

# Deep Reinforcement Learning Based Energy Scheduling of a Hybrid Electricity/Heat/Hydrogen Energy System

Xianlian Wang<sup>1</sup>, Junyi Dong<sup>1</sup>, Li Sun<sup>1\*</sup>

1 National Engineering Research Center of Power Generation Control and Safety, School of Energy and Environment, Southeast University, Nanjing, 210018, Jiangsu Province, China. (\*sunli12@seu.edu.cn)

## ABSTRACT

Hybrid electricity/heat/hydrogen energy system is a potential solution for the future low-carbon residential energy system. This paper studies the efficient energy scheduling of such system, including hydrogen production, utilization and storage processes. To overcome the problems of coupling among the multi-energy flow and the uncertainties on both sides of power and load, a deep reinforcement learning (DRL) algorithm, namely deep deterministic policy gradient (DDPG), is used to realize adaptive energy scheduling of the system. The scheduling results of simulation experiment under typical winter day scenario illustrate that, based on the pre-trained DDPG framework, the system can achieve a rapid response to the environment and optimize energy efficiency. Additionally, by appropriate power charging and discharging, the states of energy storage devices can essentially recover to their initial states, enabling the sustainable operation of the hybrid energy system.

**Keywords:** hydrogen fuel cell, DDPG, energy scheduling, hybrid energy system

## 1. INTRODUCTION

The development of hybrid energy systems (HESs) is of great significance in encouraging the utilization of renewable energy and promoting the transition to low-carbon or carbon-free systems. As the proportion of renewable energy increases, HESs can efficiently integrate various energy sources to ensure the stability and reliability of energy supply [1-2]. Besides, green hydrogen produced by renewable energy electrolysis can serve as a bridge between electricity and gas networks, expediting the process of deep cross-industry decarbonization [3]. Due to low operation noise, high power generation efficiency [4], and clean products, fuel cells make ideal devices for utilizing hydrogen energy. However, the energy supply of fuel cells is subjected to a

specific thermoelectric ratio, while the energy demand is influenced by diverse factors and presents multiple uncertainties [5]. Therefore, optimal efficiency of fuel cells and stable operation of HESs require device-level development and maintenance, coupled with system-level intelligent scheduling strategies.

To date, traditional optimization algorithms are commonly used to optimize one or multiple predefined objective functions under operational constraints of various devices in HESs [6-7]. Since these algorithms usually rely on specific operational cycles as well as complex demand prediction models, improving the algorithms' resilience to uncertainties in HESs remains a challenge. On the other hand, reinforcement learning (RL) based approaches are capable of mapping the state information to an aggregation decision-making model at the system level, which alleviates the requirement of precise device-level modeling through exploration and exploitation during the training process [8]. Besides, the integration of RL and neural networks, known as deep reinforcement learning (DRL), addressed the challenge of "curse of dimensionality" in RL, thus improving the computational efficiency of RL algorithms and expanding the scope of RL applications [9-10].

Currently, there is a research gap in applying RL to fuel cell-based HESs, especially for multi-energy flow systems that require continuous DRL algorithms to handle complex states and constraints. Furthermore, the majority of existing studies only tackle uncertainties of a particular form in HESs, disregarding the practical issues of energy coupling and efficiency enhancement under multiple uncertainties. For these reasons, this paper presents a new problem concerning the energy scheduling of a hybrid electricity/heat/hydrogen energy system, which encompasses renewable energy, electricity, heat, and hydrogen production, utilization, and storage. An efficient energy scheduling model based on a continuous DRL algorithm called deep deterministic

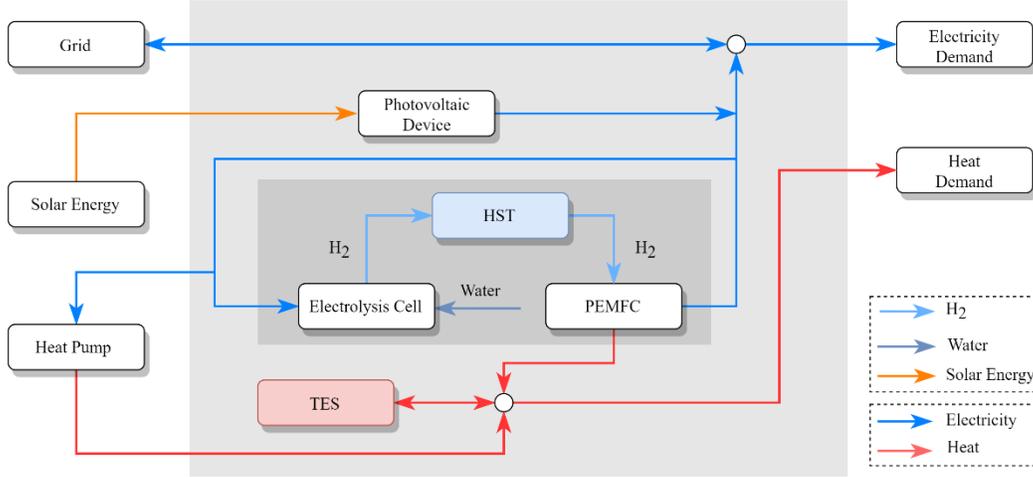


Fig. 1. Schematic diagram of the hybrid electricity/heat/hydrogen energy system

policy gradient (DDPG) is proposed to minimize operation costs and improve sustainability. The proposed model can deal with multiple uncertainties in system operation, which include photovoltaic (PV) power, electricity demand, and heat demand. Finally, the effectiveness and economy of the proposed DDPG-based energy scheduling are demonstrated in simulation under the typical winter day scenario.

## 2. SYSTEM MODELING AND OPTIMIZATION

The structure diagram of a hybrid electricity/heat/hydrogen energy system is shown in Figure 1, including fuel cell, water electrolysis cell, PV device, heat pump, hydrogen storage tank, thermal energy storage, grid, and electricity and heat demands.

### 2.1 Fuel cell model

In this paper, the proton exchange membrane fuel cell (PEMFC) is modelled based on the experiment results and performance evaluation in paper [11]. The hydrogen consumption at anode and cathode,  $n_{H_2}^{FC}$ , can be calculated by theoretical current,  $I$ , as

$$n_{H_2}^{FC} = \frac{n_{cell} I}{2F} \quad (1)$$

where  $n_{cell}$  is the number of cells and  $F$  is the Faraday constant.

PEMFC is characterized by the current density  $j$  as

$$j = \frac{I}{A_{cell}} \quad (2)$$

where  $A_{cell}$  is the effective area of one cell. The power as a function of current density is shown in Figure 2.

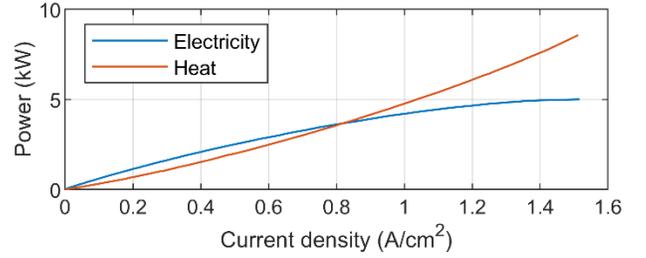


Fig. 2. Impact of current density on PEMFC powers.

### 2.2 Water Electrolysis Cell model

This study adopts the proton exchange membrane method due to its compact structure, constant electrolyte concentration, and strong energy-fluctuating adaptability, which are suitable for coordinated operation with PV power and other volatile energy sources [12]. The hydrogen production rate of the water electrolysis cell  $N_{EH}$  is positively related to current  $I_{EH}$  as

$$N_{EH} = \frac{\eta_F n_{EH} I_{EH}}{2F} \quad (3)$$

where,  $N_{EH}$  is the number of moles of hydrogen production per second,  $n_{EH}$  is the number of cells. The power curve of electrolysis cell as a function of current density is shown in Figure 3.

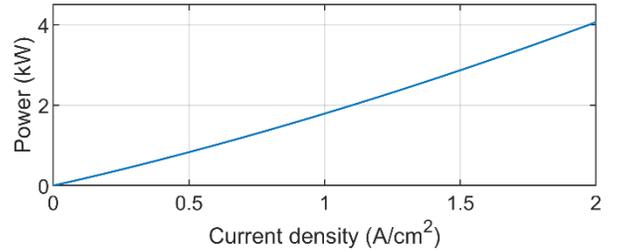


Fig. 3. Impact of the current density on power.

### 2.3 PV device model

A PV device is also included in this system as part of the power source. The power generated by the PV device can be calculated as

$$P_E^{PV} = \eta_{PV} \eta_{inv} A_{PV} I^0 \quad (4)$$

where  $\eta_{PV}$  and  $\eta_{inv}$  are respectively the PV efficiency and inverter efficiency.  $I^0$  is the solar irradiance measured in Watts per square meter ( $W / m^2$ ) and  $A_{PV}$  is effective area of solar panels.

#### 2.4 Heat Pump model

Air source heat pump (HP) is used worldwide to efficiently provide heat from electricity. The coefficient of performance (COP) is defined as the ratio of the heat generated to the electricity consumed. By investigating air source HPs of different manufacturers under a wide range of power conditions and temperature differences, an empirical relationship of COP is regressed as

$$COP = h_1 - h_2 \Delta T_{HP} + h_3 \Delta T_{HP}^2, \quad 15 \leq \Delta T_{HP} \leq 60 \quad (5)$$

where  $\Delta T_{HP}$  represents the temperature difference between the produced water and ambient air, and  $h_i (i=1,2,3)$  is the fitting coefficient [13]. The heat generated by HP can be obtained by

$$P_H^{HP} = P_E^{HP} COP \quad (6)$$

where  $P_{E,k}^{HP}$  is the electricity consumed by heat pump.

#### 2.5 Thermal energy storage model

Thermal energy storage (TES) is included in the hybrid system to decouple the mismatch between electricity and heat demands. To describe the dynamic characteristic of TES, the heat storage degree (HSD), similar to SOC of battery, is introduced and defined as

$$HSD_{t+1} = HSD_t - \frac{P_{H,t}^{TES}}{H_c} \Delta T \quad (7)$$

where  $0 \leq t \leq T$  is the time step,  $P_{H,k}^{TES}$  is the heating power (kW) of the heat charge (negative) or discharge (positive), and  $H_c$  is the full heat storage capacity. The value of HSD ranges from 0 to 1, which means the capacity of TES is between empty and full.

#### 2.6 Hydrogen Storage Tank model

This paper utilizes the hydrogen storage tank (HST) introduced in [14] to store the hydrogen generated by water electrolysis cell, while simultaneously supplying fuel to the fuel cell. The pressure dynamic of HST can be written as

$$TP_{k+1}^H = TP_k^H - z_k \frac{N_k^{HST} RT_H}{M_{H_2} V_H} \quad (8)$$

where  $TP_k^H$  and  $N_k^{HST}$  are the pressure of HST and hydrogen flow rate at time step  $k$ ,  $V_H$  is the volume,  $T_H$  is the temperature generally regarded as a constant,  $M_{H_2}$  is the Molar mass of hydrogen,  $R$  is the universal gas constant, and  $z_k$  is the compression coefficient of hydrogen which can be calculated as follows:

$$z_k(TP_k) = 1 + \sum_{i=1}^9 a_i \left( \frac{T_0}{T_H} \right)^{b_i} (TP_k)^{c_i} \quad (9)$$

where  $a_i, b_i, c_i$  are the constants and  $T_0 = 100K$ .

#### 2.7 Optimization problem formulation

The storage ability of TES and HST provides the proposed HES with sufficient scheduling flexibility. To enhance the efficiency of PV energy utilization, the PV power is directly used by electrolysis cell. Based on these premises, the scheduling optimization problem of the HES is essentially to adjust the power output of fuel cells and heat pumps, so as to minimize the operation cost and ensure the sustainability.

##### 2.7.1 Objective Function

$$\begin{aligned} J &= \min \sum_t (\lambda_1 C_t^{FC} + \lambda_2 C_t^G) \\ &= \sum_t (\lambda_1 c_0 P_{E,t}^{FC} + \lambda_2 p_t P_{E,t}^G) \end{aligned} \quad (10)$$

where  $J$  includes the operation cost of fuel cell,  $C_k^{FC}$ , and power purchasing cost of grid,  $C_k^G$ .  $c_0$  is a constant and  $p_t$  is the electricity price.

##### 2.7.2 Device Constraints

$$P_{E,\min}^{FC} \leq P_{E,t}^{FC} \leq P_{E,\max}^{FC} \quad (11)$$

$$P_{E,\min}^{EH} \leq P_{E,t}^{EH} \leq P_{E,\max}^{EH} \quad (12)$$

$$P_{E,\min}^{HP} \leq P_{E,t}^{HP} \leq P_{E,\max}^{HP} \quad (13)$$

$$HSD_{\min} \leq HSD_t \leq HSD_{\max} \quad (14)$$

$$TP_{\min} \leq TP_t \leq TP_{\max} \quad (15)$$

$$HSD_0 = 0.5 \cdot (HSD_{\min} + HSD_{\max}) \quad (16)$$

$$TP_0 = 0.5 \cdot (TP_{\min} + TP_{\max}) \quad (17)$$

where (11)-(13) are the upper and lower power constraints for fuel cell,  $P_{E,t}^{FC}$ , electrolysis cell,  $P_{E,t}^{FC}$  and heat pump,  $P_{E,t}^{HP}$ . (14)-(15) and (16)-(17) are the power constraint and initial state of TES and HST. This initial value setting way can ensure sufficient charging and discharging space of them.

##### 2.7.3 Power and Hydrogen Balances

$$P_{E,t}^{FC} + P_{E,t}^{PV} + P_{E,t}^G = P_{E,t}^L + P_{E,t}^{HP} + P_{E,t}^{EH} \quad (18)$$

$$P_{H,t}^{FC} + P_{H,t}^{HP} + P_{H,t}^{TES} = P_{H,t}^L \quad (19)$$

$$N_t^{EH} + N_t^{HST} = N_t^{FC} \quad (20)$$

where  $P_{E,t}^G$ ,  $P_{E,t}^L$  and  $P_{H,t}^L$  are respectively the power of grid, electricity demand and heat demand, and  $P_{E,t}^G > 0$  means electricity purchase from grid and  $P_{E,t}^G < 0$  means electricity sale.

The scheduling optimization problem in (10) is a cumulative objective function optimization with multiple decision variables, which requires overcoming various state and control constraints at each time step, while the variables interact in a time-state space and the feasible domain is difficult to plot. Additionally, system uncertainties, such as PV power and demands, are difficult to be handled by conventional deterministic algorithms due to their dependence on complex prediction models and cumbersome computational processes. Therefore, this paper adopts the model-free DDPG algorithm to achieve efficient energy scheduling of the HES.

### 3. DEEP DETERMINISTIC POLICY GRADIENT

#### 3.1 Markov Decision Process

The RL task is usually described by markov decision process (MDP), which can be defined by a five tuple:  $(S, A, P, R, \gamma)$ : state, action, transition matrix, reward and discount factor.

##### 3.1.1 State Space

$$S = [t, P_{PV}, P_L^E, P_L^T, TP, HSD] \quad (21)$$

where the states are respectively time, PV power, electricity demand, heat demand, pressure of HST and HSD of TES.

##### 3.1.2 Action Space

$$A = [N_{FC}, P_{HP}] \quad (22)$$

where the actions are respectively hydrogen flow rate of fuel cell and power of heat pump.

##### 3.1.3 Reward Function

$$\mathcal{R} = - \left( \begin{aligned} &\lambda_1 C^{FC} + \lambda_2 C^G + \lambda_3 P_{ex}^H + \lambda_3 P_{loss}^H \\ &+ \lambda_5 \Delta HSD_T + \lambda_6 \Delta TP_T \end{aligned} \right) \quad (23)$$

where  $C^{FC}$  and  $C^G$  have the same definition as (10).  $P_{ex}^H$  and  $P_{loss}^H$  are the heat power excess and loss at each time step.  $\Delta HSD_T$  and  $\Delta TP_T$  are the terminal state cost of TES and HST, which will be calculated after the last time step and satisfy  $\Delta HSD_T = HSD_T - HSD_0$  and  $\Delta TP_T = TP_T - TP_0$ .

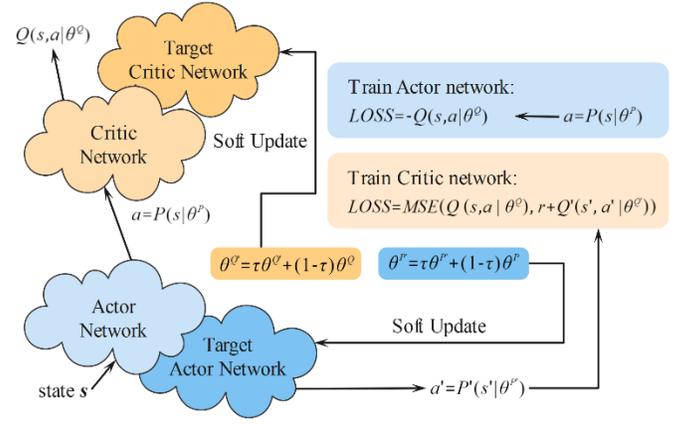


Fig. 4. Schematic diagram of DDPG algorithm

#### 3.2 DDPG algorithm

DDPG is a classical DRL algorithm which can solve the continuous control problem of equipment powers (fuel cell and heat pump) of the HES. In this paper, the Ornstein-Uhlenbeck (OU) noise is utilized for its suitability for inertial systems, which can be added to primary strategies to efficiently exploring the unknown environment. Meanwhile, the noise attenuation is considered to achieve the balance between exploration and exploitation. The schematic diagram of DDPG is shown in figure 4, where the critic network, actor network, target critic network, and target actor network are respectively marked as  $Q(s,a|\theta^Q)$ ,  $s P(s|\theta^P)$ ,  $Q'(s,a|\theta^{Q'})$ , and  $P'(s|\theta^{P'})$ , with  $\theta^Q$ ,  $\theta^P$ ,  $\theta^{Q'}$  and  $\theta^{P'}$  being the network parameters. The OU noise added on action can be expressed as

$$a = P(s|\theta^P) \quad (24)$$

$$a' = a + \mathcal{N}(s) \quad (25)$$

where  $a$  is the action output at state  $s$ ,  $\mathcal{N}(s)$  is the OU noise related to the state and  $a'$  is the new action.

### 4. SIMULATION RESULTS AND ANALYSIS

#### 4.1 System and DDPG settings

The specification and safety constraints of each device are given in Table 1.

Table 1. Main parameters of devices

Device	Specification	Constraints
Fuel cell	$n_{cell} = 12$	$P_{E,max}^{FC} = 4 \text{ kW}$
	$A_{cell} = 100 \text{ cm}^2$	$P_{E,min}^{FC} = 0 \text{ kW}$
Electrolysis cell	$n_{EH} = 10$	$P_{E,max}^{EH} = 4 \text{ kW}$
	$A_{EH} = 2.5 \text{ cm}^2$	$P_{E,min}^{EH} = 0 \text{ kW}$
HST	$V_H = 1.0 \text{ m}^3$	$TP_{max} = 1 \text{ Mpa}$
	$T_H = 300 \text{ K}$	$TP_{min} = 0.5 \text{ Mpa}$

TES	$H_c = 10 \text{ kWh}$	$HSD_{\max} = 1$ $HSD_{\min} = 0$
Heat Pump	$P_{E,r}^{HP} = 4 \text{ kW}$	$P_{E,\max}^{HP} = 4 \text{ kW}$ $P_{E,\min}^{HP} = 0 \text{ kW}$
PV	$A_{pv} = 16 \text{ m}^2$	\

In addition to data in Table 1, the Faraday constant is 96485 C/mol, the temperature of hot water  $T_{DHW}$  is  $55^\circ\text{C}$ , the molar mass of hydrogen is 2.016 g/mol and the gas constant  $R$  is 8.314 J/(mol·K).

The actor and critic network both have four layers. The actor network has five inputs (state) and two outputs (action), and the activation function of output layer is sigmoid. The critic network has seven inputs (state and action) and one output (reward). The target networks are initialized with the same structures and parameters.

#### 4.2 Simulation Results of typical day scenario

To simulate the operation of the HES, a scenario is randomly generated before each training episode according to the typical winter day scenario data [15] and the feasible operation region of the devices, which includes ambient temperature, solar irradiance, electricity demand and heat demand. The randomness in the scenario corresponds to the multiple uncertainties in system operation. The training cumulative reward curve of DDPG is shown in figure 5, where “mean” denotes the average cumulative reward of every five episodes. Based on the pre-trained networks, the power scheduling results of typical winter day scenario are shown in figure 6 and 7. Finally, the pressure and HSD scheduling results are shown in figure 8.

The results in Figure 5 illustrate the exploration and exploitation process of DDPG. With the improvement of strategy, the algorithm converges at about 160 episodes. The power flows in figure 6 and 7 show that there exists no energy excess or loss as a result of the optimized scheduling strategy of fuel cell and heat pump, which ensures the effectiveness of DDPG. Additionally, the results in figure 8 show that the excess hydrogen produced by electrolytic cell can be stored in the HST when PV power is high. Because COP increases with temperature, the heat pump will generate additional heat for TES to charge. After the typical winter day scheduling, the pressure of HST returns to 0.77, with an initial state 0.75, and HSD of TES returns to 0.49, with an initial state 0.5. The results fully guarantee the thermoelectric decoupling demand of the next day. Under the typical daily scenario, the total cost based on DDPG is 3.77, which is 8.71% lower than the 4.13 based on traditional rule-based method.

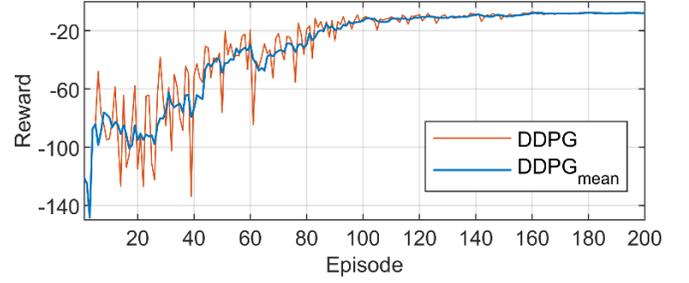


Fig. 5. Cumulative reward curves of DDPG

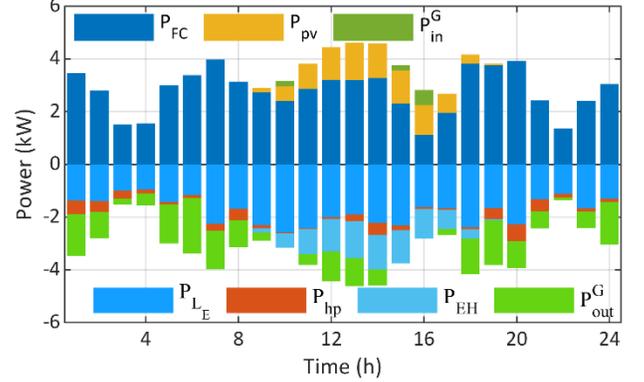


Fig. 6. Electricity power scheduling results of DDPG

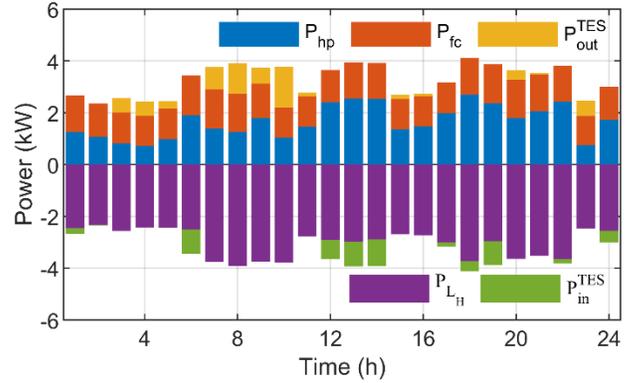


Fig. 7. Scheduling results of heat power based on DDPG

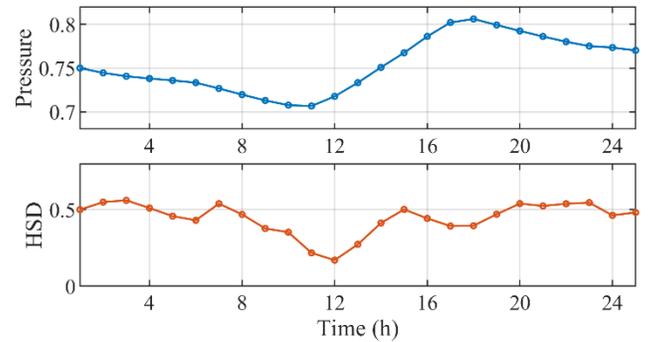


Fig. 8. Scheduling results of the pressure of HST and HSD of TES based on DDPG

## 5. CONCLUSION

This paper proposes a DRL-based energy scheduling method for hybrid electricity/heat/ hydrogen energy systems composed of a fuel cell, water electrolysis cell, PV device, heat pump, hydrogen storage tank, and thermal energy storage. Due to the problems of coupling

among the multi-energy flow and uncertainties on both sides of power and load, achieving efficient energy scheduling with traditional optimization methods remains a challenge. Therefore, the model-free DDPG algorithm is used in this paper to overcome these issues. Firstly, the optimal scheduling problem is established by considering the physical characteristics and safety constraints of the devices. Secondly, the system operation progress is described as an MDP model, including state space, action space, and reward function. Finally, the simulation experiment is carried out under the typical winter day scenario. The scheduling results of DDPG illustrate that the training task under randomly generated scenarios can be successfully completed, and the scheduling strategy of the system can effectively achieve thermoelectric decoupling with no energy excess or loss. Besides, the pressure of HST and HSD of TES can return to 102.67% and 98.00% of their initial states, respectively, which ensures the long-term sustainable operation capability of the HES.

In subsequent studies, we aim to improve the training efficiency of the DDPG algorithm. Besides, we intend to compare DDPG with other RL algorithms in long-term scenarios, so as to validate its effectiveness and superiority.

#### ACKNOWLEDGEMENT

Li Sun would like to acknowledge the support from National Natural Science Foundation of China (NSFC) under grant 51936003, and the funding from Science and Technology Department of Jiangsu Province under Grant BK20211563 & BE2022040.

#### REFERENCE

[1] Niveditha N, Singaravel M M R. Optimal sizing of hybrid PV–Wind–Battery storage system for Net Zero Energy Buildings to reduce grid burden[J]. *Applied Energy*, 2022, 324: 119713.

[2] Azad A H, Shateri H. Design and optimization of an entirely hybrid renewable energy system (WT/PV/BW/HS/TES/EVPL) to supply electrical and thermal loads with considering uncertainties in generation and consumption[J]. *Applied Energy*, 2023, 336: 120782.

[3] Kojima H, Nagasawa K, Todoroki N, et al. Influence of renewable energy power fluctuations on water electrolysis for green hydrogen production[J]. *International Journal of Hydrogen Energy*, 2023, 48(12): 4572-4593.

[4] Haseli Y. Maximum conversion efficiency of hydrogen fuel cells[J]. *international journal of hydrogen energy*, 2018, 43(18): 9015-9021.

[5] Wu J, De G, Tan Z, et al. Study on bi - level multi - objective collaborative optimization model for integrated energy system considering source - load uncertainty[J]. *Energy Science & Engineering*, 2021, 9(8): 1160-1179.

[6] Li J, Liu Y, Qin D, et al. Research on equivalent factor boundary of equivalent consumption minimization strategy for PHEVs. *IEEE Transactions on Vehicular Technology*, 2020, 69(6): 6011-6024.

[7] Wang W, Zhang Z, Shi J, et al. Optimization of a dual-motor coupled powertrain energy management strategy for a battery electric bus based on dynamic programming method. *IEEE Access*, 2018, 6: 32899-32909.

[8] Naeem M, Rizvi S T H, Coronato A. A gentle introduction to reinforcement learning and its application in different fields[J]. *IEEE access*, 2020, 8: 209320-209344.

[9] Ye Y, Qiu D, Sun M, et al. Deep reinforcement learning for strategic bidding in electricity markets. *IEEE Transactions on Smart Grid*, 2019, 11(2): 1343-1355.

[10] Fu Z, Wang H, Tao F, et al. Energy management strategy for fuel cell/battery/ultracapacitor hybrid electric vehicles using deep reinforcement learning with action trimming. *IEEE Transactions on Vehicular Technology*, 2022, 71(7): 7171-7185.

[11] Wang C, Li Q, Wang C, et al. Thermodynamic analysis of a hydrogen fuel cell waste heat recovery system based on a zeotropic organic Rankine cycle. *Energy*, 2021, 232: 121038.

[12] Liso V, Savoia G, Araya S S, et al. Modelling and experimental analysis of a polymer electrolyte membrane water electrolysis cell at different operating temperatures. *Energies*, 2018, 11(12): 3273.

[13] Akmal M, Fox B. Modelling and simulation of underfloor heating system supplied from heat pump, 2016 UKSim-AMSS 18th International Conference on Computer Modelling and Simulation (UKSim). *IEEE*, 2016: 246-251.

[14] Zheng J, Zhang X, Xu P, et al. Standardized equation for hydrogen gas compressibility factor for fuel consumption applications. *International journal of hydrogen energy*, 2016, 41(15): 6610-6617.

[15] Sun L, Jin Y, Shen J, et al. Sustainable residential micro-cogeneration system based on a fuel cell using dynamic programming-based economic day-ahead scheduling. *ACS Sustainable Chemistry & Engineering*, 2021, 9(8): 3258-3266.