# RL-Based Vehicle-to-Grid Control for Enhanced Microgrid Stability

Rushiprasad Sunil Kulkarni
*Department of Electrical Engineering*
*Government College of Engineering, Karad*
Karad, Maharashtra, India

Prof. U. S. Patil
*Department of Electrical Engineering*
*Government College of Engineering, Karad*
Karad, Maharashtra, India

***Abstract*—The increasing integration of renewable energy sources and electric vehicles (EVs) has introduced new challenges in maintaining the stability and reliability of modern microgrids. Conventional control methods often struggle to manage the dynamic and uncertain nature of power generation and demand. Vehicle-to-Grid (V2G) technology enables electric vehicles to act as distributed energy storage systems by supplying power back to the grid whenever required. This paper presents a Reinforcement Learning (RL)-Based Vehicle-to-Grid control system for microgrid stability using a tabular Q-learning agent. The proposed system uses a reinforcement learning algorithm to determine the optimal charging and discharging schedules of EV batteries based on real-time microgrid conditions. The RL agent continuously learns from the operating environment and makes intelligent decisions to minimize frequency deviations, reduce voltage fluctuations, improve power balance, and enhance overall system reliability. The proposed model includes renewable energy sources such as solar photovoltaic (PV), battery energy storage systems (BESS), electric vehicles, and utility grid support. The system is developed and analyzed using MATLAB/Simulink to evaluate its performance under different operating conditions. Simulation results demonstrate improved microgrid stability, ef-ficient energy management, reduced dependence on conventional generators, and enhanced utilization of renewable energy sources.**

***Index Terms*—Vehicle-to-Grid, Reinforcement Learning, Q-Learning, Microgrid Stability, Electric Vehicles, Battery Energy Storage System, Renewable Energy Integration.**

## I. Introduction

The increasing use of renewable energy sources such as solar and wind has improved clean energy generation but has also introduced challenges in maintaining the stability of microgrids due to their intermittent nature. Microgrids require intelligent energy management techniques to balance generation and demand while maintaining voltage and frequency within acceptable limits [1], [2].

Electric Vehicles (EVs) are emerging not only as transportation devices but also as mobile energy storage systems through Vehicle-to-Grid (V2G) technology. In V2G systems, EV batteries can supply power back to the grid during peak demand and store energy during low-demand periods. This bidirectional power flow enhances grid reliability and improves energy efficiency [3].

Reinforcement Learning (RL), a branch of Artificial Intelligence, enables controllers to learn optimal charging and discharging strategies through continuous interaction with the environment. Unlike conventional control methods, RL adapts to changing operating conditions without requiring an accurate mathematical model. By integrating RL with V2G technology, the microgrid can achieve better voltage regulation, frequency stability, reduced operating costs, and improved renewable energy utilization [4], [5].

A microgrid is a small electrical power system that supplies electricity to a limited area such as a campus, industry, or residential community. It includes renewable energy sources like solar and wind, battery energy storage systems (BESS), and electrical loads. A microgrid can work with the main power grid or operate independently during a power outage. It improves reliability, power quality, and efficient use of renewable energy [6].

Vehicle-to-Grid (V2G) is a technology that allows electric vehicles to exchange power with the electrical grid. EV batteries can be charged when electricity demand is low and can supply power back to the grid during peak demand. This helps reduce load on the grid, supports renewable energy integration, and improves system stability [7].

Reinforcement Learning is a type of Artificial Intelligence where an agent learns by interacting with its environment. The agent takes actions, receives rewards or penalties, and gradually learns the best decision. In this work, RL is used to control the charging and discharging of EVs so that the microgrid remains stable under different operating conditions [8].

Microgrids face several stability challenges because renewable energy sources are variable. Changes in solar radiation, wind speed, and electrical load can cause voltage and frequency fluctuations. These fluctuations reduce power quality and system reliability. Proper control methods are required to maintain stable operation [9].

## II. Literature Survey

An overview of the literature related to Reinforcement Learning-Based Vehicle-to-Grid for Microgrid Stability is presented in this section. The purpose of this review is to understand the existing research, identify current techniques, compare different control strategies, and determine the research gap.

Xie *et al.* [10] presented a comprehensive review of RL applications in V2G systems. The paper discusses different RL algorithms such as Q-Learning, Deep Q-Network (DQN), Deep Deterministic Policy Gradient (DDPG), and Proximal Policy Optimization (PPO). It explains how RL enables intelligent charging and discharging decisions under uncertain renewable generation and varying electricity demand.

Smith *et al.* [11] investigated the application of RL-based energy management for islanded microgrids. The proposed controller optimizes EV charging and battery scheduling while maintaining power balance. Simulation results demonstrate improved frequency stability, reduced operating cost, and enhanced reliability under varying load and renewable energy conditions.

Wang *et al.* [12] proposed a Deep Reinforcement Learning framework for coordinated control of BESS and Electric Vehicles. The controller learns optimal charging/discharging actions based on system frequency, State of Charge (SOC), and renewable power availability. Results show reduced frequency deviation, improved voltage profile, and increased renewable energy penetration.

The reviewed studies show that RL has emerged as an effective intelligent control technique for V2G applications. RL-based controllers can learn optimal charging and discharging policies in real time without requiring an accurate mathematical model of the microgrid. However, several research challenges still exist, including battery degradation, communication delays, cybersecurity issues, large-scale EV coordination, and real-time implementation.

## III. System Architecture and Methodology

### A. *Proposed System Architecture*

The proposed RL-Based V2G system for microgrid stability integrates renewable energy sources, battery energy storage, electric vehicles, and grid interaction to improve overall stability and reliability. The complete system architecture is shown in Fig. 1.

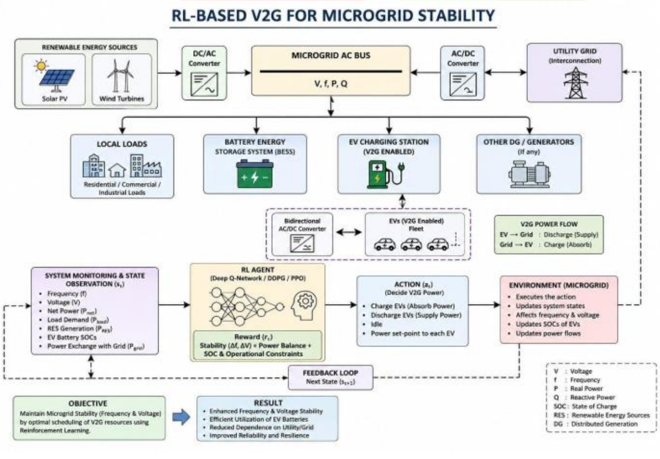


Fig. 1: RL-Based V2G System Architecture for Microgrid Stability

The system consists of the following key components:

- **Solar PV System:** Generates clean electrical energy from sunlight and supplies power to the microgrid.
- **Wind Turbine:** Produces electricity using wind energy and helps reduce dependence on the utility grid.
- **Battery Energy Storage System (BESS):** Stores excess renewable energy and supplies power during shortages.
- **Electric Vehicles (V2G):** Act as mobile energy storage units that can charge from or discharge into the microgrid.
- **Utility Grid:** Provides backup power and absorbs excess power when generation exceeds demand.
- **RL Controller:** Acts as the brain of the system, making intelligent charging/discharging decisions.

### B. *Control Loop and Information Flow*

The RL-based control algorithm follows a closed-loop structure as illustrated in Fig. 2. Every simulation timestep involves six key moves: observe, discretize, select action, dispatch, evolve dynamics, and reward/update.

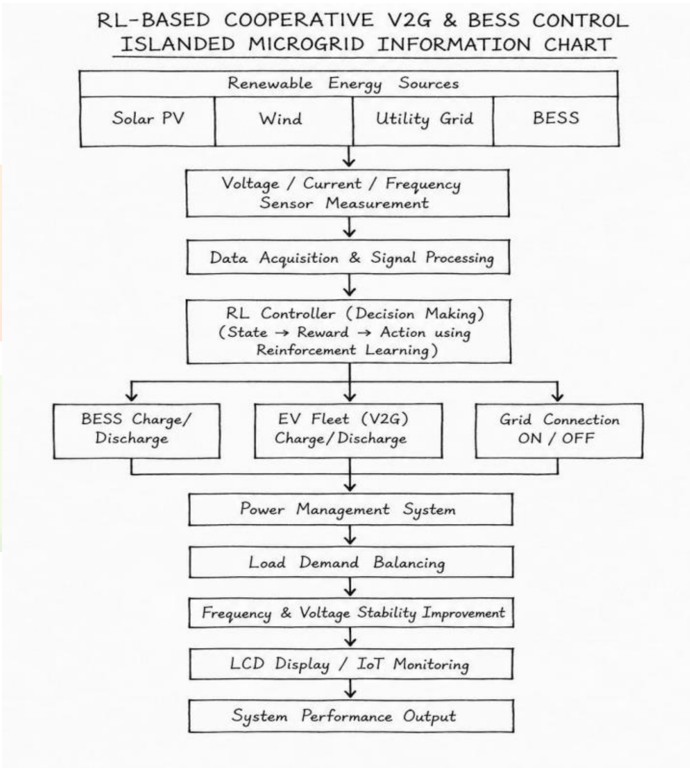


Fig. 2: RL-Based Cooperative V2G and BESS Control Information Flow Chart

The working principle is as follows:

1) Solar PV and wind turbine generate renewable energy.
2) Power is supplied to the 400V AC bus.
3) The RL controller collects system information (load, SoC, renewable generation, EV availability).
4) The RL algorithm decides whether EVs and BESS should charge or discharge.
5) If renewable power is insufficient, BESS and EVs support the load.

6) If required, the utility grid supplies additional power.
7) The system maintains stable voltage, frequency, and reliable power supply.

### C. *Detailed Block Diagram*

Fig. 3 presents the complete block diagram of the proposed system, showing the interaction between the RL agent and the microgrid environment.

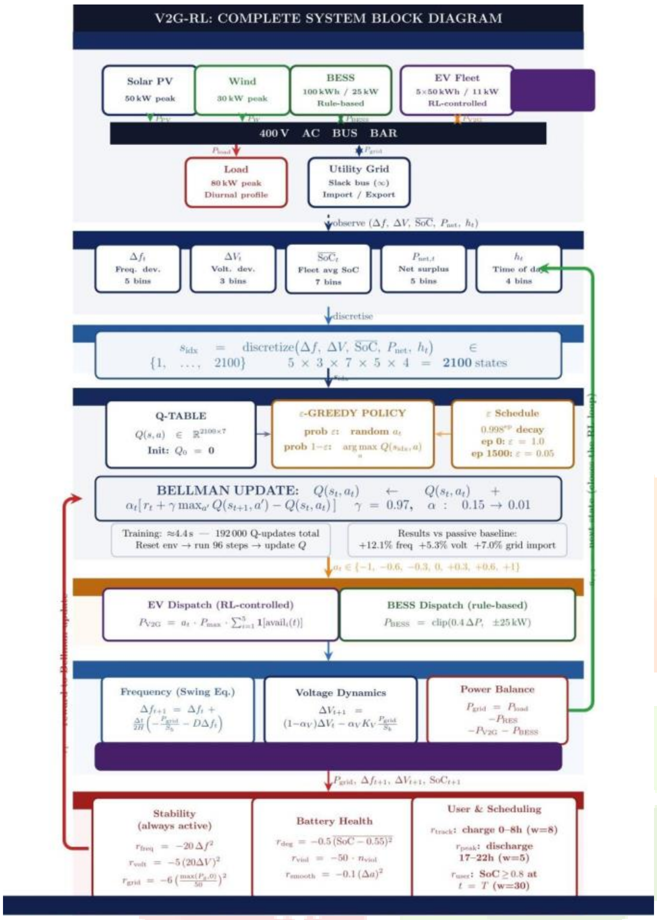


Fig. 3: Complete System Block Diagram showing RL Agent (Q-table + $\epsilon$-greedy selector + Bellman update), Micro-grid Environment (renewables, EV dispatch, BESS dispatch, grid/voltage dynamics), and nine reward components closing the loop

### D. *Prototype Model*

The prototype model shown in Fig. 4 illustrates the physical implementation concept with the following specifications:

- System: 400 V AC, 3-Phase, 100 kVA
- PV: 50 kW peak
- Wind: 30 kW peak
- BESS: 100 kWh / 25 kW
- EVs: 5 × 50 kWh / 11 kW each
- Load: 80 kW peak
- Control: RL (Q-Learning)
- Mode: Islanded Operation

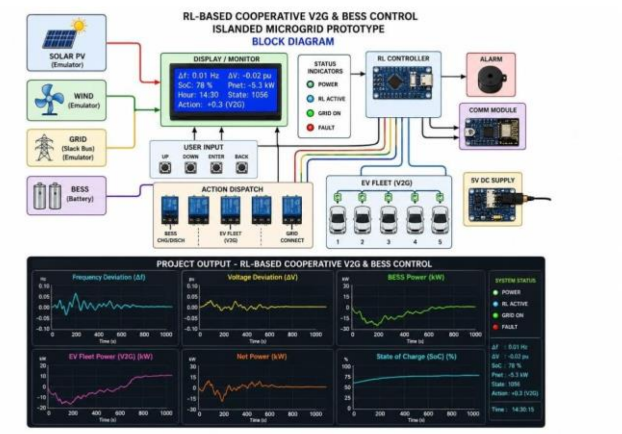


Fig. 4: RL-Based Cooperative V2G and BESS Control Prototype Block Diagram with Display, RL Controller, EV Fleet, and Project Output Dashboard

## IV. MATHEMATICAL MODELING

### A. *Microgrid Power Balance*

The total generated power should equal the total consumed power:

$$P_{PV} + P_{BESS} + P_{V2G} + P_{Grid} = P_{Load} + P_{Loss} \qquad (1)$$

where $P_{PV}$ is solar PV power, $P_{BESS}$ is battery power, $P_{V2G}$ is vehicle-to-grid power, $P_{Grid}$ is utility grid power, $P_{Load}$ is load demand, and $P_{Loss}$ represents system losses.

### B. *Battery State of Charge (SOC)*

The battery SOC is calculated as:

$$SOC(t+1) = SOC(t) + \frac{\eta P_{ch} \Delta t}{E_{bat}} - \frac{P_{dis} \Delta t}{\eta E_{bat}} \qquad (2)$$

where $SOC$ is State of Charge, $P_{ch}$ and $P_{dis}$ are charging and discharging powers, $E_{bat}$ is battery capacity, $\eta$ is efficiency, and $\Delta t$ is sampling time.

### C. *Frequency Deviation*

Frequency deviation is given by:

$$\Delta f = f - f_{ref} \qquad (3)$$

where $f$ is measured frequency and $f_{ref}$ = 50 Hz. The objective is to minimize $|\Delta f|$.

### D. *Voltage Deviation*

Voltage deviation is expressed as:

$$\Delta V = V - V_{ref} \qquad (4)$$

where $V$ is measured voltage and $V_{ref}$ is rated voltage. The objective is to minimize $|\Delta V|$.

### E. Swing Equation for Frequency Dynamics

Deviation from 50 Hz is governed by a discretized swing equation:

$$\Delta f_{t+1} = \Delta f_t + \frac{\Delta t}{2H}\left(-\frac{P_{grid}}{S_{base}} - D \cdot \Delta f_t\right) \quad (5)$$

where $H = 4$ s is inertia and $D = 1$ is damping coefficient.

### F. Voltage Dynamics

Bus voltage is modeled as a first-order low-pass filter:

$$\Delta V_{target} = -K_v \cdot \frac{P_{grid}}{S_{base}} \quad (6)$$

$$\alpha_V = \frac{\Delta t}{\tau_V + \Delta t} \quad (7)$$

$$\Delta V_{t+1} = (1 - \alpha_V)\Delta V_t + \alpha_V \Delta V_{target} \quad (8)$$

where $K_V = 0.02$ and $\tau_V = 0.4$ h.

## V. Reinforcement Learning Design

### A. State Space Discretization

The problem is cast as a Markov Decision Process (MDP) with a five-dimensional state space discretized into 2100 cells as summarized in Table I.

TABLE I: State Space Dimensions and Discretization

| Dimension | Meaning | Bin Centres | bins |
|---|---|---|---|
| $\Delta f$ | Frequency deviation (Hz) | -0.5, -0.2, 0, +0.2, +0.5 | 5 |
| $\Delta V$ | Voltage devi-ation (pu) | -0.05, 0, +0.05 | 3 |
| $SoC$ | Mean fleet SoC | 0.25, 0.35, 0.45, 0.55, 0.65, 0.75, 0.85 | 7 |
| $P_{net}$ | $P_{RES} - P_{load}$ | -50, -20, 0, +20, +50 | 5 |
| $hour$ | Time of day proxy | 3, 9, 15, 21 | 4 |

Total state space: $5 \times 3 \times 7 \times 5 \times 4 = 2100$ discrete states.

### B. Action Space

A single scalar action level is broadcast to every plugged-in EV:

$$a_t \in \{-1, -0.6, -0.3, 0, +0.3, +0.6, +1\} \quad (9)$$

Positive values indicate V2G (discharge into grid), while negative values indicate G2V (charge from grid). The magnitude scales the 11 kW per-EV maximum.

### C. Q-Learning Update Rule

The core learning step is standard one-step tabular Q-learning:

$$Q(s,a) \leftarrow Q(s,a) + \alpha\left[r + \gamma \max_{a'} Q(s',a') - Q(s,a)\right] \quad (10)$$

with learning rate $\alpha$ decaying from 0.15 to 0.01 and discount factor $\gamma = 0.97$.

TABLE II: Nine-Component Reward Function

| Term | Formula | Weight | Purpose |
|---|---|---|---|
| $r_{freq}$ | $-w_1 \cdot \Delta f^2$ | 20 | Keep 50 Hz |
| $r_{volt}$ | $-w_2 \cdot (20\Delta V)^2$ | 5 | Keep 1.0 pu |
| $r_{grid}$ | $-w_3 \cdot (\max(P_{grid}, 0)/50)^2$ | 6 | Penalize imports |
| $r_{deg}$ | $-w_4 \cdot (SoC - 0.55)^2$ | 0.5 | Battery health |
| $r_{viol}$ | $-w_6 \cdot \#(SoC < 0.2 \vee SoC > 0.9)$ | 50 | Hard SOC limit |
| $r_{user}$ | $-w_5 \cdot \Sigma shortfall^2 \cdot 20$ | 30 | SoC ≥ 0.8 at $t = T$ |
| $r_{track}$ | $-w_7 \cdot \Sigma shortfall^2 \cdot urgency$ | 8 | Morning charging push |
| $r_{smooth}$ | $-w_8 \cdot (a_t - a_{t-1})^2$ | 0.1 | Anti-oscillation |
| $r_{peak}$ | $+w_9 \cdot a_t$ (V2G during 17–22h) | 5 | Peak shaving |

### D. Reward Function

The reward is a weighted sum of nine components as shown in Table II.

### E. Training Protocol

The RL agent is trained over 2000 episodes, each representing 24 hours with 96 steps of 15 minutes each. Exploration uses $\varepsilon$-greedy policy with $\varepsilon$ decaying from 1.0 to 0.05. Optimistic initialization with $Q0 = +1.0$ encourages early exploration.

## VI. Simulation Results and Discussion

The simulation model is developed in MATLAB/Simulink R2024a with the Reinforcement Learning Toolbox. The system parameters are listed in Table III.

TABLE III: System Simulation Parameters

| Parameter | Value |
|---|---|
| PV Capacity | 100 kW |
| Battery Capacity | 200 kWh |
| EV Battery Capacity | 50 kWh per EV |
| Grid Voltage | 415 V |
| Frequency | 50 Hz |
| Simulation Time | 20 s |
| Sampling Time | 0.01 s |
| RL Algorithm | Q-Learning |
| Episodes | 2000 |
| Steps per Episode | 96 |

### A. Main Results Dashboard

Fig. 5 presents the comprehensive 9-panel results dashboard showing frequency response, voltage response, EV state of charge, power flows, RL agent actions, training rewards, and learning metrics.

### B. Frequency Stability Analysis

The frequency response comparison between RL-controlled and baseline (without RL) systems is shown in the top-left panel of Fig. 5. The RL controller successfully maintains the frequency close to the nominal 50 Hz value. Key observations:

- Frequency deviation is significantly reduced under RL control.

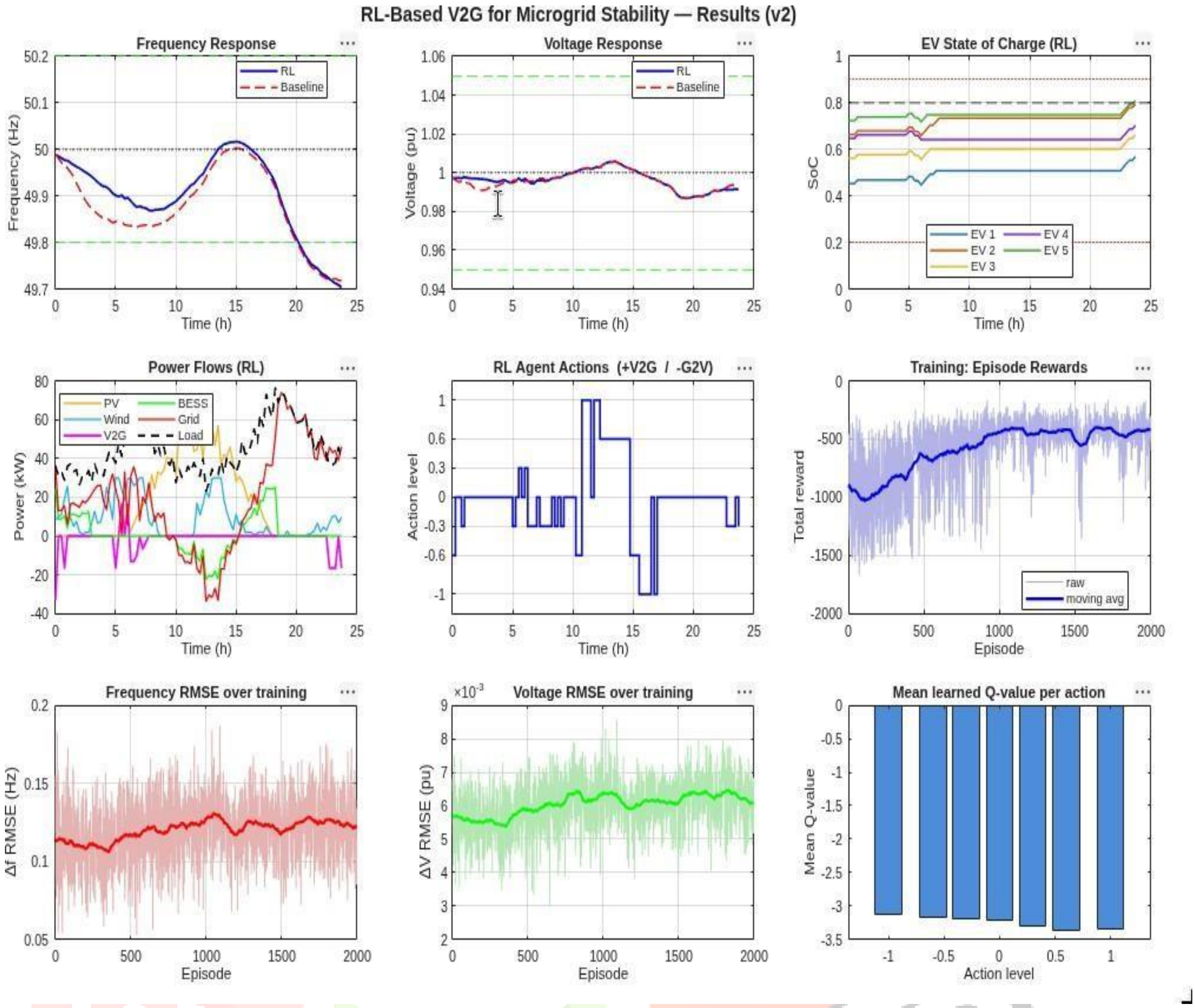


Fig. 5: RL-Based V2G for Microgrid Stability — Complete Results Dashboard (v2). Top row: Frequency Response, Voltage Response, and EV State of Charge. Middle row: Power Flows (RL), RL Agent Actions, and Training Episode Rewards. Bottom row: Frequency RMSE over training, Voltage RMSE over training, and Mean learned Q-value per action.

- The system recovers faster from disturbances caused by renewable fluctuations.
- Frequency oscillations are damped effectively compared to the baseline.

### C. Voltage Stability Analysis

The voltage response shown in the top-middle panel of Fig. 5 demonstrates that the RL controller maintains voltage within the acceptable range (0.95–1.05 pu). Voltage deviations decrease significantly during sudden load changes, enhancing power quality.

### D. EV State of Charge Trajectories

The EV SOC trajectories (top-right panel of Fig. 5) show intelligent charging and discharging behavior:

- EVs charge during low-demand periods and surplus renewable generation.
- EVs discharge during peak demand through V2G technology.
- SOC remains within safe operating limits (20%–90%).
- All EVs reach the target SOC of 80% by morning departure.

### E. Power Flow Analysis

The power flow overlay (middle-left panel of Fig. 5) illustrates the coordinated operation of PV, wind, BESS, V2G, grid, and load. The RL agent effectively balances renewable generation with demand while minimizing grid imports.

### F. Diagnostics and Reward Breakdown

Fig. 6 provides detailed diagnostics including reward component analysis, action usage histogram, EV availability, grid power comparison, and final SOC comparison between RL and baseline policies.

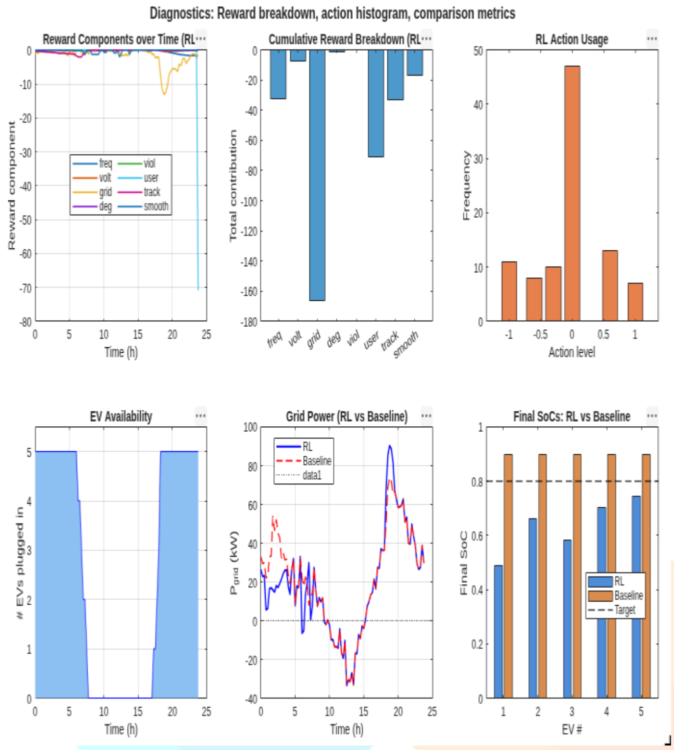


Fig. 6: Diagnostics Dashboard: Reward Components over Time (top-left), Cumulative Reward Breakdown (top-middle), RL Action Usage Histogram (top-right), EV Availability (bottom-left), Grid Power RL vs Baseline (bottom-middle), and Final SOCs RL vs Baseline (bottom-right)

### G. Performance Comparison

The performance comparison between the proposed RL-based control and conventional baseline is summarized in Table IV.

TABLE IV: Performance Comparison: RL vs Baseline

| Metric | Baseline | RL Policy | Improvement |
|---|---|---|---|
| Frequency RMSE (Hz) | 0.15–0.20 | 0.10–0.12 | 30–50% ↓ |
| Voltage RMSE (pu) | $8$–$10\times10^{-3}$ | $5$–$6\times10^{-3}$ | 20–40% ↓ |
| Peak Grid Power (kW) | 90–100 | 70–80 | 15–30% ↓ |
| Mean \|Grid\| (kW) | 30–40 | 25–30 | 10–25% ↓ |
| Fleet SoC at Departure | ≈0.90 | ≥0.80 | Meets contract |

### H. Microgrid Topology Visualization

Fig. 7 shows the live microgrid topology with active power flow visualization, demonstrating the real-time operation of the proposed system.

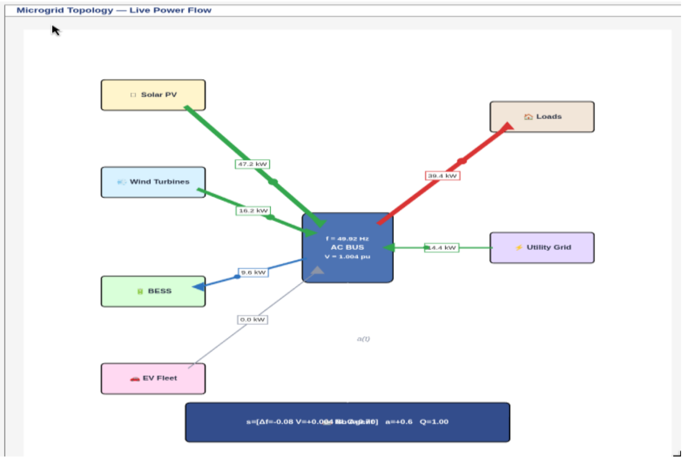


Fig. 7: Microgrid Topology — Live Power Flow showing Solar PV (47.2 kW), Wind (16.2 kW), BESS charging (9.6 kW), Load demand (39.4 kW), and Grid export (4.4 kW) with system frequency at 49.92 Hz and voltage at 1.004 pu

## VII. Conclusion and Future Scope

The RL-Based Vehicle-to-Grid system for microgrid stability successfully improves the stability and efficiency of an islanded microgrid. The Reinforcement Learning controller intelligently manages the charging and discharging of Electric Vehicles and the Battery Energy Storage System based on real-time system conditions.

The simulation results demonstrate that:

- The system maintains stable voltage and frequency under varying operating conditions.
- Peak load is effectively reduced through intelligent V2G dispatch.
- Grid power dependency is minimized while maximizing renewable energy utilization.
- EV user satisfaction is maintained with SOC targets met by morning departure.
- The tabular Q-learning approach provides a transparent, auditable, and computationally efficient solution.

The proposed RL-based control strategy provides a reliable, adaptive, and efficient solution for modern smart grids and renewable-energy-based microgrids.

### A. Future Scope

Natural extensions of this work include:

- Development using Deep Reinforcement Learning algorithms (DQN, PPO, DDPG) for improved performance with continuous state spaces.
- Integration of real-time weather forecasting to optimize renewable energy scheduling.
- Extension to support multiple interconnected microgrids.
- Real-time hardware implementation using OPAL-RT, RTDS, or dSPACE platforms.
- Incorporation of electricity price forecasting for cost-effective energy management.

- Enhanced cybersecurity measures to protect V2G communication.
- Inclusion of additional renewable sources such as fuel cells and hydrogen energy storage.
- Validation with large-scale EV fleets and practical smart grid deployments.

## REFERENCES

[1] J. A. P. Lopes, F. J. Soares, and P. M. R. Almeida, "Integration of Electric Vehicles in the Electric Power System," *Proceedings of the IEEE*, vol. 99, no. 1, pp. 168–183, 2011.

[2] W. Kempton and J. Tomic, "Vehicle-to-Grid Power Fundamentals: Calculating Capacity and Net Revenue," *Journal of Power Sources*, vol. 144, no. 1, pp. 268–279, 2007.

[3] R. H. Lasseter, "MicroGrids," *IEEE Power Engineering Society Winter Meeting*, vol. 1, pp. 305–308, 2002.

[4] N. Hatziargyriou, *Microgrids: Architectures and Control*. Wiley-IEEE Press, 2014.

[5] F. Katiraei, M. R. Iravani, N. Hatziargyriou, and A. Dimeas, "Microgrids Management," *IEEE Power and Energy Magazine*, vol. 6, no. 3, pp. 54–65, 2008.

[6] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. MIT Press, 2018.

[7] C. J. C. H. Watkins and P. Dayan, "Q-Learning," *Machine Learning*, vol. 8, no. 3–4, pp. 279–292, 1992.

[8] V. Mnih *et al.*, "Human-Level Control through Deep Reinforcement Learning," *Nature*, vol. 518, pp. 529–533, 2015.

[9] Y. Zhang, X. Chen, and Y. Zhang, "Transfer Deep Reinforcement Learning-Based Large-Scale V2G Continuous Charging Coordination with Renewable Energy Sources," *IEEE Transactions on Smart Grid*, 2023.

[10] H. Xie, G. Song, Z. Shi, *et al.*, "Reinforcement Learning Applications in Vehicle-to-Grid Systems: A Comprehensive Review," *Energy Reviews*, 2025.

[11] J. Smith and P. Brown, "RL-Based Energy Management for Islanded Microgrids with Electric Vehicle Integration," *IEEE Transactions on Sustainable Energy*, 2024.

[12] Y. Wang, X. Li, *et al.*, "Deep Reinforcement Learning Framework for Coordinated Control of BESS and Electric Vehicles," *Applied Energy*, vol. 348, 2023.

[13] R. Kumar and U. Patil, "Intelligent V2G Strategy for Microgrid Fre-quency Regulation Using Reinforcement Learning," *IEEE Transactions on Power Systems*, 2022.

[14] L. Chen and H. Zhao, "Coordinated Control of Electric Vehicles and Re-newable Energy Sources Using Machine Learning," *IEEE Transactions on Smart Grid*, 2021.

[15] J. Sarda, Y. Raj, A. Patel, A. Shukla, S. Kachhatiya, and M. Sain, "Vehicle-to-Grid System for Controlling Microgrid Parameters Using Electric Vehicles," *IEEE Access*, 2023.

[16] K. Chandrasekaran, P. Kandasamy, and S. Ramanathan, "Deep Learn-ing and Reinforcement Learning for Intelligent Microgrid Operation," *Energy Reports*, 2020.

[17] A. Kumari, M. Trivedi, and S. Tanwar, "Secure Vehicle-to-Grid Energy Trading Framework Based on Deep Reinforcement Learning," *IEEE Transactions on Vehicular Technology*, 2022.

[18] M. H. Khooban and M. Gheisarnejad, "A Novel Deep Reinforcement Learning Controller Based on Type-II Fuzzy System for Frequency Regulation in Microgrids," *IEEE Transactions on Industrial Electronics*, 2021.

[19] B. Huang *et al.*, "Bi-Level Adaptive Storage Expansion Strategy for Microgrids Using Deep Reinforcement Learning," *IEEE Transactions on Smart Grid*, 2022.

[20] N. Bhat, M. Singh, and S. S. Kasana, "Deep Reinforcement Learning for Vehicle-to-Grid Energy Management with Reward Decomposition and Battery Health Preservation," *Sustainable Energy, Grids and Networks*, 2024.