

Ateneo de Manila University

Archium Ateneo

Graduate School of Business Publications

Graduate School of Business

1-1-2023

A Data-Driven Scheduling Approach for Integrated Electricity-Hydrogen System Based on Improved DDPG

Yaping Zhao
Shenzhen University

Jingsi Huang
Peking University

Endong Xu
Shenzhen University

Jianxiao Wang
Peking University

Xiaoyun Xu
Ateneo de Manila University

Follow this and additional works at: <https://archium.ateneo.edu/gsb-pubs>





Part of the [Electrical and Computer Engineering Commons](#)

Recommended Citation

Zhao, Y., Huang, J., Xu, E., Wang, J., Xu, X.: A data-driven scheduling approach for integrated electricity-hydrogen system based on improved DDPG. *IET Renew. Power Gener.* 00, 1–14 (2023). <https://doi.org/10.1049/rpg2.12693>

This Article is brought to you for free and open access by the Graduate School of Business at Archium Ateneo. It has been accepted for inclusion in Graduate School of Business Publications by an authorized administrator of Archium Ateneo. For more information, please contact oadrcw.ls@ateneo.edu.

A data-driven scheduling approach for integrated electricity-hydrogen system based on improved DDPG

Yaping Zhao¹ | Jingsi Huang²  | Endong Xu¹ | Jianxiao Wang³  | Xiaoyun Xu⁴

¹Department of Transportation Economics and Logistics Management, College of Economics, Shenzhen University, Shenzhen, China

²Department of Industrial Engineering and Management, College of Engineering, Peking University, Beijing, China

³National Engineering Laboratory for Big Data Analysis and Applications, Peking University, Beijing, China

⁴Department of Operations and Information Technology, Graduate School of Business, Ateneo de Manila University, Quezon City, Metro Manila, Philippines

Correspondence

Jingsi Huang, Department of Industrial Engineering and Management, College of Engineering, Peking University, Beijing, China.
Email: hjsvivan1990@163.com

Funding information

National Natural Science Foundation of China, Grant/Award Number: 72001145; Ministry of Education Program in Humanities and Social Sciences, Grant/Award Number: 20YJC630226; Natural Science Foundation of Guangdong Province, Grant/Award Number: 2022A1515011235

Abstract

The involvement of hydrogen energy systems has been recognised as a promising way to mitigate climate problems. As a kind of efficient multi-energy complementary system, the hydropower-photovoltaic-hydrogen (HPH) system could be an ideal approach to combining hydrogen with an installed renewable energy system to improve the flexibility of energy management and reduce power curtailment. However, the intra-day scheduling of HPH system brings challenges due to the time-related nonlinear hydropower generation process, the complex energy conversion process and the uncertain natural resource supply. Faced with these challenges, an improved deep deterministic policy gradient (DDPG)-based data-driven scheduling algorithm is proposed. In contrast to the prevalent DDPG, two sets of actor-critic networks are properly designed based on prior knowledge-based deep neural networks for the considered complex uncertain system to search for near-optimal policies and approximate actor-value functions. In addition, customized reward functions are proposed with the consideration of interactions among different energy supplies, which helps to improve convergence speed and stability. Finally, the case study results demonstrate that the proposed system model and the optimal energy management strategy based on the improved DDPG algorithm can guide the electricity-hydrogen system to achieve rapid response and more reasonable energy management.

1 | INTRODUCTION

1.1 | Background and literature review

The depletion of traditional fossil energy and the increasingly prominent environmental problems perplex countries worldwide and restrict the further development of human society. Faced with this challenge, promoting energy transformation and consumption revolution by building green, low-carbon, reliable and clean energy systems with renewable energy as the main body is the strategic choice of China and most countries around the world. According to the data from International Energy Agency, in 2019, the global newly-installed generation

capacity of renewable energy power was 190.9 GW, of which China's newly-installed capacity was 65.4 GW, accounting for 34.3% of the global new capacity [1]. By the end of 2019, the installed capacity of renewable energy power generation in China has reached 794 GW, accounting for 39.5% of the total installed capacity, including 356 GW of hydropower, 210 GW of wind power, and 204 GW of photovoltaic power. On the other hand, the annual electricity curtailment of hydropower, wind power, and photovoltaic is 30.0 TWh, 16.9 TWh, and 4.6 TWh, respectively [2], which highlights the problem of renewable energy curtailment.

With the development of non-polluting electrolysis technologies, hydrogen plays a vital role in environment protection

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *IET Renewable Power Generation* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology.

and industry application [3]. As a flexible and environmentally friendly energy storage medium, hydrogen has the ability to achieve both short- and long-term energy storage, which can be used to compensate for the gap between energy supply and consumption [4] and decrease power curtailment. Given its high energy density (2.5-3 times more than natural gas in the same weight), hydrogen is regarded as one of the most efficient fuels [5]. Hydrogen energy is widely used in traditional industries such as oil refining and iron smelting. Besides, in the field of new energy vehicles, hydrogen energy has also become one of the most promising power fuels [6]. Electrolyzing water to produce hydrogen using electricity resources at valley prices can effectively reduce the production cost of hydrogen energy and thus become one of the production methods for the sustainable application of hydrogen energy. Therefore, it is necessary to combine a hydrogen energy system with the installed renewable energy system to improve the flexibility of energy management and promote energy utilization efficiency and system sustainability.

This paper considers an integrated hydropower-photovoltaic-hydrogen (HPH) system, and focuses on its real-time scheduling strategy. Abundant achievements have been accumulated in hydro-photovoltaic (PV) complementary systems to determine the optimal scheduling. In a recently published review article, In [7], the theory and optimal scheduling results of hydro-PV complementary systems are briefly analyzed. Given uncertainties in renewable energy systems, especially wind power and PV power, the study of [8] divides the scheduling methods into three levels, namely prediction, day-ahead scheduling, and real-time scheduling. However, most studies only focus on the first two levels. For example, the long-term and short-term scheduling strategies are considered simultaneously by [9] to decide the optimal PV plant size. A general principle for short-term scheduling of the hydro-PV complementary system is proposed in [10]. The authors of [11] construct a multi-objective optimization operation strategy for long-term hydro-PV systems. A three-layer nested framework is developed by [12] to solve the generation planning problem of cascade hydropower and PV systems considering the uncertainty from PV generation.

Given the superior capability of hydrogen in energy storage and conversion, extensive researches have been conducted on the optimal energy operation of integrated hydrogen energy systems. Among the day-ahead operation researches, the authors of [13] develop a multi-objective optimal operation model of a combined cascade hydropower and hydrogen system. To solve the model efficiently, an E-POA based algorithm is adopted. The research of [14] studies the optimal capacity programming and rule-based operation strategy of a combined PV-hydrogen system. Genetic Algorithm is adopted to optimize the storage component capacities and the operation parameters simultaneously. In [15], a ten-day forecasting model is combined with the operation model of cascade hydropower stations considering electrolyzer and hydrogen storage tank. Although there are some studies in the day-ahead optimal operation of hydropower/PV combined hydrogen system, few study real-time optimal operation of integrated hydrogen energy systems.

For real-time operation researches, model predictive control (MPC) based optimal scheduling method is considered. A bi-level MPC method of real-time optimal operation is proposed in [16] for a renewable hydrogen based microgrid. With their strategy, the system operational cost is reduced. MPC is also adopted by [17] to improve the effect of real-time operation with system uncertainties. Although the above researches can help to guide the modeling and scheduling of HPH systems from different angles, there are few studies based on data-driven real-time scheduling methods which are more realistic but more challenging.

As a kind of data-driven method, deep reinforcement learning (RL) algorithms relax the strict requirement on system dynamics needed by modeling optimization methods. This implies that RL algorithms possess obvious advantages with regard to adaptability, high encapsulation, and time-scale flexibility [18]. According to the review of [19], RL algorithms are widely adopted in load and new energy generation prediction, power market transaction decision, and load side demand response problems. Based on the updating characteristics, RL algorithms can be divided into value function-based RL algorithm, strategy-based RL algorithm and Actor-Critic (AC) algorithm. AC algorithm combines the advantages of value function-based RL algorithms and strategy-based RL algorithms. It uses two neural networks, namely actor network and critical network, to make up for the disadvantages that value function-based RL algorithm cannot deal with large or continuous action space and that strategy-based RL algorithm is difficult to converge [20]. Based on the framework of AC, Deep Deterministic Policy Gradient (DDPG) algorithm is proposed to further recover the complexity of model construction and the difficulty of algorithm convergence [20]. In recent researches of multi-energy system scheduling, a three-layer nested framework is proposed by [21] to maximize total energy production for a large hydro-photovoltaic power plant. An economical optimisation problem is explored by [22] via a CDLR-DDPG algorithm for a combined wind power and natural-gas system. Voltage deviation of a pumped storage hydro-wind-solar system is minimized by [23] via the DDPG algorithm. The authors of [24] use the DDPG framework to determine self-adaptive control strategies for a hybrid energy system with solid oxide fuel cells. The optimal operation strategy is explored by [4] for underground space based integrated hydrogen energy systems using DDPG. In [25], the CPO algorithm is used to study the optimal scheduling of microgrids. Policy updates are constrained by Kullback-Leibler (KL) divergence but suffer from disadvantages such as excessive computational effort and long training time. Therefore, it is not suitable for practical engineering. Additionally, a common problem with existing research is the over-reliance on reward functions for unconstrained exploration, which can easily lead to scheduling decisions that violate the system's operational feasibility [26].

At present, the optimal scheduling of integrated HPH systems has the following problems to be further solved: The real-time scheduling of HPH systems is seldom studied due to the time-related non-linear even non-convex hydropower

generation functions, the complex multi-energy conversion process, and the dynamic resource supply from the environment. In addition, the intra-day optimal scheduling usually requires fast and efficient optimization methods, which further increases the complexity and difficulty of the problem. Compared with the disadvantages of the existing studies, this paper formulates the real-time scheduling problem of a HPH system under a data-driven RL framework. By making full use of historical data, this framework relaxes the strict requirement on system dynamics needed by traditional optimization methods. That is, it is able to learn the dynamics by itself. In addition, the availability of more and more current data enables the proposed data-driven framework to adaptively learn and optimize the complex stochastic problem. This implies that the framework is robust to system changes. To improve the convergence speed and stability of the prevalent DDPG algorithm, the properly designed actor-critic networks and the customized reward functions are proposed with the consideration of system dynamics.

1.2 | Contribution

Through the review and discussion of the above studies, the main contributions of this paper can be summarized as following:

1. An integrated hydropower-photovoltaic-hydrogen system is developed to maximize system revenues from the cooperation of various natural resources. It properly involves in the practical physical constraints including hydrogen and hydropower production as well as the conversion among various forms of energy. Such a system is helpful for not only the increase in revenue, but also the improvement of sustainable development by virtue of resource complementarity. As shown by the case study of the Longyangxia hydro-PV plant, after optimization, daily revenue of the system increases by 1000000 RMB and power adoption rate increases by 5% to 8%.
2. An improved DDPG real-time scheduling algorithm is developed for the targeted system. Given the complexities incurred by real-time optimization, mutually coupled non-linear energy conversion and uncertain natural resource supply, it is challenging to make adaptive and stable policies. This study specifically designs value functions and deep neural networks. Rather than hand-tune control policies, the networks can flexibly generalize to different conditions with the help of data. Besides, considering the instability in decision-making caused by the influences of uncertain external factors, the reward function of the Markov decision process is customized based on the operations of the system. It is jeopardized by the penalty functions so that the policy update is always kept within a controllable range. This contributes to the improvement in convergence and stability of the proposed method.
3. After the training phase of the proposed algorithm is finished offline, the system is capable of making online deci-

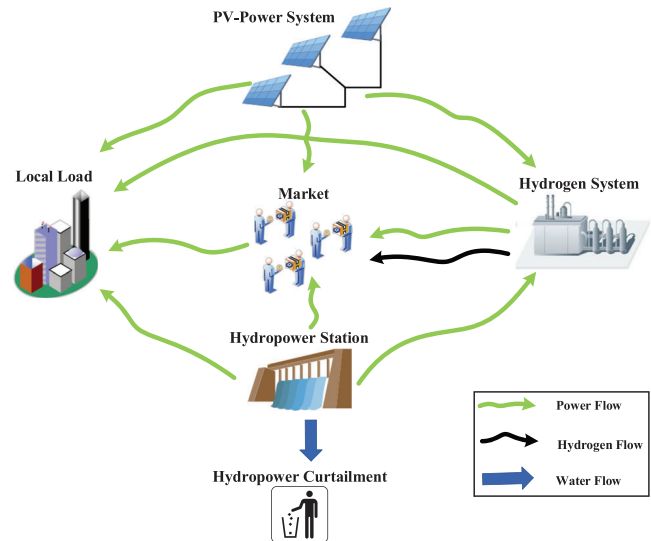


FIGURE 1 Typical system structure schematic of the integrated system.

sions, while interacting with the environment, to manage real-time changes. By applying the proposed algorithm, the negative effect of uncertainties can be curtailed and managed by practitioners. The extensive comparison experiments with other popular algorithms show that the improved DDPG algorithm is capable of making full use of historical data to yield effective and adaptive policies to maximize the revenue of the targeted system.

The rest of the paper is organized as follows: Section 2 introduces the system structure and the model; Section 3 describes the proposed improved DDPG algorithm; Section 4 conducts an empirical study using history data of the Longyangxia hydro-PV plant to demonstrate the effectiveness of the proposed algorithm; Section 5 summarizes and prospects this study.

2 | PROBLEM DESCRIPTION

2.1 | Brief of the integrated HPH system

The targeted integrated HPH system is illustrated in Figure 1. It mainly consists of a hydropower system, electric-hydrogen conversion devices (electrolyzers and hydrogen fuel cells), hydrogen energy storage tanks, photovoltaic devices, and a distribution network. The integrated system supplies power and hydrogen through renewable energy generation, various energy conversions, and flexible energy storage. Hydropower-PV generation is the main energy resource, while the hydrogen fuel cells (HFC) serve as supplementaries. Besides, due to the uncertainty and fluctuation in nature as well as demand, hydrogen storage is adopted to achieve better utilization of natural resources and match demand and supply.

Figure 2 displays the modelling framework which mainly consists of three parts, namely objective functions, decision

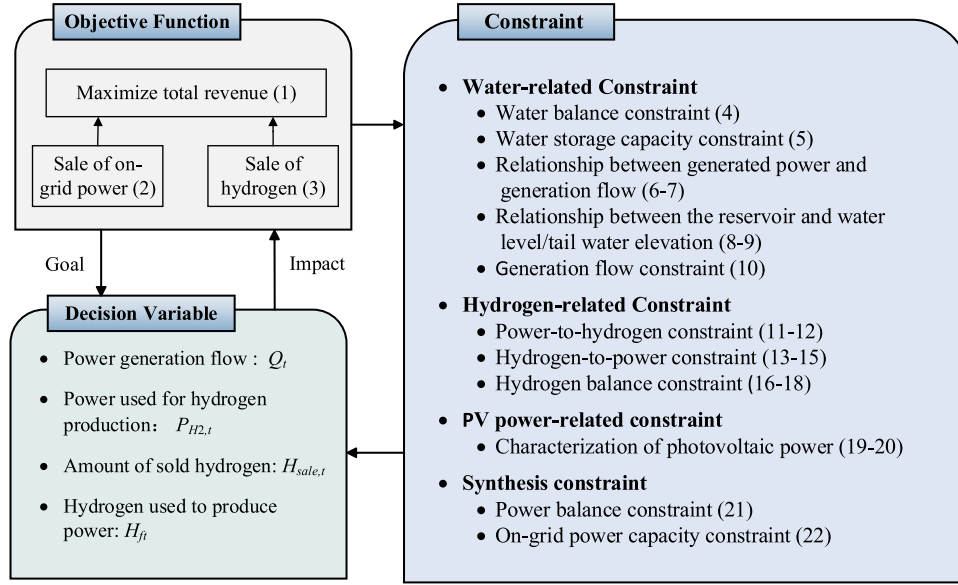


FIGURE 2 Model framework of the considered problem.

variables and constraints. The numbers after the respective objectives and constraints corresponding to the equations described in the following subsections. The optimization objective concerns the maximization of system economic benefit which is subjected to the benefits from on-grid power and hydrogen sales. To effectively operate hydropower, PV-power, and hydrogen, in each period t , one needs to decide the variables including the power generation flow Q_t , the power that is used to produce hydrogen $P_{H_2,t}$, the amount of sold hydrogen $H_{sale,t}$, and the amount of hydrogen used to generate power H_{ft} .

2.2 | Objective function

The goal of the integrated system is to optimize the total system benefit. This mainly originates from two parts. One is the revenue from the sold on-grid electricity, denoted as F_w . The other comes from the sold hydrogen, denoted as F_b . The total revenue F can be expressed as follows:

$$\text{Maximize } F = F_w + F_b, \quad (1)$$

$$F_w = \sum_{t=1}^T (p_{nt} \cdot P_{nt} \cdot \Delta t), \quad (2)$$

$$F_b = \sum_{t=1}^T (p_b \cdot H_{sale,t}). \quad (3)$$

Here, for each period t , P_{nt} is the on-grid power, p_{nt} is the real-time on-grid power price, $H_{sale,t}$ is the volume of the sold hydrogen and p_b is the unit price of the sold hydrogen.

2.3 | Constraints

2.3.1 | Water balance constraint

The reservoir needs to satisfy the following balance equation:

$$V_{t+1} = V_t + I_t - Q_t \cdot \Delta t - P_{dt}. \quad (4)$$

Here, V_t represents the initial storage capacities of the reservoir in period t (or the final volume at the end of period $t - 1$). In period t , I_t is the natural inflow, Q_t is the water flow to generate hydropower and P_{dt} is the water curtailment volume.

2.3.2 | Water storage capacity constraint

Considering the issues like safety and sustainability, the amount of water stored in the reservoir for each period should be within a pre-specified range.

$$V^{min} \leq V_t \leq V^{max}, \quad (5)$$

where V^{min} and V^{max} represent the minimum and the maximum capacity, respectively.

2.3.3 | Relationship between hydropower and generation flow

Reservoirs are helpful to reduce spatial and temporal imbalances of runoff so as to meet direct and indirect needs better. Particularly, hydropower can be generated by converting mechanical energy from water flow to electric energy. This mechanism can be described via a function among generation water flow rate, and net water head and output coefficient of the working

hydropower station [27], as shown below:

$$P_{Ht} = \eta_t \cdot Q_t \cdot H_t, \quad (6)$$

where Q_t is the flow rate to generate hydropower. η_t and H_t represent the output coefficient and the net water head of the considered hydropower station, respectively. Net water head is the difference between water level L_{up}^t and tail water elevation L_{down}^t , that is,

$$H_t = L_{up}^t - L_{down}^t. \quad (7)$$

2.3.4 | Relationship between water storage and water level/tail water elevation

Nonlinear functions $f(x)$ can be used to depict the relationship between water storage of the reservoir V_t and water level L_{up}^t which is closely related to the characteristics of the reservoir. This can be conducted with the help of historical data.

$$V_t = f(L_{up}^t) \cup L_{up}^t = f^{-1}(V_t). \quad (8)$$

Similarly, tail water elevation L_{down}^t can be calculated based on discharge-elevation curve specified by nonlinear function $g(x)$.

$$Q_t = g(L_{down}^t) \cup L_{down}^t = g^{-1}(Q_t). \quad (9)$$

2.3.5 | Generation flow constraint

Given the operating mechanism of the reservoir, the generation flow should be within its lower and upper capacity limits.

$$Q^{min} \leq Q_t \leq Q^{max}, \quad (10)$$

where Q^{min} and Q^{max} are the lower and the upper water generating flow limit, respectively.

2.3.6 | Power-to-hydrogen constraint

As stated in [28], Faraday's law manifests that the chemical conversion relationship between the used electricity and the generated hydrogen in an electrolyzer cell can be expressed as follows.

$$H_{pt} = \eta_e \frac{P_{H_2,t}}{P_{H_2}^{max}} C_{el} \Delta t, \quad (11)$$

where H_{pt} is the amount of hydrogen generated during period t , and it is proportionally related to the conversion rate η_e , the used hydropower $P_{H_2,t}$, and the capacity of the electrolyzer C_{el} . $P_{H_2}^{max}$ is the upper value of the used hydropower. The thermodynamic formula suggests that $C_{el}/P_{H_2}^{max} = 6.6$ [29].

Besides, the amount of hydropower used to generate hydrogen should be within a range as follows:

$$0 \leq P_{H_2,t} \leq P_{H_2}^{max}, \quad (12)$$

where $P_{H_2}^{max}$ is the upper bound of hydrogen production power.

2.3.7 | Hydrogen-to-power constraint

The amount of power that is generated from hydrogen by hydrogen fuel cells P_{ft} can be obtained via the following equation [30]:

$$P_{ft} = \eta_f H_{ft} / \Delta t. \quad (13)$$

Here, η_f represents the inversion rate, and H_{ft} is the hydrogen amount invested to produce electricity. The converted electricity is confined by the capacity of the hydrogen fuel cells and the stored hydrogen in the tanks, that is,

$$0 \leq P_{ft} \leq C_f, \quad (14)$$

$$0 \leq H_{ft} \leq H_{s,t-1}, \quad (15)$$

where C_f is the limit of the hydrogen fuel cells, and $H_{s,t-1}$ is the stored hydrogen at the end of period $t - 1$.

2.3.8 | Hydrogen balance constraint

For the hydrogen in the considered integrated system, the following balance equation holds:

$$H_{s,t-1} + H_{pt} = H_{s,t} + H_{sale,t} + H_{ft}. \quad (16)$$

This balance equation suggests that the total hydrogen comes from two sources, namely the stored volume $H_{s,t-1}$ at the end of period $t - 1$ and the produced volume from hydropower H_{pt} in period t . They can be used to generate power H_{ft} or sold $H_{sale,t}$ in period t . Then the leftover can be stored $H_{s,t}$. In addition, the capacity of storage tanks is limited by C_b , and the sold volume is constrained by the available amount.

$$0 \leq H_{st} \leq C_b, \quad (17)$$

$$0 \leq H_{sale,t} \leq H_{s,t-1} + H_{pt}. \quad (18)$$

2.3.9 | PV power generation

The uncertain PV power can be described by adding errors to the forecasts of its output.

$$P_{PV,t} = P_{\overline{PV}} + e, \quad (19)$$

where $P_{\overline{PV}}$ is the forecast of photovoltaic power in period t , and e represents its forecast error. It is widely recognized that

forecast errors of an uncertain parameter can be characterized via normal distributions [31–33], that is,

$$e \sim N(\mu, \sigma^2). \quad (20)$$

Here, μ and σ represent the mean and the standard deviation of the forecast error, respectively. Studies on the forecast of day-ahead PV power [34, 35] suggest that σ usually ranges from 7.0% to 13.7% of the PV installed capacity.

2.3.10 | Power balance constraint

In each period t , the amount of electricity generated should equal the amount of electricity consumed. This implies that the following dynamic power balance equation exists:

$$P_{Ht} + P_{PV,t} + P_{ft} = P_{Lt} + P_{H2,t} + P_{wt}. \quad (21)$$

This equation suggests that available electricity comes from hydropower P_{Ht} , photovoltaic power $P_{PV,t}$ and hydrogen-converted power P_{ft} . These power will be used to meet local load P_{Lt} , produce hydrogen $P_{H2,t}$ or sell to markets P_{wt} .

2.3.11 | On-grid power capacity constraint

The capacity of the on-grid power in the distribution network is limited.

$$-P_{wt}^{max} \leq P_{wt}^{max} \leq P_{wt}^{max}, \quad (22)$$

where P_{wt}^{max} is the maximum allowable capacity of on-grid power in period t .

2.4 | Markov decision process model of the integrated system

In the considered integrated system, the entire operation process is discretely partitioned in time. Besides, the storage states of hydrogen tank and reservoir water at the next time slot depend only on the storage states and decisions made at the current time slot. This suggests that the targeted problem can be transformed into a finite stochastic dynamic decision process, which is then modeled as a Markov decision process (MDP).

MDP is commonly used to describe reinforcement learning methods, and it can be characterized by a five-tuple as $(S, \mathcal{A}, f, R, \gamma)$. Here, S represents the set of possible states, and \mathcal{A} denotes the set of actions that can be chosen from. f is the transition probability function, that is, $s_{t+1} \sim f(s_t, a_t)$ represents the probability to reach state s_{t+1} when action a_t is taken at state s_t . R represents the instantaneous reward function, and $\gamma \in [0, 1]$ is the discount factor of the accumulated reward and it reflects the damping effect of future reward. For the considered system in this paper, it interacts intermittently with the environment in all decision stages, and then dynamically makes decisions regarding hydropower and hydrogen. In each period

t , the integrated system senses the environment state variable s_t , and takes action a_t in response to the interaction state variable s_t based on the policy $\pi(a_t|s_t)$. At this time, the whole system reaches a new state s_{t+1} according to the transition probability function $f(s_t, a_t)$, and generates a reward of r_t as a feedback of the decision. The MDP of the targeted problem is defined as follows.

2.4.1 | State

Local and global information of the integrated system is crucial for decision-making, and can be described by defining state variable s_t in period t . Generally speaking, it should be capable of reflecting the major characteristics of the system, and tracking the variation of its status. Besides, the definition needs to be concise to make the learning more efficient.

In the integrated system, state s_t includes period index t , reservoir water storage V_t , hydrogen storage H_{st} , natural inflow of the reservoir I_t , local load P_{Lt} , PV-power $P_{PV,t}$, and on-grid electricity price p_{wt} . They drive the optimization of related decisions on actions when changes happen to the environment of the integrated system.

Therefore, the entire state at period t can be formulated as

$$s_t = [t, V_t, H_{s,t-1}, I_t, P_{Lt}, P_{PV,t}, p_{wt}]. \quad (23)$$

2.4.2 | Action

Actions are decisions made in each period. Through the optimization of a series of actions, the integrated system is possible to achieve optimal operation regarding the objective in Section 2.2 under specified physical constraints. The action space in this study can be expressed as

$$a_t = [Q_t, P_{H2,t}, H_{sale,t}, H_{ft}]. \quad (24)$$

These actions are related to the regulation of water and hydrogen in each period t .

A policy $\pi : S \rightarrow \mathcal{A}$ denotes the mapping from each state to a specific action. This suggests that under policy π , action a_t is taken at state s_t and sequentially moves to state s_{t+1} with instantaneous reward r_t as a result.

2.4.3 | Reward

Reward function R is used as a feedback to evaluate the quality of the applied policy at each step. Its design is key to the speed and the optimality of the overall policies. Since the goal of the considered system is to achieve the maximum total revenue, a natural idea is to refer to Equations (1)–(3), the objective functions, as the reward. Further more, since the revenue is obtained mainly through the balance between water storage and hydrogen storage, it is crucial to ensure these two storages meet their physical requirements. This, however, is significantly challenging

to achieve due to the uncertainties involved in the integrated system and the coupling relationships of periods.

Given the above considerations, the reward of the MDP consists of two parts. One is the revenue that results from the objective function in Section 2.2, while the other is storage penalties derived from the constraints in Section 2.3. This leads to the following reward function.

$$r_t = p_{w,t} \cdot P_{w,t} \cdot \Delta t + p_b \cdot H_{sale,t} - \lambda_w \cdot Pen_{wt} - \lambda_b \cdot Pen_{bt}. \quad (25)$$

Here, Pen_{wt} and Pen_{bt} are the penalties resulted from water storage and hydrogen storage, respectively. Given their requirements on capacities, the penalties are defined as follows.

$$Pen_{wt} = |\max\{V_t, V_{max}\} + \min\{V_t, V_{min}\} - V_{max} - V_{min}|, \quad (26)$$

$$Pen_{bt} = |\max\{H_{st}, C_b\} + \min\{H_{st}, 0\} - C_b|, \quad (27)$$

where λ_w and λ_b are penalty coefficients.

Considering the discount factor γ of future rewards, the accumulated reward R_t from period t to the terminated period T can be expressed as

$$R_t = \sum_{i=t}^T \gamma^{i-t} r_i. \quad (28)$$

Given the uncertainties in the environment as well as decisions, the total reward under a specific policy π is defined as the action-value function $Q^\pi(s_t, a_t)$.

$$Q^\pi(s_t, a_t) = E[R_t | s_t, a_t, \pi]. \quad (29)$$

The optimal policy π^* leads to the largest reward feedback $Q^*(s_t, a_t)$, namely the optimal action-value function. $Q(s_t, a_t)$ conforms to the following Bellman equation.

$$Q^\pi(s_t, a_t) = E_{r_t, s_{t+1} \sim S}[r_t | s_t, a_t + \gamma E_{a_{t+1} \sim \pi}[Q^\pi(s_{t+1}, a_{t+1})]].$$

2.4.4 | Transition probability function

System state transition probability function f is crucial for the deduction of $Q^*(s_t, a_t)$. However, due to the complex coupling relationships and uncertainties among and in various resources, it is challenging to explicitly express f . To overcome this challenge, DDPG method is adopted in this study.

DDPG is a model-free actor-critic reinforcement learning algorithm. It takes advantage of deterministic policy gradient and neural network function approximation. By approximately representing the action-value function and the policy via properly designed neural networks, this algorithm is able to reduce the overhead of calculation and the difficulty in system dynamic modeling.

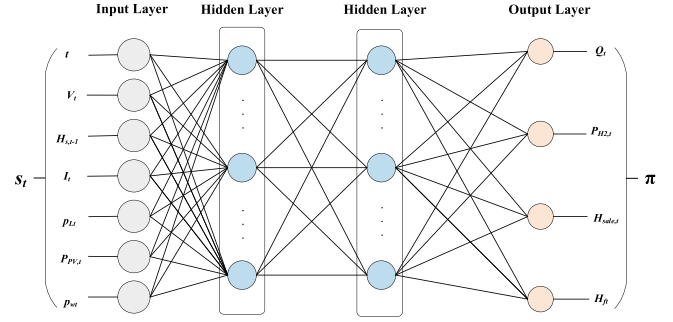


FIGURE 3 Neutral network structure of (targeted) actor network.

3 | IMPROVED DDPG DATA-DRIVEN OPTIMIZATION ALGORITHM

For the problem described in Section 2, an improved DDPG optimization algorithm is proposed in this section. DDPG is a data-driven optimisation method under the actor-critic mechanisms. It can update and optimise both the action-value functions and the policy functions based on deterministic policy gradients and deep function approximators in high-dimensional, continuous action space. This enables DDPG to deal with complex optimization problems with continuous nonlinear decision variables.

To be specific, there are four neural networks existing in the proposed improved DDPG data-driven optimization algorithm: an actor network denoted by μ with parameters θ^μ , a critic network denoted by Q with parameters θ^Q , and their corresponding target networks denoted by μ' and Q' (with parameters $\theta^{\mu'}$ and $\theta^{Q'}$, respectively). Actor networks map states to actions to take, while critic networks yield value estimates of the actions. The actor network is used to select the preferable action $a_t = \mu(s_t | \theta^\mu)$ given a specific state. The critic network is used to evaluate the performance of the action yielded by the actor network, namely, $Q(s_t, a_t | \theta^Q)$. The target networks μ' and Q' leverage time-delayed updates, and slowly track the learned networks μ and Q . Neural networks are adopted as nonlinear function approximators to adaptively learn policy functions and action-value functions. Figures 3 and 4 schematically illustrate the neural network structures of the (targeted) actor and the (targeted) critic.

The actor-critic (AC) network architecture diagram applied in this paper is further depicted in Figure 5. Given the consideration of exploration, a noise N_t obtained from a certain noise process is added to the actor neural network as well as the critic neural network. For the actor neural network, the connection layer makes s_t and N_t unified into one, and then go through three fully connection layers. The first two activation functions are leakRelu, and the last one is Sigmoid which finally leads to a_t . For the critic neural network, the connection layer unifies s_t , N_t and a_t . Then they are fully connected twice by activation functions leakRelu, and once with no activation functions. This finally leads to $Q(s_t, a_t)$.

The reduction in updating frequency is helpful to make the learning process more stable. In addition, to address the issue

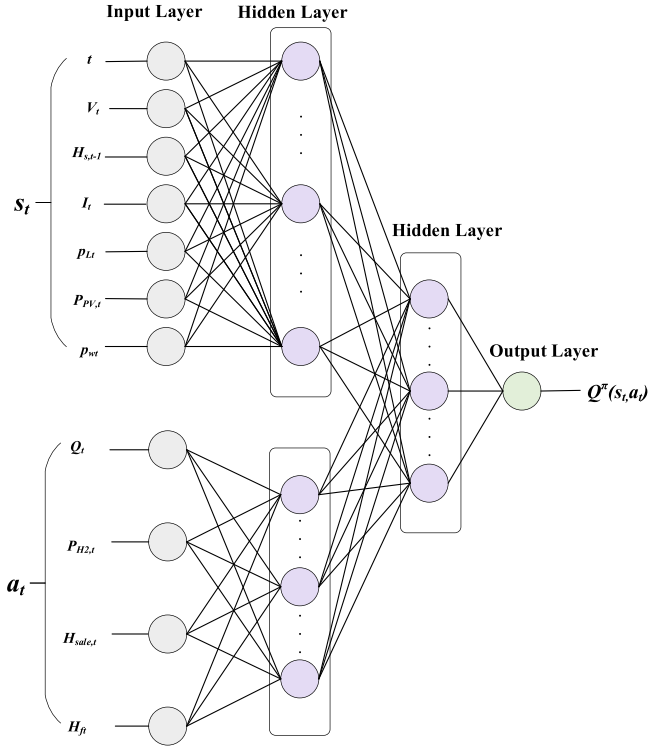


FIGURE 4 Neutral network structure of (targeted) critic network.

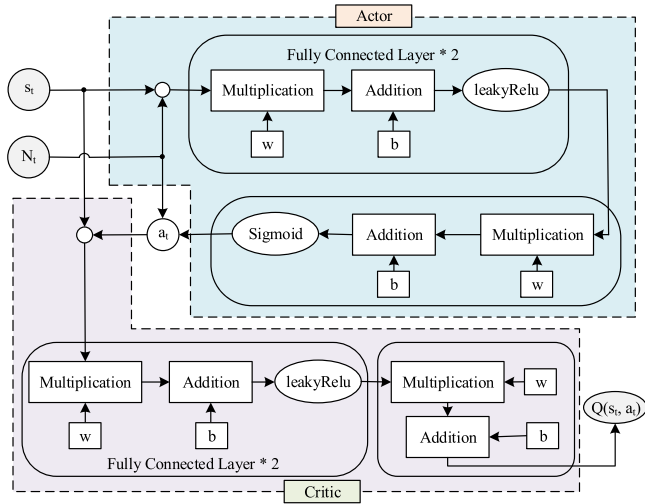


FIGURE 5 AC network architecture diagram.

that samples may not be independent identically distributed, memory replay mechanism is applied. That is, when updating the networks, a minibatch of N transitions are sampled randomly from the replay buffer, $(s_t^j, a_t^j, r_t^j, s_{t+1}^j), i = 1, 2, \dots, N$. The loss function demonstrated below is then minimized to update the critic network.

$$L = \frac{1}{N} \sum_i \left(r_t^j + \gamma Q'(s_{t+1}^j, \mu'(s_{t+1}^j | \theta^{\mu'})) | \theta^{Q'} - Q(s_t^j, a_t^j | \theta^Q) \right)^2 \quad (30)$$

where γ is the discounting factor.

Similarly, the actor network is updated by maximising the accumulation reward $J(\theta^\mu)$ using sampled policy gradient.

$$\Delta_{\theta^\mu} J = \frac{1}{N} \sum_i \Delta_a Q(s, a | \theta^Q) |_{s=s_t^j, a=\mu(s_t^j)} \Delta_{\theta^\mu} \mu(s | \theta^\mu) |_{s_t^j} \quad (31)$$

Based on Equations (30) and (31), the networks of (Q, μ) can be updated. Given these results, the target networks (Q', μ') is softly updated, that is, they are updated slowly with the tracking on the learned networks (Q, μ) via the following weighted forms.

$$\theta^{Q'} \leftarrow \tau \theta^Q + (1 - \tau) \theta^{Q'}, \quad (32)$$

$$\theta^{\mu'} \leftarrow \tau \theta^\mu + (1 - \tau) \theta^{\mu'}, \quad (33)$$

where τ is the updating smoothing coefficient.

Through iterative learning, the (near) optimal policy for the considered integrated system can be obtained. The detailed flowchart of the proposed algorithm is illustrated by Figure 6. In contrast to the prevalent DDPG, we for the first time design a customized prior knowledge-based deep neural networks for the complex uncertain energy system. In addition, the elaborately designed reward function can help the algorithm to converge to the (near) optimal policy.

4 | CASE STUDY

4.1 | Case description

Located in Qinghai province of north-west China, Longyangxia hydropower station takes the role of a leading reservoir of the upstream-Yellow River. It is a 1280 MW hydropower peaking station with four 320 MW turbines. Beside, the intensive solar radiation and long lighting time of Qinghai unable Longyangxia-PV plant to be one of the largest hydro-PV plant in the world. This PV system is connected directly to the hydropower station, and has a capacity of 320 MW, as shown in Figure 7.

With the help of historical data of Longyangxia hydro-PV station, daily natural inflow, daily PV power and daily local load in April 2014 are collected and used in the following study. This is because pre-statistical analysis indicates that the runoff and PV-power of 2014 is typical. In addition, although the flood season brings higher reservoir inflow, the analysis indicates that there is not a significant difference. Therefore, we choose a typical day in April, and reduce the case study scenarios into high and low solar intensities for the consideration of conciseness. Related parameters are listed in Table 1.

According to the consumption characteristics of electricity, one day of 24 h is divided into 3 kinds of period (i.e. peak period, normal period, and valley period). The time-of-use electricity price is used to regulate electricity consumption. Detailed price information is shown in Table 2. In addition, it is assumed that hydrogen sales price is 3.3 RMB/m³, and this equals the sales price of industrial hydrogen [36].

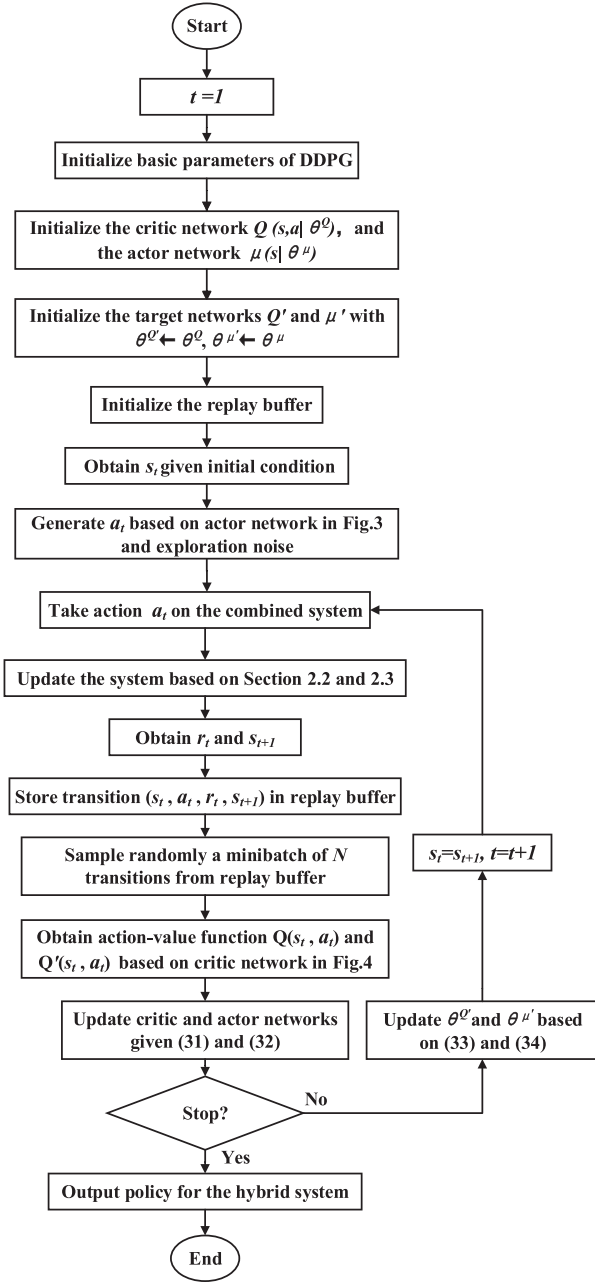


FIGURE 6 Flowchart of the proposed algorithm.



FIGURE 7 Layout of Longyangxia hydro-PV plant.

TABLE 1 Basic parameter of HPH system.

Objective	Parameter	Value
Hydropower	Hydropower output coefficient η_h	85%
	Reservoir lower bound V_{min}	$5.34 \times 10^9 \text{ m}^3$
	Reservoir upper bound V_{max}	$2.47 \times 10^{10} \text{ m}^3$
	Flow lower bound Q^{min}	$50 \text{ m}^3/\text{s}$
	Flow upper bound Q^{max}	$292 \text{ m}^3/\text{s}$
Hydrogen	Power-hydrogen coefficient η_e	80%
	Hydrogen-power coefficient η_f	60%
	Hydrogen production volume $P_{H_2}^{max}$	200 MW
	The capacity of HFC C_f	100 MW
	Hydrogen tank capacity C_b	$5 * 10^5 \text{ m}^3$

TABLE 2 Adopted time-of-use electricity price.

	Peak	Normal	Vallay
Time	9:00-12:00 18:00-23:00	13:00-17:00	1:00-8:00 24:00
Price(RMB/kWh)	0.89	0.58	0.21

TABLE 3 Key parameter of the proposed DDPG.

Parameter	Value
Smoothing coefficient τ	0.005
Discount factor γ	0.99
Replay buffer capacity	$1 * 10^5$
Minibatch size N	100
Exploration noise variance	0.10

4.2 | Tested algorithm introduction

The neural networks of the (target) actor and critic of the proposed algorithm are designed with two hidden layers, and they has 350 and 300 neurons, respectively. Other key hyperparameters are shown in Table 3. Given the novelty of the considered problem, there is no existing baseline algorithms for comparison purposes to the knowledge of the authors. To validate the effectiveness of the proposed algorithm, an myopic algorithm (Myo) and a model predictive control (MPC) approach are introduced and compared. Instead of mining all the historical data when making decisions as in the proposed algorithm, these two algorithms only use a part of the data to optimize the system.

To be specific, Myo is very short-sighted, and it makes decisions based only on the present system states and information, and considers the benefit of only the present period. That is, in each period t , this algorithm minimizes the economic benefit of the system, $p_{wt} \cdot P_{wt} \cdot \Delta t + p_b \cdot H_{sale,t}$, in this period. The optimization is subject to the constraints in Section 2.3 given the state of the system $V_t, H_{s,t-1}$, and the information $I_t, P_{L,t}, P_{PV,t}$.

MPC is a well-established technique for controlling and optimizing multi-variable systems. It can effectively deal with constraints on manipulated variables and outputs. Given a long history of success in the process industries, in recent years, this kind of algorithm is also rapidly expanding in energy management problems [16, 17]. Different from Myo, MPC utilizes the most recent data to predict future unknown information, and then optimizes the system considering the subsequent performance. It contains the basic components of prediction, optimization and receding horizon implementation. That is, in each period t , it optimizes the economic benefit of several further periods (5 periods in this case study), $\sum_{t'=t}^{t+4} (p_{wt'} \cdot P_{wt'} \cdot \Delta t + p_b \cdot H_{sale,t'})$ where the required further information $I_{t'}, P_{Lt'}, P_{PV,t'}, t' = t + 1, \dots, t + 4$ is predicted by the average of the k most recent data. To evaluate the effect of information utilization degree, a series of $k = 1, 5, 10$ are examined. Then the policy regarding period t is applied to the system, leading to new system states. The process is repeated until the terminal condition is met.

4.3 | Scheduling result and discussion

Section 4.3.1 analyzes the convergence and relative performances of the proposed algorithm first. Then, for instances coupled and uncoupled with hydrogen systems, the outcomes of intra-day real-time scheduling are further examined in Section 4.3.2. In Section 4.3.3, to take into account the impact of the uncertainties from input water, load demand and photovoltaic power, system operating results of the integrated HPH system are described and analyzed.

4.3.1 | Convergence analysis and algorithm comparison

Since the setting of penalty parameters λ_w and λ_b in Equation (25) constitute the reward function and thus can impact the convergence and performance of the proposed algorithm, a series of experiments with different penalty parameters were conducted first. The improved DDPG algorithm with a specific ratio of λ_w and λ_b was trained for 3000 episodes, and then the algorithm was tested for 100,000 steps based on the environment with the same local load, PV power and inflow water under high solar intensity scenario. Besides, the performances of the short-sighted Myo algorithm and a series of MPC algorithms are summarized in Table 4. The comparative indicators including system power generation, power adoption rate and system revenue. Among the factors, the rate of power adoption is calculated as:

$$\text{Power adoption rate} = \frac{\sum_{t=1}^{24} (P_{Lt} + P_{wt} + P_{H_2,t})}{\sum_{t=1}^{24} (P_{Ht} + P_{PV,t} + P_{ft})}. \quad (34)$$

Figure 8 shows the convergence curves obtained through the proposed improved DDPG algorithm in solving the energy scheduling optimization problem considering the uncertain-

ties of inflow water, load demand and PV power. In the optimization process, after 1000 episodes, the system revenue converges and the rewards are almost stable. Besides, the results in Table 4 show that when $\frac{\lambda_w}{\lambda_b} < 1$, the system generation level and power adoption level are higher. However, when $\frac{\lambda_w}{\lambda_b} > 1$, in general, the system revenue tends to be larger. Considering both the power adoption rate and the system revenue, the ratio of $\lambda_w : \lambda_b = 1.0 : 0.9$ is adopted in the following experiments.

As for algorithm comparisons, the results in Table 4 show that the proposed improved DDPG algorithm has the best performance compared to the Myo and MPC algorithms. It yields the most system revenue and the highest power adoption rate. Combined with the results of system generation and hydropower, this implies that the proposed algorithm is superior in resource utilization. The main reason is that the improved DDPG algorithm is data-driven, and it fully uses the historical data to implicitly model system dynamics (via the constructed neural networks) and learn the optimal decisions considering the overall performance. On the contrary, Myo and MPC only use the information in a relatively limited amount of data and focus on the performance over a short period. Particularly, Myo has the worst performance as it makes decisions only based on the present information, and ignores the time series related effects on the future. For MPC, when more information is considered via more data (increased k), the system performance tends to be better. These evidences manifest the advantages of the data-driven algorithm, that is, by making full use of historical data and learning from past practices, it is possible to effectively deal with the uncertainties in the complicated integrated system and achieve sound performances.

4.3.2 | System benefit comparison with and without hydrogen devices

To verify the advantage of coupling hydrogen device access to the system on reducing hydro-PV curtailment and increasing system revenue, two cases are designed: Uncoupling the hydrogen system, and coupling the hydrogen system.

(1) System daily scheduling when uncoupling a hydrogen system

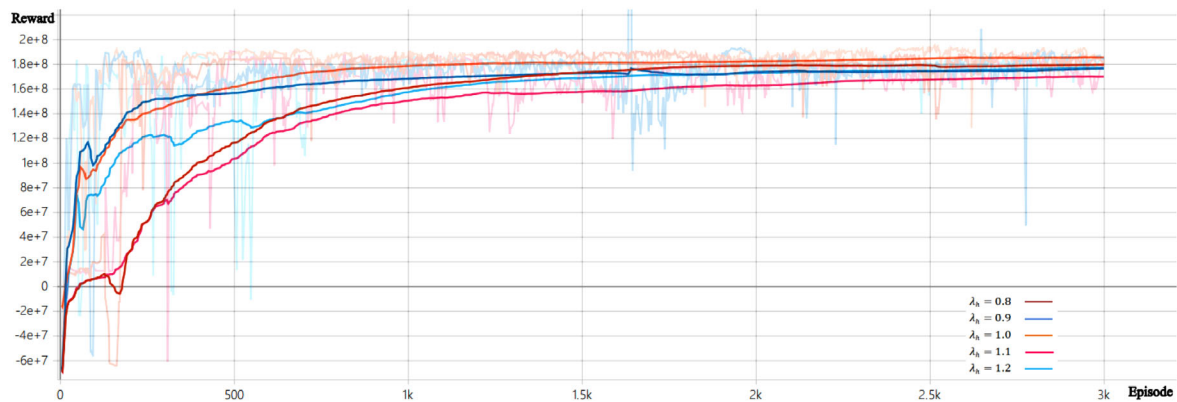
Here, we estimate the hydro-PV system's power production, power adoption rate, and revenue at its installed capacity without using hydrogen devices.

(2) System daily scheduling when coupling a hydrogen system. We assume that there are space to install electrolyzers, HFCs and hydrogen storage tanks in this scenario. Table 1 specifies the corresponding parameters.

Table 5 provides the comparison results regarding power generation, system revenue, and power curtailment under the two cases. It can be observed that under the second case of coupling a hydrogen system, the daily revenue rises by 1,000,000 RMB, and the power adoption rate increases by 5% to 8%.

TABLE 4 Results with different methods.

Method	Parameter	System generation (MWh)	Hydropower (MWh)	Power adoption rate	System revenue (RMB)
Improved DDPG	$\lambda_w : \lambda_b = 1.0:1.2$	14219	12047	0.371	6076924
	$\lambda_w : \lambda_b = 1.0:1.1$	14078	11906	0.368	6069084
	$\lambda_w : \lambda_b = 1.0:1.0$	14039	11867	0.366	6192985
	$\lambda_w : \lambda_b = 1.0:0.9$	13707	11534	0.357	6262954
	$\lambda_w : \lambda_b = 1.0:0.8$	13874	11702	0.361	6179911
Myo	-	9711	7538	0.198	3285301
MPC	$k = 1$	10200	8027	0.216	3822072
	$k = 5$	11127	8954	0.233	3949811
	$k = 10$	12231	10058	0.281	5319605

**FIGURE 8** Convergence curves of proposed DDPG approach with different λ_b .**TABLE 5** Comparison results when coupling and uncoupling a hydrogen system.

Scenario	System generation (MWh)	PV generation (MWh)	Hydropowerb (MWh)	System adoption rate	System revenue (RMB)
Case 1	High solar intensity	11,179	2173	0.28	5,039,121
	Low solar intensity	11,026	649	0.29	5,123,754
Case 2	High solar intensity	13,707	2173	0.35	6,262,954
	Low solar intensity	13,499	649	0.34	6,301,092

The experiment results implies that the coupling of hydrogen devices can effectively increase the adoption of power and further improve the system revenue. In addition, with the use of hydrogen devices, the generation capacity of the hydro-PV system is also increased.

4.3.3 | System scheduling results and corresponding analysis

This section, the optimal scheduling strategy and scheduling results from the HPH system under the representative scenar-

ios of high solar intensity and low solar intensity are described. Figures 9 and 10 display the hydropower, PV power and HFC output process when applying the proposed algorithm under high solar intensity and low solar intensity scenarios, respectively. The results show that the hydropower and hydrogen can be ideal compensations for photovoltaic power. In addition, Table 5 demonstrates that by coupling a hydrogen system, the energy adoption rate is increased both in both solar intensity scenarios.

On-grid power and the power adopted though electrolyzers under high solar intensity and low solar intensity scenarios are shown in Figures 11 and 12, respectively. It is interesting

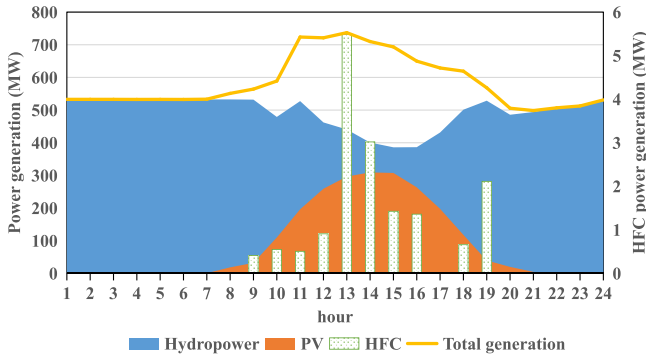


FIGURE 9 System power generation under high solar intensity scenario.

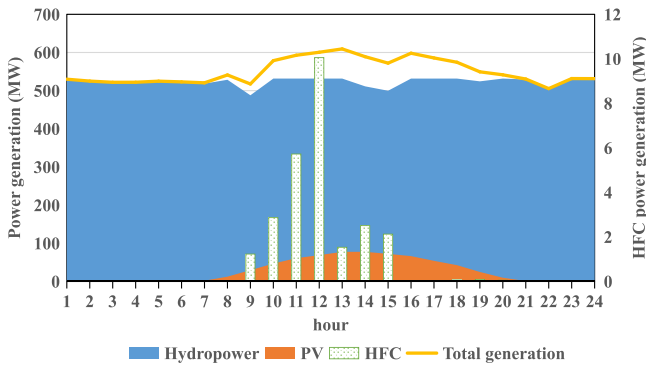


FIGURE 10 System power generation under low solar intensity scenario.

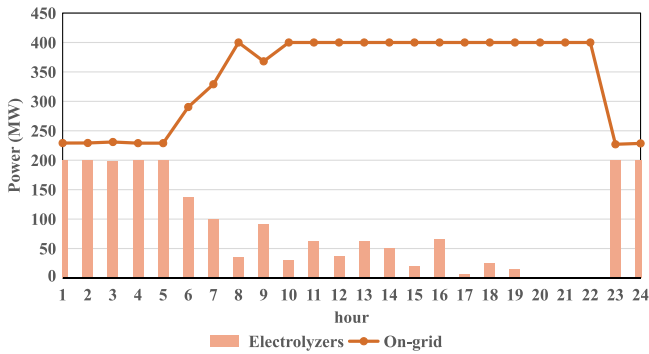


FIGURE 11 Power adoption under high solar intensity scenario.

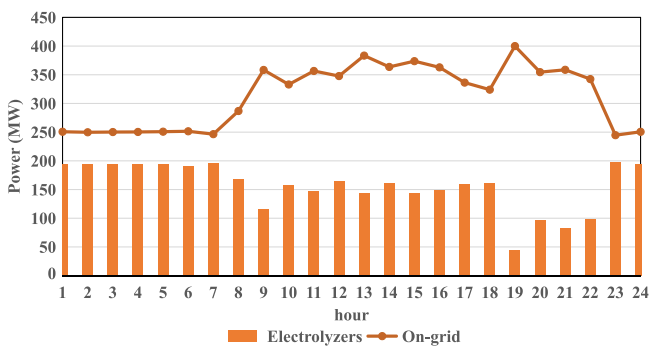


FIGURE 12 Power adoption under low solar intensity scenario.

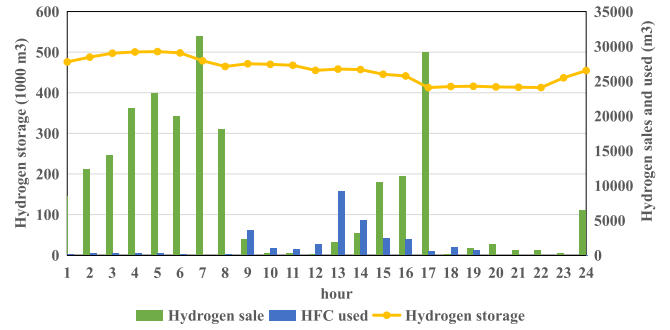


FIGURE 13 Hydrogen scheduling results under high solar intensity scenario.

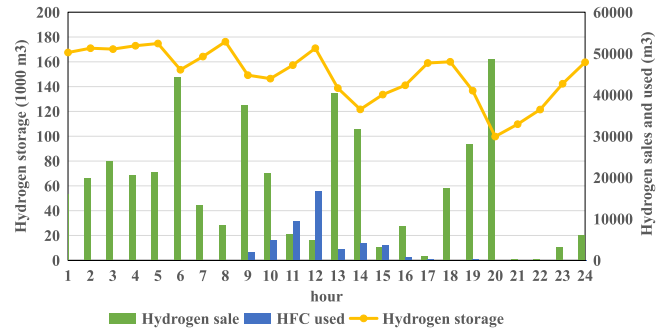


FIGURE 14 Hydrogen scheduling results under low solar intensity scenario.

that (1) the electrolyzers adopt more power under low solar intensity scenario compared with that under high solar intensity scenario. (2) The on-grid power in high solar intensity scenario is more than that in low intensity scenario. A possible reason is that hydropower station increases the output of the turbine by increasing the water discharge. In this process, to make full use of the discharged water resources and hydropower resources, the output level of the electrolyzers has been correspondingly increased, and the on-grid power is reduced. In high solar intensity scenario, to absorb the PV power, the hydropower station plays a more important role in energy storage and auxiliary output during the day-time. Therefore, the discharge flow is reduced, which affects the output level of the electrolyzers and thereby increases the on-grid power.

The hydrogen energy operation results from hydrogen sales, HFC adoption, and hydrogen storage under high solar intensity and low solar intensity scenarios are shown in Figures 13 and 14, respectively. As shown in the figures, the electrolyzers turn electric power to hydrogen, and hydrogen sales especially during valley time. The hydrogen fuel cells start-up in peak periods. As a result, the hydrogen system uses hydro-photovoltaic curtailment to produce hydrogen during the valley periods and gain the benefits of peak-valley arbitrage. In addition, through the comparison of scheduling results under the two scenarios, it can be observed that since the output level of the electrolyzers is higher under the low solar intensity scenario, the amount of hydrogen sales also increases along with the rise of hydropower generation. Besides, the revenue under the low

solar intensity scenario is slightly higher than that under the high solar intensity scenario.

In summary, hydrogen system is flexible and efficient. It does not depend on terrain and environment. It can well assist hydro-PV system to realize energy conversion and storage. Through the coupling of hydrogen system, the power curtailment is reduced with the increase of system revenue.

5 | CONCLUSION

This paper proposes an improved DDPG algorithm for the intra-day real-time scheduling of an integrated HPH system. Uncertainties from inflow water, PV generation and load demand are considered. The revenue of the system is optimized via decisions on hydropower generation flow, power used for hydrogen production, hydrogen sales amount and hydrogen used to generate power. Based on the case study of China's Longyangxia hydro-PV power plant, the following main conclusions can be made:

- (1) The coupling of the hydrogen system can reduce power curtailment of the original hydro-PV power plant along with the increase in system revenue (by 5% to 8%) and power adoption (by about 6%). In this respect, the hydrogen system can be applied as an ideal complementary of the sustainable development of multi-energy systems.
- (2) The improved DDPG algorithm that effectively optimizes the system is data-driven to adaptively learn system dynamics instead of requiring accurate models. In this way, the historical data can be fully used to deal with the 'dimensional disaster' problem occurring in modeling, if possible, the complex multi-energy system.

With the promotion and application of renewable energy systems containing hydrogen energy in a wider range, managers will make higher requests on the accuracy and speed of optimization algorithms. There are still some issues that could be addressed in the future work, and they are summarized as follows: First, more detailed transformation mechanisms between electricity and hydrogen could be involved, for example, more power generation and energy storage technologies. Second, more advanced data-driven algorithms for the electric-hydrogen system including cascade hydropower stations should be explored to solve the dimensional disaster problem in online scheduling. Finally, other scheduling objectives such as carbon emission level and system flexibility also deserve considerations.

NOMENCLATURE

Set

T set of decision time periods

Parameters

Δt period duration
 η_e conversion efficiency of power-hydrogen

η_f conversion efficiency of hydrogen-power
 η_t output coefficient of hydropower station
 C_{el} capacity of electrolyzer
 C_f maximum capacity of hydrogen fuel cells
 C_b maximum capacity of hydrogen storage tank
 I_t inflow water of reservoirs in period t
 $P_{L,t}$ local load during period t
 $P_{PV,t}$ PV power generated in period t
 $P_{H_2}^{max}$ maximum hydropower input
 $P_{w,t}^{max}$ maximum on-grid power
 p_h unit sale price of hydrogen
 $p_{w,t}$ on-grid electricity price in period t
 Q^{max} maximum power generation flow
 Q^{min} minimum power generation flow
 V^{max} minimum water storage of reservoir
 V^{min} maximum water storage of reservoir

Variables

F economic benefit of the system
 F_w income from the sale of on-grid electricity
 F_h income from the sale of hydrogen
 H_t net water head of hydropower station in period t
 H_{ft} volume of hydrogen inversed to power in period t
 H_{pt} volume of hydrogen produced through electrolyzers
 H_{st} volume of hydrogen storage at the end of period t
 $H_{sale,t}$ sales volume of hydrogen
 L_t^{up} water level in period t
 L_t^{down} tailwater elevation in period t
 P_{ft} power generated from hydrogen fuel cells
 P_{ct} water curtailment of the reservoir in period t
 P_{Ht} output hydropower during period t
 $P_{H_2,t}$ power used for hydrogen production in period t
 P_{wt} on-grid power during period t
 Pen_{wt} penalty of water storage at the end of period t
 Pen_{et} penalty of hydrogen storage at the end of period t
 Q_t power generation flow during period t
 V_t water storage of the reservoir at the beginning of period t

AUTHOR CONTRIBUTIONS

Yaping Zhao: Funding acquisition, methodology, software, writing - original draft. Jingsi Huang: Conceptualization, methodology, writing - original draft. Endong Xu: Data curation, software, validation. Jianxiao Wang: Supervision, writing - review and editing. Xiaoyun Xu: Software, supervision.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (Grant Number 72001145); the Ministry of Education Program in Humanities and Social Sciences (Grant Number 20YJC630226); and the Natural Science Foundation of Guangdong Province (Grant Number 2022A1515011235). The authors thank anonymous reviewers for their valuable suggestions and comments.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Jingsi Huang  <https://orcid.org/0000-0002-6830-6351>

Jianxiao Wang  <https://orcid.org/0000-0001-9871-5263>

REFERENCES

- IEA: Renewable energy market update: Outlook for 2020 and 2021 (2020)
- NEA: Releases the grid-connected operation of renewable energy in 2019 (2020)
- Amin, M., Mehdi, M.: A comprehensive review on coupling different types of electrolyzer to renewable energy sources. *Energy* 158, 632–655 (2018)
- Li, H., Qin, B., Jiang, Y., Zhao, Y., Shi, W.: Data-driven optimal scheduling for underground space based integrated hydrogen energy system. *IET Renewable Power Gener.* 16(12), 2521–2531 (2022)
- Zivar, D., Kumar, S., Foroozesh, J.: Underground hydrogen storage: A comprehensive review. *Int. J. Hydrogen Energy* 46(45), 23436–23462 (2021)
- Ball, M., Wietschel, M.: The future of hydrogen—opportunities and challenges. *Int. J. Hydrogen Energy* 34(2), 615–627 (2009)
- Fan, H., Yan, T., Guo, S.: Methods on scheduling of hydro-photovoltaic complementary system. In: *E3S Web of Conferences*, vol. 257, p. 01034. EDP Sciences, Les Ulis (2021)
- Guo, C., Huang, Y., Wang, L., Pan, Y., Dai, S.: Overview for collaborative dispatch of stochastic power supply and conventional power supply. *J. Electric Power Sci. Technol.* 29(04), 5–11 (2014)
- Ming, B., Liu, P., Guo, S., Zhang, X., Feng, M., Wang, X.: Optimizing utility-scale photovoltaic power generation for integration into a hydropower reservoir by incorporating long-and short-term operational decisions. *Appl. Energy* 204, 432–445 (2017)
- An, Y., Fang, W., Ming, B., Huang, Q.: Theories and methodology of complementary hydro/photovoltaic operation: Applications to short-term scheduling. *J. Renew. Sustain. Energy* 7(6), 063133 (2015)
- Li, F.F., Qiu, J.: Multi-objective optimization for integrated hydro-photovoltaic power system. *Appl. Energy* 167, 377–384 (2016)
- Ming, B., Liu, P., Cheng, L., Zhou, Y., Wang, X.: Optimal daily generation scheduling of large hydro-photovoltaic hybrid power plants. *Energy Convers. Manage.* 171, 528–540 (2018)
- Huang, J., Wu, X., Zheng, Z., Huang, Y., Li, W.: Multi-objective optimal operation of combined cascade reservoir and hydrogen system. *IEEE Trans. Ind. Appl.* 58(2), 2836–2847 (2021)
- Yang, Z., Elia, P.C., Anders, L., Jinyue, Y.: Comparative study of hydrogen storage and battery storage in grid connected photovoltaic system: Storage sizing and rule-based operation. *Appl. Energy* 201, 397–411 (2017)
- Di, L., Bende, W., Yaodong, W., Huicheng, Z., Qiuhua, L., Yong, P., et al.: Optimal operation of cascade hydropower stations using hydrogen as storage medium. *Appl. Energy* 137, 56–63 (2015)
- Petrollese, M., Valverde, L., Cocco, D., Cau, G., Guerra, J.: Real-time integration of optimal generation scheduling with mpc for the energy management of a renewable hydrogen-based microgrid. *Appl. Energy* 166, 96–106 (2016)
- Li, B., Roche, R.: Optimal scheduling of multiple multi-energy supply microgrids considering future prediction impacts based on model predictive control. *Energy* 197, 117180 (2020)
- Zhou, S., Hu, Z., Gu, W., Jiang, M., Chen, M., Hong, Q., et al.: Combined heat and power system intelligent economic dispatch: A deep reinforcement learning approach. *Int. J. Electr. Power Energy Syst.* 120, 106016 (2020)
- Jin-Quan, Z., Xue, X., Chun-Lei, X., Wei, H., Xue-Wei, S.: Review on application of new generation artificial intelligence technology in power system dispatching and operation. *Autom. Electr. Power Syst.* 44(24), 1–10 (2020)
- Schulman, J., Wolski, F., Dhariwal, P., Radford, A., Klimov, O.: Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347* (2017)
- Gong, Y., Liu, P., Liu, Y., Huang, K.: Robust operation interval of a large-scale hydro-photovoltaic power system to cope with emergencies. *Appl. Energy* 290, 116612 (2021)
- Zhang, B., Hu, W., Cao, D., Huang, Q., Chen, Z., Blaabjerg, F.: Economical operation strategy of an integrated energy system with wind power and power to gas technology—a drl-based approach. *IET Renew. Power Gener.* 14(17), 3292–3299 (2020)
- Huang, Q., Hu, W., Zhang, G., Cao, D., Liu, Z., Huang, Q., et al.: A novel deep reinforcement learning enabled agent for pumped storage hydro-wind-solar systems voltage control. *IET Renew. Power Gener.* 15(16), 3941–3956 (2021)
- Chen, T., Gao, C., Song, Y.: Optimal control strategy for solid oxide fuel cell-based hybrid energy system using deep reinforcement learning. *IET Renew. Power Gener.* 16(5), 912–921 (2022)
- Schulman, J., Levine, S., Abbeel, P., Jordan, M., Moritz, P.: Trust region policy optimization. In: *International Conference on Machine Learning*, pp. 1889–1897. International Machine Learning Society, Madison, WI (2015)
- Li, H., Wan, Z., He, H.: Constrained ev charging scheduling based on safe deep reinforcement learning. *IEEE Trans. Smart Grid* 11(3), 2427–2439 (2019)
- Cai, X., McKinney, D.C., Lasdon, L.S., Watkins Jr, D.W.: Solving large non-convex water resources management models using generalized benders decomposition. *Oper. Res.* 49(2), 235–245 (2001)
- Francesco, C., Damian, R.F., Nicola, M., Alessandro, M., Laura, V.: Polygeneration system based on pemfc, cpvt and electrolyzer: Dynamic simulation and energetic and economic analysis. *Appl. Energy* 192, 530–542 (2017)
- Amini, B.H., Mohammadmehdi, C., Aneeb, C., Hassan, Q.: A highly efficient hydrogen generation electrolysis system using alkaline zinc hydroxide solution. *Int. J. Hydrogen Energy* 44(1), 72–81 (2019)
- Li, J., Lin, J., Song, Y., Xing, X., Fu, C.: Operation optimization of power to hydrogen and heat (p2hh) in adn coordinated with the district heating network. *IEEE Trans. Sustain. Energy* 10(4), 1672–1683 (2019)
- Ding, H., Hu, Z., Song, Y.: Stochastic optimization of the daily operation of wind farm and pumped-hydro-storage plant. *Renew. Energy* 48, 571–578 (2012)
- Majidi, M., Nojavan, S., Zare, K.: Optimal stochastic short-term thermal and electrical operation of fuel cell/photovoltaic/battery/grid hybrid energy system in the presence of demand response program. *Energy Convers. Manage.* 144, 132–142 (2017)
- Zhao, T., Cai, X., Yang, D.: Effect of streamflow forecast uncertainty on real-time reservoir operation. *Adv. Water Resour.* 34(4), 495–504 (2011)
- Ogliari, E., Dolara, A., Manzolini, G., Leva, S.: Physical and hybrid methods comparison for the day ahead pv output power forecast. *Renew. Energy* 113, 11–21 (2017)
- Yang, H.T., Huang, C.M., Huang, Y.C., Pai, Y.S.: A weather-based hybrid method for 1-day ahead hourly forecasting of pv power output. *IEEE Trans. Sustain. Energy* 5(3), 917–926 (2014)
- Carl-Jochen, W.: Into the hydrogen energy economy—milestones. *Int. J. Hydrogen Energy* 30(7), 681–685 (2005)

How to cite this article: Zhao, Y., Huang, J., Xu, E., Wang, J., Xu, X.: A data-driven scheduling approach for integrated electricity-hydrogen system based on improved DDPG. *IET Renew. Power Gener.* 1–14 (2023). <https://doi.org/10.1049/rpg.2.12693>