RELIABILITY-ORIENTED MICROGRID CONTROL USING REINFORCEMENT LEARNING

O. Aliyeva[1], I. Babazade[2]

•

[1,2]Azerbaijan State Oil and Industry University, Scientific-Research İnstitute of Geo-technological Problems of Oil, Gaz and Chemistry, Baku, Azerbaijan

[1] olga_da@mail.ru , [2]ikram.babazade@asoiu.edu.az

Abstract

The growing penetration of renewable energy sources (RES) in modern power systems introduces significant challenges to reliability due to the inherent stochasticity of wind and solar generation. Microgrids, as decentralized energy systems, offer resilience and flexibility but require intelligent control strategies to balance variable generation and demand. This paper proposes a reinforcement learning (RL)-based multi-agent control framework for microgrids, formalized through a Markov Decision Process (MDP) with states reflecting generation, storage, and load variability. Deep RL algorithms—DQN, PPO, and A2C—are compared against conventional methods in a high-fidelity simulation environment with real-world load and RES data. Key reliability indicators, including Loss of Load Probability (LOLP), Mean Time Between Failures (MTBF), and the self-sufficiency ratio, are evaluated. Results demonstrate that PPO achieves a >40% reduction in LOLP compared to rule-based control, ensuring higher resilience and operational adaptability. The study further discusses computational constraints, generalization issues, and cyber-physical vulnerabilities, and proposes solutions based on federated RL, transfer learning, and anomaly detection. The findings underline RL's potential as a cornerstone for reliability-centered energy management in microgrids.

Keywords: reinforcement learning, microgrid reliability, renewable energy uncertainty, PPO, DQN, A2C, multi-agent systems

I. Introduction

The modern power industry is undergoing a transformation toward sustainable, intelligent, and decentralized energy architectures. A central driver of this transformation is the large-scale integration of renewable energy sources (RES), particularly wind and solar generation. While environmentally beneficial, the variability and intermittency of RES introduce significant uncertainty into power generation processes, directly challenging the reliability of electricity supply.

Microgrids—localized systems that combine distributed generation, storage, and controllable loads—are widely recognized as a promising solution to enhance resilience and autonomy [1, 2]. They can operate in both grid-connected and islanded modes, offering flexibility under disturbances. However, ensuring reliable performance of microgrids under high RES penetration is non-trivial. Fluctuations in generation, coupled with stochastic consumer demand, increase the probability of supply-demand imbalances, energy losses, and reliability degradation [3, 4].

Conventional optimization approaches, such as linear or stochastic programming, have limited ability to cope with such stochasticity. These methods typically rely on complete prior knowledge of system parameters and are not robust to incomplete or noisy data [5]. As a result, there is a growing interest in intelligent, adaptive control strategies. Reinforcement Learning (RL) is particularly attractive because it allows agents to learn optimal control strategies through continuous interaction with the environment, without requiring explicit system models [6].

DOI: https://doi.org/10.24412/1932-2321-2025-489-339-345

For microgrid operation, RL offers several advantages:

- learning directly from real-time system feedback,
- adaptability to dynamically changing conditions,
- robustness under incomplete or uncertain data, and
- scalability to high-dimensional, multi-agent control tasks [7].

Recent advances in Deep Reinforcement Learning (DRL) further extend these capabilities, enabling control strategies in nonlinear and complex environments [8].

Despite the growing body of research on RL in energy systems, a gap remains in linking RL methods directly to reliability indicators of microgrids. Most studies emphasize control efficiency or cost minimization, while reliability metrics such as Loss of Load Probability (LOLP), Mean Time Between Failures (MTBF), and self-sufficiency ratio are often overlooked.

The goal of this study is to address this gap by developing and evaluating a reinforcement learning-based framework for microgrid operation explicitly focused on reliability enhancement. The study investigates the performance of modern DRL algorithms—DQN, PPO, and Actor-Critic—under realistic operating conditions with variable RES and demand, and compares their ability to improve microgrid reliability relative to conventional approaches.

II. Problem statement

The integration of renewable energy sources (RES), such as photovoltaic (PV) and wind, into microgrids contributes to sustainable development but simultaneously creates critical challenges for ensuring reliability. Electricity generation from RES is inherently stochastic, while consumer demand exhibits both temporal and behavioral variability. This dual uncertainty significantly increases the probability of supply-demand imbalance, load losses, and overuse of backup units, which in turn reduces overall reliability of the microgrid.

From the reliability engineering perspective, the main negative effects can be summarized as follows: increased Loss of Load Probability (LOLP) due to mismatches between generation and demand, reduced Mean Time Between Failures (MTBF) of supply adequacy, and lower self-sufficiency ratio (SSR) as microgrids rely more frequently on external grid support. Furthermore, intensified cycling of storage systems and backup generators leads to accelerated wear and premature failures [9-11].

Conventional optimization methods, including linear programming (LP) and stochastic programming, are poorly suited for such environments. They rely on complete prior knowledge of system parameters, do not scale well in real time, and rarely integrate probabilistic reliability indices into the decision-making process. Table 1 highlights the limitations of traditional methods compared to reinforcement learning (RL), which uniquely allows direct optimization of reliability metrics.

Table 1: *Comparison of optimization methods for reliability in microgrids under uncertainty*

Method	Uncertainty Resilience	Real-Time Operation	Reliability Metrics Integration	Scalability
Linear Programming (LP)	Low	No	Not included	Low
Stochastic Optimization	Medium	Partial	Limited	Medium
Reinforcement Learning	High	Yes	Direct (LOLP, MTBF, SSR)	High

To better illustrate the challenge, Figure 1 presents a conceptual microgrid architecture where each subsystem (PV, wind turbine, storage, loads, diesel generator) is managed by an autonomous RL agent. These agents interact locally and cooperate globally to maintain reliability even under stochastic disturbances.

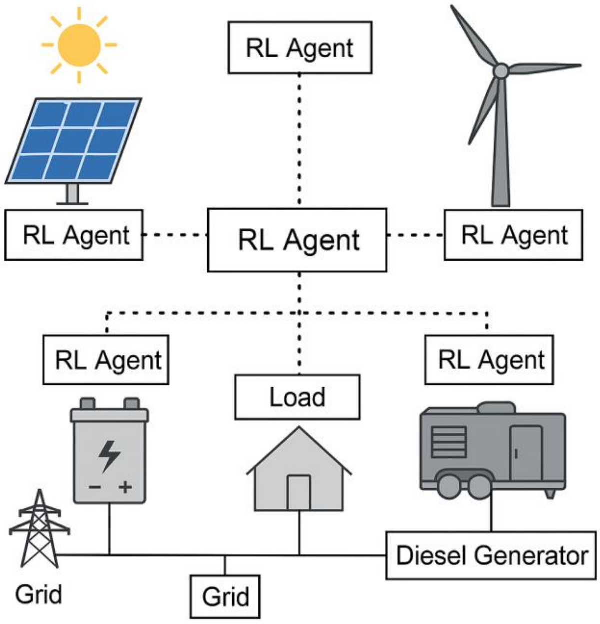

Figure 1: *Conceptual microgrid architecture with RL agents*

The reliability-centered optimization loop is further explained in Figure 2. The RL agent receives uncertain inputs (RES generation and demand fluctuations), processes them through its policy network, and outputs actions aimed at minimizing LOLP, increasing MTBF, and improving SSR. This representation makes explicit how reliability indicators are embedded into the optimization process, rather than treated as secondary performance measures.

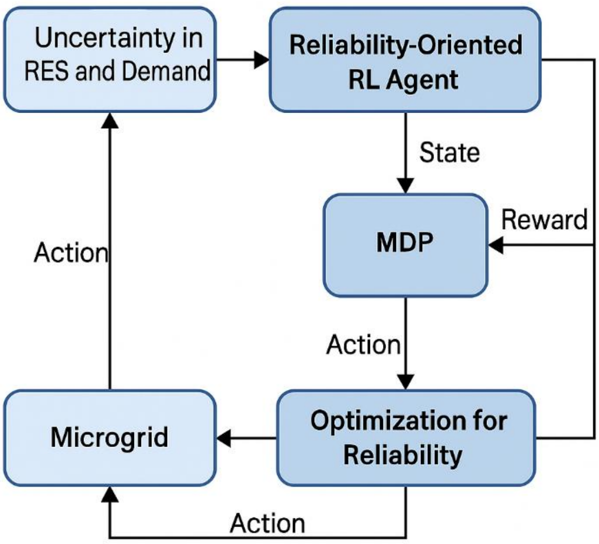

Figure 2: *Reliability-oriented optimization loop in RL*

Finally, to demonstrate the practical impact of renewable penetration on reliability, Figure 3 compares the evolution of LOLP under traditional and RL-based control strategies. While conventional methods exhibit a sharp increase in LOLP once RES penetration exceeds 40%, the RL-based strategy maintains significantly lower values across the full range.

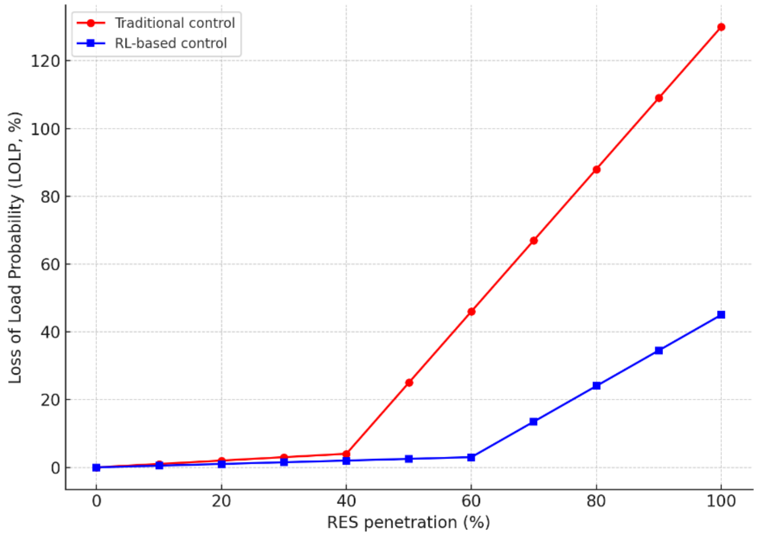

Figure 3: *Impact of RES penetration on LOLP under traditional vs. RL-based control*

III. Problem Solution

To overcome the reliability challenges identified in Section II, a reinforcement learning (RL)-based control framework was developed for microgrid operation. The problem was formalized as a Markov Decision Process, where states include renewable generation, load, and storage conditions, while actions represent redispatching storage, activating backup generation, and managing power exchange with the grid. The reward function is explicitly reliability-oriented, minimizing interruptions that increase the Loss of Load Probability (LOLP) while rewarding stable operation that extends the Mean Time Between Failures (MTBF) and improves the self-sufficiency ratio (SSR).

Three deep RL algorithms were investigated: Deep Q-Network (DQN), which is suited for discrete switching tasks such as charge/discharge control; Proximal Policy Optimization (PPO), which provides robustness and fast convergence in continuous control environments; and Advantage Actor–Critic (A2C), which combines policy and value learning for balanced adaptability. Their main characteristics are summarized in Table 2, showing that PPO is particularly well suited for reliability-centered optimization in microgrids with high renewable penetration [12, 13].

The experimental environment reproduced a representative microgrid including a 50 kW photovoltaic array, a 30 kW wind turbine, a 200 kWh lithium-ion storage system, a 60 kW diesel generator, and a ±40 kW grid connection. Load data corresponded to a hybrid residential–commercial consumer group with a daily consumption of about 850 kWh. Renewable and demand time series were taken from CAISO and NREL datasets, and RL agents were trained over 5000 daily episodes.

Table 2: *Comparison of reinforcement learning algorithms for microgrid reliability optimization*

Algorithm	Model Type	Continuous Action Handling	Convergence	Stability	Hyperparameter Sensitivity	Typical Use
DQN	Value-based	Limited	Moderate	Medium	High	Discrete ESS switching
PPO	Policy gradient	Well suited	High	High	Low	Continuous hybrid energy control
A2C	Actor–critic	Suitable	Medium	Medium	Moderate	Demand–supply balancing

The learning dynamics of the algorithms are presented in Figure 4, which compares the evolution of LOLP during training. PPO converges to a stable policy after approximately 1000 episodes, reducing LOLP by more than 40% compared to a rule-based baseline. A2C also improves reliability but at a slower rate, while DQN shows more variability due to its discrete-action design. These results confirm that RL, and especially PPO, provides a scalable and adaptive solution for enhancing microgrid reliability under stochastic renewable generation and demand uncertainty.

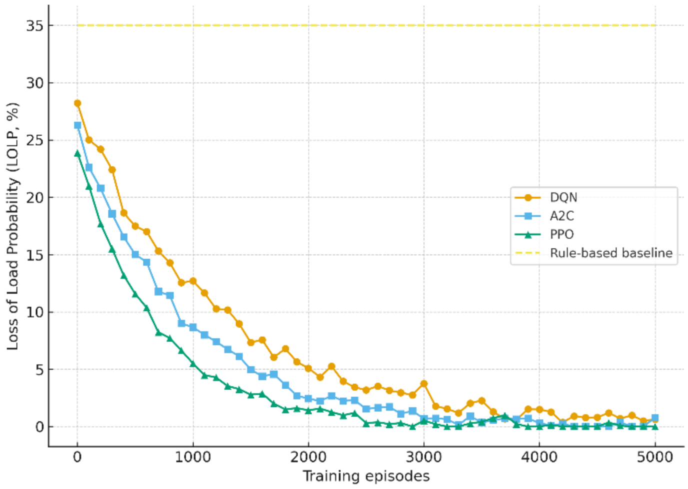

Figure 4: *Convergence of RL algorithms in reducing LOLP during training*

IV. Discussion

The results obtained in Section III confirm that reinforcement learning (RL) can significantly improve the reliability of microgrid operation under uncertain renewable generation and variable demand. Compared with traditional optimization approaches discussed in Section II, RL not only adapts to stochastic fluctuations but also directly optimizes reliability metrics. This shift is critical because reliability—expressed through indicators such as LOLP, MTBF, and SSR—is the primary concern in decentralized power systems with high shares of renewable energy.

The comparative evaluation of algorithms presented in Table 2 shows that PPO consistently outperforms DQN and A2C in terms of stability and convergence. While DQN remains effective for discrete control actions, its sensitivity to hyperparameters and limited scalability make it less attractive for complex microgrid environments. A2C demonstrates balanced performance but requires longer training to reach reliable policies. PPO, in contrast, exhibits robust convergence and strong adaptability, which makes it particularly well suited for continuous and hybrid control scenarios.

The training dynamics in Figure 4 highlight the reliability benefits of PPO. Within approximately 1000 training episodes, PPO reduces LOLP by more than 40% compared to rule-based control, achieving stable operation much earlier than the other RL algorithms. This finding is especially relevant for practical deployment, where long training periods are often infeasible. Faster convergence directly translates into more reliable operation during the early stages of learning, which is a decisive advantage for real-world microgrids.

At the same time, several limitations remain. First, the computational cost of training complex RL agents is significant, especially when scaling to multi-agent architectures. Second, generalizing trained policies to different microgrid topologies or load patterns is not trivial, which may limit transferability. Finally, the increasing autonomy of RL agents introduces new risks related to cybersecurity, as malicious interventions in control signals could compromise system reliability.

These challenges point to promising research directions. Distributed training frameworks and cloud-based computation can reduce the cost of training. Transfer learning and meta-RL approaches can improve adaptability across diverse microgrid configurations. In addition, incorporating anomaly detection mechanisms and decentralized trust protocols can mitigate cyber-physical vulnerabilities. By addressing these aspects, future RL frameworks will not only enhance operational reliability but also ensure resilience against both technical and external threats.

Overall, the discussion confirms that reinforcement learning is more than a control alternative—it is a transformative approach for microgrid management, shifting the focus from efficiency alone to explicit reliability-centered optimization. The results of this study provide strong evidence that RL, and particularly PPO, can become a practical cornerstone technology for the next generation of intelligent, decentralized energy systems.

V. Conclusion

This study has presented a reinforcement learning (RL)-based framework for improving the reliability of microgrid operation under renewable energy uncertainty. Unlike conventional optimization methods that struggle with incomplete information and rapidly changing conditions, the proposed approach integrates reliability indicators directly into the decision-making process. By formulating the problem as a Markov Decision Process with a reward function centered on LOLP, MTBF, and SSR, the framework ensures that control strategies are explicitly aligned with reliability objectives.

The results demonstrate that deep RL algorithms, and especially Proximal Policy Optimization (PPO), significantly reduce the probability of load loss while improving fault-free operational intervals and system self-sufficiency. As shown in Table 2 and Figure 4, PPO outperforms DQN and A2C by combining robust convergence with adaptability to continuous control tasks. This advantage translates into faster attainment of reliable policies and superior resilience under high penetration of renewables.

The broader implication of these findings is that reinforcement learning is not only an efficient optimization tool but also a transformative method for reliability-centered management of decentralized energy systems. By embedding reliability directly into optimization objectives, RL-based frameworks can provide adaptive, scalable, and autonomous control strategies that meet the demands of next-generation power systems.

At the same time, several challenges remain to be addressed. High computational costs, limited policy generalization across diverse microgrid topologies, and the vulnerability of autonomous agents to cyber-physical threats highlight the need for further research. Promising directions include federated and meta-RL approaches for improving adaptability, integration of forecasting and Model Predictive Control to strengthen hybrid decision-making, and the development of anomaly detection and blockchain-based trust mechanisms to ensure resilience against external disruptions.

In conclusion, reinforcement learning has the potential to become a cornerstone technology for the reliable operation of smart microgrids. By explicitly linking control strategies to reliability indices, it offers a pathway toward resilient, sustainable, and adaptive energy systems capable of meeting the challenges of the renewable-dominated future.

CONFLICT OF INTEREST

The authors declare no potential conflict of interests.

References

[1] Lund, H., Ostergaard, P. A., Connolly, D., & Mathiesen, B. V. (2017). Smart energy and smart energy systems. *Energy*, 137, 556–565.

[2] Farrokhabadi, M., Cañizares, C. A., & Bhattacharya, K. (2017). Frequency control in isolated/islanded microgrids through voltage regulation. *IEEE Transactions on Smart Grid*, 8(3), 1185–1194.

[3] Katiraei, F., & Iravani, M. R. (2006). Power management strategies for a microgrid with multiple distributed generation units. *IEEE Transactions on Power Systems*, 21(4), 1821–1831.

[4] Shahidehpour, M., Yamin, H., & Li, Z. (2002). *Market operations in electric power systems*. Wiley-IEEE Press.

[5] Bahramirad, S., Reder, W., & Khodaei, A. (2012). Reliability-constrained optimal sizing of energy storage system in a microgrid. *IEEE Transactions on Smart Grid*, 3(4), 2056–2062.

[6] Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.

[7] Wang, Y., Chen, Q., Hong, T., & Kang, C. (2019). Review of smart meter data analytics: Applications, methodologies, and challenges. *IEEE Transactions on Smart Grid*, 10(3), 3125–3148.

[8] Meng, L., Savaghebi, M., Andrade, F., Vasquez, J. C., & Guerrero, J. M. (2017). Review on control of DC microgrids and multiple microgrid clusters. *IEEE Journal of Emerging and Selected Topics in Power Electronics*, 5(3), 928–948.

[9] Rzayeva, S. V., Piriyeva, N. M., & Ismayilova, S. I. (2025). High and low voltage coordination in electrical power systems. *International Journal on Technical and Physical Problems of Engineering (IJTPE)*, 17(1), 19–31.

[10] Rzayeva, S. V., Piriyeva, N. M., & Guseynova, I. A. (2024). Analysis of reliability of typical power supply circuits. *Reliability: Theory & Applications*, 19(3), 173–178.

[11] Khan, A. A., Naeem, M., Iqbal, M., Qaisar, S., & Anpalagan, A. (2016). A compendium of optimization objectives, constraints, tools and algorithms for energy management in micro-grids. *Renewable and Sustainable Energy Reviews*, 58, 1664–1683.

[12] Wouters, C., Fraga, E., & James, A. (2015). MILP approach for the design of residential microgrids with energy interactions restrictions. *Computer Aided Chemical Engineering*, 37, 2357–2362.

[13] Delgado, C., & Domínguez-Navarro, J. A. (2014, March 25–27). Optimal design of a hybrid renewable energy system. In *Proceedings of the 9th International Conference on Ecological Vehicles and Renewable Energies (EVER)*. Monte-Carlo, Monaco.