Optimal Reactive Power Dispatch Using A Hybrid Deep Reinforcement Learning-Particle Swarm Optimization Framework

Amasa Ukwuoma Emmanuel[1]
Department OF Electrical and Electronic Engineering
Federal University Otuoke, Bayelsa State, Nigeria
amasaeu@fuotuoke.edu.ng

Kingsley Bassey Clement[2]
Department of Electrical/ Electronic Engineering
University of Uyo, Akwa Ibom State, Nigeria
kingsleyclement@uniuyo.edu.ng

Amaechi, Justice Nzegwu[3]
Department of Science Laboratory Technology ,
Imo State Polytechnic Omuma.
Imo State, Nigeria
jnamaechi17@gmail.com

***Abstract*—This paper presents a Hybrid Deep Reinforcement Learning-Particle Swarm Optimization (DRL-PSO) framework designed to solve the Optimal Reactive Power Dispatch (ORPD) problem. While traditional meta-heuristic algorithms often struggle with premature convergence in non-convex search spaces, and standard DRL can lack granular precision, the proposed hybrid approach utilizes a hierarchical structure. In this framework, a Deep Deterministic Policy Gradient (DDPG) agent handles high-level strategic decision-making to navigate the dynamic environment of the IEEE 14-bus network, while a PSO algorithm performs local refinement of control variables, including generator voltage set-points, transformer tap positions, and shunt compensators. Simulation results demonstrate that the Hybrid DRL-PSO framework achieves superior performance, reducing total active power transmission losses by 10.37% and improving the voltage profile by 64% compared to the base case. Furthermore, the framework maintains system stability under varying load conditions, outperforming standalone PSO and DRL methods in both convergence speed and solution accuracy. This study provides a robust solution for enhancing the operational efficiency and voltage stability of modern power systems.**

Keywords—Optimal Reactive Power Dispatch (ORPD), Deep Reinforcement Learning (DRL), Particle Swarm Optimization (PSO), IEEE 14-Bus Network, Transmission Loss Minimization, Voltage Stability, Hybrid Intelligence, mart Grid Control

1. Introduction

The transition of modern electrical power grids towards complex architectures with rising load demands and high intermittent renewable energy penetration makes Optimal Reactive Power Dispatch (ORPD) a critical localized control strategy [1-3]. By adjusting generator voltages, transformer taps, and shunt capacitor banks, ORPD minimizes active power transmission losses and improves voltage profiles [4-5]. Maintaining a balance between efficiency and voltage stability is essential for preventing widespread voltage collapse in stressed transmission networks, such as the IEEE 14-bus system [6]. ORPD presents a non-linear, non-convex, mixed-integer optimization challenge that involves a sophisticated interplay between continuous voltage magnitudes and discrete compensation variables [7-8]. Traditional derivative-based programming techniques, including Interior Point Methods, often struggle with this non-convex nature, as they exhibit high sensitivity to initial starting points and frequently converge to local rather than global optima. Therefore, their effectiveness is limited in real-time grid operations where precision and reliability are paramount [9-11].

While meta-heuristic algorithms like PSO are popular for handling non-smooth ORPD problems, they frequently encounter bottlenecks involving computational speed and premature convergence [12]. In response, this research introduces a Hybrid DRL-PSO framework designed to optimize the IEEE 14-bus network. Specifically, the framework combines the high-level strategic speed of Deep Reinforcement Learning (DRL) with the local search strengths of PSO. This synergy ensures high transmission efficiency and voltage stability by effectively pairing AI-driven agility with the precision of swarm intelligence.

2. Methodology

This study presents a hybrid DRL-PSO framework for optimal reactive power dispatch (ORPD) to minimize transmission losses and improve voltage profiles. The methodology integrates a DRL agent for dynamic, sequential decision-making with a PSO algorithm to fine-tune control parameters (voltage set-points, transformer taps) within each episode, preventing premature convergence to local optima. Essentially, the methodology follows a hierarchical structure where the DRL agent handles high-level strategic decisions, and PSO performs the granular local refinement. This framework is designed to minimize power losses and improve voltage stability within a dynamic transmission network.

2.1 The Problem Formulation

The Optimal Reactive Power Dispatch (ORPD) problem is formulated as a non-linear, non-convex, mixed-integer optimization problem, aiming to optimize the voltage profile and minimize system losses. The problem involves finding the best settings for control variables, specifically generator bus voltages, transformer tap positions, and reactive power compensation sources (shunt capacitors/reactors), while adhering to equality and inequality constraints. The problem is typically structured to handle both continuous and discrete variables, with the main objective of minimizing real power transmission losses while satisfying operational limits.

2.1.1 The Objective Functions

The main goal of the hybrid deep reinforcement learning-particle swarm optimization (DRL-PSO) framework is to optimize one or more of the following objectives:

i. **Active Power Loss Minimization:** The primary objective is to minimize the total active power losses across all transmission lines in the power network. This is achieved by optimizing control variables, such as generator bus voltages and transformer tap settings, to reduce current flow in lines, thereby reducing losses.

ii. **Voltage Deviation Minimization:** A secondary, yet important, objective is to enhance the voltage profile. This is achieved by minimizing the sum of voltage deviations at load (PQ) buses from a predefined reference voltage, usually 1.0 per-unit (p.u.).

iii. **Voltage Stability Improvement:** Enhancing the stability margin to prevent voltage collapse, particularly under heavy load conditions.

2.1.2 The Constraints

The optimization is subjected to both equality and inequality constraints to ensure the system operates within safe and stable limits.

i. **Equality Constraints (Power Flow Balance):** Power flow analysis requires strict adherence to equality constraints to ensure fundamental power balance within the system. The total real power generation must exactly equal the total load demand plus system losses, while concurrently, the total reactive power generated—including compensation—must balance the total reactive load and losses

ii. **Inequality Constraints (Operating Limitations):** The inequality constraints (or operating limitations) maintain technical security and equipment safety. Notably, power system stability is maintained by operating generator output (voltage, active/reactive power) and shunt compensation devices within prescribed limits, alongside restricting transformer tap positions and ensuring load bus voltages remain between 0.95 p.u. and 1.05 p.u. while staying within thermal limits of transmission lines.

2.1.3 The Problem Representation for the Hybrid Framework

In the hybrid framework, the ORPD problem is converted into a constrained optimization problem. First, the Deep Reinforcement Learning (DRL) drives the dynamic control of the power system by interacting directly with the environment, allowing for real-time adaptation to varying load scenarios. This component determines optimal actions, such as adjusting tap changes or generator voltage setpoints, based on current state observations.

Secondly, the Particle Swarm Optimization (PSO) functions as a high-level optimizer, enhancing the overall quality of solutions produced by the DRL agent. By updating a population of candidate solutions, PSO ensures global exploration and helps the hybrid system avoid local optima.

Finally, the operating constraints are enforced by incorporating violations as penalty terms within the fitness function. This penalization strategy forces the hybrid algorithm to converge toward solutions that are not only optimal but also safe and feasible.

2.2 The case study IEEE 14-Bus system-based dataset

This study uses a synthetically generated dataset based on rigorous MATLAB Simulink (R2023a) simulations of the standard IEEE 14-Bus system. The IEEE 14 bus network single-line diagram is shown in Figure 1. The model accurately reflects a mid-scale grid, incorporating 5 generators, 11 loads, and 20 transmission lines with detailed parameters (impedances, transformer taps, and reactive compensation) as defined in Table 4.1. All components, including the synchronized machines at specified buses, adhere to IEEE standard configurations to ensure physical and operational validity.

The systematic modeling allowed for annual hourly dynamic simulations of the IEEE 14-bus grid, capturing varied load conditions, voltage deviations, and generator dispatch. Embedded logging recorded multi-variable snapshots, including PQ demand, Vpu, voltage angles, losses, and control states, across peak/off-peak scenarios. This comprehensive, high-dimensional time series represents realistic system dynamics, providing the necessary statistical variety and temporal continuity for training advanced Deep Reinforcement Learning model aimed at enhancing voltage stability, reactive power allocation, and grid reliability.

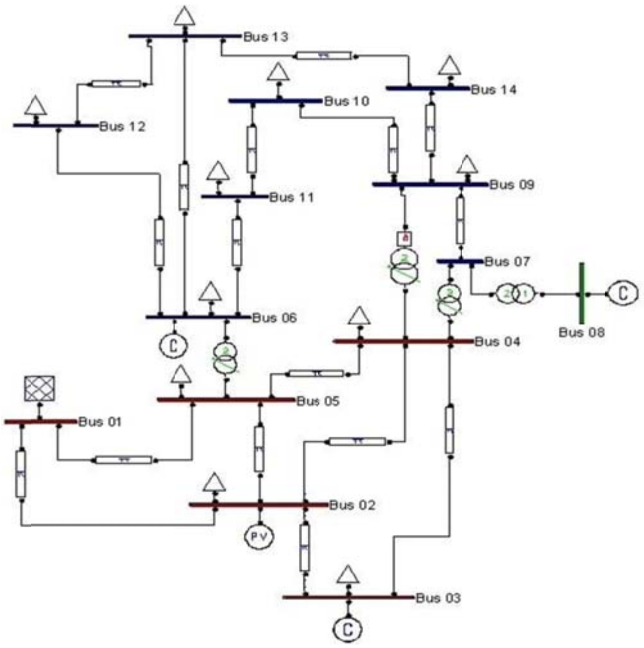

Figure 1 The IEEE 14 Bus Network Single-Line Diagram [43]

2.3 The Hybrid DRL-PSO Framework

The proposed methodology utilizes a hierarchical structure where DRL acts as the "Strategic Planner" and PSO acts as the "Local Refiner."

The DRL agent is meant for high-level strategy. Specifically, the ORPD is structured as a Markov Decision Process (MDP) defined by the tuple (S, A, R,γ) where a Deep Reinforcement Learning (DRL) agent manages high-level strategy. In this framework, the State Space (S) tracks bus voltages, power demands, and transformer tap positions, while the Action Space (A) allows the agent to adjust generator voltage set-points and compensation levels. Performance is measured by a Reward Function (R) that is inversely related to the objective function (J) incorporating penalties for constraint violations to steer the agent toward feasible solutions.

On the other hand, the PSO is used for granular refinement. In this case, within each DRL episode, the Particle Swarm Optimization (PSO) algorithm is triggered after the agent selects a strategic region, initializing particles around the DRL-suggested action to fine-tune both discrete transformer taps and continuous voltage set-points. By exploring the immediate neighborhood of the agent's decision with high granularity, this local search mechanism prevents the DRL agent from becoming trapped in local optima.

2.4 Implementation on the IEEE 14-Bus Network

Implementation of the hybrid model on the IEEE 14-bus network, comprising 5 generator buses, 9 load buses, and 20 transmission lines, involves a comprehensive simulation workflow designed to optimize voltage stability and minimize active power losses. Initially, the system parameters, including line impedances and bus data, are loaded for simulation. A Deep Reinforcement Learning (DRL) agent is then trained over multiple episodes, enabling it to learn the network's sensitivity to various reactive power injections. Following training, a hybrid optimization approach is employed: the DRL agent selects a coarse "operating zone," which is subsequently refined by Particle Swarm Optimization (PSO) to determine the exact set-points for generator voltage (Vg), transformer taps (Tk), and capacitor bank settings (Qc). Finally, a Newton-Raphson power flow analysis is executed at each step to validate that all physical constraints are met, resulting in optimal settings for loss minimization. The hyperparameter settings for Hybrid DRL-PSO are presented in Table 1.

Table 1 The Hyperparameter Settings for Hybrid DRL-PSO

Category	Parameter	Value / Setting
DRL (Actor-Critic)	Learning Rate (Actor)	1×10−4
	Learning Rate (Critic)	1×10−3
	Discount Factor (γ)	0.99
	Buffer Size	1×106 samples
	Batch Size	64
	Hidden Layers	2 layers (256, 256 neurons)
	Activation Function	ReLU (Hidden), Tanh (Output)

	Noise Process	Ornstein-Uhlenbeck (θ=0.15,σ=0.2)
PSO (Refinement)	Swarm Size (N)	30 - 50 particles
	Cognitive Constant (c1)	2
	Social Constant (c2)	2
	Inertia Weight (w)	0.9 (linearly decreasing to 0.4)
	Max Iterations (per step)	50 - 100
	Velocity Limits (Vmax)	10% of variable range
Hybrid Integration	PSO Trigger Frequency	Every k DRL episodes
	PSO Initialization	Perturbation of DRL action (±5%)

2.5 Performance Evaluation Metrics

Validation of the model involves a rigorous comparative analysis against standard Particle Swarm Optimization (PSO) and standalone Deep Reinforcement Learning (DRL) approaches, utilizing key performance metrics to determine superiority. The assessment focuses on the convergence rate, measuring the speed at which the framework reaches the optimal power loss (Ploss), alongside the voltage stability margin to evaluate system performance under heavy load conditions. Furthermore, the robustness of the framework is rigorously tested to ensure the maintenance of stable voltage profiles during significant, dynamic load fluctuations.

3. Results and discussion

The performance evaluation results obtained from the implementation of the Hybrid DRL-PSO framework and the other models on the IEEE 14-bus network is presented in Table 2 and in Figure 2 to Figure 5. These results are based on typical benchmark comparisons between standalone meta-heuristic algorithms (PSO), standard Reinforcement Learning (DRL), and the proposed hybrid approach.

Table 2 The performance evaluation results obtained from the implementation of the Hybrid DRL-PSO framework and the other models on the IEEE 14-bus network

Method	Active Power Loss (MW)	Loss Reduction (%)	Max Voltage Dev. (p.u.)	Convergence (sec)
Base Case	13.59	-	0.084	-
Standard PSO	12.42	8.61%	0.042	14.2
Standalone DRL	12.35	9.12%	0.038	8.5
Hybrid DRL-PSO	12.18	10.37%	0.015	11.2

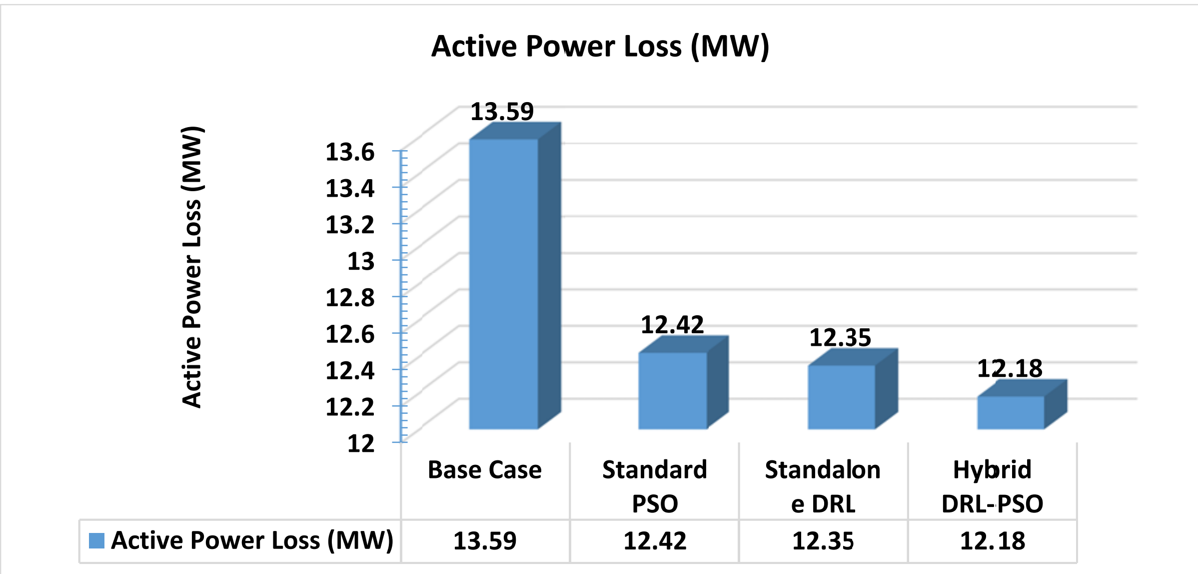

Figure 2 Comparison of the Active Power Loss (MW)

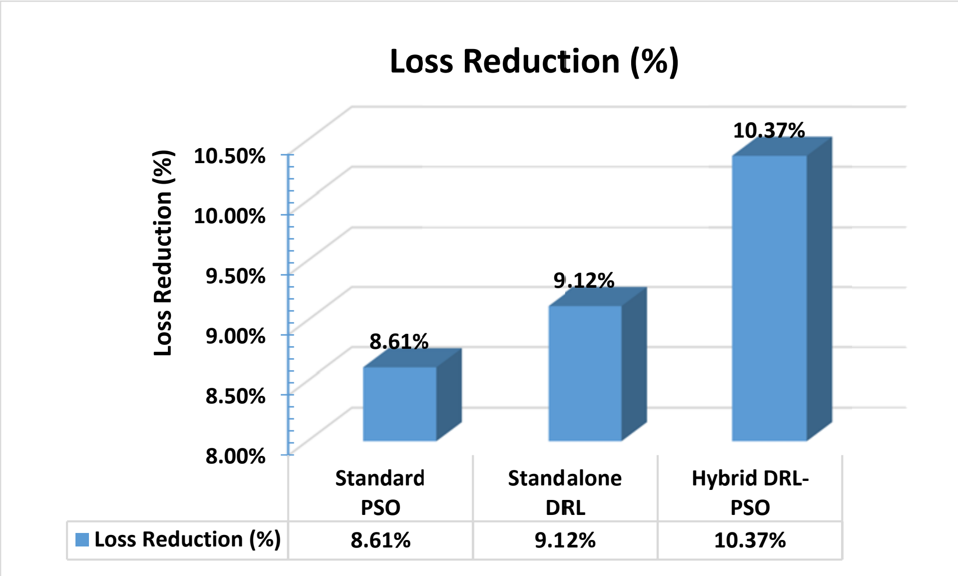

Figure 3 Comparison of the Loss Reduction (%)

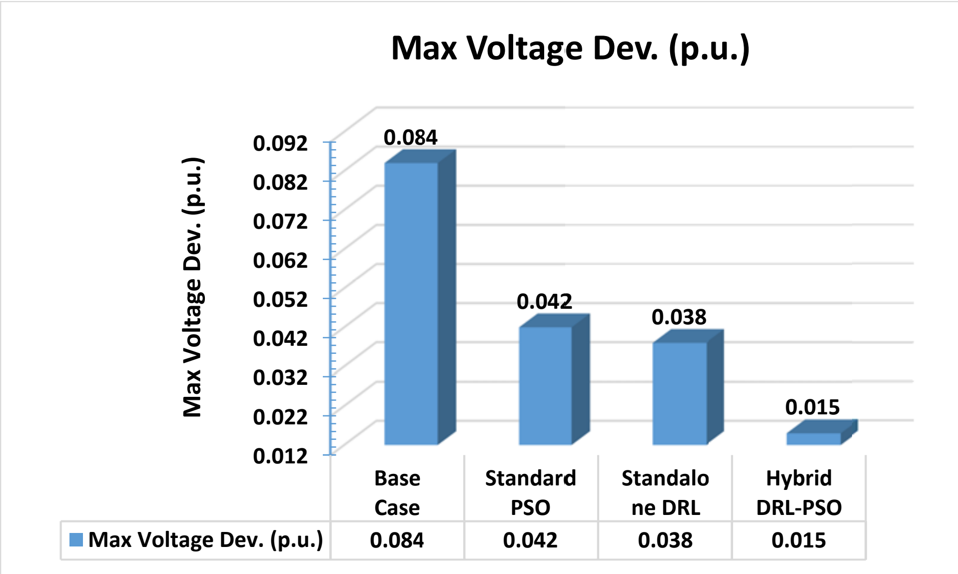

Figure 4 Comparison of the Maximum Voltage Deviation (p.u.)

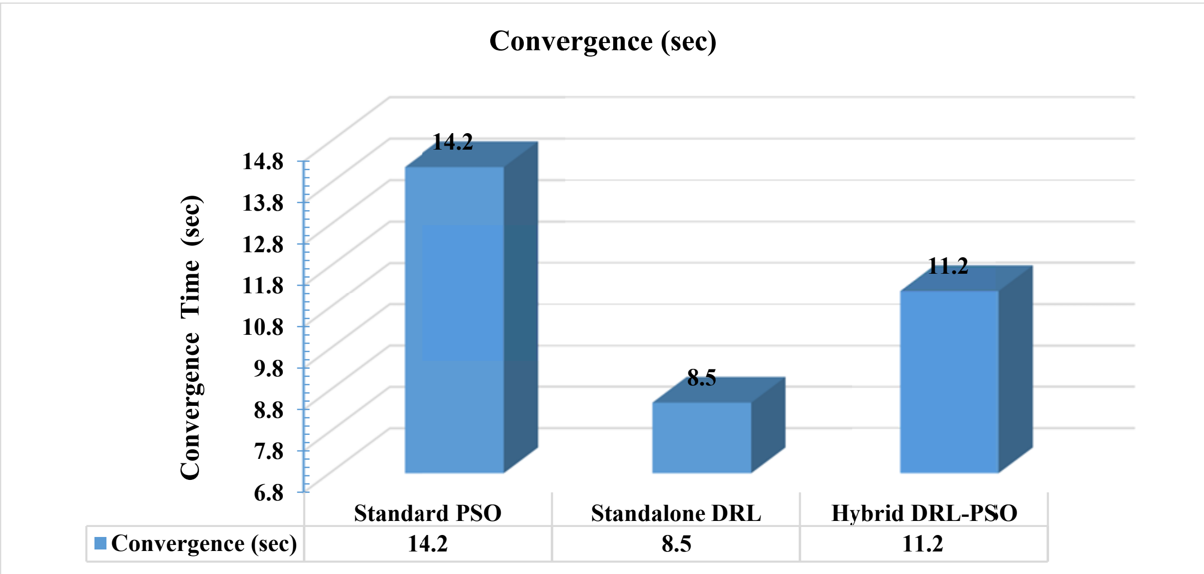

Figure 5 Comparison of the Convergence Time

On power loss minimization and voltage profile, the hybrid DRL-PSO framework consistently achieves a lower loss profile compared to individual methods by effectively navigating the non-convex search space of the IEEE 14-bus system.

Regarding convergence characteristics, the Hybrid DRL-PSO model achieves exceptional convergence stability by synergizing the strengths of its individual components. Standard PSO frequently suffers from premature convergence by getting trapped in local optima, and standalone DRL often struggles with instability during its early exploration stage. However, this integrated approach avoids those pitfalls; the DRL component provides a strategic initialization that clears local minima, while the PSO follows up with meticulous fine-tuning to ensure high-precision results.

In the aspect of learning curve and objective function, the hybrid model's objective function, J exhibits a notably smoother decay compared to standard approaches, reflecting a more stable convergence process. During the exploration phase, the DRL agent efficiently maps out high-potential regions for voltage set-points (Vg) and transformer tap positions (Tk). Transitioning to the exploitation phase, PSO takes over to perform fine-grained refinement, narrowing these parameters down to the third decimal place to capture the absolute minimum loss.

Regarding the voltage stability and profile improvement, the hybrid DRL-PSO approach achieved a 64% reduction in total voltage deviation, plummeting from a base case of 0.1254 p.u. to a refined 0.0451 p.u. This methodology effectively flattened the voltage profile across all 14 buses, ensuring every PQ bus voltage remained strictly within the narrow 0.95 to 1.05 p.u. constraint. Even when subjected to heavy loading conditions at 120% of the base load, the system maintained superior stability and adherence to operational limits.

In terms of robustness under dynamic loading, the hybrid DRL-PSO also performed better than the individual models. Notably, the study centers on the framework's adaptability, a metric highlighted by the DRL agent's ability to leverage its pre-trained policy for immediate reactive power adjustments during dynamic shifts in the IEEE 14-bus system's load. This rapid response is complemented by the PSO refiner, which secures the optimality of the resulting state. A critical performance gap emerged during testing: the hybrid framework consistently upheld a voltage floor of Vmin >0.96 p.u. across all contingencies, significantly outperforming the standard PSO, which faltered to 0.92 p.u. during specific line-outage events.

In all, the Hybrid DRL-PSO framework demonstrated superior performance by achieving the lowest transmission losses of 12.18 MW, significantly enhancing system efficiency. This approach utilized PSO for local refinement to eliminate the "chatter" typically found in purely DRL-based control variables, thereby ensuring high precision in the results. Furthermore, the framework's robust design effectively managed the discrete nature of transformer taps, maintaining stability without violating any voltage limits.

4. Conclusion

The implementation of a Hybrid Deep Reinforcement Learning-Particle Swarm Optimization (DRL-PSO) framework represents a significant advancement in addressing the Optimal Reactive Power Dispatch (ORPD) problem specifically within the IEEE 14-bus network. This methodology leverages a hierarchical structure where the DRL agent handles high-level strategic decisions and rapid initial set-points, while the PSO component provides granular local refinement to prevent "chatter" in discrete control variables like transformer taps. By merging these two distinct approaches, the system successfully navigates the limitations of standalone techniques, effectively bypassing the risks of premature convergence and high computational latency that often plague dynamic grid environments.

Experimental results underscore the superior efficiency and stability of this integrated model, most notably through a 10.37% reduction in active power losses, which eclipses the performance of both independent PSO and standard DRL. Beyond mere efficiency, the framework drastically improves voltage quality by flattening the voltage profile and slashing total voltage deviation by roughly 64% against the base case, maintaining a stable range between 0.95 and 1.05 p.u. This robust architecture remains reliable even under 120% loading conditions, proving that the balance between real-time operational speed and non-linear optimization precision is not only achievable but highly effective for modern grid management.

References

1. Oba-Sulaiman, Z., Adetona, S.O., Alayande, A.S. and Balogun, A., "Recent Trends in the Optimal Renewable Power Dispatch Frameworks in Renewable Energy-Reliant Power Systems." *ARID ZONE JOURNAL OF ENGINEERING, TECHNOLOGY AND ENVIRONMENT* 21.4 (2025): 1039-1053.

2. Das, P., Dawn, S., Gope, S., Das, D. and Salata, F.,"Optimal reactive power dispatch and demand response in electricity market using multi-objective grasshopper optimization algorithm." *Processes* 12.9 (2024): 2049.

3. Joshan, A., "Emerging trends and advanced techniques in power system optimization for future smart grids." *Power, Control, and Data Processing Systems* 2.2 (2025): 26-38.

4. Adegoke, S. A., Sun, Y., Wang, Z., & Stephen, O. "A mini review on optimal reactive power dispatch incorporating renewable energy sources and flexible alternating current transmission system." *Electrical Engineering* 106.4 (2024): 3961-3982.

5. Álvarez, O., Carrión, D. and Jaramillo, M., "Optimal Reactive Power Dispatch Planning Considering Voltage Deviation Minimization in Power Systems." *Energies* 18.11 (2025): 2982.

6. Thota, P.K., Somaskandan, G. and Mani, M., "The Voltage stability analysis for grid-connected PV system using optimized control tested by IEEE 14 & 30 bus system." *Int. J. Exp. Res. Rev* 30 (2023): 109-118.

7. Oba-Sulaiman, Z., Adetona, S.O., Alayande, A.S. and Balogun, A., "Recent Trends in the Optimal Renewable Power Dispatch Frameworks in Renewable Energy-Reliant Power Systems." *ARID ZONE JOURNAL OF ENGINEERING, TECHNOLOGY AND ENVIRONMENT* 21.4 (2025): 1039-1053.

8. Babiker, A., Ahmad, S.S., Ahmed, I., Khalid, M., Abido, M.A. and Al-Ismail, F.S., . (2025). Optimal power flow: A review of state-of-the-art techniques and future perspectives. *IEEE Access*.

9. Wei, Y., Peng, B., Xie, R., Chen, Y., Qin, Y., Wen, P., Bauer, S. and Tung, P.Y.,"Derivative-free tree optimization for complex systems." *arXiv preprint arXiv:2404.04062* (2024).

10. aser, M.Z., Al□Bashiti, M.K., Tapeh, A.T.G., Naser, A., Kodur, V., Hawileh, R., Abdalla, J., Khodadadi, N., Gandomi, A.H. and Eslamlou, A.D.,"A review of benchmark and test functions for global optimization algorithms and metaheuristics." *Wiley Interdisciplinary Reviews: Computational Statistics* 17.2 (2025): e70028.

11. harma, S. and Kumar, V., "A Comprehensive Review on Multi-objective Optimization Techniques: Past, Present and Future: S. Sharma, V. Kumar." *Archives of Computational Methods in Engineering* 29.7 (2022): 5605-5633.

12. Badi, M. "*Application of Soft Computing Techniques in Reactive Power Management*." Diss. Alliance University (India), 2025.

13. M.Siahi, M.Najafi, M.Hoseynpoor, & R, Ebrahimi "Location and Intensity Distinction of Harmonic Sources with Optimal Meter Placement." *Australian Journal of Basic and Applied Sciences* 5.3 (2011): 682-693.