta_i + (1-\tau)\theta_i'$$

  end for

end for

4.1.3 Action Space Design

At time slot t, the action in IES can be indicated by the output of the device. In this paper, the multi-agent body consists of three parts, grid agent body, heat network agent body and gas network agent body. The intelligent body action can be represented as follows:

$$
\begin{aligned}
A_t &= \{a_e^t, a_h^t\} \\
s.t. \quad a_e^t &= \{p_{CHP}(t)\} \\
a_h^t &= \{h_{GB}(t)\} \\
a_g^t &= \{p_{GT}(t)\}
\end{aligned}
\tag{20}
$$

Where, A_t is the total set of multi-agent actions; a_e^t represents grid agent actions; a_h^t represents thermal network agent actions; a_g^t represents thermal network agent actions.

4.2 Lower Layer Model Design

The objective of the lower layer model is to verify the security of the grid through power flow calculations after the upper layer has made decisions on the output of the grid and the heat network. Meanwhile, its interactive power with the superior grid is obtained. Gas network

to minimize operating costs to optimize autonomous costs. Gas source treatment of the gas network by flow calculation calibration. The results are returned to the upper DRL model to calculate the operating costs of the IES system.

4.2.1 Power Flow Calculation

The electric output of CHP units, electric boiler heat output, and gas boiler heat output are delivered to the lower model in multi-intelligent planning. By power flow calculation, the exchange power and line losses with the main network are derived. The specific calculations are as follows:

$$\Delta P = \frac{P^2 + Q^2}{U^2} R \tag{21}$$

$$\Delta Q = \frac{P^2 + Q^2}{U^2} X \tag{22}$$

4.2.2 Gas Network Flow Calculation

The calculation process is aimed at minimizing the cost of the gas network. Electric output of CHP units, heat output of electric
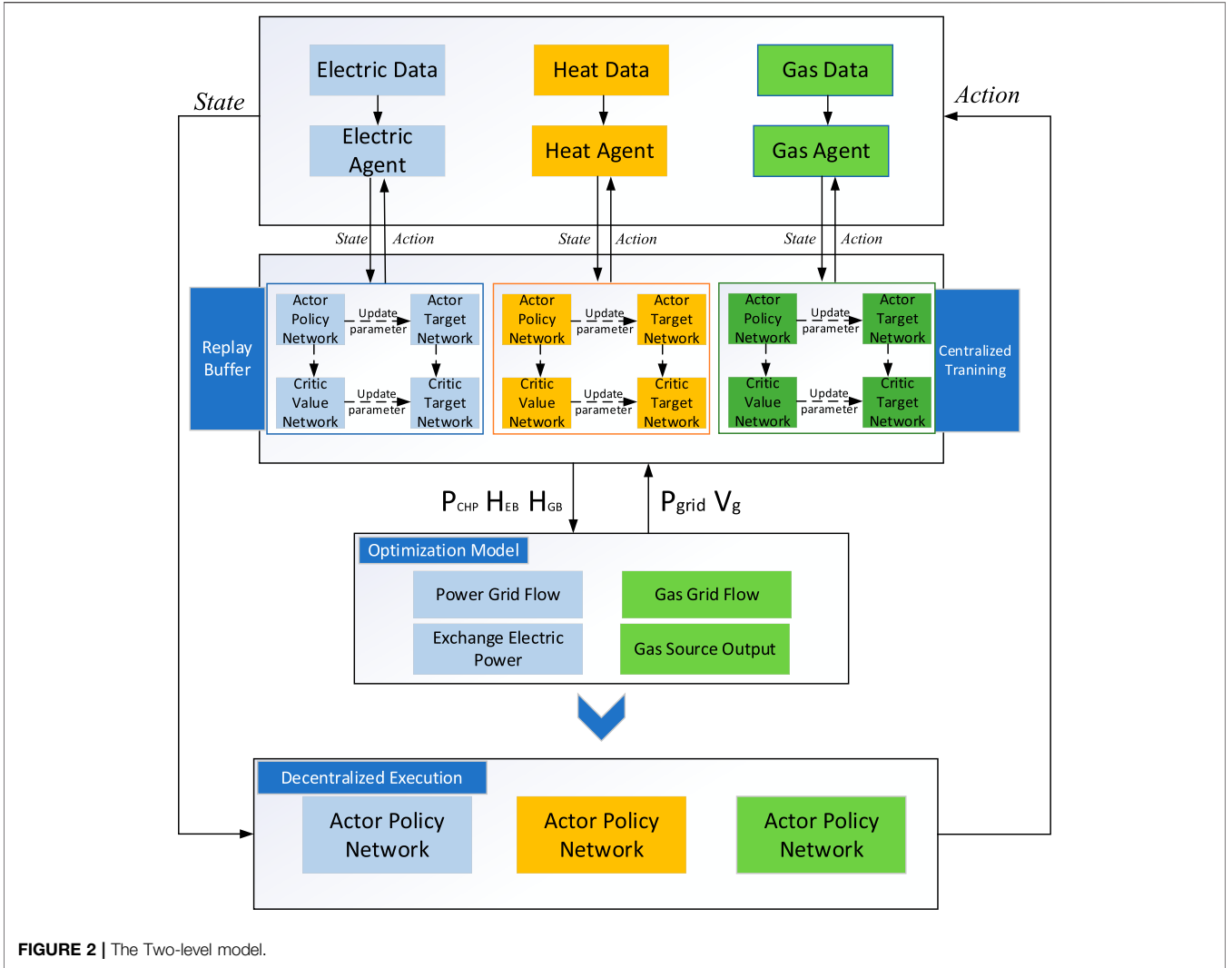
FIGURE 2 | The Two-level model.

boilers, and heat output of gas boilers are passed into the gas network flow model through the upper layer agent planning. At the same time, the calculation considers the constraints of the gas network and constructs the gas network flow model. The details are as follows:

$$w_{j,t} + \sum_{ij \in Z(j)} w_{ij,t} = \sum_{jk \in v(j)} w_{jk,t} \qquad (23)$$

$$w_{ij,t} + w_{ji,t} = 0 \qquad (24)$$

$$w_{j,t} = w_{j,t}^{\text{well}} - w_{j,t}^{\text{pgu}} - w_{j,t}^{\text{load}} \qquad (25)$$

$$w_{ij,t} = C_{ij} \sqrt{\left| \psi_{i,t}^2 - \psi_{j,t}^2 \right|} \qquad (26)$$

$$\psi_{\min} \le \psi_{i,t} \le \psi_{\max} \qquad (27)$$

$$w_{ij,\min} \le w_{ij,t} \le w_{ij,\max} \qquad (28)$$

$$0 \le w_{j,t}^{\text{loadcut}} \le w_{j,t}^{\text{load}} \qquad (29)$$

Where, $w_{ij,t}$ is the gas flow transmitted from pipe node i to node j; $w_{j,t}^{well}$ and $w_{j,t}^{pgu}$ are the source point injection gas flow and the gas consumption of the CHP unit or gas boiler at the node j,

respectively; $w_{j,t}^{load}$ is the node j gas load volume; $Z(j)$ and $v(j)$ are the set of pipes with node j as the last node and node j as the first node, respectively; $\psi_{i,t}$ is the gas pressure at node i; $w_{ij,\max}$ and $w_{ij,\min}$ are the upper and lower node gas pressure limits; C_{ij} is the constant of pipe ij constant. The constant is related to the pipe length, direct, operating temperature and the gas pressure difference between the nodes.

4.3 Model Training and Decision Algorithms

In this paper, we propose to use the MADDPG algorithm based on multi-intelligent reinforcement learning. It belongs to the centralized training-decentralized execution algorithm framework and it is a natural extension of the DDPG algorithm for multi-agent systems. The improvement consists in introducing the inputs of the current policy sampling actions of other agents as additional information in the modeling process of the Q-value function. The MADDPG algorithm has two main advantages: 1) In the training phase, each agent's participant network makes decisions based on local information (the agent's own operations and state); 2) The algorithm does not require

FIGURE 3 | Coupling network.
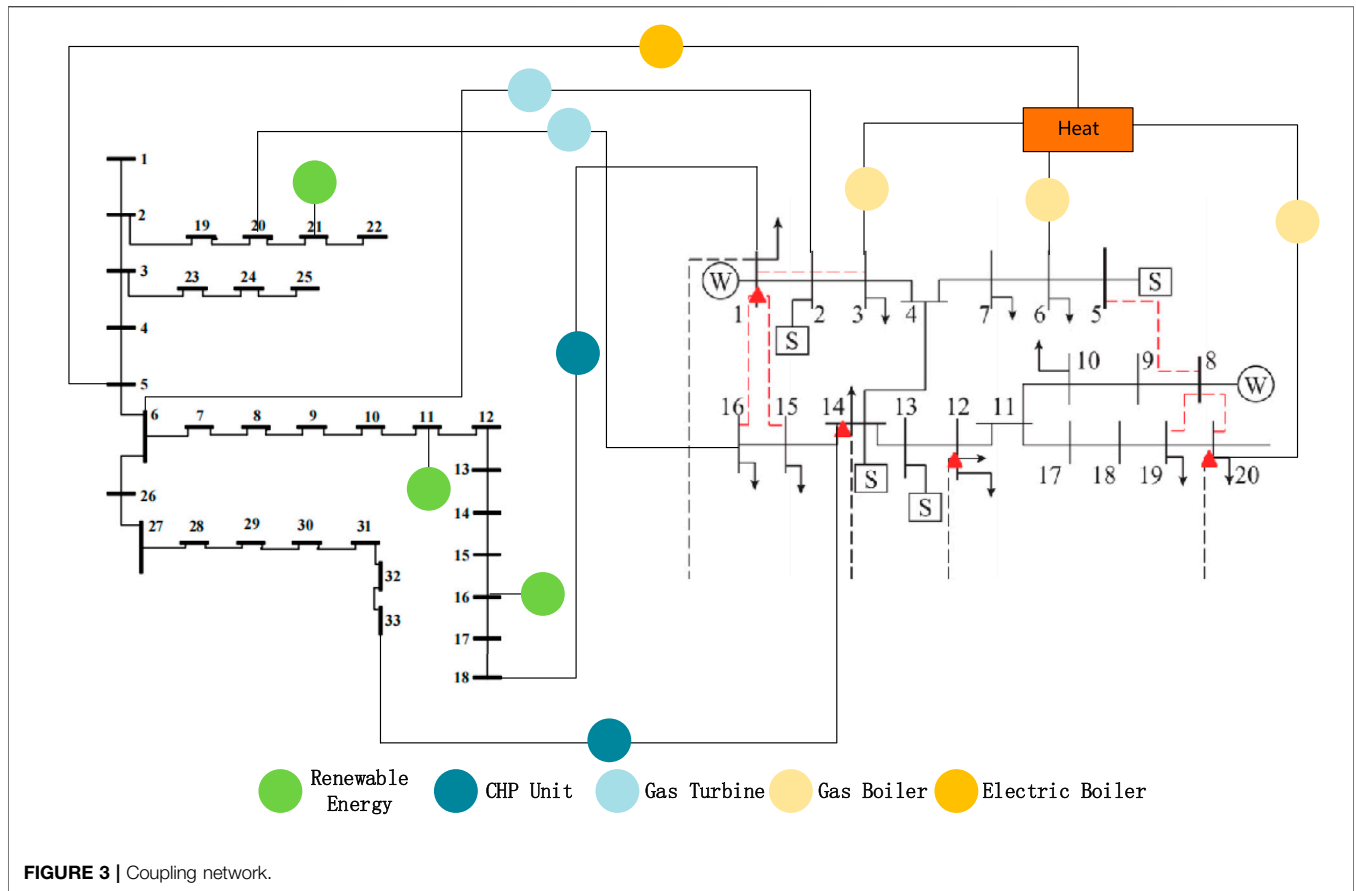
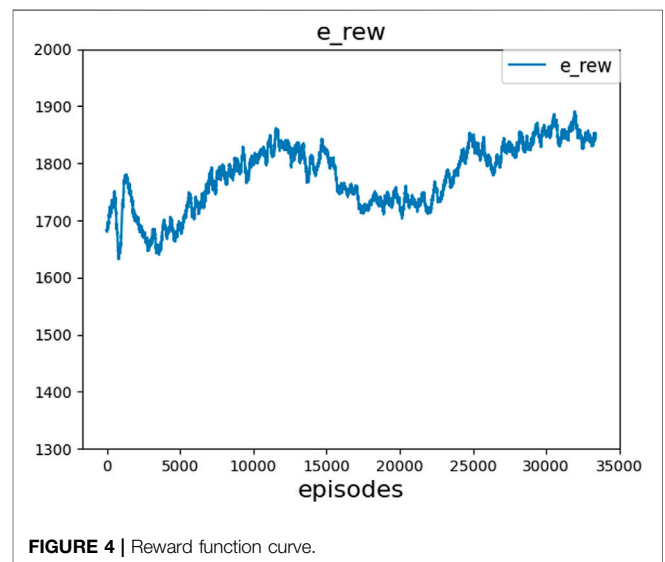TABLE 2 | Equipment parameters.

Parameter	Value	Parameter	Value
η_{CHP}	2.58	H_{EB}^{max}	1,000 kW
η_{G_E}	0.35	H_{GB}^{max}	300 kW
η_{G_H}	0.9	P_{CHP}^{max}	500 kW
η_{EB}	0.85	P_{GT}^{max}	500 kW
η_{GT}	0.8		

TABLE 3 | Price of electricity/gas/heat.

Period/h	Electricity price ($/kW h)	Gas price ($/m³)
00–08	0.08	0.06
08–20	0.16	0.12
20–24	0.12	0.09

FIGURE 4 | Reward function curve.

input information about environmental changes or linkage relationships between agents. Therefore, the algorithm is not only applicable to cooperative environments but also to competitive environments. The implementation of the MADDPG algorithm is shown in **Table 1**.

At moment t, the upper layer DRL agents observe the state of the environment. The upper layer DRL selects the action of the multi-intelligent according to the strategy and transmits this action to the lower layer. The lower layer uses the solver to calculate the flow of the grid and the gas network. The lower layer returns the upper grid output and the treatment of gas sources in the gas network and returns to the DRL model to calculate the reward. Finally, the state s_t at this moment, the state s_{t+1} at the

FIGURE 5 | Evaluation loss function curve.

FIGURE 6 | Strategy loss function curve.

FIGURE 7 | Grid operation.

FIGURE 8 | Operation of heat network.


FIGURE 9 | Operation of the gas network.
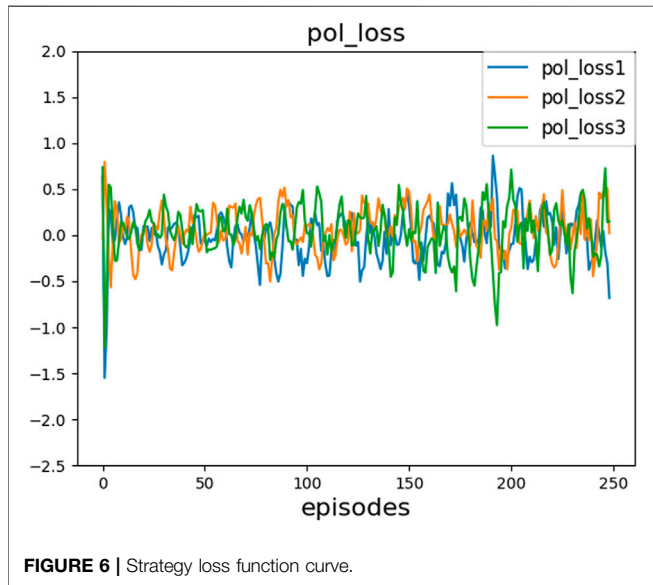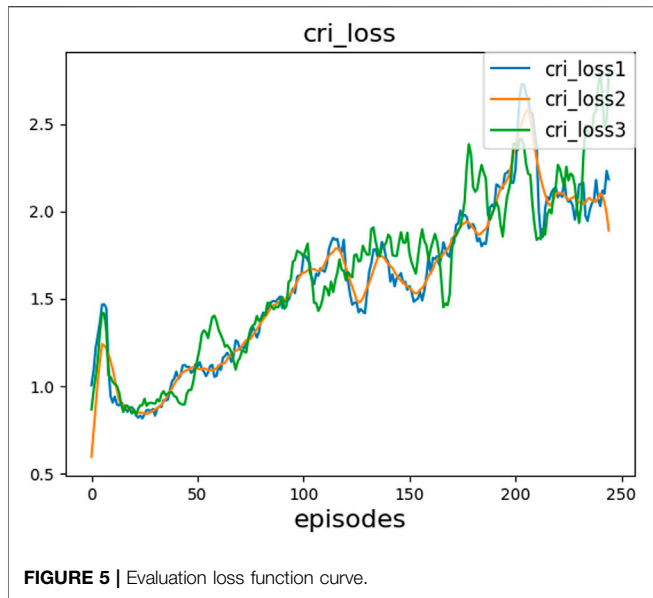
next moment, and the action A_t are stored in the experience pool for the DRL model to train and use. The specific process is shown in **Figure 2**.

5 EXPERIMENTAL VERIFICATIONS

5.1 Example Setting

In this section, arithmetic examples will be designed to verify the advantages and effectiveness of the proposed MADDPG algorithm-based two-level strategy model for integrated energy system dynamic game strategy. The code of chemical learning and trend calculation model is solved based on python programming.
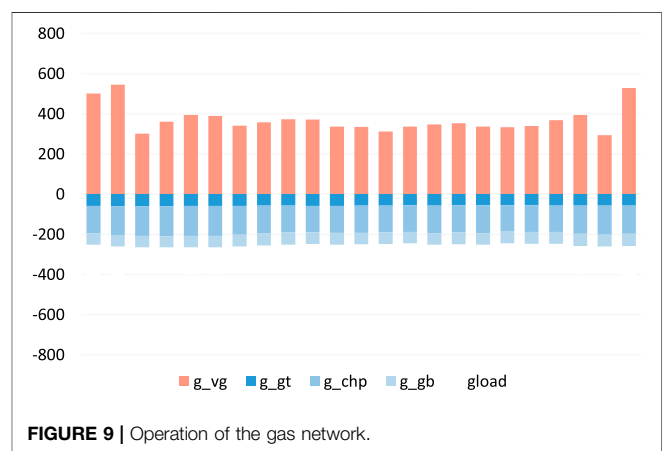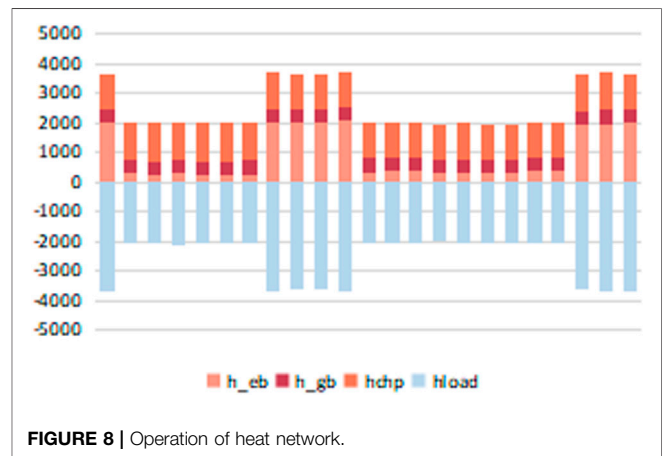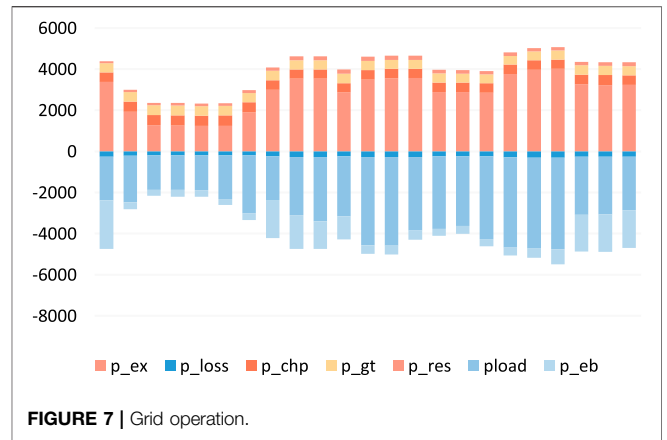
Figure 3 shows the electric-thermal-gas coupling system. The system consists of a 33-node electrical system, and a 20-node natural gas system. The heat network is converted to electrical load form connected to the grid. The power system includes 2 CHP power supplies, 2 wind power supplies and 1 PV power supply. The heating system includes 4 heat sources (1 electric
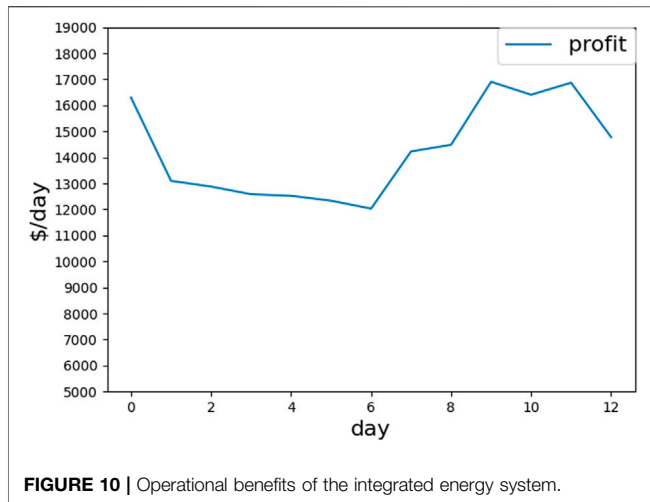
FIGURE 10 | Operational benefits of the integrated energy system.

boiler, 3 gas boilers). A natural gas system with 6 gas sources. The system scheduling time is 24 h and the interval between 2 adjacent periods is 1 h. The operating parameters of the components in the integrated energy system are shown in **Tables 2**, **3**.

5.2 Result Analysis

This section tests the ability of the proposed model to cope with system uncertainty. The model needs to consider the electrical, thermal and gas loads in the IES as well as the renewable energy output. The existence of uncertainty leads to a large number of different scenarios for IES control. This example puts 1 year of renewable energy data (wind turbine and photovoltaic), electrical load data, thermal load data and gas load data into the model to validate the strategy results. The search time step is 24 and sets for 50,000 rounds.

The results shown in **Figure 4** show the effectiveness of the control of the integrated energy system using the two-level model. Initially, due to the incomplete exploration of the action strategies of the agents, the agents choose to sacrifice profits to ensure the operational constraints. Therefore, this decision results in a low value of the reward function. At a later period, through extensive learning, the agents can make effective decisions for different scenarios. Thus, the reward function increases until it converges at about 26,000 steps. In addition, the loss functions of **Figures 5**, **6** show that the loss functions corresponding to the three agents show oscillations at the initial stage. This is due to the failure to find a better action to guarantee the function when the agent is making decisions in the early stages. However, the loss function of the agents smoothed out by continuously sampling and learning from the experience pool and adding noise to ensure data diversity.

In order to verify the convergence and strategy effectiveness under multiple scenarios, the operation of the grid, heat and gas networks for one of the days are shown in **Figures 7**, **8**, and **9**, respectively.

From the results in **Figures 7–9**, it can be seen that the full consumption of renewable energy is achieved within the

integrated energy system. The heat load in the heat network can be met by the equipment in the integrated energy system. In addition, due to the limited power supply equipment in the integrated energy system, most of the electricity is satisfied by the grid exchange power. This also makes the operational efficiency of the integrated energy system more related to the exchange of power. As can be seen from **Figure 10**, the higher demand for cooling equipment in summer leads to an increase in power consumption. The resulting increase in interactive power with the grid increases the cost of integrated energy and leads to reduced benefits. In winter, the benefits of the integrated energy system are increased by the abundance of heating equipment within the system to meet the thermal energy demand.

6 CONCLUSION

For the game relationship between different subjects in IES and the complex energy coupling relationship, this paper establishes a two-level DRL model of an integrated energy system containing three subjects: electricity, heat and gas. The upper layer of the model uses the reinforcement learning MADDPG algorithm to determine the power output of each device for electricity and heat. The lower layer calculates the interactive power with the grid that satisfies the power flow and the output power of the gas source node of the gas network. The proposed two-level model in this paper can simplify the design of reward functions for multi-intelligent reinforcement learning. The model can strengthen action constraints and effectively improve the training speed of reinforcement learning. The analysis shows that the proposed model can effectively improve the energy utilization efficiency of the integrated energy system with the full consideration of the dynamic game of electricity, heat and gas. In addition, the model combines a data-driven multi-agent reinforcement learning approach with traditional trending algorithms. This allows the model to be solved with higher efficiency, realizing the collaborative control strategies for IES, meeting the needs of various energy sources while reducing the cost and energy consumption of the integrated energy system.

In future research, on the one hand, we will further consider the coupling network with more constraints and energy storage devices to the multi-agent economy influence, on the other hand, consider the load demand of different level, such as the reducible and non-reducible loads, and improve the control strategy, to reduce the cost of the integrated energy system and increase energy efficiency.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

REFERENCES

Abdelaziz Mohamed, M., and Eltamaly, A. M. (2018). "A Novel Smart Grid Application for Optimal Sizing of Hybrid Renewable Energy Systems," in *Modeling and Simulation of Smart Grid Integrated with Hybrid Renewable Energy Systems* (Cham: Springer), 39–51. doi:10.1007/978-3-319-64795-1_4

Abdelaziz Mohamed, M., and Eltamaly, A. M. (2018). "Sizing and Techno-Economic Analysis of Stand-Alone Hybrid Photovoltaic/Wind/Diesel/Battery Energy Systems," in *Modeling and Simulation of Smart Grid Integrated with Hybrid Renewable Energy Systems* (Cham: Springer), 23–38. doi:10.1007/978-3-319-64795-1_3

Al-Ghussain, L., Ahmad, A. D., Abubaker, A. M., Abujubbeh, M., Almalaq, A., and Mohamed, M. A. (2021). A Demand-Supply Matching-Based Approach for Mapping Renewable Resources towards 100% Renewable Grids in 2050. *IEEE Access* 9, 58634–58651. doi:10.1109/ACCESS.2021.3072969

Ali, A., Usman, M., Usman, O., and Sarkodie, S. A. (2021). Modeling the Effects of Agricultural Innovation and Biocapacity on Carbon Dioxide Emissions in an Agrarian-Based Economy: Evidence from the Dynamic ARDL Simulations. *Front. Energy Res.* 8, 381. doi:10.3389/fenrg.2020.592061

Fan, H., Yu, Z., Xia, S., and Li, X. (2021). Review on Coordinated Planning of Source-Network-Load-Storage for Integrated Energy Systems. *Front. Energy Res.* 9, 641158. doi:10.3389/fenrg.2021.641158

Huang, W., and Liu, W. (2020). Multi-energy Complementary Based Coordinated Optimal Planning of Park Integrated Energy station-Network[J]. *Automation Electr. Power Syst.* 44 (23), 20–28.

Lan, T., Jermsittiparsert, K., T. Alrashood, S., Rezaei, M., Al-Ghussain, L., and A. Mohamed, M. (2021). An Advanced Machine Learning Based Energy Management of Renewable Microgrids Considering Hybrid Electric Vehicles' Charging Demand. *Energies* 14 (3), 569. doi:10.3390/en14030569

Leo Kumar, S. P. (2017). State of the Art-Intense Review on Artificial Intelligence Systems Application in Process Planning and Manufacturing. *Eng. Appl. Artif. Intell.* 65, 294–329. doi:10.1016/j.engappai.2017.08.005

Liu, H., Li, J., Ge, S., Zhang, P., and Chen, X. (2019). Coordinated Scheduling of Gridl-Connected Integrated Energy Microgrid Based on Multi-Agent Game and Reinforcement Learning[J]. *Automation Electr. Power Syst.* 43 (01), 40–48.

Liu, J., Chen, J., Wang, X., Zeng, J., and Huang, Q. (2020). Energy Management and Optimization of Multi-Energy Grid Based on Deep Reinforcement Learning[J]. *Power Syst. Technol.* 44 (10), 3794–3803.

Ma, H., Liu, Z., Li, M., Wang, B., Si, Y., Yang, Y., et al. (2021). A Two-Stage Optimal Scheduling Method for Active Distribution Networks Considering Uncertainty Risk. *Energy Rep.* 7, 4633–4641. doi:10.1016/j.egyr.2021.07.023

Makhadmeh, S. N., Khader, A. T., Al-Betar, M. A., Naim, S., Abasi, A. K., and Alyasseri, Z. A. A. (2021). A Novel Hybrid Grey Wolf Optimizer with Min-Conflict Algorithm for Power Scheduling Problem in a Smart Home. *Swarm Evol. Comput.* 60, 100793. doi:10.1016/j.swevo.2020.100793

Mohamed, M. A. (2022). A Relaxed Consensus Plus Innovation Based Effective Negotiation Approach for Energy Cooperation between Smart Grid and Microgrid. *Energy* 252, 123996. doi:10.1016/j.energy.2022.123996

Mohamed, M. A., Chabok, H., Awwad, E. M., El-Sherbeeny, A. M., Elmeligy, M. A., and Ali, Z. M. (2020). Stochastic and Distributed Scheduling of Shipboard Power Systems Using MθFOA-ADMM. *Energy* 206, 118041. doi:10.1016/j.energy.2020.118041

Nie, H., Zhang, J., Chen, Y., and Xiao, T. (2021). Real-time Economic Dispatch of Community Integrated Energy System Based on a Double-Layer Reinforcement Learning Method[J]. *Power Syst. Technol.* 45 (04), 1330–1336.

Park, S. H., and Han, K. (2018). Methodologic Guide for Evaluating Clinical Performance and Effect of Artificial Intelligence Technology for Medical Diagnosis and Prediction. *Radiology* 286 (3), 800–809. doi:10.1148/radiol.2017171920

Qiao, J., Wang, X., Zhang, Q., Zhang, D., and Pu, T. (2021). Optimal Dispatch of Integrated Electricity-Gas System with Soft Actor-Critic Deep Reinforcement Learning[J]. *Proc. CSEE* 41 (03), 819–833.

Rezaei, M., Dampage, U., Das, B. K., Nasif, O., Borowski, P. F., and Mohamed, M. A. (2021). Investigating the Impact of Economic Uncertainty on Optimal Sizing of Grid-independent Hybrid Renewable Energy Systems. *Processes* 9 (8), 1468. doi:10.3390/pr9081468

Sun, Q., Li, H., Ma, Z., Wang, C., Campillo, J., Zhang, Q., et al. (2015). A Comprehensive Review of Smart Energy Meters in Intelligent Energy Networks. *IEEE Internet Things J.* 3 (4), 464–479. doi:10.1109/JIOT.2015.2512325

Sutton, R. S., and Barto, A. G. (1998). *Introduction to Rein-Forcement learning[M]*. Cambridge, MA, USA: MIT Press.

Tan, H., Yan, W., Ren, Z., Wang, Q., and Mohamed, M. A. (2022). A Robust Dispatch Model for Integrated Electricity and Heat Networks Considering Price-Based Integrated Demand Response. *Energy* 239, 121875. doi:10.1016/j.energy.2021.121875

Walter, V., Cañizares, C. A., Trovato, M. A., and Forte, G. (2020). An Energy Management System for Isolated Microgrids with Thermal Energy Resources[J]. *Trans. Smart Grid* 2020, 1–18. doi:10.1109/TSG.2020.2973321

Weinstein, A. M. (2010). Computer and Video Game Addiction-A Comparison between Game Users and Non-game Users. *Am. J. drug alcohol abuse* 36 (5), 268–276. doi:10.3109/00952990.2010.491879

Xia, T., Rezaei, M., Dampage, U., Alharbi, S. A., Nasif, O., Borowski, P. F., et al. (2021). Techno-economic Assessment of a Grid-independent Hybrid Power Plant for Co-supplying a Remote Micro-community with Electricity and Hydrogen. *Processes* 9 (8), 1375. doi:10.3390/pr9081375

Yang, T., Zhao, L., Liu, Y., Feng, S., and Pen, H. (2021). Dynamic Economic Dispatch for Integrated Energy System Based on Deep Reinforcement Learning [J]. *Automation Electr. Power Syst.* 45 (05), 39–47.

Yao, Z., and Wang, Z. (2020). Two-level Collaborative Optimal Allocation Method of Integrated Energy System Considering Wind and Solar Uncertainty[J]. *Power Syst. Technol.* 44 (12), 4521–4531.

Ying, J., Wang, J., Xu, J., Fang, X., and Zhang, H. (2019). Real-Time Energy Management of a Microgrid Using Deep Reinforcement Learning[J]. *Energies* 12 (12). doi:10.3390/en12122291

FUNDING