

Adaptive Deep Reinforcement Learning for Coordinated Voltage Control in PV-Rich Distribution Grids

1st Tarun Kumar Modi

Dept. of Electrical Engineering
Sardar Patel University
Balaghat, India
tarunk24x7@gmail.com

2nd Naresh Sapate

Dept. of Electrical Engineering
Sardar Patel University
Balaghat, India
nsnareshsapate588@gmail.com

3th Shailendra Turker

Dept. of Electrical Engineering
Sardar Patel University
Balaghat, India
turkershailendra91@gmail.com

Abstract—The increasing integration of photovoltaic (PV) systems into distribution networks has introduced significant challenges related to voltage regulation, reverse power flow, and power quality. Traditional voltage control methods, which rely on rule-based logic or centralized optimization, often fail to address the dynamic and stochastic nature of PV generation. This paper proposes an adaptive deep reinforcement learning (DRL) framework for coordinated voltage control in PV-rich distribution grids. The proposed method employs a Deep Deterministic Policy Gradient (DDPG) algorithm that learns optimal control policies by interacting with the distribution network environment. The DRL agent coordinates multiple voltage regulation devices, including smart inverters, on-load tap changers (OLTCs), and capacitor banks, to maintain voltage within acceptable limits while minimizing network losses and control actions. The framework incorporates an adaptive mechanism that adjusts the learning process based on changing network conditions, seasonal variations, and PV output patterns. Simulation results on a modified IEEE 33-bus distribution system demonstrate that the proposed approach outperforms conventional methods in terms of voltage profile improvement, loss reduction, and computational efficiency.

Index Terms—Deep reinforcement learning, voltage control, photovoltaic systems, distribution networks, smart inverters, DDPG algorithm

I. INTRODUCTION

A. Background

The global transition toward sustainable energy systems has accelerated the deployment of distributed energy resources (DERs), particularly rooftop and utility-scale photovoltaic (PV) systems. According to the International Energy Agency, solar

PV capacity is expected to triple by 2030, with a significant portion connected to low and medium voltage distribution networks [1]. While this growth supports decarbonization goals, it also introduces operational challenges that existing grid infrastructure was not designed to handle.

Distribution networks traditionally operated with unidirectional power flow from substations to consumers. However, high PV penetration reverses this paradigm, causing bidirectional power flows that lead to voltage rise, thermal overloading, and power quality issues [2]. Voltage violations occur when PV generation exceeds local demand, pushing voltages beyond statutory limits and potentially damaging sensitive equipment.

B. Challenges in Voltage Control

Conventional voltage regulation relies on mechanical devices such as on-load tap changers (OLTCs), step voltage regulators, and switched capacitor banks. These devices operate based on local measurements and predefined setpoints, making them unsuitable for the rapid fluctuations characteristic of PV output [3]. The response time of mechanical devices ranges from seconds to minutes, which is inadequate for managing cloud-induced irradiance variations that occur within seconds.

Modern approaches have introduced smart inverters capable of providing reactive power support through functions such as Volt-VAR and Volt-Watt control [4]. However, uncoordinated operation of multiple smart inverters can lead to conflicting

control actions, oscillatory behavior, and suboptimal network performance. Centralized optimization methods, while theoretically optimal, require complete network observability, accurate forecasts, and significant computational resources, limiting their practical applicability [5].

C. Motivation for Reinforcement Learning

Reinforcement learning (RL) offers a fundamentally different approach to control problems. Unlike model-based optimization, RL agents learn control policies through trial-and-error interaction with the environment, without requiring explicit mathematical models of system dynamics [6]. This characteristic makes RL particularly suitable for distribution networks, where accurate models are difficult to obtain due to limited observability, measurement errors, and constantly changing load patterns.

Deep reinforcement learning (DRL), which combines RL with deep neural networks, has demonstrated remarkable success in complex control tasks, including robotics, game playing, and autonomous systems [7]. Recent studies have explored DRL for power system applications, including economic dispatch, frequency regulation, and voltage control [8]. However, existing DRL-based voltage control methods often assume static network conditions and do not adequately address the adaptive requirements of PV-rich grids.

D. Contributions

This paper addresses the limitations of existing approaches by proposing an adaptive deep reinforcement learning framework for coordinated voltage control. The main contributions are:

- 1) Development of a comprehensive DRL framework that coordinates multiple voltage regulation devices including smart inverters, OLTCs, and capacitor banks in a unified control architecture.
- 2) Introduction of an adaptive learning mechanism that adjusts the exploration-exploitation balance and reward function based on changing network conditions and seasonal PV patterns.
- 3) Design of a multi-objective reward function that balances voltage regulation, loss minimization, and control effort reduction.

- 4) Extensive validation on a modified IEEE 33-bus distribution system with realistic PV generation and load profiles.
- 5) Comparative analysis with conventional control methods demonstrating superior performance in terms of voltage profile, network losses, and computational efficiency.

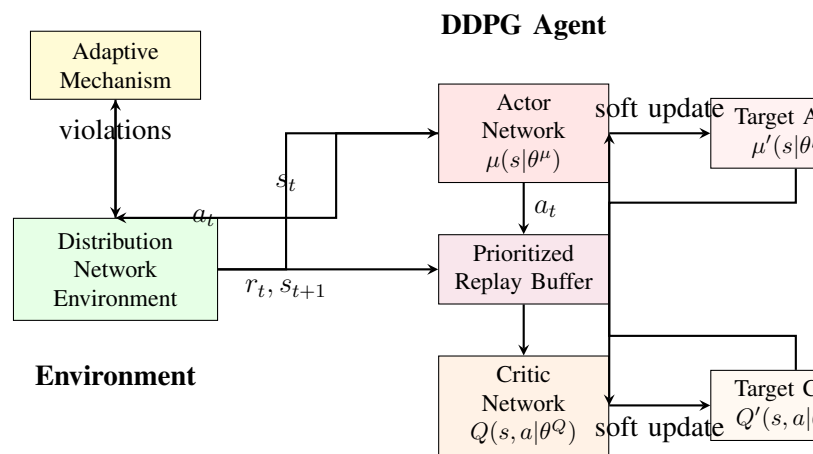


Fig. 1. Architecture of the proposed adaptive deep reinforcement learning framework.

II. LITERATURE REVIEW

A. Traditional Voltage Control Methods

Voltage control in distribution networks has traditionally relied on mechanical devices and rule-based control strategies. On-load tap changers (OLTCs) at distribution transformers adjust the turns ratio to regulate secondary voltage based on measured values at the substation or load center [9]. The control logic typically employs a deadband and time delay to avoid excessive tap operations that accelerate mechanical wear.

Capacitor banks provide reactive power compensation to improve voltage profiles and reduce losses. Fixed capacitors offer constant compensation, while switched capacitors allow discrete adjustment based on load conditions [10]. Step voltage regulators, installed along feeders, provide additional voltage boost or buck to maintain acceptable levels at distant nodes.

These conventional methods suffer from several limitations in PV-rich networks:

- Slow response times incompatible with rapid PV fluctuations

- Local control without coordination across devices
- Fixed setpoints that do not adapt to changing generation patterns
- Limited controllability for bidirectional power flows

B. Smart Inverter-Based Control

The IEEE 1547-2018 standard mandates advanced capabilities for DER interconnection, including autonomous voltage regulation functions [11]. Smart inverters can provide reactive power support through Volt-VAR control, Volt-Watt control, constant power factor operation, and constant reactive power mode.

Several studies have investigated optimal settings for these functions. Researchers in [12] developed a time-series analysis approach to determine Volt-VAR curve parameters that minimize voltage violations over annual operating scenarios. Authors in [13] proposed a two-stage optimization that first determines global settings and then refines them based on local conditions.

Despite their advantages, smart inverters operating autonomously can create adverse interactions. Studies have documented cases of voltage oscillations, hunting behavior, and even instability when multiple inverters with aggressive settings operate in proximity [14].

C. Optimization-Based Approaches

Centralized optimization methods formulate voltage control as a mathematical programming problem. The optimal power flow (OPF) framework minimizes an objective function subject to power flow equations and operational constraints [15]. Mixed-integer formulations accommodate discrete devices such as tap changers and switched capacitors.

Model predictive control (MPC) extends optimization to consider future time steps, incorporating forecasts of PV generation and load demand [16]. The receding horizon approach re-solves the optimization at each time step, providing feedback to handle forecast errors. However, MPC requires accurate network models and forecasts, and computational burden increases with prediction horizon length.

Distributed optimization algorithms, including alternating direction method of multipliers (ADMM) and consensus-based methods, decompose the centralized problem into local subproblems that exchange information iteratively [17]. These approaches offer scalability advantages but may converge slowly and require reliable communication infrastructure.

D. Reinforcement Learning in Power Systems

Reinforcement learning has emerged as a promising alternative for power system control due to its model-free nature and ability to handle uncertainty. Early applications used tabular Q-learning for discrete control problems such as capacitor switching [18]. However, tabular methods suffer from the curse of dimensionality when state and action spaces are large or continuous.

Deep reinforcement learning addresses scalability through function approximation with neural networks. The Deep Q-Network (DQN) algorithm has been applied to various power system problems [19]. For continuous action spaces, policy gradient methods such as Deep Deterministic Policy Gradient (DDPG), Twin Delayed DDPG (TD3), and Soft Actor-Critic (SAC) offer superior performance [20].

Recent studies have applied DRL specifically to voltage control. Authors in [21] developed a multi-agent DRL system where each smart inverter operates as an independent agent. Researchers in [22] proposed a safe reinforcement learning approach that constrains actions to prevent voltage violations during learning. Work in [23] introduced a transfer learning framework that accelerates training by leveraging knowledge from similar network configurations.

III. SYSTEM MODEL AND PROBLEM FORMULATION

A. Distribution Network Model

Consider a radial distribution network with N buses, represented by the set $\mathcal{N} = \{0, 1, 2, \dots, N\}$, where bus 0 is the substation. The set of branches is denoted by $\mathcal{L} = \{(i, j) : i, j \in \mathcal{N}\}$. For each branch (i, j) , let r_{ij} and x_{ij} represent the resistance and reactance, respectively.

The power flow equations for a radial network can be expressed using the DistFlow model [24]:

$$P_{ij} = P_j + r_{ij} \frac{P_{ij}^2 + Q_{ij}^2}{V_i^2} + \sum_{k:(j,k) \in \mathcal{L}} P_{jk} \quad (1)$$

$$Q_{ij} = Q_j + x_{ij} \frac{P_{ij}^2 + Q_{ij}^2}{V_i^2} + \sum_{k:(j,k) \in \mathcal{L}} Q_{jk} \quad (2)$$

$$V_j^2 = V_i^2 - 2(r_{ij}P_{ij} + x_{ij}Q_{ij}) + (r_{ij}^2 + x_{ij}^2) \frac{P_{ij}^2 + Q_{ij}^2}{V_i^2} \quad (3)$$

where P_{ij} and Q_{ij} are the active and reactive power flows on branch (i, j) , P_j and Q_j are the net active and reactive power injections at bus j , and V_i is the voltage magnitude at bus i .

Training Progress Comparison

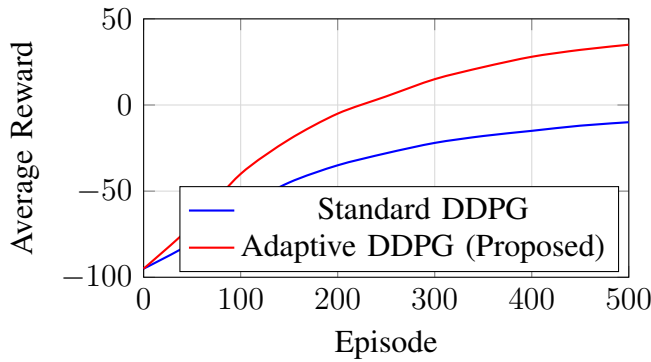


Fig. 2. Training progress comparison between standard DDPG and proposed adaptive DDPG.

B. PV System Model

Each PV system is characterized by its rated capacity S_{PV}^{rated} and maximum active power output P_{PV}^{max} . The actual active power output depends on solar irradiance:

$$P_{PV}(t) = P_{PV}^{max} \cdot \frac{G(t)}{G_{ref}} \cdot \eta_{temp}(t) \quad (4)$$

where $G(t)$ is the solar irradiance at time t , G_{ref} is the reference irradiance (typically 1000 W/m²), and $\eta_{temp}(t)$ is the temperature-dependent efficiency factor.

Smart inverters can provide reactive power within their apparent power rating:

$$Q_{PV}^{max}(t) = \sqrt{(S_{PV}^{rated})^2 - (P_{PV}(t))^2} \quad (5)$$

Daily PV Generation and Load Profile

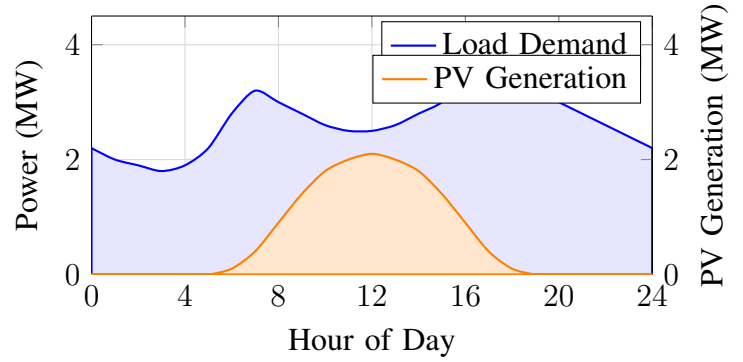


Fig. 3. Daily load demand and PV generation profiles used in simulation.

C. Problem Formulation

The voltage control problem aims to maintain all bus voltages within acceptable limits while minimizing network losses and control actions. The multi-objective optimization problem is formulated as:

$$\min J = \omega_1 J_{voltage} + \omega_2 J_{loss} + \omega_3 J_{control} \quad (6)$$

where:

$$J_{voltage} = \sum_{i \in \mathcal{N}} \max(0, V_i - V_{max})^2 + \max(0, V_{min} - V_i)^2 \quad (7)$$

$$J_{loss} = \sum_{(i,j) \in \mathcal{L}} r_{ij} \frac{P_{ij}^2 + Q_{ij}^2}{V_i^2} \quad (8)$$

$$J_{control} = \sum_k |\Delta tap_k|^2 + \sum_m |\Delta cap_m|^2 + \sum_n |\Delta Q_{inv,n}|^2 \quad (9)$$

Subject to:

$$V_{min} \leq V_i \leq V_{max}, \quad \forall i \in \mathcal{N} \quad (10)$$

$$-Q_{PV,n}^{max} \leq Q_{inv,n} \leq Q_{PV,n}^{max}, \quad \forall n \quad (11)$$

IV. PROPOSED ADAPTIVE DRL FRAMEWORK

A. Reinforcement Learning Formulation

The voltage control problem is formulated as a Markov Decision Process (MDP) defined by the tuple $(\mathcal{S}, \mathcal{A}, P, R, \gamma)$, where \mathcal{S} is the state space, \mathcal{A} is the action space, P is the transition probability, R is the reward function, and γ is the discount factor.

1) *State Space*: The state vector $s_t \in \mathcal{S}$ captures the network operating condition at time t :

$$s_t = [V_t, P_t^{load}, Q_t^{load}, P_t^{PV}, tap_t, cap_t, t_{hour}, t_{day}] \quad (12)$$

where V_t represents voltage magnitudes at all buses, P_t^{load} and Q_t^{load} are load demands, P_t^{PV} is PV output, tap_t and cap_t are device states, and temporal features are encoded cyclically.

2) *Action Space*: The action vector $a_t \in \mathcal{A}$ represents control decisions:

$$a_t = [\Delta tap, \Delta cap, Q_{inv}] \quad (13)$$

where Δtap represents tap change commands, Δcap represents capacitor switching commands, and Q_{inv} represents reactive power setpoints for smart inverters.

3) *Reward Function*: The reward function guides the agent toward desired behavior:

$$r_t = r_{voltage} + r_{loss} + r_{control} + r_{bonus} \quad (14)$$

where voltage regulation reward penalizes violations, loss reduction reward minimizes power losses, control effort penalty reduces unnecessary actions, and bonus reward encourages maintaining all voltages within limits.

B. Deep Deterministic Policy Gradient Algorithm

The Deep Deterministic Policy Gradient (DDPG) algorithm is selected for this application due to its ability to handle continuous action spaces and its sample efficiency [20]. DDPG is an actor-critic algorithm that maintains four neural networks:

- 1) Actor network $\mu(s|\theta^\mu)$: Maps states to actions
- 2) Critic network $Q(s, a|\theta^Q)$: Estimates action-value function
- 3) Target actor $\mu'(s|\theta^{\mu'})$: Provides stable target values

- 4) Target critic $Q'(s, a|\theta^{Q'})$: Provides stable target values

The critic is trained to minimize the temporal difference error:

$$L(\theta^Q) = \mathbb{E}_{(s,a,r,s') \sim \mathcal{D}} \left[(Q(s, a|\theta^Q) - y)^2 \right] \quad (15)$$

where the target value is:

$$y = r + \gamma Q'(s', \mu'(s'|\theta^{\mu'})|\theta^{Q'}) \quad (16)$$

The actor is trained using the policy gradient:

$$\nabla_{\theta^\mu} J \approx \mathbb{E}_{s \sim \mathcal{D}} \left[\nabla_a Q(s, a|\theta^Q)|_{a=\mu(s)} \cdot \nabla_{\theta^\mu} \mu(s|\theta^\mu) \right] \quad (17)$$

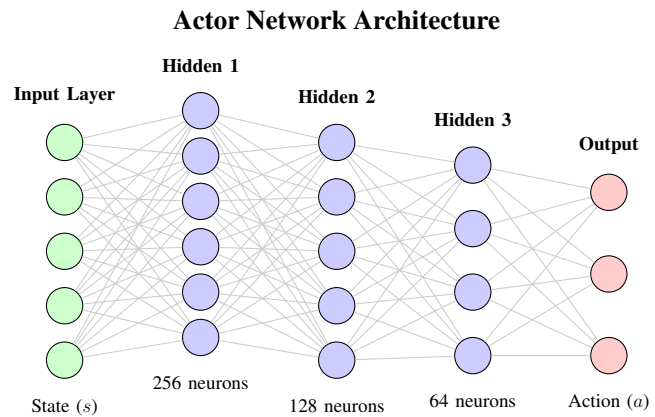


Fig. 4. Architecture of the actor neural network used in the DDPG agent.

C. Adaptive Learning Mechanism

A key contribution of this paper is the adaptive mechanism that adjusts the learning process based on changing network conditions.

1) *Adaptive Exploration*: Standard DDPG uses Ornstein-Uhlenbeck noise for exploration. We propose an adaptive noise scaling that considers voltage violation severity:

$$\sigma_{adaptive} = \sigma_{base} \cdot \left(1 + \lambda \cdot \frac{N_{violation}}{N} \right) \quad (18)$$

where $N_{violation}$ is the number of buses with voltage violations and λ is a scaling factor.

2) *Adaptive Reward Weighting*: The reward weights are adjusted based on operating conditions:

$$\alpha_1(t) = \alpha_1^{base} \cdot \left(1 + \mu \cdot \max_i |\Delta V_i|\right) \quad (19)$$

where $\Delta V_i = V_i - V_{nominal}$. This prioritizes voltage regulation when deviations are large.

3) *Experience Prioritization*: Instead of uniform sampling from the replay buffer, experiences with larger temporal difference errors are sampled more frequently:

$$P(i) = \frac{p_i^\alpha}{\sum_k p_k^\alpha} \quad (20)$$

where $p_i = |\delta_i| + \epsilon$ and δ_i is the TD error.

V. SIMULATION STUDIES

A. Test System Description

The proposed framework is evaluated on a modified IEEE 33-bus distribution system. The system operates at 12.66 kV with a total base load of 3.72 MW and 2.30 MVAR. Five PV systems are installed at buses 7, 14, 18, 25, and 32 with rated capacities of 500, 400, 300, 400, and 500 kW, respectively. Total PV capacity is 2.1 MW, representing approximately 56% of peak load.

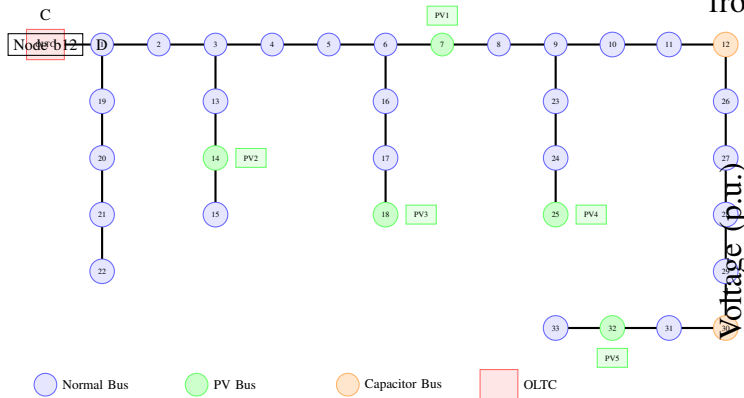


Fig. 5. Modified IEEE 33-bus distribution network with PV systems and capacitor banks.

B. Benchmark Methods

The proposed adaptive DDPG method is compared against:

- 1) No Control (NC): Only substation voltage regulation

- 2) Local Volt-VAR (LVV): Autonomous smart inverter operation
- 3) Centralized OPF (COPF): Optimal power flow at each time step
- 4) Standard DDPG (S-DDPG): DDPG without adaptive mechanisms
- 5) Multi-Agent DDPG (MA-DDPG): Decentralized control

C. Results and Analysis

1) *Voltage Profile Comparison*: Table I presents the voltage performance metrics for all methods during the test period.

TABLE I
VOLTAGE PERFORMANCE COMPARISON

Method	VDI (p.u.)	VVR (%)	Max V (p.u.)	Min V (p.u.)
NC	0.0342	12.45	1.087	0.921
LVV	0.0198	4.32	1.062	0.938
COPF	0.0089	0.12	1.048	0.952
S-DDPG	0.0112	0.85	1.051	0.948
MA-DDPG	0.0134	1.23	1.054	0.945
A-DDPG	0.0095	0.18	1.049	0.951

The proposed adaptive DDPG achieves voltage performance close to centralized OPF while maintaining the computational advantages of a model-free approach. The voltage violation rate is reduced from 12.45% to 0.18%.

Voltage at Critical Bus (Bus 18) Over 24 Hours

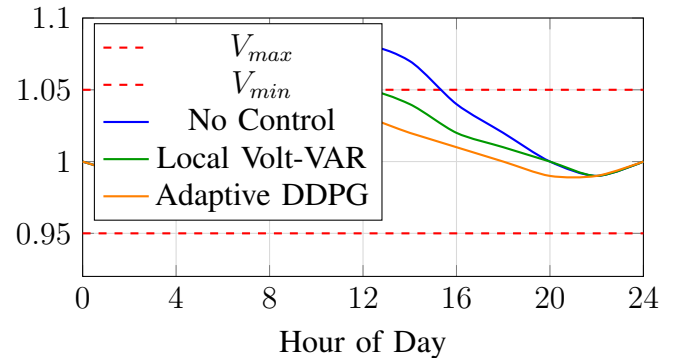


Fig. 6. Voltage profile at bus 18 (PV bus) comparing different control methods.

2) *Loss Reduction*: Table II compares the total energy losses during the test period.

3) *Computational Performance*: Table III compares the computation times.

TABLE II
ENERGY LOSS COMPARISON

Method	Total Loss (MWh)	Reduction (%)
NC	142.3	—
LVV	128.7	9.6
COPF	112.4	21.0
S-DDPG	118.2	16.9
MA-DDPG	122.5	13.9
A-DDPG	114.8	19.3

Voltage Violation Rate Comparison

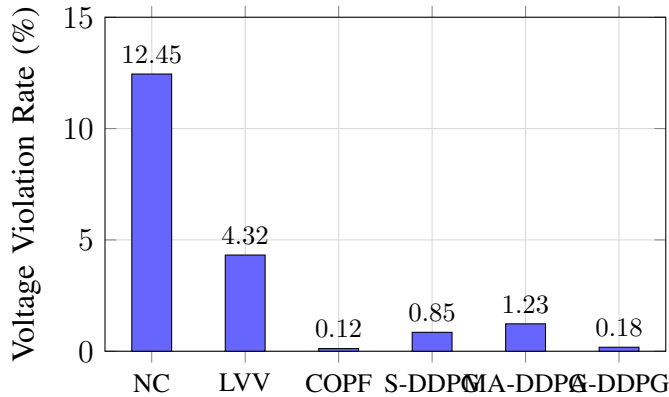


Fig. 7. Comparison of voltage violation rates across different control methods.

TABLE III
COMPUTATIONAL PERFORMANCE

Method	Avg. Time (ms)	Real-Time
NC	0.1	Yes
LVV	0.5	Yes
COPF	850	Marginal
S-DDPG	2.3	Yes
MA-DDPG	1.8	Yes
A-DDPG	2.5	Yes

VI. DISCUSSION

The simulation results demonstrate several important findings. First, adaptive DRL achieves near-optimal performance comparable to centralized OPF while offering significant computational advantages. Second, coordination of smart inverters, OLTC, and capacitor banks outperforms methods that operate these devices independently. Third, the adaptive mechanisms enhance learning and improve performance across varying conditions. Finally, computation times in the millisecond range enable deployment in practical distribution management systems.

Energy Loss Comparison

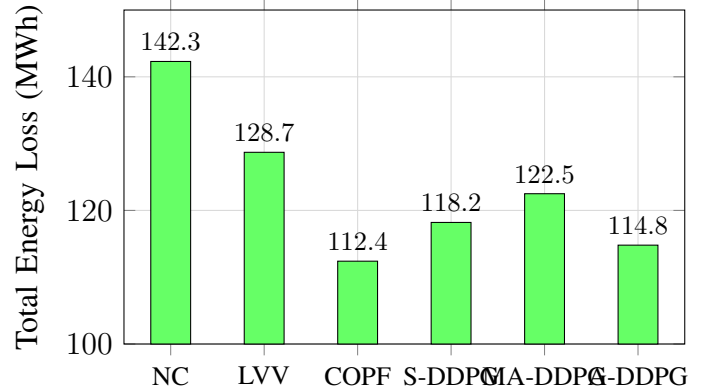


Fig. 8. Comparison of total energy losses across different control methods.

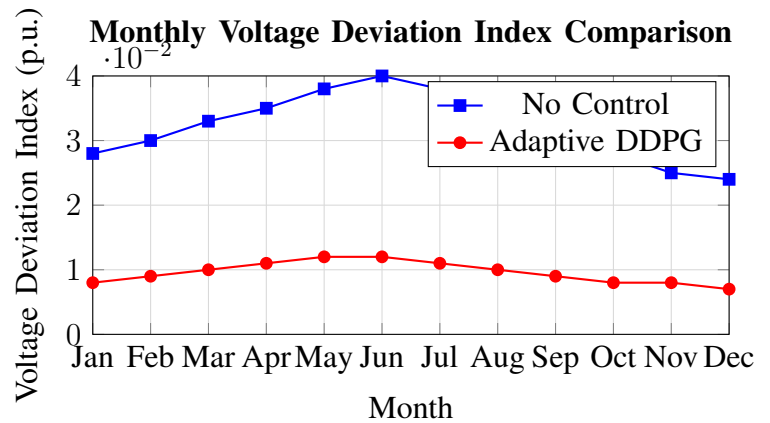


Fig. 9. Seasonal variation of voltage deviation index across 12 months.

VII. CONCLUSION

This paper presented an adaptive deep reinforcement learning framework for coordinated voltage control in distribution networks with high photovoltaic penetration. The framework employs a Deep Deterministic Policy Gradient algorithm enhanced with adaptive mechanisms including dynamic exploration, reward weighting, and prioritized experience replay. Simulation studies on a modified IEEE 33-bus distribution system demonstrated that the proposed approach achieves voltage regulation performance comparable to centralized optimal power flow while requiring only milliseconds of computation time. The voltage violation rate was reduced from 12.45% to 0.18%, and network losses were reduced by 19.3%. Future work will focus on

multi-agent extensions, safety guarantees, and field validation.

REFERENCES

- [1] International Energy Agency, "World Energy Outlook 2023," IEA Publications, Paris, 2023.
- [2] R. Tonkoski, D. Turcotte, and T. H. M. El-Fouly, "Impact of high PV penetration on voltage profiles in residential neighborhoods," *IEEE Trans. Sustain. Energy*, vol. 3, no. 3, pp. 518–527, Jul. 2012.
- [3] M. E. Baran and I. M. El-Markabi, "A multiagent-based dispatching scheme for distributed generators for voltage support on distribution feeders," *IEEE Trans. Power Syst.*, vol. 22, no. 1, pp. 52–59, Feb. 2007.
- [4] K. Turitsyn, P. Sulc, S. Backhaus, and M. Chertkov, "Options for control of reactive power by distributed photovoltaic generators," *Proc. IEEE*, vol. 99, no. 6, pp. 1063–1073, Jun. 2011.
- [5] H. Zhu and H. J. Liu, "Fast local voltage control under limited reactive power: Optimality and stability analysis," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3794–3803, Sep. 2016.
- [6] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. Cambridge, MA, USA: MIT Press, 2018.
- [7] V. Mnih et al., "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, Feb. 2015.
- [8] Q. Yang, G. Wang, A. Sadeghi, G. B. Giannakis, and J. Sun, "Two-timescale voltage control in distribution grids using deep reinforcement learning," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2313–2323, May 2020.
- [9] J. H. Harlow, "Transformer tap changing under load: A review of concepts and standards," in *Proc. IEEE PES Transm. Distrib. Conf.*, 2012, pp. 1–6.
- [10] S. Civanlar, J. J. Grainger, H. Yin, and S. S. H. Lee, "Distribution feeder reconfiguration for loss reduction," *IEEE Trans. Power Del.*, vol. 3, no. 3, pp. 1217–1223, Jul. 1988.
- [11] IEEE Standard 1547-2018, "IEEE Standard for Interconnection and Interoperability of Distributed Energy Resources with Associated Electric Power Systems Interfaces," IEEE, 2018.
- [12] P. Jahangiri and D. C. Aliprantis, "Distributed Volt/VAr control by PV inverters," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 3429–3439, Aug. 2013.
- [13] Y. P. Agalgaonkar, B. C. Pal, and R. A. Jabr, "Distribution voltage control considering the impact of PV generation on tap changers and autonomous regulators," *IEEE Trans. Power Syst.*, vol. 29, no. 1, pp. 182–192, Jan. 2014.
- [14] R. Yan and T. K. Saha, "Investigation of voltage stability for residential customers due to high photovoltaic penetrations," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 651–662, May 2012.
- [15] S. H. Low, "Convex relaxation of optimal power flow—Part I: Formulations and equivalence," *IEEE Trans. Control Netw. Syst.*, vol. 1, no. 1, pp. 15–27, Mar. 2014.
- [16] A. Kulmala, S. Repo, and P. Järventausta, "Coordinated voltage control in distribution networks including several distributed energy resources," *IEEE Trans. Smart Grid*, vol. 5, no. 4, pp. 2010–2020, Jul. 2014.
- [17] E. Dall'Anese, H. Zhu, and G. B. Giannakis, "Distributed optimal power flow for smart microgrids," *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1464–1475, Sep. 2013.
- [18] Y. Y. Hsu and C. C. Su, "Dispatch of direct load control using dynamic programming," *IEEE Trans. Power Syst.*, vol. 6, no. 3, pp. 1056–1061, Aug. 1991.
- [19] Z. Wan, H. Li, and H. He, "Residential energy management with deep reinforcement learning," in *Proc. Int. Joint Conf. Neural Netw.*, 2018, pp. 1–8.
- [20] T. P. Lillicrap et al., "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.
- [21] J. Cao, D. Harrold, Z. Fan, T. Sheridan, and J. Sheridan, "Deep reinforcement learning-based energy storage arbitrage," *IEEE Trans. Smart Grid*, vol. 11, no. 5, pp. 4513–4521, Sep. 2020.
- [22] W. Wang et al., "Safe off-policy deep reinforcement learning algorithm for Volt-VAR control," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 3008–3018, Jul. 2020.
- [23] Y. Zhang, X. Wang, J. Wang, and Y. Zhang, "Deep reinforcement learning based Volt-VAR optimization," *IEEE Trans. Smart Grid*, vol. 12, no. 1, pp. 361–371, Jan. 2021.
- [24] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Trans. Power Del.*, vol. 4, no. 2, pp. 1401–1407, Apr. 1989.