



OPEN ACCESS

EDITED BY

Wenping Zhang,
Tianjin University, China

REVIEWED BY

Linfei Yin,
Guangxi University, China
Haris M. Khalid,
University of Dubai, United Arab Emirates
Qiuye Sun,
Northeastern University, China

*CORRESPONDENCE

Lingfang Sun,
✉ sunlf@neepu.edu.cn

RECEIVED 15 July 2024

ACCEPTED 26 September 2024

PUBLISHED 10 October 2024

CITATION

Wang E, Yuan L, Zeng F, Liu X, Liu J, Sun L and Zhuang M (2024) Load frequency optimal control of the hydropower-photovoltaic hybrid microgrid system based on the off-policy integral reinforcement learning algorithm. *Front. Energy Res.* 12:1464722. doi: 10.3389/fenrg.2024.1464722

COPYRIGHT

© 2024 Wang, Yuan, Zeng, Liu, Liu, Sun and Zhuang. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Load frequency optimal control of the hydropower-photovoltaic hybrid microgrid system based on the off-policy integral reinforcement learning algorithm

Enzhong Wang¹, Lin Yuan¹, Fanfei Zeng², Xiaoheng Liu², Jiannan Liu², Lingfang Sun^{3*} and Min Zhuang³

¹Guoneng Qinghai Yellow River MaerDang Hydropower Development Co., Ltd., Qinghai, China, ²Beijing SP Zhishen Control Technology Co., Ltd., Beijing, China, ³School of Automation Engineering, Northeast Electric Power University, Jilin, China

With the promotion and development of clean energy, it is challenging to ensure the optimization of control performance in frequency control of the hydropower-photovoltaic hybrid microgrid system caused by the output power fluctuation of photovoltaic power generation. In this study, an optimal load frequency controller (LFC) for a hydropower-photovoltaic hybrid microgrid system was designed to improve the dynamic response when the load and photovoltaic output power are perturbed based on the off-policy integral reinforcement learning algorithm. First, a mechanism model of the hydropower-photovoltaic hybrid microgrid system was established. Next, the LFC problem was transformed into a zero-sum game control problem based on the characteristics of the power system. Subsequently, three neural networks were employed to approximate the Nash equilibrium solution of the zero-sum game with historical input and output data when the system dynamics are completely unknown. Finally, simulation experiments were conducted to verify the effectiveness and optimality of the proposed method. The introduction of this method provides a new perspective for frequency control for the hydropower-photovoltaic hybrid microgrid system.

KEYWORDS

hydropower-photovoltaic hybrid microgrid system, load frequency controller, off policy integral reinforce learning algorithm, data-based optimal control, neural networks

1 Introduction

With the development of the national economy and society, the contradiction between increasing energy demand and energy shortages has become increasingly obvious (Gilani et al., 2020; Patnaik et al., 2020; Zhang and Kong, 2022). Traditional thermal power generation causes problems such as the consumption of nonrenewable energy and excessive carbon emissions (Ahmad et al., 2018; Cowie et al., 2020; Olabi and Abdelkareem, 2022).

Abbreviations: LFC, load frequency controller; PV, photovoltaic; PID, proportional-integral-derivative; ADP, adaptive dynamic programming; IRL, integral reinforcement learning.

Hydropower and solar energy have attracted the attention of researchers owing to their renewable and environment-friendly nature (Gielen et al., 2019; Zepter et al., 2019).

However, photovoltaic (PV) power generation is intermittent, leading to unstable output power and microgrid frequency oscillations (Thirunavukkarasu and Sawle, 2021; Chen et al., 2022; Wu and Yang, 2023). To ensure the frequency stability of a microgrid, it is necessary to supplement controllable power sources, such as hydroelectric units or energy storage devices, to fill the power deficit, which can effectively maintain the microgrid frequency stability (Coban et al., 2022). The power quality of PV power systems can be improved by utilizing a control algorithm for controllable power sources, which is applied to obtain an optimal load frequency controller (LFC) system (Papaefthymiou et al., 2010; Ma et al., 2014; Dhundhara and Verma, 2020). Some researchers focus on the suppression of local load fluctuations and their interactions with the distribution system (Khalid et al., 2022). Additionally, the role of ancillary services and the integration of renewable energy should also be addressed upon introduction to minimize fluctuations and cover intermittency (Khalid et al., 2022; Rehman et al., 2024; Osman et al., 2022).

Owing to their simple structure and ease of implementation, proportional-integral-derivative (PID) control methods are widely used in microgrid LFC (Mohamed et al., 2020; Nisha and Jamuna, 2022). Ray et al. (2011) utilized a PI controller to regulate the frequency of a microgrid and achieve the required frequency ratings. Guha et al. (2021) designed a fractional-order PID method to solve the frequency stabilization problem of microgrid systems with uncertain parameters. Huang et al. (2021) used fuzzy reasoning in PID to improve the control performance of a hydraulic turbine regulation system.

Many practical power systems can only be partially modeled, and models of unknown parts are unavailable (Ganguly et al., 2018; Li et al., 2023; Wu and Yang, 2023). Dynamic characteristics of droop-controlled inverters are evaluated by a reduce-order small-signal transfer function model, which is designed on the basis of the Jordan continued-fraction expansion to provide a preprocessing method for real-time power system simulation (Wang et al., 2020). Therefore, owing to the insensitivity to the dynamics of the unmodeled parts of the controlled object, adaptive control methods have been proposed by continuously identifying system parameters to achieve the ideal control effect. Adaptive control methods can be used to resolve problems arising from parameter variations in the LFC of a power system. Zeng et al. (2015) designed a port-controlled Hamiltonian system that decomposed nonlinear control into stabilizing control with a given equilibrium point and proposed L-2 adaptive control for application to a hydroelectric generator unit. Fang et al. (2011) effectively improved the dynamic performance of the hydraulic turbine regulation process using an improved particle swarm optimization algorithm, which was applied to the optimal design of the parameters of a hydraulic turbine regulation system to achieve an optimal positive setting of the parameters. Tran et al. (2021) used a combination of second-order sliding film control and a state estimator for frequency regulation to reduce the number of overtones. Although these methods can achieve better control performance, they have not been widely popularized in practical power systems owing to their complexity and difficulty.

The adaptive dynamic programming (ADP) algorithm is an emerging intelligent control algorithm that solves the problem of dimensional disasters caused by the traditional dynamic programming (DP) method (Werbos, 1992; Vamvoudakis and Lewis, 2010; Lewis et al., 2012; Bellman and Dreyfus, 2015) and is suitable for systems with a high degree of nonlinearity. Shuai et al. (2020) used a hybrid ADP algorithm to achieve optimal operation of gas and electric systems. Xue et al. (2022) used ADP for the real-time scheduling of battery heat storage tank integrated heat and power systems, providing optimal economic operation strategies. The off-policy integral reinforcement learning (IRL) algorithm is proposed based on the theory of the ADP algorithm, which can explore system information with historical input and output data, thereby overcoming the difficulty of traditional ADP relays on neural network weights in the training process to find the continuous excitation function. Chai et al. (2017) used the game theory to solve multi-objective trajectory optimization problems for aerial vehicles. Song et al. (2019) proposed an off-policy IRL algorithm to solve an optimal control problem with partially known system dynamics. Based on the ADP algorithm, this paper proposes an integral reinforcement learning method that requires only the historical input-output data of the system, allowing for optimal solutions even when the system dynamics are completely unknown.

To the best of our knowledge, in the hydropower-photovoltaic hybrid microgrid system, the challenges of considering system disturbances and employing model-free methods for frequency control are quite evident. Traditional frequency control methods typically rely on a mathematical model of the system and assume that disturbances are known or predictable. However, in real microgrid systems, disturbances such as load variations and fluctuations in renewable energy output are often unpredictable, and obtaining an accurate model of the system can be difficult or complex. The existing reinforcement learning methods for frequency control in the hydropower-photovoltaic hybrid microgrid systems have not simultaneously addressed disturbances in the system and utilized the model-free approaches, which motivates our study. The focus of this paper is on how to abstract a hybrid power generation system with disturbances as a zero-sum game problem and solve it using the proposed model-free method. This approach provides a theoretical foundation and basis for the grid integration of a series of photovoltaic combined power generation systems. The main contributions of this article are as follows:

1. A hydropower-photovoltaic hybrid microgrid system model was constructed on the basis of the mechanistic modeling of the hydraulic turbine and photovoltaic power generation, meanwhile treating the photovoltaic power generation perturbed as the disturbance term.
2. Based on the power generation characteristics, the secondary frequency modulation control signal was used as the control vector, and the input system load frequency and solar energy power were used as the perturbation vectors of the hydropower-photovoltaic microgrid power system, which transforms the LFC problem into a zero-sum optimal control problem. By solving the Nash equilibrium of the zero-sum game, the optimal control rate and the maximum disturbance that the system can withstand can be obtained,

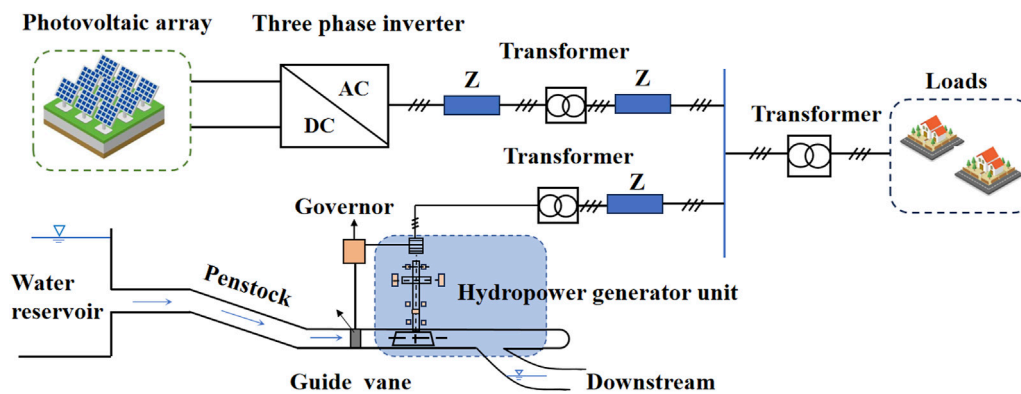


FIGURE 1
Main structure of the independent microgrid.

thereby controlling the load frequency of the hydropower-photovoltaic hybrid microgrid system.

3. An off-policy IRL algorithm was adopted to resolve the zero-sum optimal control problem in which three networks were employed to approximate the Nash equilibrium point of the zero-sum game to obtain the optimal LFC of the hybrid system. The proposed method overcomes the limitation of existing solution methods that require precise system model information.

2 Problem statement

The hydropower-photovoltaic microgrid power system effectively exploits the inherent frequency regulation advantages of hydropower units while integrating solar energy generation resources within the same regional grid. This hybrid system aims to enhance the overall frequency quality of the microgrid by balancing both renewable energy inputs and electrical load demand. However, such integration significantly increases the operational requirements for the Load Frequency Control (LFC) controller. In this study, Figure 1 outlines the core structure of the system: a power busbar connects hydropower units (HP), photovoltaic generation units (PV), and electrical loads. Specifically, the PV units are connected to the alternating current (AC) microgrid through direct current (DC) to alternating current (AC) conversion using DC/AC inverters.

The frequency stability of this isolated microgrid relies heavily on maintaining an active power balance within the network. Variations in electrical load and the intermittent, fluctuating output from photovoltaic sources can disturb this balance, leading to changes in system frequency. A central feature of the hydropower-photovoltaic microgrid system is the hydro-turbine generator, which is responsible for providing rotational reserves that help regulate frequency by adjusting the mechanical input to the turbine. This response compensates for any mismatch between generation and demand, ensuring system stability.

The hydropower units play a critical role in Load Frequency Control (LFC) tasks. The primary function of the LFC system is to regulate water flow into the turbines of the hydropower generators. It achieves this by dynamically adjusting the active power output of

the hydropower units in real-time, depending on the load and intermittent power output from the solar resources. This real-time control is vital for compensating fluctuations in both solar power production and load changes, stabilizing the generator speed, and ultimately controlling the frequency of the microgrid.

To improve the effectiveness of the system and minimize control costs, an advanced optimal load frequency controller was designed, utilizing an off-policy Inverse Reinforcement Learning (IRL) algorithm. This controller ensures the stability of the grid-connected voltage in the hydropower-photovoltaic microgrid by optimizing the dynamic allocation of power resources. In essence, it manages the trade-offs between ensuring grid frequency stability and maintaining operational cost efficiency, leading to a robust, reliable, and sustainable microgrid power system.

3 Materials and methods

3.1 Establishment of the hydro turbine group model

The hydro turbine group consisted of hydro turbines, governors, and generators. A turbine group model was established for each part.

The equations of moment and flow of the hydro turbine are expressed as Equation 1:

$$\begin{cases} \Delta m_t(t) = n_x \Delta x(t) + n_y \Delta y(t) + n_h \Delta h(t) \\ \Delta q(t) = n_{qx} \Delta x(t) + n_{qy} \Delta y(t) + n_{qh} \Delta h(t), \end{cases} \quad (1)$$

where $\Delta m_t(t)$ and $\Delta q(t)$ represent the increments in torque and flow, respectively; $\Delta y(t)$, $\Delta h(t)$, and $\Delta x(t)$ indicate the relative deviation variables of the guide vane opening, water head, and hydro turbine speed, respectively; n_x , n_y , and n_h represent the transfer coefficients of the hydro turbine torque to the speed, guide vane opening and water head, respectively; n_{qx} , n_{qy} and n_{qh} are transfer coefficients of the hydro turbine flow to the speed, guide vane opening and the water head, respectively. Under stable conditions, each transfer coefficient is regarded as a constant.

When $n = \frac{n_{qy} n_h}{n_y} - n_{qh}$, the transfer function from the guide vane increment $\Delta y(t)$ to the torque increment $\Delta m_t(t)$ of the hydro turbine is as follows:

$$G_{ym}(s) = \frac{\Delta m(s)}{\Delta y(s)} = n_y \frac{1 - nT_w s}{1 + n_{qh}T_w s}, \quad (2)$$

where T_w indicates the inertia time constant of the water flow.

The hydro turbine governor can be simplified as a first-order inertial link by ignoring the nonlinear factors, as Equation 3:

$$G_y(s) = \frac{\Delta y(s)}{u(s)} = \frac{1}{T_y s + 1}, \quad (3)$$

where T_y represents the response time constant of the hydro-turbine governor. The corresponding differential equation is as follows:

$$\frac{d\Delta y(t)}{dt} = -\frac{1}{T_y} \Delta y(t) + \frac{1}{T_y} u(t) \quad (4)$$

Equation 2 is substituted into Equation 4 after the Laplace transformation. The differential equation is obtained by the Laplace transformation of Equation 2, and Equation 4 is substituted into the differential equation to obtain the hydro turbine differential equation as follows:

$$\frac{d\Delta m_t(t)}{dt} = \frac{1}{n_{qh}T_w} \left[-\Delta m_t(t) + \left(n_y + n_y n \frac{T_w}{T_y} \right) \Delta y(t) - n_y n \frac{T_w}{T_y} u(t) \right] \quad (5)$$

The second-order model of the generator includes the rotor rotation motion equation and the equation that characterizes the relationship between the power angle and speed, as follows:

$$\begin{cases} \frac{d\Delta\delta(t)}{dt} = \omega_0 \Delta x(t) \\ \frac{dx(t)}{dt} = \frac{1}{T_J} [\Delta m_t(t) - \Delta P_e(t) - D\Delta x(t)], \end{cases} \quad (6)$$

where δ indicates the generator power angle, which remains constant under stable operating conditions, and ΔP_e indicates the electromagnetic power increment, which is equivalent to the relative value of the load increment ΔP_L . $\omega_0 = 100\pi$ indicates the synchronous electrical angular velocity, D is the damping coefficient, and T_J indicates the inertia time constant of the

generator. ΔP_S is the sum of the incremental relative values of power supplied by all sources within the microgrid. From the Equation 6, we get $\Delta P_S - \Delta P_L = T_M \frac{d\Delta f(t)}{dt} + D_M \Delta f(t)$. T_M represents the equivalent inertia time coefficient of the system, which is equal to the weighted sum of the inertia coefficients of all the generators in the system. D_M represents the equivalent load-damping factor of the system.

By combining Equations 4–6, the following mathematical model of the hydro-turbine group can be obtained:

$$\begin{cases} \frac{d\delta(t)}{dt} = \omega_0 \Delta x(t) \\ \frac{d\Delta x(t)}{dt} = \frac{1}{T_J} [\Delta m_t(t) - \Delta P_e(t) - D\Delta x(t)] \\ \frac{d\Delta m_t(t)}{dt} = \frac{1}{n_{qh}T_w} \left[-\Delta m_t(t) + \left(n_y + n_y n \frac{T_w}{T_y} \right) \Delta y(t) - n_y n \frac{T_w}{T_y} u(t) \right] \\ \frac{d\Delta y(t)}{dt} = -\frac{1}{T_y} \Delta y(t) + \frac{1}{T_y} u(t) \end{cases} \quad (7)$$

3.2 Establishment of the photovoltaic model

PV panels convert solar energy into electrical energy based on PV effects. The main body of the frequency control in this study was the hydropower unit. Therefore, in this subsection, a first-order model with time constant T_{solar} was used to express the frequency characteristics of the PV power as follows:

$$\frac{d\Delta P_{PV}}{dt} = -\frac{1}{T_{solar}} \Delta P_{PV} + \frac{1}{T_{solar}} \Delta P_{solar}, \quad (8)$$

where T_{solar} represents the time constant of the PV power system, ΔP_{PV} represents the output power of the PV power system, and ΔP_{solar} represents the solar power. Solar power has obvious volatility, as PV panels are easily affected by light intensity and ambient temperature. Therefore, solar power is regarded as a disturbance term. A block diagram of the model is shown in the Figure 2.

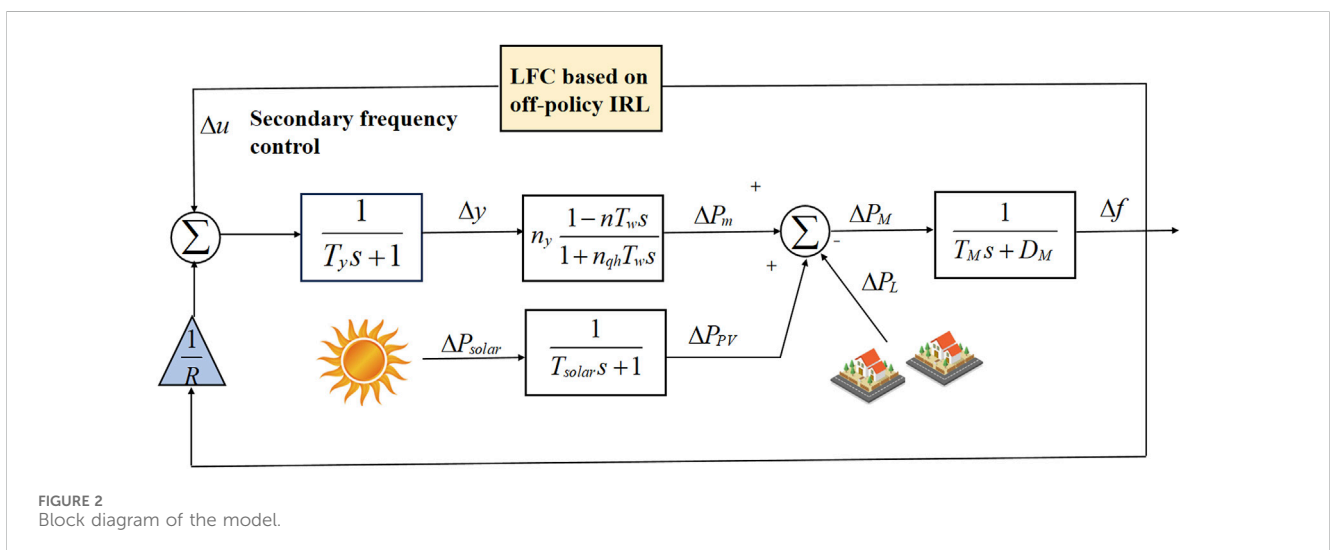


FIGURE 2 Block diagram of the model.

3.3 Establishment of the hydropower-photovoltaic microgrid power system model

The transient changes in the voltage and power angle of the system can be ignored in the frequency control analysis. Therefore, in the analysis process of LFC, $\Delta m_t(t) = \Delta P_m(t)$ is the incremental relative value of the output power of the hydro turbine group, the relative value of the electromagnetic power increment $\Delta P_e(t)$ is replaced by the relative value of the load increment $\Delta P_L(t)$, and the relative value of the speed deviation $\Delta x(t)$ is equal to the relative value of the frequency deviation $\Delta f(t)$.

By combining Equations 7, 8, the hydro-photovoltaic microgrid power system can be derived as Equation 9:

$$\begin{cases} \frac{d\delta(t)}{dt} = w_0 \Delta f(t) \\ \frac{d\Delta f(t)}{dt} = \frac{1}{T_f} [\Delta P_m(t) - \Delta P_L(t) - D\Delta f(t)] \\ \frac{d\Delta P_m(t)}{dt} = \frac{1}{n_{gh}T_w} \left[-\Delta P_m(t) + \left(n_y + n_y n \frac{T_w}{T_y} \right) \Delta y(t) - n_y n \frac{T_w}{T_y} u(t) \right] \\ \frac{d\Delta y(t)}{dt} = -\frac{1}{T_y} \Delta y(t) + \frac{1}{T_y} u(t) \\ \frac{d\Delta P_{PV}}{dt} = -\frac{1}{T_{solar}} \Delta P_{PV} + \frac{1}{T_{solar}} \Delta P_{solar} \end{cases} \quad (9)$$

Here, $x(t) = [x_1(t) \ x_2(t) \ x_3(t) \ x_4(t) \ x_5(t)]^T = [\delta(t) \ f(t) \ \Delta P_m(t) \ \Delta y(t) \ \Delta P_{PV}]^T$ are state variable. The load frequency $\Delta P_L(t)$ and solar energy ΔP_{solar} of the input system are considered as the elements in the disturbance vector, and the PV output power fluctuation and load power change cause the power supply and demand of microgrid to lose balance. The system disturbance vector can be obtained as follows: $w(t) = [\Delta P_L(t) \ \Delta P_{solar}(t)]^T$; $u = u(t)$ is the control vector signal. Therefore, the load frequency model of the hydropower-photovoltaic system can be obtained as follows:

$$\dot{x} = Ax + Bu + Fw, \quad (10)$$

where the system state variable $x \in R^{5 \times 1}$ is the system control variable and $u \in R^{1 \times 1}$ is the disturbance variable $w \in R^{2 \times 1}$. The system matrix is expressed as follows:

$$A = \begin{bmatrix} 0 & w_0 & 0 & 0 & 0 \\ 0 & -\frac{D}{T_f} & \frac{1}{T_f} & 0 & 0 \\ 0 & 0 & -\frac{1}{n_{gh}T_w} & \frac{n_y}{n_{gh}T_w} \left(1 + n \frac{T_w}{T_y} \right) & 0 \\ 0 & 0 & 0 & -\frac{1}{T_y} & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{T_{solar}} \end{bmatrix},$$

$$B = \begin{bmatrix} 0 \\ 0 \\ \frac{n_y n}{n_{gh}T_y} \\ \frac{1}{T_y} \\ 0 \end{bmatrix}, F = \begin{bmatrix} 0 & 0 \\ -\frac{1}{T_f} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & \frac{1}{T_{solar}} \end{bmatrix}$$

Thus far, the load frequency control problem of the hydropower-photovoltaic microgrid power system has been transformed into a zero-sum game optimal control problem, wherein the input of the governor was taken as the control variable and the load frequency and solar power were taken as the disturbance variables.

4 Results of the optimal controller based on off-policy IRL algorithm

The hydropower-photovoltaic microgrid power system model was established using Equation 10, where x, u and w are the state, control input, and disturbance input of the system, respectively. $x = 0$ is the equilibrium point of the hydropower-photovoltaic microgrid power system. The infinite-horizon performance index function can be designed as follows:

$$J(x(0), u, w) = \int_0^\infty \{r(x(t), u, w)\}, \quad (11)$$

where in the utility function can be described as Equation 12:

$$r(x(t), u, w) = x^T R x + u^T S u - w^T T w, \quad (12)$$

where the coefficient matrices R, S and T are real symmetric positive definite matrices. Please note that this performance indicator has clear economic significance, in which $x^T R x$ is the penalty cost when the quality of power supply at the node i deviates from the system's stable value and $u^T S u - w^T T w$ is the control costs and disturbance costs incurred by node i to reduce penalty costs.

The purpose of the zero-sum game is to solve for an optimal control that satisfies Equation 13.

$$V^*(x(0)) = \inf_u \sup_w \{J(x(0), u, w)\} \quad (13)$$

The zero-sum game selects to minimize the player set u and maximize the player set w , the saddle point u^* and w^* must satisfy the following inequality Equation 14:

$$J(x, u^*, w) \leq J(x, u^*, w^*) \leq J(x, u, w^*) \quad (14)$$

When there is a unique set of solutions that satisfy the following Nash equilibrium condition Equation 15:

$$V^*(x) = \inf_u \sup_w \{J(x, u, w)\} = \sup_w \inf_u \{J(x, u, w)\}, \quad (15)$$

the cost function of every player can be written as Equation 16:

$$V(x(t)) = \int_t^\infty \{x^T R x + u^T S u - w^T T w\} dt \quad (16)$$

Using the Leibniz formula and differentiating Equation 6, the Bellman equation of the zero-sum game can be obtained as follows:

$$H(x, \nabla V, u, w) = x^T R x + u^T S u - w^T T w + \nabla V^T (Ax + Bu - Fw), \quad (17)$$

where $\nabla V = \frac{\partial V}{\partial x}$. The control and disturbance inputs can be obtained as Equations 18, 19:

$$\frac{\partial H}{\partial u} = 2uS + \nabla V^T B = 0 \tag{18}$$

$$\frac{\partial H}{\partial w} = -2wT + \nabla V^T F = 0 \tag{19}$$

The optimal control policy u^* and the optimal disturbance w^* can be derived as follows:

$$u^* = -\frac{1}{2}S^{-1}B^T \nabla V \tag{20}$$

$$w^* = \frac{1}{2}T^{-1}F^T \nabla V \tag{21}$$

The Hamilton-Jacobi-Bellman equation can be obtained by substituting Equations 20, 21 into Equation 17 as follows:

$$0 = x^T R x + \nabla V^T A x - \frac{1}{4} \nabla V^T B S^{-1} B^T \nabla V + \frac{1}{4} \nabla V^T F T^{-1} F^T \nabla V \tag{22}$$

The following equations were used to update the control and disturbance policies as Equations 23, 24:

$$u^{[k+1]} = -\frac{1}{2}S^{-1}B^T \nabla V^{[k]} \tag{23}$$

$$w^{[k+1]} = \frac{1}{2}T^{-1}F^T \nabla V^{[k]}, \tag{24}$$

where the superscript $[k]$ represents the number of steps in the iteration process.

The Equation 11 can be transformed as Equation 25:

$$\dot{x} = Ax + Bu^{[k]} + Fw^{[k]} + B(u - u^{[k]}) + F(w - w^{[k]}) \tag{25}$$

The Equation 22 can be rewritten as follows:

$$\begin{aligned} & V^{[k]}(x(t + \Delta t)) - V^{[k]}x(t) \\ &= \int_t^{t+\Delta t} \Delta V^{[k]T} \dot{x} d\tau \\ &= \int_t^{t+\Delta t} \Delta V^{[k]T} (Ax + bu + Fw) d\tau \\ &+ \int_t^{t+\Delta t} \Delta V^{[k]T} (B(u - u^{[k]}) + F(w - w^{[k]})) \\ &= - \int_t^{t+\Delta t} (x^T R x + u^{[k]T} S u^{[k]} - w^{[k]T} T w) d\tau \\ &+ \int_t^{t+\Delta t} \Delta V^{[k]T} (B(u - u^{[k]}) + F(w - w^{[k]})) \end{aligned} \tag{26}$$

By deriving Equation 23, we get $\nabla V^{[k]T} B = -2u^{[k+1]}S$ and $\nabla V^{[k]T} F = 2w^{[k+1]}T$; upon substituting these parameters into Equation 26, the following equation was obtained:

$$\begin{aligned} & V^{[k]}(x(t + \Delta t)) - V^{[k]}x(t) \\ &= - \int_t^{t+\Delta t} (x^T R x + u^{[k]T} S u^{[k]} - w^{[k]T} T w) d\tau \\ &+ \int_t^{t+\Delta t} -2(u^{[k+1]T} S (u - u^{[k]}) - w^{[k+1]T} T (w - w^{[k]})) \end{aligned} \tag{27}$$

From Equations 26, 27, the system dynamic matrices A , B , and F are replaced. Equation 27 overcomes the difficulty of obtaining the dynamic information of the system in practical applications. $(V^{[k]}, u^{[k+1]}, w^{[k+1]})$ is a unique solution for the off-policy IRL algorithm. Three neural networks were employed to solve the solution $(V^{[k]}, u^{[k+1]}, w^{[k+1]})$; the expressions are as follows:

$$\hat{V}^{[k]} = P_V^T(x) \hat{\theta}_V^{[k]} \tag{28}$$

$$u^{[k]}(x) = (P_u^a(x))^T \hat{\theta}_{u_a}^{[k]} \tag{29}$$

$$w^{[k]}(x) = (P_w^b(x))^T \hat{\theta}_{w_b}^{[k]}, \tag{30}$$

$P_V^T(x)$, $P_u^T(x)$, and $P_w^T(x)$ satisfy $P_V^T(0) = 0$, $P_u^T(0) = 0$ and $P_w^T(0) = 0$, respectively, and are linearly independent. $\hat{V}^{[k]}$ is approached by a critical neural network (CNN), and $u^{[k]}(x)$ and $w^{[k]}(x)$ are approached by an action neural network (ANN) and a disturbance neural network (DNN), respectively. Here, $\hat{\theta}_V^{[k]}$, $\hat{\theta}_{u_a}^{[k]}$, and $\hat{\theta}_{w_b}^{[k]}$ indicate the weights of CNN, ANN, and DNN, respectively.

According to Equation 28, the residual can be written as follows:

$$\begin{aligned} \epsilon^{[k]}(x, u, w) &= \hat{V}^{[k]}(x(t)) - \hat{V}^{[k]}(x(t + \Delta t)) \\ &- \int_t^{t+\Delta t} (x^T R x + u^{[k]T} S u^{[k]} - w^{[k]T} T w) d\tau \\ &+ \int_t^{t+\Delta t} -2(u^{[k+1]T} S (u - u^{[k]}) - w^{[k+1]T} T (w - w^{[k]})) \end{aligned} \tag{31}$$

Substituting Equations 28–30 into Equation 31 yields

$$\begin{aligned} \vartheta^{[k]}(x, u, w) &= [P_V(x(t)) - P_V(x(t + \Delta t))]^T \hat{\theta}_V^{[k+1]} - \int_t^{t+\Delta t} x^T R x d\tau \\ &- \sum_{a=1}^n \int_t^{t+\Delta t} \hat{\theta}_{u_a}^{[k]T} P_u^a(x(t)) S (P_u^a(x(t)))^T \hat{\theta}_{u_a}^{[k]} d\tau \\ &+ \sum_{b=1}^m \int_t^{t+\Delta t} \hat{\theta}_{w_b}^{[k]T} P_w^b(x(t)) T (P_w^b(x(t)))^T \hat{\theta}_{w_b}^{[k]} d\tau \\ &+ 2 \sum_{a=1}^n \int_t^{t+\Delta t} \hat{\theta}_{u_a}^{[k]T} P_u^a(x(t)) S (P_u^a(x(t)))^T \hat{\theta}_{u_a}^{[k+1]} d\tau \\ &- 2 \sum_{a=1}^n \int_t^{t+\Delta t} u^T S (P_u^a(x(t)))^T \hat{\theta}_{u_a}^{[k+1]} d\tau \\ &- 2 \sum_{b=1}^m \int_t^{t+\Delta t} \hat{\theta}_{w_b}^{[k]T} P_w^b(x(t)) T (P_w^b(x(t)))^T \hat{\theta}_{w_b}^{[k+1]} d\tau \\ &+ 2 \sum_{b=1}^m \int_t^{t+\Delta t} w^T T (P_w^b(x(t)))^T \hat{\theta}_{w_b}^{[k+1]} d\tau. \end{aligned} \tag{32}$$

In order to simplify Equation 32, the following parameters are defined as Equations 33–39:

$$Q_{FA}(x(t)) = P_V^T(x(t)) - P_V^T(x(t + \Delta t)) \tag{33}$$

$$Q_{FB}(x(t)) = 2 \sum_{a=1}^n \int_t^{t+\Delta t} \hat{\theta}_{u_a}^{[k]T} P_u^a(x(t)) S (P_u^a(x(t)))^T d\tau \tag{34}$$

$$Q_{FC}(x(t), u) = 2 \sum_{a=1}^n \int_t^{t+\Delta t} u^T S (P_u^a(x(t)))^T d\tau \tag{35}$$

$$Q_{FD}(x(t)) = -2 \sum_{b=1}^m \int_t^{t+\Delta t} \hat{\theta}_{w_b}^{[k]T} P_w^b(x(t)) T (P_w^b(x(t)))^T d\tau \tag{36}$$

$$Q_{FE}(x(t), w) = -2 \sum_{b=1}^m \int_t^{t+\Delta t} w^T T (P_w^b(x(t)))^T d\tau \tag{37}$$

$$Q_M(x(t)) = \int_t^{t+\Delta t} x^T R x d\tau \tag{38}$$

$$\begin{aligned} Q_N(x(t)) &= \sum_{a=1}^n \int_t^{t+\Delta t} \hat{\theta}_{u_a}^{[k]T} P_u^a(x(t)) S (P_u^a(x(t)))^T \hat{\theta}_{u_a}^{[k]} d\tau \\ &- \sum_{b=1}^m \int_t^{t+\Delta t} \hat{\theta}_{w_b}^{[k]T} P_w^b(x(t)) T (P_w^b(x(t)))^T \hat{\theta}_{w_b}^{[k]} d\tau \end{aligned} \tag{39}$$

Then, Equation 32 can be written as Equation 40:

$$\begin{aligned} \vartheta^{[k]}(x, u, w) &= Q_{FA}(x(t)) \hat{\theta}_V^{[k+1]} + Q_{FB}(x(t)) \hat{\theta}_u^{[k+1]} \\ &- Q_{FC}(x(t), u) \hat{\theta}_u^{[k+1]} + Q_{FD}(x(t)) \hat{\theta}_w^{[k+1]} \\ &- Q_{FE}(x(t), w) \hat{\theta}_w^{[k+1]} - Q_M(x(t)) - Q_N(x(t)) \end{aligned} \tag{40}$$

The Equations 41–43 are then generated to obtain the optimal solutions:

$$\hat{W}^{[k]} = \left[\hat{\theta}_V^{[k+1]}, \hat{\theta}_{u1}^{[k+1]} \dots \hat{\theta}_{ua}^{[k+1]}, \hat{\theta}_{w1}^{[k+1]} \dots \hat{\theta}_{wb}^{[k+1]} \right] \quad (41)$$

$$X_A^{[k]}(x, u, w) = \left[Q_{FA}(x(t)), Q_{FB}^l(x(t)) - Q_{FC}^l(x(t), u) \dots, Q_{FB}^n(x(t)) - Q_{FC}^n(x(t)), Q_{FD}^l(x(t)) - Q_{FE}^l(x(t), w), \dots, Q_{FD}^n(x(t)) - Q_{FE}^n(x(t), w) \right] \quad (42)$$

$$X_B^{[k]} = Q_M(x(t)) + Q_N(x(t)) \quad (43)$$

Finally, Equation 32 can be written as follows:

$$\vartheta^{[k]}(x, u, w) = X_A^{[k]}(x, u, w)\hat{W}^{[k]} - X_B^{[k]} \quad (44)$$

In order to solve weight $\hat{W}^{[k]}$, the residual error $\vartheta^{[k]}(x, u, w)$ is approximated to zero. The inner product is applied to solve $\hat{W}^{[k]}$ using Equation 45 as follows:

$$\langle d\vartheta^{[k]}(x, u, w)/d\hat{W}^{[k]}, d\vartheta^{[k]}(x, u, w) \rangle_D = 0, \quad (45)$$

upon substituting Equation 45 is substituted in Equations 44, 46 is obtained as follows:

$$\langle X_A^{[k]}, X_A^{[k]} \rangle_D \hat{W}^{[k]} - \langle X_A^{[k]}, X_B^{[k]} \rangle_D = 0 \quad (46)$$

$\hat{W}^{[k]}$ can be calculated as follows:

$$\hat{W}^{[k]} = \langle X_A^{[k]}, X_A^{[k]} \rangle_D^{-1} \langle X_A^{[k]}, X_B^{[k]} \rangle_D \quad (47)$$

Various numerical integrals in domain D were acquired to calculate $\langle X_A^{[k]}, X_A^{[k]} \rangle_D$ and $\langle X_A^{[k]}, X_B^{[k]} \rangle_D$. The Monte Carlo integration was used to resolve this calculation. When $\mu_D \triangleq \int_D d(x, u, w)$, and $Q_M \triangleq [(x_i, u_i, w_i) | (x_i, u_i, w_i) \in D, i = 1, 2, \dots, M]$, all of them are the sets sampled on domain D . M indicates the dimensions of the Q_M , which should be as wide as possible to ensure that the sample set provides full coverage of D . Therefore, $\langle X_A^{[k]}, X_A^{[k]} \rangle_D$ can be obtained as follows:

$$\begin{aligned} \langle X_A^{[k]}, X_A^{[k]} \rangle_D &= \int_D (X_A^{[k]}(x_i, u_i, w_i))^T (X_A^{[k]}(x_i, u_i, w_i)) d(x_i, u_i, w_i) \\ &= \frac{\mu_D}{M} \sum_{i=1}^M ((X_A^{[k]}(x_i, u_i, w_i))^T (X_A^{[k]}(x_i, u_i, w_i))) \\ &= \frac{\mu_D}{M} (y^{[k]})^T y^{[k]} \end{aligned} \quad (48)$$

in which, $y^{[k]} = [(X_A^{[k]T}(x_1, u_1, w_1), X_A^{[k]T}(x_2, u_2, w_2), \dots, X_A^{[k]T}(x_M, u_M, w_M))]$.

$$\begin{aligned} \langle X_A^{[k]}, X_B^{[k]} \rangle_D &= \frac{\mu_D}{M} \sum_{i=1}^M ((X_A^{[k]}(x_i, u_i, w_i))^T (X_B^{[k]}(x_i))) \\ &= \frac{\mu_D}{M} (y^{[k]})^T \beta^{[k]} \end{aligned} \quad (49)$$

where $\beta^{[k]} = [X_B^{[k]}(x_1), X_B^{[k]}(x_2), \dots, X_B^{[k]}(x_M)]^T$.

Upon substituting Equations 48, 49 in Equation 47, the following equation is obtained:

$$\hat{W}^{[k]} = \left[(y^{[k]})^T y^{[k]} \right]^{-1} (y^{[k]})^T \beta^{[k]} \quad (50)$$

The zero-sum problem can be solved using Algorithm 1 as follows:

Step 1: Start with the signals u and w as well as collecting the hydropower-photovoltaic cogeneration system data (x_p, u_p, w_p) to build the set Q_M ; then, calculate the $Q_{FA}(x), Q_{FB}(x), Q_{FC}(x, u), Q_{FD}(x), Q_{FE}(x, w), Q_M(x)$ and $Q_N(x)$.

Step 2: The values of cost function, control, and disturbance are set initial admissible weight vectors as $\hat{\theta}_V^{[0]}, \hat{\theta}_{u1}^{[0]} \dots \hat{\theta}_{ua}^{[0]}$, and $\hat{\theta}_{w1}^{[0]} \dots \hat{\theta}_{wb}^{[0]}$, separately.

Step 3: Calculate the $y^{[k]}$, and $\beta^{[k]}$ to renew the $\hat{W}^{[k]}$ by the Equation 50.

Step 4: Let $k = k + 1$, return to step 3, and go on.

Step 5: Until $\|\hat{W}^{[k+1]} - \hat{W}^{[k]}\| \leq \alpha$, where in α is a small positive constant. Then, the iteration is stopped, and $\hat{W}^{[k]}$ is used to acquire the control policy using Equation 41.

Algorithm 1 Off-policy IRL method to solve the optimal control problem.

It is worth mentioning that the input and output data of the hydropower-photovoltaic microgrid power system are necessary to solve the zero-sum problem when the system dynamics are completely unknown.

5 Discussion

The hydropower-photovoltaic microgrid power system model was established, the proposed Algorithm 1 was utilized to solve the LFC control, and the simulation was realized in the MATLAB platform. Simulation results verified that the microgrid can maintain frequency stability despite local load and PV disturbances. The control and disturbance curves eventually approach to 0, as shown in Figure 3. The Figure 3 illustrates the behavior of two variables over a period of 10 time steps, designated on the x -axis. The y -axis represents the Control Value ranging from -0.5 to 1.0 . The graph features two sets of trajectories for the control u and ω , each represented by both initial estimated values and adjusted values obtained using the Algorithm 1. The dashed and solid lines indicate the approximation curves under initial admissible control and Algorithm 1, respectively. It can be seen that the convergence speed of Algorithm 1 is better than that of the initial admissible control method. The frequency finally stabilized. For variable ω , it starts from a lower value and similarly converge towards zero. Overall, the obtained trajectories using Algorithm 1 exhibit a more rapid convergence towards 0 for both u and ω compared to their respective initial trajectories. Demonstrating the enhanced performance of Algorithm 1 over the initial admissible control method.

The weight convergence curves of the three networks are shown in Figures 4–6. These three figures illustrates the convergence of weights for every seven different networks $\theta_v, \theta_u,$ and θ_ω over 15 iterations, represented on the x -axis. The y -axis denotes the Weight of Value Function, with values ranging from -10 to 15 . At the beginning of the iterations, the weights of these networks start from various initial values. Some weights exhibit significant fluctuations, particularly up to around the 5th iteration, where noticeable oscillations are apparent. As iterations proceed, the weights for all networks gradually stabilize and converge to constant values between 0 and 1, indicating that the values no

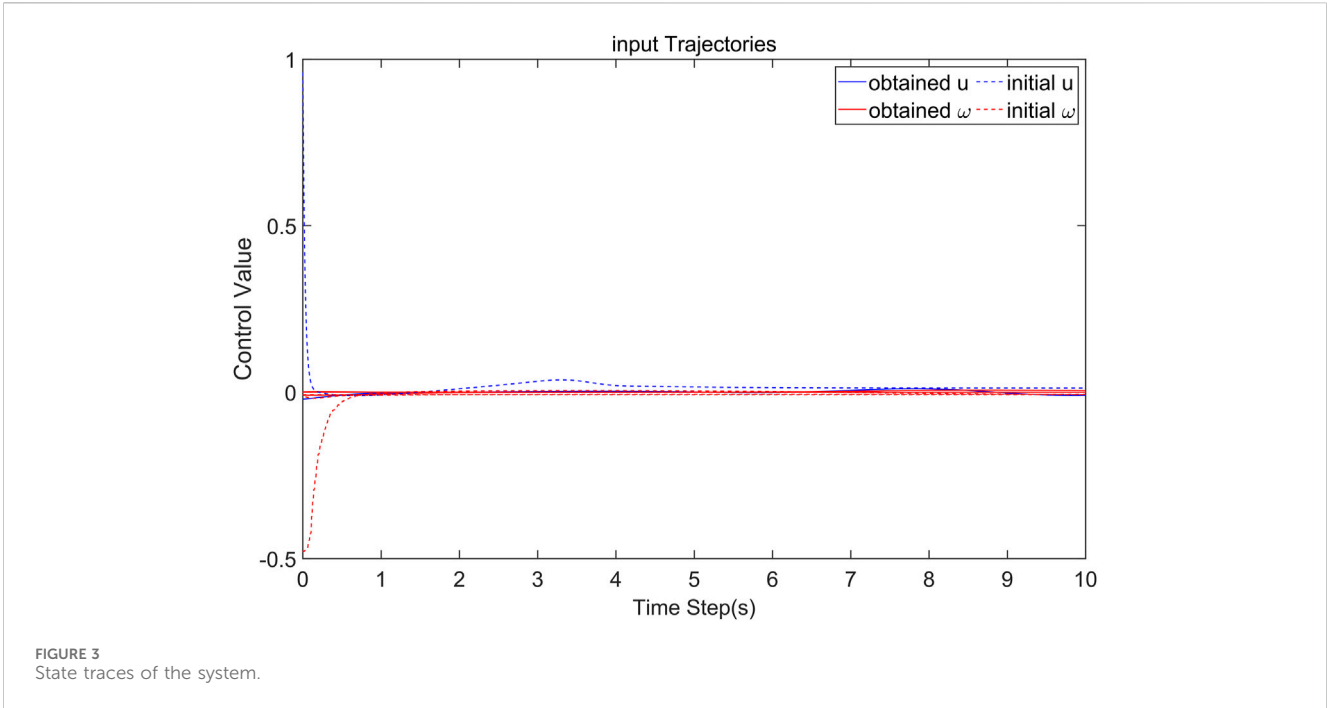


FIGURE 3 State traces of the system.

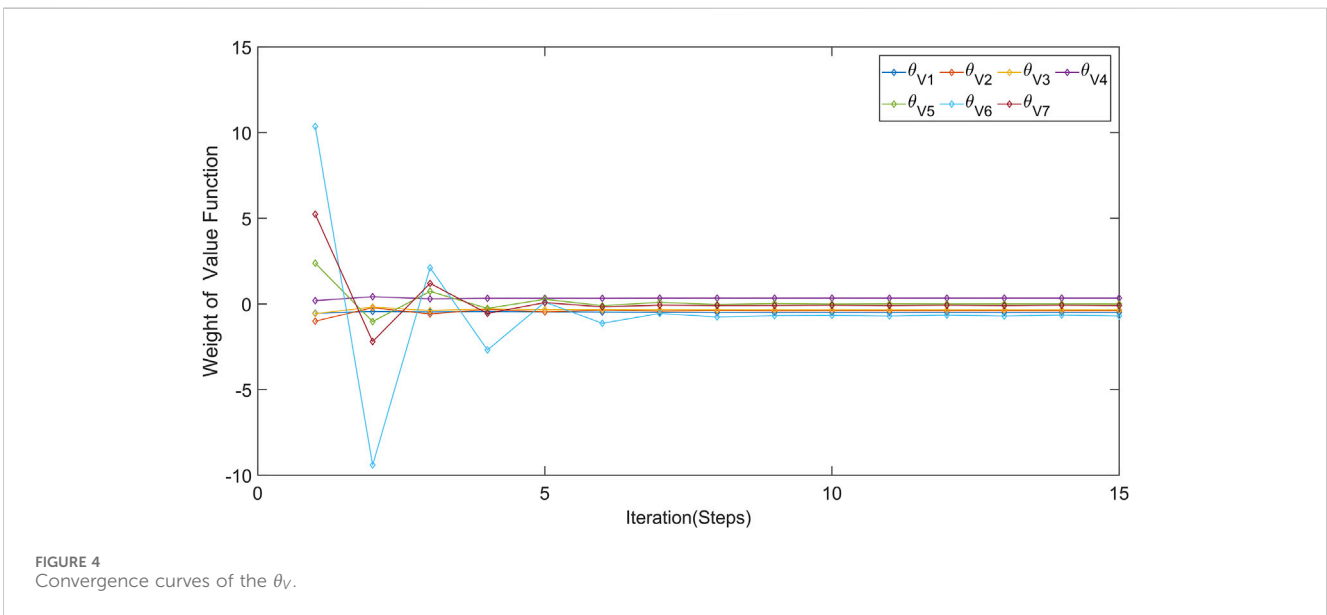


FIGURE 4 Convergence curves of the θ_v .

longer change significantly with additional iterations. By the 15th step, all weights have reached their steady-state values. The convergence of these weights within 15 iterations suggests that the optimization process is effective, allowing the system to arrive at an optimal control that approximates the Nash equilibrium point for the zero-sum game problem.

Compared to traditional Dynamic Programming methods, the proposed method effectively overcomes the “curse of dimensionality,” significantly reducing the computational burden when solving high-dimensional matrices. In contrast to previous reinforcement learning approaches for controlling the optimal

frequency of hydropower-photovoltaic microgrid power systems, this method incorporates the consideration of disturbance factors, providing a robust theoretical basis for the grid integration of hybrid power generation systems.

The desired voltage and current is 50 hz sinusoidal waves, such that the system is dynamic with high frequency. Yet the IRL method depends on a process to collect the control and states data from the system under a quasi-optimal control, which may lead to power oscillation and need more time to turn the system from the transient state to steady state. Therefore, the limitation of this method is that it is currently only applicable to offline systems.

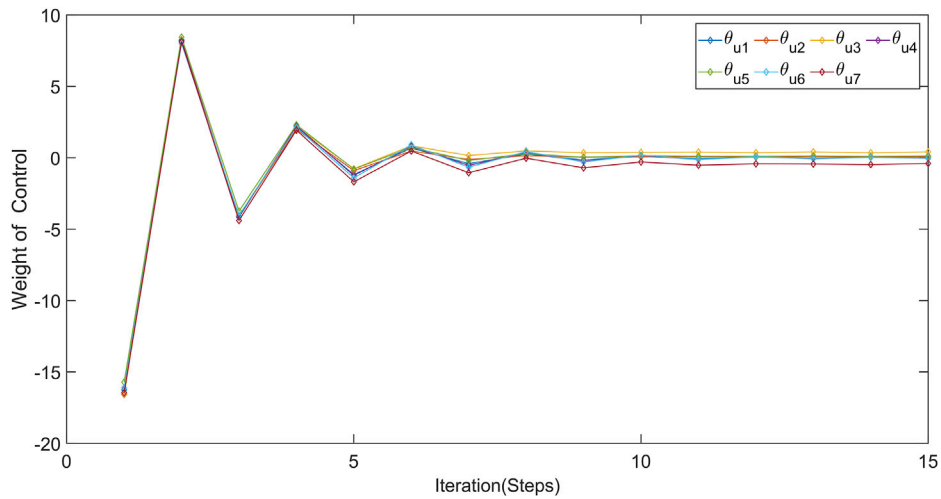


FIGURE 5 Convergence curves of the θ_u .

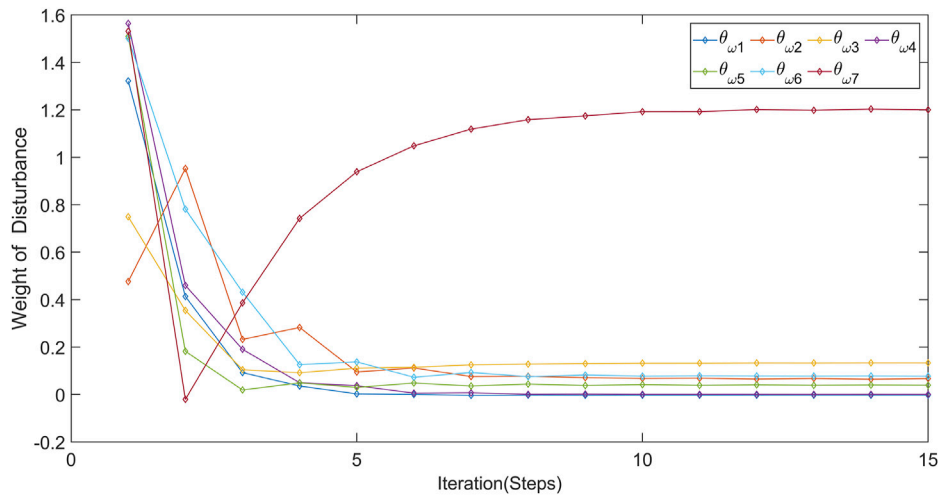


FIGURE 6 Convergence curves of the θ_w .

6 Conclusion

This paper focused on the hydropower-photovoltaic hybrid microgrid system and designed an optimal LFC using the IRL algorithm. First, the mechanism models of the hydro turbine generator and the photovoltaic generator were established, respectively. Second, a state-space model of the hydropower-photovoltaic hybrid microgrid system was developed, and based on the power generation characteristics, it was transformed in solving a zero-sum game problem. Third, the IRL algorithm was employed to approximate the Nash equilibrium point of the zero-sum game problem using three neural networks. Finally, the simulation experiments were conducted to verify the effectiveness of the proposed method.

Data availability statement

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

Author contributions

EW: Conceptualization, Writing–original draft, Writing–review and editing. LY: Methodology, Writing–review and editing. FZ: Writing–review and editing. XL: Writing–review and editing. JL: Writing–original draft. LS: Writing–original draft. MZ: Writing–original draft.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by a grant for the project “Research and engineering demonstration of a safe, autonomous, and controllable intelligent control system for 10 million kilowatts of clean energy” (CSIEKJ220700539).

Conflict of interest

Authors EW and LY were employed by Guoneng Qinghai Yellow River MaerDang Hydropower Development Co., Ltd.

Authors FZ, XL, and JL were employed by Beijing SP Zhishen Control Technology Co., Ltd.

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

Publisher's note

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

References

- Ahmad, S., Ahmad, A., and Yaqub, R. (2018). Optimized energy consumption and demand side management in smart grid. *Smart Grid as a Solut. Renew. Effic. Energy*, 1–25. doi:10.4018/978-1-5225-0072-8.ch001
- Bellman, R. E., and Dreyfus, S. E. (2015). *Applied dynamic programming*. Princeton, New Jersey: Princeton University Press.
- Chai, R., Savvaris, A., Tsourdos, A., and Chai, S. (2017). Multi-objective trajectory optimization of space manoeuvre vehicle using adaptive differential evolution and modified game theory. *Acta Astronaut.* 136, 273–280. doi:10.1016/j.actaastro.2017.02.023
- Chen, Z., Chen, J., Fu, K., and Xue, L. (2022). Power coordination control strategy of microgrid based on photovoltaic generation. *MATEC Web Conf.* 355, 03065. doi:10.1051/mateconf/202235503065
- Coban, H. H., Rehman, A., and Mousa, M. (2022). Load frequency control of microgrid system by battery and pumped-hydro energy storage. *Water* 14, 1818. doi:10.3390/w14111818
- Cowie, P., Townsend, L., and Salemin, K. (2020). Smart rural futures: will rural areas be left behind in the 4th Industrial Revolution?. *J. Rural. Stud.* 79, 169–176. doi:10.1016/j.jrurstud.2020.08.042
- Dhundhara, S., and Verma, Y. P. (2020). Application of micro pump hydro energy storage for reliable operation of microgrid system. *IET Renew. Power Gener.* 14, 1368–1378. doi:10.1049/iet-rpg.2019.0822
- Fang, H., Chen, L., and Shen, Z. (2011). Application of an improved PSO algorithm to optimal tuning of PID gains for water turbine governor. *Energy Convers. Manag.* 52, 1763–1770. doi:10.1016/j.enconman.2010.11.005
- Ganguly, S., Shiva, C. K., and Mukherjee, V. (2018). Frequency stabilization of isolated and grid connected hybrid power system models. *J. Energy Storage*. 19, 145–159. doi:10.1016/j.est.2018.07.014
- Gielen, D., Boshell, F., Saygin, D., Bazilian, M. D., Wagner, N., and Gorini, R. (2019). The role of renewable energy in the global energy transformation. *Energy Strategy Rev.* 24, 38–50. doi:10.1016/j.esr.2019.01.006
- Gilani, M. A., Kazemi, A., and Ghasemi, M. (2020). Distribution system resilience enhancement by microgrid formation considering distributed energy resources. *Energy* 191, 116442. doi:10.1016/j.energy.2019.116442
- Guha, D., Roy, P. K., and Banerjee, S. (2021). Equilibrium optimizer-tuned cascade fractional-order 3DOF-PID controller in load frequency control of power system having renewable energy resource integrated. *Int. Trans. Electr. Energy Syst.* 31, e12702. doi:10.1002/2050-7038.12702
- Huang, Z., Liu, X., Fu, H., and Du, Z. (2021). A novel parameter optimisation method of hydraulic turbine regulating system based on fuzzy differential evolution algorithm and fuzzy PID controller. *Int. J. Bio Inspired Comput.* 18, 153–164. doi:10.1504/IJBIC.2021.119203
- Khalid, H. M., Muyeen, S. M., and Kamwa, I. (2022). An improved decentralized finite-time approach for excitation control of multi-area power systems. *Sustain. Energy, Grids Netw.* 31, 100692. doi:10.1016/j.segan.2022.100692
- Lewis, F. L., Vrabie, D., and Syrmos, V. L. (2012). *Optimal control*. Hoboken, New Jersey: John Wiley & Sons.
- Li, J., Guo, W., and Liu, Y. (2023). Nonlinear state feedback-synergetic control for low frequency oscillation suppression in grid-connected pumped storage-wind power interconnection system. *J. Energy Storage*. 73, 109281. doi:10.1016/j.est.2023.109281
- Ma, T., Yang, H., Lu, L., and Peng, J. (2014). Technical feasibility study on a standalone hybrid solar-wind system with pumped hydro storage for a remote island in Hong Kong. *Renew. Energy*. 69, 7–15. doi:10.1016/j.renene.2014.03.028
- Mohamed, R., Helaimi, M., Taleb, R., Gabbar, H. A., and Othman, A. M. (2020). Frequency control of microgrid system based renewable generation using fractional PID controller. *IJECS* 19, 745–755. doi:10.11591/ijeecs.v19.i2.pp745-755
- Nisha, G., and Jamuna, K. (2022). Frequency stabilization of stand-alone microgrid with tuned PID controller. *ECS Trans.* 107, 773–782. doi:10.1149/10701.0773ecst
- Olabi, A. G., and Abdelkareem, M. A. (2022). Renewable energy and climate change. *Renew. Sustain. Energy Rev.* 158, 112111. doi:10.1016/j.rser.2022.112111
- Osman, N., Khalid, H. M., Tha'er, O. S., Abuashour, M. I., and Muyeen, S. M. (2022). A PV powered DC shunt motor: study of dynamic analysis using maximum power Point-Based fuzzy logic controller. *Energy Convers. Manag.* X 15, 100253. doi:10.1016/j.ecmx.2022.100253
- Papaefthymiou, S. V., Karamanou, E. G., Papanthassiou, S. A., and Papadopoulos, M. P. (2010). A wind-hydro-pumped storage station leading to high RES penetration in the autonomous island system of Ikaria. *IEEE Trans. Sustain. Energy*. 1, 163–172. doi:10.1109/TSTE.2010.2059053
- Patnaik, B., Mishra, M., Bansal, R. C., and Jena, R. K. (2020). AC microgrid protection-A review: current and future prospective. *Appl. Energy*. 271, 115210. doi:10.1016/j.apenergy.2020.115210
- Ray, P. K., Mohanty, S. R., and Kishor, N. (2011). Proportional-integral controller based small-signal analysis of hybrid distributed generation systems. *Energy Convers. Manag.* 52, 1943–1954. doi:10.1016/j.enconman.2010.11.011
- Rehman, A. U., Ullah, Z., Qazi, H. S., Hasanien, H. M., and Khalid, H. M. (2024). Reinforcement learning-driven proximal policy optimization-based voltage control for PV and WT integrated power system. *Renew. Energy* 227, 120590. doi:10.1016/j.renene.2024.120590
- Shuai, H., Ai, X., Fang, J., Ding, T., Chen, Z., and Wen, J. (2020). Real-time optimization of the integrated gas and power systems using hybrid approximate dynamic programming. *Int. J. Electr. Power Energy Syst.* 118, 105776. doi:10.1016/j.ijepes.2019.105776
- Song, R., Wei, Q., and Li, Q. (2019). Off-policy integral reinforcement learning method for multi-player non-zero-sum games. *Adapt. Dyn. Program. Single Multiple Control.* 227–249. doi:10.1007/978-981-13-1712-5_12
- Thirunavukkarasu, M., and Sawle, Y. (2021). Smart microgrid integration and optimization. *Act. Electr. Distrib. Netw. Smart Approach*, 201–235. doi:10.1002/9781119599593.ch11
- Tran, A. T., Minh, B. L. N., Huynh, V. V., Tran, P. T., Amaefule, E. N., Phan, V. D., et al. (2021). Load frequency regulator in interconnected power system using second-order sliding mode control combined with state estimator. *Energies* 14, 863. doi:10.3390/en14040863
- Vamvoudakis, K. G., and Lewis, F. L. (2010). Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem. *Automatica* 46, 878–888. doi:10.1016/j.automatica.2010.02.018
- Wang, R., Sun, Q., Pinjia, Z., Yonghao, G., Dehao, Q., and Peng, W. (2020). Reduced-order transfer function model of the droop-controlled inverter via Jordan continued-fraction expansion. *IEEE Trans. Energy* 35, 1585–1595. doi:10.1109/TEC.2020.2980033

- Werbos, P. (1992). Approximate dynamic programming for real-time control and neural modeling. *Handb. Intelligent Control*.
- Wu, J., and Yang, F. (2023). A dual-driven predictive control for photovoltaic-diesel microgrid secondary frequency regulation. *Appl. Energy*. 334, 120652. doi:10.1016/j.apenergy.2023.120652
- Xue, X., Ai, X., Fang, J., Yao, W., and Wen, J. (2022). Real-time schedule of integrated heat and power system: a multi-dimensional stochastic approximate dynamic programming approach. *Int. J. Electr. Power Energy Syst.* 134, 107427. doi:10.1016/j.ijepes.2021.107427
- Zeng, Y., Zhang, L. X., Guo, Y. K., and Qian, J. (2015). Hamiltonian stabilization additional L₂ adaptive control and its application to hydro turbine generating sets. *Int. J. Control Autom. Syst.* 13, 867–876. doi:10.1007/s12555-013-0460-7
- Zepter, J. M., Lüth, A., Crespo del Granado, P. C., and Egging, R. (2019). Prosumer integration in wholesale electricity markets: synergies of peer-to-peer trade and residential storage. *Energy Build.* 184, 163–176. doi:10.1016/j.enbuild.2018.12.003
- Zhang, D., and Kong, Q. (2022). Green energy transition and sustainable development of energy firms: an assessment of renewable energy policy. *Energy Econ.* 111, 106060. doi:10.1016/j.eneco.2022.106060