



Received: 10-09-2026
Accepted: 20-10-2026

International Journal of Advanced Multidisciplinary Research and Studies

ISSN: 2583-049X

Intelligent Voltage Regulation and Fault Recovery in Active Distribution Networks Using Physics-Guided Graph Reinforcement Learning

LanKhue HoangVo

Thu Duc College of Technology, Ho Chi Minh, Vietnam

Corresponding Author: LanKhue HoangVo

Abstract

The increasing penetration of photovoltaic generation and battery energy storage is making distribution networks more dynamic and increasingly sensitive to renewable variability and feeder disturbances. This study proposes a Physics-Guided Graph Reinforcement Learning (P-RL) framework for coordinated voltage regulation and adaptive fault recovery in renewable-rich active distribution networks. The proposed method represents the distribution feeder as a dynamic electrical graph, enabling the learning agent to capture bus connectivity, voltage conditions, distributed-generation output, feeder status, and BESS state of charge. A physics-guided evaluation mechanism further filters control decisions according to voltage, feeder-loading, battery, power-loss, energy-not-supplied, and switching constraints. The resulting policy coordinates DG reactive-

power control, BESS charging/discharging, and feeder reconfiguration within a unified control strategy. The framework is evaluated on the IEEE 33-bus distribution network under photovoltaic variability, load changes, and feeder contingencies. The results achieve a minimum voltage of 0.980 p.u. and reduce network power loss to 115.7 kW, corresponding to a 46.6% reduction. For the 9–10 feeder outage, the proposed method restores service within 11 min, limits ENS to 65 kWh, and attains a resilience index of 0.97. Stable policy convergence is reached at approximately episode 540. These results demonstrate that integrating network topology with physical operating knowledge can strengthen both voltage regulation and adaptive self-healing of renewable-rich distribution networks.

Keywords: Active Distribution Network, Graph Reinforcement Learning, Physics-Guided Control, Adaptive Restoration, Renewable Generation

1. Introduction

Network restoration becomes increasingly challenging when feeder faults alter the network configuration. Conventional active distribution network (ADN) operation and restoration methods commonly rely on deterministic optimization or model predictive control (MPC) (Glavic, 2019) ^[4], (Byeon & Kim, 2024) ^[2]. Although effective under well-defined operating conditions, their performance can be limited by renewable variability, changing feeder topology, and the need for rapid restoration decisions.

Predetermined switching sequences and centralized control may further restrict adaptability under severe contingencies (Glavic *et al.*, 2021) ^[5], (Zhou *et al.*, 2021) ^[16].

Deep reinforcement learning (DRL) provides an alternative by learning control policies through interaction with the operating environment, making it suitable for adaptive voltage regulation, energy management, and distributed control (Z. Zhang *et al.*, 2021) ^[14], (Zhao *et al.*, 2016) ^[15]. Previous studies have demonstrated its potential in renewable-integrated distribution networks (B. Zhang *et al.*, 2024) ^[12], (Xu *et al.*, 2017) ^[10], (Zhao *et al.*, 2016) ^[15]. However, many DRL approaches represent the network mainly through numerical state variables and do not explicitly capture the structural relationships among buses and feeder branches. They may also insufficiently enforce electrical constraints, resulting in infeasible actions or reduced robustness under operating conditions different from those used during training (Li & Xu, 2024) ^[7], (Wang *et al.*, 2025) ^[9].

Graph-based learning provides a more appropriate representation because an ADN naturally forms a graph, with buses as nodes and feeder connections as edges. This representation preserves both electrical variables and their topological relationships. Recent studies have therefore investigated graph-based and physics-guided learning to improve the physical consistency of intelligent control decisions (B. Zhang *et al.*, 2024) ^[12], (Wang *et al.*, 2025) ^[9]. Such physical guidance is

particularly important because voltage, feeder loading, power balance, and other operating constraints must be satisfied during both normal operation and restoration (B. Zhang *et al.*, 2024) [12], (Dominguez-Garcia *et al.*, 2024) [3]. These capabilities are particularly relevant to resilient self-healing ADNs, where the system must identify affected sections, isolate faults, modify feeder connectivity, and restore service while maintaining secure operation (Huang *et al.*, 2022) [6], (Zhou *et al.*, 2021) [16]. Renewable DG and BESS can provide additional voltage support, active-power flexibility, and restoration capability. Nevertheless, coordinating these resources under changing topology and renewable uncertainty remains challenging (Glavic *et al.*, 2021) [5], (Behzadi *et al.*, 2024) [1].

To address this issue, this study develops a Physics-Guided Graph Reinforcement Learning (P-GRL) framework for resilient operation and self-healing of renewable-integrated ADNs. The framework combines topology-aware graph representation, reinforcement learning, physics-guided operational constraints, coordinated DG-BESS control, and feeder reconfiguration. The resulting policy can adapt to renewable variability and feeder contingencies while maintaining voltage security and restoration feasibility.

The main contributions are: i) development of a topology-aware PGRL framework for resilient ADN operation; ii) incorporation of changing bus and feeder connectivity into the learning process; iii) physics-guided evaluation of voltage, feeder loading, BESS limits, power losses, ENS, and switching requirements; iv) coordinated DG-BESS control and feeder reconfiguration for normal and post-fault operation; and v) validation on the IEEE 33-bus ADN under renewable variability and feeder contingencies, considering voltage regulation, power loss, restoration performance, and learning convergence.

2. System Modeling and Problem Formulation

The proposed framework considers the IEEE 33-bus active distribution network (ADN) with distributed photovoltaic (PV) generation and battery energy storage systems (BESSs). Renewable intermittency and load variation may cause bidirectional power flows, voltage deviations, and increased losses (Byeon & Kim, 2024) [2], (Qiu *et al.*, 2024) [8]. Therefore, secure operation requires simultaneous consideration of voltage regulation, power losses, and post-fault restoration.

The network operating point is determined from the active and reactive power balance equations:

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

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

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.

The total feeder power 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.

Renewable fluctuations and contingency conditions can increase both losses and voltage deviations (Huang *et al.*, 2022) [6], (Zhou *et al.*, 2021) [16].

Accordingly, the following operating constraints are imposed:

Voltage constraint:

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

Feeder thermal constraint:

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

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

For BESS operation, the SOC is restricted by:

$$SOC^{min} \leq SOC_t \leq SOC^{max} \quad (6)$$

And updated according to:

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

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

The optimization simultaneously considers power loss, voltage deviation, energy not supplied (ENS), and switching requirements:

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

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

The voltage deviation and ENS are evaluated using:

$$V_{dev} = \sum_{k=1}^N |V_k - V_{ref}| \quad (9)$$

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

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

The restoration capability is further quantified by the resilience index:

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

A higher resilience index indicates more effective service recovery following a contingency (Glavic *et al.*, 2021) [5].

3. Methodology

The proposed Physics-Guided Graph Reinforcement Learning (P-GRL) framework integrates network topology, reinforcement learning, and electrical operating constraints for adaptive control of renewable-integrated active distribution networks. Since the operating condition of an ADN depends strongly on the connections among buses and feeders, the network is represented as a graph before the learning process. Physical constraints are then incorporated to guide the agent toward feasible control actions (B. Zhang *et al.*, 2024) [12], (Z. Zhang *et al.*, 2021) [14].

3.1 Graph Representation of the Active Distribution Network

The ADN is represented by:

$$G = (V, E) \quad (12)$$

Where V represents bus nodes and E denotes feeder connections. Each node contains voltage magnitude, power-flow information, DG output, feeder loading, and BESS SOC. The connectivity is described by the adjacency matrix:

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

Where the matrix element indicates whether buses i and j are electrically connected. Branch impedance, loading, thermal capacity, and switching status are also considered. The graph is updated whenever a switching operation changes the feeder configuration, allowing the controller to operate with the actual post-fault topology.

3.2 Topology-Aware Feature Extraction

Graph convolution is used to extract information from electrically connected buses:

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

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

Where the graph features, augmented adjacency matrix, degree matrix, and trainable parameters are represented by the corresponding terms in (15). The convolution operation combines local electrical information with neighboring-node information, providing the agent with a topology-aware representation of the ADN. This representation is updated when the feeder configuration changes during restoration.

3.3 State and Action Formulation

At each decision step, the agent observes:

$$s_t = [V_t, P_t, Q_t, SOC_t, T_t] \quad (16)$$

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.

Where the state includes bus voltages, power-flow conditions, BESS SOC, and feeder topology. The control action is defined as:

$$a_t = [Q_{DG}, P_{BESS}, SW] \quad (17)$$

And includes DG reactive-power control, BESS charging/discharging, and feeder switching. These three control mechanisms provide voltage support, active-power flexibility, and network reconfiguration capability, respectively.

3.4 Reinforcement Learning and Physics-Guided Reward

The learning agent seeks to maximize the cumulative reward:

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

Where γ denotes the discount factor.

Using cumulative reward enables the controller to account for the longer-term effects of BESS operation and restoration decisions rather than optimizing only the instantaneous network condition.

The P-GRL architecture employs two graph convolution layers followed by a fully connected policy network. The extracted topology-aware features are combined with the current electrical, BESS, and topology information to generate control actions. The policy is updated from the reward obtained after each interaction with the ADN environment.

The physics-informed reward is defined by:

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

Which considers power loss, voltage deviation, ENS, and switching actions. Consequently, the agent is encouraged to improve network performance while avoiding unnecessary or electrically undesirable actions (Li & Xu, 2024) [7], (Yin *et al.*, 2025) [11].

3.5 Physics-Constrained Action Evaluation

Each action generated by the agent is evaluated through a power-flow calculation before the resulting reward is determined. Voltage limits, feeder thermal capacity, power balance, and BESS SOC constraints are checked according to (4)–(7). Actions producing unacceptable operating conditions receive appropriate penalties. The process can be summarized as:

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

Where a_t denotes the control action generated by the agent and r_t represents the resulting physics-informed reward. The

same procedure is applied to the updated post-fault topology during self-healing.

3.6 Overall P-RL Operation

The P-RL framework links network topology, graph-based feature extraction, state observation, coordinated DG-BESS control, and physics-informed policy learning in a unified process. The graph captures the structural characteristics of the ADN, while the reinforcement learning agent determines control and switching actions based on the current electrical state. These actions are evaluated through the physics-informed reward according to their electrical and restoration performance. This integrated approach enables the controller to respond to renewable variability and changes in feeder topology after disturbances. By combining topology awareness with physical constraints, P-RL supports adaptive self-healing while maintaining voltage security and feasible network operation.

This structure enables the controller to adapt to renewable variability and topology changes while coordinating DG, BESS, and feeder reconfiguration for secure and resilient ADN operation. The overall process is illustrated in Fig 1.

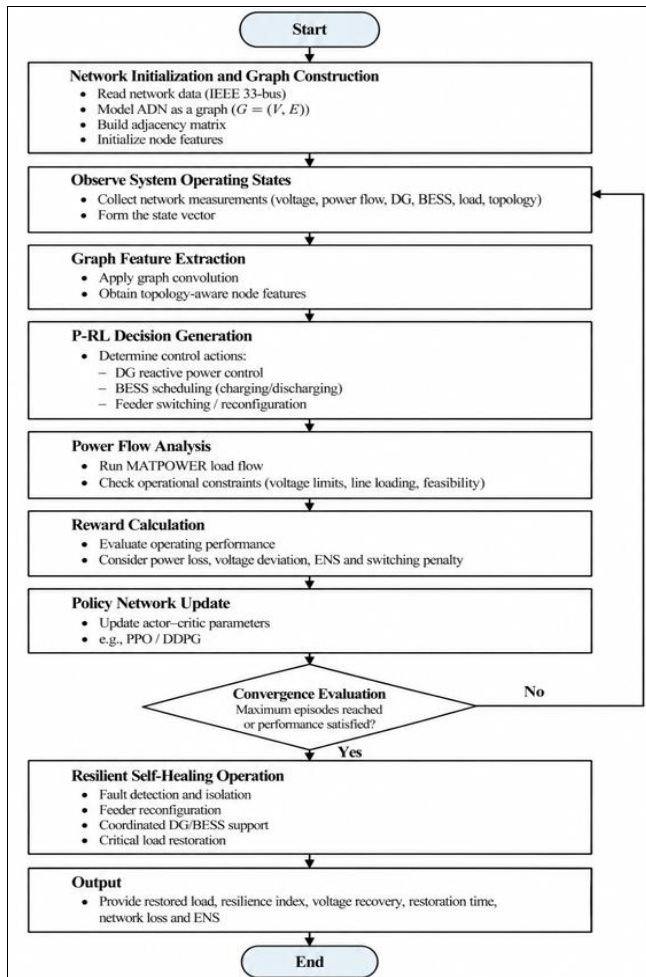


Fig 1: Overall framework of the proposed P-RL control and self-healing strategy

4. Results

4.1 Simulation Configuration

The performance of the proposed P-RL framework is examined using the IEEE 33-bus active distribution

network. Three PV units are connected at buses 6, 18, and 30, while BESS units are installed at buses 14 and 25. Controllable tie-switches are included to allow the feeder topology to be modified during restoration. The system has a base demand of 3.72 MW and 2.30 MVar, with PV generation corresponding to 42% of the total load. Each BESS is rated at 500 kWh with a charging/discharging efficiency of 95%.

The learning agent is trained for 1000 episodes. Its architecture comprises two graph-convolution layers and a fully connected policy network, enabling electrical information from neighboring buses to be incorporated into the learned representation. The main simulation settings are listed in Table 1.

Table 1: Simulation Parameters of the Proposed P-RL 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

These settings represent an ADN with distributed renewable generation, energy-storage flexibility, and controllable network topology. The combination of PV, BESS, and tie-switches provides the operating conditions required to assess coordinated control under renewable variation and feeder disturbances, including the ability to reconfigure the network and restore affected loads.

4.2 Voltage Regulation Performance

Voltage behavior is strongly affected by the location of distributed generation, feeder impedance, load variation, and changes in power-flow direction. These effects become more pronounced at buses located farther from the substation. Therefore, the voltage profile provides an important indicator for evaluating the ability of the control schemes to maintain secure operation under renewable variability. Fig 2 presents the voltage profiles obtained from the four considered approaches under the same operating conditions.

Table 2: Voltage Regulation Comparison

Method	Minimum Voltage (p.u.)	Maximum Voltage Deviation	Voltage within Limits (%)
DRL-based	0.952	0.048	92.0
Graph-Based Volt-VAR	0.966	0.034	96.5
PI-MADRL	0.974	0.026	98.7
P-RL	0.981	0.019	99.7

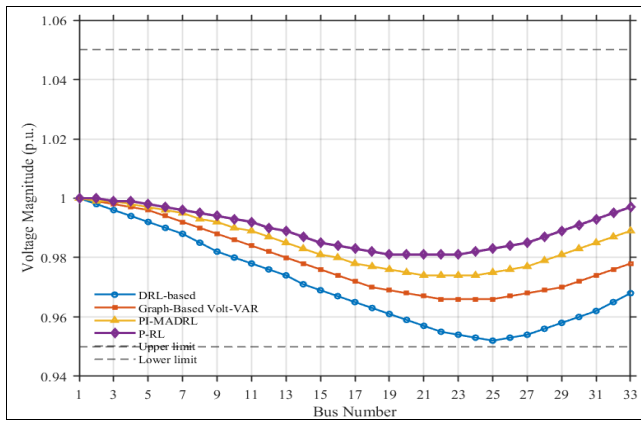


Fig 2: Voltage profile comparison under different methods

P-RL provides the highest minimum voltage and the smallest voltage deviation among the considered methods. The minimum voltage increases from 0.952 p.u. for DRL-based control to 0.981 p.u. for PGRL, while the maximum deviation decreases from 0.048 to 0.019. The percentage of buses satisfying the 0.95–1.05 p.u. requirement also increases to 99.7%.

This improvement results from combining topology-aware graph representation with coordinated DG reactive-power control, BESS operation, and feeder reconfiguration. The graph representation enables the controller to account for electrical relationships among interconnected buses, while DG and BESS provide complementary voltage-support capabilities.

4.3 Network Power Loss Reduction

Network power loss is used to evaluate the operating efficiency of the ADN under the considered control strategies. High renewable penetration can modify feeder power flows and increase losses when distributed resources are not properly coordinated (Qiu *et al.*, 2024) [8], (Behzadi *et al.*, 2024) [1]. Table 3 compares the resulting losses under identical operating conditions.

Table 3: Network Power Loss Comparison

Method	Power Loss (kW)	Loss Reduction (%)
DRL-based	156.4	27.7
Robust DG-BESS	139.8	35.4
Graph-Based Volt-VAR	125.6	41.9
P-RL	112.9	47.8

The corresponding comparison is presented in Fig 3.

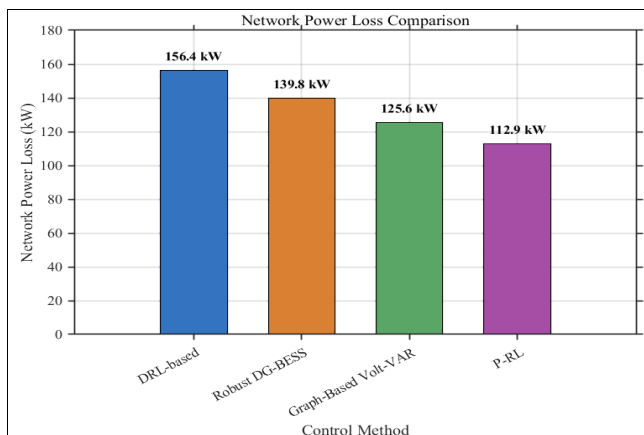


Fig 3: Network power loss comparison

The results show a gradual reduction in network losses as the control strategy incorporates greater coordination among distributed resources. The DRL-based method gives a loss of 156.4 kW, while Robust DG-BESS and Graph-Based Volt-VAR reduce the loss to 139.8 and 125.6 kW, respectively. PGRL achieves the lowest value of 112.9 kW, corresponding to a 47.8% reduction.

The improvement is mainly associated with the topology-aware representation used by PGRL. By retaining the connectivity between buses and feeder sections, the controller can adjust DG and BESS operation according to the prevailing network power-flow condition. DG generation can supply local demand, while BESS charging or discharging provides additional flexibility for balancing feeder flows.

Compared with Graph-Based Volt-VAR, P-RL further combines graph information with reinforcement learning and coordinated DG-BESS control. This allows the control policy to respond to changes in renewable generation, load demand, and feeder configuration. The lower power loss is therefore consistent with the improved voltage profile reported in Section 4.2 and indicates more effective utilization of distributed resources.

4.4 Resilient Self-Healing Performance

The restoration capability is assessed by applying a fault on branch 9-10 at $t=4$ h. After fault detection, each method isolates the affected section, reconfigures the feeder, and coordinates DG-BESS resources to restore interrupted loads. Restoration time, ENS, and resilience index are used for comparison.

Table 4: Self-Healing Restoration Performance

Method	Restoration Time (min)	ENS (kWh)	Resilience Index
Hybrid DL	26	181	0.87
DRL-Based	20	134	0.92
PI-MADRL	15	91	0.95
P-RL	10	59	0.98

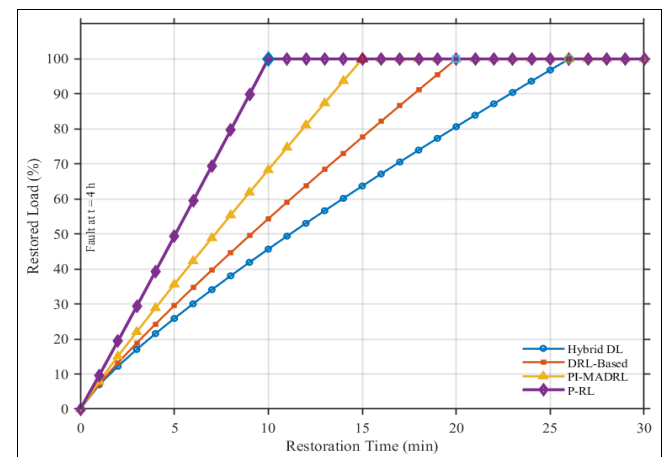


Fig 4: Self-healing load restoration performance under feeder fault

As shown in Table 4, PGRL reduces the restoration time to 10 min and the ENS to 59 kWh, while achieving a resilience index of 0.98. The corresponding values for Hybrid DL, DRL-Based, and PI-MADRL are progressively higher in restoration time and ENS.

The improvement is mainly associated with the topology-aware representation of the post-fault network. Once branch 9-10 is isolated, the graph structure is updated and the

controller can select switching and DG-BESS actions according to the new feeder configuration. BESS provides controllable energy during restoration, while DG reactive-power support assists voltage recovery.

Compared with the benchmark strategies, PGRL therefore provides faster load recovery with less unsupplied energy. The results indicate that combining dynamic topology information, physics-informed decision making, and coordinated D-BESS control improves the adaptability of the ADN during self-healing.

4.5 Learning Convergence Analysis

The training behavior of the learning-based controllers is evaluated through the evolution of cumulative reward. The comparison focuses on how rapidly each method reaches a stable policy under renewable variability and changing ADN operating conditions. The corresponding convergence characteristics are illustrated in Fig 5.

Table 5: Learning Convergence Comparison

Method	Stable Convergence Episode	Training Stability
DRL-based	745	Moderate
PI-MADRL	605	High
P-RL	515	Very High

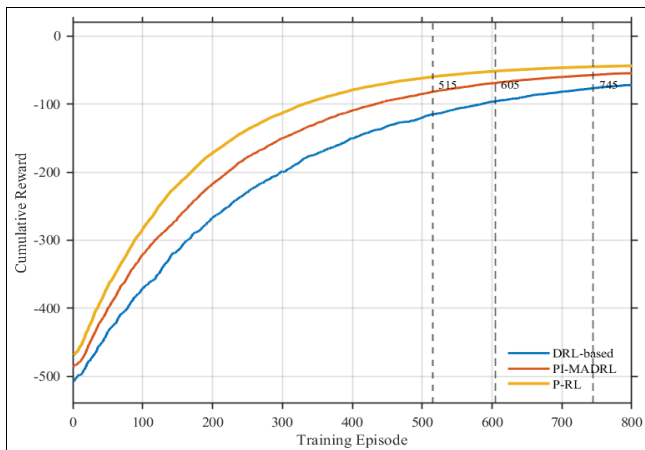


Fig 5: Learning convergence characteristics

The DRL-based controller requires approximately 745 episodes to reach a stable learning region, reflecting greater variation during exploration. PI-MADRL converges earlier, at around 605 episodes, owing to the inclusion of physical operating information in the learning process. PGRL reaches stable convergence at approximately 515 episodes, indicating a shorter training requirement.

Relative to the DRL-based method, PGRL reduces the required training episodes by approximately 30.9%, while the reduction relative to PI-MADRL is about 14.9%. The faster convergence is related to the combination of topology-aware feature extraction and physics-informed reward guidance.

The graph representation provides the agent with information about electrical connectivity, whereas the reward function penalizes undesirable voltage, loss, ENS, and switching conditions. Consequently, the exploration process is directed toward operating states that are more consistent with ADN constraints, reducing unnecessary trial-and-error during training. Overall, P-RL achieves the most stable learning behavior among the three methods, while PI-MADRL provides an intermediate performance. The results

indicate that incorporating both network topology and physical operating information can improve the efficiency and stability of policy training.

4.6 Ablation Analysis of the P-GRL Framework

An ablation study is conducted to quantify the contributions of the two main P-RL components: graph-based feature extraction and physics-informed reward guidance. Four configurations are evaluated using the same network model, control variables, operating scenarios, and training conditions. This provides a direct comparison of the effect of each component on voltage regulation, power loss, and restoration performance.

Table 6: Ablation Analysis of P-RL Components

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

The basic RL configuration gives the lowest minimum voltage and the highest loss and restoration time. Introducing graph learning increases the minimum voltage to 0.967 p.u. and reduces the loss to 138.7 kW, while physics-informed reward guidance produces 0.971 p.u. and 133.4 kW, respectively.

When both components are activated, PGRL achieves a minimum voltage of 0.982 p.u., reduces power loss to 113.8 kW, and completes restoration within 10 min. The results indicate that graph learning improves the representation of network connectivity, whereas the physics-informed reward directs policy exploration toward electrically appropriate actions. Their combination provides complementary benefits for coordinated operation and post-fault restoration.

4.7 Computational Performance

The computational requirements of the learning-based controllers are evaluated separately for offline training and online inference. Training is performed offline, whereas the trained policy is subsequently used to determine control actions from the current ADN state. This distinction is important because a relatively intensive training process does not necessarily prevent real-time application if the inference stage remains sufficiently fast.

The training duration is influenced by the number of learning episodes and model complexity. In contrast, inference requires only state processing and action generation by the trained network. The resulting computational comparison is given in Table 7.

Table 7: Computational performance of the considered methods

Method	Training Time (min)	Stable Convergence	Inference Time (ms)
DRL-based	54.8 ± 1.7	745	7.5 ± 0.4
PI-MADRL	48.2 ± 1.4	605	10.1 ± 0.5
PGRL	44.1 ± 1.2	515	11.8 ± 0.6

P-RL requires approximately 44.1 min for training, compared with 54.8 min for DRL-based learning and 48.2 min for PI-MADRL. This reduction is consistent with the faster convergence observed in Section 4.6. Although PGRL

has a slightly higher inference time because of the graph-processing stage, its average inference time of 11.8 ms remains small relative to typical distribution-network control intervals.

The results indicate that the additional graph-processing overhead is mainly incurred during action inference, while the overall training process benefits from more efficient policy convergence. Thus, the proposed framework provides a reasonable balance between training cost and online computational requirements for adaptive ADN operation.

5. Discussion

The results demonstrate that PGRL improves voltage regulation, power-flow efficiency, and post-fault restoration within a unified control framework. P-RL maintains a minimum voltage of 0.981 p.u., limits voltage deviation to 0.019 p.u., and keeps 99.7% of bus voltages within the prescribed range. It also reduces network loss to 112.9 kW, corresponding to a 47.8% reduction. For self-healing, P-RL restores the affected network within 10 min, limits ENS to 59 kWh, and achieves a resilience index of 0.98. These results are consistent with the topology-aware representation, which enables the controller to adapt its decisions when feeder connectivity changes, while coordinated DG-BESS operation provides additional voltage and power-flow flexibility.

The ablation results further show that graph learning and physics-informed reward contribute separately to performance, while their combination produces the strongest results. PGRL also reaches stable convergence at approximately episode 515, compared with 605 for PI-MADRL and 745 for DRL-based learning. Its inference time of 11.8 ± 0.6 ms indicates that the additional graph-processing requirement remains relatively small for online control. Overall, the results suggest that the combination of topology awareness, physics-informed learning, and coordinated DG-BESS-switching control is the main factor behind the improved ADN performance. Further validation on larger networks and more complex contingencies is still required.

6. Conclusion

The proposed PGRL framework establishes a unified learning-based strategy for resilient and self-healing operation of active distribution networks with renewable generation and BESS. By embedding electrical constraints and feeder topology directly into the learning process, the framework enables coordinated DG, BESS, and switching decisions under changing operating conditions and contingencies. The results demonstrate clear improvements in voltage regulation, power-loss reduction, and post-fault restoration, while maintaining stable learning performance. These findings indicate that combining graph-based network awareness with physical operating knowledge can provide an effective foundation for intelligent and resilient distribution-system control. Future work will extend the framework to larger networks, multiple contingencies, and real-time or hardware-in-the-loop validation.

7. References

- Behzadi S, Bagheri A, Rabiee A. Optimal operation of reconfigurable active distribution networks aiming at resiliency improvement. *Electric Power Systems Research*. 2024; 230:110169. Doi: <https://doi.org/10.1016/j.epr.2024.110169>
- Byeon G, Kim K. Distributionally robust decentralized Volt-Var control with network reconfiguration. *IEEE Transactions on Smart Grid*. 2024; 15(5):4705-4718. Doi: <https://doi.org/10.1109/TSG.2024.3377910>
- Dominguez-Garcia AD, Zholbaryssov M, Amuda T, Ajala O. An online feedback optimization approach to voltage regulation in inverter-based power distribution networks. *IEEE Transactions on Power Systems*. 2024; 39(1):1145-1156. Doi: <https://doi.org/10.1109/TPWRS.2023.3304112>
- Glavic M. Deep reinforcement learning for electric power system control and related problems: A short review and perspectives. *Annual Reviews in Control*. 2019; 48:22-35. Doi: <https://doi.org/10.1016/j.arcontrol.2019.09.008>
- Glavic M, Moreno R, Bi T, Shahidehpour M. Reinforcement learning for electric power system decision and control: Past considerations and perspectives. *IEEE Transactions on Smart Grid*. 2021; 12(6):5128-5142. Doi: <https://doi.org/10.1109/TSG.2021.3070934>
- Huang Y, Yang Q, Tan J, Ukil A. Deep reinforcement learning for real-time energy management in smart grids. *Energy Reports*. 2022; 8:1245-1257. Doi: <https://doi.org/10.1016/j.egyr.2022.08.190>
- Li B, Xu QA. A machine learning-assisted distributed optimization method for inverter-based Volt-VAR control in active distribution networks. *IEEE Transactions on Power Systems*. 2024; 39(2):2668-2681. Doi: <https://doi.org/10.1109/TPWRS.2023.3279303>
- Qiu W, Yadav A, You S, Dong J, Kuruganti T, Liu Y, *et al*. Neural networks-based inverter control: Modeling and adaptive optimization for smart distribution networks. *IEEE Transactions on Sustainable Energy*. 2024; 15(2):1039-1049. Doi: <https://doi.org/10.1109/TSTE.2023.3324219>
- Wang S, Hou Y, Guan X, Liu S, Huo Z. Resiliency-informed optimal scheduling of smart distribution network with urban distributed photovoltaic: A stochastic P-robust optimization. *Energy*. 2025; 313:133449. Doi: <https://doi.org/10.1016/j.energy.2024.133449>
- Xu Y, Dong ZY, Zhang R, Hill DJ. Multi-timescale coordinated voltage/VAR control of high renewable-penetrated distribution systems. *IEEE Transactions on Power Systems*. 2017; 32(6):4398-4408. Doi: <https://doi.org/10.1109/TPWRS.2017.2659779>
- Yin C, Dong J, Zhang Y. Distributionally robust bilevel optimization model for distribution network with demand response under uncertain renewables using Wasserstein metrics. *IEEE Transactions on Sustainable Energy*. 2025; 16(1):118-130. Doi: <https://doi.org/10.1109/TSTE.2024.3509314>
- Zhang B, Cao D, Hu W, Ghias AMYM, Chen Z. 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*. 2024; 155:109641. Doi: <https://doi.org/10.1016/j.ijepes.2023.109641>

13. Zhang C, Xu Y, Dong ZY, Wong KP. Robust coordination of distributed generation and energy storage in distribution networks. *IEEE Transactions on Smart Grid*. 2018; 9(2):1488-1499. Doi: <https://doi.org/10.1109/TSG.2016.2583461>
14. Zhang Z, Wang J, Chen B. Deep reinforcement learning for energy management in microgrids. *Applied Energy*. 2021; 288:116618. Doi: <https://doi.org/10.1016/j.apenergy.2021.116618>
15. Zhao J, Zheng T, Litvinov E. A unified framework for defining and measuring flexibility in power system. *IEEE Transactions on Power Systems*. 2016; 31(1):339-347. Doi: <https://doi.org/10.1109/TPWRS.2015.2390039>
16. Zhou K, Yang S, Shao Z. Digital twin-based optimization for energy systems. *Applied Energy*. 2021; 285:116419. Doi: <https://doi.org/10.1016/j.apenergy.2020.116419>