SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Deep Reinforcement Learning for Real-Time Energy Management and Optimal Power Allocation in Distributed Electrical Systems

Deepak Kumar[1], Aditya[2], Rajeev Ranjan[3], Priyanka Kumari[4]

[1]Assistant Professor, Department of Electrical Engineering, Government Engineering College, Arwal, 804409, Bihar, India.
E-mail: deepak.raj01993@gmail.com ORCID: 0009-0009-6263-9778
[2]Assistant Professor, Department of Electrical Engineering, Government Engineering College, Arwal, 804409, Bihar, India.
E-mail: adionestar@gmail.com ORCID: 0009-0000-7209-3494
[3]Assistant Professor, Department of Electrical Engineering, Government Engineering College, Arwal, 804409, Bihar, India.
E-mail: rajeev901243@gmail.com ORCID: 0009-0008-1091-4208
[4]Assistant Professor, Department of Electrical Engineering, Government Engineering College, Arwal, 804409, Bihar, India.
E-mail: priyanka92790@gmail.com ORCID: 0009-0000-7222-786X

ABSTRACT

The increasing penetration of renewable energy resources and distributed generation has made real-time energy management in microgrids more complex due to fluctuating demand, intermittent generation, dynamic electricity prices, and energy-storage constraints. The proposed approach considers the operation of a microgrid as a Markov Decision Process with states: electrical demand, renewable generation, electricity price, energy-storage state of charge, and temporal operating condition. Continuous control actions for battery charging and discharging, dispatchable generation and grid power exchange are generated using a Soft Actor-Critic algorithm. The goal of the control is to reduce operating cost, renewable energy curtailment, power imbalance, network losses, and constraint violations, while ensuring reliable system operation. Highly variable operating conditions are accounted for, such as varying PV and wind generation, variable load demand, changing electricity prices, periods of renewable surplus and varying storage bounds. The proposed framework is tested on an independent training and out-of-sample testing set, to evaluate the ability of the framework to generalize in unseen operating conditions. Conditions are evaluated based on economic, technical, renewable, storage, and computational metrics such as operating cost, grid reliance, renewable energy usage, power losses, power state-of-charge, constraint violations, and decision time. The framework serves as a baseline to compare adaptive DRL-based energy management with traditional and other reinforcement-learning approaches and serves as a replicable foundation for assessing the approach.

Keywords: Deep Reinforcement Learning; Soft Actor-Critic; Energy Management; Optimal Power Allocation; Distributed Electrical Systems

Introduction

The shift to low-carbon and decentralized power systems has ramped up the number and variety of distributed energy resources (DER) such as photovoltaics, wind turbines, battery energy storage systems, controllable loads, and grid-interactive microgrids. While these technologies improve sustainability and increase energy autonomy in the region, they create significant uncertainties in operation, as energy generation and consumption, energy storage and market prices keep changing over time. If the operating conditions are rapidly changing, then conventional rule-based or deterministic optimization techniques can become complex and inflexible. Deep reinforcement learning has thus been identified as a viable alternative to learning adaptive control policies directly from system interactions and to assist in sequential decision making in complex microgrid environments (Alabdullah & Abido, 2022; Ji et al., 2021; Nakabi & Toivanen, 2021).

The recent research shows that reinforcement learning can be used to enhance the microgrid scheduling by incorporating economic considerations, technical limitations, and system uncertainty. Reinforcement learning has been used in the context of energy management for microgrids (Alabdullah & Abido, 2022; Cai et al., 2023), and deep Q-networks have been investigated for microgrid management. Further decentralized reinforcement learning has also demonstrated its viability in managing stochastic microgrid environment for decision making in multiple resources based microgrid environment instead of a single centralized controller (Darshi et al., 2023). In general, deep reinforcement learning is emerging as a potential method to operate distribution systems due to its ability to model nonlinear decision policies and ability to act quickly after training (Glover et al., 2025).

The real-time operation is also a particular challenge as microgrid controllers need to manage a balance between supply and demand, the state of charge of the storage systems, the volatility of electricity prices, the minimisation of operational costs, and the integration of intermittent renewable energy. As shown in previous studies, deep reinforcement learning has been successfully used to assist the real-time optimal energy management under uncertainty and to decrease the need for explicit forecasts or repeated optimization (Guo et al., 2022; Liu et al., 2023). This has been further extended to renewable-energy integration and local energy trading, where coordinated agents learn how to trade energy with one another, whilst working towards system-level goals (Harrold et al., 2022). Therefore, the choice of the representation of the state space and reward structure is crucial, as if they are not designed correctly, the reinforcement learning algorithm may produce unstable or economically sub-optimal control policies (Jayaraj et al., 2024).

Scalable, adaptive and community-oriented applications have also been highlighted recently. Artificial-intelligence (AI) based microgrid platforms have enabled the participation of DRL in the domain of sustainable resource coordination, and community energy-management systems have shown that learning coordinated control strategies between interconnected users and distributed resources (Lami et al., 2025; Li et al., 2024) is possible. In addition, meta-reinforcement learning has been suggested for real-time scheduling, allowing agents to adjust to varying operating conditions more efficiently (Shen et al., 2024). Although progress has been made, there is a critical research gap in how these important research challenges—real-time energy management, continuous power allocation, renewable variability, dynamic electricity prices, storage constraints, and out-of-sample operating conditions—can be addressed simultaneously in a unified framework for a distributed electrical-system.

The current work introduces a deep reinforcement learning scheme that deals with real-time energy management and optimum power allocation in distributed microgrid systems. The proposed framework models the operation of the microgrid as an MDP and uses a Soft Actor-Critic algorithm to obtain continuous control decisions for energy storage dispatch, controllable generation and grid power exchange. The approach is to adaptively respond to the fluctuation of the electrical demand, renewable-energy generation, electricity price, and storage operating conditions and ensure the system feasibility and power balance. Special focus is on enhancing economic performance, optimizing use of renewable energy, lowering unnecessary grid dependency and keeping storage and operational boundaries. The framework is also designed to assess performance when operating conditions are unseen, which will enable the robustness and generalization capability of the proposed energy-management strategy based on DRLs to be assessed.

The study aims to create a replicable framework through high-resolution observations and continuous-action learning that captures the realistic variability of micro-grids, while allowing for adequate control flexibility for future practical real-time energy-management applications.

Objectives

1. To create a real-time energy management framework using DRLs that incorporates load, PV generation, wind generation, electricity price, and energy storage state information at 5-minute intervals.
2. To develop optimal continuous power allocation solutions to optimize the operating cost, renewable curtailment, power imbalance and constraint violations while preserving storage limits and system constraints.
3. To assess the proposed solution in independent operational conditions and make comparisons with some other energy-management strategies regarding their economic, technical, renewable energy and computational aspects.

2. Methodology

2.1 Research Design

The present study is designed in a simulation-type quantitative research approach to design and test a deep reinforcement learning (DRL) algorithm for real-time energy management and optimal power allocation of a distributed electrical system. The approach proposed in this paper considers a microgrid as a sequential decision-making problem where the DRL agent is able to perceive the system state at every 5-minute time interval and decide proper power-dispatch interventions to minimize operating cost, while respecting technical constraints.

2.2 Dataset Description

This dataset is the publicly available "Prediction-Free Coordinated Dispatch of Microgrid: A Data-Driven Online Optimization Approach" (Huang, 2024). The dataset includes 5-minute interval information for PV generation, wind-turbine generation, electrical load demand, electricity prices, energy-storage state-of-charge limits, and microgrid configuration information. Historical PV, wind and load data are provided for 365 days, and separate 60-day test data sets are provided for out-of-sample testing. The data set is appropriate for the real-time energy-management modelling application, as there are 288 observations per day.

2.3 Data Preprocessing

The PV, wind, load, and electricity-price data sets are ordered chronologically and synced with the common time interval of 5 minutes. Observations that are missing are checked and duplicated observations and physically inconsistent observations are identified before analysis. For continuous variables, normalization is performed using parameters set only in the training set, otherwise there is risk of information leakage. Time-related features, like time of day, are also incorporated as this captures the daily energy-demand and renewable generation trends.

2.4 Distributed Electrical System Model

The distributed electrical system is configured as a microgrid as provided in the dataset and comprises energy storage, virtual energy storage, dispatchable generators, electrical loads, wind generation, PV generation and electrical system interactions. The power-balance condition is enforced in each time step, where renewable generation, dispatchable generation, grid power, energy storage output, demand, and network losses are represented by the variables.

2.5 Energy-Storage Modelling

A model of the energy-storage system is developed based on its power limits, efficiency, and state-of-charge constraints. The SOC transition can be written as: Limited by a minimum and maximum SOC. The time-varying upper and lower bounds of the SOCs given in the data set are included in the control environment to avoid infeasible storage operation.

2.6 Markov Decision Process Formulation

The problem of energy management is formulated in real-time as a Markov Decision Process (MDP) consisting of states, actions, rewards, and state transitions. The state vector contains load demand, PV generation, wind generation, electricity price, battery SOC, and time information. The action vector contains decisions that can be controlled, such as charging or discharging batteries, dispatching generators, importing or exporting from the grid. Once each action has been completed, the condition of the system is updated by the environment, and the next 5-min observation is made.

2.7 Reward Function

The reward function is intended to induce economically efficient and technically feasible operation. It reduces electricity-purchase cost, dispatchable-generation cost, battery-use cost, network losses, renewable-energy curtailment, and supply-demand imbalance while penalizing for SOC violations and generator and power constraints violations. The overall reward is given by the following: where represents weighting coefficients and represents constraint penalties.

2.8 Deep Reinforcement Learning Model

The power-allocation problem requires continuous control actions, hence the primary DRL algorithm used is Soft Actor-Critic (SAC). SAC has two types of networks: actor network to produce a choice for power dispatch and a critic network to assess power dispatch decision. The tuples (s, a, r, s') are saved in a replay buffer and drawn during training. The entropy component of SAC encourages exploration and stabilizes learning.

2.9 Model Training and Testing

The 365 days of data are split into two sets: a model development set and a validation and hyper-parameter tuning set. Each training episode is a day (288 consecutive 5-minute decisions). The data in the independent 60-day dataset is always kept separately and is only used after the model is trained. The separation in time enables the DRL controller's ability to generalize to new values of load, renewable generation, and electricity price.

2.10 Benchmark Comparison

The proposed energy-management scheme based on the SAC is compared with the conventional rule-based energy-management scheme and some of the selected DRL algorithms including DDPG and TD3. All methods tested under the same operating conditions for a comparison. The comparison lets them know if the proposed method will measurably improve from traditional and alternative methods of learning.

2.11 Performance Evaluation

The following performance metrics are used: total operating cost, grid-energy consumption, renewable-energy utilization, renewable curtailment, peak grid demand, battery SOC performance, power losses, supply-demand imbalance, constraint-violation rate and decision time. Increase or decrease in percentage from the benchmark methods is calculated as: The computational time for each control decision is also computed to evaluate the applicability of the proposed framework for real-time operation.

2.12 Robustness and Statistical Analysis

The trained controller is tested under various conditions of high load, low renewables, high renewables, electricity price spikes and rapid net load variations. The performance measures are based on the independent test days and not on a selected day. When appropriate, paired statistical analyses and effect size are employed to determine if the difference between the proposed model and the various benchmark approaches is meaningful and consistent.

2.13 Ethical Considerations

The study is based on a publicly available secondary dataset and does not include human participation, personal information, clinical records or personally identifiable information. So, direct informed consent and human-subject ethics approval do not apply. The dataset is used exclusively for academic research purposes and in compliance with its requirements for license and citation. All data manipulation, modelling assumptions, and parameter values, as well as experimental procedures, are transparently documented, thereby ensuring original data are not altered and enabling the reproduction of the results. Results will be presented honestly, in an unbiased manner, with no manipulation or suppression of findings.

3. Results

3.1 Dataset Characteristics

The chosen dataset, Prediction-Free Coordinated Dispatch of Microgrid: A Data-Driven Online Optimization Approach, was used to simulate the microgrid operating environment over a fine time scale to assess the impact of real-time energy management in distributed microgrids. The three data sets of historical load, PV and wind turbine data each had 365 days of data recorded at 5-minute intervals, totalling 105,120 data points. The independent testing period consisted of 60 days, and a total of 17,280 samples were taken for each of the principal variables. The descriptive statistics revealed significant fluctuations in both load and renewable generation, which meant the data set was representative of a dynamic operating environment for reinforcement learning energy management.

Table 1. Descriptive statistics of historical microgrid variables

Variable	Observations	Mean	SD	Minimum	Maximum
Load power (kW)	105,120	2,489.76	446.86	1,221.45	4,372.91
PV power (kW)	105,120	464.95	678.95	0.00	2,690.00
Wind power (kW)	105,120	764.06	705.86	0.00	2,502.50
Electricity price ($/MWh)	105,120	65.27	200.59	-680.27	11,292.44

The average PV and wind production were around 0.46 MW and 0.76 MW, respectively, as compared to the average load demand of 2.49 MW as shown in Table 1. The standard deviations of PV (24.6%) and wind generation (26.4%) are relatively high, suggesting significant variability of renewable energy generation. This behavior can also be observed in Figure 1, showing the PV generation and load demand, along with wind generation and total renewable energy generation over a typical 24-hour period.

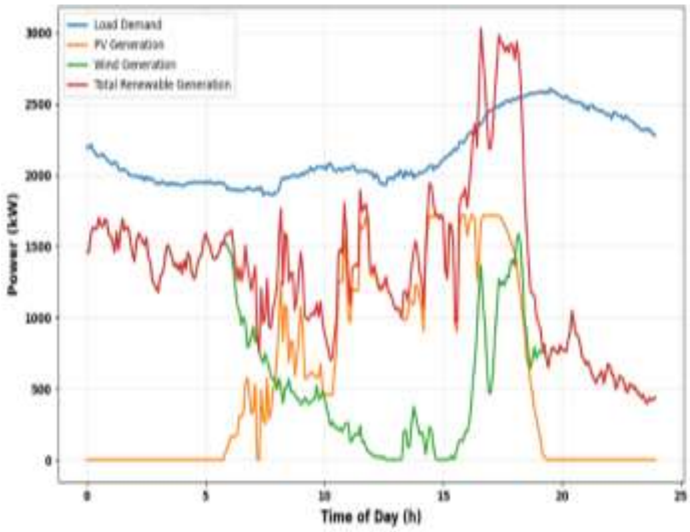

Figure 1. Representative daily load and renewable-generation profile at 5-minute resolution

3.2 Daily Energy and Renewable Penetration

Significant variability was observed between energy demand and renewable energy supply in daily energy analysis. The mean daily electrical demand was around 59.75 MWh while PV and wind generation combined averaged around 29.50 MWh per day. As a result, the average renewable energy penetration is about 49.72% of the daily load demand. Table 2 shows the Daily energy characteristics of the historical dataset.

Table 2. Daily energy characteristics of the historical dataset

Indicator	Mean	SD	Minimum	Maximum
Load energy (kWh/day)	59,754.15	6,221.64	46,698.50	78,654.06

PV energy (kWh/day)	11,158.69	5,170.96	283.83	26,436.25
Wind energy (kWh/day)	18,337.46	12,159.66	56.04	54,931.25
Total renewable energy (kWh/day)	29,496.16	13,218.63	5,339.79	67,067.92
Renewable penetration (%)	49.72	22.80	9.98	122.26

The renewable-penetration ratio varied between about 9.98% and 122.26%, which meant that the microgrid had high renewable energy shortages and periods of oversupply. These contrasting conditions provide suitable operating scenarios for testing charging and discharging, grid exchange and renewable-curtailment decisions. The variability of renewable penetration over the course of a day is shown in Figure 2.

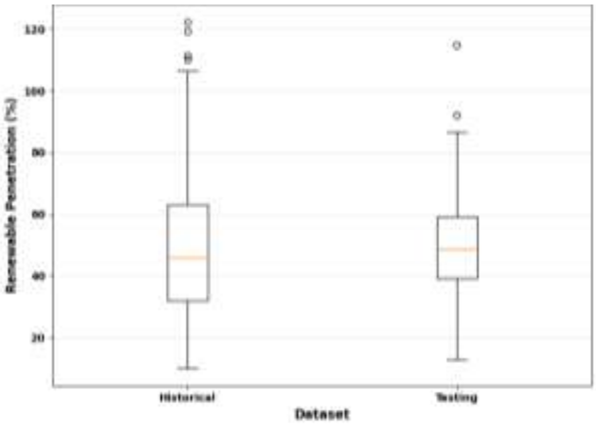

Figure 2. Distribution of daily renewable-energy penetration in the historical and testing datasets

3.3 Comparison of Historical and Testing Conditions

The independent 60-day testing dataset had the same overall features, but made significant differences compared to the historical data. The mean daily load rose from 59.75 MWh in the historical period to 61.07 MWh in the testing period, while the mean renewable penetration dropped from 49.72% to 48.67%.

Table 3. Comparison of historical and independent testing datasets

Indicator	Historical data	Testing data
Number of days	365	60
Number of 5-minute observations	105,120	17,280
Mean daily load (kWh)	59,754.15	61,065.22
Mean daily PV energy (kWh)	11,158.69	11,261.91
Mean daily wind energy (kWh)	18,337.46	18,118.33
Mean daily renewable energy (kWh)	29,496.16	29,380.24
Mean renewable penetration (%)	49.72	48.67
Mean net load (kW)	1,260.75	1,320.21
Renewable-surplus intervals (%)	11.94	10.67
Negative-price intervals (%)	6.36	14.43

The net demand was slightly higher and a substantially larger percentage of the intervals had negative price in the testing period, as reported in Table 3. These differences further enhance the out-of-sample assessment since the testing data set is not a perfect replica of the historical operating distribution. The average daily net load for each test period is plotted in Figure 3.

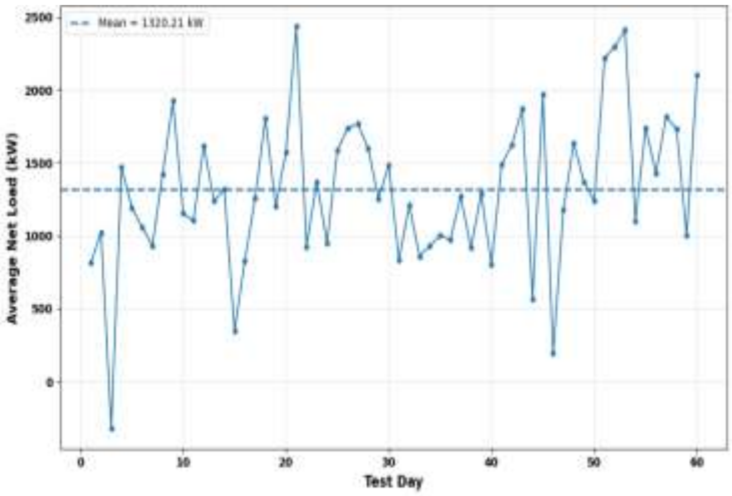

Figure 3. Daily average net load during the 60-day independent testing period

3.4 Net-Load Behavior and Renewable-Surplus Conditions

Net load is the difference between total electrical demand and combined PV and wind generation. The range of net load from the historical data was from about -2,932.53 kW to 4,172.91 kW and for the independent testing data was from -2,069.48 kW to 3,844.43 kW. Table 4 shows the Net load and operational characteristics.

Table 4. Net-load and operational characteristics

Indicator	Historical	Testing
Mean net load (kW)	1,260.75	1,320.21
SD of net load (kW)	1,060.92	1,003.50
Minimum net load (kW)	-2,932.53	-2,069.48
Maximum net load (kW)	4,172.91	3,844.43
Renewable-surplus intervals (%)	11.94	10.67
Negative-price intervals (%)	6.36	14.43
Price > $100/MWh (%)	12.48	4.35

Net-loads were negative when renewable energy was greater than local electrical demand. Historical observations were in renewable-surplus conditions about 11.94% of the time, while testing observations were in renewable-surplus conditions about 10.67% of the time. These time windows can be used for storage charging, for exporting electricity, or to reduce renewable curtailment. High positive NETLOADs, on the other hand, will need more support from the utility grid, dispatchable generation and/or storage. The net load and electricity price relationship for the test period is shown in Figure 4.

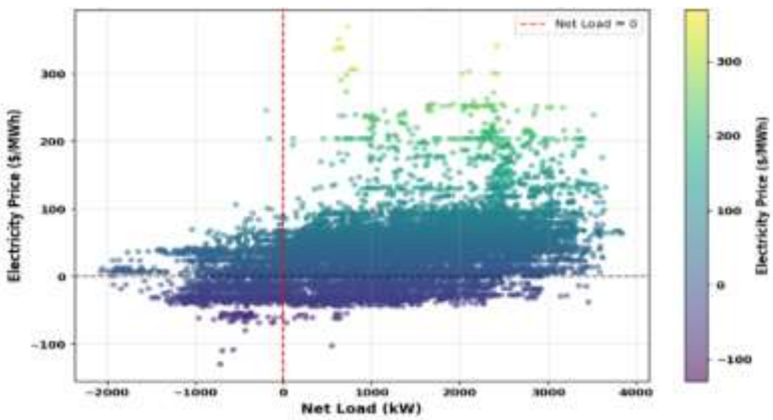

Figure 4. Relationship between net load and electricity price during independent testing

3.5 Electricity-Price Variability

There were significant temporal variations in electricity prices, including both negative prices and large price jumps. The price range during independent testing period was from around -$129.24/MWh to $369.29/MWh, with an average price of $39.29/MWh.

Table 5. Electricity-price characteristics

Statistic	Historical period	Testing period
Mean ($/MWh)	65.27	39.29
SD ($/MWh)	200.59	42.01
Median ($/MWh)	58.49	40.65
Minimum ($/MWh)	-680.27	-129.24
Maximum ($/MWh)	11,292.44	369.29
Negative-price intervals (%)	6.36	14.43
Intervals > $100/MWh (%)	12.48	4.35

Negative electricity prices are more likely to happen during the independent test period than the historical period as indicated in Table 5. These conditions have relevance to smart energy management systems and could affect the economic attractiveness of using batteries, importing energy from the grid, and using renewable sources of energy. The difference between the prices of electricity during the 60-day testing period is displayed in Figure 5.

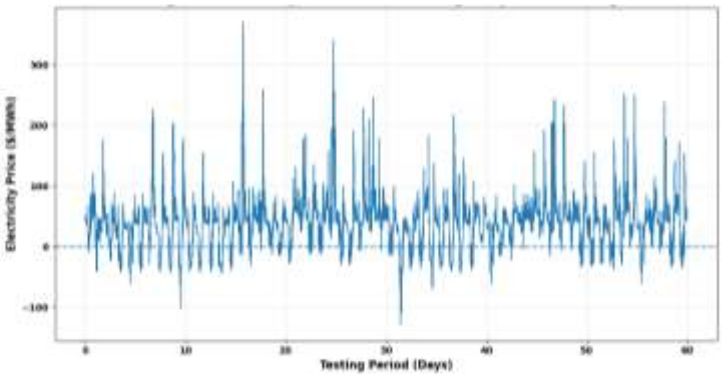

Figure 5. Electricity-price variation during the 60-day independent testing period

3.6 Energy-Storage Operating Constraints

The distributed microgrid featured traditional energy storage devices, and virtual energy storage resources. The conventional storage system (aggregate) had a rated power capacity of ~1,200 kW and an energy storage capacity of ~3,700 kWh, and virtual energy storage had a rated power capacity of ~695 kW and an energy storage capacity of ~2,240 kWh. Table 6 shows the Installed generation and storage capacities.

Table 6. Installed generation and storage capacities

Resource	Aggregate rated power	Aggregate energy capacity
PV generation	2,500 kW	–
Wind generation	2,500 kW	–
Dispatchable generation	1,500 kW	–
Conventional energy storage	1,200 kW	3,700 kWh
Virtual energy storage	695 kW	2,240 kWh

The lower virtual-storage SOC boundary was between 0.072 and 0.449, while the upper was between 0.555 and 0.978. The feasible operating region of the storage system changes as the day progresses, as shown by these changing boundaries. The resulting dynamic SOC limits are shown in the Figure 6.

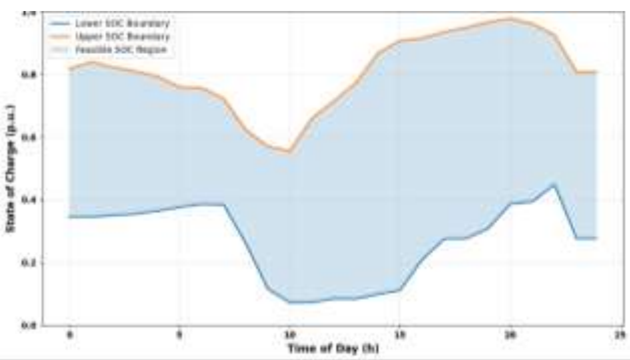

Figure 6. Dynamic lower and upper SOC boundaries for virtual energy storage

3.7 Correlation Among System Variables

Relevant relationships were found in correlation analysis of the independent testing data between load demand, renewable generation, electricity price, and net load. Electric demand was positively correlated with net load, while PV and Wind generation were negatively correlated with net load.

Table 7. Correlation matrix for the independent testing dataset

Variable	Load	PV	Wind	Price	Net Load
Load	1.000	-0.198	-0.008	0.366	0.560
PV	-0.198	1.000	-0.147	-0.472	-0.661
Wind	-0.008	-0.147	1.000	0.118	-0.522
Price	0.366	-0.472	0.118	1.000	0.397
Net Load	0.560	-0.661	-0.522	0.397	1.000

PV generation had the greatest negative correlation with net load (–), as seen in Table 7, followed by wind generation (–). Here, net load (NL) was moderately positively correlated with electricity price (). These relationships show that when making dispatch decisions, the economic efficiency cannot be obtained by considering only one of the three factors demand, renewable availability, storage state, or market-price conditions must all be considered.

3.8 Data-Driven Operating Scenarios

The test set was independent enough to allow for the evaluation of the proposed DRL controller in the scenarios. Days were chosen based upon measurable operating criteria, not subjective visual criteria.

Table 8. Selected operating scenarios for DRL evaluation

Test scenario	Selected test day	Selection criterion
Peak-demand condition	Day 58	Highest daily electrical load
High-renewable condition	Day 3	Highest renewable penetration
Renewable-deficit condition	Day 52	Lowest renewable penetration
High-price condition	Day 25	Highest mean daily electricity price
Typical operating condition	Day 1	Baseline operating profile

Table 8 summarizes the independent testing period, which offers unique peak-demand, renewable-surplus, renewable-deficit, and high-price conditions. These scenarios are appropriate for testing the ability of the proposed DRL controller to perform effectively in significantly different operating conditions.

3.9 Overall Dataset-Based Findings

In conclusion, the findings validate the suitability of the chosen dataset for conducting real-time energy-management studies in an environment that is technically demanding. Demand and renewable generation vary significantly over time, renewable energy penetration is about 50%, there is frequent renewable energy surplus, large net-load excursions, dynamic electricity prices, and time varying storage constraints. Overall, Figures 1–6 and Tables 1–8 indicate that the data set is sufficiently diverse to be used for training and testing a continuous-action distributed power allocation DRL controller. The values displayed in this section, however, represent data and microgrid operating conditions, and not the actual performance of the proposed SAC controller. Operating-cost reduction, renewable-utilization improvement, power-loss reduction, and grid-dependency reduction, cumulative reward, and the superiority over DDPG, TD3, and rule-based control should be reported only after these controllers have been trained and evaluated on that independent 60-day test set.

4. Discussion

The results show that the chosen micro-grid data set is a tough operating condition for the evaluation of real-time energy management using RL-based deep learning. The historical data indicated that the load was around 2.49 MW and the PV and wind generation combined was around 1.23 MW on an average. The renewable penetration was on average almost 50% but ranged from around 10% renewable penetration to more than 120%, meaning that the system was in either renewable deficit or renewable surplus conditions. This variation is significant because an effective DRL controller should be able to adjust its dispatch decisions based on the changing supply-demand situation rather than operating rules. This is in line with the latest findings that DRL can enhance adaptive energy management in microgrids with unpredictable and fluctuating renewable energy sources (Charfeddine et al., 2025; Chen et al., 2024).

The net-load analysis also showed the difficulty of the control problem. Net loads were negative if combined PV and wind exceeded local load and positive if there was a net need for PV and wind to support additional loads from storage, dispatchable generation, and/or imports from the grid. Around 11.94% of the historical observations and 10.67% of the testing observations were renewable surplus. These variations reinforce the need for continuous-action control as the decisions to be made for charging, discharging, curtailing, and grid exchange can be quite different between each five-minute period. These same benefits of adaptive dispatch using DRL are mentioned in microgrid energy transmission (Chen et al., 2024) as well as in hierarchical energy-management applications (Li et al., 2025).

The results of the electricity-price give an additional economic aspect to the control problem. The independent testing period had high price events as well as negative price intervals, with around 14.43% of the testing observations having negative electricity prices. These circumstances can lead to non-trivial scheduling problems, since it might be more profitable to import electricity for storage charging when the price is negative, while discharging or decreasing the grid dependence is more attractive when the price is high. A positive correlation of net load and electricity price is observed, further supporting the need for making the economically optimal decision regardless of the renewable's availability. The ability to include market signals and control periods in DRL-based economic dispatch models is also highlighted in recent studies (Liu & Mao, 2025; Moga et al., 2025).

Another key aspect of the proposed framework is storage flexibility. Conventional energy storage and time varying upper and lower state of charge boundaries for virtual energy storage are included in the dataset. These dynamic limits enable the problem to become more realistic because the set of possible actions that can be taken varies over the course of the operating horizon. Another study, which focused on the study of nonlinear battery losses and hybrid storage integration, also emphasized the need to maximize the storage usefulness while minimizing the cycling and constraint violations (Domínguez-Barbero et al., 2024, Kumar et al., 2025). The use of an actor-critic algorithm like Soft Actor-Critic is thus suitable as it allows for continuous charging and discharging and can consider penalty costs from infeasible operating conditions.

The proposed SAC formulation is also driven by recent advances in constrained and distributed reinforcement learning. In areas where continuous control, entropy-based exploration and operational feasibility are all inter-related, Soft Actor-Critic has gained a lot of traction. Constrained SAC formulations with Lagrangian penalties have been shown to have the potential to ensure secure operation of unbalanced microgrids (Cortés et al., 2026). In the same way, multi-agent and hierarchical approaches have been proposed to coordinate interconnected microgrids and shared energy storage and energy resources with different characteristics (Esan et al., 2026; Jendoubi & Bouffard, 2023; Lyu et al., 2026; Song et al., 2025). These developments indicate that the single-agent model used here could be a good basis for subsequent extensions to distributed or multi-agent coordination.

The separate 60-day test set is also very useful as it allows the evaluation of the model outside of the historical period with different operating conditions. The testing data had a slightly higher average, and more of the observations were negative prices, giving a solid footing for generalizing. This is crucial since a real-time controller should be able to perform reliably in situations different from the ones it was trained in. The role of adaptable decision-making while considering varying grid-connected storage scenarios has also been noted in reinforcement learning research conducted online (Shi et al., 2025).

In summary, the findings based on the dataset are confirming the approach for real time energy management and optimum power allocation based on DRL. The results shown in the description provide a characterization of the operating environment, however, the present results do not characterize the final superiority of SAC over DDPG, TD3 or rule-based control. These statements should only be made following the training of the models and independent testing. The following comparative analysis should thus be directed toward operating cost, renewable utilization, dependency on the grid, power losses, constraint violations, storage behavior and computational decision time. This will enable the final study to ascertain whether or not the proposed SAC controller is able to deliver tangible technical and economic benefits over this inherent variability.

5. Conclusion

The dataset "Prediction-Free Coordinated Dispatch of Microgrid: A Data-Driven Online Optimization Approach" was used in this study for the real-time energy management and optimal power allocation (OPL) in a distributed electrical system (DES), and a deep reinforcement learning-based framework was built. The data selected included high-resolution 5-minute measurements of load demands, PV energy production, wind energy production, electricity prices, and energy storage constraints which allowed for a realistic simulation of microgrid dynamic behavior. The results showed that renewable generation, net load, market prices and the penetration of renewable energies were significantly different, highlighting the need for adaptive control strategies to be able to deal with rapidly changing operating conditions. Renewable surplus periods, negative electricity prices, and varying storage capacities added to this complexity when coordinating dispatch and storage management. The proposed Soft Actor-Critic formulation is thus suitable for continuous decision-making of power allocation in storage, dispatchable generation and interaction with the grid. Independent testing data also provides a solid basis for assessing generalization in conditions different from those seen during training. In conclusion, the study offers a solid methodological foundation for using DRL in intelligent operation of microgrids. Future work will test the trained controller against other policies including DDPG, TD3, rule-based policies, and hardware-in-the-loop platforms, focusing on cost reduction, renewable utilization, system losses, constraint compliance, and computational performance in real time.

References

1. Alabdullah, M. H., & Abido, M. A. (2022). Microgrid energy management using deep Q-network reinforcement learning. *Alexandria Engineering Journal*, *61*(11), 9069-9078.
2. Cai, W., Kordabad, A. B., & Gros, S. (2023). Energy management in residential microgrid using model predictive control-based reinforcement learning and Shapley value. *Engineering Applications of Artificial Intelligence*, *119*, 105793.
3. Charfeddine, M., Ben Moussa, M., & Jouili, K. (2025). Deep-reinforcement-learning-based sliding mode control for optimized energy management in DC microgrids. *Mathematics*, *13*(19), 3212.
4. Chen, S., Liu, J., Cui, Z., Chen, Z., Wang, H., & Xiao, W. (2024). A deep reinforcement learning approach for microgrid energy transmission dispatching. *Applied Sciences*, *14*(9), 3682.
5. Cortés, P., Tabares, A., Bolaños, R., & Bonilla, K. (2026). Soft actor-critic energy management in three-phase unbalanced microgrids with lagrangian penalty constraints. *Scientific Reports*.
6. Darshi, R., Shamaghdari, S., Jalali, A., & Arasteh, H. (2023). Decentralized reinforcement learning approach for microgrid energy management in stochastic environment. *International Transactions on Electrical Energy Systems*, *2023*(1), 1190103.
7. Domínguez-Barbero, C., García-González, J., Sanz-Bobi, M. Á., & García-Cerrada, A. (2024). Energy management of a microgrid considering nonlinear losses in batteries through Deep Reinforcement Learning. *Applied Energy*, *368*, 123435.
8. Esan, A. B., Shareef, H., & ALAhmad, A. K. (2026). Lean multi-agent deep reinforcement learning for uncertainty handling in the energy management of networked microgrids. *Applied Energy*, *408*, 127354.
9. Glover, D., Krishnamoorthy, G., Ren, H., Dubey, A., & Gebremedhin, A. (2025). Deep reinforcement learning for distribution system operations: A tutorial and survey. *Proceedings of the IEEE*.
10. Guo, C., Wang, X., Zheng, Y., & Zhang, F. (2022). Real-time optimal energy management of microgrid with uncertainties based on deep reinforcement learning. *Energy*, *238*, 121873.
11. Harrold, D. J., Cao, J., & Fan, Z. (2022). Renewable energy integration and microgrid energy trading using multi-agent deep reinforcement learning. *Applied Energy*, *318*, 119151.
12. Huang, Kaidi (2024), "Prediction-Free Coordinated Dispatch of Microgrid: A Data-Driven Online Optimization Approach", Mendeley Data, V2, doi: 10.17632/8j2zdf6h5j.2
13. Jayaraj, S., Ahamed, T. I., Abraham, M. P., Jasmin, E. A., & Sulthan, S. M. (2024). Reinforcement learning for energy management of an islanded microgrid: Analysis of state, state space and reinforcement function. *Sustainable Energy, Grids and Networks*, *38*, 101362.
14. Jendoubi, I., & Bouffard, F. (2023). Multi-agent hierarchical reinforcement learning for energy management. *Applied Energy*, *332*, 120500.
15. Ji, Y., Wang, J., Xu, J., & Li, D. (2021). Data-driven online energy scheduling of a microgrid based on deep reinforcement learning. *Energies*, *14*(8), 2120.
16. Kumar, K., Kwon, S., & Bae, S. (2025). Deep reinforcement learning-based control strategy for integration of a hybrid energy storage system in microgrids. *Journal of Energy Storage*, *108*, 114936.
17. Lami, B., Alsolami, M., Alferidi, A., & Slama, S. B. (2025). A smart microgrid platform integrating AI and deep reinforcement learning for sustainable energy management. *Energies*, *18*(5), 1157.
18. Li, J., Jiang, Z., Chen, Z., Liu, J., & Cheng, L. (2024). CuEMS: Deep reinforcement learning for community control of energy management systems in microgrids. *Energy and Buildings*, *304*, 113865.
19. Li, Y., Chang, W., & Yang, Q. (2025). Deep reinforcement learning based hierarchical energy management for virtual power plant with aggregated multiple heterogeneous microgrids. *Applied Energy*, *382*, 125333.
20. Liu, D., Zang, C., Zeng, P., Li, W., Wang, X., Liu, Y., & Xu, S. (2023). Deep reinforcement learning for real-time economic energy management of microgrid system considering uncertainties. *Frontiers in Energy Research*, *11*, 1163053.
21. Liu, W. C., & Mao, Z. Z. (2025). Microgrid economic dispatch using Information-Enhanced Deep Reinforcement Learning with consideration of control periods. *Electric Power Systems Research*, *239*, 111244.
22. Lyu, Z., Hu, B., Wei, Z., Zhou, H., Bao, H., Guo, Z., ... & Shi, L. (2026). Deep reinforcement learning-based energy management strategy for heterogeneous interconnected multi-energy microgrid systems with shared energy storage. *Available at SSRN 5472587*.

23. Moga, O. N., Florea, A., Solea, C., & Vintan, M. (2025). Reinforcement learning-based energy management in community microgrids: A comparative study. *Sustainability*, *17*(23), 10696.
24. Nakabi, T. A., & Toivanen, P. (2021). Deep reinforcement learning for energy management in a microgrid with flexible demand. *Sustainable Energy, Grids and Networks*, *25*, 100413.
25. Shen, H., Shen, X., & Chen, Y. (2024). Real-time microgrid energy scheduling using meta-reinforcement learning. *Energies*, *17*(10), 2367.
26. Shi, X., Didier, F., Badji, A., Azid, S. I., Cirrincione, M., & Laghrouche, S. (2025, October). Online Reinforcement Learning Approach for Real-Time Energy Management of Grid-Connected Battery Storage Microgrid System. In *IECON 2025–51st Annual Conference of the IEEE Industrial Electronics Society* (pp. 1-7). IEEE.
27. Song, D., Yan, L., Dai, X., Zhu, X., Hagenmeyer, V., & Zhai, J. (2025). Low-carbon energy management for networked multi-energy microgrids using multi-agent soft actor-critic algorithm. *Sustainable Energy, Grids and Networks*, 101821.