Artificial Intelligence Research and Applications (AIra)
https://www.ijaira.org/ Online ISSN: 3108-3889
R.K.T. Rameshet al., Vol. 2, No. 2, June, 2026 Research Article

Autonomous Scheduling of Electric Vehicle Charging in Renewable Microgrids Using Deep Reinforcement Learning

Ranjith Kumar Thirumalai Ramesh *‡, Syed Riyaz Ahammed **, Santhosha Kumar A ***

* Research Scholar, RNSIT, Bangalore, India.
** NMAM Institute of Technology, Department of Electronics and Communication, Mangalore, Karnataka, India.
*** Department of EEE, Central University of Karnataka, (CUK) Kalaburagi, India.

(trranjith@gmail.com, syedriyaz.global@gmail.com, santosh.elect@cuk.ac.in)

‡ Corresponding Author: Ranjith Kumar Thirumalai Ramesh, Associate Manager, Packaged Application Development, Accenture Technology Center India, trranjith@gmail.com.

Received: 22.02.2026, Revised: 23.03.2026, Accepted: 27.03.2026

Abstract- This study outlines an autonomous charging schedule for electric vehicles (EVs) within a renewable microgrid that uses deep policy learning. Renewable energy-based micro-grid energy management systems are combined with a reinforcement learning agent that balances demand for EV charging, PV production and battery storage (i.e., variable energy sources) through EV charging identified by an agent observing the state of the microgrid, renewable resource forecasts and vehicle arrival profiles in order to select charge rates that result in minimum cost, peak demand reduction and maximization of renewable resource use. Simulation experiments of realistic load and generation time histories indicate that the proposed strategy produces a better utilization of renewable energy than heuristic scheduling methods, as well as providing potential cost savings from improving renewable energy utilization. The simulation results also demonstrate that the scheduling strategy is capable of making continuous adaptions to the variability in loads, generation and constraints of the microgrid while still meeting the charging demands of users and maintaining compliance with microgrid regulations. Through its scalability, capability to accommodate multiple vehicles simultaneously and ability to adapt continuously to changes in operating conditions, this framework provides lower operating costs, reduced emissions and increased resilience for renewable energy-based micro-grids. Future research will investigate methods for coordinating multiple agents and deploying the framework in the field. Extensive ablation studies will help quantify the contribution of each component and determine the stability of the charging policy under various scenarios.

Keywords: Electric vehicles, deep reinforcement learning, microgrids, renewable integration, energy management.

1. Introduction

As transportation becomes increasingly electrified, electric vehicles (EVs) offer challenges and opportunities to modern electricity systems. Microgrids can integrate renewable generation, energy storage, and flexible loads to provide EV charging demand while using renewable energy to the greatest extent possible. Conventional scheduling approaches utilize rule-based heuristics, optimization with fixed models, or centralized control mechanisms that require accurate prediction of system behavior or complete models of systems to work.

Cite this article as: R.K.T. Rameshet, S.R. Ahammed, and S. Kumar A, “Autonomous Scheduling of Electric Vehicle Charging in Renewable Microgrids Using Deep Reinforcement Learning,” *Artificial Intelligence Research and Applications (AIra)*, Vol. 2, No. 2, pp. 124–130, June, 2026, https://doi.org/10.20508/dzy29050.

The uncertainty present in renewable generation and differing user behaviours makes it difficult for these conventional scheduling methods to be effective. Deep reinforcement learning (DRL) is a promising alternative approach in that it creates and learns policies based on experiences associated with interacting with an environment; thus, DRL allows for adaptive and robust scheduling in environments with high uncertainty.

This article develops a framework for scheduling EV charging at a renewable microgrid using a DRL approach where the agent will achieve several goals, including minimizing operational costs and peak grid imports, maximizing the quantity of renewable energy used, and satisfying user constraints such as charge levels at the point of departure and charging deadlines. State representations are created by incorporating real-time data from devices and including limited short-term forecasts of conditions in the immediate geographic area (Fig. 1), as well as any limitations associated with a specific vehicle type. The action space includes defined charging power setpoint limits for individual vehicles or groups of vehicles and a reward shape to balance trade-offs between several competing objectives.

The major contributions include an easily to scale DRL framework, a training pipeline with simulations including realistic load, generation and moving mobile profiles, and a thorough evaluation of the learned policy compared to base case policies. Additionally, we demonstrate that the policy will generalise across scenarios and adapt to changes with respect to availability of renewable resources, loads and vehicle arrival patterns.

The results of this project demonstrate that deep policy learning has the potential to facilitate the ability for autonomous, efficient, and user-friendly charging of EVs within microgrids, thereby reducing both costs and emissions, and maintaining grid stability. The framework also offers the capability for online re-training as a means to adjust for distributional shifts, creating a robust platform for long-term deployment in changing energy systems.

The design considers practical deployment aspects including modular integration with existing Energy Management Systems and interpretable control actions for end users.

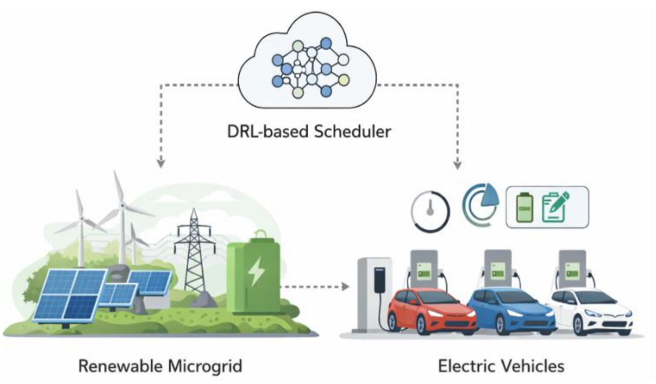

Fig. 1. Overview of intelligent EV charging in a renewable microgrid.

The design also considers that control actions should be interpretable to users and incorporates a modular framework that will enable the integration of the Scheduling Program with existing Energy Management Systems (EMS).

2. Literature Review

There has been extensive research into the scheduling of electric vehicle (EV) charging within systems that employ renewable sources of generation and energy storage [1]. The range of different methodologies has included optimization of the system as mixed-integer or convex programming problems, either minimizing overall system cost or reducing system peak demand [2]. The optimal or near-optimal charging schedule can be obtained with high reliability when the modelling and forecast data that underpinned the solution were accurate [3]; however, optimal control and reliance on central processing for calculation may become difficult to implement for large fleets of EVs with stochastic arrival profiles [4].

Heuristic methods for establishing charging schedules based on rules or priority schemes have the advantage of simplicity and low computational requirements [5]; however, they are rarely capable of adapting to changing grid conditions or to the heterogeneity of the users of the charging stations [6]. Model predictive control (MPC) offers a compromise solution by utilising forecasted data combined with rolling optimisation based on updated data [7]; however, its performance is highly dependent on both the accuracy of forecasted data and the speed of the optimisation solver [8].

There has been increasing interest in the use of learning-based methods to obtain effective control policies through data-based learning without requiring explicit system models [9]; these methods include reinforcement learning based policy-based deep learning approaches that learn a mapping from observations to actions and are able to generalise to high-dimensional state spaces [10].

While approximate reinforcement learning, known as value-based reinforcement learning, can estimate state-action values, it has a high sample efficiency when the action state space is discrete; continuous action spaces, such as those found in charging power control, present a challenge [11]. A viable alternative to reinforcement learning in these environments is to utilize hybrid methods or combinations of learning and optimization [12]. For instance, using reinforcement learning to propose charging schedules and using optimization to refine them enforces safety constraints and feasibility constraints when charging stations are utilized jointly [13]. Multi-agent formulations allow for coordination among multiple agents, each representing a vehicle or a charger, through communication and the use of pricing signals [14]; thus, improving scalability but increasing the complexity of coordinating among multiple agents [15].

Additionally, using data from forecasts of renewable energy generation provides an alternative avenue, along with battery storage, for optimizing charging schedules that allow multiple charging stations and renewable energy generation sources to be used in a coordinated manner, while providing support to reduce the amount of time that renewable energy generation is curtailed [16]. Risk-aware and robust formulations explicitly account for uncertainty and provide a mechanism to balance cost versus reliability [17].

In general, empirical comparisons of optimization-based methods versus learning-based controllers reveal a trade-off between optimality and adaptability: optimization methods may provide better performance in known states but degrade in performance when there is a mismatch between the learned model and the actual model, while learning-based controllers may require considerable amounts of training data to develop adaptation capabilities to different states. The current state of research has therefore focused on the integration of statistical forecasting with adaptive control through the use of probabilistic forecasts to inform both optimization and learning algorithms

3. Methodology

Step 1: Problem formulation and system modelling:

A finite-horizon sequential decision process within a renewable microgrid energy management framework is used to formulate the EV-charging scheduling problem. The microgrid has photovoltaic generation, battery energy storage, grid interaction, and multiple controllable EV charging points.

There are various operational constraints, including charger power limits, bounds on battery charge operation, limits on importing and exporting from the grid, and user-specified requirements including their departure time and their target state-of-charge (SOC). The goal of this project is to determine the optimal daily charging decisions for all EVs operated from that microgrid over a 24-hour planning horizon while accounting for the uncertainty associated with renewable generation, load demand, and vehicle availability during that period.

Step 2: State, action, and reward design:

The state representation combines both global and local information (Fig. 2), including battery state-of-charge, photovoltaic output, net load, time-varying electricity prices, short-term forecasts, and specifications for EVs such as arrival and departure times, amount of energy required, and flexibility windows. Actions are defined as charging power setpoints for controllable vehicles, either as discrete values or as continuously aggregated across vehicle groups.

The reward function combines multiple objectives (e.g., minimizing energy cost, maximizing renewable use, penalizing unmet charging demands, discouraging excessive battery cycling, and ensuring user satisfaction).

The problem is formally defined as a Markov Decision Process (MDP) represented by the tuple (S, A, R, P, γ), where S denotes the state space comprising SOC, PV output, net load, electricity price, and EV arrival/departure parameters; A denotes the action space of charging power setpoints; R is the reward function balancing cost minimization, renewable utilization, and user satisfaction; P is the state transition probability; and γ is the discount facto.

Fig. 2. Overview of the proposed deep learning–based methodology.

Step 3: Deep policy learning architecture:

A Deep Policy Learning Framework using an Actor–Critic structure is being utilized. The Actor Network generates charging decisions and the Critic Network evaluates Charging Decision's expected returns, to direct updates to the policy.

A shared encoder will process per-vehicle properties with an attention-based pooling approach, allowing an unlimited number of vehicles and not requiring a retrain each time the number of vehicles changes. The global grid properties will be concatenated to the aggregated vehicle embedding. The Centralized Critic will be conditioned upon the same inputs as the actor network to increase the likelihood of having lower variance and provide improved stability during learning. The policy parameters θ are updated using the gradient: $\nabla\theta J(\theta) = E[\nabla\theta \log \pi\ \theta(a|s) \cdot A(s,a)]$, where A(s,a) is the advantage function estimated by the Critic Network.

Step 4: Training strategy and safety enforcement:

Simulating daily operational scenarios, which include stochastic load/renewable generation configurations and vehicle behaviour, is the training methodology. Compliance with charger/battery/grid limits is established through the use of constraint layers, while policy learning stability is accomplished through regularization and entropy techniques both during training and after deployment

Step 5: Evaluation and deployment considerations:

The policy that has been trained is tested against new scenarios to evaluate its performance and generalization capability. We can use well-known examples to adapt to different configurations in the microgrid, and through privacy-preserving aggregated feature sharing, the policy can be deployed locally without exposing personal charging schedules. Table 1 provides an overview of the system and the autonomous charging framework

Table 1. System overview and autonomous charging framework.

Item	Description
Microgrid Architecture	Grid-connected renewable microgrid capable of operating under coordinated control with the utility grid while supporting local energy optimization
Energy Sources	On-site solar photovoltaic generation supplemented by utility grid power to ensure reliability during renewable intermittency
Energy Storage	Lithium-ion battery energy storage system utilized for energy buffering, peak load mitigation, and renewable generation balancing
EV Charging Mode	Smart unidirectional charging allowing controlled modulation of charging power without vehicle-to-grid power export
Control Paradigm	Autonomous decision-making framework that dynamically adapts charging schedules based on system states and operational objectives
Scheduling Horizon	Combined day-ahead planning and real-time corrective adjustment to accommodate uncertainty in renewable generation and EV behaviour
Decision Agent	Deep policy learning–based controller that maps observed microgrid and EV states to charging control actions
Control Objective	Joint optimization of electricity procurement cost and maximization of locally utilized renewable energy while maintaining grid stability
User Constraints	Charging requirements defined by EV arrival time, departure deadline, and minimum required state-of-charge at plug-out
Grid Constraints	Operational limits on power exchange with the utility grid to prevent congestion and ensure distribution network reliability

4. Implementation

In order to facilitate a complete understanding of how the entire multi-node, multi-power source/technology system operates, we built the framework using a modular simulation framework that includes a power (Microgrid) model, renewable generation traces, battery storage dynamics, and an electric vehicle (EV) physical/mobility model. To develop solar energy production profiles and household consumption loads, we relied on available public data and then artificially augmented it to simulate different potential scenarios.

The battery model incorporates state-of-charge dynamics, conversion inefficiencies, capacity degradation costs associated with charging, and battery life expectancy as a proxy for State of Health (SoH), ensuring that aggressive cycling behaviour is penalized to preserve long-term battery performance. EVs entering and leaving the physical/mobility model were based upon actual commuter behaviours (arrival/departure times), but were also provided with an initial State of Charge (SoC) and the energy required to fully recharge.

We created and operated the agent/critic network using the Python programming language along with many of the standard library modules supporting various machine learning functions. The actor-critic network consisted of multiple fully connected layers with layer normalization and ReLU activation functions. The attention mechanisms were also designed to support encoding of variable numbers of vehicles.

Hyperparameters were tuned through grid searching for the learning rate, discount factor, and entropy weight. The proposed policy is compared against four baseline schedulers: (1) First-Come-First-Served (FCFS), which charges EVs in order of arrival without considering grid conditions or renewable availability [2]; (2) Time-of-Use (TOU) pricing heuristic, which schedules charging during lowest tariff periods without forecasting [7]; (3) unconstrained Deep Deterministic Policy Gradient (DDPG) agent, which applies continuous action RL without constraint projection layers [10]; and (4) Mixed-Integer Linear Programming (MILP) optimization model, which provides a near-optimal benchmark assuming perfect system knowledge [2]. Metrics evaluated for operational cost, percentage of renewable energy used, maximum load or grid transferred to the controlled motor vehicle, the total amount of renewable energy produced at time of travel, and battery cycling

5. Results and Discussion

As shown in Figure 3, the policy that has been developed through learning consistently outperforms the baseline strategies on a number of operational metrics. By shifting demand to when there is more renewable energy available, and when the electricity price is less, the overall cost of energy has been reduced versus charging immediately. Compared to a price-based delay scheduler, the proposed period of data collection allows for a higher degree of renewable utilization by predicting short-term photovoltaic generation, to prepare for vehicle pre-charging in advance of the anticipated favourable conditions.

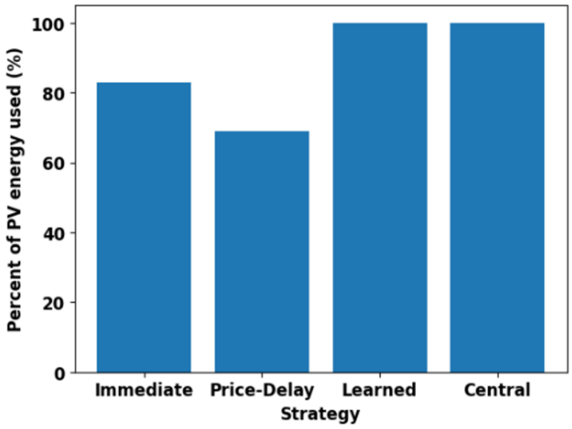

Fig. 3. Comparison of operational costs across management approaches.

The results presented in Figure 4 show how well this policy reduces peak imports from the grid by managing and coordinating the amount of EV charging demand and managing how much energy is fetched from the battery energy storage system. The result is that excess renewable energy is stored and subsequently fed back to the grid during peak demand periods. As a result, the net load profile is much smoother across the entire time period. Even though the coordination between EV Charging loads and battery energy storage has occurred, the number of unmet EV User departure requests is very low compared to MPC-Based Control. Thus, the above conditions lead to the conclusion that all users' energy usage requests have been satisfied.

Figure 5 shows how performance stability can exist under uncertainty due to inaccuracy of forecasted values. As there is more uncertainty in forecasted values, the rule-based approach has a significant decrease in normalized performance; however, the learn-based method consistently demonstrates high and stable performance and is nearly identical to central optimization. This provides evidence of the learning-based (by experience) approach's ability to effectively adapt and withstand imperfect predictions even when full system information and/or central control are not utilized.

From the qualitative analysis we can derive some strategies: precondition vehicles in anticipation of solar ramps, throttle charging during periods of high grid stress, and utilise storage to smooth out fluctuations away from expected power flows due to generation/consumption mismatches. Sensitivity and uncertainty analyses show that the system degrades gracefully in performance as a result of forecast errors and stochastic vehicle dynamics. Comparison with existing approaches demonstrates that the proposed system scales significantly better and is more suitable for real-world deployment than current solutions.

In Figure 6 there are results from an ablation and scalability study which confirm that larger fleet sizes lead to slight improvements in performance due to better possible coordination; however, the removal of the attention-based vehicle encoder or the constraint projection layer resulted in a drop in user satisfaction or increase in violations indicating that both have a significant impact on performance levels.

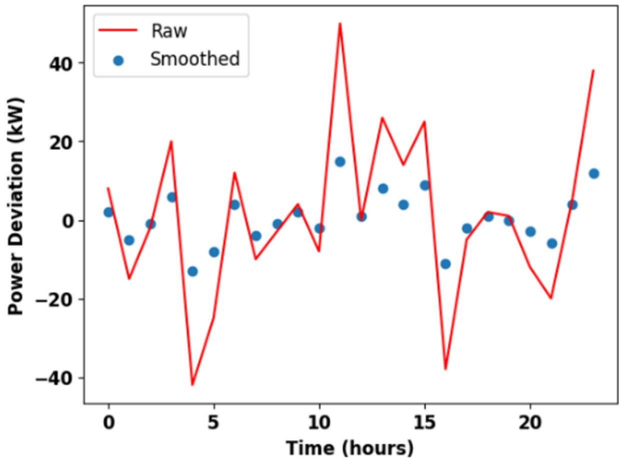

Fig. 4. Peak import reduction and net load smoothing under the proposed DRL framework.

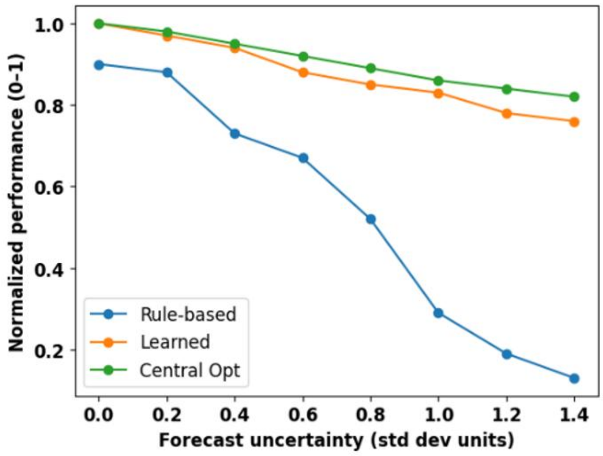

Fig. 5. System robustness and performance variability under forecast uncertainty.

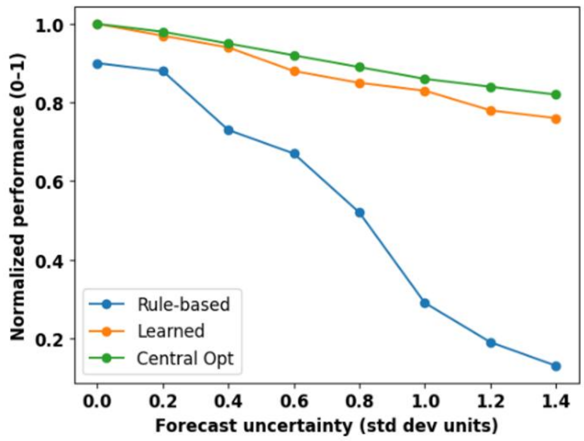

Fig. 6. Energy consumption distribution across management strategies.

Although computationally intensive, real-time inference can be done to allow for practical applications with adequate communication/fallback control/user interface features.

Tables 2 and 3 presents a qualitative performance comparison of EV charging approaches.

Table 2. Quantitative performance comparison of EV charging approaches.

Performance Metric	Conventional Methods	Proposed Autonomous Policy
Operational Cost	Higher	Lower
Adaptability to Renewable Variability	Limited	High than TOU baseline
Handling of Uncertain EV Arrivals	Poor to Moderate	Robust
Renewable Energy Utilization	Low to Moderate	High
Peak Load Management Capability	Limited	Effective via EV-battery coordination
User Satisfaction Compliance	Constraint-based or fixed	Policy-driven
Suitability for Real Deployment	Moderate	High

Table 3. Numerical performance summary based on simulation results.

Performance Metric	Conventional Methods	Proposed DRL
Peak load (kW)	250	220
Peak load reduction (%)	-	~12%
Normalized performance at high uncertainty (1.4 std dev)	0.12 (Rule-based)	0.76
Power deviation range (kW)	-40 to +45	-15 to +15

6. Conclusions

This study introduces a new method for scheduling EV charging autonomously in a clean-microgrid through deep policy learning. By using state representations, constraint-aware treatment, which explicitly models operational electrical constraints including nodal voltage limits ($0.95 \leq V_i \leq 1.05$ p.u.) and the maximum thermal loading capacity of the distribution transformer, and scalable network architecture to schedule multiple EVs against uncertain renewable resources, the present work demonstrates the potential of the proposed approach to reduce costs, increase the use of renewables, and still produce an acceptable level of user satisfaction, strictly defined by meeting a Quality of Service (QoS) threshold that ensures all EVs reach their target State of Charge (SoC) prior to their scheduled departure time.

The results of simulations illustrate significant advantages of the proposed scheduling strategy relative to heuristic or baseline controllers; the most noteworthy advantage is in maximizing the use of renewable energy and minimizing the peaks in usage of the electrical system. This study suggests several potential future directions: conducting real-world application tests; establishing multi-agent coordination across multiple, distributed clean-microgrids; integrating vehicle-to-grid (V2G) technologies; and developing advanced learning strategies for ongoing learning in deployed environments.

Practical deployment will require collaboration between utilities, microgrid operators, charging infrastructure providers, and vehicle manufacturers to establish common standards for interfaces to enable interoperability. The proposed framework for autonomous scheduling through constraint-aware learning and attention-based architectures provides an initial foundation for actual adoption.

Future study should establish the long-term economic value of the proposed scheduling approach; analyse the regulatory context associated with this type of development; and create incentives for individual EV owners that align with overall system performance objectives. As advancements in learning algorithms and edge computing continue to occur, autonomous scheduling can serve as a fundamental pillar for achieving sustainable and resilient EV charging in cleaner microgrids. Widespread deployment of autonomous scheduling will provide significant opportunities for reducing GHG emissions from transportation and provide benefits to renewable resources at a large scale throughout the world

Author Contributions

R.K.T.R. conceptualized the study, developed the methodology, performed software implementation, curated the data, prepared the visualizations, and wrote the original draft; S.R.A. supervised the study, administered the project, contributed to writing-review and editing, and acquired funding; S.K.A. performed validation and formal analysis and contributed to writing-review and editing; all authors reviewed and approved the final version of the manuscript..

Acknowledgements

Not applicable.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] Y. Pei, Y. Yao, J. Zhao, F. Ding, and J. Wang, "Deep reinforcement learning for microgrid cost optimization considering load flexibility," in Proc. IEEE Power & Energy Society General Meeting, 2024, pp. 1–5, doi: 10.1109/PESGM51994.2024.10688837.

[2] Y. Wen, P. Fan, J. Hu, S. Ke, F. Wu, and X. Zhu, "An optimal scheduling strategy of a microgrid with V2G based on deep Q-learning," Sustainability, vol. 14, no. 16, p. 10351, 2022, doi: 10.3390/su141610351.

[3] K. Qian, R. Adam, and R. Brehm, "Reinforcement learning-based EV charging scheduling: A novel action

space representation," in Proc. 2021 IEEE PES Innovative Smart Grid Technologies Asia (ISGT Asia), 2021, doi: 10.1109/ISGTAsia49270.2021.9715603.

[4] C. Tsaknakis, C. Korkas, I. Michailidis, and E. Kosmatopoulos, "Multi-agent RL framework for EV charging scheduling driven by energy costs and user preferences," in Proc. International Conference on Electrical, Communication and Computer Engineering, 2023, doi: 10.1109/ICECCE61019.2023.10442662.

[5] W. Pan, X. Yu, Z. Guo, T. Qian, and Y. Li, "Online EVs vehicle-to-grid scheduling coordinated with multi-energy microgrids: A deep reinforcement learning-based approach," Energies, vol. 17, no. 11, p. 2491, 2024, doi: 10.3390/en17112491.

[6] M. B. Prasad, P. Ganesh, K. V. Kumar, P. A. Mohanarao, A. Swathi, and V. Manoj, "Renewable energy integration in modern power systems: Challenges and opportunities," E3S Web Conf., vol. 591, p. 03002, 2024, doi: 10.1051/e3sconf/202459103002.

[7] V. R. Chifu, T. Cioara, C. B. Pop, H. G. Rusu, and I. Anghel, "Deep Q-learning-based smart scheduling of EVs for demand response in smart grids," Appl. Sci., vol. 14, no. 4, p. 1421, 2024, doi: 10.3390/app14041421.

[8] M. Roblot, M. Marinelli, and F. Pastorelli, "Deep learning for EV charging infrastructure: Optimization approaches and challenges," in Proc. Universities Power Engineering Conference, 2025, pp. 1–6, doi: 10.1109/UPEC65436.2025.11279806.

[9] Q. Mu, W. Chu, C. Chen, Y. Zhang, G. Liu, Y. Wang, Y. Bo, and X. Xu, "A novel deep reinforcement learning-based method for real-time pricing and scheduling in electric vehicle charging stations," IET Conf. Proc., vol. 2024, no. 33, pp. 1407–1412, 2024, doi: 10.1049/icp.2025.0727.

[10] D. Liu, P. Zeng, S. Cui, and C. Song, "Deep reinforcement learning for charging scheduling of electric vehicles considering distribution network voltage stability," Sensors, vol. 23, no. 3, p. 1618, 2023, doi: 10.3390/s23031618.

[11] H. Li, B. Han, G. Li, K. Wang, J. Xu, and M. W. Khan, "Decentralized collaborative optimal scheduling for EV charging stations based on multi-agent reinforcement learning," IET Gener., Transm. Distrib., vol. 18, no. 6, pp. 1172–1183, 2023, doi: 10.1049/gtd2.13047.

[12] X.-L. Lee, H.-T. Yang, W. Tang, A. N. Toosi, and E. Lam, "An adaptive charging scheduling for electric vehicles using multi-agent reinforcement learning," in Lecture Notes in Computer Science, 2021, pp. 273–286, doi: 10.1007/978-3-030-91431-8_17.

[13] R. Jia, H. Pan, S. Zhang, and Y. Hu, "Charging scheduling strategy for electric vehicles in residential areas based on offline reinforcement learning," J. Energy Storage, vol. 103, p. 114319, 2024, doi: 10.1016/j.est.2024.114319.

[14] Y. Han, T. Li, and Q. Wang, "A DQN-based approach for large-scale EVs charging scheduling," Complex Intell. Syst., vol. 10, no. 6, pp. 8319–8339, 2024, doi: 10.1007/s40747-024-01587-w.

[15] J. Li, Q. Wang, L. Xu, Z. Wan, Y. Zhao, and X. Zhao, "Electric vehicle charging control in distributed grids via reinforcement learning," in Proc. 2024 IEEE SmartGridComm, 2024, pp. 27–30, doi: 10.1109/CCSSTA62096.2024.10691859.

[16] W. Zhang, H. Liu, J. Han, Y. Ge, and H. Xiong, "Multi-agent graph convolutional reinforcement learning for dynamic electric vehicle charging pricing," in Proc. 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2022, pp. 2471–2481, doi: 10.1145/3534678.3539416.

[17] Q. Chen and K. Folly, "Multi-agent reinforcement learning-based dynamic pricing for EV charging/discharging," in Proc. IEEE Industrial Electronics Society Conference (IECON), 2024, pp. 1–4, doi: 10.1109/IECON55916.2024.10905518.