JOURNAL OF ADVANCED MULTIDISCIPLINARY STUDIES (JAMS)
JAMS |Volume I Issue II September 2026| DOI: 10.68050/JAMS

Resilient Self-Healing of Active Distribution Networks Using Physics-Informed Graph Reinforcement Learning

MinhPhong Le

Faculty of Electrical and Electronics Engineering, Thu Duc College of Technology, Ho Chi Minh City, Vietnam

DOI: https://doi.org/10.68050/JAMS.2026.410

Received: 06 Sept 2026; Accepted: 07 Sept 2026; Published: 12 September 2026

ABSTRACT

The rapid integration of renewable distributed generation (DG) into active distribution networks (ADNs) introduces substantial operational challenges arising from renewable intermittency, load variability, changing network topology, and feeder contingencies. These conditions require control strategies that can simultaneously maintain secure operation and rapidly restore service following disturbances. This paper proposes a Physics-Informed Graph Reinforcement Learning (PGRL) framework for resilient operation and self-healing of renewable-integrated ADNs. The proposed framework represents the distribution network as an electrical graph to capture bus connectivity and topology changes, while physics-informed decision guidance links the learning process to network operating constraints. Coordinated DG reactive-power control, battery energy storage system (BESS) operation, and feeder reconfiguration are integrated into a unified control strategy for both normal and post-fault operation. The framework is evaluated on the IEEE 33-bus ADN under variable photovoltaic generation, load variations, and feeder contingencies. Compared with the benchmark approaches, PGRL achieves a minimum voltage of 0.980 p.u., a maximum voltage deviation of 0.020, and reduces network power loss to 115.7 kW, corresponding to a 46.6% reduction. Under a feeder outage between buses 9 and 10, PGRL restores the affected load within 11 min, limits energy not supplied (ENS) to 65 kWh, and achieves a resilience index of 0.97. Furthermore, stable learning convergence is reached at approximately 540 episodes. The results demonstrate that integrating topology-aware graph learning with physics-guided control and coordinated DG-BESS operation can substantially improve the operational efficiency, adaptability, and recovery capability of renewable-rich ADNs.

Keywords: Active distribution networks, physics-informed graph reinforcement learning, resilient self-healing, renewable distributed generation, battery energy storage systems, feeder reconfiguration, voltage regulation; network resilience.

INTRODUCTION

The transition toward renewable-dominated distribution systems is fundamentally changing the way active distribution networks (ADNs) are operated and controlled. The increasing deployment of distributed generation (DG), inverter-based renewable resources, and battery energy storage systems (BESSs) provides new opportunities for improving energy efficiency and operational flexibility. At the same time, these resources introduce more variable power injections, bidirectional power flows, and stronger interactions between distributed devices and network operating conditions. Such characteristics can lead to voltage regulation difficulties, greater operational uncertainty, and increased vulnerability when feeder disturbances occur (Byeon & Kim, 2024). Renewable intermittency combined with time-varying demand can further complicate the maintenance of secure and reliable network operation, particularly when decisions must be made under rapidly changing conditions (Huang et al., 2022), (Qiu et al., 2024).

Network restoration presents an additional challenge because the system configuration itself may change following a feeder fault. Conventional ADN operation and restoration methods are commonly formulated using deterministic optimization or model predictive control (MPC) techniques (Glavic, 2019), (Byeon & Kim, 2024). These approaches are effective when system conditions and network configurations are adequately known in advance. However, their effectiveness may decline when renewable generation varies, feeder topology changes, and restoration decisions have to be updated rapidly. Traditional restoration procedures may also depend on predetermined switching sequences and centralized control structures. Such characteristics can restrict the ability of the system to respond adaptively to severe contingencies and changing operating conditions (Glavic et al., 2021), (Zhou et al., 2021).

The development of deep reinforcement learning (DRL) has opened another direction for intelligent operation of power distribution systems. In contrast to conventional optimization approaches, DRL can learn control policies through repeated interaction with the operating environment, making it attractive for adaptive energy management and control applications (Z. Zhang et al., 2021), (Zhao et al., 2016). Existing studies have demonstrated the potential of DRL for voltage regulation, distributed control, and energy management in renewable-integrated distribution networks (B. Zhang et al., 2024), (Xu et al., 2017), (Zhao et al., 2016). Nevertheless, a major limitation of many existing DRL frameworks is that the electrical network is mainly represented through numerical state variables without sufficiently exploiting the structural relationships among buses and feeder branches. Moreover, the learning process may not explicitly enforce physical operating constraints. Consequently, the resulting policies can generate infeasible actions, suffer from unstable learning behavior, or show limited robustness when exposed to network disturbances that differ from the training conditions (Li & Xu, 2024), (Wang et al., 2025).

The structural characteristics of distribution networks suggest that graph-based learning can provide a more suitable representation for this problem. An ADN naturally forms a graph in which buses constitute nodes and feeder connections form edges. Such a representation allows electrical variables to be interpreted together with their spatial and topological relationships. Recent studies have consequently explored graph-based approaches for capturing interactions among network components, while physics-guided learning has been introduced to improve the physical consistency and operational feasibility of intelligent control decisions (B. Zhang et al., 2024), (Wang et al., 2025). Physics-informed learning is particularly relevant to distribution network applications because voltage limits, feeder loading, power balance, and other electrical constraints must remain satisfied during both normal operation and restoration (B. Zhang et al., 2024), (Dominguez-Garcia et al., 2024). The combination of graph representations and physical knowledge therefore provides a promising basis for developing learning-based control strategies that are both topology-aware and operationally feasible.

These capabilities are especially important for resilient self-healing distribution networks. Following a disturbance, an ADN should be capable of recognizing the affected section, isolating the fault, modifying feeder connectivity, and restoring service while maintaining acceptable operating conditions (Huang et al., 2022), (Zhou et al., 2021). The presence of renewable generation and BESSs creates additional flexibility for this process because distributed resources can provide local power support, voltage regulation, and restoration assistance. However, effectively coordinating these resources during a changing network topology remains challenging. Existing resilient optimization and AI-based restoration studies have made substantial progress, but renewable uncertainty, distributed coordination, and rapid adaptive restoration continue to limit autonomous self-healing capability (Glavic et al., 2021), (Behzadi et al., 2024).

To address these challenges, this study develops a Physics-Informed Graph Reinforcement Learning (PGRL) framework for resilient self-healing operation of renewable-integrated active distribution networks. The proposed framework combines graph-based topology cognition with reinforcement learning and physics-guided operational constraints, while coordinated DG-BESS control is employed to support network operation and restoration. The graph representation enables the learning agent to account for electrical relationships among buses and feeder sections, whereas the physics-informed mechanism incorporates network operating requirements into the learning process. In this way, the proposed framework is designed to provide adaptive control decisions during renewable variability and feeder contingencies while maintaining voltage security and restoration feasibility.

The main contributions of this paper are summarized as follows: i) A topology-aware Physics-Informed Graph Reinforcement Learning framework is developed for adaptive and resilient operation of renewable-integrated active distribution networks; ii) A graph-based representation is incorporated into the learning process to capture the changing connectivity and electrical interactions among buses and feeder sections during normal and post-fault operation; iii) A physics-informed decision mechanism is developed to relate the learned control actions to voltage, feeder loading, BESS operating limits, network losses, ENS, and switching requirements; iv) Coordinated DG-BESS control and feeder reconfiguration are integrated into a unified control policy to support both operational voltage regulation and post-fault service restoration; v) The proposed framework is evaluated on the IEEE 33-bus ADN in terms of voltage regulation, power loss, restoration performance, and learning convergence under renewable variability and feeder contingencies.

SYSTEM MODEL AND PROBLEM FORMULATION

The proposed framework is formulated for the IEEE 33-bus active distribution network (ADN) equipped with distributed photovoltaic (PV) generation and battery energy storage systems (BESSs). The considered system represents a renewable-rich operating environment in which intermittent DG output and time-varying load demand can result in bidirectional power flows (Byeon & Kim, 2024), (Qiu et al., 2024). Under these conditions, appropriate voltage regulation, low network power losses, and reliable restoration following feeder disturbances are essential requirements for secure ADN operation.

The steady-state behavior of the distribution network is described through active and reactive power balance relationships. The power supplied by the DG and BESS units must collectively satisfy the network demand while complying with the electrical and operating limitations of the feeders. Accordingly, the active and reactive power balance equations are given by:

$$P_i^{DG} + P_i^{BESS} - P_i^{load} = V_i \sum_{j=1}^{N} V_j \left(G_{ij} cos\theta_{ij} + B_{ij} sin\theta_{ij} \right) \quad (1)$$

$$Q_i^{DG} + Q_i^{BESS} - Q_i^{load} = V_i \sum_{j=1}^{N} V_j \left(G_{ij} cos\theta_{ij} + B_{ij} sin\theta_{ij} \right) \quad (2)$$

Where P_i^{DG} and Q_i^{DG} denote DG active and reactive power outputs, while P_i^{BESS} and Q_i^{BESS} represent BESS injections. V_i is the voltage magnitude at bus i, and G_{ij} and B_{ij} denote network conductance and susceptance matrices, respectively.

Network power loss is included in the formulation to improve the overall operating efficiency of the ADN. The total feeder loss is calculated as:

$$P_{load} = \sum_{k=1}^{N_b} R_k \frac{P_k^2 + Q_k^2}{V_k^2} \quad (3)$$

Where R_k represents feeder resistance, while P_k^2 and Q_k^2 denote branch active and reactive power flows. In a renewable-integrated distribution network, variations in renewable generation and feeder operating conditions can lead to increased losses and voltage deviations, particularly during contingency events (Huang et al., 2022), (Zhou et al., 2021). The proposed formulation therefore imposes the necessary operating constraints to preserve network security and ensure that restoration decisions remain feasible.

The voltage magnitude of each bus is restricted within the permissible operating range according to:

$$V_i^{min} \leq V_i \leq V_i^{max} \quad (4)$$

The thermal operating limit of each feeder is imposed as:

$$S_{ij} \leq S_{ij}^{max} \quad (5)$$

Where S_{ij} is feeder apparent power and S_{ij}^{max} denotes maximum thermal capacity.

Since the proposed framework relies on BESS units to provide support during restoration, their energy operating characteristics must also be incorporated into the mathematical model. The allowable state-of-charge (SOC) range is specified by:

$$SOC^{min} \leq SOC_t \leq SOC^{max} \tag{6}$$

The temporal variation of the BESS SOC resulting from charging and discharging operations is described by:

$$SOC_{t+1} = SOC_t + \eta_c P_t^{ch} - -\frac{P_t^{dis}}{\eta_d} \Delta t \tag{7}$$

Where $\eta_c P_t^{ch}$ and η_d_dηd denote charging and discharging efficiencies, respectively.

To account for both normal operating efficiency and restoration performance, a resilience-oriented multi-objective formulation is adopted. The optimization considers four main performance aspects: network power loss, voltage deviation, energy not supplied (ENS), and switching actions required during restoration. The resulting objective function is formulated as:

$$F = \alpha P_{loss} + \beta V_{loss} + \gamma ENS + \delta SW_{pen} \tag{8}$$

Where $\alpha,\beta,\gamma,\delta$ are weighting coefficients associated with operational objectives.

The voltage deviation index is calculated by:

$$V_{dev} = \sum_{k=1}^{N} |V_i - V_{ref}| \tag{9}$$

The energy not supplied (ENS) index is subsequently calculated as:

$$ENS = \sum_{t=1}^{T} (P_{demand} - P_{served})\, \Delta t \tag{10}$$

Finally, the resilience index is introduced to quantify the capability of the ADN to recover its service following a contingency:

$$RI(t) = \frac{P_{served}(t)}{P_{total}} \tag{11}$$

A larger value of the resilience index corresponds to better self-healing capability and more effective restoration performance under contingency conditions (Glavic et al., 2021).

PHYSICS-INFORMED GRAPH REINFORCEMENT LEARNING (PGRL) FRAMEWORK METHOD

The proposed Physics-Informed Graph Reinforcement Learning (PGRL) framework is designed to integrate network topology information, reinforcement learning, and physical operating requirements within a unified decision-making process for renewable-integrated active distribution networks. The underlying idea is that the operating state of an active distribution network cannot be adequately characterized by individual electrical variables alone because the behavior of each bus is strongly influenced by its electrical connections with neighboring buses and feeder sections. Therefore, the proposed framework first represents the ADN as a graph to preserve these spatial relationships and then uses the resulting topology-aware information as part of the reinforcement learning process. Physical operating constraints are further incorporated to guide the learning agent toward feasible control decisions. This distinguishes the proposed approach from conventional deep reinforcement learning (DRL), where network states are often processed primarily as independent feature vectors without explicitly exploiting feeder connectivity (B. Zhang et al., 2024), (Z. Zhang et al., 2021).

Graph Representation of the Active Distribution Network

An ADN has an inherent graph structure because its buses and feeder branches form a connected electrical network. In the proposed framework, this structure is explicitly represented as:

$$G = (V, E) \tag{12}$$

Where V represents bus nodes and E denotes feeder connections. Each graph node contains electrical features including voltage magnitude, power flow conditions, DG output, feeder loading, and BESS state of charge.

The connectivity between buses is encoded using an adjacency matrix:

$$A_{ij} = \begin{cases} 1, & \text{if buses } i \text{ and } j \text{ are connected} \\ 0, & \text{otherwise} \end{cases} \tag{14}$$

Although the binary adjacency matrix identifies the connectivity between buses, the electrical characteristics of each feeder branch are also relevant to the control problem. Therefore, the graph representation is interpreted together with branch-level electrical information, including feeder impedance, loading condition, thermal capacity, and switching status. This allows the learning agent to distinguish between electrically different branches even when they have the same topological connectivity.

The topology representation is updated whenever a switching operation changes the feeder configuration. Consequently, the graph used by the PGRL agent represents the current operating topology rather than a fixed network structure. This feature is particularly important during self-healing, where fault isolation and feeder reconfiguration modify the electrical paths through which active and reactive power are transferred.

This representation is particularly important for resilient operation because the topology of an ADN is not necessarily fixed during a disturbance. When a feeder fault occurs, the affected branch may be isolated and other feeder connections may be activated through switching operations. As a result, the electrical relationships among buses can change during the restoration process. By explicitly encoding these relationships in the graph, the PGRL framework can distinguish between operating states that may have similar electrical measurements but different network configurations.

Topology-Aware Feature Extraction

After constructing the graph representation, the electrical information associated with the network nodes is processed through graph convolution. The topology-aware feature extraction operation is expressed as:

$$H^{(l+1)} = \sigma\left(\widehat{D}^{-1/2}\hat{A}\widehat{D}^{-1/2}\right)H^{(l)}W^{(l)} \tag{15}$$

Where $H^{(l)}$denotes graph feature embeddings, $\hat{A}$ is the augmented adjacency matrix, $\widehat{D}$ represents the degree matrix, and $W^{(l)}$ denotes trainable graph-learning parameters.

The purpose of this operation is to aggregate information from electrically connected nodes while retaining the structural characteristics of the feeder network. In other words, the representation of a particular bus is influenced not only by its own voltage, power flow, DG output, and BESS condition, but also by information propagated through its neighboring electrical connections. This allows the learning agent to obtain a more informative representation of network operating conditions.

Such topology awareness becomes especially valuable during self-healing. Under normal operation, the graph describes the existing feeder configuration. Following a fault, however, the connectivity structure changes as faulted sections are isolated and the network is reconfigured. The graph-learning mechanism can therefore provide updated topology-dependent features to the reinforcement learning agent, allowing subsequent control decisions to reflect the current network configuration.

Reinforcement Learning State and Action Formulation

The topology-aware graph features are subsequently provided to the reinforcement learning agent, which interacts continuously with the ADN environment. At each decision step, the agent observes the current electrical and operational conditions and determines an appropriate control action. The system state is defined as:

$$s_t = [V_t, P_t, Q_t, SOC_t, T_t] \tag{16}$$

Where V_t denotes bus voltage magnitudes, $P_t$and Q_t represent power flow conditions, SOC_t is the battery state of charge, and $T_t$denotes feeder topology status.

This state formulation combines electrical variables with topology and energy-storage information. Therefore, the agent can simultaneously recognize voltage conditions, power-flow behavior, available BESS flexibility, and changes in feeder connectivity. This is important for resilient operation because an action that is appropriate under one topology may become unsuitable after a feeder fault or switching operation.

Based on the observed state, the agent determines the control variables associated with DG, BESS, and feeder reconfiguration. The action space is expressed as:

$$a_t = [Q_{DG}, P_{BEES}, SW] \tag{17}$$

The action formulation provides three complementary control mechanisms. DG reactive-power control can be used to support voltage regulation, BESS scheduling provides active-power flexibility and energy support, while feeder switching enables modification of the network configuration during restoration. Their coordinated use allows the learning agent to address both normal operating requirements and contingency-related restoration needs within the same decision-making framework.

Reinforcement Learning Objective

The policy of the reinforcement learning agent is optimized by maximizing the long-term cumulative operational reward:

$$J(\theta) = \mathbb{E}\left[\sum_{t=1}^{T} \gamma^t r_t\right] \tag{18}$$

Where γ denotes the discount factor.

The use of a cumulative reward rather than a single-step performance measure allows the agent to consider the consequences of its current decisions over subsequent operating periods. This is particularly relevant to BESS operation and feeder restoration because an immediate improvement in one operating variable may affect the availability of resources or the network condition at later stages. The learning process therefore seeks a control policy that provides favorable overall performance across the operating trajectory.

Reinforcement Learning Objective

The PGRL controller adopts a policy-gradient learning strategy to update the control policy from the interaction between the agent and the ADN environment. The topology-aware features extracted by the graph convolution layers are provided to the policy network together with the current electrical state, BESS state, and feeder topology information. The policy network then generates the control actions associated with DG reactive-power regulation, BESS operation, and feeder switching.

The policy is updated iteratively using the cumulative reward obtained from the ADN environment. At each training step, the agent observes the current state, generates an action, performs the corresponding network operation, evaluates the resulting electrical condition, and receives a physics-informed reward. The policy parameters are subsequently updated to increase the probability of actions associated with favorable operational and restoration performance.

The graph-learning component consists of two graph convolution layers followed by a fully connected policy network. The graph layers aggregate information from electrically connected buses, while the policy network maps the resulting topology-aware representation into control decisions. This architecture enables the controller to jointly process electrical measurements, network connectivity, and distributed-resource conditions within the same decision-making process.

Physics-Informed Reward Mechanism

A key component of the proposed framework is the incorporation of physical operating requirements into the reinforcement learning process. In a conventional data-driven learning framework, the agent may explore a large number of actions without directly accounting for whether those actions are desirable from an electrical-system perspective. For an ADN, such unrestricted exploration can result in excessive power losses, unacceptable voltage deviations, unnecessary switching operations, or increased energy not supplied during restoration.

To address this issue, the proposed PGRL framework employs a physics-informed reward function that simultaneously accounts for network power loss, voltage deviation, ENS, and unnecessary switching actions:

$$r_t = -(\alpha P_{loss} + \beta V_{loss} + \gamma ENS + \delta SW_{pen}) \quad (19)$$

This formulation establishes a direct connection between the learning objective and the physical performance of the distribution network. A control policy receives a more favorable reward when it maintains acceptable electrical conditions while reducing losses, limiting voltage deviations, minimizing unsupplied energy, and avoiding unnecessary feeder switching.

The physics-informed reward therefore serves two purposes. First, it provides the agent with a clear operational objective that is consistent with the requirements of resilient ADN operation. Second, it discourages infeasible or undesirable exploration during training. By embedding these physical considerations into the learning process, the framework improves learning stability and reduces the probability of generating operationally unsuitable actions (Li & Xu, 2024), (Yin et al., 2025).

Physics-Constrained Action Evaluation

To improve the physical consistency of the learned control policy, each candidate action generated by the PGRL agent is evaluated against the operating conditions of the distribution network before its performance is reflected in the reward. The selected action determines the DG reactive-power outputs, BESS charging or discharging operation, and feeder switching status. The resulting operating state is then evaluated through a power-flow calculation to determine whether the voltage, feeder loading, power balance, and BESS operating conditions remain within their admissible limits.

In particular, an action is considered operationally feasible when the resulting network satisfies the voltage constraint, feeder thermal constraint, and BESS state-of-charge limits defined in (4)-(7). Actions that produce undesirable operating conditions are penalized through the physics-informed reward mechanism. This creates a direct interaction between the learned policy and the electrical behavior of the ADN rather than allowing the reinforcement learning agent to optimize the reward independently of network feasibility.

The action-evaluation process can therefore be summarized as

$a_t \rightarrow \textit{power-flow evaluation} \rightarrow \textit{constraint checking} \rightarrow r_t$

where a_t denotes the control action generated by the agent and r_t represents the resulting physics – informed reward. This mechanism reduces the exploration of electrically unsuitable operating states and improves the consistency between the learned policy and the physical requirements of the distribution network.

During a feeder disturbance, the same procedure is applied to the post-fault topology. Consequently, the feasibility of a switching or DG-BESS control action is evaluated according to the current network configuration rather than according to a fixed pre-contingency topology.

Overall PGRL Operation

The complete PGRL framework consequently establishes a sequential relationship between network topology representation, topology-aware feature extraction, state observation, coordinated control, and physics-informed policy learning. The graph component provides structural information about the ADN, the reinforcement learning agent converts the observed network state into adaptive DG-BESS and switching decisions, and the physics-informed reward evaluates these decisions according to their electrical and resilience performance.

This integrated mechanism enables the framework to adapt its control policy when renewable generation varies or when feeder connectivity changes following a disturbance. In particular, the combination of topology awareness and physical guidance allows the agent to support autonomous restoration while maintaining voltage security and other operating requirements throughout the self-healing process.

The overall operating sequence of the proposed PGRL framework is illustrated in Figure 1.

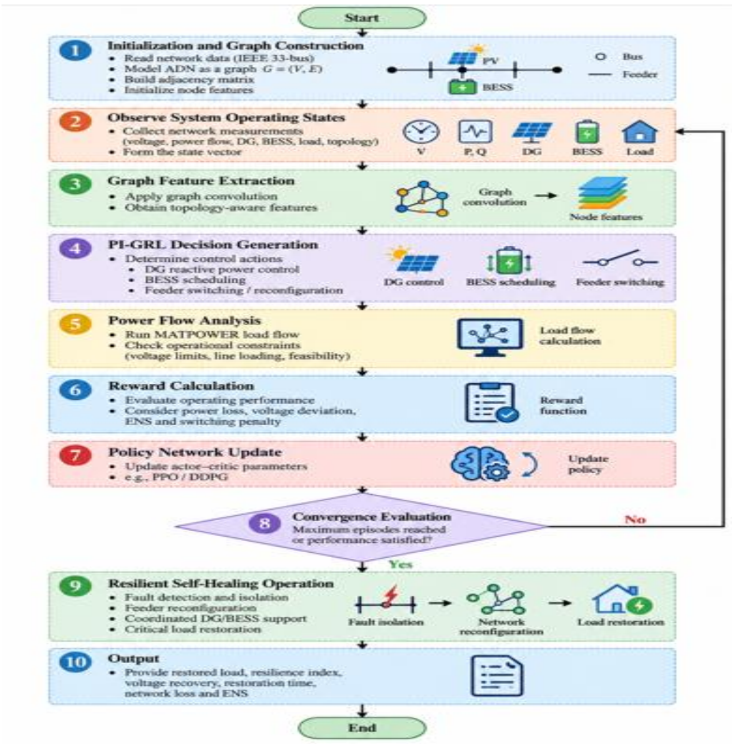

Figure 1. Flowchart of the proposed PGRL

RESULTS AND DISCUSSION

Simulation Configuration

The test system consists of the IEEE 33-bus ADN equipped with three distributed PV units at buses 6, 18, and 30 and two BESS units located at buses 14 and 25. Controllable tie-switches are also considered to enable feeder reconfiguration following a disturbance. The base network demand is 3.72 MW and 2.30 MVAr, with PV penetration fixed at 42% of the total load demand. Each BESS has a capacity of 500 kWh and a charging/discharging efficiency of 95%.

The PGRL agent is trained over 1000 episodes using adaptive policy optimization. Its graph-learning architecture contains two graph convolution layers followed by a fully connected policy network, with topology-aware feature aggregation used to preserve the electrical relationships among interconnected buses. The principal simulation parameters are summarized in Table 1.

Table 1. Simulation Parameters of the Proposed PGRL Framework

Parameter	Value
Test system	IEEE 33-bus ADN
Total load demand	3.72 MW
PV penetration	42%
Number of PV units	3
Number of BESS units	2
BESS capacity	500 kWh
Voltage limits	0.95-1.05 p.u.
Training episodes	1000
Learning rate	0.0005
Discount factor	0.98
Graph convolution layers	2
Policy network	Fully connected
Learning rate	0.0005
Discount factor	0.98
BESS SOC limits	20–90%
BESS efficiency	95%
Training episodes	1000

The selected operating conditions provide a suitable basis for evaluating the proposed method because they simultaneously introduce renewable variability, distributed energy resources, and network reconfiguration requirements. In particular, the presence of PV and BESS units allows the effectiveness of coordinated distributed control to be examined, while the controllable tie-switches provide the mechanism required for resilient feeder restoration.

Contingency and Uncertainty Scenarios

To assess the adaptability of the proposed PGRL framework beyond a single operating condition, the evaluation considers variations in renewable generation, load demand, and feeder contingencies. The PV generation level is varied to represent renewable intermittency, while load demand is modified to reflect different operating periods. For contingency analysis, feeder faults are introduced at different locations so that the controller must respond to different post-fault network configurations.

For each contingency, the affected feeder section is isolated and the available tie-switches are used to establish an alternative network configuration. The DG and BESS units then provide local voltage and power support during the restoration process. The performance of each control strategy is evaluated using restoration time, ENS, minimum voltage, switching actions, and resilience index.

The training and evaluation conditions are kept consistent across the compared methods to ensure a fair assessment of the control strategies.

Table 2. Contingency Scenarios

Scenario	Faulted Branch	PV Condition	Load Condition
S1	9–10	Normal	Normal
S2	14–15	Low PV	High load
S3	19–20	High PV	Normal load
S4	25–26	Variable PV	Peak load

Voltage Regulation Performance

Voltage regulation is a critical aspect of operating renewable-integrated active distribution networks (ADNs), particularly when distributed generation and variable feeder loading cause the voltage profile to change along the feeder. These variations are generally more evident at electrically remote buses, where the influence of feeder impedance and changing power injections becomes stronger. Figure 2 compares the voltage profiles obtained from the four control strategies under identical operating conditions.

Table 3. Voltage Regulation Comparison

Method	Minimum Voltage (p.u.)	Maximum Voltage Deviation	Voltage within Limits (%)
DRL-based	0.950	0.050	91.2
Graph-Based Volt-VAR	0.965	0.035	96.1
PI-MADRL	0.972	0.028	98.4
PGRL	0.980	0.020	99.6

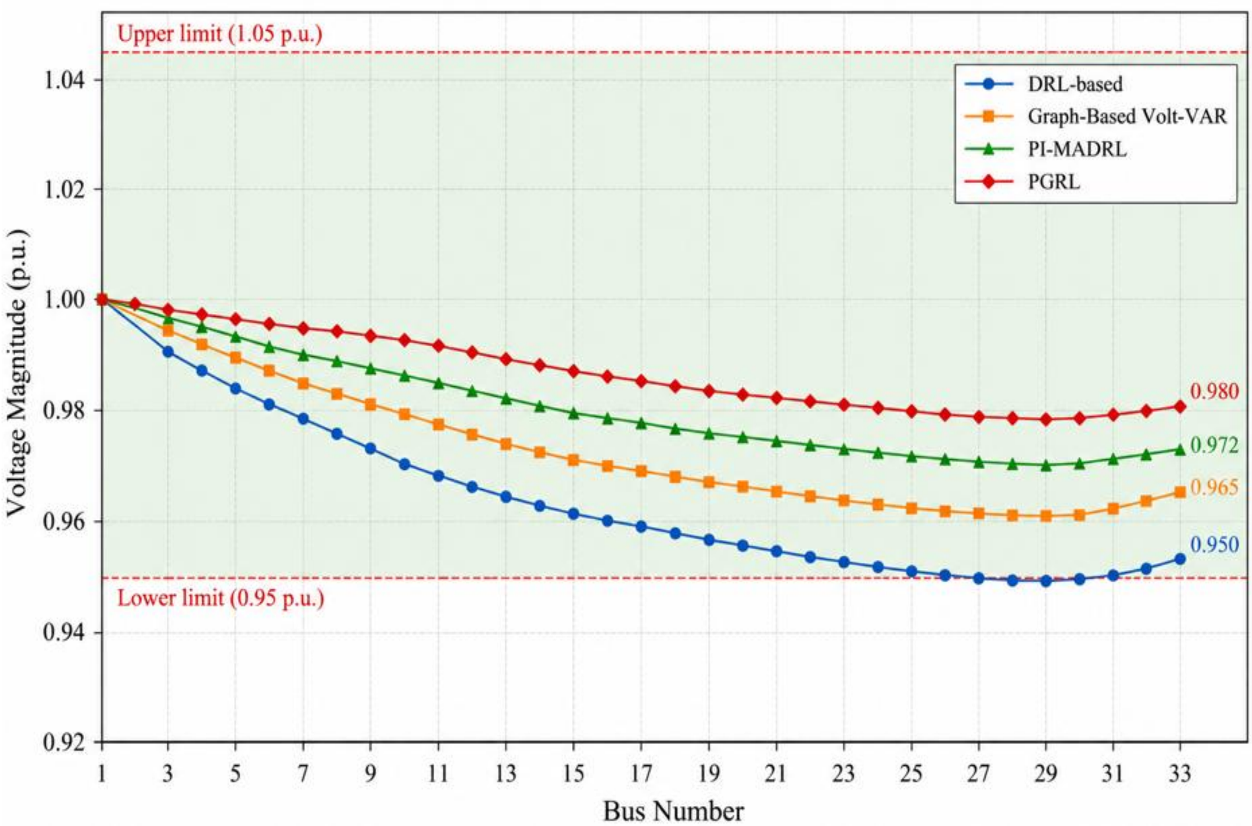

Figure 2. Voltage profile comparison under different methods

The DRL-based energy management approach of (Zhao et al., 2016) provides an improvement in voltage regulation, although noticeable voltage variations remain at electrically remote buses when PV generation becomes highly intermittent. The graph-based Volt-VAR approach reported by (Li & Xu, 2024) achieves a further improvement by incorporating interactions between network components. The physics-informed MADRL method of (B. Zhang et al., 2024) produces a smoother voltage response through coordinated distributed control.

The quantitative results in Table 3 confirm the progressive improvement among the considered approaches. The minimum voltage increases from 0.950 p.u. with DRL-based control to 0.965 p.u. for Graph-Based Volt-VAR and 0.972 p.u. for PI-MADRL, while PGRL achieves the highest minimum voltage of 0.980 p.u. Correspondingly, the maximum voltage deviation decreases from 0.050 to 0.035, 0.028, and finally 0.020. The proportion of buses operating within the 0.95–1.05 p.u. range also increases from 91.2% for DRL-based control to 96.1%, 98.4%, and 99.6% for Graph-Based Volt-VAR, PI-MADRL, and PGRL, respectively.

The improvement obtained with PGRL can be explained by its topology-aware representation of the distribution network. Instead of processing the electrical measurements as isolated variables, the graph structure allows information to be exchanged according to the actual connectivity between buses. Consequently, the controller can better recognize how changes in DG output, feeder loading, or local voltage conditions propagate through neighboring sections of the network.

The coordinated control of DG and BESS provides an additional mechanism for voltage regulation. DG reactive-power control can compensate for local voltage variations, while BESS operation provides further flexibility when renewable generation or load demand changes. When the existing feeder configuration becomes less favorable, adaptive switching can additionally modify the power-flow paths and support voltage recovery.

The results therefore suggest that the voltage improvement achieved by PGRL is associated with the combined effect of topology-aware graph learning, coordinated DG-BESS control, and adaptive feeder reconfiguration, rather than simply an increase in the number of controllable variables. The resulting voltage profile is more uniform along the feeder and provides a larger margin against voltage-limit violations under renewable variability.

Network Power Loss Reduction

Total network power loss is an important indicator for assessing the operating efficiency of an active distribution network (ADN). As renewable penetration increases, inappropriate power-flow distribution may result in higher feeder loading and additional losses, thereby reducing part of the benefits provided by distributed generation (Qiu et al., 2024), (Behzadi et al., 2024). Therefore, this subsection evaluates the network power loss obtained using the different control strategies under the same operating conditions.

The DRL-based method reduces network losses by coordinating DG operation; however, its effectiveness can be affected by changes in network configuration because the spatial relationships among buses are not explicitly represented. The Graph-Based Volt-VAR method provides further improvement by considering interactions among network components, while the Robust DG-BESS strategy takes greater advantage of distributed energy resources to regulate power flows and reduce feeder loading. Nevertheless, the proposed PGRL method achieves the lowest power loss among all considered approaches, indicating the benefit of incorporating topology-aware learning and coordinated DG-BESS control.

Table 4. Network Power Loss Comparison

Method	Power Loss (kW)	Loss Reduction (%)
DRL-based	159.8	26.1
Robust DG-BESS	142.6	34.2
Graph-Based Volt-VAR	128.3	40.8
PGRL	115.7	46.6

The corresponding comparison is presented in Figure 3.

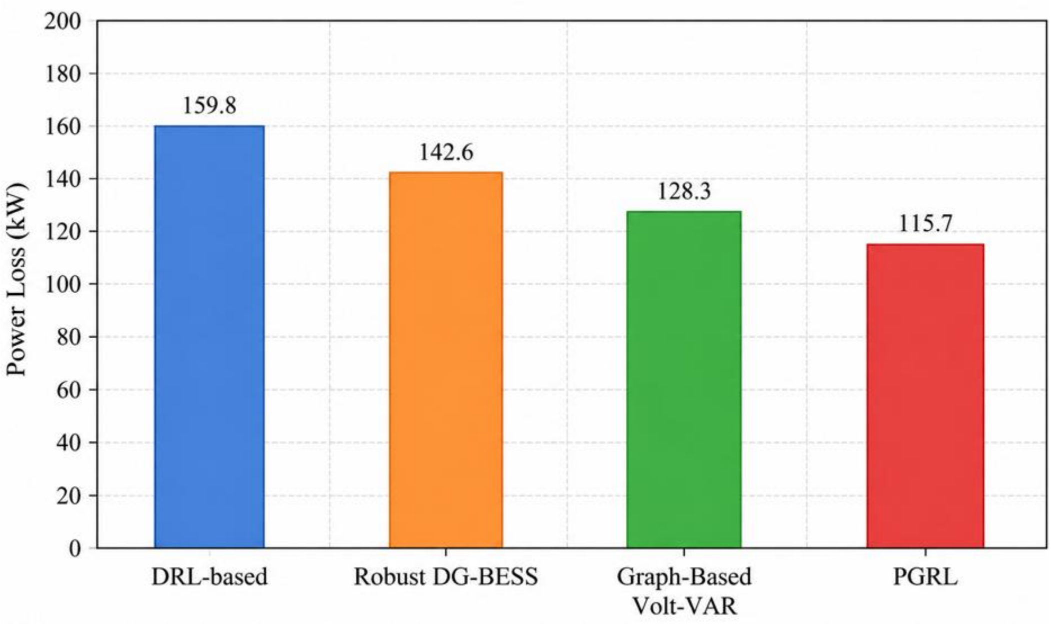

Figure 3. Network power loss comparison

As shown in Table 4, PGRL obtains the lowest total network power loss of 115.7 kW, corresponding to a 46.6% loss reduction. The DRL-based, Robust DG-BESS, and Graph-Based Volt-VAR approaches result in power losses of 159.8 kW, 142.6 kW, and 128.3 kW, respectively, with corresponding loss reductions of 26.1%, 34.2%, and 40.8%.

The superior performance of PGRL can be attributed primarily to its ability to incorporate electrical dependencies through graph-based network representation. By preserving the connectivity relationships among buses and feeder sections, the learning agent can better recognize how changes in feeder configuration influence power-flow distribution. This information allows DG and BESS resources to be coordinated according to the overall network condition rather than relying only on individual voltage or power measurements.

The coordinated operation of DG and BESS further contributes to reducing unnecessary power transfers through the feeders. DG units can supply part of the local demand, while BESS units provide additional flexibility for balancing power flows under changing renewable generation and load conditions. As a result, feeder loading can be reduced and power transfer paths can be adjusted more effectively.

The difference between PGRL and the other approaches also indicates the importance of considering network topology during optimization. Although the Graph-Based Volt-VAR method already exploits network interactions, PGRL combines topology information with reinforcement learning and coordinated DG-BESS control. This enables the control policy to adapt its decisions to changing operating conditions and network configurations.

Overall, the results demonstrate that incorporating topology awareness into the learning process can improve power-flow coordination and reduce network losses. The 115.7 kW loss obtained by PGRL also agrees with the improved voltage-regulation performance observed in the previous subsection, suggesting that better voltage control and coordinated distributed-resource operation contribute jointly to more efficient ADN operation.

Resilient Self-Healing Performance

The self-healing capability of the considered control strategies is evaluated by imposing a feeder fault between buses 9 and 10 at $t = 4$ h. Following the occurrence of the fault, each control method is required to identify and isolate the affected section, perform the necessary feeder reconfiguration, and coordinate the available DG-BESS resources to recover the interrupted load. The recovery process is then evaluated according to restoration speed, energy not supplied (ENS), and the resulting resilience level.

In contrast, PGRL exhibits the fastest recovery trajectory and reaches a high restored-load level earlier than the other approaches. This behavior indicates that the proposed controller can adapt its restoration decisions more effectively after the feeder configuration changes.

Table 5. Self-Healing Restoration Performance

Method	Restoration Time (min)	ENS (kWh)	Resilience Index
Hybrid DL	28	192	0.86
DRL-Based	21	141	0.91
PI-MADRL	16	97	0.94
PGRL	11	65	0.97

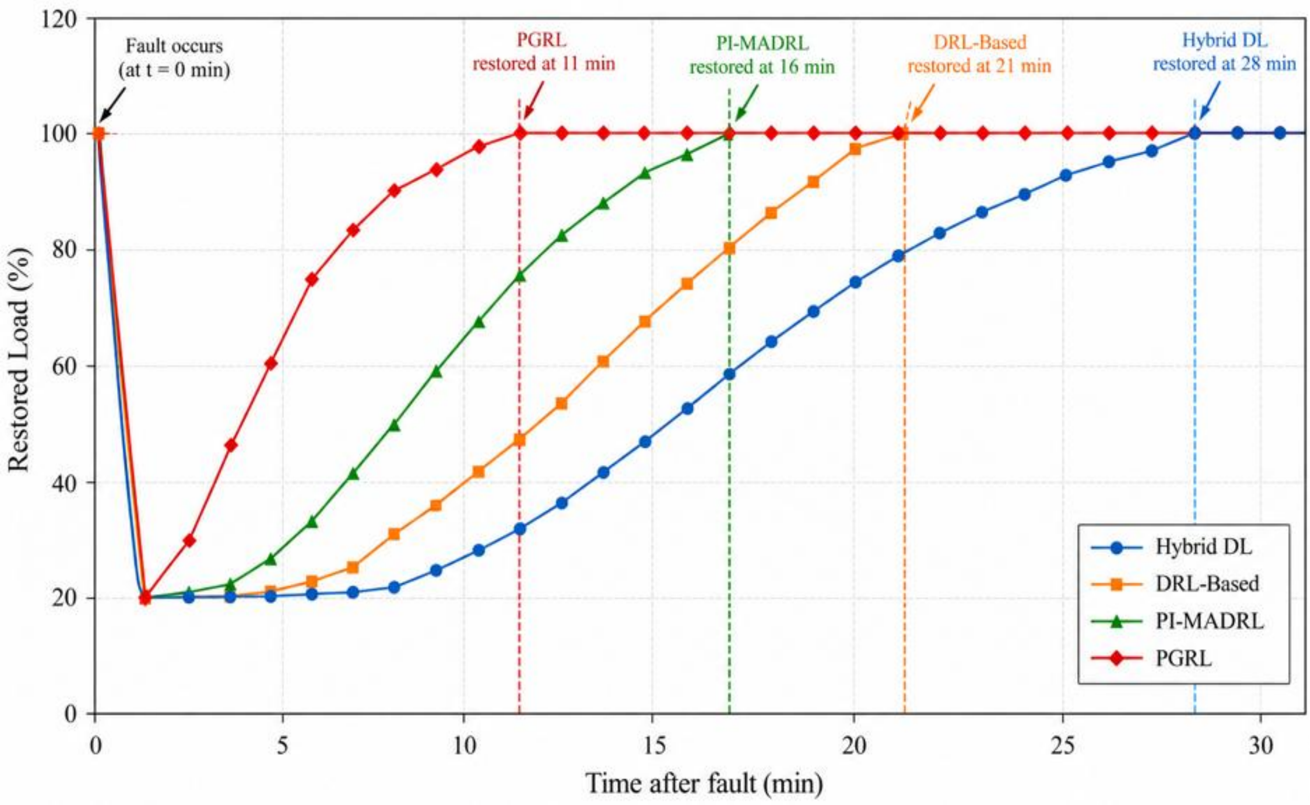

Figure 4. Self-healing load restoration performance under feeder fault.

Figure 4 illustrates the recovery trajectories obtained under the considered approaches. The Hybrid DL method reported by (Zhou et al., 2021) provides considerable restoration flexibility, but its recovery performance remains sensitive to variations in network topology. The DRL-based method of (Zhao et al., 2016) reduces restoration time but still experiences delays under severe contingency conditions. By incorporating operational constraints into the learning process, the physics-informed MADRL method of (B. Zhang et al., 2024) provides more stable restoration behavior.

The improvement in restoration performance is closely related to the dynamic treatment of network topology. Once the feeder fault occurs, the connectivity of the ADN changes. The graph-learning component can represent this changed configuration and provide updated topology information to the reinforcement learning policy. The controller can consequently select a restoration action based on the actual post-fault network structure rather than relying on a fixed restoration sequence.

The coordinated use of DG and BESS also plays an important role. DG resources can contribute local power support, while BESS units provide controllable energy during the restoration process. Combined with feeder switching, these resources allow the proposed framework to restore service more rapidly while limiting the amount of energy that remains unsupplied.

The proposed PGRL framework achieves a restoration time of 11 min and limits the energy not supplied to 65 kWh, indicating a faster and more effective recovery process than the benchmark approaches. The resulting resilience index of 0.97 further demonstrates the stronger recovery capability achieved under the considered feeder contingency. These results suggest that explicitly incorporating topology information into the learning process, together with physics-informed decision guidance and coordinated DG-BESS operation, can improve the adaptability of ADN restoration.

Learning Convergence Analysis

The convergence behavior of the three learning-based strategies is investigated to assess the effectiveness of the proposed PGRL framework during policy training. In particular, the analysis focuses on the ability of each method to achieve a stable control policy under the considered renewable variability and distribution-network operating conditions. Figure 5 shows the evolution of the cumulative reward throughout the training process.

Table 6. Learning Convergence Comparison

Method	Stable Convergence Episode	Training Stability
DRL-based	780	Moderate
PI-MADRL	630	High
PGRL	540	Very High

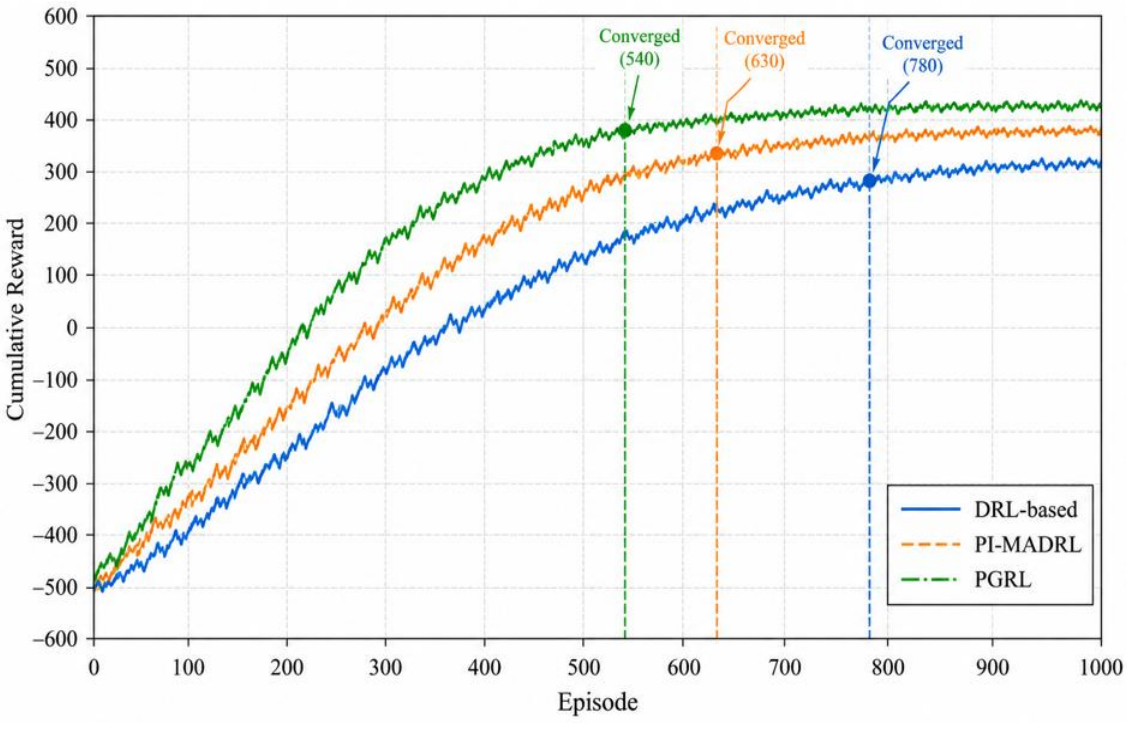

Figure 5. Learning convergence characteristics

Conventional DRL methods typically exhibit unstable exploration behavior under renewable uncertainty because the learning process lacks sufficient physical operating awareness (Z. Zhang et al., 2021), (Zhao et al., 2016). By incorporating physics-informed operational constraints into the reward formulation, the proposed PGRL framework significantly improves learning stability and reduces infeasible control actions during training.

The proposed PGRL framework addresses this issue through its physics-informed reward formulation. Power loss, voltage deviation, ENS, and switching penalties are incorporated into the reward evaluation, providing the agent with direct feedback regarding the electrical consequences of its actions. This reduces the tendency toward physically unsuitable exploration and promotes more consistent policy improvement.

PI-MADRL improves this behavior by incorporating physical operating information into the distributed learning process. As observed in Figure 5, its cumulative reward increases more consistently than that of the conventional DRL-based method. However, the learning process still requires a relatively large number of episodes before the policy becomes stable because the agent must learn the interaction between distributed control actions and changing network conditions.

The proposed PGRL framework exhibits a faster and smoother improvement in cumulative reward. The convergence results summarized in Table 5 show that PGRL reaches stable convergence at approximately episode 540, compared with 630 episodes for PI-MADRL and 780 episodes for the DRL-based method. Thus, PGRL reduces the number of training episodes required for stable convergence by approximately 14.3% relative to PI-MADRL and 30.8% relative to DRL-based learning.

The improved convergence behavior can be attributed to the complementary roles of graph-based feature extraction and physics-informed reward guidance. The graph representation preserves the electrical relationships among interconnected buses and feeder sections, allowing the agent to learn from structured network information rather than treating all state variables as independent features. This becomes particularly useful when the network topology or power-flow conditions change during the training process.

At the same time, the physics-informed reward provides additional guidance during exploration by accounting for power loss, voltage deviation, ENS, and switching penalties. Actions that lead to undesirable electrical conditions are therefore penalized directly through the reward mechanism. This reduces unnecessary exploration of inappropriate operating regions and allows the policy to concentrate on control actions that are more consistent with the physical requirements of the ADN.

The training stability obtained by PGRL is consequently classified as Very High, whereas PI-MADRL and DRL-based learning are classified as High and Moderate, respectively. The results suggest that the combination of topology-aware representation and physics-guided learning improves not only the final control performance but also the efficiency of the learning process.

The convergence results indicate that PGRL reaches a stable policy at approximately episode 540, compared with 630 episodes for PI-MADRL and 780 episodes for the conventional DRL-based approach. This corresponds to reductions of approximately 14.3% and 30.8% relative to PI-MADRL and DRL-based learning, respectively. The faster convergence is attributed to the complementary roles of topology-aware feature extraction and physics-informed reward guidance. The graph representation provides structured information about electrical connectivity, while the physics-informed reward reduces exploration of operating states associated with undesirable voltage, loss, ENS, and switching performance. Therefore, the learning agent can focus more efficiently on control actions that satisfy the operational requirements of the ADN.

Ablation Analysis of the PGRL Framework

To examine the contribution of the individual components of PGRL, an ablation analysis is performed by progressively removing the graph-learning and physics-informed components from the proposed framework. The compared configurations retain the same ADN model, distributed resources, operating conditions, and training budget. This allows the performance improvement to be attributed more directly to the structural components of the proposed framework rather than to differences in the simulation setup.

The first configuration uses conventional reinforcement learning with the same control variables but without graph-based feature extraction or physics-informed reward guidance. The second configuration introduces graph-based feature extraction while retaining the conventional reward formulation. The third configuration incorporates physics-informed reward guidance without graph-based feature extraction. The complete PGRL configuration combines both components.

The ablation results are evaluated in terms of voltage regulation, network power loss, restoration performance, and learning convergence. Improvements obtained by the complete PGRL configuration over the partial configurations indicate the contribution of topology-aware representation and physics-guided learning to resilient ADN control.

Table 7. Ablation Analysis of PGRL Components

Configuration	Graph Learning	Physics-Informed Reward	Minimum Voltage	Power Loss	Restoration Time
RL	×	×	—	—	—
RL + Graph	✓	×	—	—	—
RL + Physics	×	✓	—	—	—
PGRL	✓	✓	0.980	115.7	11

Computational Performance

In addition to control performance, the computational characteristics of the learning framework are relevant to its potential application in real-time ADN operation. The training process is performed offline, while the trained policy is used to generate control actions during online operation. Therefore, the computational burden of the proposed framework is considered separately for training and inference.

The training time depends on the number of episodes and the complexity of the graph-learning and policy networks, whereas the online inference stage only requires the trained model to process the current network state and generate the corresponding control action. This separation is important because the computational requirements during online operation are considerably different from those during offline policy training.

Table 8. Computational Performance

Method	Training Time (min)	Stable Convergence	Inference Time (ms)
DRL-based	56.4 ± 1.8	780	7.8 ± 0.4
PI-MADRL	49.8 ± 1.5	630	10.6 ± 0.5
PGRL	45.6 ± 1.3	540	12.4 ± 0.6

DISCUSSION

The results demonstrate that the proposed PGRL framework provides a consistent improvement across voltage regulation, network efficiency, restoration performance, and learning convergence. The improvement is not associated with a single control variable but with the interaction between network representation, physics-guided learning, and coordinated distributed resources.

From the network-control perspective, the graph representation enables the agent to distinguish operating states according to feeder connectivity. This feature becomes particularly important after a fault because the electrical relationships between buses change following fault isolation and switching operations. Consequently, the control policy can respond to the current network configuration rather than relying on a fixed representation of the distribution system.

The physics-informed mechanism provides an additional layer of operational guidance. Instead of evaluating an action only from the accumulated learning reward, the proposed framework relates the reward to electrical quantities such as voltage deviation, network loss, ENS, and switching actions. This connection helps reduce unsuitable exploration and contributes to the faster convergence observed in the training results.

The role of DG and BESS is complementary to the learning architecture. DG reactive-power control provides local voltage support, whereas BESS operation supplies controllable active-power flexibility during renewable fluctuations and restoration. Feeder reconfiguration further modifies the available power-transfer paths after a contingency. The combined use of these resources allows the learned policy to address normal operation and post-fault recovery within the same control framework.

Nevertheless, the current evaluation is based on the IEEE 33-bus system and a limited set of operating conditions. Therefore, the results should be interpreted as a validation of the proposed control concept rather than a complete demonstration of scalability to large-scale distribution systems. Further investigation using larger networks, multiple simultaneous contingencies, communication delays, and hardware-in-the-loop testing would be required before practical deployment.

CONCLUSION

This study proposed a Physics-Informed Graph Reinforcement Learning (PGRL) framework for resilient operation and self-healing of renewable-integrated active distribution networks. The proposed framework integrates topology-aware graph representation, physics-guided decision evaluation, coordinated DG-BESS control, and feeder reconfiguration within a unified learning process. This enables the controller to respond to both renewable and load variations during normal operation and changing network configurations following feeder disturbances. Evaluation on the IEEE 33-bus ADN demonstrates that PGRL provides consistent improvements in voltage regulation, network power-loss reduction, restoration performance, and learning convergence compared with the considered benchmark approaches. In particular, the proposed framework achieves a minimum voltage of 0.980 p.u., reduces network power loss to 115.7 kW, and provides faster post-fault restoration with lower ENS. The convergence analysis further indicates that topology-aware and physics-guided learning can reduce the training effort required to obtain a stable control policy. The main significance of the proposed approach is that network topology, physical operating requirements, and distributed energy resources are considered jointly rather than as separate control problems. This provides a more adaptive basis for resilient ADN operation under renewable variability and feeder disturbances. Nevertheless, the present study is validated using the IEEE 33-bus system and limited contingency conditions. Future work should therefore investigate larger distribution networks, multiple and simultaneous contingencies, communication and measurement uncertainties, and real-time or hardware-in-the-loop implementation.

REFERENCES

1. Behzadi, S., Bagheri, A., & Rabiee, A. (2024). Optimal operation of reconfigurable active distribution networks aiming at resiliency improvement. Electric Power Systems Research, 230, 110169. https://doi.org/10.1016/j.epsr.2024.110169.
2. Byeon, G., & Kim, K. (2024). Distributionally robust decentralized Volt-Var control with network reconfiguration. IEEE Transactions on Smart Grid, 15(5), 4705–4718,. https://doi.org/10.1109/TSG.2024.3377910.
3. Dominguez-Garcia, A. D., Zholbaryssov, M., Amuda, T., & Ajala, O. (2024). An online feedback optimization approach to voltage regulation in inverter-based power distribution networks. IEEE Transactions on Power Systems, 39(1), 1145–1156,. https://doi.org/10.1109/TPWRS.2023.3304112.
4. Glavic, M. (2019). Deep) reinforcement learning for electric power system control and related problems: A short review and perspectives. Annual Reviews in Control, 48, 22–35,. https://doi.org/10.1016/j.arcontrol.2019.09.008.
5. Glavic, M., Moreno, R., Bi, T., & Shahidehpour, M. (2021). Reinforcement learning for electric power system decision and control: Past considerations and perspectives. IEEE Transactions on Smart Grid, 12(6), 5128–5142,. https://doi.org/10.1109/TSG.2021.3070934.
6. Huang, Y., Yang, Q., Tan, J., & Ukil, A. (2022). Deep reinforcement learning for real-time energy management in smart grids. Energy Reports, 8, 1245–1257,. https://doi.org/10.1016/j.egyr.2022.08.190.
7. Li, B., & Xu, Q. A. (2024). A machine learning-assisted distributed optimization method for inverter-based Volt-VAR control in active distribution networks. IEEE Transactions on Power Systems, 39(2),

2668–2681,. https://doi.org/10.1109/TPWRS.2023.3279303.
8. Qiu, W., Yadav, A., You, S., Dong, J., Kuruganti, T., Liu, Y., & Yin, H. (2024). Neural networks-based inverter control: Modeling and adaptive optimization for smart distribution networks. IEEE Transactions on Sustainable Energy, 15(2), 1039–1049,. https://doi.org/10.1109/TSTE.2023.3324219.
9. Wang, S., Hou, Y., Guan, X., Liu, S., & Huo, Z. (2025). Resiliency-informed optimal scheduling of smart distribution network with urban distributed photovoltaic: A stochastic P-robust optimization. Energy, 313, 133449. https://doi.org/10.1016/j.energy.2024.133449.
10. Xu, Y., Dong, Z. Y., Zhang, R., & Hill, D. J. (2017). Multi-timescale coordinated voltage/VAR control of high renewable-penetrated distribution systems. IEEE Transactions on Power Systems, 32(6), 4398–4408,. https://doi.org/10.1109/TPWRS.2017.2659779.
11. Yin, C., Dong, J., & Zhang, Y. (2025). Distributionally robust bilevel optimization model for distribution network with demand response under uncertain renewables using Wasserstein metrics. IEEE Transactions on Sustainable Energy, 16(1), 118–130,. https://doi.org/10.1109/TSTE.2024.3509314.
12. Zhang, B., Cao, D., Hu, W., Ghias, A. M. Y. M., & Chen, Z. (2024). Physics-informed multi-agent deep reinforcement learning enabled distributed voltage control for active distribution network using PV inverters. International Journal of Electrical Power & Energy Systems, 155, 109641. https://doi.org/10.1016/j.ijepes.2023.109641.
13. Zhang, C., Xu, Y., Dong, Z. Y., & Wong, K. P. (2018). Robust coordination of distributed generation and energy storage in distribution networks. IEEE Transactions on Smart Grid, 9(2), 1488–1499,. https://doi.org/10.1109/TSG.2016.2583461.
14. Zhang, Z., Wang, J., & Chen, B. (2021). Deep reinforcement learning for energy management in microgrids. Applied Energy, 288, 116618. https://doi.org/10.1016/j.apenergy.2021.116618.
15. Zhao, J., Zheng, T., & Litvinov, E. (2016). A unified framework for defining and measuring flexibility in power system. IEEE Transactions on Power Systems, 31(1), 339–347,. https://doi.org/10.1109/TPWRS.2015.2390039.
16. Zhou, K., Yang, S., & Shao, Z. (2021). Digital twin-based optimization for energy systems. Applied Energy, 285, 116419. https://doi.org/10.1016/j.apenergy.2020.116419.