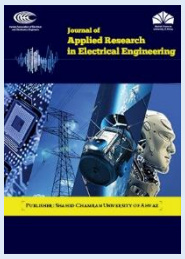
Shahid Chamran University of Ahvaz

Iranian Association of Electrical and Electronics Engineers

Journal of Applied Research in Electrical Engineering

E-ISSN: 2783-2864
P-ISSN: 2717-414X
Homepage: https://jaree.scu.ac.ir/



Research Article

Distribution Network Reconfiguration and Radiality Detection using Deep Reinforcement Learning

Amirhossein Ghaemipour[1], Habib Rajabi Mashhadi[1,2,*], and Seyed Hossein Mostafavi[1]

[1] *Department of Electrical Engineering, Faculty of Engineering, Ferdowsi University of Mashhad, Mashhad, Iran*

[2] *Center of Excellence on Soft Computing and Intelligent Information Processing (SCIIP), Faculty of Engineering, Ferdowsi University of Mashhad, Mashhad, Iran*

** Corresponding Author: h_mashhadi@um.ac.ir*

Abstract: Distribution network reconfiguration (DNR) is one of the most effective methods for reducing losses and damages in distribution networks, resulting in the least investment for power companies. One of the most critical constraints of the distribution network is that it remains radial during operation. In this paper, a novel method is presented for detecting the radiality of distribution networks. Then, this novel method is combined with the Double Deep Q Network algorithm, which is used to determine the optimal structure of the IEEE 33- and 69-bus distribution networks. This approach is implemented to minimize power losses and voltage deviation. Additionally, distributed generation sources, such as wind turbines and photovoltaic cells under fixed and uncertain capacity with varying load conditions, are utilized to investigate their impact on this approach.

Keywords: Distribution network radial constraint, optimal distribution network structure, deep reinforcement learning

Article history
Received 12 October 2025; Revised 31 January 2026; Accepted 25 February 2026; Published online 26 July 2026.
© 2026 Published by Shahid Chamran University of Ahvaz & Iranian Association of Electrical and Electronics Engineers (IAEEE)
How to cite this article
A. Ghaemipour, H. Rajabi Mashhadi, and S. H. Mostafavi, " Distribution Network Reconfiguration and Radiality Detection using Deep Reinforcement Learning," *J. Appl. Res. Electr. Eng.*, vol. 4, no. 2, pp. 165-177, 2025. DOI: https://doi.org/10.22055/jaree.2026.50841.1267

Check for updates

1. INTRODUCTION

Electricity is one of the most vital needs of humans, and meeting this need is essential. Providing electricity to customers involves three stages: generation, transmission, and distribution. Numerous studies have been conducted to increase efficiency [1][2] and maintain reliability [3][4] in each of these sectors. Considering the advances made in the field of algorithms, artificial intelligence is one of the most recent methods for solving problems. Also, this is applicable in both the academic and industrial worlds, with numerous applications across multiple industries, particularly in the electricity sector. For example, deep learning is used to predict wind direction [5], to forecast wind turbine power generation [6], to fault location [7], and to predict customers' electricity prices, achieving perfect accuracy. Considering the applicability of these algorithms and the sensitivity of the distribution sector due to its proximity to subscribers [8], this research has focused on combining these two areas. The optimal arrangement of the distribution network (DN) has a significant impact on sensitive conditions for DN operators, enabling them to minimize power losses by operating optimally. The following section discusses previous research that utilized different algorithms for distribution network reconfiguration (DNR).

In [9], a genetic algorithm (GA) was used to rearrange the DN to minimize power losses and voltage deviation. This was tested on 16- and 33-bus DNs. In [10], the rearrangement of 33-, 85-, and 295-bus DNs was carried out. This approach utilized the Harris Hawks optimization (HHO) algorithm to minimize voltage deviation and power losses. In [11], the reconfiguration of 33-, 69-, and 119-bus DNs was performed using the improved cuckoo search algorithm (ICSA). This sought to minimize voltage deviation and power losses. In [12], the selective firefly algorithm was used for DNR on 33-, 70-, and 84-bus DNs. This method was implemented to minimize power losses. In [13], the DNR was implemented by determining the capacity of distributed generation (DG) in DNs with 16-, 33-, 69-, 83-, 119-, and 202-buses. This aims to minimize power losses and voltage deviation. This approach was modeled using mixed-integer linear programming (MILP). In [14], the combined algorithm grey

wolf optimizer (GWO) and particle swarm optimizer (PSO) was utilized to DNR and find the optimal capacity and location of DGs. This process was tested on the DNs of 33- and 69-bus. Also, it is intended to minimize power losses. In [15], the ant colony (AC) algorithm was used for DNR of 33- and 69-buses to decrease power losses and voltage deviation. In [16], DNR was applied to 33-, 69-, and 119-bus DNs to reduce power losses and voltage deviation. This study modeled the system using feasibility-preserving evolutionary optimization (FPEO). In [17], DNR for 33-, 84-, 119-, and 136-bus DNs has been modeled. This method has been implemented using the chaotic stochastic fractal search algorithm (CSFSA). This approach seeks to minimize power losses and voltage deviations. Dynamic distribution network reconfiguration (DDNR) is implemented on a 95-bus DN in reference [18]. This is evaluated by considering DGs and electric vehicle (EV) charging stations. This evaluation seeks to minimize operating costs and energy not supplied (ENS). DNR is evaluated in a multi-objective manner for the 33-bus DN in [19]. It is modeled based on graph theory and seeks to minimize power losses, voltage deviations, ENS, and load shedding.

In [20], the DNR was modeled using the switch opening and exchange (SOE) method and mathematical programming (MP) to minimize power losses and voltage deviations. That study tested the model on 33-, 84-, 119-, 136-, and 417-bus DNs. In [21], the DNR was performed to reduce power losses and voltage deviation. This modeling was performed under a binary group search optimization (BGSO) on 33- and 69-bus DNs. In [22], a modified sequential switch opening and exchange (MSSOE) method was used to evaluate the reconfiguration of 33-, 69-, and 119-bus DNs. The objective was to minimize power losses and voltage deviation. In [23], the harmonic search algorithm (HSA) was employed to DNR and DG locating in 33- and 69-bus DNs. This approach was simulated to reduce power losses and voltage deviation. In [24], the DNR was utilized to minimize energy losses in a 33-bus DN. This approach was modeled using mixed-integer non-linear programming (MINLP), which employed wind turbines and photovoltaic (PV) cells. DNR has been done in [25] for 33- and 70-bus DNs. This reference is based on the Improved PSO (IPSO), seeking to minimize the ENS and enhance the voltage stability index (VSI). The effect of DGs has been evaluated in this modeling. The DDNR introduced in [26] is applied to a 95-bus DN. It seeks to reduce ENS, operating costs, and energy losses. This evaluation is based on a hybrid IPSO and artificial bee colony optimization algorithm that considers the impact of DGs and EVs.

In [27], DNR for 33- and 69-bus DNs has been carried out to minimize power losses and voltage deviations. This method has been modeled using the NoisyNet deep Q network (NoisyNet-DQN) algorithm under uncertainty in wind turbines. Reconfiguration of 33-, 119-, and 136-bus DNs in [28] has been carried out to decrease power losses and voltage deviations. This process has been conducted using the deep Q network (DQN) and dueling deep Q network (Dueling DQN) algorithms. This method assumes that the problem dimensions have been reduced. In [29], a DNR model is introduced for smaller-scale DNs, such as a 16-bus DN. This model aims to decrease power losses and voltage deviations. In [30], the reconfiguration of 33- and 136-bus DNs has been modeled using the DQN, Dueling DQN, DQN with prioritized experience replay, proximal policy optimization (PPO), and soft actor-critic (SAC) algorithms. This modeling has been investigated to reduce power losses and voltage deviations by assuming a decreased action space. DNR based on maintaining stability and reliability for 14- and 123-bus DNs is presented in [31]. The primary basis of this model is DQN, with uncertainty in DGs. The DQN model in [32] is oriented for DDNR and seeks to maximize reliability while decreasing voltage deviation and power losses. This investigates the impact of DGs on 33- and 118-bus DNs. The DQN suggested in [33] is for DDNR and seeks to minimize the total cost of power loss and switching. This analyzes the impact of DGs and different loads in 15- and 33-bus DNs. DRL based on explainer neural networks is presented in [34] . It looks to reduce power losses and voltage deviations in the 33- and 136-bus DNs through daily analysis.

In the above explanations, several prior research studies in this field were described. None of them specifically mentions the radiality constraint of the DNs. It should be noted that during the operation of the DNs, the final structure of the DNs should remain radial. In this paper, an experimental method is employed to detect the radial constraint of the DN. This method is utilized to determine the structure of the IEEE 33- and 69-bus DNs using the double deep Q network (DDQN) algorithm. A comparison of the suggested modeled structure with similar previous DRLs is given in Table 1.

The rest of the paper is organized as follows. In Section 2, DRL and the agent learning method are described. In Section 3, the proposed method is modeled using an empirical approach, along with DRL. Section 4 introduces the details of the DNs used in this article. In Section 5, the results of applying the suggested method to the IEEE 33- and 69-bus DN are presented. Finally, a conclusion is drawn in the final section.

Table 1: Comparison of the suggested design and the previous DRL.

Ref	Year	Daily load	Reward		DG		SA	State	Power flow software
			losses	VD	Fixed	Uncertainty			
[27]	2021	✓	✓	✓	×	✓	×	p_t, q_t	Pypower
[28]	2024	✓	✓	✓	×	×	✓	α_t, p_t, q_t	Pandapower
[29]	2025	×	✓	✓	✓	×	×	α_t, v_t	Pandapower
[30]	2023	✓	✓	✓	×	×	×	α_t, p_t, q_t	Pandapower
[34]	2024	✓	✓	✓	×	×	×	α_t, p_t, q_t	Pandapower
Proposed method	2026	✓	✓	✓	✓	✓	✓	α_t, p_t, q_t	Pandapower, Matpower

p_t, q_t: Active & reactive power of buses, α_t: Condition of lines, v_t: Voltage magnitude of buses, VD: Voltage deviation, SA: Sensitivity analysis.

2. DEEP REINFORCEMENT LEARNING

DRL combines the decision-making capability of reinforcement learning (RL) with the function approximation power of deep neural networks. In classical RL, storing and updating a value for every state-action pair becomes infeasible in high-dimensional or continuous environments. Neural networks address this limitation by generalizing across similar states and learning compact representations, which enables agents to operate in complex environments.

One of the basic and important DRL algorithms is DQN. It uses a neural network to estimate the Q-value function $Q(s,a)$, which tells the agent how good an action is in a given state. The goal of DQN is to reduce the difference between the predicted Q-value and the target Q-value, which is calculated using the Bellman equation. The loss function for DQN is [35]:

$$L_{DQN}(\theta) = \mathbb{E}_{(s,a,r,s')}\left[\frac{1}{2}(y^{DQN} - Q(s,a;\theta))^2\right] \quad (1)$$

$$y^{DQN} = r + \gamma \max_{a'} Q(s', a'; \theta^-) \quad (2)$$

where θ represents the parameters of the main Q-network, and θ^-denotes the parameters of the target network, which is updated less frequently to stabilize training. DQN often gives Q-values that are too high because it uses the max operator when calculating the target value. To fix this problem, a new method called DDQN was introduced. In DDQN, choosing the best action and measuring its value are done separately. The main network picks the action, and the target network gives its value. The loss function for DDQN is [36]:

$$L_{DDQN}(\theta) = \mathbb{E}_{(s,a,r,s')}\left[\frac{1}{2}(y^{DDQN} - Q(s,a;\theta))^2\right] \quad (3)$$

$$y^{DDQN} = r + \gamma Q(s', \underset{a'}{\operatorname{argmax}}(s', a', \theta); \theta^-) \quad (4)$$

This formulation leads to more accurate value estimates and often improves learning stability and performance. In practical applications, a replay buffer is used to train the neural network. The experiences stored in this memory are represented as tuples of four elements (s_j, a_j, r_j, s'_j), and during training, a random mini-batch is sampled from this memory as follows:

$$L_{DDQN}(\theta) = \mathbb{E}_{(s,a,r,s')\sim u(D)}\left[\frac{1}{2}(y^{DDQN} - Q(s,a;\theta))^2\right] \quad (5)$$

In DRL, the gradient descent method is employed to train the neural network. The gradient of the network parameters is generally expressed as follows:

$$\nabla_\theta L_{DDQN}(\theta) = \mathbb{E}_{(s,a,r,s')\sim u(D)}[-(y^{DDQN} - Q(s,a;\theta))\nabla_\theta Q(s,a;\theta)] \quad (6)$$

Finally, utilizing the gradient descent method, the loss function is updated as follows:

$$\theta = \theta - \eta \nabla_\theta L_{DDQN}(\theta) \quad (7)$$

where η is the learning rate. this paper uses the DDQN algorithm for DNR.

3. PROBLEM MODELING

3.1. Experimental method for detecting radiality of distribution network.

The radiality of the DN is one of the most critical constraints that must be considered when operating the DN. Most articles do not directly address how to detect and model the radiality of the DN during operation. Therefore, in this research, we have developed an experimental method for the modeling process, which we refer to as the NAN method. This method is based on the implementation of power flow on the DN, which has been analyzed in MATLAB and Python environments using the MATPOWER and Pandapower packages [37], respectively. The key selection process in the DN is such that the following technical and security constraints must be met:

- No islands are formed in the DN, and no bus is isolated.
- All customers' demands are supplied, and there is no subscriber without electricity.
- No loops exist during the operation of the DN, and the DN has a radial structure in the end.

The process of understanding each of these cases after implementing power flow with these packages from the DN is illustrated in Fig. 1. Initially, all the tie switches are closed. The process mentioned is repeated until the correct switches are selected to open, and there are no loops in the DN. It can then be concluded that the DN has a radial structure and that the switches chosen are technically suitable.

To better understand how the NAN method operates, we provide an example using the IEEE 69-bus DN. In this method, first, all tie switches are closed, and for example assumed that lines 5 and 7 were opened by the DN operator. By performing the power flow analysis using PandaPower or MATPOWER, the output data from buses and lines is presented in tabular form. As illustrated in Fig. 2, opening lines 5 and 7 results in the formation of a yellow-colored island.

- Based on Fig. 1, when using PandaPower, if in the output results for the buses you observe "NAN" values in the columns vm_pu (voltage magnitude in p.u.) and va_degree (voltage angle in degrees), it indicates that some buses are isolated. In this example, buses 6 and 7 are isolated. Furthermore, in the output results for the lines, except for those lines explicitly opened by the operator, if you observe a value of zero in the pl_mw (active power passing of lines) and ql_mvar (reactive power passing of lines) columns or "NAN" in the i_ka (current passing of lines) column, it indicates that an island has formed in the DN. Consequently, in this example, an island is created which includes line 6 within it.
- In MATPOWER modeling, "NAN" values in columns 8 and 9 (bus voltage magnitude and angle) indicate isolated buses. In this example, buses 6 and 7 are isolated. Similarly, "NAN" values in columns 14 through 17 of the line results (active and reactive power flow) signify the formation of an island within the DN, which includes line 6.

In summary, by utilizing the NAN method, we can evaluate the DN's necessary constraints and ensure its operational feasibility. The application of this method within the proposed algorithm is explained in the following section and Algorithm 1.

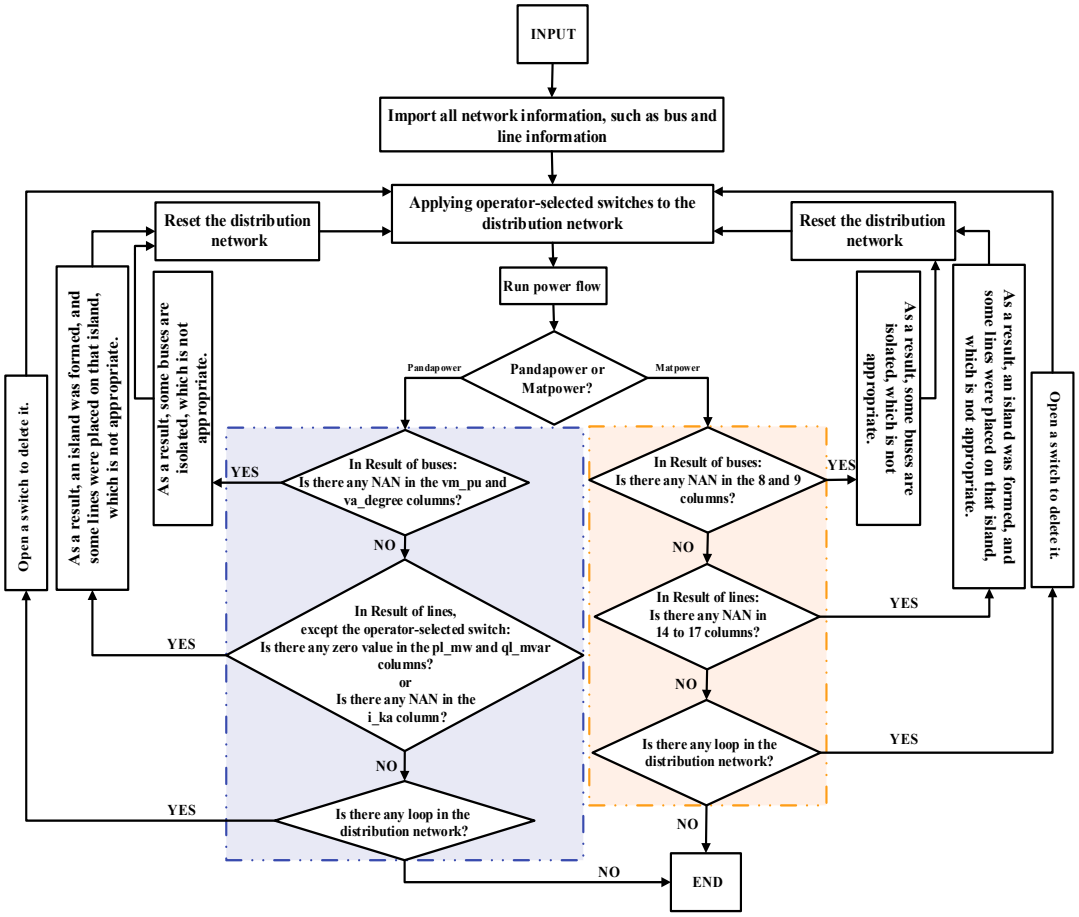


Fig. 1: The overall process of implementing the NAN method to detect the radiality of the DN.

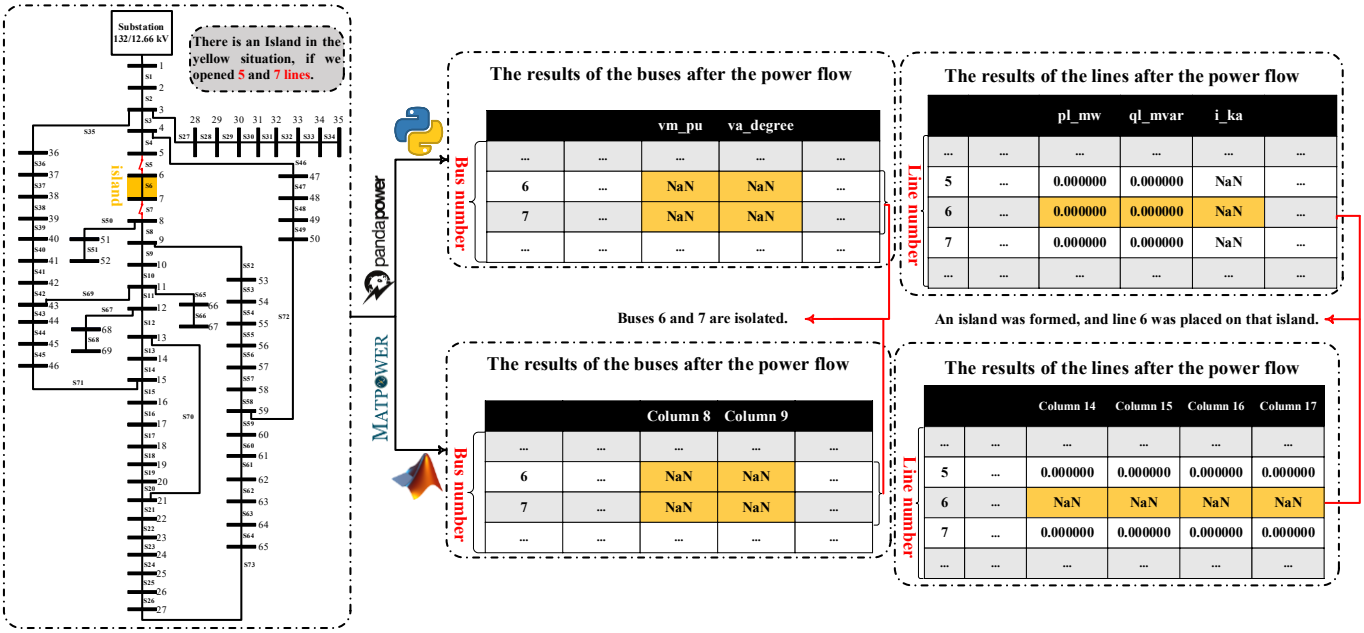


Fig. 2: The example of radiality detection using NAN method in the IEEE 69-bus DN.

3.2. Problem modeling based on deep reinforcement learning

In this modeling, the DN serves as the environment, and the DN operator acts as the DRL agent. The lines in the DN represent the actions that the DN operator is testing to determine the optimal structure for operating the DN. The modeled states include $s_t = [\alpha_t, p_t, q_t]$, where α_t indicates the disconnection or connection of the DN lines, p_t indicates the active power of all buses, and q_t indicates the reactive power of all buses.

Choosing good actions depends on how well each action is evaluated. Therefore, defining a proper evaluation function is essential. This function should reflect the system's main goals and guide the decision-making process toward better solutions. One key goal in DNs is to reduce total power loss. High losses lead to lower efficiency and increased operational

costs. To reflect this, the first part of the reward function is defined as:

$$r_p(s_t, a_t) = -(P_{t+1} - P_{init}) \quad (8)$$

where P_{t+1} represents the total active power loss in the DN at time step $t+1$, and P_{init} is the total active power loss (MW) under the DN's initial structure. In fact, this expression aims to minimize the power loss relative to the DN's initial structure. Another critical aspect of a reliable DN is maintaining voltage levels within acceptable limits. Voltage violations can lead to poor power quality and potential damage to equipment. Thus, the second component of the reward function focuses on voltage and is given by:

$$r_v(s_t, a_t) = \begin{cases} if\ \min(V_{t+1}) < \min(V_{init}) : 0 \\ else: \min(V_{t+1}) \end{cases} \quad (9)$$

where V_{init} is the voltage vector of the DN under the initial structure, and V_{t+1} is the voltage vector (p.u.) at time step $t+1$. This expression is designed to decrease voltage deviation compared to the initial state. By combining these two components, we obtain the final reward function used in our method:

$$r(s_t, a_t) = \lambda_1 * r_p(s_t, a_t) + \lambda_2 * r_v(s_t, a_t) \quad (10)$$

where λ_1 and λ_2 are weighting factors that determine the contribution of power loss and voltage deviation, respectively, in the overall reward function.

The general structure of the proposed DDQN algorithm is illustrated in Fig. 3 and the implementation of the DDQN algorithm is available in Algorithm 1. According to this Algorithm, we have to choose actions until the DN is not radial. For example, in the IEEE 33-bus DN, five actions must be selected, and by choosing each action, one switch will open. The NAN method in this algorithm is used for two purposes:

- We prevent the DN from forming islands in the middle of the episode.
- We prevent the DN from forming a non-radial structure as the final structure.

In the aforementioned process, we open a switch at each time step. As a result, we should calculate vector Q_t each time. If the action with the highest Q-value is valid according to the NAN method, it will be selected. If not, the next action with the second-highest value will be checked, and if it is valid, it will be executed. This process continues until the optimal valid action is selected. In the mentioned algorithm, Q_t shows the value of the available actions, and $\mathcal{A}_t$ represents a list of actions that are sorted according to their values. In summary, the main elements of the algorithm are as follows:

- Environment: the environment is the DN.
- Agent: the agent is the DN operator that tries to open the suitable switches to form a radial structure.
- Reward function: the reward is calculated according to Equations (8-10).
- Action space: the action space at each time step is the switches that operator can open.
- State: the state is $s_t = [\alpha_t, p_t, q_t]$.

Also, in this study, the weights of the main neural network are updated according to Equation (7), and after a certain number of time steps, the weights of the main network are copied to the target network.

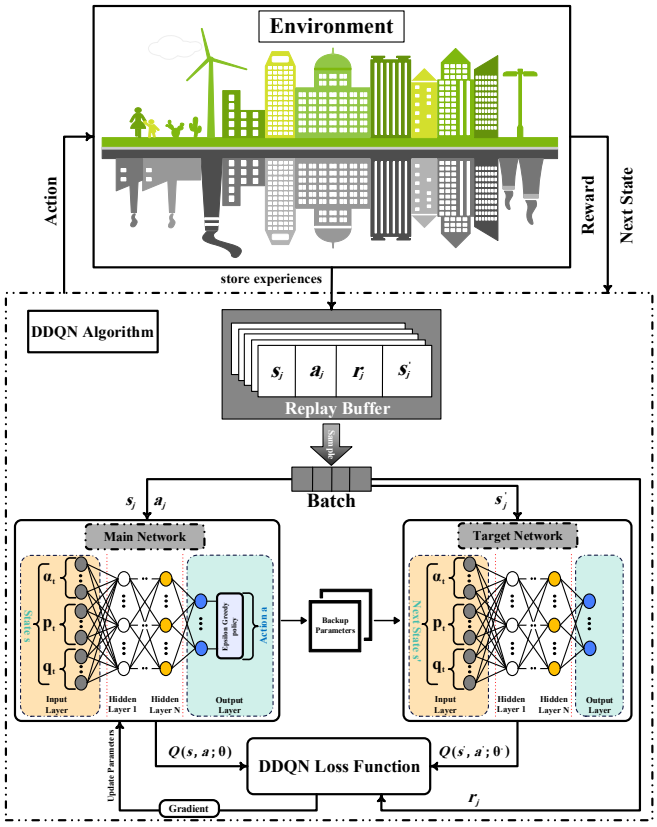


Fig. 3: Structure of the proposed method.

Algorithm 1 DDQN Algorithm using NAN method for DNR

Input: Learning rate η, discount factor γ, target network update period F, replay buffer D, and network weights θ and θ'
for episode $\in \{1,2,\dots,M\}$**:**
  Initialize state s_0
  while distribution network is not radial:
    set flag is_valid = false
    with probability of $1-\varepsilon$:
      Set values of available actions $Q_t = Q(s_t;\theta)$
      Set $\mathcal{A}_t$= argsort (Q_t, descending)
      action_index = 0
      while is_valid = false:
        action_index += 1
        $a_t = \mathcal{A}_t$[action_index]
        if a_t is valid according to NAN method:
          is_valid = true
    otherwise:
      while is_valid = false:
        choose a random a_t
        if a_t is valid according to NAN method:
          is_valid = true
    Perform the action a_t and receive r_t and s'_t
    Store transition (s_t, a_t, r_t, s'_t) in D
    set $s_t = s'_t$
    Sample random minibatch of transitions (s_j, a_j, r_j, s'_j) from D
    Perform a gradient descent step on $L_{DDQN}(\theta)$ for minibatch
    Update ε
  end
  if episode%$F == 0$**:**
    $\theta^- = \theta$
  end
end

4. CASE STUDY

The proposed method has been tested on IEEE 33- and 69-bus DNs.

4.1. IEEE 33-bus distribution network

The IEEE 33-bus DN is studied as the first test DN. This DN has a rated power of 100 MVA and a rated voltage of 12.66 kV. This test DN has 33 buses and 37 lines, of which lines 1 to 32 are equipped with normally closed switches. Also, lines 33 to 37 are fitted with tie switches. This DN's active power losses and minimum voltage in the initial state are 0.20267 MW and 0.913 p.u., respectively. The active and reactive demand in the initial state are 3.715 MW and 2.3 MVAR, respectively. The general structure of this DN is depicted in Fig. 4.

4.2. IEEE 69-bus distribution network

The IEEE 69-bus DN is studied as the second test DN. This DN has a rated power of 100 MVA and a rated voltage of 12.66 kV. This test DN has 69 buses and 73 lines, of which lines 1 to 68 are equipped with normally closed switches. Furthermore, lines 69 to 73 are fitted with tie switches. This DN's active power losses and minimum voltage in the initial state are 0.225 MW and 0.909 p.u., respectively. The active and reactive demand in the initial state are 3.802 MW and 2.695 MVAR, respectively. The general structure of this DN is illustrated in Fig. 5.

The cases designed in this paper are as follows:

- Case 1: DNR without installing any DGs.
- Case 2: DNR with the installation of wind turbines.
- Case 3: DNR with the installation of PV cells.

Information on DGs installed on the 33- and 69-bus DNs is shown in Table 2.

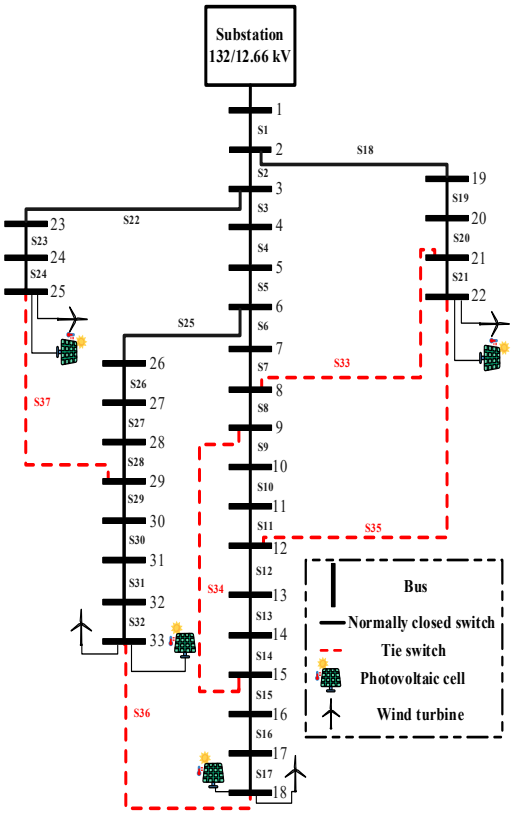


Fig. 4: IEEE 33-bus DN structure.

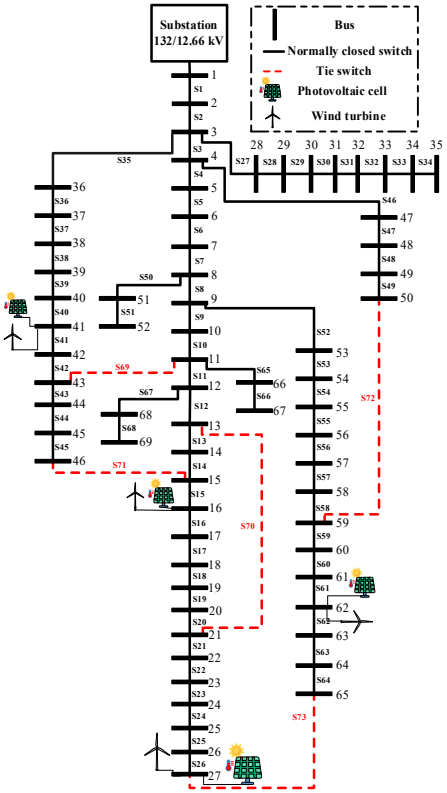


Fig. 5: IEEE 69-bus DN structure.

Table 2: Details of DGs installed on DNs.

DN	DG	IB	C (MVA)	PF (p.u.)
33-bus	PV	18, 22, 25, 33	0.100	1
	Wind	18, 22, 25, 33	0.1175	0.85
69-bus	PV	16, 27, 41, 62	0.100	1
	Wind	16, 27, 41, 62	0.1175	0.85

IB: Installed bus, C: Capacity, PF: Power factor.

5. SIMULATION AND RESULTS

The convergence diagram of the DDQN agent in the 33-bus DN is shown in Fig. 6, where a 50-point moving average window was applied and averaged over 10 runs. The light-colored areas represent the standard error. Additionally, the bus voltage profile and line power losses after reaching the optimal policy are presented in Figs. 7(a) and 7(b), respectively.

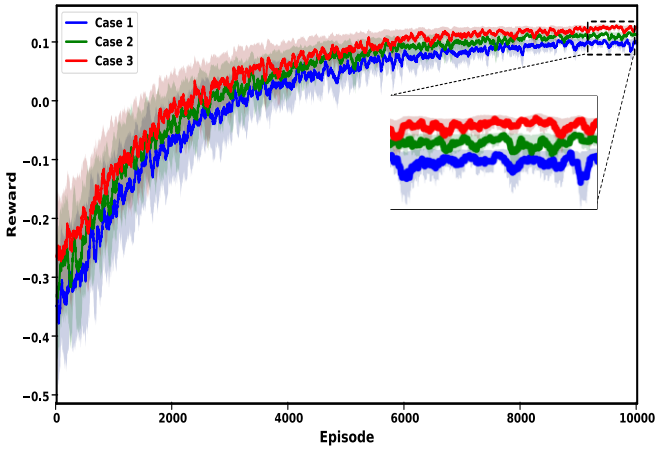


Fig. 6: Convergence diagram in the 33-bus DN.

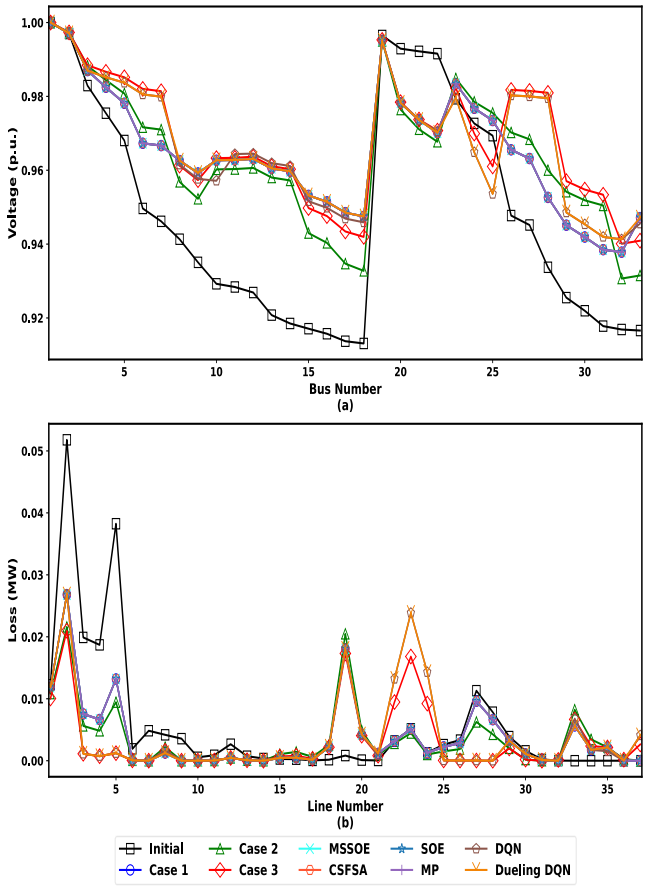


Fig. 7: Results of different algorithms in the 33-bus DN: (a) Bus voltage profile (p.u.), (b) line power losses (MW).

In addition, the numerical results of different algorithms' output in the 33-bus DN are listed in Table 3. Compared to the DQN and Dueling DQN algorithms, the superiority of case 1 in reducing power losses is evident. Compared to the other references listed in this Table, the superiority of the proposed method is demonstrated in the following tables. Additionally, the effect of DGs utilized in cases 2 and 3, in terms of minimizing power losses, is compared to case 1 in this Table, which performed more successfully than case 1.
The convergence diagram after applying the proposed method in the 69-bus DN is displayed in Fig. 8, where a 50-point moving average window was applied and averaged over 10 runs. The light-colored areas denote the standard error. Additionally, the bus voltage profile and line power losses after reaching the optimal policy are presented in Figs. 9(a) and 9(b), respectively.

Table 3: Comparison of numerical results in the 33-bus DN.

Ref	Open switches	PL (MW)	AVD (p.u.)
Initial	33, 34, 35, 36, 37	0.20267	0.0515
Case 1	7, 9, 14, 32, 37	0.13955	0.0347
Case 2	7, 9, 14, 31, 37	0.12568	0.0358
Case 3	7, 9, 14, 31, 28	0.11430	0.0316
MSSOE [22]	7, 9, 14, 32, 37	0.13955	0.0347
CSFSA [17]	7, 9, 14, 32, 37	0.13955	0.0347
SOE [20]	7, 9, 14, 32, 37	0.13955	0.0347
MP [20]	7, 9, 14, 32, 37	0.13955	0.0347
DQN [28]	7, 10, 14, 28, 32	0.14070	0.0329
Dueling DQN [28]	7, 9, 14, 28, 32	0.13997	0.0326

AVD: Average voltage deviation, PL: Power loss.

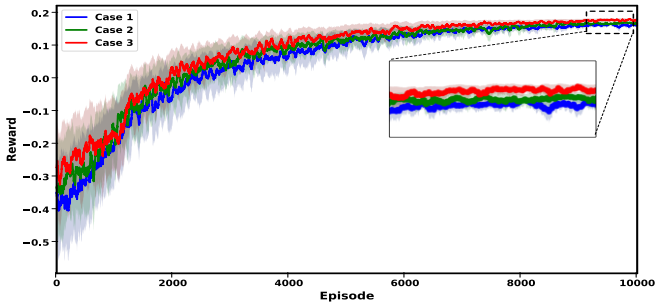


Fig. 8: Convergence diagram in the 69-bus DN.

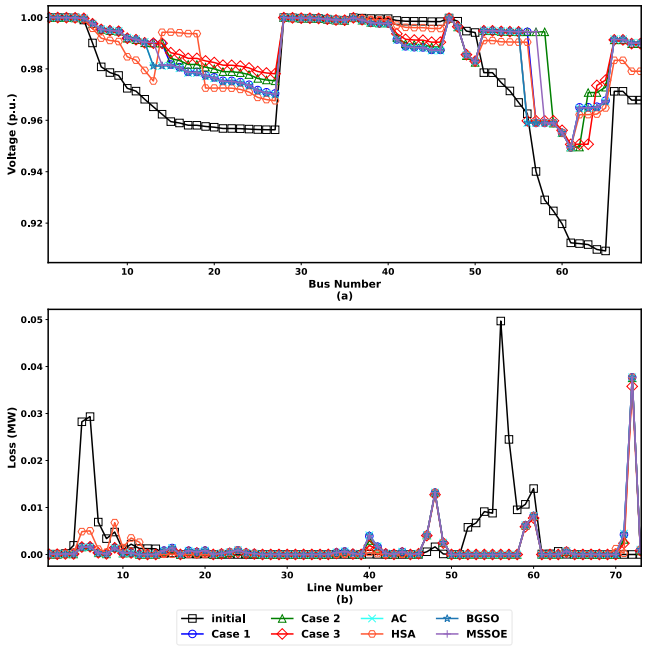


Fig. 9: Results of different algorithms in the 69-bus DN: (a) Bus voltage profile (p.u.), (b) line power losses (MW).

The numerical results obtained from different algorithms on the 69-bus DN are presented in Table 4. Compared to the DN's initial state, all algorithms have successfully minimized power losses and average voltage deviation. Still, the proposed method in this article has contributed a higher percentage to reducing power losses compared to AC, HSA, BGSO, and MSSOE. That demonstrates the optimality of the DRL agent modeling type. In the comparison between the modeled cases, it is observed that all cases have been able to significantly reduce power losses and average voltage deviation compared to the DN's initial state. Furthermore, the effectiveness of wind turbines and PV cells is evident in their greater impact on minimizing power losses and average voltage deviation compared to case 1.

Table 4: Comparison of numerical results in the 69-bus DN.

Ref	Open switches	PL (MW)	AVD (p.u.)
Initial	69, 70 ,71 ,72 ,73	0.22500	0.0266
Case 1	14, 56, 61, 69, 70	0.09860	0.0133
Case 2	14, 58, 62, 69, 70	0.09350	0.0115
Case 3	14, 55, 63, 69, 70	0.08677	0.0125
AC [15]	12, 55, 61, 69, 70	0.09880	0.0141
HSA [23]	13, 18, 56, 61, 69	0.10522	0.0139
BGSO [21]	12, 55, 61, 69, 70	0.09880	0.0141
MSSOE [22]	13, 57, 61, 69, 70	0.09869	0.0129

AVD: Average voltage deviation, PL: Power loss.

The comparison between the computational and training time of the modeling performed, as compared to previous research, is listed in Tables 5 to 8. Compared to the MSSOE, CSFSA, SOE, and MP methods, which had similar results to our proposed method in Table 3, the superiority of our method in terms of computational time is evident. In addition, our method has performed better than the MSSOE method in Table 4. Compared to previous DRL methods, such as those in [27], [28], and [30], the proposed method in this paper demonstrates superiority in terms of computational and training time.

5.1. Daily Analysis

We examined the effects of the proposed algorithm in the 33- and 69-bus DNs with fixed DGs in the previous section. In this section, we will evaluate the effect of uncertainty in DGs and load variety [38]. The output power of PV cells' production is determined by the following equation [24]:

$$P_t^{PV,max} = N_{PV}.FF.V_{g,t}.I_{g,t} \quad (11)$$

$$FF = \frac{V_{MPP}.I_{MPP}}{V_{OC}.I_{SC}} \quad (12)$$

$$V_{g,t} = V_{OC} - N_{PV}T_{cg,t} \quad (13)$$

$$T_{cg,t} = T_{a,t} - SI_t\left(\frac{N_{OT}-20}{0.8}\right) \quad (14)$$

Where N_{PV}, FF, $V_{g,t}$, and $I_{g,t}$ represent the number of PV cells, the filling factor of the PV array, and the output voltage and current of the PV cell, respectively. Furthermore, V_{MPP}, I_{MPP}, V_{OC}, and I_{SC} denote maximum power point (MPP) for voltage and current, and PV arrays' open circuit voltage and short circuit current, respectively. Finally, $T_{cg,t}$, $T_{a,t}$, SI_t, and N_{OT} demonstrate PV arrays' temperature, ambient temperature, solar radiation intensity, and PV arrays' normal operation temperature, respectively. Furthermore, the output production of wind turbines is evaluated by following equation [24]:

$$P_t^{WT,max} = \begin{cases} 0; & V_{w,t} < V_{cut-in} \\ P_{rated}\left(\frac{V_{w,t}-V_{cut-in}}{V_{rated}-V_{cut-in}}\right); & V_{cut-in} \le V_{w,t} < V_{rated} \\ P_{rated}; & V_{rated} \le V_{w,t} < V_{cut-out} \\ 0; & V_{w,t} \ge V_{cut-out} \end{cases} \quad (15)$$

Where V_{cut-in}, $V_{cut-out}$, and $V_{w,t}$ represent the wind turbine's cut-in speed, cut-out speed, and current wind speed, respectively. Additionally, P_{rated} and V_{rated} denote the rated power and rated wind speed of the wind turbine.

The daily analysis cases designed in this article are as follows:

- Case 4: DNR analysis under daily operations with varying load conditions.
- Case 5: DNR analysis under daily operations, incorporating both varying loads and uncertainty in wind turbines.
- Case 6: DNR analysis under daily operations, considering varying loads and uncertainty in PV cells.

After implementing the proposed method under daily analysis using cases 4 to 6, the results are demonstrated for both the 33-bus and 69-bus DNs in Figs. 10 and 11.

The proposed method demonstrates adaptability under various operation conditions, effectively synchronizing with DNs operating in different scenarios to achieve optimal configurations. Analysis shows that it reduces power losses and average voltage deviations compared to the DN's initial state across all hours. Furthermore, considering these figures, DGs significantly reduce power losses and average voltage deviations compared to Case 4 over most hours.

Table 5: Computational time comparison in the 33-bus DN.

Ref	Computational time (s)
Proposed method	0.0880
MSSOE [22]	0.2330
CSFSA [17]	49.59
DQN [30]	0.207
Dueling DQN [30]	0.502
SAC [30]	0.453
PPO [30]	0.206
SOE [20]	1
MP [20]	19
DQN [27]	0.3438
DDQN [27]	0.3125
NoisyNet-DQN [27]	0.4844
DQN [28]	0.21
Dueling DQN [28]	0.21

Table 6: Computational time comparison in the 69-bus DN.

Ref	Computational time (s)
Proposed method	0.0950
MSSOE [22]	0.4325
DQN [27]	0.4063
DDQN [27]	0.4219
NoisyNet-DQN [27]	0.4375

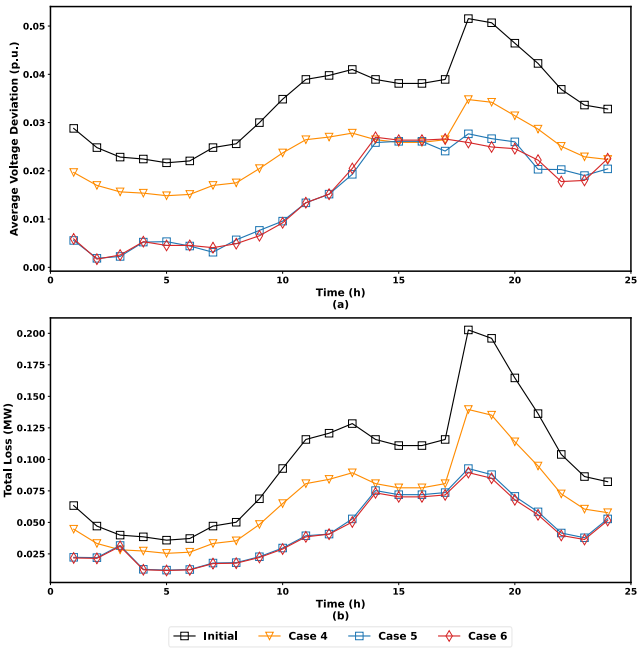


Fig. 10: Daily analysis' results in the 33-bus DN: (a) Average voltage deviation (p.u.), (b) total power losses (MW).

Table 7: Training time comparison between DRL algorithms in the 33-bus DN.

Ref	Training time (s)
Proposed method	960.04
All algorithm in [30]	5727
DQN [27]	1075.8
DDQN [27]	1086.5
NoisyNet-DQN [27]	1201.8
DQN [28]	15145
Dueling DQN [28]	15145

Table 8: Training time comparison between DRL algorithms in the 69-bus DN.

Ref	Training time (s)
Proposed method	1093.48
DQN [27]	1239.0
DDQN [27]	1246.5
NoisyNet-DQN [27]	1276.7

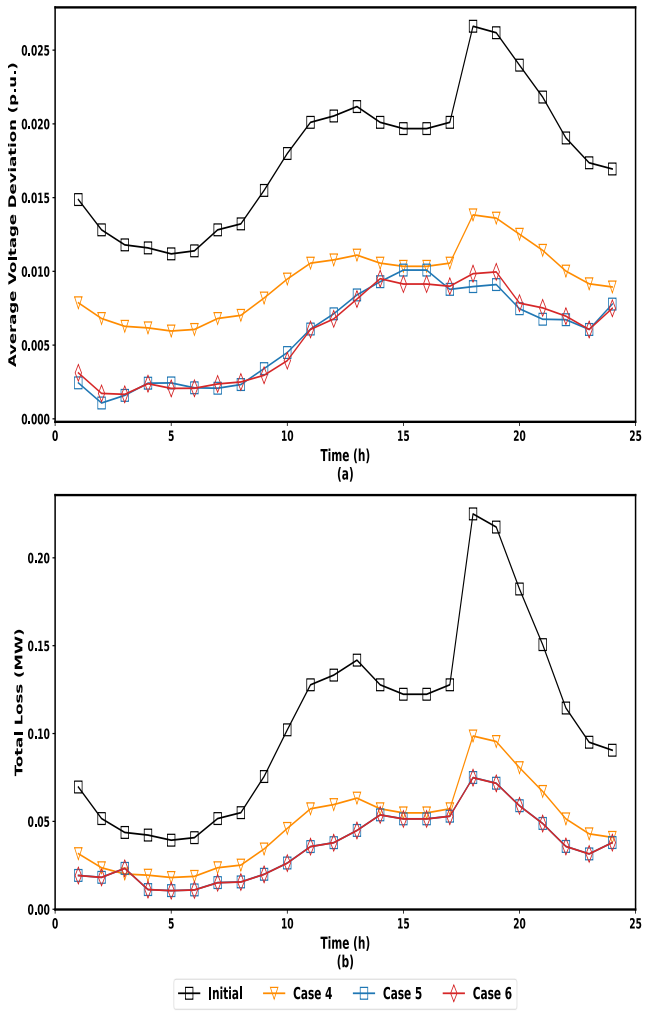


Fig. 11: Daily analysis' results in the 69-bus DN: (a) Average voltage deviation (p.u.), (b) total power losses (MW).

5.2. Sensitivity Analysis

In our experiments, we used a replay buffer with a capacity of 2500, and the batch size was set to 64. For the neural network, we used two hidden layers, each containing 64 units. Moreover, the discount factor γ was set to 0.99 to account for future rewards. The target network was updated every 100 episodes to maintain stability. We used $\lambda_1 = 1$ and $\lambda_2 = 0.05$ to balance the different components of the reward function. To encourage exploration, we applied an epsilon-greedy strategy, where epsilon followed the equation below:

$$\varepsilon = \varepsilon_0 * \varepsilon_{\text{decay}}^{\text{episode}} \quad (16)$$

where ε_0 is the initial epsilon value, and $\varepsilon_{\text{decay}}$ represents the decay factor. These values are assumed to be 1 and 0.9996, respectively. Fig. 12 presents a plot of epsilon values over different episodes.

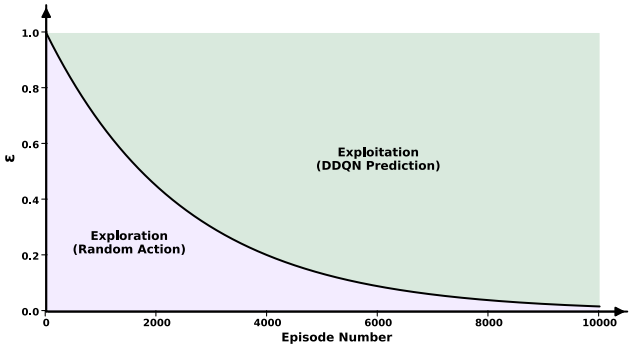


Fig. 12: Epsilon value across episodes, illustrating how exploration decreases over time according to the epsilon-greedy strategy.

One of the most critical parameters in training neural networks is the learning rate, which determines the step size at each iteration as the model navigates toward minimizing the loss function. To assess the impact of different learning rates on the algorithm's performance, we tested four values—0.0005, 0.005, 0.01, and 0.025—on both the 33- and 69-bus DNs. Before analyzing the results, a moving average with a window length of 50 was applied to Figs. 13 and 14 to smooth out short-term fluctuations. Additionally, the data in these figures were averaged across 10 distinct executions to ensure robustness and reduce the impact of random variations. The figures reveal how varying learning rates influence both the network's reward and total power loss. Among the tested values, a learning rate of 0.0005 demonstrated superior training effectiveness. This value resulted in the lowest power losses and the highest reward compared to other learning rates, indicating its capability to optimize network performance while mitigating overfitting or instability. Given these findings, all simulations in this study utilized a learning rate of 0.0005.

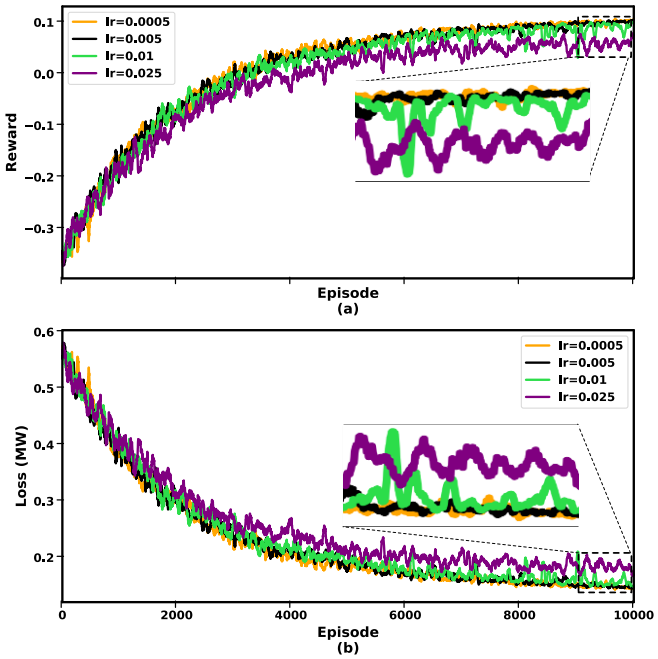


Fig. 13: Agent learning process in the 33-bus DN with different learning rate values: (a) Reward, (b) total loss (MW).

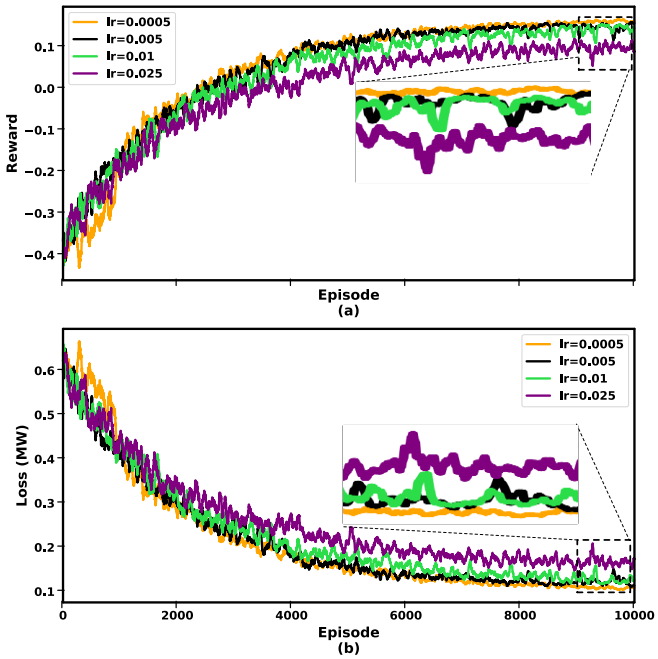


Fig. 14: Agent learning process in the 69-bus DN with different learning rate values: (a) Reward, (b) total loss (MW).

The replay buffer capacity was set to 2500 in this paper. Furthermore, we conducted experiments with four different capacities—500, 1000, 2500, and 5000—to assess the impact of this hyperparameter. To ensure clarity and reduce short-term fluctuations, a moving average with a window length of 50 was applied to both Figs. 15 and 16. Furthermore, the data in these figures were averaged across 10 distinct executions to enhance robustness and minimize the effects of random variations. Our analysis reveals that, in both DNs examined, the replay buffer size has a negligible impact on performance. Consequently, we have chosen 2500 as the replay buffer capacity, which appears to be a suitable selection based on our findings.

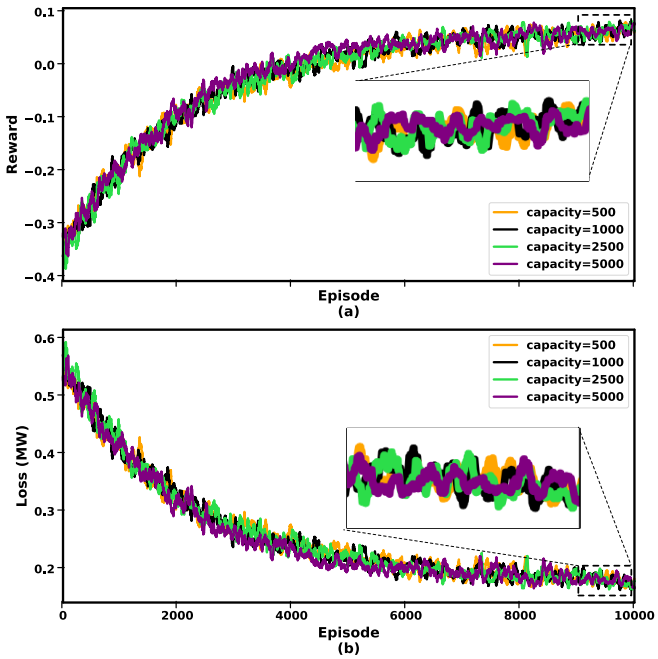


Fig. 15: Agent learning process in the 33-bus DN with different replay buffer capacities: (a) Reward, (b) total loss (MW).

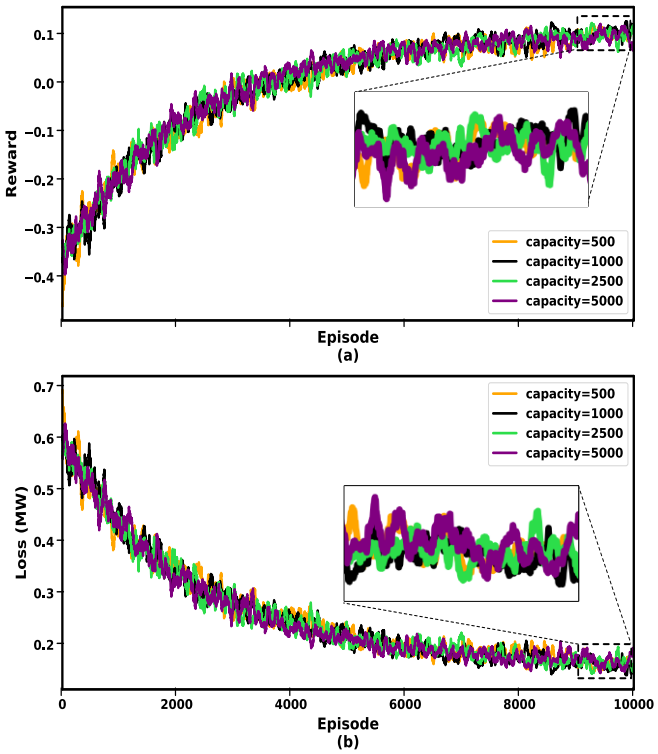


Fig. 16: Agent learning process in the 69-bus DN with different replay buffer capacities: (a) Reward, (b) total loss (MW).

Finally, Figs. 17 and 18 demonstrate the effect of batch size value on the training process in the 33- and 69-bus DNs. As shown, the batch size value does not have such a significant impact on the training process, and the chosen value of 64 in this study seems appropriate.

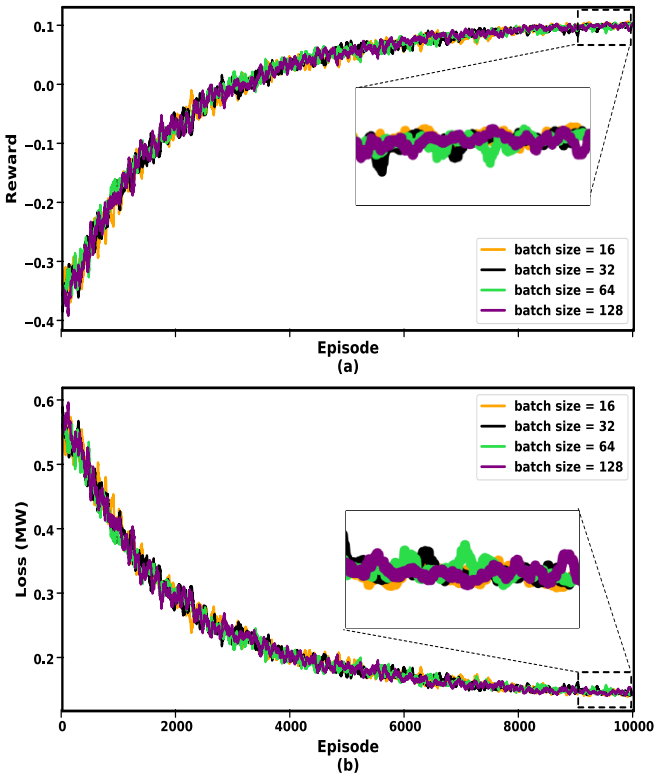


Fig. 17: Agent learning process in the 33-bus DN with different batch size: (a) Reward, (b) total loss (MW).

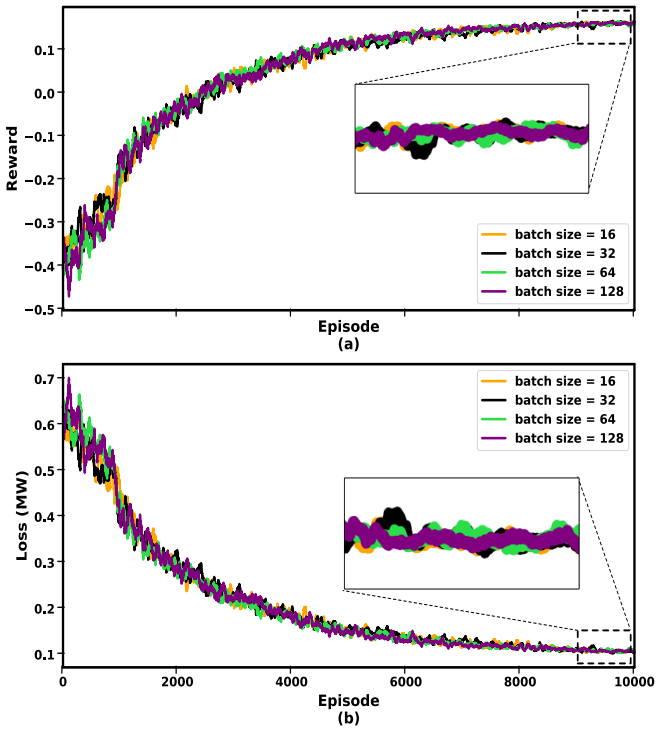


Fig. 18: Agent learning process in the 69-bus DN with different batch size: (a) Reward, (b) total loss (MW).

6. Conclusion

In this paper, a novel method for detecting the radiality of DNs is introduced, which is integrated with the DDQN algorithm to determine the optimal DN structure. The proposed approach is tested on the IEEE 33- and 69-bus DNs to minimize power losses and voltage deviation. The effectiveness of the proposed method in achieving the highest reward by the DRL agent, compared to the references mentioned in this paper's results, is evident. Additionally, the impact of wind turbines and PV cells under fixed and uncertain capacity with a variety of loading on the DNR is analyzed separately using the proposed method. The simulation results showed that even under conditions where DG sources were used, the proposed algorithm was able to minimize power losses and voltage deviation. In future research, we will investigate additional algorithms for radiality detection and explore DRL models across a wider range of simulation scenarios for DNR.

Credit Authorship Contribution Statement

Amirhossein Ghaemipour: Conceptualization, Investigation, Methodology, Software, Validation, Visualization, Roles/Writing - original draft. **Habib Rajabi Mashhadi**: Supervision, Writing - review & editing. **Seyed Hossein Mostafavi**: Conceptualization, Investigation, Methodology, Software, Validation, Visualization, Roles/Writing - original draft,

Declaration Of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The ethical issues; including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy has been completely observed by the authors.

References

[1] A. Rouhipour, E. Mashhour, and M. Saniei, "Optimal Switch Placement in Radial Distribution Networks to Reduce Energy Loss and Improve Network Security in Khuzestan Province Conditions," *J. Appl. Res. Electr. Eng.*, vol. 1, no. 2, pp. 149–158, 2022.

[2] A. Ghafari, M. Saniei, M. Razzaz, and A. Saffarian, "Investigation of the Operation of Active Superconducting Fault Current Limiters in Distribution Networks Connected to Microgrids," *J. Appl. Res. Electr. Eng.*, vol. 2, no. 1, pp. 19–25, 2023.

[3] A. Ghaedi and M. Mahmoudian, "Reliability Evaluation of Power Systems Including Pumped-Storage Generation Units.," *J. Appl. Res. Electr. Eng.*, vol. 3, no. 1, 2024.

[4] A. Ghaedi and M. Mahmoudian, "Reliability Modelling of Central Receiver Power Plants.," *J. Appl. Res. Electr. Eng.*, vol. 3, no. 2, 2024.

[5] A. Karamali, A. Daeichian, and S. Rezaei, "Using Long Short-Term Memory Networks as Virtual Wind Direction Sensors for Improved Wind Farm Turbines Orientation," in *2024 9th International Conference on Technology and Energy Management (ICTEM)*, 2024, pp. 1–5.

[6] A. H. Karamali, A. R. Danesh, M. Kalantar, M. R. Moghadasian, and S. Rezaei, "Enhancing Wind Power Forecasting Through a Hybrid Deep Learning Approach: Long Short-Term Memory Integrated with Attention Mechanism," in *2025 12th Iranian Conference on Renewable Energies and Distributed Generation (ICREDG)*, 2025, pp. 1–6.

[7] S. Ansari, H. Samet, and T. Ghanbari, "Fault Location in Solar Farms," *IEEE Syst. J.*, vol. 15, no. 3, pp. 4003–4012, 2021.

[8] S. A. Miri, M. Rahimi, and A. Bagheri, "An Optimal Conservative Voltage Reduction Strategy in Active Distribution Networks Using Coordinated Operation of RES, OLTC, ESS, and SCB," in *2025 12th Iranian Conference on Renewable Energies and Distributed Generation (ICREDG)*, 2025, pp. 1–6.

[9] J. Z. Zhu, "Optimal reconfiguration of electrical distribution network using the refined genetic algorithm," *Electr. Power Syst. Res.*, vol. 62, no. 1, pp. 37–42, 2002.

[10] A. M. Helmi, R. Carli, M. Dotoli, and H. S. Ramadan, "Efficient and Sustainable Reconfiguration of Distribution Networks via Metaheuristic Optimization," *IEEE Trans. Autom. Sci. Eng.*, vol. 19, no. 1, pp. 82–98, 2022.

[11] T. T. Nguyen and T. T. Nguyen, "An improved cuckoo search algorithm for the problem of electric distribution network reconfiguration," *Appl. Soft Comput.*, vol. 84, p. 105720, 2019.

[12] C. Gerez, L. I. Silva, E. A. Belati, A. J. S. Filho, and E. C. M. Costa, "Distribution Network Reconfiguration Using Selective Firefly Algorithm and a Load Flow Analysis Criterion for Reducing the

Search Space," *IEEE Access*, vol. 7, pp. 67874–67888, 2019.

[13] L. A. Gallego Pareja, J. M. López-Lezama, and O. Gómez Carmona, "A Mixed-Integer Linear Programming Model for the Simultaneous Optimal Distribution Network Reconfiguration and Optimal Placement of Distributed Generation," 2022.

[14] M. F. Abd El-salam, E. Beshr, and M. B. Eteiba, "A new hybrid technique for minimizing power losses in a distribution system by optimal sizing and siting of distributed generators with network reconfiguration," *Energies*, vol. 11, no. 12, p. 3351, 2018.

[15] S. H. Mirhoseini, S. M. Hosseini, M. Ghanbari, and M. Ahmadi, "A new improved adaptive imperialist competitive algorithm to solve the reconfiguration problem of distribution systems for loss reduction and voltage profile improvement," *Int. J. Electr. Power Energy Syst.*, vol. 55, pp. 128–143, 2014.

[16] A. Landeros, S. Koziel, and M. F. Abdel-Fattah, "Distribution network reconfiguration using feasibility-preserving evolutionary optimization," *J. Mod. Power Syst. Clean Energy*, vol. 7, no. 3, pp. 589–598, 2019.

[17] T. Tran The, D. Vo Ngoc, and N. Tran Anh, "Distribution network reconfiguration for power loss reduction and voltage profile improvement using chaotic stochastic fractal search algorithm," *Complexity*, vol. 2020, no. 1, p. 2353901, 2020.

[18] H. Lotfi, M. H. Nikkhah, and M. E. Hajiabadi, "Dynamic Reconfiguration for Energy Management in EV and RES-Based Grids Using IWOA," 2025.

[19] H. Parsadust, M. E. Hajiabadi, and H. Lotfi, "Bi-Level Graph-Based Optimisation for Distribution Network Reconfiguration and Optimal Placement of TCLBS and DC Switches," *IET Gener. Transm. Distrib.*, vol. 19, no. 1, p. e70144, 2025.

[20] J. Zhan, W. Liu, C. Y. Chung, and J. Yang, "Switch Opening and Exchange Method for Stochastic Distribution Network Reconfiguration," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 2995–3007, 2020.

[21] S. Teimourzadeh and K. Zare, "Application of binary group search optimization to distribution network reconfiguration," *Int. J. Electr. Power Energy Syst.*, vol. 62, pp. 461–468, 2014.

[22] V. Vai, S. Suk, R. Lorm, C. Chhlonh, S. Eng, and L. Bun, "Optimal reconfiguration in distribution systems with distributed generations based on modified sequential switch opening and exchange," *Appl. Sci.*, vol. 11, no. 5, p. 2146, 2021.

[23] R. S. Rao, K. Ravindra, K. Satish, and S. V. L. Narasimham, "Power Loss Minimization in Distribution System Using Network Reconfiguration in the Presence of Distributed Generation," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 317–325, 2013.

[24] S. A. Miri, Z. Azimi, and A. Bagheri, "Dynamically Reconfigurable Active Distribution Networks for Maximum Renewables Integration at the Presence of Voltage-Dependent Loads," in *2024 28th International Electrical Power Distribution Conference (EPDC)*, IEEE, 2024, pp. 1–6.

[25] H. Lotfi, A. A. Shojaei, V. Kouhdaragh, and I. Sadegh Amiri, "The impact of feeder reconfiguration on automated distribution network with respect to resilience concept," *SN Appl. Sci.*, vol. 2, no. 9, p. 1590, 2020.

[26] M. Noruzi Azghandi, A. A. Shojaei, S. Toosi, and H. Lotfi, "Optimal reconfiguration of distribution network feeders considering electrical vehicles and distributed generators," *Evol. Intell.*, vol. 16, no. 1, pp. 49–66, 2023.

[27] B. Wang, H. Zhu, H. Xu, Y. Bao, and H. Di, "Distribution Network Reconfiguration Based on NoisyNet Deep Q-Learning Network," *IEEE Access*, vol. 9, pp. 90358–90365, 2021.

[28] N. Gholizadeh and P. Musilek, "A Generalized Deep Reinforcement Learning Model for Distribution Network Reconfiguration with Power Flow-Based Action-Space Sampling," 2024.

[29] A. Ghaemipour, H. R. Mashhadi, and S. H. Mostafavi, "Distribution Network Reconfiguration in the Presence of Distribution Generation Using Deep Reinforcement Learning," in *2025 6th International Conference on Optimizing Electrical Energy Consumption (OEEC)*, 2025, pp. 1–5.

[30] N. Gholizadeh, N. Kazemi, and P. Musilek, "A Comparative Study of Reinforcement Learning Algorithms for Distribution Network Reconfiguration With Deep Q-Learning-Based Action Sampling," *IEEE Access*, vol. 11, pp. 13714–13723, 2023.

[31] S. H. Oh, Y. T. Yoon, and S. W. Kim, "Online reconfiguration scheme of self-sufficient distribution network based on a reinforcement learning approach," *Appl. Energy*, vol. 280, p. 115900, 2020.

[32] S. Malekshah, A. Rasouli, Y. Malekshah, A. Ramezani, and A. Malekshah, "Reliability-driven distribution power network dynamic reconfiguration in presence of distributed generation by the deep reinforcement learning method," *Alexandria Eng. J.*, vol. 61, no. 8, pp. 6541–6556, 2022.

[33] O. B. Kundačina, P. M. Vidović, and M. R. Petković, "Solving dynamic distribution network reconfiguration using deep reinforcement learning," *Electr. Eng.*, vol. 104, no. 3, pp. 1487–1501, 2022.

[34] N. Gholizadeh and P. Musilek, "Explainable reinforcement learning for distribution network reconfiguration," *Energy Reports*, vol. 11, pp. 5703–5715, 2024.

[35] V. Mnih *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[36] H. Van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double q-learning," in *Proceedings of the AAAI conference on artificial*

intelligence, 2016.

[37] L. Thurner *et al.*, "Pandapower—An Open-Source Python Tool for Convenient Modeling, Analysis, and Optimization of Electric Power Systems," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6510–6521, 2018.

[38] A. Bagheri, S. Mobayen, S. Behzadi, and S. A. Miri, "Probabilistic optimal co-planning of distributed series reactor and dynamic line thermal rating for congestion management of highly-renewable-penetrated electric power systems," *Electr. Power Syst. Res.*, vol. 248, p. 111924, 2025.

BIOGRAPHY



Amirhossein Ghaemipour was born in Mashhad, Iran, in 2000. He received his B.Sc. degree in Electrical Power Engineering from Shahid Montazeri University, Mashhad, Iran, in 2021. He is working toward the M.Sc. degree in power systems at Ferdowsi University of Mashhad, Mashhad, Iran. He received the prestigious Best Paper Award at the 6th International Conference on Optimizing Electrical Energy Consumption (OEEC) on 25-26 February 2025. His research interests include the application of AI in electrical engineering, deep reinforcement learning, deep learning, smart grid, power systems reliability and resilience, power system operation and planning, and power markets.



Habib Rajabi Mashhadi was born in Mashhad, Iran, in 1967. He received the B.Sc. and M.Sc. degrees with honors from the Ferdowsi University of Mashhad, Mashhad, Iran, both in electrical engineering, and the Ph.D. degree from the Department of Electrical and Computer Engineering of Tehran University, Tehran, Iran, under joint cooperation of Aachen University of Technology, Germany, in 2002. He is a professor of electrical engineering at Ferdowsi University of Mashhad and is with the Center of Excellence on Soft Computing and Intelligent Information Processing, Ferdowsi University of Mashhad. His research interests are power system operation and planning, power system economics, and biological computation.



Seyed Hossein Mostafavi received his B.Sc. degree in electrical engineering from Shahid Rajaee University, Tehran, Iran, in 2022. He is currently pursuing the M.Sc. degree at Ferdowsi University of Mashhad, Mashhad, Iran. His research interests include power systems, deep learning, deep reinforcement learning, and control theory.

Copyrights
© 2026 by the author(s). Licensee Shahid Chamran University of Ahvaz, Ahvaz, Iran. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution –NonCommercial 4.0 International (CC BY-NC 4.0) License (http://creativecommons.org/licenses/by-nc/4.0/).