

Deep Reinforcement Learning and NSGA-II hybrid model Optimization of Reactive Power on a power transmission system: Case Study of the IEEE 14-bus network

Sam, Bassey Asuquo¹

Department of Electrical /Electronic Engineering,
University of Uyo, Akwa Ibom State, Nigeria
greatbasam@gmail.com

Njoku Chukwudi Aloziem²

Department of Computer Engineering,
University of Uyo,
Akwa Ibom State, Nigeria

Glory Bassey Sam³

Department of Computer Engineering,
University of Uyo, Akwa Ibom State, Nigeria

Abstract Deep Reinforcement Learning and Non-Dominated Sorting Genetic Algorithm II (NSGA-II) hybrid model optimization of reactive power on a power transmission system: Case Study of the IEEE 14-bus network is presented. The resultant model from the DRL and NSGA-II is referred to as the DRL + NSGA-II hybrid model. The DRL + NSGA-II hybrid model introduces a multi-objective optimization strategy that enhances the energy management capabilities of conventional DRL models. While DRL learns policies through interaction with the environment, NSGA-II facilitates Pareto-optimal selection of dispatch configurations, simultaneously minimizing both active and reactive losses. The results show that the slack generator, Bus 1, maintains the highest active power surplus at 91.899 MW. This trend continues with moderate surpluses across: Bus 2: 43.974 MW, Bus 3: 38.878 MW, Bus 6: 33.591 MW, Bus 8: 33.685 MW. On the deficit side (negative values of P_{mismatch}), buses 4 to 14 show remarkably consistent levels of power demand, with most values clustered tightly around -29 MW . This consistency indicates that the hybrid optimization strategy has successfully pushed the power system toward global load balance, with minimal deviation among buses. In all the results indicate that the hybrid optimization strategy has successfully pushed the power system toward global load balance, with minimal deviation among buses. Also, the results show that the DRL + NSGA-II hybrid model provided well-balanced dispatch strategy, devoid of congestion or over-saturation on critical links, which indicates effective exploration of state-action spaces by the

DRL agent and optimal tuning from the NSGA-II layer.

Keywords — Optimization of Reactive Power, Deep Reinforcement Learning (DRL), IEEE 14-Bus Network, Non-Dominated Sorting Genetic Algorithm II (NSGA-II), Power Transmission System

1. Introduction

As the both population grows in every nation across the globe, there comes growing demand for energy to support the growing population [1,2,3]. Also, the advancements in technologies have also increased the dependence on energy dependent technologies [4,5]. This makes it more imperative that efficient power generation, transmission and distribution systems are required to satisfy the growing demand [6,7].

In this work, the focus is on addressing the challenges that emanate from imbalance in the reactive power in the transmission network [8,9]. The reactive power problem can lead to high power loss, poor voltage profile and also instability in the power system [10]. The solution to reactive power control has been on for several years. However, in recent years, machine learning models and other intelligent algorithms have been widely adopted in power system solutions development [11,12]. In this work, Deep Reinforcement Learning and Non-Dominated Sorting Genetic Algorithm II (NSGA-II) hybrid model is presented for optimizing the reactive power in a transmission network [13,14]. In the DRL + NSGA-II hybrid model, the NSGA-II which is a multi-objective evolutionary algorithm, complements the learning

dynamics of Deep Reinforcement Learning (DRL) by guiding the model toward a Pareto-optimal frontier that balances active and reactive power mismatch across the grid.

2. Methodology

The study presents an approach to reactive power optimization for the IEEE 14-bus network using a combination of the Deep Reinforcement Learning (DRL) and Non-Dominated Sorting Genetic Algorithm II (NSGA-II) models. The hybrid model derived from the integration of the DRL and NSGA-II algorithm is referred to as the DRL + NSGA-II model. In this case, the NSGA-II (Non-dominated Sorting Genetic Algorithm II), a multi-objective evolutionary algorithm, complements the learning dynamics of Deep Reinforcement Learning (DRL) by guiding the model toward a Pareto-optimal frontier that

balances active and reactive power mismatch across the grid.

The system model for the hybrid intelligent mechanism for optimizing the reactive power in distributed grid networks is as shown in Figure 1. The methodology is structured into three interconnected phases: (i) development of predictive models using advanced machine learning technique based on the Deep Reinforcement Learning (DRL) model; (ii) implementation of standalone metaheuristic optimization algorithms based on Non-dominated Sorting Genetic Algorithm II (NSGA-II); and (iii) integration of predictive learning with optimization strategies to enhance grid stability and control accuracy. Each phase is designed to contribute incrementally to the overall goal of achieving efficient and dynamic reactive power management under varying grid conditions.

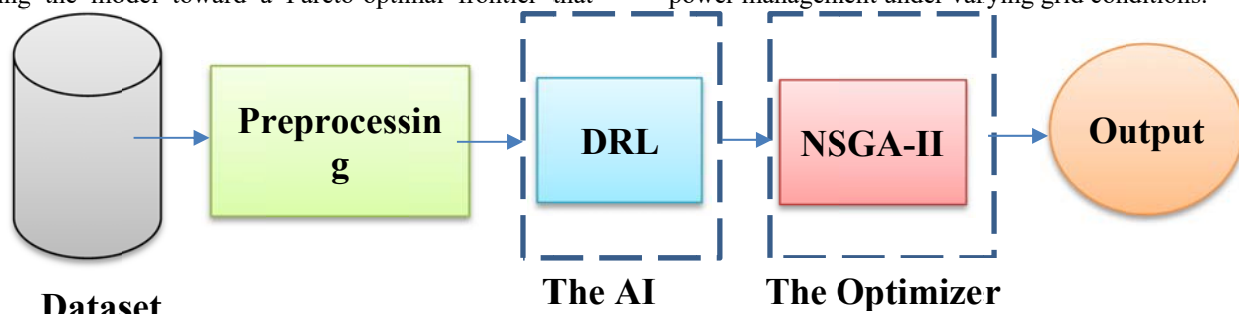


Figure 1: The system model

2.1 Dataset Pre-processing

The effectiveness of learning-based reactive power optimization mechanisms depends critically on the quality and structure of input data. One of the data processing action performed is the missing or noisy data handling. In real-world power system data, missing or anomalous readings can distort optimization outcomes. Let the raw dataset be denoted as:

$$\mathcal{D} = \{(x_i, y_i)\}_{i=1}^N \quad (1)$$

Where, $x_i \in R^d$ is the feature vector (e.g., voltage V , active power P , reactive power Q), $y_i \in R$ is the target quantity at timestamp i . The missing values $x_i(j) = NaN$ are interpolated as:

$$x_i(j) = x_{i-1}(j) + \frac{x_{i+1}(j) - x_{i-1}(j)}{2} \quad (2)$$

If prior data is available, then forward fill can be applied as:

$$x_i(j) = x_{i-1}(j) \quad (3)$$

The noise $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$ in time series is mitigated using a Savitzky–Golay filter, denoted as:

$$\hat{x}_i = \sum_{k=-w}^w c_k x_{i+k} \quad (4)$$

Where, w is the window size and c_k are smoothening polynomial coefficients.

To ensure learning stability and convergence across optimization models, features are normalized to comparable scales. Let $x^{(j)}$ represent j^{th} feature, then min-max normalization is given as:

$$x_i^{(j)'} = \frac{x_i^{(j)} - \min(x^{(j)})}{\max(x^{(j)}) - \min(x^{(j)})}, \quad x_i^{(j)'} \in [0, 1] \quad (5)$$

Where, $x_i^{(j)'}$ is the normalized j^{th} feature at timestamp i . The normalized features include: Voltage magnitudes $V_i \in [0.95, 1.05]$, Power Demand $P_{D,i}$, $Q_{D,i}$, and Power Generation $P_{G,i}$, $Q_{G,i}$. This transformation is essential for LSTM input gates and DRL network gradients.

2.2 The Deep Reinforcement Learning (DRL) Model

The diagram of the framework for the DRL model is shown in Figure 2. The framework in Figure 2 identifies the agent and the environment as key components that engage in iterative process of capturing State of the environment, formulating control Action and Reward in response to the State and reward policy. The Action performed alters the State which is updated and the process repeats.

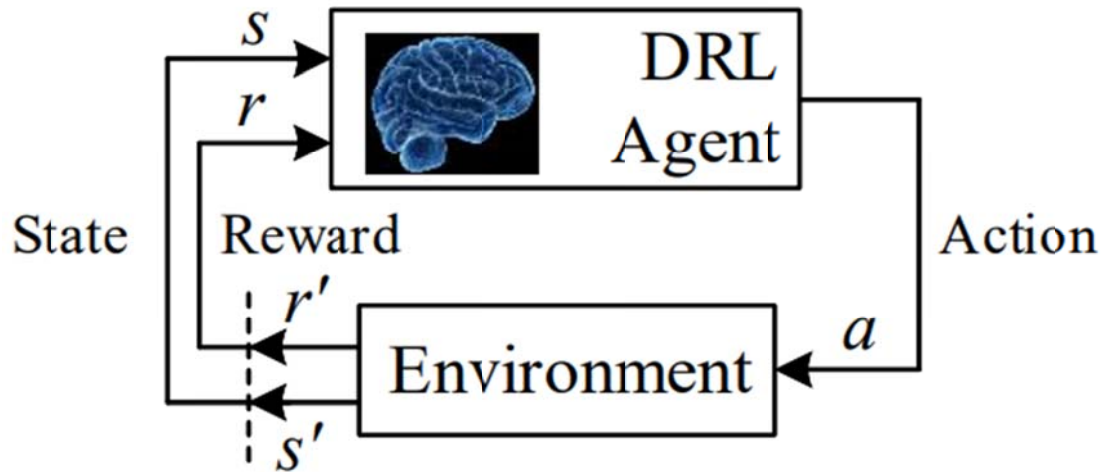


Figure 2 The diagram of framework for the DRL Model

Specifically, the key components and procedure of the DRL model can be explained as follows:

- i. **The environment:** The environment represents a dynamic system with parameters that need to be monitored and control. In this case the power network like the IEEE 14-bus network is the dynamic system or the physical system with parameters that are being monitored and controlled by the DRL.
- ii. **The agent:** This is what carries out the continuous monitoring of the environment. It interacts with the dynamic system as well as collect requisite information and also passes information to the system
- iii. **The state:** This is the set of information that describe the condition or state of the environment. The agent collects the State of the environment.
- iv. **The action:** The agent provides the control signal which specifies the control action to be carried out in response to the received state of the environment and the reward policy or function. After performing the action, new state is generated and the corresponding reward. With new set of state and reward, the agent updates the DRL framework and continues the process again (iteratively) until the desired result is obtained.

2.3 The Deep Reinforcement Learning (DRL) and Non-Dominated Sorting Genetic Algorithm II (NSGA-II) Hybrid Model

In modern smart grids, the reactive power control problem is inherently multi-objective. Operators must balance conflicting objectives such as: minimization of power losses (to reduce operational cost and thermal stress); voltage profile regulation (to ensure stability and reliability); and minimization of control effort or capacitor switching frequency. The traditional single-objective

optimization fails to capture these trade-offs. Thus, a Multi-Objective Evolutionary Algorithm (MOEA) is employed to obtain a Pareto-optimal set of non-dominated solutions. The MOEA approach used in this work is NSGA-II which stands for Non-dominated Sorting Genetic Algorithm II. In this case, the DRL learns control policies $\pi(s) = a$ from experience, while the NSGA-II performs offline multi-objective tuning of

- i. Action selections or ΔQ vectors
- ii. Reward shaping parameters
- iii. Policy hyperparameters

Specifically, the DRL learns value function $Q(s, a; \theta)$ and extracts policy candidate $\{a_i\}$ from policy network, then finally define multi-objective function over each a_i as;

$$\left. \begin{aligned} f_1(a_i) &= \text{Simulated power loss from applying } a_i \\ f_2(a_i) &= \text{Voltage deviation induced by } a_i \\ f_3(a_i) &= \text{Control effort (e.g. capacitor operation)} \end{aligned} \right\} \quad (6)$$

Hence, the NSGA-II filter optimal action is taken as:

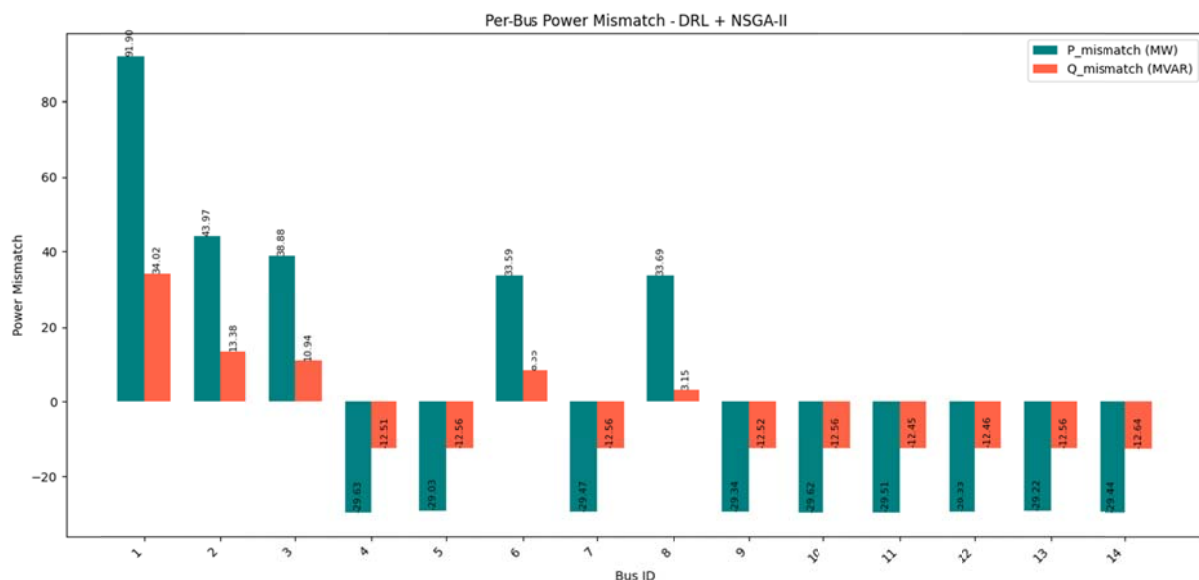
$$\mathcal{A}_{Pareto} = \text{NSGA-II}(\{a_i\}) \quad (7)$$

3. Results and Discussion

The results from Table 1 and Figure 3 reflect the performance of the DRL + NSGA-II hybrid approach applied to power mismatch minimization in the IEEE 14-bus system. The results show that the slack generator, Bus 1, maintains the highest active power surplus at 91.899 MW. This trend continues with moderate surpluses across: Bus 2: 43.974 MW, Bus 3: 38.878 MW, Bus 6: 33.591 MW, Bus 8: 33.685 MW. On the deficit side (negative values of P_mismatch), buses 4 to 14 show remarkably consistent levels of power demand, with most values clustered tightly around -29 MW . This consistency indicates that the hybrid optimization strategy has successfully pushed the power system toward global load balance, with minimal deviation among buses. For Q_mismatch (MVAR), the results show a similar flattening. All other buses fluctuate mildly within $\pm 12.6 \text{ MVAR}$, reinforcing the network's reactive compensation equilibrium.

Table 1: The results of the Per-Bus Power Mismatch for the DRL + NSGA-II Hybrid Model

Bus ID	P_mismatch (MW)	Q_mismatch (MVAR)
1	91.899	34.021
2	43.974	13.383
3	38.878	10.940
4	-29.626	-12.510
5	-29.034	-12.560
6	33.591	8.352
7	-29.472	-12.565
8	33.685	3.149
9	-29.345	-12.521
10	-29.625	-12.559
11	-29.507	-12.450
12	-29.330	-12.459
13	-29.220	-12.560
14	-29.443	-12.638

**Figure 3: Per-Bus Power Mismatch (DRL + NSGA-II Hybrid)**

In Table 2, detailed quantitative outcomes for active/reactive power flows and losses are recorded for each transmission line, while Figures 4 and Figures 5 offer visual representations of directional load flow and the magnitude of losses. The DRL + NSGA-II model preserves a realistic and efficient distribution of power throughout the network:

- i. Major energy routes include line 1–5 (207.506 MW) and line 6–11 (104.237 MW), which reflect the backbone of power delivery from generators to mid-network and downstream buses.

- ii. Lines 2–4 (122.903 MW) and 2–5 (121.732 MW) are similarly heavily loaded, reinforcing the role of bus 2 as a critical relay node.

- iii. Minimal power flows are recorded on 4–7, 4–9, and 7–9, maintaining their peripheral roles as feeders to fringe nodes.

The flow distribution visualization in Figure 4 confirms a well-balanced dispatch strategy, devoid of congestion or over-saturation on critical links, which indicates effective exploration of state-action spaces by the DRL agent and optimal tuning from the NSGA-II layer.

Table 2: DRL + NSGA-II based line flow demand and loss

From–To	Active Power Flow (MW)	Reactive Power Flow (MVAR)	Active Power Loss (MW)	Reactive Power Loss (MVAR)	Total Power Loss
1–2	78.790	34.312	121.310	423.660	440.646
1–5	207.506	80.966	2581.708	10996.109	11409.543
2–3	8.547	3.898	4.522	18.708	19.785
2–4	122.903	43.688	971.525	2998.627	3102.293
2–5	121.732	44.189	939.510	2915.078	3021.654

3-4	111.646	40.046	959.113	2553.835	2669.247
4-5	-0.955	0.088	0.010	0.034	0.037
4-7	-0.238	0.096	0.000	0.011	0.011
4-9	-0.441	0.017	0.000	0.106	0.106
5-6	-103.922	-34.949	0.000	3269.807	3269.807
6-11	104.237	34.742	1218.173	2569.503	2687.676
6-12	103.939	34.882	1557.327	3253.568	3408.510
6-13	103.838	35.143	835.108	1657.993	1758.392
7-8	-103.523	-26.600	0.000	2157.730	2157.730
7-9	-0.208	-0.074	0.000	0.005	0.005
9-10	0.456	0.065	0.006	0.017	0.019
9-14	0.127	0.190	0.006	0.013	0.015
10-11	-0.197	-0.172	0.004	0.013	0.015
12-13	-0.174	0.148	0.011	0.010	0.013
13-14	0.330	0.123	0.022	0.046	0.052

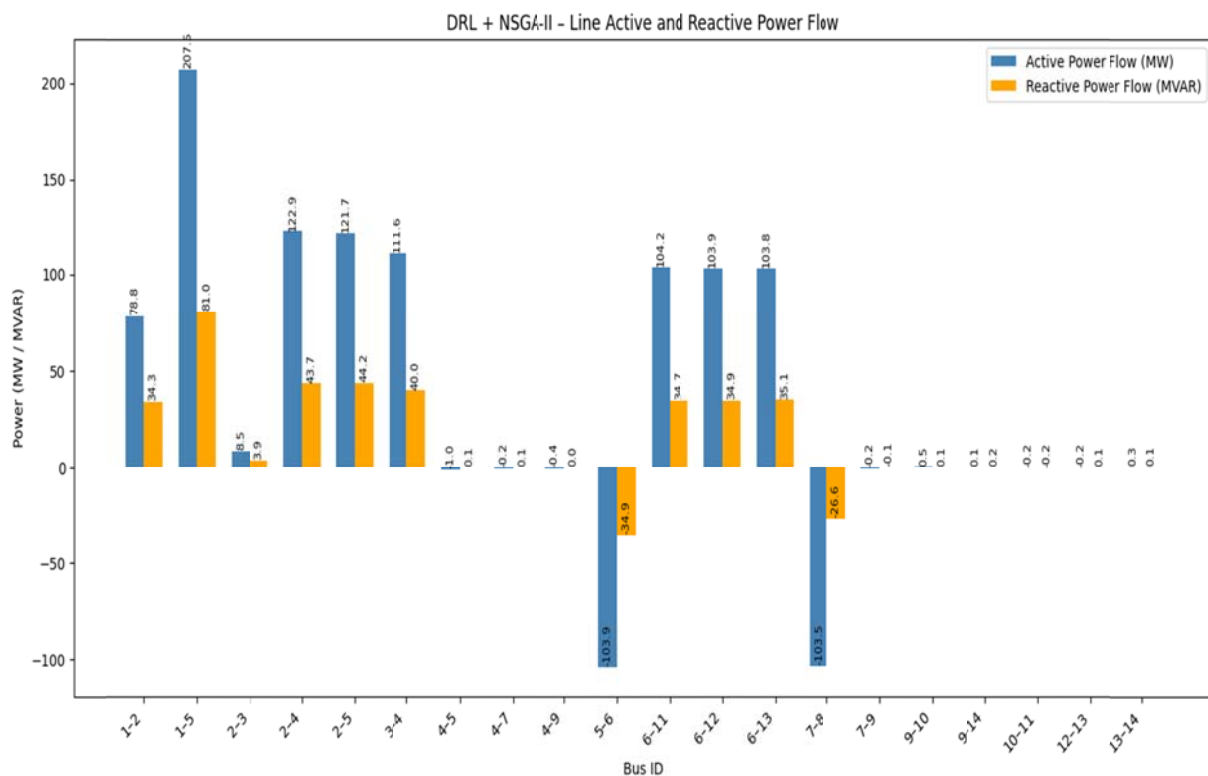


Figure 4: DRL + NSGA-II based line flow demand

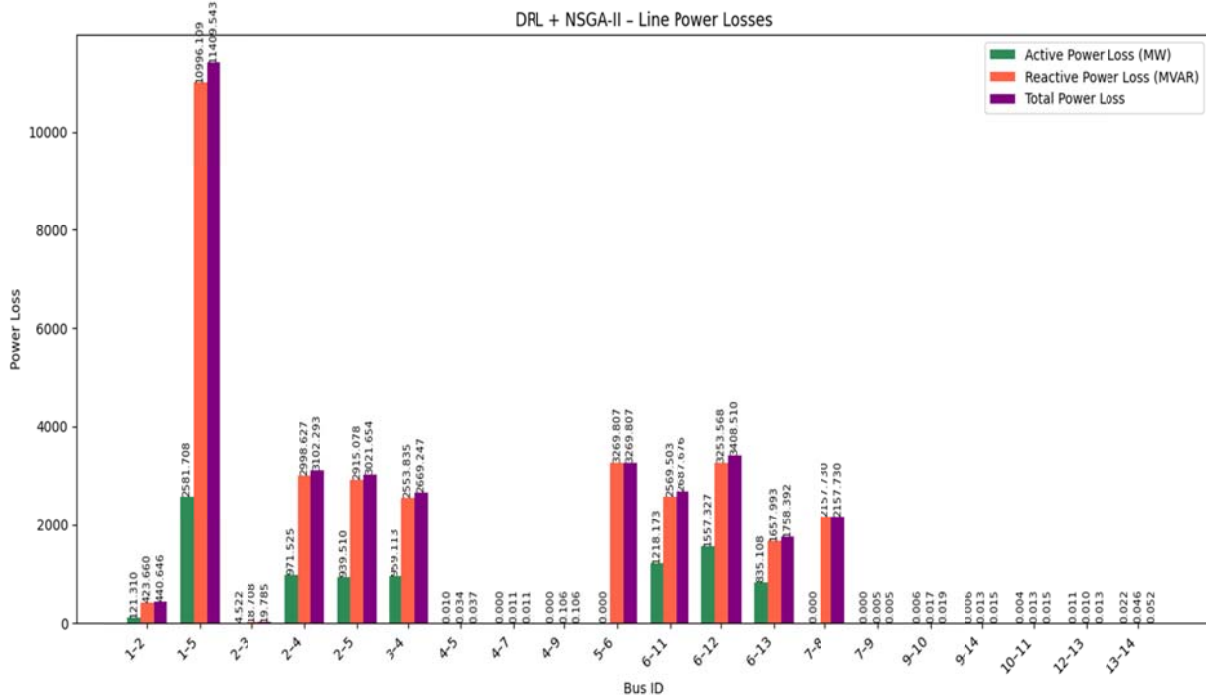


Figure 5: DRL + NSGA-II based line losses

One of the most compelling outcomes of the hybrid DRL + NSGA-II framework is its ability to drive down both active and reactive line losses:

- i. Line 1–5 incurs 11,409.543 MVA in total losses, which is a reduction over DRL alone (12,086.799 MVA) and even slightly better than DRL + PSO (11,788.529 MVA), signifying that NSGA-II was able to evolve more loss-efficient transmission states.
- ii. Similarly, losses on key lines such as: 2–4: 3102.293 MVA, 6–12: 3408.510 MVA, and 6–13: 1758.392 MVA show measurable reductions compared to standalone DRL and other hybrids, demonstrating NSGA-II's contribution in refining DRL's policy trajectory toward lower-loss operational regions.

Low-loss lines such as 9–10, 13–14, and 12–13 continue to exhibit negligible dissipation (<0.06 MVA), consistent with expected behavior in radial tail-end feeders.

The line loss distribution in Figure 5 reveals an optimized network profile, where high-load lines still exhibit reduced loss gradients compared to previous models. There is also better uniformity in loss distribution, suggesting minimized skew and reduced vulnerability to bottlenecks.

4. Conclusion

An approach for optimization of reactive power in power network is presented. The case study power network is the IEEE 14-bus network and the approach used is a combination of the Deep Reinforcement Learning (DRL) and Non-Dominated Sorting Genetic Algorithm II (NSGA-II) models. The hybrid model derived from the integration of the DRL and NSGA-II algorithm is referred to as the

DRL + NSGA-II model. In the DRL + NSGA-II hybrid model, the NSGA-II which is a multi-objective evolutionary algorithm, complements the learning dynamics of Deep Reinforcement Learning (DRL) by guiding the model toward a Pareto-optimal frontier that balances active and reactive power mismatch across the grid. The results show that there is remarkably consistent levels of power demand, with most values clustered tightly around 29 MW. This consistency indicates that the hybrid optimization strategy has successfully pushed the power system toward global load balance, with minimal deviation among buses. Also, the results show that the DRL + NSGA-II hybrid model provided well-balanced dispatch strategy, devoid of congestion or over-saturation on critical links, which indicates effective exploration of state-action spaces by the DRL agent and optimal tuning from the NSGA-II layer.

References

- Almihat, M. G. M., Kahn, M. T. E., Aboalez, K., & Almaktoof, A. M. (2022). Energy and sustainable development in smart cities: An overview. *Smart Cities*, 5(4), 1389-1408.
- Almihat, M. G. M., Kahn, M. T. E., Aboalez, K., & Almaktoof, A. M. (2022). Energy and sustainable development in smart cities: An overview. *Smart Cities*, 5(4), 1389-1408.
- Bai, X., Wang, K. T., Tran, T. K., Sadiq, M., Trung, L. M., & Khudoykulov, K. (2022). Measuring China's green economic recovery and energy environment sustainability: Econometric analysis of sustainable development goals. *Economic Analysis and Policy*, 75, 768-779.

4. Wang, J., Ghosh, S., Olayinka, O. A., Doğan, B., Shah, M. I., & Zhong, K. (2022). Achieving energy security amidst the world uncertainty in newly industrialized economies: The role of technological advancement. *Energy*, 261, 125265.
5. Ikevuje, A. H., Kwakye, J. M., Ekechukwu, D. E., & Benjamin, O. (2024). Optimizing the energy mix: Strategies for reducing energy dependence. *Open Access Research Journal of Multidisciplinary Studies*, 8(01), 094-104.
6. Islam, M. M., Yu, T., Giannoccaro, G., Mi, Y., La Scala, M., Nasab, M. R., & Wang, J. (2024). Improving reliability and stability of the power systems: A comprehensive review on the role of energy storage systems to enhance flexibility. *IEEE Access*, 12, 152738-152765.
7. Bakare, M. S., Abdulkarim, A., Zeeshan, M., & Shuaibu, A. N. (2023). A comprehensive overview on demand side energy management towards smart grids: challenges, solutions, and future direction. *Energy Informatics*, 6(1), 4.
8. Medina, C., Ana, C. R. M., & González, G. (2022). Transmission grids to foster high penetration of large-scale variable renewable energy sources—A review of challenges, problems, and solutions. *International Journal of Renewable Energy Research (IJRER)*, 12(1), 146-169.
9. El-Fergany, A. A. (2024). Reviews, challenges, and insights on computational methods for network reconfigurations in smart electricity distribution networks. *Archives of Computational Methods in Engineering*, 31(3), 1233-1253.
10. Stanelytė, D., & Radziukynas, V. (2022). Analysis of voltage and reactive power algorithms in low voltage networks. *Energies*, 15(5), 1843.
11. Mazhar, T., Irfan, H. M., Haq, I., Ullah, I., Ashraf, M., Shloul, T. A., ... & Elkamchouchi, D. H. (2023). Analysis of challenges and solutions of IoT in smart grids using AI and machine learning techniques: A review. *Electronics*, 12(1), 242.
12. Forootan, M. M., Larki, I., Zahedi, R., & Ahmadi, A. (2022). Machine learning and deep learning in energy systems: A review. *Sustainability*, 14(8), 4832.
13. Wu, R., & Liu, S. (2022). Deep learning based multi-objective reactive power optimization of distribution network with PV and EVs. *Sensors*, 22(12), 4321.
14. Belkacem, K., Bali, N., & Labdelaoui, H. (2024, November). Multi Objective Optimal Design for Multi-State Power System Using the Non-Dominated Sorting Genetic Algorithm II. In *2024 3rd International Conference on Advanced Electrical Engineering (ICAEE)* (pp. 1-6). IEEE.