# Coordinated, Data-Driven Voltage Regulation in Wind-plus-Solar Power Distribution Systems with Energy Storage

Miguel Jimenez-Aparicio
*Electric Power Systems Research*
*Sandia National Laboratories*
Albuquerque, NM, USA

Rachid Darbali-Zamora
*Renewable Energy and Distributed Systems Integration*
*Sandia National Laboratories*
Albuquerque, NM, USA

2025 Resilience Week (RWS) | 979-8-3315-5687-7/25/$31.00 ©2025 IEEE | DOI: 10.1109/RWS66711.2025.11304349

***Abstract*—A data-driven, multi-agent Reinforcement Learning (RL)-based controller is presented for coordinated voltage regulation in power distribution systems integrating wind, solar, and energy storage resources. The proposed approach aims to provide feeder-wide voltage control, optimize energy storage scheduling, and mitigate Distributed Energy Resource (DER) curtailment, thereby enhancing power quality and system stability. Through simulations on a real-time platform, the RL controller demonstrates significant improvements in voltage deviation reduction and standard deviation at critical points, effectively shifting voltage distributions closer to nominal values. The results highlight the economic benefits of optimal energy storage operation and the controller's ability to leverage coordinated reactive power contributions from all DERs to address sudden generation swings and maintain grid stability.**

***Index Terms*—Distributed Energy Resources, Voltage Regulation, Reinforcement Learning, Energy Storage Systems.**

## I. Introduction

The integration of Distributed Energy Resources (DERs) like wind, solar, and energy storage into power distribution systems presents significant challenges, particularly concerning voltage regulation and grid stability. Advanced control and energy management strategies are crucial for ensuring the safe and efficient operation of these increasingly smarter power grids. For instance, energy storage can absorb excess generation during peak production and release it during periods of lower generation. For real-time voltage regulation, reactive power compensation devices - such as capacitor banks - and on-load tap changers are essential for dynamic adaptation to changing conditions. Recent industry standards, such as IEEE 1547-2018, encourage DERs to provide voltage regulation at their Point Of Interconnection (POI) [1]. Relying on DER grid support functionalities is vital for a reliable transition to an inverter-dominated energy future. Hybrid systems, which often combine complementary sources like wind and solar, offer resilient, local, and relatively constant electricity at an affordable cost. Microgrids, for example, can be strategically placed in weak areas of the power grid to defer or even eliminate the need for system upgrades [2]. An Distributed Energy Resources Management System (DERMS) is necessary to ensure coordination among DERs.

This article has been authored by an employee of National Technology & Engineering Solutions of Sandia, LLC under Contract No. DE-NA0003525 with the U.S. Department of Energy (DOE). The employee owns all right, title and interest in and to the article and is solely responsible for its contents. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this article or allow others to do so, for United States Government purposes. The DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan https://www.energy.gov/downloads/doe-public-access-plan.

The authors would like to acknowledge the financial support provided by the U.S. Department of Energy (DOE), Office of Energy Efficiency (EERE), Wind Energy Technology Office (WETO) for the Wind Hybrid Integration Platform (WHIP) project under Work Breakdown Structure No. 1.2.6.701 as well as the contributions from Bret Barker and Patrick Gilman.

## II. Background

### A. Voltage Regulation

Traditionally, distribution voltage regulation has been managed using step voltage regulators [3]. These devices adjust taps every few hours and incorporate a response delay to prevent unnecessary changes from temporary conditions and excessive wear [4]. However, this delay renders them less effective in systems with high renewable energy penetration due to the dynamic nature of renewable generation.

More recently, the effectiveness and timely response of autonomous voltage regulation provided by DER inverters have been demonstrated in simulation [5], [6] and experimentally [7], [8]. The paper in [9] discusses the advantages of both autonomous and coordinated controls. Droop-based control, for instance, does not necessitate communication between Inverter-Based Resources (IBRs). However, it faces challenges such as load-dependent frequency/voltage deviation, issues with non-linear loads, and poor reactive power sharing in the presence of unequal buses. Conversely, droop-free control demands robust communication between IBRs, making it susceptible to voltage and frequency instability in the event of communication failures. Coordinated controls of solar and wind inverters have also been proposed to mitigate voltage deviations across distribution feeders [10], [11]. Furthermore, a normalized sign least mean fourth control has been shown to offer power quality solutions including harmonic suppression, voltage regulation, and reactive power compensation [12].

Advanced control methodologies for voltage regulation are increasingly leveraging Artificial Intelligence (AI) and Ma-

Authorized licensed use limited to: Sandia National Laboratories. Downloaded on January 06,2026 at 07:20:32 UTC from IEEE Xplore. Restrictions apply.

chine Learning (ML) algorithms, such as Multi-Agent Reinforcement Learning (MARL) [11], [13], [14]. However, it is noted that these algorithms often do not explicitly consider the integration of energy storage and the associated challenges in voltage regulation.

### B. Energy Storage Scheduling in Wind-Solar Microgrids

The optimized utilization of energy storage technologies offers significant cost-savings in various smart power grid applications, enabling charging during low electricity retail prices and discharging when prices are higher [15]. Managing a hybrid microgrid with an Energy Storage System (ESS) presents a complex optimization problem. Reference [16] explores a capacity allocation method for ESS that minimizes costs by considering comprehensive installation and operation costs, as well as power purchase costs. This approach also aims to reduce renewable energy curtailment through the use of a genetic algorithm to solve the optimization problem.

The work in [17] proposes utilizing typically inactive Photovoltaic (PV) inverters during night-time to charge an ESS device. If the Point of Common Coupling (PCC) voltage exceeds a threshold, ESS charging is activated. During the day, the battery supplies power when renewable energy production is insufficient to meet the load. Other approaches include environmental and economic dispatch optimization using a grey wolf group optimization technique, coupled with a non-parametric kernel density estimation method to derive typical wind and solar data [18]. Additionally, an ant colony algorithm has been employed for optimizing day-ahead scheduling in wind-solar-storage microgrids, incorporating investment and amortization costs in the constraints and demonstrating potential cost reductions exceeding 20% [19].

### C. Contributions

This paper main contribution is the development of a multi-agent, data-driven, MARL-based voltage regulation controller specifically for systems integrating wind, solar, and ESS resources. The controller is designed for coordinated, real-time, feeder-wide voltage control, extending beyond just the DER interconnection point. It considers multi-day ahead energy storage scheduling and aims to avoid DER curtailment, ultimately improving power quality and system stability through AI-driven control. The emphasis on joint operation highlights the need for real-time control to address inaccuracies in forecasting and manage sudden transients or drops in generation.

## III. Methodology

### A. AI-driven Voltage Regulation

RL is the branch of AI/ML that explores the training of agents according to a reward function. Applications with more than one agent are commonly referred to as MARL. The agents can interact with the environment and receive a reward. The goal of the training is to find the actions that, given an observation, optimize the long-term reward. MARL agent's training presents significant challenges, such as 1) the inherent environment non-stationarity from any individual agent's perspective, as other agents' policies continuously evolve, 2) the intricate credit assignment problem in shared reward scenarios, and 3) the complexities of coordination and communication in partially observable environments.

The coordinated voltage regulation using DERs' reactive power control is modeled as a MARL problem. Each DER in the system is an independent agent, and the available action is the reactive power injection or absorption (assuming active power priority, which limits the reactive power contribution in order not to exceed the device's power rating). Deep Deterministic Policy Gradient (DDPG) is the selected RL algorithm due to the continuous action space. For training stability purposes, and given that all inverter-based DERs have similar dynamics, all agents actually share the same neural network structure and weights. Agents do not know of other agents' action, but they do have access to a global weighted average voltage metric that is calculated by the DERMS. Agents are trained over hundreds of episodes with different generation and load conditions to enable generalization.

DDPG is an off-policy, actor-critic algorithm that employs two main neural networks: the Actor $\mu$ and the Critic $\Upsilon$, each with a primary network and a "target" network. These target networks are updated slowly using a "soft update" mechanism to stabilize training. The Actor network is the agent's "policy." Its primary purpose is to learn a deterministic policy that maps an observed state to a specific, continuous action. On the other hand, the Critic network is the agent's "evaluator." Its purpose is to estimate the action-value function (Q-function), which predicts the expected cumulative future reward (Q-value) for taking a specific action in a given state. The critic provides a quality score for the actor's chosen actions, guiding the actor's learning. Algorithm 1 describes the DDPG training procedure. The input arguments are the experience or replay buffer length $D$, the number of training episodes $M$, the number of time steps per episode $T$, the batch size $Z$, the number of agents $H$, the discount factor $\gamma$, and the target network's update rate $\tau$. The agents' state S is comprised of 6 measurements: voltage $V$, frequency $f$, active power $P$ injection, reactive power $Q$ contribution, the group reactive power contribution $Q_g$, and a global, weighted voltage average $V_{avg}$ that is defined as:

$$V_{avg} = \sum_{n=1}^{N} \omega_n \cdot V_n \tag{1}$$

where $N$ is the number of system-wide measurements, $V$ is the measured voltage value, and $\omega$ is the weight for each sample. All of the measurements in $S$ are in per unit (p.u.) values. Note that the first 4 measurements can be provided by any DER device according to the IEEE 1547-2018 standard. The group reactive power contribution and the global voltage average are calculated by the DERMS in every control time step and passed to the individual DER agents. The action $A$ is a reactive power setpoint $Q \in [-1, 1]$ p.u. The structure and layer size of the actor and critic networks are shown in Fig. 1.

Authorized licensed use limited to: Sandia National Laboratories. Downloaded on January 06,2026 at 07:20:32 UTC from IEEE Xplore. Restrictions apply.

The reward function $R_h$ is calculated in every control time step $i$ for each agent $h$, and is expressed as:

$$R_{h,i} = -\big[A_i(V_{avg}) + B_i(Q_{h,i},\ Q_{g,i}(h)) + C_i(Q_{h,i}, Q_{h,i-1})\big] \tag{2}$$

where:

$$A_i(V_g) = a_1 \cdot V_{avg}^2 + a_2 \cdot V_{avg} + a_3 \tag{3}$$

$$\begin{gathered} B_i(Q_{h,i},\ Q_{g,i}(h)) = b_1 \cdot Q_{dev}^2 + b_2 \cdot Q_{dev} + b_3 \\ Q_{dev} = |Q_{h,i} - Q_{g,i}(h)| \\ Q_{g,i}(h) = \sum_{k=1}^{K} Q_{k,i} \text{ if } k \in h\text{'s group} \end{gathered} \tag{4}$$

$$\begin{gathered} C_i(Q_{h,i}, Q_{h,i-1}) = c_1 \cdot Q_{var}^2 + c_2 \cdot Q_{var} + c_3 \\ Q_{var} = |Q_{h,i} - Q_{h,i-1}| \end{gathered} \tag{5}$$

The main training objective is to maintain the average weighted voltage metric as close to 1 p.u. as possible, as formulated in Eq. 3. In addition, the RL controller needs to ensure system stability by minimizing unnecessary voltage oscillations and contradictory behaviors. This is expressed as a penalization on significant reactive power mismatches with nearby DER devices, as shown in Eq. 4, and on abrupt changes in consecutive actions, as described in Eq. 5.

### B. Use Case

The test system is the SWiFT facility in Lubbock, Texas (USA). Fig. 1 shows the one-line diagram of the system. The distribution feeder contains both wind and solar generation. A group of 3 WTGs is connected to Bus 4 [20]. A solar PV array, a variable load, and a battery ESS are located at the end of the feeder in Bus 10 [21], [22]. Table I summarizes the power rating of all devices and their group assignment. A total of 62 hours of solar irradiance, wind speed, and load data are available with a 1-second resolution. To speed up the simulation process, we use a real-time simulation platform for both training and testing. The digital twin of a relevant distribution feeder is simulated in OPAL-RT. Meteorological data, observations, and actions are exchanged every 1 second using a Modbus I/O interface [23]. Under certain generation and demand patterns, severe voltage deviations are experienced

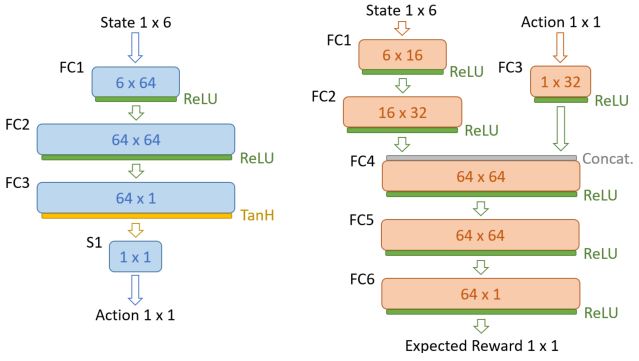



Fig. 1: Actor (left) and critic (right) networks.

**Algorithm 1** Multi-Agent DDPG Training

**procedure** TRAINDDPGAGENTS($D, M, T, Z, H, \gamma, \tau$)
  Initialize actor network $\mu$ and critic network $\Upsilon$ with random weights $\theta$, $\phi$, respectively.
  Initialize target networks $\mu'$ and $\Upsilon'$ with weights $\theta_t \leftarrow \theta$ and $\phi_t \leftarrow \phi$, respectively.
  Initialize replay buffer $\mathcal{B}$ with capacity $D$.
  Initialize Ornstein-Uhlenbeck noise process $\mathcal{N}$.
  **for** episode $m \leftarrow 1$ **to** $M$ **do**
    Initialize wind speed, irradiance, load, and ESS (dis)charging profile arrays of size $T$.
    **for** time step $i \leftarrow 1$ **to** $T$ **do**
      **for** agent $h \leftarrow 1$ **to** $H$ **do**
        Observe state $S_{i,h}$.
        Select action (with noise) $A_{i,h} = \pi(S_{i,h}|\theta) + \mathcal{N}_i$.
        Execute action $A_{i,h}$.
        Observe reward $R_{i,h}$.
        Receive next state $S_{i+1,h}$.
        Store transition $(S_{i,h}, A_{i,h}, R_{i,h}, S_{i+1,h})$ in $\mathcal{B}$.
      **end for**
      Sample a random batch of $Z$ transitions $(S_j, A_j, R_j, S_{j+1})$ from $\mathcal{B}$.
      Calculate $y_j = r_j + \gamma \Upsilon_t(S_{j+1}, \pi_t(S_{j+1}|\theta_t)|\phi_t)$ using Bellman equation.
      Update critic by minimizing the loss $L = \frac{1}{Z}\sum_j (y_j - Q(S_j, A_j|\phi))^2$.
      Update actor using the sampled policy gradient $\nabla_\theta J \approx \frac{1}{Z}\sum_{i=1}^{Z} \nabla_A Q(S_i, A_i|\phi) \nabla_\theta \pi(S_i|\theta)$.
      Update target networks $\theta_t \leftarrow \tau\theta + (1-\tau)\theta_t$, $\phi_t \leftarrow \tau\phi + (1-\tau)\phi_t$.
    **end for**
  **end for**
**end procedure**

at the end of the feeder. Voltage measurements are available on every bus. In order to calculate Eq. 1, the buses' voltage's weight is $\omega = 1$, the DER's voltage is weighted $\omega = 2$, and the load's voltage weight is $\omega = 16$. The load is considered a critical service.

### C. Battery Scheduling Methodology

The battery schedule is generated using linear programming optimization (in particular, with Python's *PuLP* library) and considering constraints such as the battery (dis)charging limits and energy balance to achieve an efficient and sustainable

TABLE I: Device Enrollment

| Device | Power rating (kVA) | Capacity (kWh) | Group |
|---|---|---|---|
| WTG1 | 200 | - | 1 |
| WTG2 | 200 | - | 1 |
| WTG3 | 200 | - | 1 |
| PV1 | 150 | - | 2 |
| ESS1 | 60 | 130 | 2 |

Authorized licensed use limited to: Sandia National Laboratories. Downloaded on January 06,2026 at 07:20:32 UTC from IEEE Xplore. Restrictions apply.

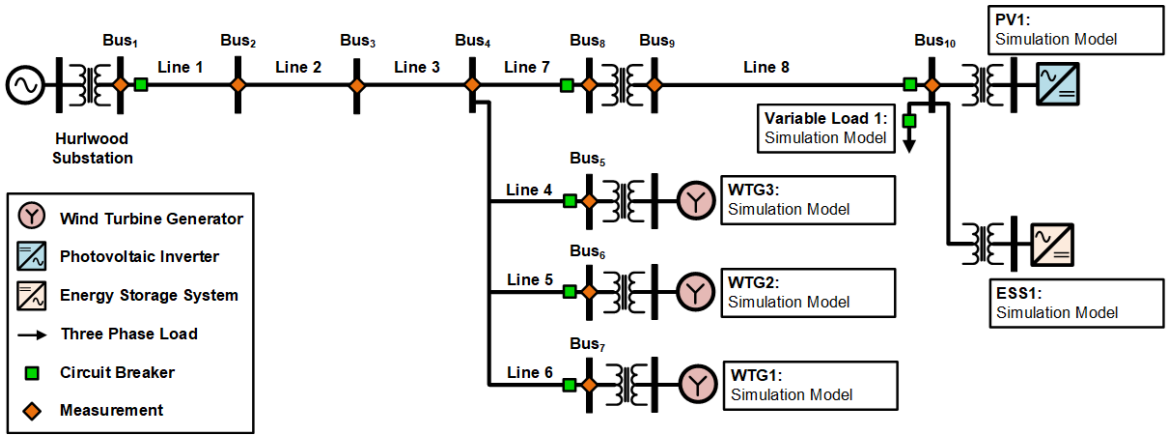



Fig. 2: One-line diagram of the SWiFT facility [1].

energy management strategy. Hour-ahead forecasting is necessary for determining uncertainty in wind and solar generation. Taking into account hourly solar irradiation, wind speed, and load forecasts, optimal ESS schedules are generated.

The goal of the optimization problem, described in Eq. 6, is to minimize the cost of importing power while maximizing the gains of exporting power. This power exchange $P_{EXCH}$ can be either positive (importing) or negative (exporting). The energy price $\rho$ is time-dependent. Eq. 7 summarizes the energy balance in the system. As shown in Eqs. 8-9, charging and discharging setpoints are set to 0 for the first hour and last hour in the scheduling period. Eq. 10 represents ESS' SOC equation. The ESS efficiency $\eta$ is considered during both charging $P_{ESS,C}$ and discharging $P_{ESS,D}$ operations. Finally, Eqs. 11-17 contains the lower and upper bounds for each of the DERs' active power and the ESS' SOC. The formulation of this optimization problem can be used for any number of hours as long as forecasted generation and load data are available.

$$\min \sum_{h=0}^{H} (\rho[h] \cdot P_{EXCH}[h]) \tag{6}$$

subject to:

$$\begin{aligned} &P_{WTG1}[h] + P_{WTG2}[h] + P_{WTG3}[h] + P_{PV1}[h] \\ &+ P_{ESS,D}[h] - P_{ESS,C}[h] + P_{EXCH}[h] \\ &= P_{LOAD}[h] \quad \forall h \in \{0, 1, ..., H\} \end{aligned} \tag{7}$$

$$P_{ESS,D}[H] = P_{ESS,D}[0] = 0 \tag{8}$$

$$P_{ESS,C}[H] = P_{ESS,C}[0] = 0 \tag{9}$$

$$SOC[h] = SOC[h-1] + \eta \cdot P_{ESS,C}[h] - \frac{P_{ESS,D}[h]}{\eta} \tag{10}$$

$$0 \leq P_{WTG1}[h] \leq P_{WTG1,Nom}[h] \quad \forall h \in \{0, 1, ..., H\} \tag{11}$$

$$0 \leq P_{WTG2}[h] \leq P_{WTG2,Nom}[h] \quad \forall h \in \{0, 1, ..., H\} \tag{12}$$

$$0 \leq P_{WTG3}[h] \leq P_{WTG3,Nom}[h] \quad \forall h \in \{0, 1, ..., H\} \tag{13}$$

$$0 \leq P_{PV1}[h] \leq P_{PV1,Nom}[h] \quad \forall h \in \{0, 1, ..., H\} \tag{14}$$

$$0 \leq P_{ESS,C}[h] \leq P_{ESS,C,Nom}[h] \quad \forall h \in \{0, 1, ..., H\} \tag{15}$$

$$0 \leq P_{ESS,D}[h] \leq P_{ESS,D,Nom}[h] \quad \forall h \in \{0, 1, ..., H\} \tag{16}$$

$$SOC_{\min} \leq SOC[h] \leq SOC_{\max} \quad \forall h \in \{0, 1, ..., H\} \tag{17}$$

## IV. Results

Training parameters are gathered in Table II. The reward function coefficients are determined heuristically. The evolution of the episodic reward for the RL controller training system-wide voltage regulation is shown in Fig. 3. For training purposes, the available 62-hour time series with 1-second resolution are distributed across 200 episodes, each 100 seconds long, in which irradiance, wind speed, and load profiles are randomly extracted from the original data to guarantee generalization. The battery (dis)charging setpoints are sampled from the generated optimal scheduling for this 62-hour interval, which is shown in Fig. 4.

Note that the forecasting in Fig. 4a estimates very high wind generation around t=15 hours, coupled with a relatively cloudy sky and mild solar generation. After that initial period, the rest of the time series indicates periods of mild wind generation

TABLE II: Training Parameters

| Parameter(s) | Value(s) |
|---|---|
| $D$ | 30e3 |
| $M$ | 200 |
| $T$ | 90 |
| $Z$ | 64 |
| $H$ | 5 |
| $\gamma$ | 0.99 |
| $\tau$ | 0.01 |
| $a_1, a_2, a_3$ | 30.6e3, -61.4e3, 30.7e3 |
| $b_1, b_2, b_3$ | 240, 0, 0 |
| $c_1, c_2, c_3$ | 240, 0, 0 |

Authorized licensed use limited to: Sandia National Laboratories. Downloaded on January 06,2026 at 07:20:32 UTC from IEEE Xplore. Restrictions apply.

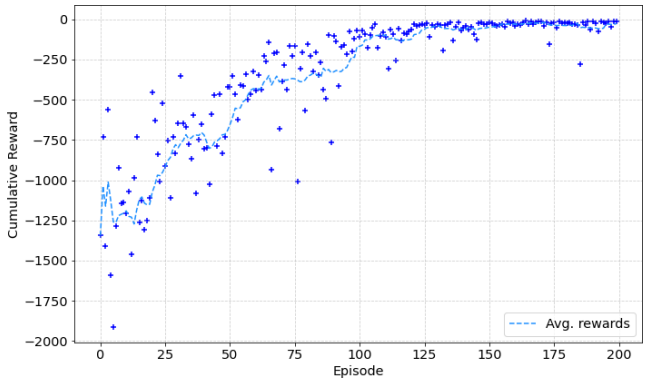


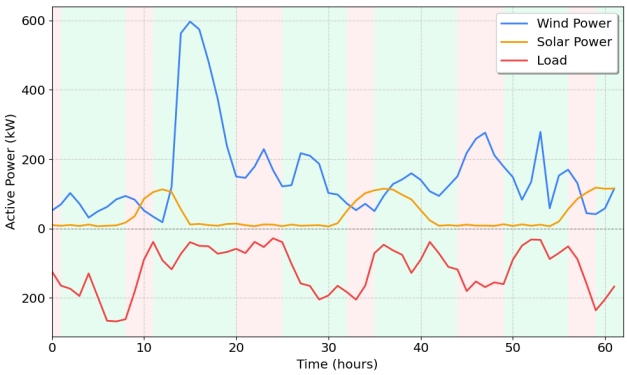

Fig. 3: Episodic reward.

and optimal solar generation. The estimated load profile foresees a larger demand in the early hours of the day, which could be coupled with electric vehicle charging. The green and red backgrounds indicate electricity retail prices, \$0.25/kWh and \$0.15/kWh, respectively. As seen in Fig. 4b, the optimal ESS schedule involves 2 full charging and discharging cycles per day. Looking at the results, charging happens during periods of low electricity prices, while discharging tends to happen when purchasing electricity is more expensive. The overall power import and export balance is shown in Fig. 4c. The effect of the ESS operation can be noted as increased import during cheaper electricity periods (note, for example, the demand shifting from $t = 33$ hours to $t = 31$ hours) and maximizing power exports during high electricity price periods (as in $t = 21$ hours and $t = 45$ hours). The system demonstrated a significant economic advantage with the integration of the ESS device over the 62-hour period. A revenue increase of 11.84% was observed, indicating the profitability of incorporating an ESS into the system. The total gains achieved with ESS optimization were \$405.25, compared to \$362.34 without ESS, further underscoring its economic viability.

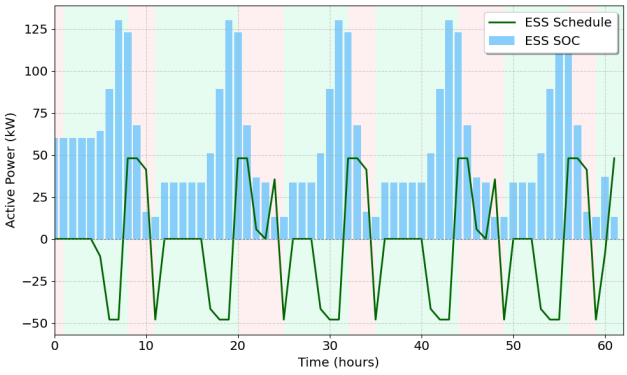
### *A. Overall Testing Performance*

The analysis of the RL controller's performance over 100 test episodes demonstrates a significant improvement in voltage regulation. Table III and Fig. 5 summarize the voltage distribution at two significant locations: the Load1 POI and the WTG1 POI (where the wind farm connects). Table III shows that the RL controller effectively brings voltages closer to their nominal values and substantially reduces the standard deviation. Specifically, the average voltage at Load1 increases by almost 0.6%, effectively halving the voltage deviation. The standard deviation at Load1 is also reduced by more than half. For WTG1, while the average voltage doesn't decrease significantly compared to the base case, its standard deviation is also cut by half.

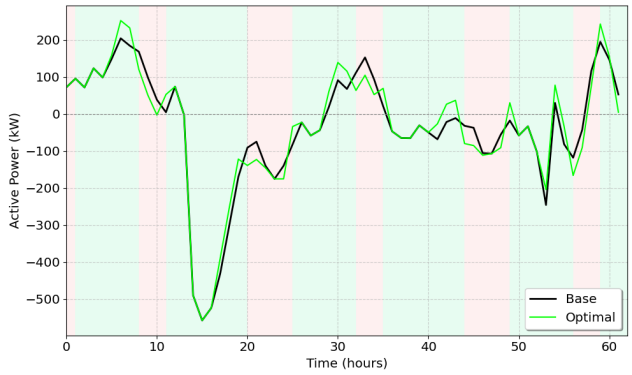
Fig. 5 provides further insight into the relationship between WTG1 and Load1. In severe under-voltage events, the voltage at Load1 can drop to 0.92 p.u., exceeding the ANSI limit of ±5%. However, with the RL controller, the voltage distribution is shifted much closer to the nominal value, effectively avoiding extreme voltage deviations.




(a) Forecasted wind power, solar power, and load profiles.




(b) Optimal ESS schedule and estimated SOC.




(c) Estimated feeder-wise power balance.

Fig. 4: Results of the ESS scheduling for a 62-hour period.

TABLE III: Testing Results for Voltage Regulation

| Case | Location | Mean (p.u.) | STD (p.u.) |
|---|---|---|---|
| Base | WTG1 | 1.0132 | 0.0107 |
| | Load1 | 0.9864 | 0.0286 |
| RL | WTG1 | **1.0119** | **0.0051** |
| | Load1 | **0.9921** | **0.0121** |

Authorized licensed use limited to: Sandia National Laboratories. Downloaded on January 06,2026 at 07:20:32 UTC from IEEE Xplore. Restrictions apply.

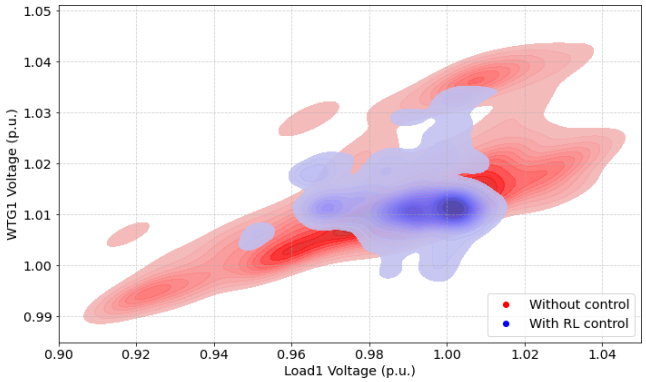



Fig. 5: Voltage distribution at WTG1's and Load1's POIs.

### B. Scenario Performance

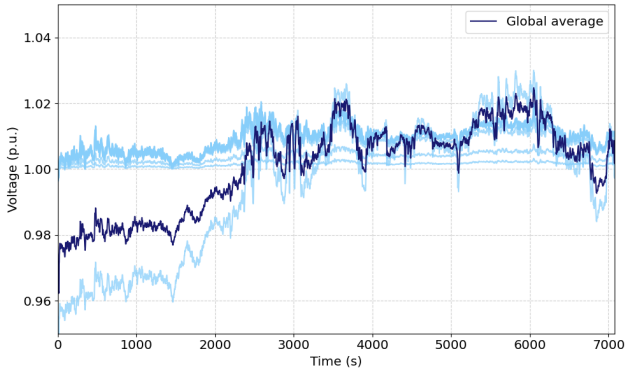
Over a 2-hour operational period, starting at $t = 9$ hours in Fig. 4, the voltage regulation results are consistent with the overall system performance. A comparison of voltage profiles is illustrated in Fig. 6. Detailed statistical analysis of voltage at Bus 10 and the global voltage across the feeder for both the base case and the RL-controlled scenario is provided in Table IV. The statistics show a reduction in the standard deviation for both Bus 10 and global voltage under RL control, indicating improved voltage stability and reduced fluctuations.

The active and reactive power contributions of WTGs, solar PV, and the ESS are presented in Fig. 7 and Fig. 8, respectively. The scenario focuses on dynamic voltage regulation over a 2-hour period highlights how voltage deviations, particularly those caused by sudden drops in solar generation (e.g., due to cloud cover), are effectively avoided through coordinated DER control actions. Even when not all DER devices are physically located in areas prone to voltage deviations, the RL control, utilizing global measurements, leverages the regulating capabilities of all available DERs to mitigate both system-wide and localized voltage deviations. For example, each WTG consistently injects approximately 0.1 p.u. of reactive power to elevate feeder bus voltages, even if their local voltages are not below 1 p.u. The solar PV system injects approximately 0.4 p.u. of reactive power during the nighttime period until sunrise. Subsequently, it switches to absorbing reactive power to prevent over-voltages during the day. Reactive power injection from the PV system is highly responsive to sudden drops in solar generation, proactively mitigating voltage swings.

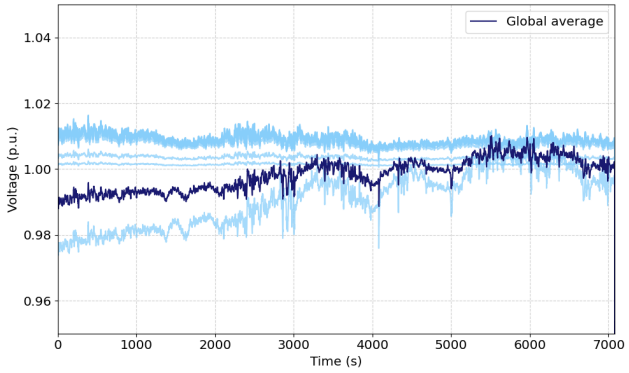
The ESS generally tends to absorb reactive power. In the initial minutes of the simulation, it exhibits a behavior opposite to that of the solar PV, suggesting that the RL agent's commands may vary based on active power contribution and/or the time of day (nighttime vs. daytime). After $t = 2500$ seconds (during daytime), the PV and ESS systems become highly coordinated, especially when both DERs are actively providing power to the grid. Turning off the battery at $t = 4100$ seconds introduces a noticeable offset in reactive power contribution, with the solar PV becoming significantly more proactive in voltage support. The individual contributions of voltage, active power, and reactive power from each DER all influence the commands issued by the RL agent.

TABLE IV: Testing Results for Voltage Regulation

| Case | Location | Mean (p.u.) | STD (p.u.) | Min (p.u.) | Max (p.u.) |
|---|---|---|---|---|---|
| Base | $V_{avg}$ | **1.0011** | 0.0130 | 0.9750 | 1.0246 |
| | Load1 | **0.9946** | 0.0212 | 0.9539 | 1.0299 |
| RL | $V_{avg}$ | 0.9981 | **0.0047** | **0.9881** | **1.0101** |
| | Load1 | 0.9897 | **0.0083** | **0.9737** | **1.0116** |




(a) Base case.




(b) RL case.

Fig. 6: Voltage profiles for the 2-hour scenario.

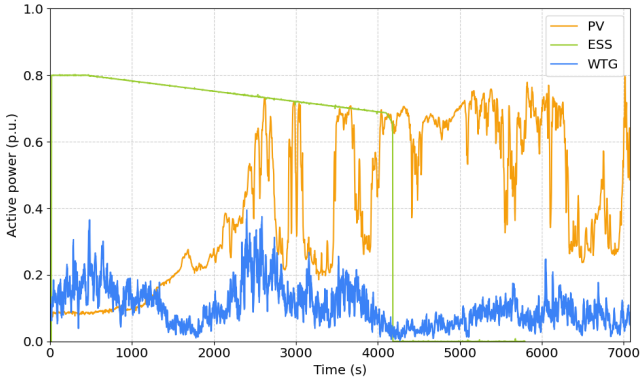



Fig. 7: Active power profiles for the 2-hour scenario.

Authorized licensed use limited to: Sandia National Laboratories. Downloaded on January 06,2026 at 07:20:32 UTC from IEEE Xplore. Restrictions apply.

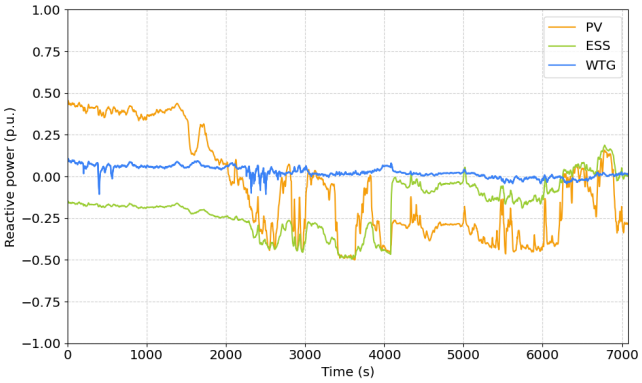



Fig. 8: Reactive power profiles for the 2-hour scenario.

## V. Conclusion

In conclusion, this paper successfully developed and validated a multi-agent, data-driven Reinforcement Learning controller that effectively manages voltage regulation in power distribution systems with high penetration of wind, solar, and energy storage. The controller's ability to coordinate DERs across the feeder, optimize battery scheduling, and mitigate voltage deviations, even during rapid generation changes, significantly improves power quality and system stability. Specifically, the average voltage at Load1 increased by almost 0.6%, effectively halving its voltage deviation, and the standard deviation at both Load1 and WTG1 was reduced by more than half. Furthermore, the integration of the ESS resulted in an 11.84% revenue increase over a 62-hour period. The demonstrated economic benefits from optimal energy storage utilization further underscore the practical value of this AI-driven approach for a more resilient and efficient renewable energy integrated grid.

Future work includes addressing cyber-physical resiliency challenges (such as communication loss, measurement errors, or cyber-attacks) and validating the approach on larger systems.

## References


[1] "IEEE Standard for Interconnection and Interoperability of Distributed Energy Resources with Associated Electric Power Systems Interfaces," *IEEE Std 1547-2018 (Revision of IEEE Std 1547-2003)*, pp. 1–138, 2018.

[2] J. O. Petinrin and M. Shaaban, "A hybrid solar PV/wind energy system for voltage regulation in a microgrid," in *2013 IEEE Student Conference on Research and Developement*, pp. 545–549, 2013.

[3] J. Dong, Y. Xue, M. Olama, T. Kuruganti, J. Nutaro, and C. Winstead, "Distribution Voltage Control: Current Status and Future Trends," in *2018 9th IEEE International Symposium on Power Electronics for Distributed Generation Systems (PEDG)*, pp. 1–7, 2018.

[4] M. Hartung, E.-M. Baerthlein, and A. Panosyan, "Comparative Study of Tap Changer Control Algorithms for Distribution Networks With High Penetration of Renewables," 2014.

[5] R. Darbali-Zamora, S. Ojetola, F. Wilches-Bernal, and J. Berg, "Development of a Wind Turbine Generator Volt-Var Curve Control for Voltage Regulation in Grid Connected Systems," in *2022 North American Power Symposium (NAPS)*, pp. 1–6, 2022.

[6] R. Darbali-Zamora and M. J. Aparicio, "Volt-var curve control using wind turbine generators and photovoltaic inverters for voltage regulation," in *2023 IEEE Kansas Power and Energy Conference (KPEC)*, pp. 1–6, 2023.

[7] J. Johnson, A. Summers, R. Darbali-Zamora, J. Hernandez-Alvidrez, J. Quiroz, D. Arnold, and J. Anandan, "Distribution voltage regulation using extremum seeking control with power hardware-in-the-loop," *IEEE Journal of Photovoltaics*, vol. 8, no. 6, pp. 1824–1832, 2018.

[8] R. Darbali-Zamora and J. C. Berg, "Development of a wind turbine generator volt-var curve control for voltage regulation using power hardware-in-the-loop," in *2023 IEEE PES Innovative Smart Grid Technologies Latin America (ISGT-LA)*, pp. 280–284, 2023.

[9] L. Aththanayake, N. Hosseinzadeh, A. Mahmud, A. Gargoom, and E. M. Farahani, "Challenges to voltage and frequency stability of microgrids under renewable integration," in *2020 Australasian Universities Power Engineering Conference (AUPEC)*, pp. 1–6, 2020.

[10] Y. Pei, Y. Yao, J. Zhao, F. Ding, and J. Wang, "Multi-Agent Deep Reinforcement Learning for Realistic Distribution System Voltage Control using PV Inverters," in *2022 IEEE Power Energy Society General Meeting (PESGM)*, pp. 1–5, 2022.

[11] M. Jimenez-Aparicio and R. Darbali-Zamora, "Cooperative Multi-Agent Reinforcement Learning Method for Fast Voltage Regulation using Distributed Wind Turbine Generators," in *2023 IEEE PES Innovative Smart Grid Technologies Latin America (ISGT-LA)*, pp. 1–5, 2023.

[12] B. Singh, B. K. Panigrahi, and G. Pathak, "Control of wind-solar microgrid for rural electrification," in *2016 IEEE 7th Power India International Conference (PIICON)*, pp. 1–5, 2016.

[13] M. Jimenez-Aparicio and R. Darbali-Zamora, "Reactive Power Control for Fast-Acting Voltage Regulation of Distributed Wind Turbines Using Reinforcement Learning," in *2023 IEEE Kansas Power and Energy Conference (KPEC)*, pp. 1–6, 2023.

[14] Y. Pei, J. Zhao, Y. Yao, and F. Ding, "Multi-Task Reinforcement Learning for Distribution System Voltage Control With Topology Changes," *IEEE Transactions on Smart Grid*, vol. 14, no. 3, pp. 2481–2484, 2023.

[15] M. J. Aparicio and S. Grijalva, "Economic Assessment of V2B and V2G for an Office Building," in *2020 52nd North American Power Symposium (NAPS)*, pp. 1–6, 2021.

[16] Z. Huang and G. Liu, "Capacity and Power Optimization of Energy Storage System in Microgrid with Wind and Solar Power Station," in *2023 5th International Academic Exchange Conference on Science and Technology Innovation (IAECST)*, pp. 1372–1376, 2023.

[17] V. Khadkikar, R. K. Varma, and R. Seethapathy, "Grid voltage regulation utilizing storage batteries in PV solar — Wind plant based distributed generation system," in *2009 IEEE Electrical Power Energy Conference (EPEC)*, pp. 1–6, 2009.

[18] M. Gu and B. Zhao, "Multi-Objective Optimum Scheduling of Microgrids Accounting for Wind and Solar Scenario Uncertainty," in *2024 11th International Forum on Electrical Engineering and Automation (IFEEA)*, pp. 586–590, 2024.

[19] H. Su, D. Yang, W. Song, J. Tuo, and Y. Liu, "Robust Optimal Scheduling of "Wind Storage" Multi-Energy Complementary Integrated Energy Microgrid Considering wind Power Consumption," in *2024 4th International Conference on Energy Engineering and Power Systems (EEPS)*, pp. 533–536, 2024.

[20] G. E. North Piegan, R. Darbali-Zamora, and J. C. Berg, "Development and Validation of a Wind Turbine Generator Simulation Model," in *2022 North American Power Symposium (NAPS)*, pp. 1–6, 2022.

[21] R. Darbali-Zamora, "Development of a dynamic photovoltaic inverter model with grid-support capabilities for power system integration analysis," in *2023 IEEE 50th Photovoltaic Specialists Conference (PVSC)*, pp. 1–8, 2023.

[22] R. Darbali-Zamora and E. I. Ortiz-Rivera, "Implementation of an analytical active power curtailment method for photovoltaic system applications," in *2024 IEEE 52nd Photovoltaic Specialist Conference (PVSC)*, pp. 0742–0749, 2024.

[23] M. Jimenez-Aparicio and R. Darbali-Zamora, "Increasing the Penetration of Distributed Wind in an Isolated Microgrid Through Data-driven Frequency Regulation using Reinforcement Learning," in *2024 IEEE Kansas Power and Energy Conference (KPEC)*, pp. 1–6, 2024.


Authorized licensed use limited to: Sandia National Laboratories. Downloaded on January 06,2026 at 07:20:32 UTC from IEEE Xplore. Restrictions apply.