Multi-Objective Optimization of Electric Distribution Systems with integrated distributed Generation using Deep Reinforcement Learning

Trieu Ngoc Ton
Faculty of Electrical and Electronics, Thu Duc College of Technology, Ho Chi Minh City, Vietnam
tonngoctrieu@gmail.com (corresponding author)

Loc Huu Pham
Faculty of Electrical and Electronics, Thu Duc College of Technology, Ho Chi Minh City, Vietnam
phamhuuloc@tdc.edu.vn

Phong Minh Le
Faculty of Electrical and Electronics, Thu Duc College of Technology, Ho Chi Minh City, Vietnam
phongleminh@tdc.edu.vn

Tan Minh Le
Faculty of Electrical and Electronics, Thu Duc College of Technology, Ho Chi Minh City, Vietnam
leminhtan@tdc.edu.vn

Received: 26 January 2025 | Revised: 28 February 2025 | Accepted: 6 March 2025

ABSTRACT

This paper proposes a method for optimizing the placement and capacity of Distributed Generators (DGs) in distribution systems based on Deep Reinforcement Learning (DRL). The objective of the method is to minimize power losses, investment costs, voltage deviations, and CO_2 emissions while ensuring strict compliance with system operating constraints. The proposed approach leverages the robust capabilities of DRL to handle nonlinear and complex-constrained problems, making it highly adaptable to various operational scenarios. Experimental results on standard distribution systems demonstrate that the proposed method outperforms traditional algorithms, significantly improving operational efficiency and enhancing the integration of renewable energy sources. This contributes to the development of smart grid systems and promotes sustainable energy solutions.

Keywords-distributed generator; reinforcement learning; multi-objective optimization; carbon emission reduction; loss minimization

I. INTRODUCTION

In the context of continuously increasing global energy demand and growing pressure to mitigate environmental impacts, the integration of Distributed Generators (DGs) into Electric Distribution Systems (EDSs) has received significant attention from researchers and the power industry [1]. DGs, including Renewable Energy Sources (RESs) such as solar and wind, as well as small-scale fossil fuel-based sources, provide many critical benefits such as reducing power losses, improving voltage profiles, enhancing reliability, and reducing the load on transmission EDSs. However, integrating DGs into EDSs is not an easy task, especially when technical, economic, and environmental factors have to be optimized simultaneously [2]. One of the main challenges of this problem is to determine the optimal location and capacity of DGs to balance objectives such as minimizing power losses, reducing investment and operating costs, maintaining voltage quality, and reducing CO_2 emissions. Traditional optimization methods, including mathematical programming techniques [3] and metaheuristic algorithms such as the Coyote Optimization Algorithm (COA) [4], Salp Swarm Algorithm (SSA) [5], Genetic Algorithm (GA) [6], and Particle Swarm Optimization (PSO) [7], have been successfully applied in many studies [8]. However, these methods often face difficulties in handling nonlinear and complex search spaces or in addressing the stringent operational constraints in EDSs. Reinforcement Learning (RL), particularly Deep Reinforcement Learning (DRL) [9], has emerged as a powerful tool for solving complex optimization problems. With its ability to learn from experience and make decisions based on dynamic environments, DRL has proven effective in tackling nonlinear and multi-objective optimization problems [10]. Unlike traditional metaheuristic algorithms, DRL does not require an explicit formulation of the objective function, but instead learns directly from the data and the environment. This is particularly advantageous in the context of modern EDSs, where uncertainties such as load fluctuations, renewable sources, and complex operational constraints pose significant challenges [10].

This paper proposes an optimization method for DG integration based on DRL, specifically utilizing the Deep Q-Network (DQN) model. The method is designed to simultaneously minimize power losses, investment costs, voltage deviation, and CO_2 emissions while ensuring operational constraints such as voltage and current constraints in distribution systems. The proposed approach is validated on 33-bus and 69-bus EDSs, with results demonstrating superior performance compared to traditional algorithms such as COA [11] and Nondominated Sorting Genetic Algorithm-II (NSGA-II) [12]. The findings not only highlight the potential of reinforcement learning for optimizing the integration of DGs into EDSs but also introduce a novel approach to addressing nonlinear optimization problems in the power industry.

II. PROBLEM DESCRIPTION

The optimization problem for integrating DGs into EDSs aims to improve operational efficiency and minimize environmental impact. To achieve this, it is necessary to determine the optimal location and capacity of DGs so that technical, economic, and environmental objectives are balanced while adhering to the operational constraints of the system. The optimization problem for DG integration is formulated as a multi-objective function that includes the following components:

- Minimizing active power losses (P_{loss}): Reducing losses in the EDS to enhance operational efficiency, as in (1).

$$P_{\text{loss}} = \sum_{k\in\text{branches}} R_k \cdot \frac{I_k^2}{V_k^2} \qquad (1)$$

- Minimizing installation and operating costs of DGs (C_{DG}): Optimizing the investment and operating costs of DGs, as in (2).

$$C_{\text{DG}} = \sum_{i=1}^{N_{\text{DG}}} P_{\text{DG},i} \cdot \text{cost}_i \qquad (2)$$

- Minimizing voltage deviation ($V_{\text{deviation}}$): Ensuring that the voltage profile remains within permissible limits while minimizing the voltage deviation at the nodes, as in (3).

$$V_{\text{deviation}} = \frac{1}{N_{\text{bus}}} \sum_{i=1}^{N_{\text{bus}}} | V_i - V_{\text{base}} | \qquad (3)$$

- Minimizing CO_2 emissions (E_{emission}): Reducing emissions from DGs using fossil fuels, as in (4).

$$E_{\text{emission}} = \sum_{i=1}^{N_{\text{DG}}} P_{\text{DG},i} \cdot \text{EF}_i \qquad (4)$$

where:

R_k: Resistance of branch k.

I_k: Current through branch k.

V_k: Voltage at branch k.

V_i: Voltage at node i.

V_{base}: Base voltage (1.0 p.u.).

$P_{\text{DG},i}$: Power output of the i-th DG.

cost_i: Investment cost per MW of the i-th DG.

EF_i: Emission factor of the i-th DG (kg CO_2/MW).

The composite objective function is constructed from the above components with corresponding weights to balance the objectives:

$$F = w_1 \cdot \frac{P_{\text{loss}}}{P_{\text{base}}} + w_2 \cdot \frac{C_{\text{DG}}}{C_{\text{base}}} + w_3 \cdot \frac{V_{\text{deviation}}}{V_{\text{base}}} + w_4 \cdot \frac{E_{\text{emission}}}{E_{\text{base}}} \qquad (5)$$

where w_1, w_2, w_3, w_4 are the weights for the respective objectives, and P_{base}, C_{base}, V_{base}, E_{base} are the normalized values for power loss, cost, voltage deviation, and CO_2 emissions, respectively. The problem must satisfy the following operational constraints:

1. Voltage at nodes: $V_{\min} \le V_i \le V_{\max}$, $\forall i \in$ nodes
2. Current through branches: $I_k \le I_{\max}$, $\forall k \in$ branches
3. Power output of DGs: $P_{\text{DG}}^{\min} \le P_{\text{DG}} \le P_{\text{DG}}^{\max}$
4. CO_2 emissions limit: $E_{\text{emission}} \le E_{\max}$

III. PROPOSED METHOD

The DRL method offers significant advantages in solving the optimization problem of integrating DGs. Firstly, DRL has the ability to learn autonomously from environmental data without requiring an explicit formulation of the objective function, thereby reducing the complexity of programming and modeling. This is particularly beneficial when dealing with problems involving nonlinear search spaces or stringent constraints. Secondly, DRL adapts well to dynamic conditions, such as load variations or fluctuating power output from renewable sources, thanks to its ability to continuously learn and make optimal real-time decisions. Additionally, this method effectively handles multi-objective problems by flexibly integrating the objectives into the reward model, achieving a balance between minimizing power losses, costs, voltage deviations, and CO_2 emissions. Lastly, experiments show that DRL not only achieves higher performance than traditional algorithms, but also ensures stability and better convergence in complex operating environments. With these outstanding advantages, DRL becomes a promising tool for integrating sustainable energy into modern power systems.

To solve the optimization problem of DG placement and sizing in EDS, this paper proposes a DRL-based method. This DRL method leverages the power of reinforcement learning, in particular the DQN model, to learn how to optimize in nonlinear and constrained environments. DRL is a combination of RL and Deep Neural Networks (DNN). The DRL model is capable of learning optimal actions through interaction with the environment, using:

- State (S): Describes the current state of the system (e.g., node voltages, power losses, DG power outputs).
- Action (A): Represents the possible actions, such as selecting the location and capacity of DGs.
- Reward (R): The feedback value from the environment that helps the model learn optimal actions (e.g., reducing power losses, costs, and emissions).
- Policy (π): Determines the best action for each state.

The proposed method consists of the following steps:

1. Environment modeling: The distribution system is modeled as an environment with states (S), actions (A), and rewards (R):
   - State: $S = [V_1, V_2, \ldots, V_n, P_{\text{loss}}, E_{\text{emission}}]$, where V_1, V_2,..., V_n are the node voltages, P_{loss} is the power loss, and E_{emission} is the CO_2 emission.
   - Action: $A = \{\text{locations of DGs, capacities of DGs}\}$.
   - Reward: $R = -F$, where F is the value of the composite objective function.
2. Applying DQN: DQN uses a DNN to approximate the Q-value function, as in (6).

   $$Q(S, A) = R + \gamma \cdot \max_{A'} Q(S', A') \quad (6)$$

   where S' is the new state after performing action A, A' is the next action, and γ is the discount factor.
3. Optimization process:
   - Step 1: Initialize the initial state (S_0) and select a random action.
   - Step 2: Execute the action, update the state (S'), and calculate the reward (R).
   - Step 3: Update the DQN network using experience replay.
   - Step 4: Repeat until convergence.
4. Result evaluation:
   - After training, the DQN model is used to determine the optimal locations and capacities of DGs.
   - The results are compared with traditional algorithms such as COA, SSA, and NSGA-II.

IV. TEST RESULTS

To evaluate the effectiveness of DRL in optimizing the location and capacity of DGs, experiments were conducted on two distribution systems: 33-bus and 69-bus. These are common systems in power grid optimization research and represent two different levels of complexity in terms of scale and structure. The 33-bus system has a simpler configuration with fewer nodes and connections, while the 69-bus system has a higher connection density and greater complexity, posing more challenges in balancing optimization objectives. The experiments were conducted using MATLAB software for power flow simulation and optimization environment development. The hardware used included a computer with the following configuration: Intel Core i7 processor (3.2 GHz) and 16 GB of RAM. Table I provides the assumed parameters used to optimize the integration of DGs into the EDS.

TABLE I. ASSUMED PARAMETERS FOR THE OPTIMIZATION PROBLEM

Parameter	33-bus	69-bus
P_{load}	3.72 MW	3.8 MW
Q_{load}	2.30 MVAR	2.69 MVAR
Number of DGs	3	3
P_{DG}	0.1-2.0 MW	0.1-2.0 MW
Renewable DG	80%	80%
Fossil DG	20%	20%
Cost of renewable DG	80,000 USD/MW	80,000 USD/MW
Cost of fossil DG	100,000 USD/MW	100,000 USD/MW
CO_2 emissions (renewable DG)	0 kg/MW	0 kg/MW
CO_2 emissions (fossil DG)	500 kg/MW	500 kg/MW
V_{base}	1.0 p.u.	1.0 p.u.
V_{min}, V_{max}	0.95, 1.05 p.u.	0.95, 1.05 p.u.
P_{loss}	202.69 kW	224.89 kW
$w_1, w_2, , w_3, w_4$	0.4, 0.3, 0.2, 0.1	0.4, 0.3, 0.2, 0.1
P_{base}	100 kW	100 kW
C_{base}	100,000 USD	100,000 USD
E_{base}	500 kg	500 kg

A. 33-Bus Distribution System

Figure 1 illustrates the single-line diagram of the 33-bus system with 37 branches, with node and line data referenced from [13, 14]. Table II shows the results for the 33-bus EDS, demonstrating that the DRL method outperforms COA, SSA, and NSGA-II combined with Differential Evolution (DE) in optimizing the location and capacity of DGs. DRL achieves the lowest power loss (65.8 kW) compared to COA (71.46 kW) and SSA (72.45 kW) and significantly reduces the DG cost to the lowest level (230,000 USD) due to its ability to effectively optimize both the location and capacity of DGs. The voltage deviation of DRL (0.008 p.u.) is also the smallest, ensuring better voltage quality than other methods. Although the CO_2 emission level of DRL (293 kg) is comparable to other methods, this emission is minimal as most DGs are from RESs.

The execution time of DRL (50 s) is higher compared to COA (33.16 s) and SSA (30.12 s). However, the superior performance of DRL in terms of power loss reduction, cost minimization, and voltage quality fully compensates for the longer computation time. This result confirms that DRL is a promising method, especially suitable for modern power systems with multi-objective optimization requirements and high precision demands. In the future, improving the DRL algorithm to reduce the execution time could enhance its feasibility for practical applications.

TABLE II. OPTIMIZATION RESULTS FOR THE 33-BUS DISTRIBUTION SYSTEM

Method	Node (P_{DG}-MW)	F	P_{loss} (kW)	C_{DG} (USD)	$V_{deviation}$ (p.u.)	CO_2 (kg)	Time (s)
COA	14 (1.07) 24 (0.75) 30 (1.1)	0.256	71.46	240,000	0.010	293	33.16
SSA	14 (1.05) 24 (0.78) 30 (1.12)	0.261	72.45	245,000	0.015	293	30.12
NSGA-II + DE	14 (1.1) 24 (0.72) 30 (1.15)	0.230	68.30	235,000	0.009	293	40.25
DRL	14 (1.08) 24 (0.74) 30 (1.11)	0.220	65.80	230,000	0.008	293	50.00

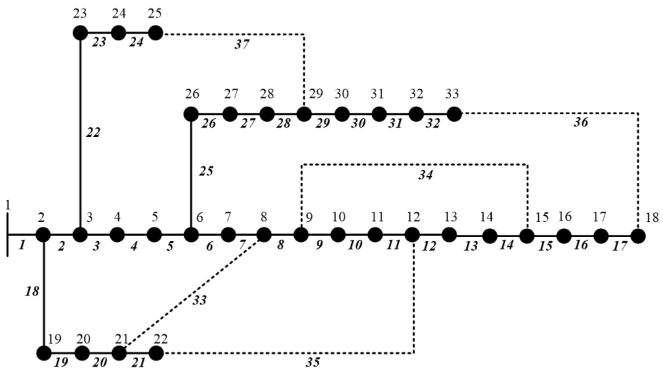

Fig. 1. Single-line diagram of the 33-bus distribution system.

Figure 2 compares the optimization methods (COA, SSA, NSGA-II + DE, DRL) based on key criteria: power loss, cost, voltage deviation, and execution time. The results show that DRL performs best in terms of power loss (63.50 kW), cost (245,000 USD), and voltage deviation (0.007 p.u.), despite having the longest execution time (150 s). This highlights DRL's superior advantages in multi-objective optimization, making it particularly suitable for systems requiring high precision.

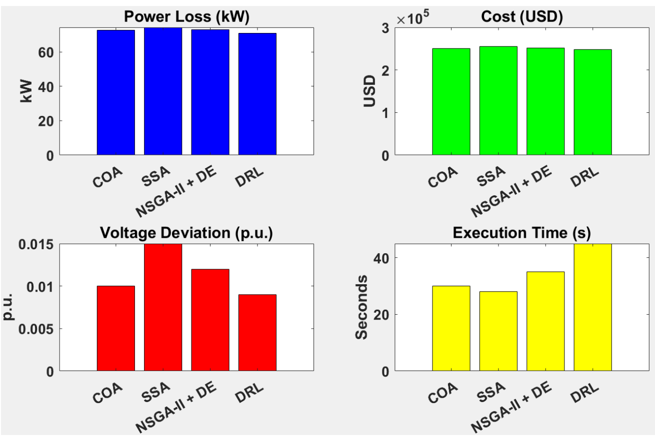

Fig. 2. Comparison of algorithms for the 33-bus system.

B. 69-Bus Distribution System

The 69-bus distribution system, shown in Figure 3, consists of 73 branches and is referenced from [13, 14]. The results for the 69-bus EDS, as shown in Table III, indicate that the DRL method outperforms traditional approaches such as COA, SSA, and NSGA-II combined with DE in optimizing the location and capacity of DGs. Specifically, DRL achieves the lowest power loss (63.5 kW), significantly lower than COA (69.39 kW) and SSA (70.45 kW). Additionally, DRL ensures the lowest DG cost (245,000 USD) by effectively optimizing both the capacity and location of DGs, and it minimizes the voltage deviation to the smallest value (0.007 p.u.) compared to other methods. However, the CO_2 emission level of DRL (320 kg) is comparable to NSGA-II with DE, but slightly higher than COA (293 kg) and SSA (300 kg). The main reason for this is the use of 20% fossil fuel-based DGs, which remains a very small portion when compared to the total system capacity.

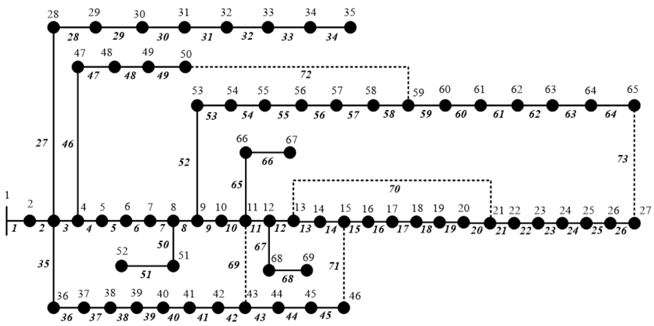

Fig. 3. Single-line diagram of the 69-bus distribution system.

Compared to other methods, the technical and economic advantages of this approach are very significant. This result confirms that DRL is a powerful tool with great potential for application in modern power grid optimization, especially in the context of increasingly complex and sustainability-oriented systems. To reduce the execution time, future research could focus on optimizing the algorithm or utilizing more powerful computing hardware.

Figure 4 compares the performance of the algorithms (COA, SSA, NSGA-II with DE, DRL) on the 69-bus system based on criteria such as power loss, cost, voltage deviation, and execution time. DRL outperforms others with the lowest power loss (63.50 kW), lowest cost (245,000 USD), and lowest voltage deviation (0.007 p.u.), demonstrating its superior optimization capabilities. However, DRL's execution time (150 s) is significantly higher than other methods, reflecting the higher computational cost to achieve optimal technical and economic performance.

TABLE III. OPTIMIZATION RESULTS FOR THE 69-BUS DISTRIBUTION SYSTEM

Method	**Node (P_{DG}-MW)**	***F***	**P_{loss} (kW)**	**C_{DG} (USD)**	**$V_{deviation}$ (p.u.)**	**CO_2 (kg)**	**Time (s)**
COA	12 (1.15) 33 (0.85) 55 (1.2)	0.265	69.39	255,000	0.01	293	105.16
SSA	12 (1.12) 33 (0.88) 55 (1.18)	0.270	70.45	260,000	0.015	300	102.89
NSGA-II + DE	12 (1.18) 33 (0.8) 55 (1.25)	0.240	66.80	250,000	0.009	320	120.00
DRL	12 (1.14) 33 (0.84) 55 (1.22)	0.230	63.50	245,000	0.007	320	150.00

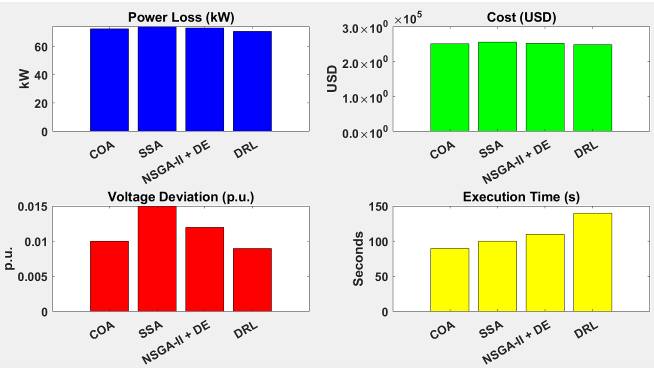

Fig. 4. Comparison of algorithms for the 69-bus system.

In both Table II and Table III, the execution time of DRL is longer than other methods due to its complex training process, which requires multiple iterations and the use of DNNs to approximate the Q-value function. In particular, the 69-bus system's complex structure increases the convergence time. However, DRL's superior technical and economic performance fully compensates for this computational cost.

DRL excels in each criterion thanks to a flexible reward model that allows simultaneous integration of objectives such as reducing power losses, optimizing costs, and improving voltage quality. Specifically, the reward model is designed to prioritize actions that reduce losses and costs, resulting in more balanced results across objectives compared to traditional algorithms that typically address each objective separately. However, DRL faces the challenge of high computational cost due to the DNN training process and optimization over a large search space. To overcome this, the algorithm can be improved by applying techniques to reduce training iterations, accelerate convergence, or utilize more powerful hardware. These enhancements not only reduce execution time but also increase the practical value of DRL, particularly in increasingly complex power systems.

V. CONCLUSION

This paper has demonstrated the effectiveness of Deep Reinforcement Learning (DRL) in optimizing the location and capacity of Distributed Generators (DGs) in Electric Distribution Systems (EDSs). Experimental results on the 33-bus and 69-bus systems show that DRL outperforms traditional algorithms such as Coyote Optimization Algorithm (COA), Salp Swarm Algorithm (SSA), and Nondominated Sorting Genetic Algorithm-II (NSGA-II) in reducing power losses, optimizing investment costs, and ensuring voltage quality. Compared to other optimization methods, DRL does not require a predefined objective function, but learns autonomously through interaction with the environment. Although the training time of DRL may be longer, it provides more comprehensive and flexible optimal solutions. The results demonstrate the great potential of DRL in optimizing and operating modern power systems. Future research will focus on accelerating the training speed and expanding DRL applications to more complex power systems.

ACKNOWLEDGMENT

We would like to express our deep gratitude to Thu Duc College of Technology (Ho Chi Minh City, Vietnam) for their enthusiastic support and for providing favorable conditions during the implementation of this research. The college's support, not only in terms of facilities but also in terms of expertise, played a crucial role in ensuring the quality and progress of the study. Additionally, we sincerely thank our colleagues and sponsoring organizations for their meaningful contributions, from providing resources to sharing in-depth knowledge. These contributions have significantly contributed to the success of the research and paved the way for further developments in future studies.

REFERENCES

[1] R. P. Payasi, A. K. Singh, and D. Singh, "Review of distributed generation planning: objectives, constraints, and algorithms," *International Journal of Engineering, Science and Technology*, vol. 3, no. 3, pp. 133-153, Jul. 2011, https://doi.org/10.4314/ijest.v3i3.68430.

[2] W. L. Theo, J. S. Lim, W. S. Ho, H. Hashim, and C. T. Lee, "Review of distributed generation (DG) system planning and optimisation

techniques: Comparison of numerical and mathematical modelling methods," *Renewable and Sustainable Energy Reviews*, vol. 67, pp. 531–573, Jan. 2017, https://doi.org/10.1016/j.rser.2016.09.063.

[3] G. S. Naik, D. K. Khatod, and M. P. Sharma, "Distributed Generation Impact on Distribution Networks: A Review," *International Journal of Electronics and Electical Engineering*, vol. 2, no. 3, Jan. 2014, Art. no. 10, https://doi.org/10.47893/IJEEE.2014.1096.

[4] J. Pierezan and L. Dos Santos Coelho, "Coyote Optimization Algorithm: A New Metaheuristic for Global Optimization Problems," in *2018 IEEE Congress on Evolutionary Computation*, Rio de Janeiro, Brazil, 2018, pp. 1–8, https://doi.org/10.1109/CEC.2018.8477769.

[5] S. S. Kola and Jayabarathi, "Optimal Allocation of Renewable Distributed Generation and Capacitor Banks in Distribution Systems using Salp Swarm Algorithm," *International Journal of Renewable Energy Research (IJRER)*, vol. 9, no. 1, pp. 96–107, Mar. 2019.

[6] D. Singh, D. Singh, and K. S. Verma, "GA based energy loss minimization approach for optimal sizing & placement of distributed generation," *International Journal of Knowledge-based and Intelligent Engineering Systems*, vol. 12, no. 2, pp. 147–156, May 2008, https://doi.org/10.3233/KES-2008-12206.

[7] F. Tooryan, E. R. Collins, A. Ahmadi, and S. S. Rangarajan, "Distributed generators optimal sizing and placement in a microgrid using PSO," in *2017 IEEE 6th International Conference on Renewable Energy Research and Applications*, San Diego, CA, USA, 2017, pp. 614–619, https://doi.org/10.1109/ICRERA.2017.8191133.

[8] P. Paliwal, N. P. Patidar, and R. K. Nema, "Planning of grid integrated distributed generators: A review of technology, objectives and techniques," *Renewable and Sustainable Energy Reviews*, vol. 40, pp. 557–570, Dec. 2014, https://doi.org/10.1016/j.rser.2014.07.200.

[9] T. T. Nguyen, N. D. Nguyen, P. Vamplew, S. Nahavandi, R. Dazeley, and C. P. Lim, "A multi-objective deep reinforcement learning framework," *Engineering Applications of Artificial Intelligence*, vol. 96, Nov. 2020, Art. no. 103915, https://doi.org/10.1016/j.engappai.2020.103915.

[10] V.-H. Bui, A. Hussain, and H.-M. Kim, "Double Deep Q -Learning-Based Distributed Operation of Battery Energy Storage System Considering Uncertainties," *IEEE Transactions on Smart Grid*, vol. 11, no. 1, pp. 457–469, Jan. 2020, https://doi.org/10.1109/TSG.2019.2924025.

[11] Z. Yuan, W. Wang, H. Wang, and A. Yildizbasi, "A new methodology for optimal location and sizing of battery energy storage system in distribution networks for loss reduction," *Journal of Energy Storage*, vol. 29, Jun. 2020, Art. no. 101368, https://doi.org/10.1016/j.est.2020.101368.

[12] W. Sheng, K.-Y. Liu, Y. Liu, X. Meng, and Y. Li, "Optimal Placement and Sizing of Distributed Generation via an Improved Nondominated Sorting Genetic Algorithm II," *IEEE Transactions on Power Delivery*, vol. 30, no. 2, pp. 569–578, Apr. 2015, https://doi.org/10.1109/TPWRD.2014.2325938.

[13] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Transactions on Power Delivery*, vol. 4, no. 2, pp. 1401–1407, Apr. 1989, https://doi.org/10.1109/61.25627.

[14] T. N. Ton, T. T. Nguyen, A. V. Truong, and T. P. Vu, "Optimal Location and Size of Distributed Generators in an Electric Distribution System based on a Novel Metaheuristic Algorithm," *Engineering, Technology & Applied Science Research*, vol. 10, no. 1, pp. 5325–5329, Feb. 2020, https://doi.org/10.48084/etasr.3372.