



# Automatic generation control of is-landed micro-grid using integral reinforcement learning-based adaptive optimal control strategy

Rasananda Muduli<sup>1</sup> · Debashisha Jena<sup>1</sup> · Tukaram Moger<sup>1</sup>

Received: 24 April 2024 / Accepted: 29 July 2024

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2024

## Abstract

Microgrids serve an essential role in the smart grid infrastructure, facilitating the seamless integration of distributed energy resources and supporting the increased adoption of renewable energy sources to satisfy the growing demand for sustainable energy solutions. This paper presents an application of integral reinforcement learning (IRL) algorithm-based adaptive optimal control strategy for automatic generation control of an is-landed micro-grid. This algorithm is a model-free actor-critic method that learns the critic parameters using the recursive least square method. The actor is straightforward and evaluates the action from the critic directly. The robustness of the proposed control technique is investigated under various uncertainties arising from parameter uncertainty, electric vehicle (EV) aggregator, and renewable energy sources. This study incorporates case studies and comparative analyses to demonstrate the control performance of the proposed control strategy. The effectiveness of the technique is evaluated by comparing it with deep Q-learning (DQN) control techniques and PI controllers. The proposed controller significantly improves performance metrics compared to the DQN and PI controllers. It reduces the peak frequency deviation by 6% and 14%, respectively, compared to the DQN and PI controllers. When subjected to multiple-step load disturbances, the proposed controller reduces the mean square error by 28% and 42%, respectively, while lowering both the integral absolute error and the integral time absolute error by 21% and 35% compared to the DQN and PI controllers. Additionally, when operating with renewable energy sources, the proposed controller decreases the standard deviation in the frequency deviation by 17% compared to the DQN controller and 23% compared to the PI controller.

---

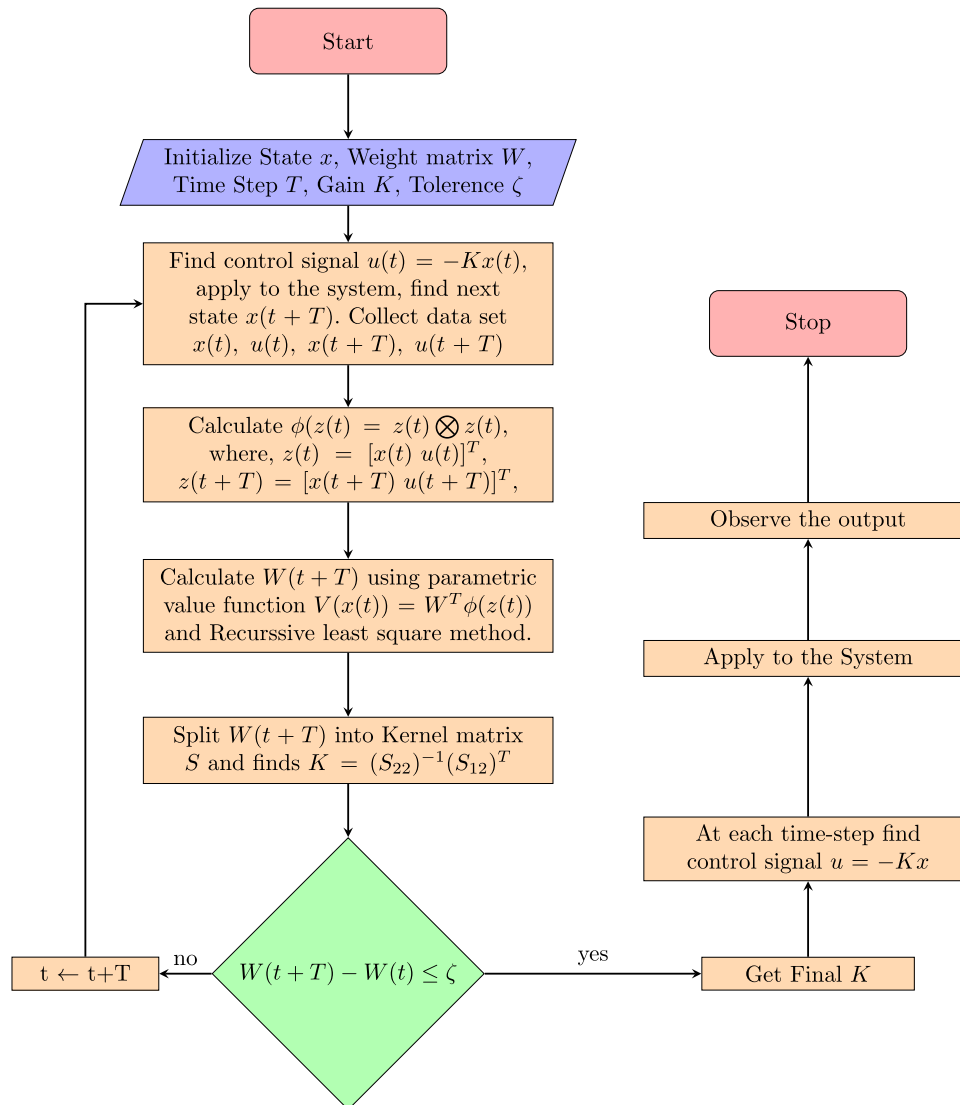
✉ Rasananda Muduli  
erasa.knp@gmail.com

Debashisha Jena  
djena.2490ee@nitk.edu.in

Tukaram Moger  
tukaram@nitk.edu.in

<sup>1</sup> Department of Electrical and Electronics Engineering,  
National Institute of Technology, Karnataka, Surathkal  
575025, India

## Graphical abstract



**Keywords** Automatic generation control · Integral reinforcement learning · Optimal control

## List of symbols

$J(x, u)$	Quadratic cost function	$\phi(z)$	Quadratic polynomial basis set
$Q(x, u)$	Q-value function	$\otimes$	Kronecker product
$A, B$	System matrices	$T_g$	Time constant of governor
$\mathbf{Q}, \mathbf{R}$	Time invariant weights matrices	$T_t$	Time constant of turbine
$x$	System state	$T_{\text{BESS}_i}$	Time constant of $i$ th BESS
$u$	Controller output	$T_P$	Time constant of power system
$V(x)$	Value function	$T_{PV}$	Time constant of PV
$Q, Q^-$	Current, target network	$T_{WT}$	Time constant of Wind
$\theta, \theta^-$	Parameters of current, target network	$T_{\text{FESS}}$	Time constant of FESS
$L(\theta), R_1$	Loss function, reward for DQN	$T_e$	Time constant of EV aggregator
$\gamma, \alpha_1$	Discount factor, learning rate	$K_P$	Gain of power system
$W$	Weight matrix	$\alpha$	Participation factor of EV aggregator
$z$	$[x \ u]^T$	$\alpha_i$	Participation factor of $i$ th BESS

$\zeta$	Tolerance
$K_I$	Integral gain
$T$	Time step
$K_{ei}$	Gain of the $i$ th EV
$N$	Total number of EVs
$N_1$	No. of EVs in SOC idle mode
$N_2$	No. of EVs in SOC Controllable mode
$\eta_2$	$N_2/N$
$\eta$	Maximum of $\eta_2$
$\Delta X_g$	Deviation in governor position
$\Delta P_g$	Deviation in generator power
$\Delta f$	Deviation in frequency
$\Delta I$	Incremental change in integral control
$\Delta P_d$	Change in load disturbance
$R$	Speed regulation of the governor
$K_e$	Gain of EV aggregator
SOC	State of charge
SOC <sub>low (high),i</sub>	Low (High) battery SOC ith EV
SOC <sub>max(min),i</sub>	Max (Min) battery SOC ith EV
$c$	Delayed time-step in DQN pre-learning
$\nabla_{\theta_t} L(\theta_t)$	Gradient of loss function
$\Delta P_{\text{Solar}}$	Solar power change
$\Delta P_{\text{Wind}}$	Wind power change
$\Delta P_{\text{PV}}$	Output solar power change
$\Delta P_{\text{W}}$	Output wind power change
$\Delta P_{\text{BESS}}$	Output battery power change
$\Delta P_{\text{FESS}}$	Output flywheel power change
$\Delta P_e$	EV aggregator power change
$\Gamma$	Disturbance matrix
HJB	Hamilton–Jacobi–Bellman
ARE	Algebraic Recati equation
TD	Temporal difference
DQN	Deep Q-network
DDQN	Double deep Q-network
DDPG	Deep deterministic policy gradient
SAC	Soft actor-critic
IRL	Integral reinforcement learning
EV	Electric vehicle
RLS	Recursive least square
PV	Photovoltaic
BESS	Battery energy storage system
FESS	Flywheel energy storage system
PI	Proportional integral controller
PID	Proportional integral derivative
MSE	Mean squared error
IAE	Integral absolute error
ITAE	Integral time absolute error
AGC	Automatic generation control
DG	Distributed generation
LFC	Load frequency control
RL	Reinforcement learning
ADP	Adaptive dynamic programming

## 1 Introduction

Renewable energy is a long-term and potentially sustainable strategy for reducing green house gas emissions significantly. Due to environmental concerns, economic expansion, and the energy crisis problem, incorporating renewable energy sources into the power system is now getting accelerated. The transition from centralized to distributed generation (DG) has enhanced the desirability and suitability of micro-grids for integrating renewable energy sources [1]. Micro-grid is a small-scale integrated energy system with multifarious distribution configuration consisting of several interconnected distributed energy resources and loads situated within a specific local area and operating within a well-defined electrical boundary [2].

Microgrids may be viewed as a more developed form of distributed generation system with minimal influence from the stochastically distributed generations through efficient management of the storage systems and dispatched loads [3]. Generally, a microgrid operates in grid-connected mode or autonomous mode. However, due to their lower equivalent inertia compared to the main grid, the management and controls of autonomous microgrids are often more complex than those of grid-connected ones. Consequently, even moderate or minor disruptions in the power supply can cause issues in power quality and stability, specifically, a deterioration in the quality of the voltage and frequency [4]. Incorporating energy storage systems into microgrids can maintain the instantaneous power balance and enhance the microgrid's dynamic performance, which is crucial in is-land operation mode. The integration of multiple energy storage devices, such as batteries and flywheel, has the potential to enhance the stability of microgrids. However, this integration also introduces increased complexity of control issues.

Frequency regulation is a significant challenge in autonomous microgrids due to the low inertia of the system and the large share of intermittent renewable sources [5]. Renewable energy sources and energy storage systems must co-ordinate in autonomous microgrids to limit the frequency deviations by minimizing the mismatch between generation and demand [6–8]. When there is an imbalance between generated power and demand power, different control strategies are needed to maintain the frequency stability of the system [9]. To stabilize the system, primary droop controllers are widely employed. Still, frequency deviations from the nominal value are observed in the steady state. Consequently, a secondary control strategy is employed to reach optimal frequency synchronization. Generally, the primary droop and secondary frequency control are called automatic generation control (AGC). In generally, it is also known as load frequency control (LFC). This paper mainly focuses on the automatic generation control of micro-grid supported by renewable energy and storage systems.

Demand response, grid integration of electric vehicles, smart homes, presence of noise on the sensor, etc. increase the unpredictability in the power system operation [10]. Huge integration of renewable energy sources increases the uncertainty of the system due to their intermittent nature. During the parameter measurement of the microgrid, the value of internal parameters of the system may deviate in a small range from their nominal values. The presence of parameter uncertainties and load disruptions introduces unpredictability and uncertainty into the power system [11]. Due to the aforementioned reasons, the dynamics of the power system varies with time, and an accurate simulation model cannot always be found. Conventional controllers such as Proportional-Integral-Derivative (PID), Linear Matrix Inequality (LMI), Sliding Mode Control (SMC), and Model Predictive Control (MPC) are developed using model-based techniques to obtain the optimal control performance. These model-based control techniques guarantee optimal the performance when applied to a precise system model with no uncertainty. However, the interest of this paper is to design of controller for unknown system dynamics.

Reinforcement learning (RL) and adaptive dynamic programming (ADP) can solve the optimal control problem from data without having the knowledge of the system dynamics [12–14]. RL and ADP fill the gap between conventional optimal control and adaptive control techniques. The optimal solution is obtained by learning online the solution to the Hamilton–Jacobi–Bellman (HJB) equation [15]. The HJB equation reduces to an algebraic Riccati equation (ARE) for linear systems. RL adopts the temporal difference (TD) approach. TD method can learn directly from the raw experience without having a model of the environment dynamics. TD method implements mainly two steps [15, 16]: (1) to solve the temporal difference equation (known as policy evaluation), (2) to find the optimal control strategy (known as policy improvement). These steps are analogous in solving Hamilton–Jacobi–Bellman (HJB) equation for the optimal control problem. Policy and value iteration methods determine the sequence in which the TD equations can be solved, and the corresponding optimal policies can be obtained.

In recent years, extensive research on RL-based AGCs has been conducted, including DQN-based AGC [17], DDQN based AGC [18], DDPG-based AGC [19, 20], SAC-based AGC [21], and others. However, these controllers often require multiple neural networks for designing the control strategy. Each algorithm presents its unique advantages; however, increased use of neural networks increases the complexity of the design. Considering the challenges associated with design complexity and popularity of PID controllers for AGC, authors of [22] propose RL-based adaptive PID for AGC. The above discussion reveals that adaptive control strategies to reduce design complexity have room for AGC.

In [23, 24], the authors present an integral reinforcement learning (IRL) algorithm-based optimum control method for automatic generation control in power systems. Papers [23] present two different implementations of online IRL controllers, each employing separate neural networks for actor and critic networks, it lacks discussion on the neural network configurations, raising concerns about the complexity of the controller design. Conversely, the recursive least square-based IRL [25] approach offers a solution to the optimal control problem without relying on neural networks. The authors in [23] utilized the gradient method to update neural network parameters by computing the cost function gradient for each parameter. However, using the recursive least square method, the parameters of the IRL-based optimal controller can be evaluated directly, eliminating the need for a gradient descent approach [25]. This approach effectively eradicates suboptimal control policies arising from local optima.

The integration of renewable energy sources poses a significant challenge to traditional power system control methods, necessitating innovative approaches to ensure efficient and reliable operation. While the integral reinforcement learning (IRL) algorithm-based optimum control method proposed by [23] and [24] shows promise in addressing uncertainties and unknown dynamics in power systems, its efficacy in the context of renewable energy sources remains largely unexplored. Furthermore, IRL-based AGC needs to investigate the impact of uncertainty and unpredictability resulting from parameter uncertainty and external disturbances such as renewable integration and electric vehicle (EV) integration. Moreover, it is crucial to emphasize that the application of IRL controller for AGC of microgrids is still limited, and further research is needed in this area. This is because the IRL-based optimal controller offers notable advantages: (1) similar to value-based model-free RL algorithms, it does not rely on the discretization of action and state spaces, and (2) akin to deep RL algorithms, it avoids the necessity of training multiple neural networks. The proposed work can be viewed as an energy management problem. A comparison table of the proposed method with traditional methods is provided in Table 1.

The main contributions of the paper are as follows:

- This paper investigates the automatic generation control of an isolated microgrid. The investigation focuses on the impact of an EV aggregator, parameter uncertainty, and the uncertainty associated with renewable energy sources on the frequency stability of the system.
- This paper introduces an adaptive optimal control technique for secondary frequency control using the integral reinforcement learning algorithm. Like other model-free RL algorithms, it can learn the system without prior knowledge of the system dynamics. During pre-learning,

**Table 1** Comparison of recent research on energy management

References	Source	Energy storage	Methodology	Adaptive
[26]	Wind turbine, CHP-based bio-waste	Hydrogen, Compressed air, Thermal storage	Bi-level optimization, Unscented transformation	No
[27]	Wind turbine, PVs, CHP-based bio-waste	Electrical energy storage, Thermal energy storage	Unscented transformation	No
[28]	Wind turbine, EVs, CHP-based bio-waste	Compressed air storage, Thermal energy storage	Ant lion optimization	No
[29]	Wind turbine, PVs, Bio-waste	Hydrogen storage	Stochastic, Optimization	No
[30]	Wind, Demand response	Energy storage system	Uncertainty-based robust, Optimization	No
[31]	Micro-grids	EV parking, Energy storage system	Fuzzy hybrid, Stochastic robust programming	No
[32]	Renewable energy sources, Demand response	EV parking	Stochastic programming	No
[33]	PV power, Fuel cell	Battery water tank, Fuel tank	Switched-model Predictive control	No
[34]	PV power, Utility grid, Diesel generator	Battery	Extended optimal, $\epsilon$ -Variable method	No
Proposed	PV wind power, EV-aggregator, Diesel generator	Battery energy storage, Flywheel energy Storage	Integral reinforcement Learning algorithm	Yes

the recursive least square (RLS) method is implemented to estimate the controller parameters.

- For comparison purposes, the effectiveness of the IRL controller is compared with the performance of deep Q-learning (DQN) techniques and PI controller.

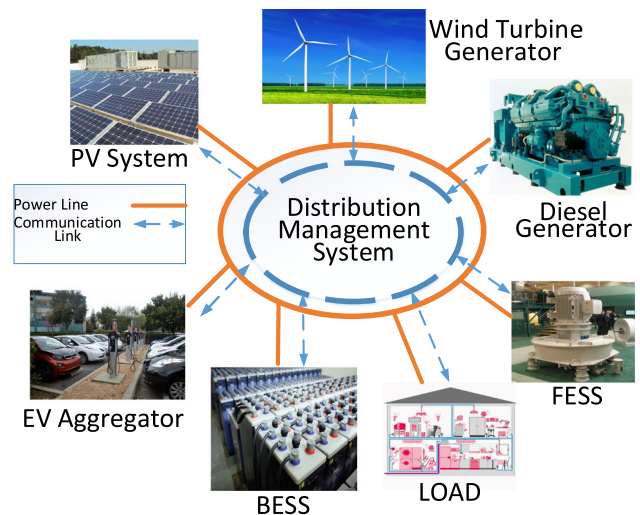
The rest parts of the paper are organized as follows. Section 2 describes the dynamic model of the system components. In Sect. 3, the mathematical formulation of the proposed control strategy is provided, followed by Sect. 4. Section 4 presents the simulation results. Finally, Sect. 5 concludes the paper.

## 2 Dynamic model of autonomous microgrid understudy

The proposed microgrid is comprised of various system components, including photovoltaic (PV) and wind power generation, a diesel engine system, an electric vehicle (EV) aggregator, and energy storage systems such as flywheel energy storage systems (FESS) and battery energy storage systems (BESS). This section presents the dynamic model of these components and the microgrid’s state-space representation. Figure 1 illustrates the schematic diagram of the system.

### 2.1 Diesel generator system

Diesel generators can provide reliable and stable power to remote areas. The dynamic model of the diesel generator can



**Fig. 1** General schematic of proposed micro-grid

be expressed as follows [35]:

$$\begin{aligned}
 \Delta \dot{X}_g(t) &= -\frac{1}{T_g} \left( \Delta X_g(t) + \frac{1}{R} \Delta f(t) - \Delta I(t) - u(t) \right) \\
 \Delta \dot{P}_g(t) &= -\frac{1}{T_t} (\Delta P_g(t) - \Delta X_g(t)) \\
 \Delta \dot{f}(t) &= -\frac{1}{T_p} (\Delta f(t) - K_p \Delta P_g(t) + K_p \Delta P_d(t)) \\
 \Delta \dot{I}(t) &= K_I \Delta f(t)
 \end{aligned} \tag{1}$$

where  $\Delta X_g(t)$ ,  $\Delta P_g(t)$ ,  $\Delta f(t)$ ,  $\Delta I(t)$  and  $\Delta P_d$  denote deviation in governor position, deviation in generator power, deviation in frequency, incremental change in integral control and change in load disturbance, respectively.

### 2.2 Renewable energy sources

The renewable energies integrated to the micro-grid are PV and wind power. The first-order dynamic equations of PV and wind are given in Eqs. 2 and 3, respectively [40–42].

$$\Delta \dot{P}_{PV}(t) = -\frac{1}{T_{PV}} \Delta P_{PV}(t) + \frac{1}{T_{PV}} \Delta P_{Solar}(t) \quad (2)$$

$$\Delta \dot{P}_W(t) = -\frac{1}{T_{WT}} \Delta P_W(t) + \frac{1}{T_{WT}} \Delta P_{Wind}(t) \quad (3)$$

### 2.3 Storage systems

In the proposed micro-grid energy storage systems, battery energy storage system and flywheel energy storage systems (BESS and FESS) are used. The capacity of the BESS should reach a certain scale to significantly influence the system frequency [36]. However, the capital cost of a large BESS is relatively high. To get a large-scale BESS, aggregated multiple small-scale BESS is implemented in [37] for LFC. In this study, multiple small-scale BESS are aggregated for primary frequency control of the micro-grid. The first-order dynamic equations of  $i$ th BESS and FESS are given in Eqs. 4, and 5, respectively [38, 39].  $\alpha_i$  represents the participation factor of each BESS and  $i = 1, 2, 3$ .

$$\Delta \dot{P}_{BESSi}(t) = -\frac{1}{T_{BESSi}} \Delta P_{BESSi}(t) + \frac{\alpha_i}{T_{BESSi}} \Delta f(t) \quad (4)$$

$$\Delta \dot{P}_{FESS}(t) = -\frac{1}{T_{FESS}} \Delta P_{FESS}(t) + \frac{1}{T_{FESS}} \Delta f(t) \quad (5)$$

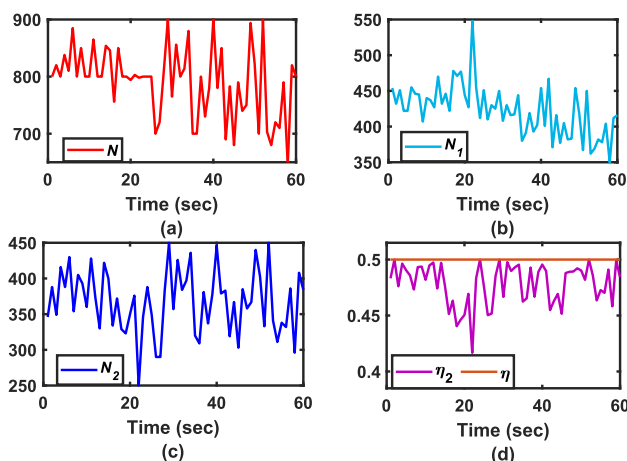
### 2.4 EV Aggregator

Assuming that all EVs participating in the EV aggregator have same time constant ( $T_e$ ), the first-order dynamic model of the EV aggregator is expressed in Eq. 6 [43].

$$\Delta \dot{P}_e(t) = -\frac{1}{T_e} \Delta P_e(t) + \frac{K_e}{T_e} \alpha P_{ce}(t) \quad (6)$$

where  $K_e = \sum_{i=1}^N K_{ei}/N$ ,  $i = 1, \dots, N$ .  $K_{ei}$  is the gain of the  $i$ th EV.  $\alpha$  is the the participation factor of EV aggregator.  $N$  is the total number of EVs. The gain ( $K_e$ ) depends on the  $SOC$  of battery. An EV can take part in LFC with  $SOC$  controllable mode or  $SOC$  idle mode. In  $SOC$  idle mode, an EV can absorb or discharge power without considering its battery  $SOC$ . The gain ( $K_{ei}$ ) of the  $i$ th EV is considered as  $\bar{K}_e = 1$  in  $SOC$  idle mode. When  $SOC$  is considered, the gain of  $i$ th EV can be modified as  $K_{ei} = \bar{K}_e - \bar{K}_e g_i(t)$ , where

$$g_i(t) = \left( \frac{SOC_i - SOC_{low (high),i}}{SOC_{max(min),i} - SOC_{low (high),i}} \right)^2 \quad (7)$$



**Fig. 2** EV aggregator parameter **a** total number of EVs, **b** number of EVs in Idle mode, **c** number of EVs in SOC controllable mode, and, **d**  $\eta_2$  and its maximum value

When the  $SOC_i \geq SOC_{high,i}$ , the  $i$ th EV will only deliver power with  $K_{ei} = \bar{K}_e$ . When the  $SOC_i \leq SOC_{low,i}$ , the  $i$ th EV will only absorb power with  $K_{ei} = 0$ . For other cases,  $K_{ei}$  is calculated using Eq. 7.

Assume that among  $N$  EVs,  $N_1$  and  $N_2 (= N - N_1)$  are participating in LFC with  $SOC$  idle mode and  $SOC$  controllable mode, respectively. Then the gain of the aggregated EVs can be written as

$$\begin{aligned} K_e &= \sum_{i=1}^N K_{ei}/N \\ &= \frac{N_1}{N} \bar{K}_e + \frac{1}{N} \bar{K}_e \left( N_2 - \sum_{i=N_1+1}^N g_i(t) \right) \\ &= \bar{K}_e + \Delta K_e \\ &= \bar{K}_e - \eta_2 g_0(t) \bar{K}_e \\ &= \bar{K}_e (1 - \eta g(t)) \end{aligned} \quad (8)$$

where  $\eta_2$  is  $N_2/N$ ,  $\eta = \max(\eta_2)$ ,  $g_0(t) = \sum_{i=N_1+1}^N g_i(t)/N_2$  and  $g(t) = (\eta_2/\eta)g_0(t)$ .  $N$ ,  $N_1$  and  $N_2$  are varying with time. Figure 2a–c represents the number of plug-in EVs participating in LFC. The maximum and minimum numbers for the total number of available EVs ( $N$ ) have been set at 650 and 900, respectively. The value of  $N$ , depicted in Fig. 2a, is determined randomly between the minimum and maximum values. Figure 2b presents the number of vehicles in idle mode. Similarly, Fig. 2c depicts the number of EVs in  $SOC$  controllable mode and it is set in such way that the  $\eta_2$  value should be less than given  $\eta$ . Given  $\eta$  is 0.5, Fig. 2d shows  $\eta$  and  $\eta_2 = N_2/N$ . Based on these variables,  $g_0(t)$  and  $g(t)$  of Eq. 8 are shown in Fig. 3a and b. The time-dependent gain ( $K_e$ ) is illustrated in Fig. 3c.



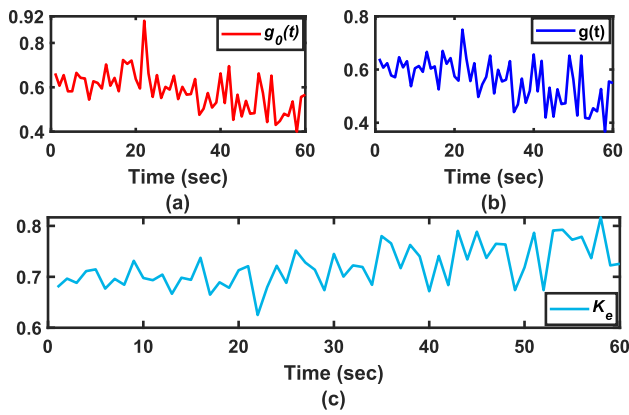


Fig. 3 EV aggregator parameter: a  $g_0$ , b  $g$ , and c  $K_e$

**Remark** In [43], the total number of EVs of the aggregator is experiencing rapid fluctuations with time between the minimum and maximum values. However, maintaining such variation during all times may not be feasible in practice. Therefore, in this paper, the total number of EVs is adjusted slowly over time and fluctuates within the minimum and maximum values. A wide range of variations in total number of EVs, including slow and fast variations and constant fluctuations, are considered.

The complete state space model of the proposed micro-grid is

$$\dot{x} = [A_1; A_2]x + Bu + \Gamma \Delta P_d \tag{9}$$

where  $x = [\Delta X_g, \Delta P_t, \Delta f, \Delta I, \Delta P_{PV}, \Delta P_W, \Delta P_{BESS1}, \Delta P_{BESS2}, \Delta P_{BESS3}, \Delta P_{FESS}, \Delta P_e]^T$ .  $u$  is the output of the controller.  $\Delta P_d = [\Delta P_{Solar}, \Delta P_{Wind}]^T$ .  $A_1$ ,  $A_2$ ,  $B$ , and  $\Gamma$  are as follows:

$$A_1 = \begin{bmatrix} -\frac{1}{T_g} & 0 & \frac{-1}{RT_g} & \frac{-1}{T_g} & 0 & 0 \\ \frac{1}{T_i} & \frac{-1}{T_i} & 0 & 0 & 0 & 0 \\ 0 & \frac{K_P}{T_P} & \frac{-1}{T_P} & 0 & \frac{K_P}{T_P} & \frac{K_P}{T_P} \\ 0 & 0 & 0 & K_I & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{T_{PV}} & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{1}{T_{WT}} \\ 0 & 0 & \frac{\alpha_1}{T_{BESS1}} & 0 & 0 & 0 \\ 0 & 0 & \frac{\alpha_2}{T_{BESS1}} & 0 & 0 & 0 \\ 0 & 0 & \frac{\alpha_3}{T_{BESS1}} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{T_{FESS}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \frac{K_P}{T_P} & \frac{K_P}{T_P} & \frac{K_P}{T_P} & \frac{K_P}{T_P} & \frac{K_P}{T_P} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{T_{BESS1}} & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{T_{BESS2}} & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{T_{BESS3}} & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{T_{FESS}} & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{T_e} \end{bmatrix}$$

$$B = \begin{bmatrix} \frac{1}{T_g} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\alpha K_e}{T_e} \end{bmatrix}^T$$

$$\Gamma = \begin{bmatrix} 0 & 0 & 0 & \frac{1}{T_{PV}} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{T_{WT}} & 0 & 0 & 0 & 0 \end{bmatrix}^T$$

This paper aims to design a controller that can effectively determine the  $u$  value, as per Eq. 9, for the above system.

### 3 Adaptive controller design

#### 3.1 IRL controller

In this section, an adaptive control strategy based on the IRL algorithm is presented. The IRL is an online learning technique of internal system dynamics [25]. In this article, the IRL algorithm is used for secondary frequency control. Consider the linear continuous-time system

$$\dot{x} = Ax + Bu \tag{10}$$

with the  $Q$ -value function

$$\begin{aligned} Q(x(t), u(t)) &= \int_t^\infty J(x(\tau), u(\tau)) d\tau \\ &= \int_t^{t+T} J(x(\tau), u(\tau)) d\tau \\ &\quad + Q(x(t+T), u(t+T)) \end{aligned} \tag{11}$$

where  $x \in R^n$  is the system state,  $u \in R^m$  is the control input,  $A \in R^{n \times n}$  and  $B \in R^{n \times m}$ .  $J(x, u)$  is the quadratic cost function. It is assumed that  $(A, B)$  is controllable. The cost function  $J(x, u)$  is defined as  $J(x, u) = x^T Qx + u^T R u$ .  $Q = Q^T \geq 0 \in R^{n \times n}$  and  $R = R^T > 0 \in R^{m \times m}$  are time-invariant weight matrices. By the Bellman's optimal principle [15], the  $K$  value can be found as

$$K = R^{-1} B^T P \tag{12}$$

where  $P$  is the unique positive definite solution of the algebraic Riccati equation (ARE)

$$A^T P + P A + Q - P B R^{-1} B^T P = 0. \tag{13}$$

The optimal continuous time  $Q$ -value function can be presented in quadratic form [44] as shown in Eq. 14.

$$\begin{aligned} Q^*(x, u) &= [x^T \ u^T] [S^*] \begin{bmatrix} x \\ u \end{bmatrix} \\ &= [x^T \ u^T] \begin{bmatrix} S_{11}^* & S_{12}^* \\ S_{21}^* & S_{22}^* \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \\ &= [x^T \ u^T] \begin{bmatrix} A^T P^* + P^* A + Q & P^* B \\ B^T P^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \end{aligned} \tag{14}$$

It can be noticed that the  $S^*$  is associated with  $P^*$  in ARE. By minimizing  $Q^*(x, u)$  with respect to  $u$ , the optimal solution  $u^*$  can be found as follows:

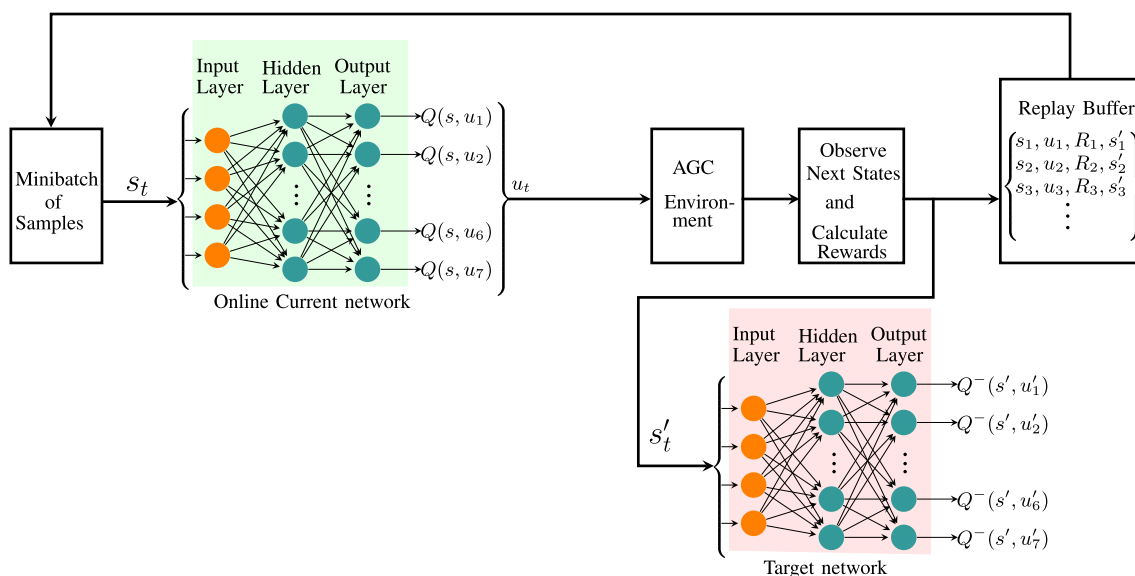
$$\begin{aligned} u^* &= -K^* x \\ &= -(S_{22}^*)^{-1} (S_{12}^*)^T x \end{aligned} \tag{15}$$

**Problem Statement:** Solving the value function (Eq. 14) yields the optimal policy ( $u$ ); however, the system matrix  $B$  is involved in this case, and it is unknown. The final objective is to implement the online adaptive IRL algorithm to get the optimal policy  $u^*$  (as per Eq. 15) without involving matrix  $B$ .

**Pseudo code of pre-learning of IRL algorithm:**

1. Initialize Time step ( $t$ ), Step size ( $T$ ),  $K(t)$ ,  $W(t)$ ,  $W(t - T)$  and  $x(t)$
2. Calculate  $u(t) = -K(t)x(t)$ . Apply the current policy to the system and observe the next state  $x(t + T)$ . Then calculate  $u(t + T) = -K(t)x(t + T)$ . Collect dataset  $(x(t), x(t + T), u(t), u(t + T))$  and compute  $\phi(z(t))$ ,  $\phi(z(t + T))$
3. Define the value function in parametric form as  $Q(x(t), u(t)) = W^T \phi(z(t))$ , where  $W$  is weight matrix,  $\phi(z(t))$  is the quadratic polynomial basis set,  $z(t)$  is the vector  $[x(t)^T \ u(t)^T]$ .  $\phi(z(t)) = z(t) \otimes z(t)$  and  $\otimes$  represents the Kronecker product [45]. The number of elements of  $W$  is  $n(n + 1)/2$ , where  $n$  represents the number of elements in  $z(t)$ .
4. Calculate the weight matrix as follows
 
$$W(t + T)^T = \Delta\phi^{-1} \int_t^{t+T} J(x, u) d\tau \tag{16}$$

where  $\Delta\phi = \phi(z(t)) - \phi(z(t + T))$ . The inverse of  $\Delta\phi$  cannot be found directly as it is a vector. In this study, the inverse is determined using the recursive least square (RLS) method [46]. So,  $\Delta\phi^{-1} = \Delta\phi^T (\Delta\phi\Delta\phi^T)^{-1}$ .
5. Unpack the vector  $W(t + T)$  into matrix  $S$  (shown in Eq. 14) and find  $K = (S_{22})^{-1} (S_{12})^T$ .
6. Check condition  $\|W(t + T) - W(t)\| < \zeta$  (Tolerance). If not so, set  $W(t) \leftarrow W(t + T)$ ,  $t \leftarrow t + T$ . Go to Step 2. Else Go to Step 7.
7. Get the final  $K$ .



**Fig. 4** Pictorial view of training process of DQN at time step  $t$



8. STOP

3.2 Deep Q-learning Controller

The DQN [47] uses two networks, namely the current ( $Q$ ) and target ( $Q^-$ ) networks. These networks are the same in structure.  $\theta$  and  $\theta^-$  are the parameter of the current network and target network, respectively. The current network selects the action ( $\arg \max Q(s, u; \theta)$ ) associated with the highest  $Q$  value. The target network evaluates the  $Q$  value ( $Q(s'_t, u'_t)$ ) of the target state ( $s'_t$ ). In this paper, these networks are three-layer backpropagation neural networks, with an input layer of four dimensions and an output layer of seven.

DQN employs a replay buffer to collect the experiences at each time step of pre-learning. At each time step ( $t$ ), a mini-batch of samples is drawn from the buffer and passed through the current neural network. The optimal action, obtained from greedy policy, is then executed in the environment, leading to the observation of next state ( $s'$ ) and the calculation of reward ( $R_t$ ). These experiences are stored as tuples containing the current state, action taken, reward obtained, and resulting next state. For example, the tuple for the  $i$ th sample is  $(s_i, u_i, R_t, s'_i)$ . Subsequently, the immediate next state is provided to the target network to determine the maximum  $Q^-$ -value. The pictorial view of pre-learning of DQN at time step ( $t$ ) is shown in Fig. 4. After pre-learning the current network is used as controller.

The loss function of  $i$ th sample of the mini-batch is expressed in Eq. 17.

$$L_i(\theta) = \left[ (R_t + \gamma \max Q_i^-(s', u'; \theta^-) - Q_i(s, u; \theta))^2 \right] \tag{17}$$

Here,  $R_t$  and  $\gamma$  are reward function and discount factor, respectively. The reward ( $R$ ) is chosen using Eq. 10.

$$R_t = -10 \Delta f^2 \tag{18}$$

The parameters  $\theta$  are updated using Eq. 19, where  $N$  is the number of samples in a mini-batch.

$$\theta_{t+1} = \theta_t + \alpha_l \frac{1}{N} \sum_i \nabla_{\theta_t} L_i(\theta_t) \tag{19}$$

where  $\nabla_{\theta_t} L(\theta_t)$  is the gradient of loss function with respect to parameter of current network  $\theta$  at time step  $t$ . After each  $c$  time step, the critic network's parameters are modified to match the actor network's.  $\theta^-$  is updated using Eq. 20.  $\theta_{t+c}$  is the critic parameter

$$\theta_t^- = \theta_{t+c} \tag{20}$$

Table 2 Parameters of the micro-grid

Parameter	Value	Parameter	Value	Parameter	Value
$T_g$	0.8 s	$T_{BESS_2}$	0.8 s	$T_P$	20
$T_t$	0.3 s	$T_{BESS_3}$	0.8 s	$\alpha_1$	0.3
$T_{PV}$	1 s	$T_{FESS}$	0.8 s	$\alpha_2$	0.5
$T_{WT}$	1 s	$T_e$	1 s	$\alpha_3$	0.2
$T_{BESS_1}$	0.8 s	$K_P$	120	$\alpha$	0.2
$K_I$	0.5			$\gamma$	0.9

Table 3 Initial K value

Simulations	$K = K_1, K_2, K_3, K_4$
Simulation1 (S1)	{0.8267, 1.000, 1.0, 1.4142}
Simulation2 (S2)	{0.267, .0500, 1.592, 2.4142}
Simulation3 (S3)	{0.701, 0.998, 4.380, 1.660}
Simulation4 (S4)	{3.082, 0.169, 0.491, 0.838}

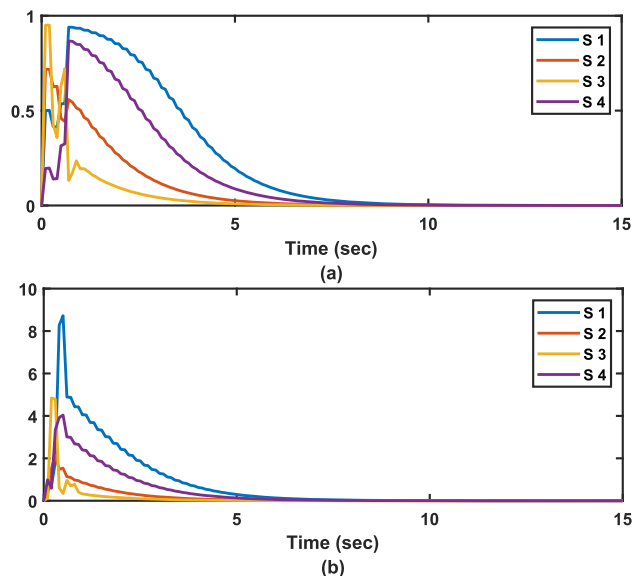


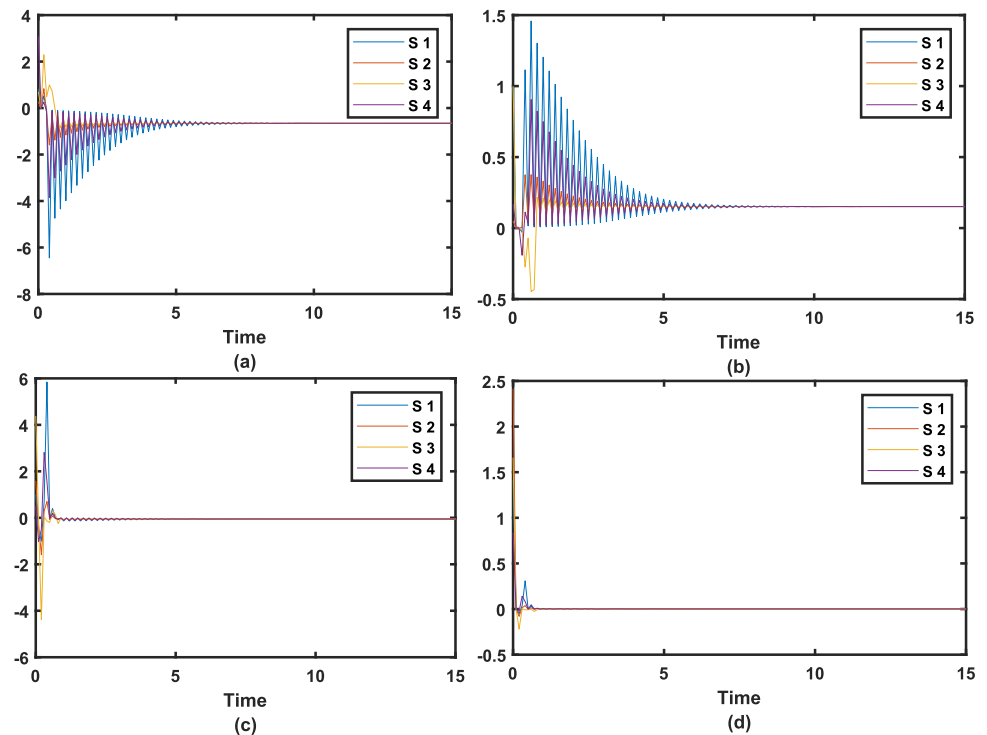
Fig. 5 Convergence Curves: a  $\|W_{t+T} - W_t\|$ , b  $\|K_{t+T} - K_t\|$

4 Simulation

This section demonstrates the performance of the proposed control technique for automatic generation control of the proposed micro-grid subject to unpredictability and uncertainty associated with load disturbances and an EV aggregator. Table 2 displays the parameters related to the micro-grid.

Pre-learning of proposed technique: Algorithm 1 depicts the pre-learning procedure of the IRL controller. It is assumed that the states ( $x = [\Delta X_g, \Delta P_t, \Delta f, \Delta I(t)]^T$ ) are observable.  $Q$  and  $R$ , involved in cost function  $J$ , are chosen as  $0.001I_{4 \times 4}$  and 0.08, respectively. The initial state  $x_0$  is  $[0.1, 0, 0, 0]^T$ . Step size ( $T$ ) is chosen as 0.1 s. Four differ-

**Fig. 6** Convergence Curves: **a**  $K_1$ , **b**  $K_2$ , **c**  $K_3$ , **d**  $K_4$



ent initial  $K$  values are chosen to show the convergence of the controller. The values are given in Table 3. The corresponding convergence curves are shown in Fig. 7, 5, and 6. Figure 7 shows the 2- norm of difference of weight matrix ( $W$ ) and the 2-norm of difference of  $K$  over consecutive time steps. The convergence of elements of  $K$  are illustrated in Fig. 5a–d. Figure 6a–o illustrates the convergence of elements of weight matrix ( $W$ ). Step size ( $T$ ) is chosen as 0.1 s. Based on algorithm 1, the kernel matrix  $S$  and optimal  $K$  values are found and expressed in Eqs. 21 and 22, respectively. Likewise, Fig. 8 depicts the convergence curve of the DQN in response to the sinusoidal load disturbance, as illustrated in Fig. 9.

#### 4.1 Case 1: Performance comparison under step load disturbance

This case study aims to investigate the performance of the proposed controller under step load disturbance. Notably, this analysis does not consider the uncertainties corresponding to the PV and wind systems. To appraise the effectiveness and adaptive performance of the proposed control strategy, a step load disturbance of 0.01 pu is applied to the system. Figure 10 depicts the outcomes of the simulation. The numerical values of the peak and peak time of the controllers are provided in Table 4. The peak value of the proposed method is reduced by 6% and 14% compared to DQN and PI controllers.

$$S = \begin{bmatrix} -0.16104 & 0.03718 & -0.012587 & 0.00035414 & 0.13471 \\ 0.03718 & -0.0085838 & 0.002906 & -8.1762e-05 & -0.031099 \\ -0.012587 & 0.002906 & -0.00098383 & 2.768e-05 & 0.010529 \\ 0.00035414 & -8.1762e-05 & 2.768e-05 & -7.788e-07 & -0.00029623 \\ 0.13471 & -0.031099 & 0.010529 & -0.00029623 & -0.20515 \end{bmatrix} \quad (21)$$

$$K = [-0.6566 \ 0.1516 \ -0.0513 \ 0.0014] \quad (22)$$

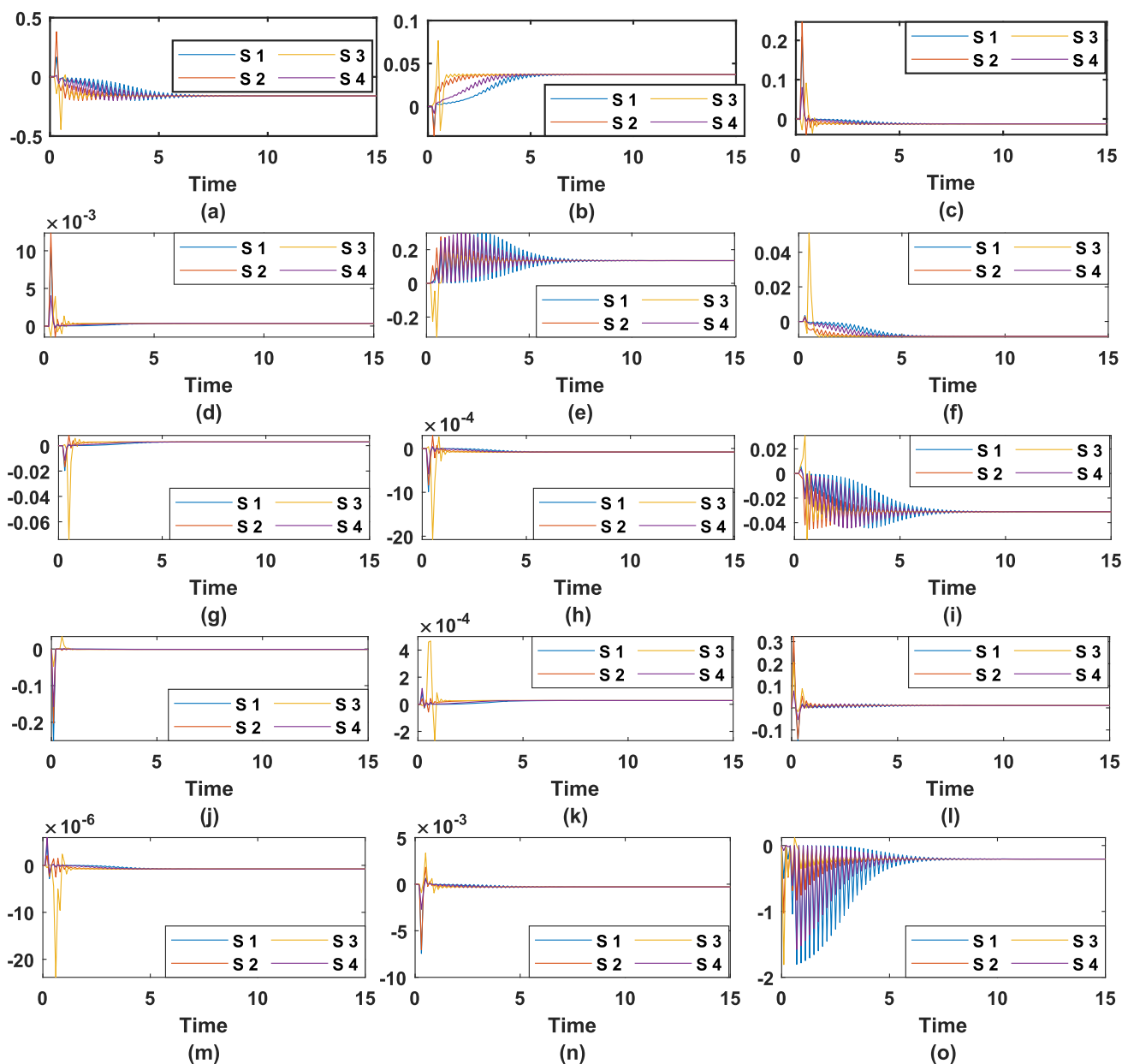


Fig. 7 Convergence Curves: a W1, b W2, c W3, d W4, e W5, f W6, g W7, h W8, i W9, j W10, k W11, l W12, m W13, n W14, and o W15

From the simulation result, it is observed that the proposed control enhances the transient performance by reducing the peak of the frequency deviation. The proposed controller shows a faster response compared to the DQN and PI controller.

### 4.2 Case 2: Performance comparison under multiple step load change

This case study evaluates the performance of the IRL controller under step frequent load changes. The effect of PV and wind power is not considered. The system is applied to

multi-step load perturbations, and the corresponding change in frequency is illustrated in Fig. 11. The numerical value of applied step load disturbances is given in Table 5.

The result demonstrates that the proposed controller helps the system to reach its steady state faster despite frequent disturbances. For performance evaluation, this case study considers three performance indices: mean squared error (MSE), integral absolute error (IAE), and integral time absolute error (ITAE) [48]. Tracking accuracy is evaluated using the MSE index. The IAE index evaluates the system’s overall overshoot. The ITAE index is used to assess the transient response time. The numerical results for these indices of

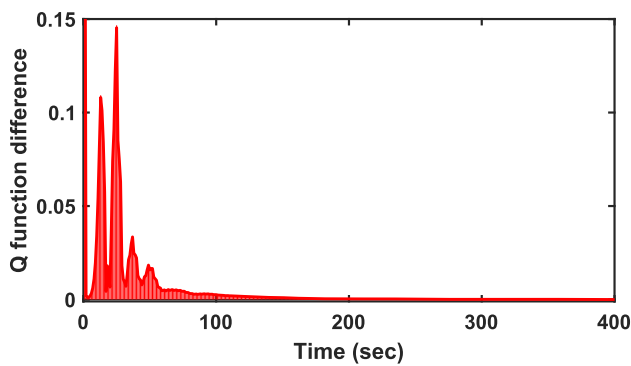


Fig. 8 Convergence curve for DQN

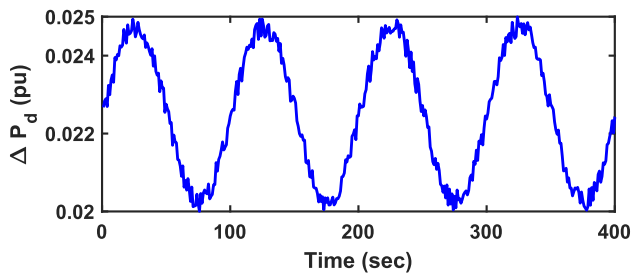


Fig. 9 Change in load for pre-learning of DQN

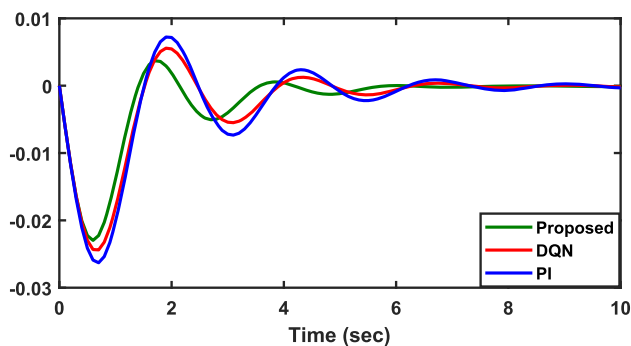


Fig. 10 Performance under step load disturbance: change in frequency ( $\Delta f$ ) in pu

frequency deviation are given in Table 6. The MSE for the proposed controller is decreased by 28% and 42% compared to the DQN and PI controller, respectively. Similarly, the IAE and ITAE for the proposed controller are approximately decreased by 21% and 35% percent compared to the DQN and PI controller, respectively. From these numerical performance values, it is observed that the proposed controller achieves better tracking performance with smaller tracking errors, smaller overshoots, and faster responses than other controllers.

Table 4 Step response of the system

Controller	Proposed	DQN	PI
Peak (pu)	0.0230	0.0244	0.0263
Peak time (s)	0.7	0.8	0.8

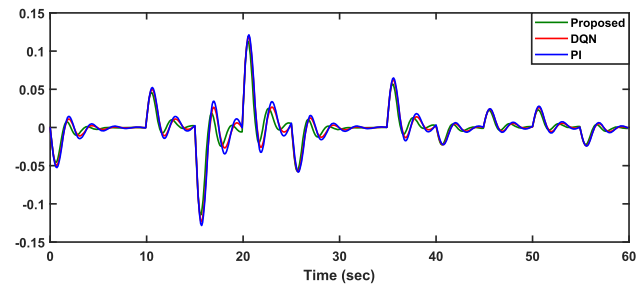


Fig. 11  $\Delta f$  (pu) Under multi-step load disturbances

### 4.3 Case 3: Performance comparison under renewable sources

This case study investigates the robustness of the proposed controller towards the PV and wind power uncertainty under random load disturbances. The system subjected to random load disturbances ( $\Delta P_d \in [0.04, 0.08]$  pu), as depicted in Fig. 12a. The perturbations in wind and solar power applied to the system are illustrated in Fig. 12b and c. The dynamic performances of the system are shown in Fig. 13. For performance evaluation, this case study considers the normal distribution of the frequency deviation. Figure 14 illustrates this. The proposed controller reduces the standard deviation in the frequency curve by 17% and 23% compared to DQN and PI controller. It is observed that the proposed controller exhibits a lower standard deviation compared to others, indicating less deviation frequency from the mean value. Furthermore, compared to other controllers, the mean value of frequency deviation for the proposed controller is closer to zero, which signifies a performance improvement. Therefore, it is observed that the probability of getting the frequency deviation close to zero is highest for the proposed controller. From the above discussion, it is concluded that the proposed controller effectively minimizes frequency deviations against the uncertainties imposed by PV and wind power.

### 4.4 Case 4: Sensitivity analysis

Theoretical examinations indicate that in cases where the power system exhibits inherent stability, variations in frequency remain relatively minor and constrained across varying load-damping coefficients. Conversely, within an unstable power system, disturbances can lead to amplified frequency deviations over time [49]. To demonstrate the

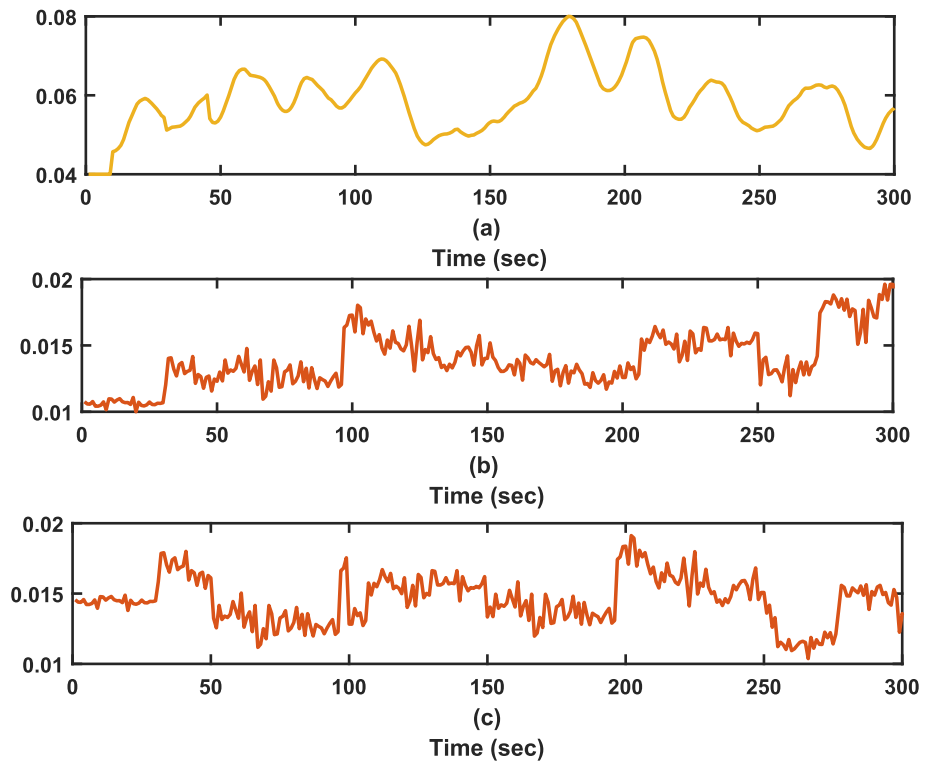
**Table 5** Applied step load to system

Time (s)	0–10	15–20	25–35	30–40	40–45	50–55	Else
Load (pu)	0.022	0.01	0.005	-0.005	0.015	0	0

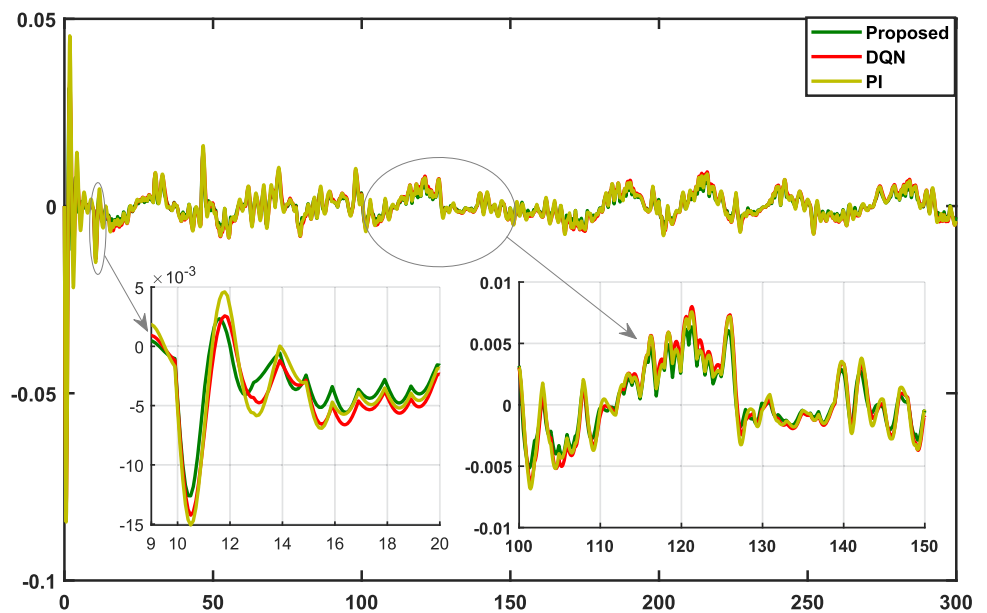
**Table 6** Numerical performance results

Performance indices	Proposed	DQN	PI
MSE	0.00044324	0.00057169	0.00063065
IAE	6.0292	7.311	8.1612
IATE	147.51	179.08	200.05

**Fig. 12** **a** Change in Load (pu), **b** Change in PV power (in pu) and **c** Change in Wind power (in pu)



**Fig. 13**  $\Delta f$  (pu) Under renewable sources



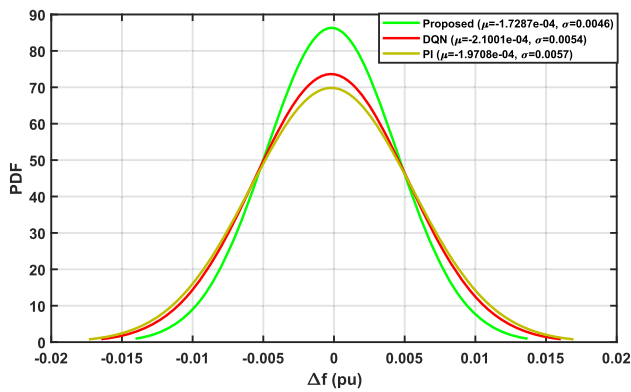


Fig. 14 Probability density of  $\Delta f$

robustness and efficacy of the proposed IRL controller in a more challenging condition, some critical parameters of the power system are varied and corresponding deviation in frequency ( $\Delta f$ ) is illustrated in Fig. 15. The change in parameters are shown in Table 7. At a time of 2 s in this case study, a step load disturbance of 0.02 per unit is applied. Figure 15 unveils the adaptability and robustness of the IRL controller to the variations in system parameters. Case 3 corresponds to no change in parameters. Despite changes to the system parameters, the desired outputs are bounded with small fluctuations.

Based on the analysis of case studies 1 and 2, it has been determined that the controller proposed is highly effective in reducing frequency deviations when faced with step load disturbances, even in the presence of EV uncertainties. Similarly, in case study 3, the proposed controller has demonstrated effectiveness in minimizing frequency deviations despite the uncertainties introduced by PV and wind power. The controller has also been shown to be robust against parameter changes, which further highlights its reliability and efficiency.

In the given study, the target network of DQN updates its parameters without requiring the gradient of the loss function; instead, it directly copies parameters from the current network. In contrast, the current network requires the gradient of the loss function to update its parameters. As shown in Fig. 4, the current networks employ gradient ascent to update 77 (weights of the current network) parameters at each time step. However, the proposed controller adopts the RLS method, requiring only 15 parameter updates. In this paper, the simulation results show that the proposed controller performs better than the DQN algorithm. However, DQN may achieve enhanced performance, particularly with deep neural networks featuring more hidden layers and a larger discretized action set. However, the complexity of the controller will increase, causing an increase in the number of parameter updates and learning time. Moreover, the proposed controller needs no discretization of action set.

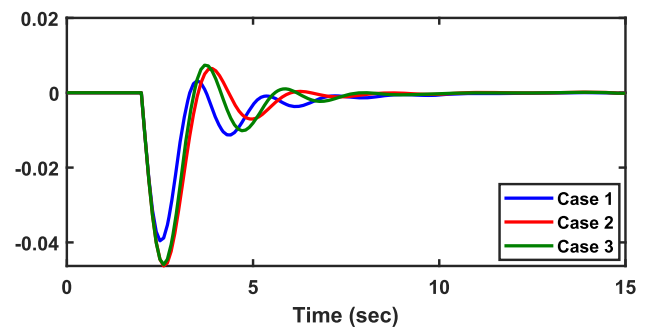


Fig. 15  $\Delta f$  (pu) under step load disturbance

Table 7 Change in parameters of the power system

Parameters	Variations (%)	Parameters/.	Variations (%)
<i>Case 1</i>			
$R$	-42	$K_P$	25
$T_g$	-50	$T_p$	25
$T_l$	33		
<i>Case 2</i>			
$R$	40	$K_P$	-25
$T_g$	-25	$T_p$	33
$T_l$	-17		

## 5 Conclusion

An optimal control technique that uses integral reinforcement learning has been successfully implemented for automatic generation control of EV, renewable, and storage system integrated is-landed micro-grid. The proposal includes four case studies that analyze the robustness of the controller against uncertainties due to load disturbances, EV aggregator, renewable integration, and parameter uncertainty. The case studies show that the proposed IRL controller effectively minimizes the frequency deviation. The proposed controller markedly enhances performance metrics relative to the DQN and PI controllers. Specifically, it achieves a reduction in peak frequency deviation by 6% and 14% compared to the DQN and PI controllers, respectively. Under multiple-step load disturbances, the controller decreases the mean square error by 28% and 42%, respectively, while reducing both the integral absolute error and the integral time absolute error by 21% and 35% in comparison to the DQN and PI controllers. Furthermore, in scenarios involving renewable energy sources, the proposed controller lowers the standard deviation in frequency deviation by 17% compared to the DQN controller and by 23% compared to the PI controller. The proposed method is a policy iteration-based algorithm. The main limitation of the proposed method is it needs permissible initial conditions, which have been determined based on human expertise in this study. This paper applies the method to



a single-area microgrid, but future work could potentially expand to interconnected multi-microgrids.

## References

1. Yang F, Feng X, Li Z (2019) Advanced Microgrid Energy Management System for Future Sustainable and Resilient Power Grid. *IEEE Trans Ind Appl* 55(6):7251–7260. <https://doi.org/10.1109/TIA.2019.2912133>
2. Kerdphol T, Rahman FS, Watanabe M, Mitani Y, Turschner D, Beck H-P (2019) Enhanced virtual inertia control based on derivative technique to emulate simultaneous inertia and damping properties for microgrid frequency regulation. *IEEE Access* 7:14422–14433. <https://doi.org/10.1109/ACCESS.2019.2892747>
3. Tang X, Deng W, Qi Z (2014) Investigation of the dynamic stability of microgrid. *IEEE Trans Power Syst* 29(2):698–706. <https://doi.org/10.1109/TPWRS.2013.2285585>
4. Nikkhajoei H, Lasseter RH (2009) Distributed generation interface to the CERTS microgrid. *IEEE Trans Power Delivery* 24(3):1598–1608. <https://doi.org/10.1109/TPWRD.2009.2021040>
5. Farrokhabadi M et al (2020) Microgrid stability definitions, analysis, and examples. *IEEE Trans Power Syst* 35(1):13–29. <https://doi.org/10.1109/TPWRS.2019.2925703>
6. Khooban M-H, Dragicevic T, Blaabjerg F, Delimar M (2018) Ship-board microgrids: a novel approach to load frequency control. *IEEE Trans Sustain Energy* 9(2):843–852. <https://doi.org/10.1109/TSTE.2017.2763605>
7. Muduli R, Jena D, Moger T (2023) Application of expected Sarsa-learning for load frequency control of multi-area power system. In: 2023 5th International conference on energy, power and environment: towards flexible green energy technologies (ICEPE), Shillong, India, pp. 1–6 <https://doi.org/10.1109/ICEPE57949.2023.10201593>
8. Sadasiva Behera, Nalin B, Dev Choudhury (2022) Modelling and simulations of modified slime mould algorithm based on fuzzy PID to design an optimal battery management system in microgrid. *Cleaner Energy Syst* 3:100029. <https://doi.org/10.1016/j.cles.2022.100029>
9. Adibi M, van der Woude J (2022) Secondary frequency control of microgrids: an online reinforcement learning approach. *IEEE Trans Autom Control* 67(9):4824–4831. <https://doi.org/10.1109/TAC.2022.3162550>
10. Guo W, Liu F, Si J, He D, Harley R, Mei S (2016) Online supplementary ADP learning controller design and application to power system frequency control with large-scale wind energy integration. *IEEE Trans Neural Netw Learn Syst* 27(8):1748–1761. <https://doi.org/10.1109/TNNLS.2015.2431734>
11. Singh V, Moger T, Jena D (2022) Uncertainty handling techniques in power systems: a critical review. *Electr Power Syst Res* 203:107633
12. Li H, Liu D, Wang D (2014) Integral reinforcement learning for linear continuous-time zero-sum games with completely unknown dynamics. *IEEE Trans Autom Sci Eng* 11(3):706–714. <https://doi.org/10.1109/TASE.2014.2300532>
13. Liu D, Xue S, Zhao B, Luo B, Wei Q (2021) Adaptive dynamic programming for control: a survey and recent advances. *IEEE Trans Syst, Man, Cybernet: Syst* 51(1):142–160. <https://doi.org/10.1109/TSMC.2020.3042876>
14. Muduli R, Nair N, Kulkarni S, Singhal M, Jena D, Moger T (2023) Load frequency control of two-area power system using an actor-critic reinforcement learning method-based adaptive PID controller. In: 2023 IEEE 3rd international conference on sustainable energy and future electric transportation (SEFET), Bhubaneswar, India, pp. 1–6 <https://doi.org/10.1109/SeFeT57834.2023.10245225>
15. Zhu LM, Modares H, Peen GO, Lewis FL, Yue B (2015) Adaptive suboptimal output-feedback control for linear systems using integral reinforcement learning. *IEEE Trans Control Syst Technol* 23(1):264–273. <https://doi.org/10.1109/TCST.2014.2322778>
16. Abouheaf M, Gueaieb W, Sharaf A (2019) Load frequency regulation for multi-area power system using integral reinforcement learning. *IET Generat, Transm Distrib* 13(19):4311–4323
17. Xi L, Zhou L, Liu L, Duan D, Xu Y, Yang L, Wang S (2020) A deep reinforcement learning algorithm for the power order optimization allocation of agc in interconnected power grids. *CSEE J Power Energy Syst* 6(3):712–723
18. Xi L, Yu L, Xu Y, Wang S, Chen X (2020) A Novel Multi-Agent DDQN-AD Method-Based Distributed Strategy for Automatic Generation Control of Integrated Energy Systems. *IEEE Transactions on Sustainable Energy* 11(4):2417–2426. <https://doi.org/10.1109/TSTE.2019.2958361>
19. Li J, Yu T, Zhang X (2022) Coordinated load frequency control of multi-area integrated energy system using multi-agent deep reinforcement learning. *Appl Energy* 306:117900
20. Khooban MH, Gheisarnejad M (2021) A novel deep reinforcement learning controller based type-II fuzzy system: frequency regulation in microgrids. *IEEE Trans Emerg Top Computat Intell* 5(4):689–699. <https://doi.org/10.1109/TETCI.2020.2964886>
21. Zheng Y, Tao J, Sun H, Sun Q, Chen Z, Dehmer M, Zhou Q (2021) Load frequency active disturbance rejection control for multisource power system based on soft actor-critic. *Energies* 14(16):4804
22. Muduli R, Jena D, Moger T (2024) Application of reinforcement learning-based adaptive pid controller for automatic generation control of multi-area power system. *IEEE Trans Automat Sci Eng.* <https://doi.org/10.1109/TASE.2024.3359219>
23. Abouheaf Mohammed, Gueaieb Wail, Sharaf Adel (2019) Load frequency regulation for multi-area power system using integral reinforcement learning. *IET Generat, Trans Distrib* 13(19):4311–4323
24. Ou R, Xu Y, Li Z, Hu R, Huang J, Yu J (2021) Integral reinforcement learning-based adaptive optimal automatic generation control. In: 2021 4th International conference on energy, electrical and power engineering (CEEPE), pp. 643–649 <https://doi.org/10.1109/CEEPE51765.2021.9475666>
25. Vrabie D, Pastravanu O, Abu-Khalaf M, Lewis FL (2009) Adaptive optimal control for continuous-time linear systems based on policy iteration. *Automatica* 45(2):477–484
26. Qu Z, Xu C, Yang F, Ling F, Pirouzi S (2023) Market clearing price-based energy management of grid-connected renewable energy hubs including flexible sources according to thermal, hydrogen, and compressed air storage systems. *J Energy Storage* 69:107981
27. Zhang X, Yu X, Ye X, Pirouzi S (2023) Economic energy management of networked flexi-renewable energy hubs according to uncertainty modeling by the unscented transformation method. *Energy* 278:128054
28. Khalafian F, Iliiae N, Diakina E, Parsa P, Alhaider MM, Masali MH, Zhu M (2024) Capabilities of compressed air energy storage in the economic design of renewable off-grid system to supply electricity and heat costumers and smart charging-based electric vehicles. *J Energy Storage* 78:109888
29. Liang H, Pirouzi S (2024) Energy management system based on economic Flexi-reliable operation for the smart distribution network including integrated energy system of hydrogen storage and renewable sources. *Energy* 293:130745
30. Pirouzi S (2023) Network-constrained unit commitment-based virtual power plant model in the day-ahead market according to energy management strategy. *IET Generat, Trans Distrib* 17(22):4958–4974

31. Norouzi M, Aghaei J, Niknam T, Pirouzi S, Lehtonen M (2022) Bi-level fuzzy stochastic-robust model for flexibility valorizing of renewable networked microgrids. *Sustain Energy, Grids Netw* 31:100684
32. Norouzi M, Aghaei J, Pirouzi S, Niknam T, Fotuhi-Firuzabad M (2022) Flexibility pricing of integrated unit of electric spring and EVs parking in microgrids. *Energy* 239:122080
33. Cavus M, Allahham A, Adhikari K, Giaouris D (2024) A hybrid method based on logic predictive controller for flexible hybrid microgrid with plug-and-play capabilities. *Appl Energy* 359:122752
34. Cavus M, Allahham A, Adhikari K, Zangiabadi M, Giaouris D (2023) Energy management of grid-connected microgrids using an optimal systems approach. *IEEE Access* 11:9907–9919
35. Wang C, Mi Y, Fu Y, Wang P (2018) Frequency control of an isolated micro-grid using double sliding mode controllers and disturbance observer. *IEEE Trans Smart Grid* 9(2):923–930. <https://doi.org/10.1109/TSG.2016.2571439>
36. Mahmoud TS, Ahmed BS, Hassan MY (2019) The role of intelligent generation control algorithms in optimizing battery energy storage systems size in microgrids: a case study from Western Australia. *Energy Convers Manage* 196:1335–1352
37. Chen S, Zhang T, Gooi HB, Masiello RD, Katzenstein W (2016) Penetration Rate and Effectiveness Studies of Aggregated BESS for Frequency Regulation. *IEEE Transactions on Smart Grid* 7(1):167–177. <https://doi.org/10.1109/TSG.2015.2426017>
38. Bevrani H, Habibi F, Babahajyani P, Watanabe M, Mitani Y (2012) Intelligent frequency control in an AC microgrid: online PSO-based fuzzy tuning approach. *IEEE Trans Smart Grid* 3(4):1935–1944. <https://doi.org/10.1109/TSG.2012.2196806>
39. Oshnoei A, Kheradmandi M, Muyeen SM (2020) Robust control scheme for distributed battery energy storage systems in load frequency control. *IEEE Trans Power Syst* 35(6):4781–4791. <https://doi.org/10.1109/TPWRS.2020.2997950>
40. Khooban M-H (2018) Secondary load frequency control of time-delay stand-alone microgrids with electric vehicles. *IEEE Trans Industr Electron* 65(9):7416–7422. <https://doi.org/10.1109/TIE.2017.2784385>
41. Javanmardi H, Dehghani M, Mohammadi M, Siamak S, Hesamzadeh MR (2022) BMI-based load frequency control in microgrids under false data injection attacks. *IEEE Syst J* 16(1):1021–1031. <https://doi.org/10.1109/JSYST.2021.3054947>
42. Rajesh KS, Dash SS (2019) Load frequency control of autonomous power system using adaptive fuzzy based PID controller optimized on improved sine cosine algorithm. *J Ambient Intell Humaniz Comput* 10:2361–2373
43. Pham TN, Nahavandi S, Hien LV, Trinh H, Wong KP (2017) Static output feedback frequency stabilization of time-delay power systems with coordinated electric vehicles state of charge control. *IEEE Trans Power Syst* 32(5):3862–3874. <https://doi.org/10.1109/TPWRS.2016.2633540>
44. Li H, Liu D, Wang D (2014) Integral reinforcement learning for linear continuous-time zero-sum games with completely unknown dynamics. *IEEE Trans Autom Sci Eng* 11(3):706–714. <https://doi.org/10.1109/TASE.2014.2300532>
45. Schacke K (2004) On the Kronecker product. Master's thesis, University of Waterloo
46. Lewis JM, Lakshminarayanan S, Dhall S (2006) Dynamic data assimilation: a least squares approach, vol 13. Cambridge University Press, Cambridge
47. Mnih V, Kavukcuoglu K, Silver D, Graves A, Antonoglou I, Wierstra D, Riedmiller M (2013) Playing atari with deep reinforcement learning. arXiv preprint [arXiv:1312.5602](https://arxiv.org/abs/1312.5602)
48. Stojanovic V (2023) Fault-tolerant control of a hydraulic servo actuator via adaptive dynamic programming. *Math Model Control* 3(3):181–191. <https://doi.org/10.3934/mmc.2023016>
49. Huang H, Li F (2013) Sensitivity analysis of load-damping characteristic in power system frequency regulation. *IEEE Trans Power Syst* 28(2):1324–1335. <https://doi.org/10.1109/TPWRS.2012.2209901>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.