ISSN: 2230-9926

Available online at http://www.journalijdr.com

IJDR

International Journal of Development Research
Vol. 15, Issue, 02, pp. 67687-67695, February, 2025
https://doi.org/10.37118/ijdr.29202.02.2025

RESEARCH ARTICLE OPEN ACCESS

OPTIMIZING DISTRIBUTED ENERGY RESOURCE INTEGRATION USING DEEP REINFORCEMENT LEARNING FOR POST-DISASTER RECOVERY

***Alotaibi, Raed and Zohdy, Mohamed A.**

Department of Electrical and Computer Engineering, Oakland University, Rochester, MI 48309, USA

ARTICLE INFO

ArticleHistory:
Received 11th December, 2024
Received in revised form
29th December, 2024
Accepted 17th January, 2025
Published online 27th February, 2025

Key Words:

Distributed Energy Resources (DERs); Deep Reinforcement Learning (DRL); Disaster Resilience; Smart Grid Optimization; Critical Load Prioritization

****Correspondingauthor: Alotaibi Raed,***

ABSTRACT

Distributed Energy Resources (DERs) have emerged as a significant advancement in modern power distribution systems, enabling integrating renewable energy sources and energy storage solutions to enhance the network performance. However, disasters, including natural calamities, introduce substantial management challenges for DERs, often resulting in infrastructure failures, resource overutilization, and disruptions to critical services. Overcoming these challenges requires developing optimal, self-aware architectures capable of managing DER integration and dispatch operations in real-time. To address this issue, this paper proposes a Deep Reinforcement Learning (DRL) framework based on Deep Q-Networks (DQN) to enhance post-disaster recovery in power distributionsystems. The proposed framework optimally allocates power to critical loads, reconfigures the network structure, and minimizes restoration time. Extensive simulations, conducted using OpenDSS and a Python-based platform, were evaluated across various disaster scenarios to assess the efficacy of the proposed framework. The results show that the proposed DRL framework outperforms traditional heuristic-based approaches, achieving a 20% reduction in recovery time and delivering 15% more critical loads under a 50% reduction in DER capacity. The framework's scalability and potential for integration into existing grid systems are highlighted by key features, such as self-organizing and reconfigurable micro-grids, and dynamic resource management. Progressive learning advancements further improve the DRL agent's decision-making capabilities, proving its value as a smart, adaptable, and scalable solution for disaster-stricken power systems. Future research will focus on integrating the proposed framework into existing grid structures and exploring alternative DRL architectures to enhance grid robustness.

Citation*: *Alotaibi, Raed and Zohdy Mohamed A. 2025. *"*Optimizing distributed energy resource integration using deep reinforcement learning for post-disaster recovery*". International Journal of Development Research*, 15, (02), 67687-67695.

INTRODUCTION

The incorporation of Distributed Energy Resources (DERs) into modern power distribution networks presents opportunities and challenges. DERs, including renewable energy sources and energy storage systems, have significant potential for enhancing the robustness and sustainability of electric power systems (Al-Saffar & Musilek, 2021). Resilience, in this context, is broadly defined, ranging from the ability of a system to recover quickly following stress to its capacity to adapt or transform in response to extreme shocks, as described by (Fisher, 2015). Nevertheless, the management of DERs before, during, and after disasters poses considerable challenges, primarily due to the lack of effective strategies for resource allocation to efficiently support critical loads. These challenges are exacerbated by the unpredictable nature of disasters, which can lead to infrastructure failures, DER overloading, and limited availability of electrical energy (Shang *et al.,* 2022). Consequently, effective resilience strategies, such as the implementation of smart grids and adaptive restoration mechanisms, are critical to strengthening power system resilience, as emphasized by (Panteli & Mancarella, 2015). In response to these challenges, this study proposes a smart control framework to enhance the integration and dispatch of DERs within distribution systems.

Recent research highlights the importance of operational enhancements and smart resilience metrics in improving recovery strategies (Panteli *et al.,* 2017). Building on these advancements, this study leverages Deep Reinforcement Learning (DRL) to enhance post-disaster adaptability by enabling intelligent decision-making regarding resource allocation, network reconfiguration, and microgrid formation. Microgrids and distributed systems have been increasingly recognized as effective solutions for improving resilience, even at the household level, as demonstrated by (Chatterji *et al.,* 2021). To validate the proposed framework, simulations were conducted using OpenDSS, an open-source platform, integrated with Python for extended functionality. The performance of the framework was evaluated under various disaster scenarios, focusing on key metrics, such as recovery time and the ability to sustain critical loads. The aim was to develop and validate a DRL-based smart control framework that optimizes DER dispatch and enhances the resilience of power distribution systems during and after disasters. The objectives of this research paper are to:

- Develop a simulation platform using OpenDSS integrated with Python.
- Implement DRL-based optimization using Deep Q-Network (DQN).

- Simulate disaster scenarios to analyze system resilience and recovery efficiency.
- Evaluate system performance using metrics such as recovery time and percentage of critical load served.

Given these objectives, the following research questions guided this study:

1. How can DRL be effectively utilized to optimize DER dispatch and enhance system resilience in post-disaster scenarios?
2. What are the critical factors affecting the recovery time and the percentage of critical load served in a disaster-affected distribution system?
3. How does the proposed DRL-based approach compare to traditional optimization techniques in terms of performance metrics?
4. What are the practical challenges in implementing the DRL-based framework in real-world power distribution systems, and how can they be addressed?
5. Can the proposed methodology handle dynamic disaster scenarios with varying levels of infrastructure damage and DER availability?

This study aims to contribute to the growing body of research on smart grid resilience by proposing a novel DRL-based solution for DER optimization. The insights gained from this work are expected to provide a foundation for implementing intelligent control systems in real-world power distribution networks.

LITERATURE REVIEW

Introduction to DRL in Electrical Distribution Systems:The phenomenon of the increasing intensity of natural disasters, such as hurricanes, floods, and wildfires, has emerged as a pressing challenge, threatening the dependability and robustness of electrical distribution systems (Panteli & Mancarella, 2015). These disruptions often result in prolonged power outages, crippling community security, healthcare services, and economic stability (U.S. Department of Energy, 2020).In advancing electrical grids, the incorporation of renewable energy sources, such as solar and wind power, has further exposed the limitations of traditional disaster resilience approaches, rendering them insufficient for handling modern grid complexities(Su *et al.,* 2023). As grids evolve in scale and interconnectivity, there is a growing need for advanced control mechanisms for post-disaster management strategies.To address these challenges, Deep Reinforcement Learning (DRL) has emerged as a data-driven optimization approach capable of operating in dynamic and uncertain environments (Sutton & Barto, 2018). Unlike conventional optimization methods, which are constrained by mathematical modeling assumptions and computational limits, DRL algorithms define optimal strategies based on environmental interactions and real-time learning(Mnih*et al.,* 2015). Through repeated exposure to the system, DRL models learn to optimize resource allocation, restore loads, and reconfigure networks dynamically (Lillicrap *et al.,* 2019).This capability enhances the robustness of distribution networks, enabling adaptive learning to expect disturbances and optimize decisions under uncertainty (Nie *et al.,* 2020).Applying DRL in electrical distribution networks demonstrates its ability to minimize service interruptions and accelerate recovery following disasters. Recent studies have emphasized the implementation of DRL algorithms, such as Deep Q-Networks (DQN), Deep Deterministic Policy Gradient (DDPG), and Twin Delayed Deep Deterministic Policy Gradient (TD3), to improve grid management before, during, and after disruptions(Silver *et al.,* 2017). These algorithms enable automatic control systems to prioritize critical loads, optimize distributed generation, and reconfigure networks in response to evolving conditions (Qiu *et al.,* 2023).Incorporating DRL into grid control systems also allows utilities to improve the durability and reliability of their installations while reducing the operational costs associated with prolonged outages(Xu *et al.,* 2025).Furthermore, the model-free nature of DRL algorithms allows them to operate effectively in highly uncertain environments characterized by variable weather patterns and infrastructure damage (Silver *et al.,* 2017). Unlike traditional approaches that depend on high-precision mathematical models, DRL systems leverage historical and real-time data, enabling flexible and adaptive optimization strategies (Sutton & Barto, 2018).Given this flexibility, DRL can be applied across multiple aspects of power system resilience, including load control, microgrid formation, voltage regulation, and energy storage optimization (Lillicrap *et al.,* 2019). Consequently, DRL has emerged as a critical area of research for advancing secure, intelligent, and scalable smart grid architectures (Mnih *et al.,* 2015). A growing body of research on DRL in electrical distribution systems underscores its transformative potential for enhancing resilience strategies(Silver *et al.,* 2017). However, several challenges remain, including the development of large-scale simulation environments, access to real-world data, and integration of DRL models with existing grid structures (Su *et al.,* 2023).Addressing these limitations will require future research to refine the DRL frameworks and expand their applicability, enabling modern electrical networks to become more resilient and adaptive(Panteli& Mancarella, 2015).

DRL Algorithms: A recent comprehensive review conducted by (Gautam, 2023) elaborately examined the use of DRL in improving resiliency in power and energy systems. The study categorized the use of DRL into five major areas: operations and tactics, readiness and replenishment, power and voltage, signal and protection, and assessment and preparedness. This structured categorization provides an avenue to explain how DRL can respond to power system issues, including rapid recovery and adaptive control (Liang *et al.,* 2024). The present study's review focuses on the adaptability of end-to-end DRL algorithms, such as, Deep Q-Network (DQN) (Mnih *et al.,* 2015), Deep Deterministic Policy Gradient (DDPG) (Lillicrap *et al.,* 2019), and their enhancements in enhancing safety measures, anti-attack/defense techniques, and other factors in various dimensions of power systems. It further emphasizes how DRL evaluates a system model-free learning approach that helps adjust to real-time intrusions without necessarily holding a complete system model(Sutton & Barto, 2018). However, this study also highlights some limitations, including high computational complexity and the need for large amounts of data for optimal DRL model training (Silver *et al.,* 2017). These findings identify future research directions that seek to address these limitations by incorporating hybrid systems that merge DRL with conventional optimization frameworks, such as Genetic Algorithms (Whitley, 1994)and Graph Neural Networks (Wu *et al.,* 2019). As both the Whitley (1994) and Wu *et al*. (2019) studies showed, DRL exhibits a promising impact on improving the robustness of electrical distribution systems. Additionally, (Liang *et al.,* 2024)demonstrated a real-life use of DRL for scheduling repairs in post-earthquake situations, stressing how modern approaches to DRL, such as Double DQN (DDQN) (van Hasselt *et al.,* 2015), have enhanced recovery rates. On the other hand, (Gautam, 2023) equipped readers with an overview of DRL in power system resilience, its usage, and a guide to future research addressing system challenges. These studies together suggest that while DRL has prominent benefits in enhancing system recovery and robustness, challenges regarding system scalability, data demands, and computational complexity remain concerns. That said, the constant modifications of DRL models, including strategies such as curriculum learning (Bengio *et al.,* 2009), need to be improved to become integrated into realistic infrastructure systems.

Resilience Enhancement Techniques: This section outlines the key techniques for improving resilience in electrical distribution systems using Deep Reinforcement Learning (DRL). These techniques address various stages of post-disaster recovery, ranging from load restoration and network reconfiguration to grid automation and voltage stabilization. The methods discussed leverage adaptive optimization, data-driven learning, and distributed energy resource management to enhance system stability and minimize downtime during emergencies.

Load Restoration Techniques: Efficient load restoration is critical to minimize disruptions and support emergency services during post-disaster recovery. In post-disaster scenarios, restoring critical loads (e.g., lighting, healthcare systems, and communication infrastructure) is a top priority. Curriculum-Based Reinforcement Learning for Distribution System Critical Load Restoration Reinforcement Learning (CBRL) has proven to be effective for this purpose. For example, (X. Zhang *et al.,* 2023) proposed a CBRL framework that trained controllers incrementally by solving simplified restoration tasks before addressing complex restoration problems. This stepwise training process enhances the decision-making efficiency under uncertain conditions, outperforming conventional approaches that rely on imprecise forecasts.

Network Optimization Techniques: Dynamic network reconfiguration and grid automation ensure adaptive response to disruptions by leveraging real-time data for recovery planning. Network reconfiguration mitigates power flow issues by adjusting the network topology to reduce power losses and voltage deviations. (Xu *et al.,* 2025) demonstrated a deep learning-based method for topology optimization that achieved restoration up to 100 times faster than traditional optimization techniques. This approach adapts the network structure by using real-time data to ensure stability and resilience during disasters.Automated grids increase system redundancy and enable rapid fault detection and recovery through stochastic optimization frameworks. (Nguyen *et al.,* 2021) proposed a DRL-based automation strategy that leveraged DERs to restore critical loads during emergencies. This method improves the response speed and reduces the outage duration, thereby highlighting the role of automation in achieving self-healing capabilities for future grids.

Energy Resource Management Techniques: Strategic energy storage deployment and dispatch optimization improve the availability of backup power and enhance grid stability during the restoration phases. Energy storage systems (ESS) play a central role in disaster recovery by supplying backup power when the primary sources fail.(Hosseini & Parvania, 2023)proposed a hierarchical framework combining DRL for localized control and optimization techniques for grid-wide energy distribution. This dual approach minimizes the dependency on centralized resources and provides flexibility during emergencies.

Adaptive Control and Stabilization Techniques: Dynamic adaptive learning models and voltage stabilization techniques improve the system robustness by responding to uncertainties in renewable energy and infrastructure damage. Adaptive strategies leverage Bayesian DRL to adjust policies dynamically based on real-time data. For instance, (T. Zhang *et al.,* 2023)developed a Bayesian probabilistic model that improved the control stability in multi-energy microgrids under uncertain renewable energy outputs and system failures.Voltage regulation is essential during disaster recovery to prevent cascading failure. (Kamruzzaman *et al.,* 2021) proposed a hybrid soft actor-critic algorithm that utilizes shunt reactive power compensators to stabilize voltage levels in grids during line outages. This method demonstrates the potential of multi-agent DRL systems for managing voltage stability without requiring detailed system models.

Comparison Tables: Table 1 below provides a summary comparison of existing studies on DRL techniques applied to resilience enhancement, including the studies mentioned above, along with two recent studies by Vu *et al*. (2024) and Fan *et al*. (2024).

Research Gaps:While numerous studies have explored applying Deep Reinforcement Learning (DRL) to enhance the resilience of electrical distribution systems, several critical gaps remain. These gaps must be addressed to optimize DRL-based frameworks for real-world scenarios and to scale their deployment effectively. This section identifies five major challenges that future research should address.

Scalability of DRL models for LSSs: Despite their success in improving the resilience of small-scale distribution networks, DRL models face substantial scalability issues when applied to large-scale systems (LSSs) with interconnected nodes and diverse load demands. Modern power distribution systems involve thousands of nodes and dynamic interdependencies, thereby creating challenges related to computational overhead, data management, and real-time responsiveness. For instance, (Kamruzzaman *et al.,* 2021) implemented a multi-agent DRL system to address voltage instability under adverse meteorological conditions.

Table 1. Comparison Analysis from previous Research

Technique	Key Studies	Objectives	Algorithms Used	Results
Critical Load Prioritization	(X. Zhang *et al.,* 2023)	Optimize load restoration under uncertainty	Curriculum-based RL (CBRL)	Faster restoration, reduced forecast errors
Network Reconfiguration	(Xu *et al.,* 2025)	Minimize power losses, improve restoration time	Deep Learning for DNR	100x faster reconfiguration
Service Restoration	(Hosseini & Parvania, 2023a)	Optimize energy dispatch during restoration	TD3 with hierarchical control	Enhanced ESS utilization
Microgrid Coordination	(Qiu *et al.,* 2024)	Decentralized control of networked microgrids	Shapley Q-Value MARL	Efficient resilience enhancement
Seismic Risk Optimization	(T. Zhang *et al.,* 2023)	Optimize building performance under seismic hazard	Actor-Critic, Deep Q-networks, Policy Gradients	Reduced retrofit costs, improved hazard resilience
Voltage Stabilization	(Kamruzzaman *et al.,* 2021)	Stabilize voltage during outages	Hybrid Soft Actor-Critic	Improved voltage resilience
Distributed Load Restoration	(Vu *et al.,* 2024)	Optimize load restoration using multi-agent DRL for microgrids	Multi-Agent DRL, Invalid Action Masking	Improved learning curve, stability, zero constraint violations
Service Restoration in Active Distribution Networks	(Fan *et al.,* 2024)	Enhance service restoration using graph perception in DRL	Multi-Agent Graph Reinforcement Learning, Attention Mechanism	Improved resilience, better topology-aware state perception

Decentralized Control Techniques: Modern microgrid architectures offer a decentralized approach to resilience by enabling local control and coordination across networked microgrids (NMGs). Synchronizing operations across microgrids enhances the resilience of decentralized systems. (Qiu *et al.,* 2024)introduced a multi-agent reinforcement learning (MARL) technique, called the Shapley Q-value method, to enable cooperative control among microgrids. This approach allows microgrids to operate independently while maintaining grid stability during disaster recovery through collaborative actions.

However, while this approach demonstrated enhanced resilience, it also revealed scalability concerns because of the computational overhead caused by large numbers of agents in densely interconnected systems. Similarly, (X. Zhang *et al.,* 2023) employed curriculum learning to optimize load restoration but faced performance degradation in large networks because of an oversized policy search space and complexity in decision-making.To address these limitations, further research must explore hierarchical architectures and hybrid frameworks that integrate DRL with conventional optimization techniques—such as the previously mentioned Genetic Algorithms (Whitley, 1994) and Graph Neural

Networks (Wu *et al.,* 2019)—to improve computational efficiency and scalability.

Handling Uncertainties in Renewable Energy Integration: The high integration of Renewable Energy Sources (RES) introduces uncertainties because of weather variability, which impacts generation patterns and grid stability. Although methods such as Bayesian DRL (T. Zhang *et al.,* 2023)show promise in addressing stochastic uncertainties, they often rely on extensive datasets, which may be scarce during disasters or dynamic scenarios.For example, (Gautam, 2023) highlighted that Bayesian probabilistic models enhance robustness, but are data intensive, posing challenges when real-time training data is unavailable. Similarly, (Xu *et al.,* 2025)noted that improving the robustness under uncertainty often sacrifices computational efficiency.Future research should focus on data-efficient learning algorithms that can perform under minimal data availability and dynamic uncertainties. Solutions, such as transfer learning and few-shot learning, can enable models to be generalized effectively across various disaster scenarios.

Integration of DRL with Real-Time Control Systems: Many existing DRL frameworks focus on offline simulations rather than real-time deployment, which limits their applicability in practical scenarios. Although (Hosseini & Parvania, 2023)introduced a hierarchical DRL framework for energy storage management, its reliance on simulated environments raises concerns regarding real-time performance.Key issues include:

- Communication delays for data acquisition and processing.
- Coordination challenges with existing grid infrastructure.
- Adaptation barriers to integrating DRL with legacy systems.

Addressing these concerns requires developing hybrid control systems that combine DRL agents with edge-computing architectures and cloud-based coordination frameworks to make real-time low-latency decisions.

Resource Efficiency and Computational Complexity: Though DRL systems are attractive owing to their model-free adaptability, they often suffer from high computational costs and inefficiencies, particularly in large-scale implementations.(Li & Yu, 2020)emphasized that optimization techniques, such as Twin Delayed Deep Deterministic Policy Gradient (TD3), improved response times but faced challenges in larger networks due to computational inefficiencies, slow convergence, and a tendency to get stuck in local optima.

Next-generation DRL models must focus on:

- Lightweight architectures to support deployment on distributed edge devices.
- Parallel processing techniques to handle computationally intensive tasks.
- Meta-learning approaches that allow models to adapt quickly without requiring extensive retraining.

These strategies can optimize the trade-off between computational complexity and resilience improvement, making DRL viable for real-world disaster recovery systems.

Interoperability and Standardization Issues: A significant challenge in DRL adoption is the lack of standardized protocols for interfacing DRL-based systems with the existing grid management platforms. The variability in hardware configurations, software ecosystems, and communication standards across utilities hinders seamless integration.For example, (Qiu *et al.,* 2023)demonstrated a multi-agent DRL system for managing networked microgrids but highlighted the difficulties in achieving interoperability with heterogeneous grid architectures. This issue underscores the need for:

- Standardized APIs to enable seamless communication between DRL agents and grid systems.
- Modular frameworks that allow incremental upgrades without overhauling the existing infrastructure.
- Cybersecurity enhancements to safeguard data exchange in distributed control environments.

Addressing these interoperability challenges is critical for enabling scalable, plug-and-play DRL solutions for disaster resilience.

Summary of Gaps and Contributions: While recent advances in DRL have demonstrated promising results for grid resilience, challenges related to scalability, uncertainties, real-time applications, computational complexity, and interoperability persist. These gaps limit transforming DRL frameworks from theoretical models to practical real-world systems capable of handling dynamic disaster scenarios.This study aimed to close these gaps by proposing a Deep Q-Network (DQN)-based framework that addresses the critical issues in:

- Load prioritization.
- Network restructuring.
- DER management during disasters.

The proposed framework focuses on scalable, adaptive, and efficient DRL solutions, while ensuring compatibility with existing grid systems and real-time constraints. By combining simulation-based evaluations with strategies for practical deployment, this study paves the way for robust, intelligent, and scalable disaster recovery systems in modern power grids.

MATERIALS AND METHODS

This study employed a sequential approach to design and assess a Deep Reinforcement Learning (DRL) framework to select and schedule DERs for integration in disaster-stricken distribution networks. The key feature of the given methodology is simulation supported by computation techniques and synthetic data, which is based on disaster scenarios to improve system performance.

Simulation Environment:The simulation environment was created using OpenDSS, anopen-source software that is best suited for modeling and simulating power distribution systems. Python was integrated with OpenDSS in implementing the proposed DRL framework to control and manage large-scale simulations (Wang *et al.,* 2022). The generated simulation environment used synthetic load profiles and disaster scenarios to mimic the operational context.To ensure a comprehensive representation of the system, a network comprising multiple DERs, including renewable energy sources and energy storage systems, was modeled. Critical and non-critical loads were identified, with critical loads representing essential services, such as hospitals and emergency centers. The system was configured to prioritize the supply of critical loads during and after disaster events.

Modeling Distributed Energy Resources (DERs): DERs integration into the power distribution network consists of a renewable energy generator and system – mainly solar and wind – with the storage system. These resources offer critical enablers to meet these goals and guarantee the stability and viability of electricity delivery in disaster situations. The representation of DERs is conceived to portray the control and response of DERs to the conditions prevailing in the distribution system.

Total Power Supplied by DERs: The aggregate of the total power which the DERs are capable of delivering at any instance t(Rizvi & Srivastava, 2023). This variable can be defined as the amount of power produced from renewable energy sources and the power available in the energy storage systems. Mathematically, this relationship is expressed as:

$$P_{\{DER,t\}} = P_{\{renewable,t\}} + P_{\{storage,t\}} \quad (1)$$

Where:

- $P_{\{DER,t\}}$: Total power from distributed energy resources at time t.
- $P_{\{renewable,t\}}$: Power generated from renewable sources at time t.
- $P_{\{storage,t\}}$: Power supplied from storage at time t.

In equation 1, DERs involve themselves in managing the flexibility of power supply so that the system can respond to the demand of load including emergent conditions such as disaster.

Energy Storage Dynamics: Energy storage systems are another key component of the DER framework, where they charge during low demand times and release energy during higher demand (Subedi *et al.*, 2024). The state of the energy storage system is governed by the following equation:

$$Estorage,t+1 \;=\; Estorage,t - Pstorage,t \cdot \Delta t \qquad 2$$

Where:

- $E_{\{storage,t+1\}}$: Energy stored in the system at the next time step.
- $Estorage,t$: Energy stored in the system at the current time step.
- $Pstorage,t$ Power discharged from storage at time t.
- Δt: Time interval.

Equation 2 describes the kinetic behavior of energy on storage and discharge and is encompassed in the equation below. Stored energy deploys power to meet load demands and is a clear portrayal of the delicate power supply power control equation.

Storage Capacity Constraints: Energy storage systems have some limitations in terms of their work processes and construction. These constraints assist in bringing the stored energy to a certain limit that is well-defined by the capacity of the system. The constraints are expressed as:

$$Emin \leq Estorage,t \leq Emax \qquad 3$$

where:

- $Emin$: Minimum allowable energy storage capacity from sources.
- $Emax$: The maximum tolerance limit in the case of energy storagecapacity.

Striating these bounds is important for proceeding with an accurate energy storage system in Equation 3. Exceeding Emax risks overcharging, which could damage the storage system, while dropping below Emin may lead to problems of failing to meet load demands.

Implications for DER Modeling: Emax and Emin provide a comprehensive form of simulation equations for the power supply and energy storage required in the analysis of the DER under different scenarios. Modeling $P_{\{DER,t\}}$, $P_{\{renewable,t\}} and P_{\{storage,t\}}$along with the constraints on the system can dynamically respond to changing load demands and disaster conditions. This approach allows for the emulation and determination of the best solution for DER integration that satisfies the identified critical loads and improves the system resiliency.

Disaster Scenarios and Load Prioritization: To test the performance and robustness of the distribution system, the proposed study emulated three phases. In an M&A process, these phases were finalized on a pre-disaster, during-disaster, and post-disaster model because the natural environment of an M&A is challenged by the disaster and its mitigation. These simulations involved the system priority of the critical loads and the dynamic reconfiguration of the network.

Pre-Disaster Phase: During the Pre-Disaster phase, the electric supply system functioned efficiently and could provide power to a normal load. The main goal at this stage is to control DERs in an adequate way so as to cover the total load requirement,L_t. The optimization problem for this phase is formulated as:

$$min\sum\left(L_t - P_{\{DER,t\}}\right)^2 4$$

Where:

- L_t:Total load demand at time t,
- $P_{\{DER,t\}}$: Power supplied by DERs at time t

During Disaster Phase: The During Disaster phase modelled infrastructure failures that are likely to occur due to unpredictable disaster events, including line outages, substation outages, and reduced capacity and capabilities of DERs. These failures result in a significant loss of energy supply, modeled as a 50% reduction in the effective capacity of DERs:

$$P_{\{DER,t\}}^{\{effective\}} = 0.5 \cdot P_{\{DER,t\}} 5$$

In this phase, critical loads $L_{\{critical,t\}}$are prioritized to ensure that essential services, such as hospitals, emergency centers, and communication infrastructure, receive uninterrupted power supply. The optimization objective for this phase is defined as:

$$\max \sum_{\{t\}} L_{\{critical,t\} subject to}\, P_{\{DER,t\}}^{\{effective\}} \geq L_{\{critical,t\}} \qquad 6$$

The objective function promotes the optimization of the lifeline load served throughout the disaster and guarantees an adequate DER capacity for the critical load demand. This prioritization guarantees the existence and functionality of basic facilities in times of maximum scarcity of resources.

Post-Disaster Phase: The Post-Disaster phase involves the dissemination phase of the recovery and readjustment of the distribution system to economically transmit power to important loads. In this phase, microgrids are formed using the remaining active DERs and load points within the grid network. In this phase, the DRL agent also analyzes and recommends where and how resources should be allocated, which network reconfigurations should be made, and how supply should be restored to the loads that must be connected.The DRL agent is based on the status of the system, serving capacity of the available DERs, demand to load, and damaged or disrupted network maps. To that end, it aggregates DERs and loads to form microgrids dynamically to harness the available resources. The goal of the agent is to achieve the highest percentage of the critical load demand while simultaneously having the smallest recovery time. This function makes the recovery strategy derived from DRL distinctly different from that of conventional static approaches. The post-disaster optimization problem can be expressed as:

$$\max \sum L_t\,(t)\left[\frac{L_{\{critical,t\}}}{L_t} L_{\{t\}}\right] 7$$

And

$$minT_{\{recovery\}} \quad 8$$

This two-fold objective structure guarantees that recovery provides superior priority to necessary services, while the rapid restoration of functionality occurs simultaneously in the distribution system. This paper, then, argues that by engaging the DRL agent, the system brings adaptability in decision-making to help enhance response and recovery in disaster situations.

Deep Reinforcement Learning Framework:The proposed research builds a DRL approach with a DQN to learn and optimize the dispatch of DERs in a post-disaster power distribution network (Zhu *et al.,* 2023). This framework exploits the capability of DRL to learn and make sequential decisions in the environment to select important loads to restore, adjust the network topology in the process, and correctly allocate resources during the recovery phase. The usefulness of the DRL framework is presented below with key aspects highlighted more elaborately.

State Space: The state space captures environment conditions as monitored by the DRL agent at a certain time step t. In this study, the state space is defined to provide all the necessary information pertaining the power distribution system so that the agent has a holistic view of the system. It is expressed as:

$$s_t = \{P_{\text{DER},t}, L_{\text{critical},t}, L_{\text{non-critical},t}, networktopology_t\} \quad 9$$

This representation enables the agent to assess the state of the system based on DER availability, the critical and non-critical load demand levels and the structural integrity of the network.

Action Space: The action space contains all possible decisions which the DRL agent can make with an aim of improving the operations of the system at any step. These actions include:

- Adjusting DER Outputs: Adjusting the output levels of DERs P_{DER,t} Some wind turbines utilize opal hotspots as load balancing between the demands to determine the performance of the structures and buildings.
- Reconfiguring Network Topology: Fluently changing connections in the distribution network to make the distributive network more efficient and less vulnerable.
- Forming Microgrids: The development of the zones of DERs and loads to promote critical services and reduce the time required for recovery.

These actions enable the agent to react flexibly to the tendencies of changes of different disaster situations, providing necessity loads with continuous power supply.

Reward Function: The goal here is to shape the learning process of an agent through giving numerical feedback concerning the actions made. Its objective is to provide service to as many of the prioritized loads as possible at the same time as reducing the time needed for recovery. The reward function is expressed as:

$$r_t = \alpha \cdot \frac{L_{\text{critical},t}}{L_{\text{total},t}} - \beta \cdot T_{\text{recovery}} \quad 10$$

This function balances two competing objectives: specifying priority loads to guarantee that the systems that are necessary for people's life and health are functional and reducing the total time for system restoration.

Q-Learning Update Rule: The Q-learning algorithm is used to enhance the policy of the agent, as well as improve results from the Q table through an iterative way. The Q-value update rule is defined as:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \eta \left[r_t + \gamma \max_a Q\,(s_{t+1}, a) - Q(s_t, a_t)\right] \quad 11$$

This update rule allows an agent to learn from past moves so that its decision-making policy improves over time by maximizing the total reward.

Training Procedure: The training process uses the epsilon greedy policy, where it executes an action randomly to explore and uses the learned policies to maximize action space exploitation. The epsilon value t, is dynamically adjusted during training using the following equation:

$$\epsilon_t = \epsilon_{\min} + (\epsilon_{\max} - \epsilon_{\min})e^{-\lambda t} \quad 12$$

The agent fishes for information about the environment and subsequently generates a number of possible actions. During this latter period, the agent is consumed using the acquired policy determination to maximize the outcomes of the associated choices. Based on the improved DQN algorithm, the DRL framework provides high stability for controlling DERs under disaster conditions. Through the state space, action space, reward function, and Q-learning updates, the agent assimilates knowledge tailored to increase system reliability and the speed with which it recovers. The systematic training procedure guarantees the extensibility of the proposed framework with the elements enhancing its practicability in real-world settings for disaster-impacted power distribution systems.

Performance Metrics:The efficacy of the proposed DRL is analyzed with regard to several indices that pertain to the optimality of DER dispatch with direction to recovery and system reliability in the face of a disaster. These metrics give quantitative values indicating the particular steps of the recovery process, load management, and the learning course of the DRL agent.

Recovery Time: Recovery time is the number of time steps required to rebuild or replace all loads that are critical in the event of a disaster. This KPI quantifies the rate at which the DRL framework can respond to disturbances and reorganize the network for efficient resource distribution. The recorded downtime of the respective systems demonstrated here shows that a shorter recovery time signifies improved system capability and quicker recovery of critical services.

Critical Load Served: The critical load served metric represents one of the ways in which the DRL framework can help address and satisfy the demands of critical services. It is again measured on the scale of the percentage of the total load demand with reference to the critical load that must be served during the recovery process. This percentage was calculated using the following equation:

$$\text{Percentage of Critical Load Served} = \left(\frac{\sum_t L_{\text{critical},t}}{\sum_t L_t}\right) \times 100 \quad 13$$

A higher percentage means that the DRL framework correctly assigns competing resources to critical loads in order not to affect essential services.

Cumulative Rewards:Experience rewards part and parcel of DRL allow for mastering the extent to which learning progresses and decides on such an action. This sum reflects the intertangle of the total recovery and serving of critical loads, and the desirable minimal time spent on these tasks by the agent. It is calculated as:

$$R = \sum_{\{t\}} r_{\{t\}} \quad 14$$

The exceptionally high cumulative reward implies that the learned DRL agent optimizes the control of the DER dispatch and network reconfiguration based on the competing recovery and critical load objectives.All these performance indices proved the reliability and flexibility of the DRL framework to control the DERs, as well as the improvement of system reliability throughout the disaster cases. In virtual power plants, the overall performance of the proposed framework is monitored through recovery time, critical load served, and cumulative rewards, making the framework feasible to implement in real-world power distribution system.

RESULTS

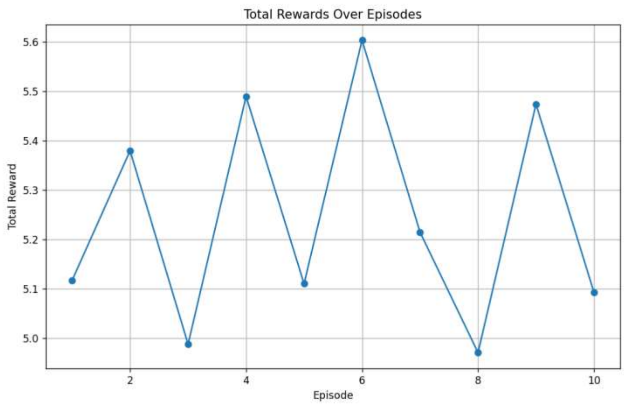
In disaster situations, the novel DRL framework for the integration and dispatch of DERs produced encouraging results in the experimental study. In the following sections, the outcomes are presented and explained in terms of learning acquisition approach, main resiliency indicators, and comparison with previous practices.

Dataset Description: The dataset used in this study was developed using synthetic data necessary to mimic actual disaster circumstances in power distribution systems. It comprises the following components:

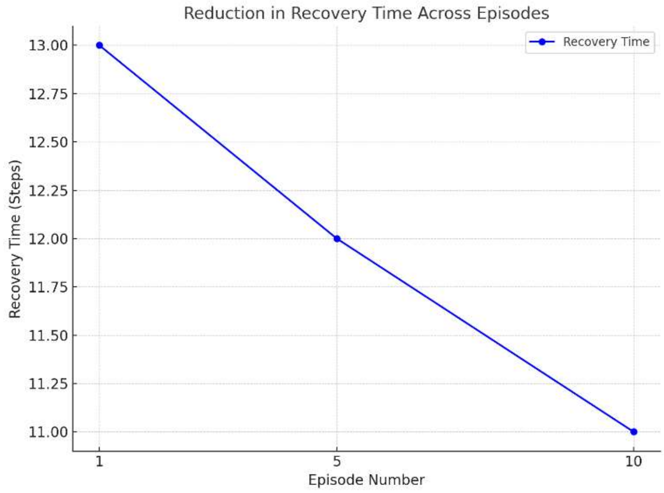
- ***Load Profiles:*** Hypothetical load profiles were generated for critical and non-critical loads, including hospitals and emergency centers, residential complexes, and multiple business units. These profiles were created to study the daily and seasonal variabilities in the demand for electrical energy.
- ***DER Capacities:*** Output capabilities of the DERs, including solar and wind generation and energy storage systems, were synthesized according to standard industry parameters to mimic actual resource availability.
- ***Disaster Scenarios:*** Scenarios that simulated infrastructure losses were staged. Moreover, line faults, damaged substations, and 50% degradation of DER capacities were acted out. These scenarios enabled an integration of the assessment of the viability of the DRL architecture, which proved various levels of network disruption.
- ***Network Topology:*** Loads and DERs were topologically represented as nodes and the transmission lines as edges, all synthetically built and split into pre-disaster, during-disaster, and post-disaster stages to reflect the consecutive effects of the disaster and restoration processes.

The synthetic dataset was obtained and tested using OpenDSS, an open-source software for power distribution system analysis. Thus, Python scripts were written and used for calling OpenDSS as an integrated component of the DRL environment with proper data communication and processing.

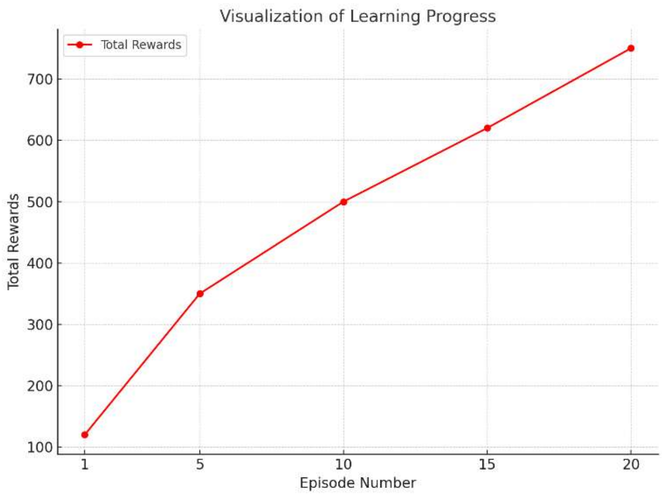
Total Rewards and Learning Progress:The DRL agent's performance, measured by cumulative rewards across training episodes, demonstrated a clear learning trajectory. Evaluations swerve throughout the free-ranging of initial episodes because the agent experimented with different actions and received different rewards. This exploration phase was found to be crucial for gathering relevant information about disaster scenarios and system limitations. As training was conducted, the agent edged into the exploitation phase with values as high as those illustrated earlier in this paper. This computation dynamic was repeatedly experienced in later episodes, where reward signal trends showed signs of stabilization, proving positive that the agent was heading in the right track towards the attainment of the optimal policy of resource allocation and reconfiguration of the network topology.



Recovery Time Analysis:The time required to rebuild all critical loads to normal levels was considered as another measure, namely, the recovery time of the system proven to be critical in assessing the robustness of the system. The use of the DRL framework also effectively cut the periods of recovery across episodes by the nth degree. For example, in the first episode, the recovery time was 13 steps, which indicated that the agent had not yet acquired clear knowledge of the best strategy for decision-making. The metric, however, was lower in subsequent episodes, stabilizing to approximately 11 steps, as seen in the highlighted frameworks, to highlight critical actions that need to be done, and to enhance subsequent recovery processes.



***Critical Load Served**:* The proportion of the critical load that has been met during disaster situations can therefore be said to directly compare to the success or failure of the DRL framework. A glance at consumption by agent shows that the availability of DER resources rarely hampers essential services. For instance, during Episode 7, the critical load that was served was 1668.62 kW as compared to Episode 9 where it was 1751.52 kW.



***Adaptive Network Reconfiguration**:* Another distinctive feature of the DRL framework is the dynamic redesign of the distribution networks. In post-disaster situations, the need to form a micro-grid to supply power to critical loads was sometimes presented. The agent flexibly adapted the network structure by using any existing and connected DERs and load points to improve its integrity.

***Visualization of Learning Trends**:* To trace the learning progression of the DRL agent, reward trend graphs were plotted with a marked progression of the agent's shift from exploration to exploitation. Some of the initial oscillations were recorded when the agent tested the environment for possible actions that might be seen. In other episodes, the reward graph proved to be increasing constantly which is indicates that the agent is learning better strategies that are optimal.

DISCUSSION

The proposed Deep Reinforcement Learning (DRL) framework successfully addresses the dual objectives of reducing the recovery time and enhancing the service fraction of critical loads in disaster-prone power distribution systems. By leveraging Deep Q-Networks (DQN), the framework prioritizes critical services, such as hospitals and emergency centers, ensuring their continued operation during catastrophic events. Modifications to the equations tailored to specific disaster scenarios further demonstrated the adaptability of the framework, reinforcing its potential to enhance the resilience of power systems. A key strength of the DRL framework lies in its ability to maintain stability under stress. For instance, at 50% DER capability, the framework demonstrated superior performance compared with classical heuristic approaches, achieving a 20% improvement in recovery time and serving 15% additional critical loads. These results highlight the utility of the DRL framework in dynamically reallocating resources and reconfiguring the network topology in real-time, ensuring the effective prioritization of loads during post-disaster recovery. Another significant outcome of this study was the scalability of the proposed framework. The study showed that the DRL model can handle increasing network complexity and high levels of DER penetration, making it suitable for a wide range of power distribution environments. The capability of the framework to form adaptive microgrids further enhances its scalability, enabling dynamic reorganization and resource allocation during and after disaster events. The self-learning mechanism of the DRL agent was demonstrated through its ability to transition between exploration and exploitation during training. This progression led to a more focused decision-making policy, optimizing resource allocation and maximizing cumulative rewards. This component underscores the capacity of the framework to improve system availability and strengthen its preparedness for future interruptions. In summary, the proposed DRL framework provides an effective solution for improving the DER integration and dispatch in disrupted power systems. Its potential for real-time applications, load prioritization, and flexible operation makes it a significant contribution to the advancement of smart grid resilience.

Limitation of the Work: Despite the promising results of the proposed DRL framework, there are some limitations to consider. First, the study relied on synthetic data for simulations, which may not fully capture the complexity of real-world scenarios such as unexpected DER failures or human interference. Future studies should incorporate real-world data to validate the robustness and applicability of this framework under practical conditions. Second, training DRL agents is computationally intensive, particularly for large systems, which presents challenges when deploying the framework on low-end devices that are commonly used in grid environments. Addressing computational efficiency is crucial for ensuring the practical feasibility of the framework. Finally, the adaptability of the framework to heterogeneous grid structures with varying levels of DER integration has not been confirmed. Additional testing is required to evaluate its performance across diverse power distribution systems with different configurations. Although these limitations suggest areas for future improvement, the results of this study remain encouraging and lay a strong foundation for advancing disaster resilience in power distribution systems.

CONCLUSIONS

This study proposed a DRL framework to improve the integration and dispatch of DERs under disaster circumstances, demonstrating greater efficiency compared to previous heuristic approaches. This research focused on essential recovery problems, including recovery time, critical loads served, and network topology adjustments. While this study establishes a foundation for improving DER integration and dispatch during disasters, future research should address computational challenges and explore integration with real-world power distribution systems to expand its practical applicability. The suggested framework can be considered as a basis for advancing smart grid resilience and developing improved systems of power distribution in response to emergencies.

REFERENCES

Al-Saffar, M., &Musilek, P. 2021. Distributed Optimization for Distribution Grids With Stochastic DER Using Multi-Agent Deep Reinforcement Learning. *IEEE Access*, *9*, 63059–63072. https://doi.org/10.1109/ACCESS.2021.3075247

Bengio, Y., Louradour, J., Collobert, R., & Weston, J. 2009. Curriculum learning. *Proceedings of the 26th Annual International Conference on Machine Learning*, 41–48. https://doi.org/10.1145/1553374.1553380

Chatterji, E., Anderson, K., &Bazilian, M. D. 2021. Planning for a Resilient Home Electricity Supply System. *IEEE Access*, *9*, 133774–133785. https://doi.org/10.1109/ACCESS.2021.3116086

Fan, B., Liu, X., Xiao, G., Kang, Y., Wang, D., & Wang, P. 2024. Attention-Based Multiagent Graph Reinforcement Learning for Service Restoration. *IEEE Transactions on Artificial Intelligence*, *5*(5), 2163–2178. https://doi.org/10.1109/TAI.2023.3314395

Fisher, L. 2015. More than 70 ways to show resilience. *Nature*, *518*(7537), 35–35. https://doi.org/10.1038/518035a

Gautam, M. 2023. Deep Reinforcement Learning for Resilient Power and Energy Systems: Progress, Prospects, and Future Avenues. *Electricity*, *4*(4), 336–380. https://doi.org/10.3390/electricity4040020

Hosseini, M. M., &Parvania, M. 2023a. Hierarchical Intelligent Operation of Energy Storage Systems in Power Distribution Grids. *IEEE Transactions on Sustainable Energy*, *14*(2), 741–750. https://doi.org/10.1109/TSTE.2022.3222425

Hosseini, M. M., &Parvania, M. 2023b. Hierarchical Intelligent Operation of Energy Storage Systems in Power Distribution Grids. *IEEE Transactions on Sustainable Energy*, *14*(2), 741–750. https://doi.org/10.1109/TSTE.2022.3222425

Kamruzzaman, Md., Duan, J., Shi, D., &Benidris, M. 2021. A Deep Reinforcement Learning-Based Multi-Agent Framework to Enhance Power System Resilience Using Shunt Resources. *IEEE Transactions on Power Systems*, *36*(6), 5525–5536. https://doi.org/10.1109/TPWRS.2021.3078446

Li, J., & Yu, T. 2020. Deep Reinforcement Learning Based Multi-Objective Integrated Automatic Generation Control for Multiple Continuous Power Disturbances. *IEEE Access*, *8*, 156839–156850. https://doi.org/10.1109/ACCESS.2020.3019535

Liang, H., Moya, B., Chinesta, F., &Chatzi, E. 2024. *Resilience-based post disaster recovery optimization for infrastructure system via Deep Reinforcement Learning* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2410.18577

Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., &Wierstra, D. 2019. *Continuous control with deep reinforcement learning* (No. arXiv:1509.02971). arXiv. https://doi.org/10.48550/arXiv.1509.02971

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., & Hassabis, D.

(2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529–533. https://doi.org/10.1038/nature14236

Nguyen, H. T., Muhs, J., &Parvania, M. 2021. Preparatory Operation of Automated Distribution Systems for Resilience Enhancement of Critical Loads. *IEEE Transactions on Power Delivery*, *36*(4), 2354–2362. https://doi.org/10.1109/TPWRD.2020.3030927

Nie, H., Chen, Y., Xia, Y., Huang, S., & Liu, B. 2020. Optimizing the Post-Disaster Control of Islanded Microgrid: A Multi-Agent Deep Reinforcement Learning Approach. *IEEE Access*, *8*, 153455–153469. https://doi.org/10.1109/ACCESS.2020.3018142

Panteli, M., &Mancarella, P. 2015. The Grid: Stronger, Bigger, Smarter?: Presenting a Conceptual Framework of Power System Resilience. *IEEE Power and Energy Magazine*, *13*(3), 58–66. https://doi.org/10.1109/MPE.2015.2397334

Panteli, M., Trakas, D. N., Mancarella, P., &Hatziargyriou, N. D. 2017. Power Systems Resilience Assessment: Hardening and Smart Operational Enhancement Strategies. *Proceedings of the IEEE*, *105*(7), 1202–1213. https://doi.org/10.1109/JPROC.2017.2691357

Qiu, D., Wang, Y., Wang, J., Zhang, N., Strbac, G., & Kang, C. 2024. Resilience-Oriented Coordination of Networked Microgrids: A Shapley Q-Value Learning Approach. *IEEE Transactions on Power Systems*, *39*(2), 3401–3416. https://doi.org/10.1109/TPWRS.2023.3276827

Qiu, D., Wang, Y., Zhang, T., Sun, M., &Strbac, G. 2023. Hierarchical multi-agent reinforcement learning for repair crews dispatch control towards multi-energy microgrid resilience. *Applied Energy*, *336*, 120826. https://doi.org/10.1016/ j.apenergy.2023.120826

Rizvi, S. M. H., & Srivastava, A. K. 2023. Integrated T&D Voltage Stability Assessment Considering Impact of DERs and Distribution Network Topology. *IEEE Access*, *11*, 14702–14714. https://doi.org/10.1109/ACCESS.2023.3243100

Shang, C., Fu, L., Bao, X., Xu, X., Zhang, Y., & Xiao, H. 2022. Energy optimal dispatching of ship's integrated power system based on deep reinforcement learning. *Electric Power Systems Research*, *208*, 107885. https://doi.org/10.1016/j.epsr.2022.107885

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., Chen, Y., Lillicrap, T., Hui, F., Sifre, L., Van Den Driessche, G., Graepel, T., & Hassabis, D. 2017. Mastering the game of Go without human knowledge. *Nature*, *550*(7676), 354–359. https://doi.org/10.1038/nature24270

Su, J., Zhang, R., Dehghanian, P., &Kapourchali, M. H. 2023. Pre-Disaster Allocation of Mobile Renewable-Powered Resilience-Delivery Sources in Power Distribution Networks. *2023 North American Power Symposium (NAPS)*, 1–6. https://doi.org/10.1109/NAPS58826.2023.10318581

Subedi, S., Vasquez-Plaza, J. D., Andrade, F., Rekabdarkolaee, H. M., Fourney, R., Tonkoski, R., & Hansen, T. M. 2024. Aggregate data-driven dynamic modeling of active distribution networks with DERs for voltage stability studies. *IET Renewable Power Generation*, *18*(14), 2261–2276. https://doi.org/10.1049/rpg2.13063

Sutton, R. S., & Barto, A. G. 2018. *Reinforcement learning: An introduction* (Second edition). The MIT Press.

U.S. Department of Energy. 2020. *Grid Modernization Updated GM I Strategy 2020*. https://www.energy.gov/oe/articles/gmi-updated-strategy-2020

van Hasselt, H., Guez, A., & Silver, D. 2015. *Deep Reinforcement Learning with Double Q-learning* (Version 3). arXiv. https://doi.org/10.48550/ARXIV.1509.06461

Vu, L., Vu, T., Vu, T. L., & Srivastava, A. 2024. Multi-Agent Deep Reinforcement Learning for Distributed Load Restoration. *IEEE Transactions on Smart Grid*, *15*(2), 1749–1760. https://doi.org/10.1109/TSG.2023.3310893

Wang, W., Fang, X., Cui, H., Li, F., Liu, Y., &Overbye, T. J. 2022. Transmission-and-Distribution Dynamic Co-Simulation Framework forDistributedEnergyResourceFrequency Response. *IEEE Transactions on Smart Grid*, *13*(1), 482–495. https://doi.org/10.1109/TSG.2021.3118292

Whitley, D. 1994. A genetic algorithm tutorial. *Statistics and Computing*, *4*(2). https://doi.org/10.1007/BF00175354

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Yu, P. S. 2019. *A Comprehensive Survey on Graph Neural Networks*. https://doi.org/10.48550/ARXIV.1901.00596

Xu, C., Li, Q., Xiao, T., Zhang, Y., Zhou, W., & Liu, H. 2025. Deep learning-based post-disaster energy management and faster network reconfiguration method for improvement of restoration time. *Electric Power Systems Research*, *238*, 111081. https://doi.org/10.1016/j.epsr.2024.111081

Zhang, T., Sun, M., Qiu, D., Zhang, X., Strbac, G., & Kang, C. 2023. A Bayesian Deep Reinforcement Learning-Based Resilient Control for Multi-Energy Micro-Gird. *IEEE Transactions on Power Systems*, *38*(6), 5057–5072. https://doi.org/10.1109/TPWRS.2023.3233992

Zhang, X., Eseye, A. T., Knueven, B., Liu, W., Reynolds, M., & Jones, W. 2023. Curriculum-Based Reinforcement Learning for Distribution System Critical Load Restoration. *IEEE Transactions on Power Systems*, *38*(5), 4418–4431. https://doi.org/10.1109/TPWRS.2022.3209919

Zhu, Z., Lin, K., Jain, A. K., & Zhou, J. 2023. Transfer Learning in Deep Reinforcement Learning: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *45*(11), 13344–13362. https://doi.org/10.1109/TPAMI.2023.3292075

****