American Journal of Neural Networks and Applications
2026, Vol. 12, No. 2, pp. 40–56
https://doi.org/10.11648/j.ajnna.20261202.11

Research Article

Deep Reinforcement Learning-Based Energy Management of Battery-Supercapacitor Hybrid Storage Systems in Renewable Microgrids

Daniel Kumi Owusu*

Department of Electrical and Electronic Engineering, Takoradi Technical University, Takoradi, Ghana

Abstract

The increasing integration of renewable energy sources into microgrids has intensified the need for intelligent energy management strategies capable of addressing the intermittency of solar and wind generation while ensuring reliable and cost-effective operation. Although Rule-Based Control (RBC) methods are straightforward to implement, their limited adaptability often leads to suboptimal utilisation of Hybrid Energy Storage Systems (HESS). This study develops and evaluates a Deep Reinforcement Learning (DRL)-based energy management system employing a Deep Q-Network (DQN) to coordinate battery–supercapacitor operation within a renewable microgrid. A Gymnasium-compatible simulation environment was constructed using a publicly available time-series dataset comprising renewable generation, load demand, electricity prices, battery state of charge (SoC), and supercapacitor SoC. Feature engineering, incorporating sinusoidal temporal representations and Min-Max normalisation, was applied to enhance learning stability and capture cyclical demand and generation patterns. The DQN agent was trained over 50 episodes and benchmarked against a conventional RBC strategy under identical operating conditions. Training performance demonstrated progressive policy improvement, with cumulative rewards increasing from approximately -1200 to -400, indicating enhanced decision-making capability during learning. The learned controller exhibited adaptive energy scheduling through dynamic utilisation of the supercapacitor and selective grid interaction in response to varying operating conditions, whereas the RBC followed a deterministic control strategy with limited flexibility. However, comparative evaluation revealed that the DQN did not consistently outperform the RBC in cumulative economic performance, suggesting the need for further refinement of the reward function, training process, and hyperparameter configuration. Nevertheless, the proposed framework demonstrates the feasibility of applying deep reinforcement learning to coordinated battery–supercapacitor energy management and highlights its potential to enhance operational flexibility and intelligent resource utilisation in renewable microgrids. The study contributes a dataset-driven reinforcement learning framework that provides a foundation for future research on advanced AI-based energy management systems and the integration of more sophisticated reinforcement learning algorithms for resilient and sustainable microgrid operation.

Keywords

Adaptive Energy Scheduling, Deep Reinforcement Learning, Hybrid Energy Storage Systems, Microgrid, Renewable Energy, Smart Energy Systems, Supercapacitor

*Correspondence: Daniel Kumi Owusu (daniel.owusu@ttu.edu.gh)

Received: 3 July 2026; **Accepted:** 20 July 2026; **Published:** 17 August 2026

1. Introduction

The growing integration of renewable energy sources, particularly solar and wind, into contemporary power systems has accelerated the advancement of microgrids as a practical solution for decentralised and sustainable energy generation [1]. Microgrids facilitate localised production and consumption of energy, thereby improving reliability, resilience, and efficiency [2]. Nevertheless, the inherent intermittency and stochastic behaviour of renewable energy sources present considerable challenges in sustaining a stable equilibrium between supply and demand [3]. Variations in generation, combined with dynamic load profiles and fluctuating electricity prices, necessitate the adoption of advanced energy management strategies to ensure optimal system performance [4]. Within this framework, energy storage systems assume a pivotal role in mitigating variability and reinforcing grid stability [5].

HESS, which integrate complementary technologies such as batteries and supercapacitors, have emerged as an effective means of overcoming the limitations associated with single-storage solutions [6]. Batteries provide high energy density and are well suited to long-term energy storage, while supercapacitors deliver high power density and are particularly effective in managing short-term fluctuations [7]. The combination of these technologies enhances operational flexibility and efficiency in microgrid applications [8]. However, the successful deployment of hybrid systems necessitates advanced control mechanisms capable of dynamically allocating energy resources in response to changing system conditions [9].

Conventional energy management strategies, particularly RBC approaches, are widely adopted owing to their simplicity and ease of implementation [10]. These methods operate on predefined heuristics, such as charging during periods of surplus generation and discharging during deficits [11]. While effective in straightforward scenarios, RBC strategies are inherently reactive and lack the adaptability required for complex and dynamic operating environments [12]. Consequently, they often result in suboptimal energy utilisation, increased reliance on the grid, inefficient State-of-Charge (SoC) management, and elevated operational costs [13]. These limitations underscore the need for intelligent, data-driven approaches capable of learning optimal control policies through system interactions [14].

The objective of this study is to design and evaluate a DQN-based energy management system for the coordinated operation of battery-supercapacitor storage within a renewable microgrid. Specifically, it seeks to model a dataset-driven microgrid environment, formulate an adaptive reward function to enable intelligent energy scheduling, investigate the operational behaviour of the learned control policy under varying renewable generation and load conditions, and assess its performance against a conventional RBC.

The significance of this research lies in its contribution to advancing intelligent energy management in microgrids through the integration of artificial intelligence techniques. The proposed system facilitates adaptive decision-making in complex and uncertain environments by employing reinforcement learning, thereby enhancing energy efficiency, reducing operational costs, and improving the overall reliability of microgrid operations [15]. Moreover, the study enriches the growing body of knowledge on Artificial Intelligence (AI) applications in power systems, particularly in the optimisation of hybrid energy storage systems [16].

The key contributions and anticipated outcomes of this study include the development of a Gymnasium-compatible reinforcement learning environment for modelling hybrid energy storage dynamics, the implementation of a DQN-based EMS capable of learning optimal control policies, and a comprehensive comparative analysis against a rule-based baseline. The results reveal distinct operational behaviours between the DQN and Rule-Based Controller, particularly in grid interaction patterns and energy storage utilisation. The findings confirm the feasibility of reinforcement learning for hybrid energy storage management, while also highlighting the need for further optimisation to enhance policy performance. These findings highlight the effectiveness of the proposed method in achieving both economic and operational optimisation.

The scope of this study is confined to the design and evaluation of an AI-driven EMS for a microgrid integrating renewable energy sources with a hybrid battery-supercapacitor storage system. The analysis is conducted using time-series data derived from publicly available datasets, encompassing renewable generation, load demand, and electricity price signals. The reinforcement learning framework is implemented within a simulated environment using a discrete action space, and its performance is assessed through key metrics, including operational cost, grid imports, and state-of-charge behaviour.

The novelty of this study lies in the integration of Deep Q-Network-based reinforcement learning with hybrid energy storage management in microgrids, complemented using engineered temporal features to enhance learning performance. In contrast to conventional approaches, the proposed method enables proactive and adaptive energy scheduling by capturing both short-term and long-term system dynamics. The demonstrated adaptive control behaviour and dynamic utilisation of hybrid energy storage resources highlight the potential of deep reinforcement learning as a robust approach to intelligent microgrid energy management.

1.1. Related Works

To provide a comprehensive overview of recent advances in intelligent microgrid energy management, a systematic literature review was conducted using Web of Science, Scopus, ScienceDirect, and IEEE Xplore. The search employed key terms including *deep reinforcement learning*, *energy management*, *hybrid energy storage system*, *battery-supercapacitor*, and *renewable microgrid*. The review was restricted to studies published between 2024 and 2026, thereby capturing the most recent developments in this rapidly evolving field. The search identified more than 150 publications, which were subsequently screened for relevance, originality, methodological rigour, and alignment with the objectives of this study. Following this screening process, 14 high-quality papers were selected for detailed examination and comparative analysis, as summarised in Table 1.

1.2. Research Gap

Despite growing interest in Deep Reinforcement Learning (DRL) for microgrid energy management, comparatively little attention has been directed towards the coordinated control of battery–supercapacitor hybrid storage systems within unified, dataset-driven simulation environments. Existing research often concentrates on battery-only configurations, simplified state representations, or theoretical control frameworks that fail to adequately reflect the temporal variability of renewable generation and load demand. Moreover, the operational behaviour of DQN controllers in managing complementary storage technologies remains insufficiently explored under consistent renewable generation, load, pricing, and storage conditions. To address these gaps, this study develops a Gymnasium-compatible reinforcement learning environment for a renewable energy microgrid incorporating a battery-supercapacitor HESS. The framework integrates engineered temporal features through sinusoidal time representations to enhance the agent's ability to capture cyclical generation and demand patterns. Unlike previous studies that focus predominantly on battery-only systems or theoretical DRL implementations, the proposed approach explicitly models the coordinated operation of battery and supercapacitor resources and evaluates a DQN-based Energy Management System against a conventional RBC strategy under identical time-series datasets and operating conditions. This enables a more comprehensive assessment of adaptive energy scheduling, storage utilisation, and grid interaction behaviour in renewable microgrids.

Table 1*. Literature review of related works.*

SN	Methods Employed	Description	Results	Constraints	Ref.
1.	PSO and neural networks combined hierarchically to decouple fast and slow storage dynamics	Four-bus ringmain DC microgrid with hybrid storage, simulated in MATLAB and validated experimentally	Hybrid PSO-NN framework improved voltage regulation, reduced battery stress, and enhanced transient response	Required extensive offline neural training, with performance dependent on dataset quality and representativeness	[17]
2.	PRISMA-guided review synthesising 214 DRL studies across hybrid power system applications	Analysis of DRL algorithms, applications, verification platforms, challenges, and prospects	DRL achieved near-optimal energy efficiency, strong adaptability, and superior robustness compared to traditional methods	Deployment hindered by low sample efficiency, safety concerns, interpretability issues, and reality gaps	[18]
3.	Integrated quantum particle swarm optimisation with deep reinforcement learning for dynamic-aware EMS	It investigated microgrid energy management using BESS under uncertain inertia, damping, and transient conditions	The hybrid framework improved economic efficiency by 52.2% and enhances BESS charging/discharging performance	DRL faces challenges from non-Markovian dynamics, high-dimensional state spaces, and computational training complexity	[19]
4.	It employed multi-objective mixed-integer nonlinear programming with metaheuristic algorithms and deep reinforcement learning	It investigated techno-economic optimisation of off-grid hybrid renewable microgrids using dual-battery storage	The framework reduced life-cycle costs by over 20%, emissions by 30%, and battery degradation by 10%	Limitations from simplified simulations, incomplete degradation modeling, and scarce real-world reinforcement learning deployments.	[20]
5.	Applied bi-LSTM forecasting with DQN and SAC reinforcement learning for HRES control	Examined supervisory control of solar, wind, biomass, and hydrogen storage within a unified HRES	SAC agent achieved superior adaptability, reducing energy imbalance below 0.5 MWh and improving stability	Reliance on a short 24-hour dataset and absence of conventional baseline comparisons	[21]
6.	Reviewed smart grid and microgrid advances using state-of-the-art literature	Researched on planning, operation, economics, and resilience of distributed energy systems	Highlighted on synergistic integration of distributed generation and storage improving efficiency and resilience	Reliance on simulated case studies and incomplete real-world validation across diverse contexts	[22]

SN	Methods Employed	Description	Results	Constraints	Ref.
7.	Analysed 101 selected articles using structured keyword searches and comparative synthesis	Examined microgrid energy management systems integrating machine learning and IoT technologies	Highlighted EMS roles in optimising microgrid stability, efficiency, and reliability through advanced technologies	Renewable variability, cybersecurity risks, and deployment complexities limiting microgrid performance	[23]
8.	Analysed diverse optimisation strategies for CPS-based microgrids using literature synthesis	Surveyed forecasting, demand management, economic dispatch, and unit commitment in CPS microgrids	Multi-agent and meta-heuristic approaches outperformed conventional methods in decentralised, dynamic CPS microgrids	Advanced optimisation techniques remained underutilised in forecasting and demand management, limiting scheduling accuracy	[24]
9.	Reviewed microgrid control strategies and IoT-enabled monitoring systems	Examined AC, DC, and hybrid microgrid modes with distributed energy resources integration	Emphasised IoT-based monitoring improved stability, efficiency, and resilience of microgrid operations	Intermittency of renewables, cybersecurity risks, and complexity of advanced control strategies	[25]
10.	Reviewed centralised, decentralised, and hierarchical microgrid control strategies	Studied renewable energy-based microgrid systems, IoT monitoring, and emerging control technologies	Advanced controllers and IoT integration significantly improved microgrid stability, scalability, and operational efficiency	System complexity, scalability issues, cybersecurity risks, and limitations of conventional linear controllers	[26]
11.	The study implemented a DQN-based reinforcement learning algorithm on TI C2000 controller	Explored adaptive real-time smart charging of supercapacitors using reinforcement learning techniques	The RL framework achieved high capacitance retention after 8,000 cycles with reduced degradation	Focused on short-term cycling tests without extensive validation across diverse operating conditions	[27]
12.	Reviewed conventional and advanced microgrid control strategies, including IoT integration	Analysed microgrid control methods, energy management systems, and emerging technologies for resilience and efficiency	The review highlighted strengths and limitations of droop, PID, MPC, ANN, and IoT - enabled control frameworks	Challenges of scalability, computational demands, cybersecurity risks, and limited real-world validation of advanced techniques	[28]
13.	Developed and simulated a communication-free SOC-based cooperative control strategy for hybrid energy storage	Studied a solar-PV DC microgrid integrating batteries and supercapacitors with droop-controlled converters	The proposed SOC-based control achieved seamless mode switching, stable voltage regulation, and extended battery lifespan	The approach did not address aging effects, SOC estimation errors, or declining storage performance over time	[29]
14.	Conducted a review of IoT, AI, blockchain, and digital twin applications in microgrids	Analysed Industry 4.0 technologies for enhancing microgrid resilience, efficiency, and sustainable energy integration	Demonstrated that IoT, cloud computing, AI, and blockchain significantly improved microgrid monitoring, control, and security	Scalability, cybersecurity, data management, and limited real-world implementation of advanced technologies	[30]

2. Materials and Methods

2.1. Dataset Description and Preprocessing

The foundation of this study is a comprehensive time-series dataset obtained from the Kaggle repository Hybrid Energy Storage Dataset (Ziya07, Kaggle). This dataset provides the essential environmental variables for a HESS, which regulates the interaction between a battery (characterised by high energy density) and a supercapacitor (characterised by high power density). A HESS is modelled as a Gymnasium compatible reinforcement learning environment for microgrid energy management. The seven-dimensional state space consists of normalised renewable generation, load demand, electricity price, battery SoC, supercapacitor SoC, and two cyclical temporal features, sin (hour) and cos (hour). The discrete actions of charging, discharging, and idle are optimised through a reward function that penalises grid imports, excessive degradation, and violations of SoC limits, thereby ensuring reliable and cost-efficient operation.

To enhance training stability and enable the Reinforcement Learning (RL) agent to effectively capture and interpret temporal patterns, the following preprocessing procedures were applied:

1) Feature engineering: to provide the model with a more comprehensive understanding of the environment, several additional features were engineered from the raw data. Combined_Renewable_kW was obtained by summing the Solar_Power_kW and Wind_Power_kW columns, thereby consolidating the primary renewable energy sources into a single metric. The Timestamp column was utilised to generate cyclical features, sin_hour and cos_hour. These features are particularly important as they capture the periodic nature of energy generation and consumption (for example, solar output typically peaks at midday), enabling the agent to adopt a proactive rather than purely reactive control strategy; and
2) Data Scaling: The continuous observation features, namely Combined_Renewable_kW, Load_Demand_kW, and Electricity_Price_USD per kWh, were normalised using a MinMaxScaler. This transformation standardises the data within a uniform range (0 to 1), thereby preventing features with larger numerical magnitudes from exerting disproportionate influence on the neural network's learning process and enhancing overall training stability.

2.2. Architecture of the Microgrid System

The grid-connected microgrid integrates renewable generation (solar PV and wind), hybrid storage and utility grid support (battery-supercapacitor system and utility grid), and a DQN-based energy management system that coordinates generation, demand, pricing, and storage states to ensure reliable and optimised operation, as illustrated in Figure 1. This study considers two renewable generation sources namely, solar photovoltaic and wind power, together with complementary storage technologies comprising a battery and a supercapacitor. The DQN framework is structurally extensible, as additional generation or storage units can be incorporated by expanding the state vector, action space, operating constraints, and reward function. However, the reported results are confined to the solar-wind and battery-supercapacitor configuration represented in the selected dataset, and the controller's performance with other technologies was not evaluated in this study.

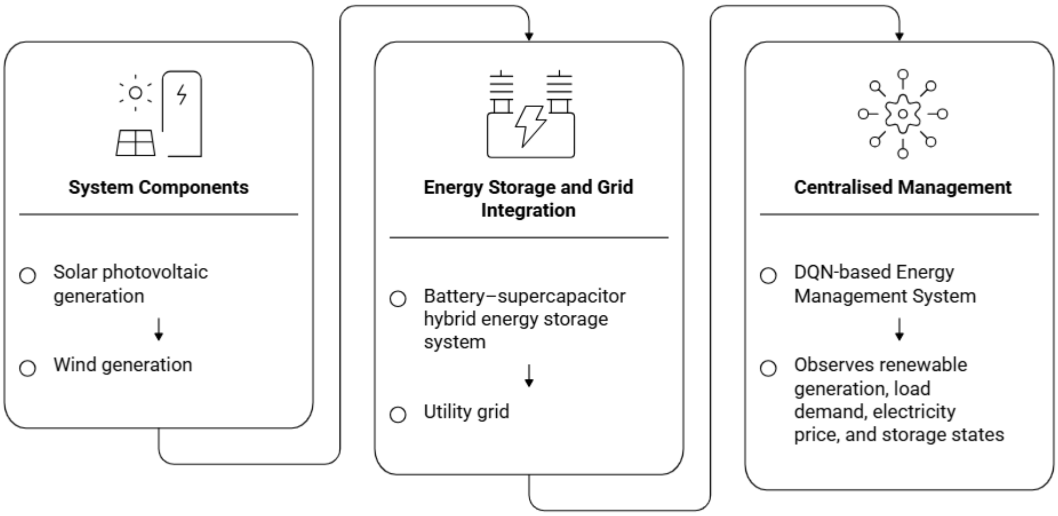

Figure 1. *Schematic Representation of the Microgrid System Architecture.*

2.3. Representation and Impact of Load Demand

Load demand was treated as an aggregated time-series variable rather than separated into residential, commercial, industrial, critical, or flexible categories. At each decision step, the normalised demand formed part of the DQN state vector and directly influenced battery and supercapacitor actions. Higher demand relative to renewable generation created an energy deficit, prompting storage discharge or grid import. Lower demand or higher renewable output produced surplus energy, which was directed to storage charging. Demand variability also influenced storage coordination. Gradual and sustained changes were mainly managed by the battery, while rapid fluctuations were handled by the supercapacitor. The dataset, however, represents demand only at an aggregated level and excludes load criticality, demand response, power factor, harmonic distortion, and appliance-level behaviour. As a result, the effects of distinct load categories were not evaluated.

2.4. Deep Reinforcement Learning Process

Figure 2 presents the process (flowchart) of the DRL framework. At each timestep, the agent observes the current microgrid state and selects one of 25 combined battery–supercapacitor actions using an ε-greedy policy. The environment then updates the storage states, computes grid exchange and operating penalties, and returns the subsequent state and reward. Each transition is stored in replay memory and utilised to update the online DQN through mini-batch learning. To enhance training stability, the target network is periodically updated. Upon completion of the training episodes, the learned policy is evaluated and benchmarked against the RBC under identical data and system constraints.

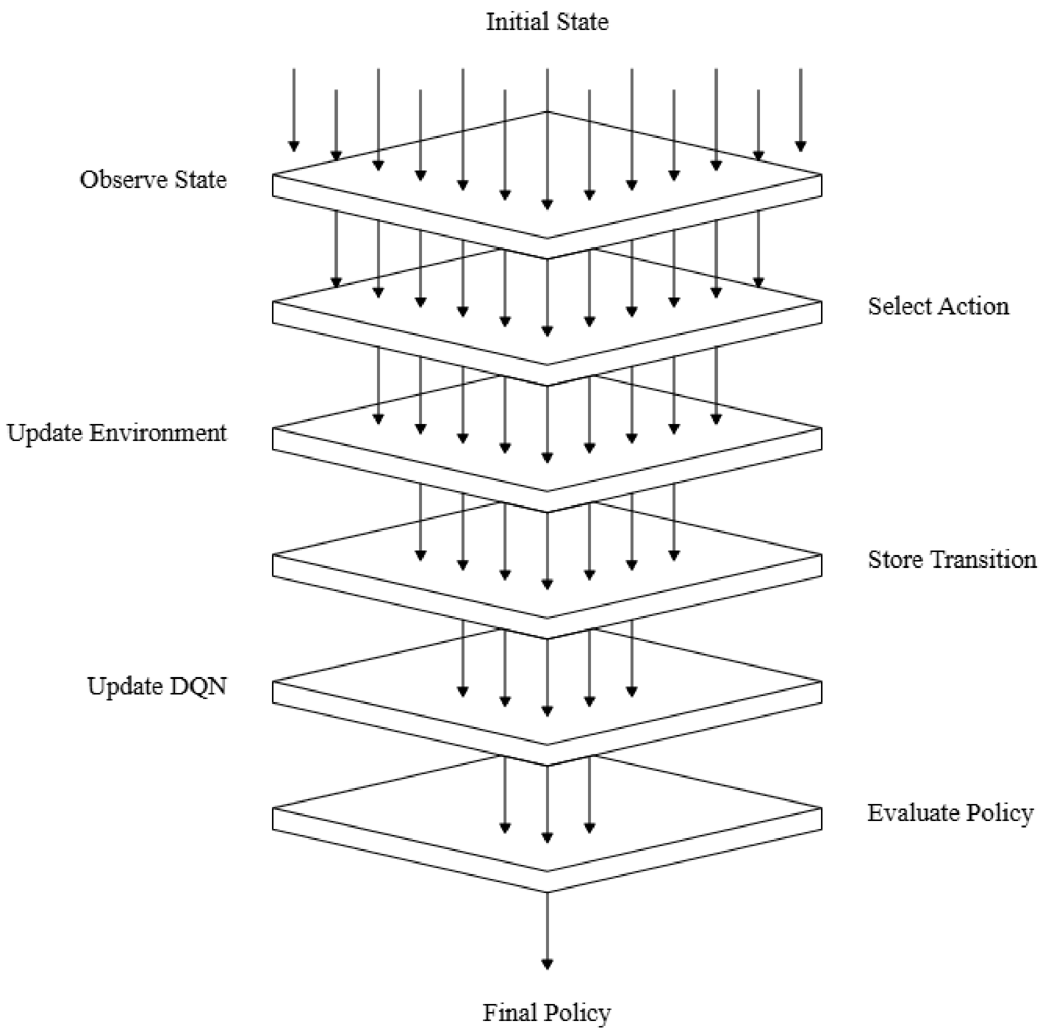

Figure 2. *Flowchart of the Deep Reinforcement Learning Process.*

2.5. Mathematical Formulation of the Energy Management Framework

2.5.1. Sample Combined Renewable Power

Solar and wind generation were aggregated into one renewable power variable, enabling unified evaluation at each decision-making interval, expressed as:

$$P_{renew}(t) = P_{solar}(t) + P_{wind}(t) \quad (1)$$

where, $P_{renew}(t)$ is the total renewable power available at time t, $P_{solar}(t)$ solar power generation at time t, $P_{wind}(t)$ is the wind power generation at time t, and t is the step within the simulation period.

2.5.2. Cyclical Temporal Feature Representation

Energy generation and consumption follow distinct daily cycles. To preserve the continuity of time and avoid artificial breaks between successive hours, the hour variable was expressed using sine and cosine transformations, as follows:

$$\text{sin\_hour} = \sin\left(\frac{2\pi h}{24}\right)$$
$$\text{cos\_hour} = \cos\left(\frac{2\pi h}{24}\right) \tag{2}$$

where, h is the hour of the day (0 - 23), π is a mathematical constant (3.14159), sin_hour is the sine representation of hourly periodicity, and cos_hour is the cosine representation of hourly periodicity.

2.5.3. Data Normalisation

To maintain numerical consistency across input variables and enhance neural network convergence during training, all continuous variables were normalised using the Min-Max scaling method, expressed as:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{3}$$

where, X is the original feature value, X_{min} is the minimum value of the feature, X_{max} is the maximum value, and X_{norm} is the normal feature value.

2.5.4. State Space Representation

The reinforcement learning agent engages with the microgrid through a seven-dimensional state space that encapsulates the prevailing operating conditions of the system, defined as:

$$S_t = \left[P_{renew,t}, L_t, EP_t, SoC_{bat,t}, \sin(h_t), \cos(h_t)\right] \tag{4}$$

where, S_t is the state vector at time t, $P_{renew,t}$ is the renewable power generation, L_t is the load demand, EP_t is the electricity price, $SoC_{bat,t}$ is the battery state of charge, $\sin(h_t)$ is the sine-based temporal feature, and $\cos(h_t)$ is the cosine-based temporal feature.

2.5.5. Reward Function Formulation

The objective of the DQN-based Energy Management System is to minimise operational costs while ensuring reliable microgrid operation. To achieve this, a reward function was designed to penalise undesirable actions, including excessive grid imports, storage degradation, and violations of storage operating limits, as expressed in:

$$G_t = \max\left(0, P_{load}(t) - P_{renew}(t) - P_{bat}(t) - P_{sc}(t)\right) \times EP(t) \tag{5}$$

where G_t is the grid import penalty at time step t that denotes the monetary expense or operational drawback associated with importing electricity from the external main grid at a given time step, $\max(0,.)$ is the operator that establishes a boundary condition that restricts grid penalty evaluation to instances where the load demand exceeds the aggregate contribution of renewable generation and storage, $P_{load}(t)$ is the load demand that represents the microgrid's electricity usage requirement at time step t, $P_{renew}(t)$ is the combined renewable generation that represents the total power output from local solar and wind resources at time step t, as defined earlier in Equation (1), $P_{bat}(t)$ is the battery power output that represents the commanded charging or discharging power of the battery system, $P_{sc}(t)$ is the supercapacitor power output that represents the transient charging or discharging power of the supercapacitor system, and $EP(t)$ is the electricity price that shows the time-dependent grid electricity tariff at time step t (USD/kWh).

Also, the Storage Degradation Penalty is expressed as in:

$$D_t = \omega_{bat}\left(P_{bat}(t)\right)^2 + \omega_{sc}\left(P_{sc}(t)\right)^2 \tag{6}$$

where D_t is the storage degradation penalty at time step t that represents the abstract penalty imposed on the agent to deter intensive usage behaviours that accelerate the physical ageing of hybrid storage assets, ω_{bat} is the battery wear coefficient that represents a fixed scaling factor that captures the elevated long-term degradation costs of electrochemical batteries operating under high-current conditions, ω_{sc} is the supercapacitor wear coefficient that represents a structural weighting factor assigned to the supercapacitor, $\left(P_{bat}(t)\right)^2$ and $\left(P_{sc}(t)\right)^2$ are the squared power outputs that represent the quadratic penalty on power terms, guiding the DQN agent towards low-amplitude operation and thereby reducing stress-induced battery degradation.

Likewise, the State-of-Charge Violation Penalty is expressed as in:

$$V_t = \sum_{i\in\{bat,sc\}} [\max\left(0, SoC_i(t) - SoC_{i,max}\right) + \max\left(0, SoC_{i,min} - SoC_i(t)\right)] \quad (7)$$

where V_t is the State-of-Charge Violation Penalty at time step *t*, representing a strict penalty triggered when the state of charge breaches safe limits, $\sum_{i\in\{bat,sc\}}$ is the Summation operator aggregating boundary checks for battery and supercapacitor into a single penalty, $SoC_i(t)$ is the current state of charge, representing the real-time percentage or energy level of the storage asset at time step t, $SoC_{i,\,max}(t)$ is the Maximum Safe State of Charge, representing the upper physical limit of the storage unit, $SoC_{i,\,min}(t)$ is the Minimum Safe State of Charge, representing the lower physical limit of the storage unit, and $\max(0,.)$ is the set of conditions that function as conditional logical gates.

2.5.6. Deep Reinforcement Learning Model Configuration

Table 2 outlines the architectural configuration and training hyperparameters employed in the DQN agent developed for optimising energy management within the hybrid battery-supercapacitor storage system. These parameters define the learning dynamics, control strategy, and convergence behaviour of the reinforcement learning framework, thereby facilitating adaptive energy scheduling under fluctuating renewable generation, load demand, and electricity pricing conditions. The chosen configuration was designed to balance computational efficiency, training stability, and optimal decision-making within the microgrid environment.

Table 2. *Configuring Hyperparameters for the Deep Q-Networks.*

SN	Parameter	Value
1.	Input State Dimension	7
2.	Hidden Layers	2
3.	Neurons per Hidden Layer	128
4.	Activation Function	ReLU
5.	Output Actions	25 (5 battery power levels x 5 supercapacitor power levels)
6.	Optimiser	Adam
7.	Learning Rate	0.001
8.	Discount Factor (γ)	0.99
9.	Loss Function	Smooth L1 (Huber)
10.	Gradient Clipping (max norm)	5.0
11.	Replay Memory Capacity	50,000
12.	Mini-Batch Size	128
13.	Initial Exploration Rate (ε)	1.0
14.	Final Exploration Rate (ε)	0.05
15.	Exploration Strategy	ε-Greedy
16.	Training Episodes	50
17.	Warm-up Steps (random actions)	2,000
18.	Training Start Threshold	1000 steps
19.	Target Network Update	Every 1000 steps

3. Results and Discussion

3.1. Exploratory Data Analysis

Figure 3 illustrates the time-series profiles of solar generation, wind generation, and load demand within the microgrid. Solar output follows a distinct diurnal cycle, with negligible production at night and peak generation around midday. Wind output displays stochastic fluctuations throughout the observation period, providing a complementary contribution to solar energy. Load demand varies dynamically, with several intervals exceeding the combined renewable supply, thereby necessitating support from hybrid energy storage systems or grid imports. On the contrary, periods of surplus renewable generation create opportunities for charging storage assets. The figure highlights the inherent mismatch between supply and demand and emphasises the need for coordinated energy management to ensure reliable and efficient microgrid operation.

Figure 3. *Renewable Supply and Load Demand.*

Also, Figure 4 depicts the temporal evolution of the battery SoC and SC energy within the HESS. The battery SoC (%) shows a gradual and steady upward trajectory across the simulation horizon, reflecting sustained charging and its role in long-term energy balancing with minimal short-term variability. In contrast, the SC energy (kWh) exhibits rapid, high-frequency fluctuations, highlighting its function in mitigating transient power imbalances and delivering short-term dynamic support. The complementary behaviour of these two storage components demonstrates effective coordination within the HESS, thereby enhancing system stability, operational reliability, and efficient energy management in the microgrid.

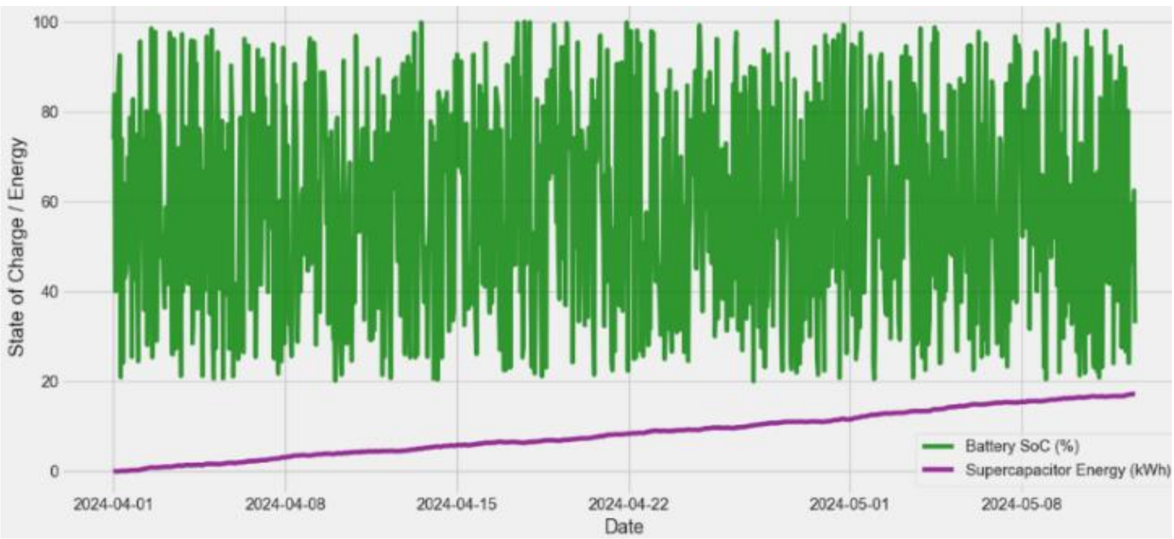

Figure 4. *Battery SoC and Supercapacitor Energy Profiles.*

Likewise, Figure 5 presents the daily average profiles of solar generation, wind generation, and microgrid load demand. Solar generation output follows a distinct diurnal, bell-shaped pattern, with negligible production during the early morning and evening, a steady increase after sunrise, and a pronounced midday peak of approximately 30 kW before declining towards sunset. In contrast, wind generation exhibits a stochastic and continuous profile characterised by moderate fluctuations and multiple local peaks, with no strict diurnal dependency and generally lower magnitude than solar output. Load demand remains relatively smooth and stable, with modest increases during morning and evening periods, occasionally surpassing renewable generation. The observed temporal mismatch between generation and demand underscores the necessity of hybrid energy storage systems, advanced energy management strategies, and potential grid support to ensure reliable, efficient, and balanced microgrid operation.

Figure 5. *Daily Solar - Wind - Load Trends.*

Similarly, Figure 6 presents the hourly averages derived from the source dataset. The solar-power series does not display the typical night-time minimum and midday maximum characteristic of measured photovoltaic generation. This atypical pattern may result from the nature of the dataset, aggregation effects, or limitations in its temporal resolution. Consequently, the profile is treated as a dataset-specific input rather than a physically validated photovoltaic generation curve.

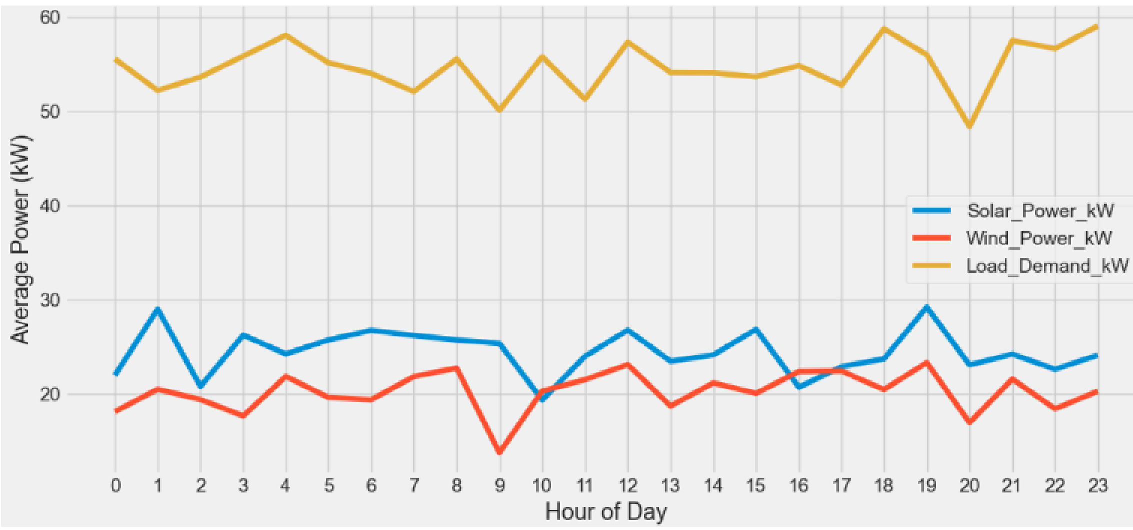

Figure 6. *Hourly Average Solar, Wind, and Load Demand Profile.*

3.2. Training Performance

The training results indicate that the DQN agent progressively improved its policy across the evaluated episodes. However, the learned policy should not be considered globally or economically optimal. Figure 7 illustrates the cumulative reward progression across 50 episodes. The reward begins at approximately -1200 in the early episodes, reflecting suboptimal decision-making during the initial learning phase. As training progresses, the reward increases steadily, despite fluctuations, reaching around -400 by the final episodes. This upward trajectory demonstrates the agent's improving performance, reduced penalties, and gradual convergence towards more effective energy management strategies. The figure highlights the reinforcement learning process, in which exploration and adaptation drive enhanced reliability, efficiency, and overall optimisation of microgrid operation.

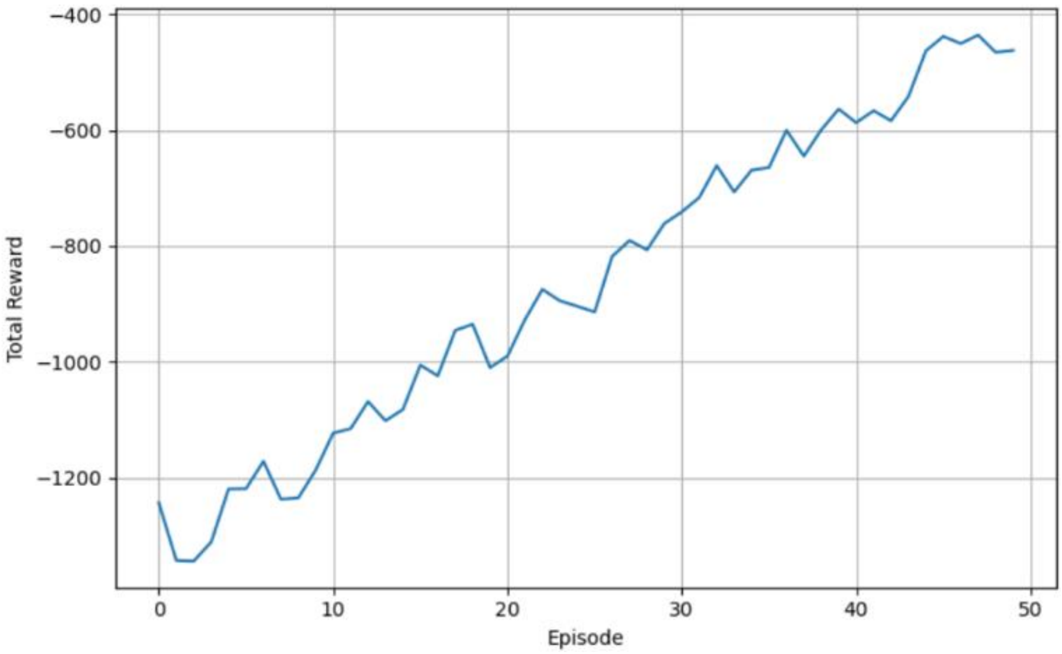

Figure 7*. Training Reward per Episode.*

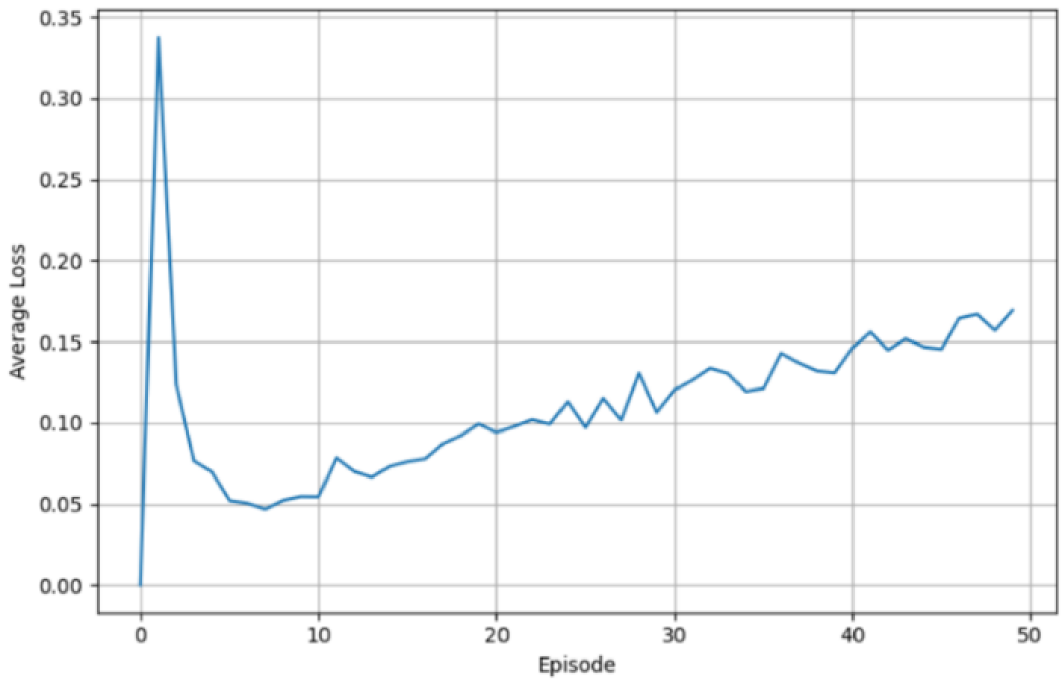

Figure 8*. Training Loss per Episode.*

The evolution of the average training loss is illustrated in Figure 8. The graph depicts the evolution of average loss across 50 episodes. At the outset, the loss rises sharply to approximately 0.33, reflecting instability during the initial learning phase. It then declines rapidly within the first 10 episodes, falling below 0.05, which indicates that the agent quickly adapts and enhances its decision-making efficiency. Following this sharp reduction, the loss gradually increases with fluctuations, reaching around 0.17 by episode 50. This subsequent rise suggests potential instability in training, possible overfitting, or continued exploration of alternative strategies, all of which introduce variability into the learning process.

The performance of the heuristic-based RBC is illustrated in Figure 9. The battery begins with a relatively high charge level of about 50 kWh and gradually decreases over time, with minor fluctuations, until it reaches approximately 10 kWh by the end of the simulation. This steady decline reflects its role in providing long-term energy balancing, discharging progressively to meet sustained demand. The supercapacitor starts near 0 kWh, rises quickly to around 5 kWh, and then remains almost constant throughout the simulation. This behaviour highlights its function as a fast-response storage device, primarily absorbing short-term fluctuations and stabilising transient power imbalances rather than supplying extended energy. Grid import remains consistently close to 0 kW across all timesteps, indicating that under the rule-based control strategy the microgrid operates with near-complete independence from the external grid, relying instead on renewable generation and the hybrid energy storage system to meet demand.

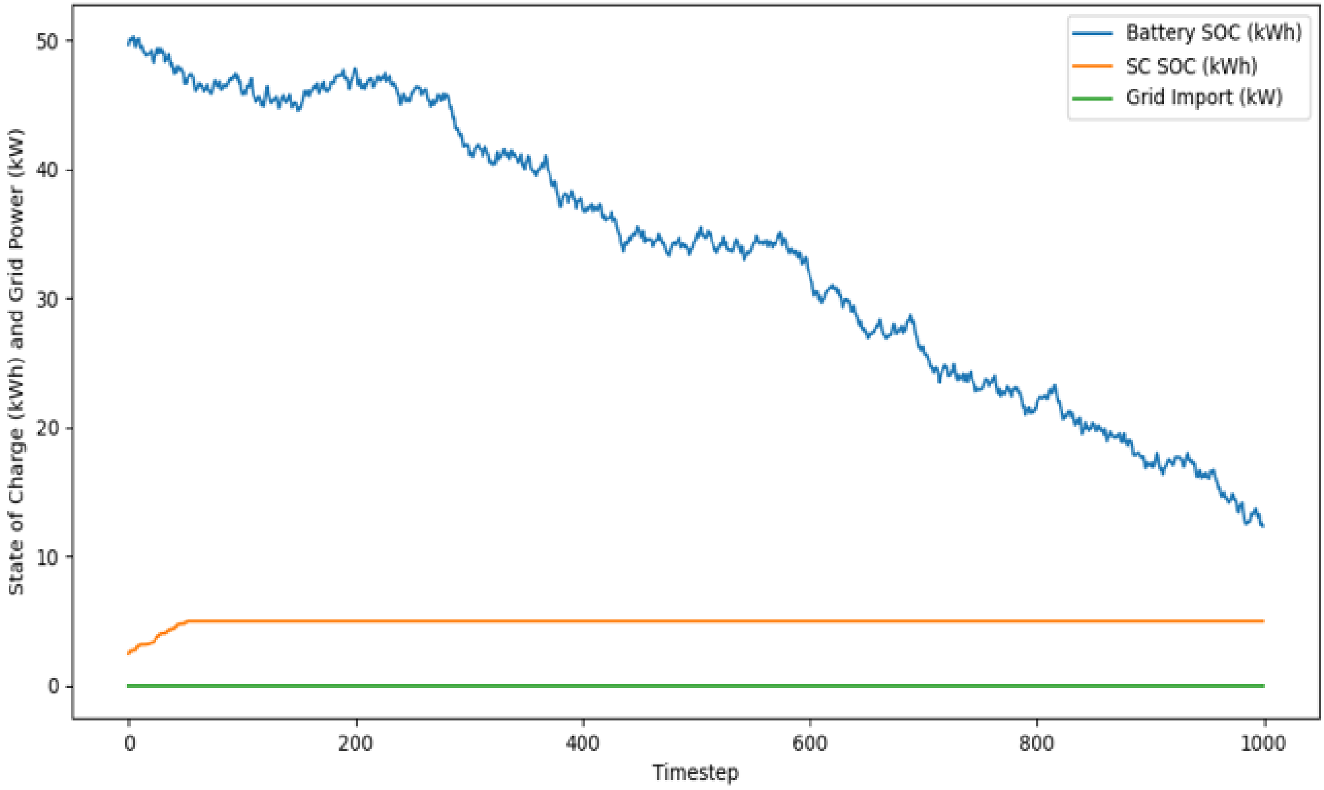

Figure 9. *Battery and SC SoC under RBC.*

3.3. Performance Comparison with Rule-Based Control

Figure 10 illustrates the grid import profiles obtained under the RBC and DQN strategies across 1,000 timesteps. The RBC profile remains consistently close to zero throughout the simulation, indicating that the controller predominantly relies on local renewable generation and the hybrid energy storage system to meet load demand, thereby minimising dependence on the external grid. In contrast, the DQN profile displays frequent fluctuations and intermittent spikes in grid import, reaching values of approximately 0.85 kW. This behaviour suggests that the DQN adopts a more dynamic operational strategy, selectively importing electricity from the grid when renewable generation and stored energy are insufficient to satisfy demand. The observed differences highlight that the DQN actively adapts to changing operating conditions, whereas the RBC follows a more deterministic and conservative control policy. Although the DQN results in greater grid interaction, this behaviour reflects its adaptive decision-making capability within the microgrid environment.

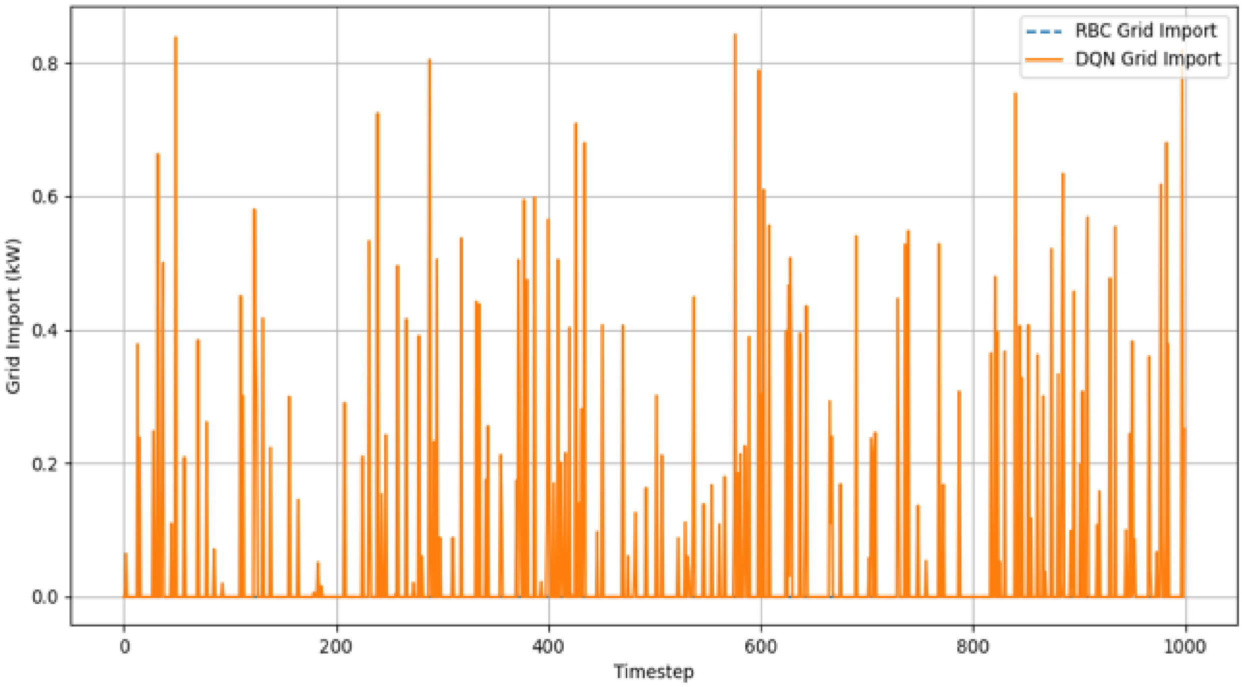

Figure 10. RBC and DQN Grid Imports.

Figure 11 compares the battery SoC trajectories under RBC and DQN control. The RBC maintains a relatively high charge level throughout the simulation, gradually declining from approximately 50 kWh to around 15 kWh by the end of the evaluation period. This pattern reflects the continuous use of the battery as a primary energy-balancing resource. In contrast, the DQN-controlled battery remains close to its lower operating limit, fluctuating only slightly between 0 and 2 kWh. This behaviour indicates that the DQN places less emphasis on battery utilisation, instead preserving capacity by limiting charge–discharge cycling. Such a strategy may mitigate degradation associated with frequent cycling, but it also represents a fundamentally different energy management philosophy compared with the RBC approach.

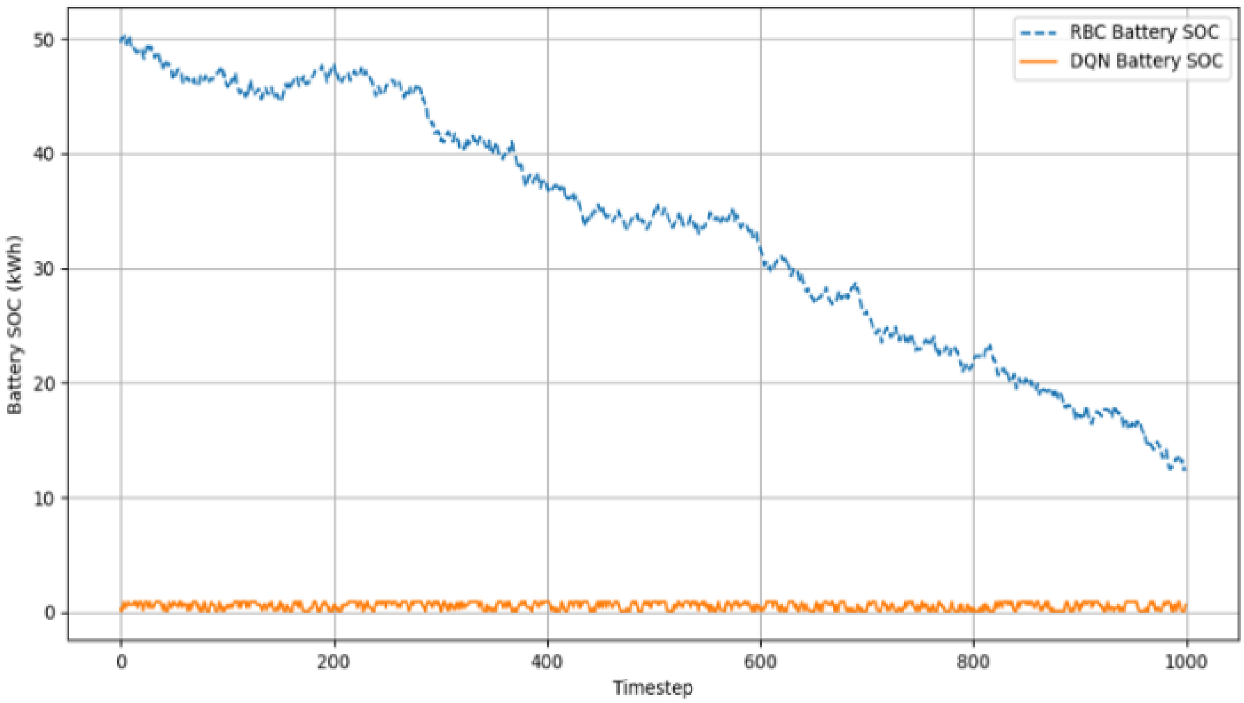

Figure 11. Battery SoC: RBC vs. DQN.

Figure 12 illustrates the supercapacitor SoC under both control strategies. The RBC profile remains almost constant throughout the simulation, indicating limited utilisation of the supercapacitor and reflecting a largely static control approach. By contrast, the DQN profile exhibits pronounced fluctuations with a generally upward trend over time. The repeated charging and discharging behaviour demonstrate that the DQN actively exploits the fast-response characteristics of the supercapacitor to accommodate short-term variations in generation and demand. This dynamic utilisation underscores the reinforcement learning agent's ability to interact with diverse energy storage resources and adapt its actions to prevailing system conditions.

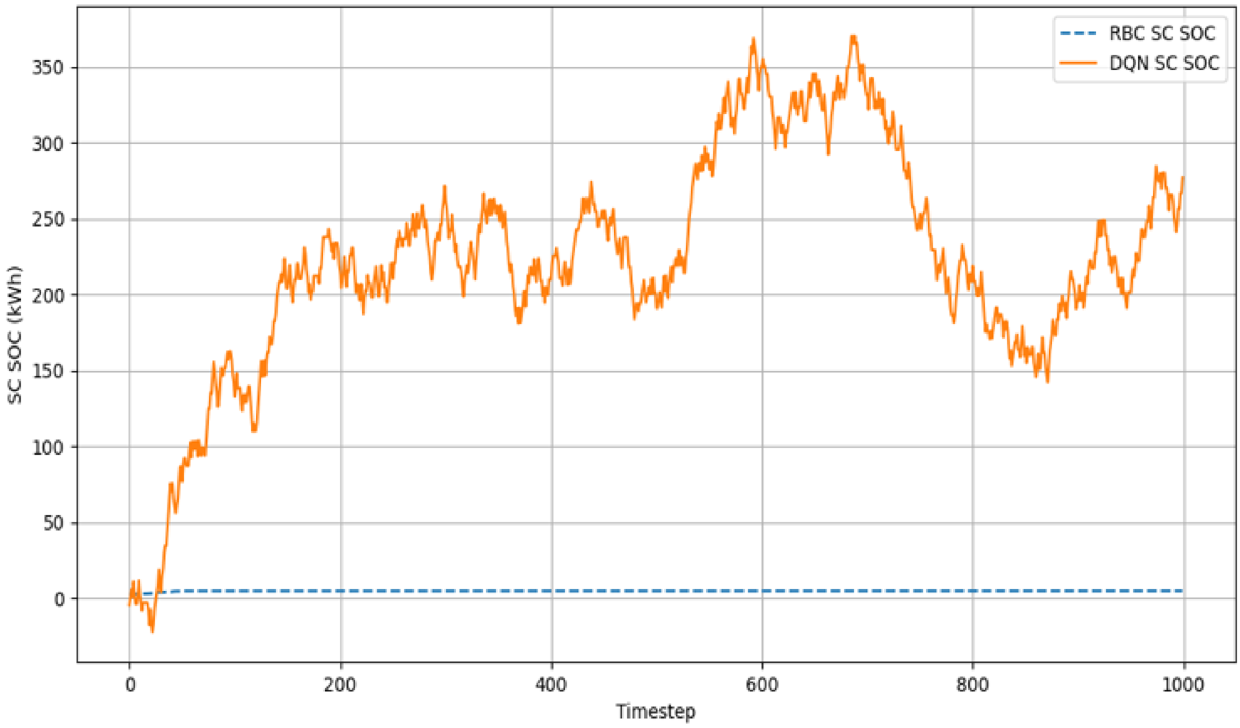

Figure 12. *Supercapacitor SoC: RBC vs. DQN.*

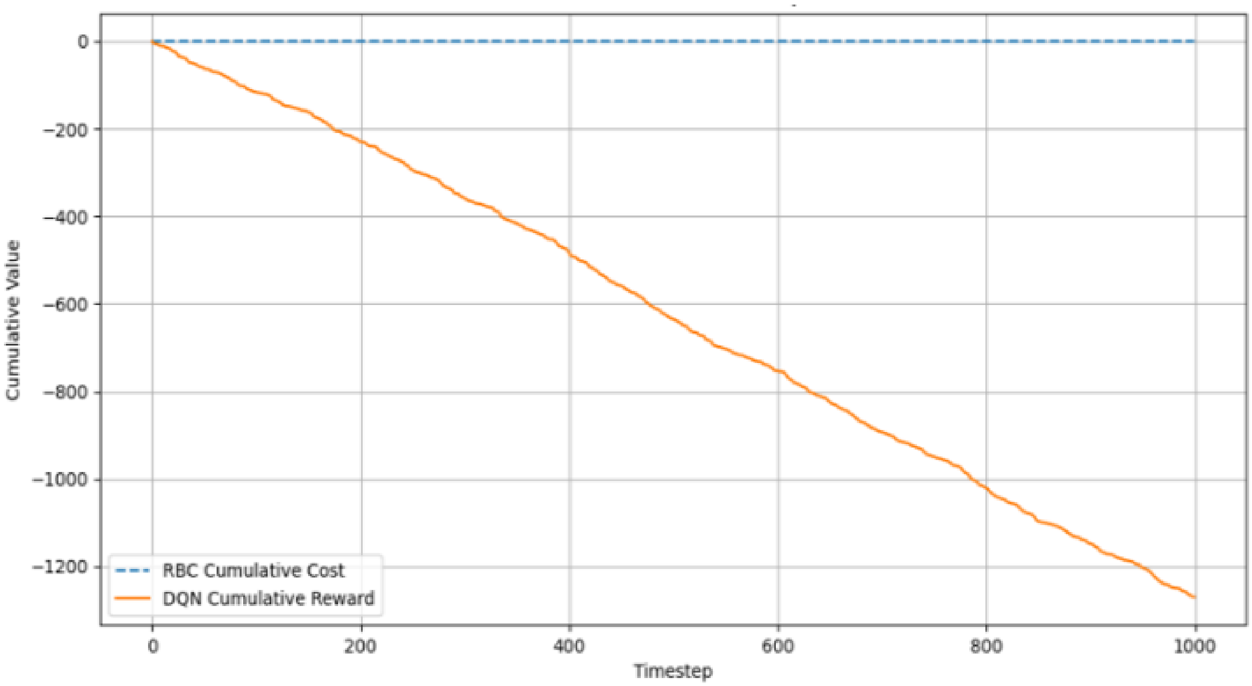

Figure 13. *Cumulative Cost and Reward: RBC vs. DQN.*

Figure 13 presents the cumulative performance profiles of the RBC and DQN strategies over 1,000 timesteps. The RBC curve remains relatively stable and close to zero throughout the simulation, indicating that cumulative penalties or costs remain largely unchanged during operation. By contrast, the DQN cumulative reward declines progressively, reaching approximately -1200 by the end of the evaluation period. This downward trend shows that the DQN accumulates negative rewards over time, suggesting that the operational decisions generated by the learned policy incur penalties associated with grid imports, storage utilisation, or other constraints embedded within the reward function. While the DQN demonstrates adaptive and dynamic control behaviour, the cumulative reward profile indicates that the learned policy has not yet converged to an economically optimal operating strategy under the evaluated conditions. These findings imply that further refinement of the reward structure, training process, and hyperparameter configuration may be necessary to enhance overall policy performance.

4. Conclusions

This study aimed to design and assess an intelligent energy management framework for hybrid battery-supercapacitor storage systems in renewable energy microgrids, employing a DRL approach through a DQN. The results demonstrated that the proposed agent progressively acquired effective charge-discharge strategies and reduced operational penalties. The DQN agent developed a more dynamic control policy than the deterministic RBC, although this did not consistently yield superior economic performance.

The findings revealed marked distinctions between the Deep Q-Network and heuristic control strategies. While the rule-based controller followed static, deterministic patterns with limited adaptability, the Deep Q-Network demonstrated dynamic utilisation of the supercapacitor, selective grid interactions, and adaptive scheduling in response to fluctuating operating conditions. These behaviours emphasise the capacity of reinforcement learning to capture complex temporal dynamics and to develop policies that enhance operational flexibility, efficiency, and reliability within microgrid environments.

This study contributes to the expanding body of knowledge on data-driven optimisation in power systems by incorporating artificial intelligence techniques into hybrid energy storage management. It advances both theory and practice by demonstrating the viability of reinforcement learning for adaptive energy scheduling, thereby facilitating the seamless integration of renewable resources into future smart energy systems. The practical implications include enhanced utilisation of complementary storage technologies, provided evidence of adaptive storage coordination, although reduced grid dependence and lower operating costs were not consistently demonstrated, while the theoretical contribution lies in demonstrating the feasibility of reinforcement learning as an alternative control framework that requires further reward and training optimisation.

While this study demonstrates the feasibility of DQN-based energy management for hybrid battery-supercapacitor systems, certain limitations remain. First, the proposed controller was evaluated solely against a conventional RBC, and therefore does not establish superiority over other contemporary DRL algorithms such as Double DQN, Dueling DQN, Proximal Policy Optimisation (PPO), or Soft Actor-Critic (SAC). In addition, the experiments were conducted within a simulated environment using a publicly available dataset, which may not fully capture the uncertainties and operational complexities of real-world microgrids. Furthermore, although the reward function effectively guided policy learning, further investigation into reward-weight optimisation and sensitivity analysis could enhance controller performance. Future research should therefore benchmark the proposed approach against multiple DRL algorithms, validate the framework using real-world microgrid data, refine reward-function design, and explore continuous-action reinforcement learning methods to achieve improved operational outcomes.

The demonstrated adaptive control behaviour and dynamic resource utilisation affirm the potential of deep reinforcement learning as a transformative approach to intelligent microgrid energy management. This study establishes a strong foundation for advancing AI-driven strategies that will play a pivotal role in enabling resilient, efficient, and sustainable energy systems in the future.

Abbreviations

AI	Artificial Intelligence
DQN	Deep Q - Network
DRL	Deep Reinforcement Learning
EMS	Energy Management System
HESS	Hybrid Energy Storage System
IoT	Internet of Things
PPO	Proximal Policy Optimisation
RL	Reinforcement Learning
SAC	Soft Actor-Critic
SoC	State of Charge

Author Contributions

Daniel Kumi Owusu: Conceptualization, Formal Analysis, Investigation, Methodology, Visualization, Writing – review & editing

Data Availability Statement

The datasets employed for this research can be found at: https://www.kaggle.com/datasets/ziya07/hybrid-energy-storage-dataset.

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] Mulaudzi, S. T., Rasimphi, T., Tinarwo, D. Assessing the Contribution of Microgrids to Sustainable Agriculture and Community Development at the Masia Development Centre, Clean Energy Technologies. 2026, 2(1), 1–16. https://doi.org/10.54963/cet.v2i1.1945

[2] Atef, M., Alahakoon, S., Mumtahina, U., Wolfs, P., Khatib, T., Uddin, M. A Review of Microgrid Energy Management Systems: Methods, Challenges, and Future Directions, International Journal of Ambient Energy. 2026, 47(1), 1–28. https://doi.org/10.1080/01430750.2025.2595117

[3] Sargentis, G., Papadodimas, N., Benekos, I., Katsoulakos, N. M., Dimitriadis, P., Tepetidis, N., Koutsoyiannis, D. Stochastic Assessment of Renewable Energy Reliability: A Case Study of North Euboea, Greece, Journal of Energy and Power Technology. 2026, 8(2), 1–13. https://doi.org/10.21926/jept.2602008

[4] Dawn, S., Vadlamudi, B., Vital, M. L. N., Das, S. S., Rao, K. D., Al Mansur, A., Ustun, T. S. Enhancing Grid Stability and Sustainability in Electrical Markets: A Review on the Synergy of Renewable Energy and Electric Vehicles, Energy Exploration and Exploitation. 2026, 44(2), 1021–1066. https://doi.org/10.1177/01445987251383177

[5] Dawn, S., Manideep, S., Rekha, S. S., Rao, C. R., Rao, K. D., Al Mansur, A., Ustun, T. S. Advancing Renewable Energy Integration in Deregulated Markets: The Role of Energy Storage, EVs, and Policy Frameworks, Energy Exploration and Exploitation. 2025, 43(6), 2745–2797. https://doi.org/10.1177/01445987251348052

[6] Gowtham, M., Suresh, S. Supercapacitor Integration in Hybrid Energy Storage Systems for Electric Vehicles: A Comprehensive Review. In Proceedings of the 2025 International Conference on Intelligent Computing, Information and Control Systems. IEEE, Lalitpur, Nepal, 2025; pp. 170–176. https://doi.org/10.1109/ICOIICS67115.2025.11390385

[7] Ghoudelbourk, S., Yousfi, R., Benbouhenni, H. Energy Management of an Autonomous Hybrid Renewable Power System. In Proceedings of the 2025 IEEE 13th International Conference on Smart Energy Grid Engineering. IEEE, Oshawa, Canada, 2025; pp. 190–194. https://doi.org/10.1109/SEGE65970.2025.11203682

[8] Srivastava, A., Jha, S. K., Gupta, D., Singh, S. A Critical Review of Microgrid Technologies: Storage and Control Systems with Effective Optimisation Techniques for Modeling Wind, Solar, Biomass, and Battery Energy Storage Systems, Cureus Journals. 2026, 3(1), 1–23. https://doi.org/10.7759/s44388-025-00041-y

[9] Eslami, F., Gangineni, M., Ebrahimi, A., Rathnayake, M., Patel, M., Lavrova, O. A Review on Protection and Cybersecurity in Hybrid AC/DC Microgrids: Conventional Challenges and AI/ML Approaches, Energies. 2026, 19(3), 1–38. https://doi.org/10.3390/en19030744

[10] Guo, Z., Grano, E., de Carvalho Pinheiro, H., Carello, M. A Comparative Evaluation of Rule-Based Strategies, ECMSs, and MPC Strategies for Fuel Cell Hybrid LCV Energy Management, World Electric Vehicle Journal. 2026, 17(3), 1–25. https://doi.org/10.3390/wevj17030163

[11] Kavaliauskas, Ž., Milieška, M., Blažiūnas, G., Gecevičius, G., Zhairabany, H. Optimisation of Hybrid Energy System Control Using MPC and MILP, Applied Sciences. 2026, 16(8), 1–22. https://doi.org/10.3390/app16083690

[12] Kırant-Mitić, T., Voss, K. A Rule-Based Predictive Control Framework for Market and Grid-Oriented Operation in Thermally Activated Buildings, Journal of Building Performance Simulation. 2026, 19(2), 209–238. https://doi.org/10.1080/19401493.2025.2565652

[13] Ioannou, I., Javaid, S., Tan, Y., Vassiliou, V. A Hierarchical Predictive-Adaptive Control Framework for State-of-Charge Balancing in Mini-Grids Using Deep Reinforcement Learning, Electronics. 2026, 15(1), 1–47. https://doi.org/10.3390/electronics15010061

[14] Wang, X., Zhang, H., Liu, H., Lewis, F. L. Control-Oriented Reinforcement Learning: A Survey of Recent Progress and Applications, International Journal of Robust and Nonlinear Control. 2026, 36(2), 694–718. https://doi.org/10.1002/rnc.70152

[15] Hu, P., Ukil, A. Enhanced Multi-Agent Reinforcement Learning for Power Quality Enhancement and False Data Injection Defense in Multi-Microgrid Systems. In Proceedings of the IECON 51st Annual Conference of the IEEE Industrial Electronics Society. IEEE, Madrid, Spain, 2025; pp. 1–6. https://doi.org/10.1109/IECON58223.2025.11221977

[16] Ketjoy, N., Muna, Y. B., Kaewpanha, M., Chamsa-ard, W., Suriwong, T., Termritthikun, C. Grid-Scale Battery Energy Storage and AI-Driven Intelligent Optimization for Techno-Economic and Environmental Benefits: A Systematic Review, Batteries. 2026, 12(1), 1–39. https://doi.org/10.3390/batteries12010031

[17] Banka, S., Kumar, D. V. A. Hierarchical Data-Driven and PSO-Based Energy Management of Hybrid Energy Storage Systems in DC Microgrids, Automation. 2026, 7(2), 1–21. https://doi.org/10.3390/automation7020050

[18] Li, Z., Long, W., Tian, H. Current Applications and Future Prospects of Deep Reinforcement Learning in Energy Management for Hybrid Power Systems, Energies. 2026, 19(9), 1–50. https://doi.org/10.3390/en19092216

[19] Asadi, Y., Eskandari, M., Mansouri, M. Revolutionising Battery Energy Storage Systems Energy Management: Dynamic-Aware Solutions with Integrated Quantum Particle Swarm Optimisation and Deep Reinforcement Learning, IET Energy Systems Integration. 2026, 8(1), 1–15. https://doi.org/10.1049/esi2.70038

[20] Khan, W., Renhai, F., Aziz, A., Yousaf, M. Z., Cai, Z., Iqbal, M. U., Geremew, M. S. Deep Reinforcement Learning-Based Energy Management for Design and Control of Off-Grid Renewable Microgrids With Dual-Battery Storage. Energy Exploration and Exploitation. 2026, 44(2), 821–869. https://doi.org/10.1177/01445987251387368

[21] Kavaliauskas, Ž., Milieška, M., Blažiūnas, G., Gecevičius, G., Zhairabany, H. Optimisation of Control for a Hybrid Renewable Energy System with Energy Storage Using Deep Reinforcement Learning Methods, Sustainability. 2026, 18(11), 1–43. https://doi.org/10.3390/su18115443

[22] Zhou, Y. Advances in Smart Grids and Microgrids: Distributed Generation and Energy Storage Systems, Processes. 2026, 14(9), 1–9. https://doi.org/10.3390/pr14091460

[23] Tasmant, H., Bossoufi, B., Alaoui, C., Siano, P. A Review of Machine Learning and IoT-Based Energy Management Systems for AC Microgrids, Computers and Electrical Engineering. 2025, 127, 1–23. https://doi.org/10.1016/j.compeleceng.2025.110563

[24] Behera, S., Khare, M. A Survey on Optimisation Strategies for Energy Storage and Microgrid Energy Management in Cyber-Physical Systems Based Microgrids, Engineering Research Express. 2025, 7(4), 1–30. https://doi.org/10.1088/2631-8695/ae28b7

[25] Ojo, K. E., Saha, A. K., Srivastava, V. M. Microgrids' Control Strategies and Real-Time Monitoring Systems: A Comprehensive Review, Energies. 2025, 18(13), 1–34. https://doi.org/10.3390/en18133576

[26] Ojo, K. E., Saha, A. K., Srivastava, V. M. Review of Advances in Renewable Energy-Based Microgrid Systems: Control Strategies, Emerging Trends, and Future Possibilities, Energies. 2025, 18(14), 1–26. https://doi.org/10.3390/en18143704

[27] Arunkumar, R., Rahman, A. F., Nathan, V. B., Krishnamoorthy, K., Gangatharan, N. Reinforcement Learning-Assisted Smart Charging Algorithm to Extend Lifespan of Supercapacitor-Based Energy Storage Systems. In Proceedings of the International Conference on Recent Innovation in Science Engineering and Technology. IEEE, Chennai, India, 2025; pp. 1–8. https://doi.org/10.1109/ICRISET64803.2025.11253581

[28] Almihat, M. G. M., Munda, J. L. Review on Recent Control System Strategies in Microgrid, Edelweiss Applied Science and Technology. 2024, 8(6), 5089–5111. https://doi.org/10.55214/25768484.v8i6.3116

[29] Gajjar, R. R., Giri, N. C., Patel, U., Gajjar, R. C., Dave, D., Aly, A. M. State of Charge Control Based Improved Hybrid Energy Storage System for DC Microgrid, Bulletin of Electrical Engineering and Informatics. 2024, 13(6), 3779–3788. https://doi.org/10.11591/eei.v13i6.7759

[30] Negi, G. S., Singh, R., Gehlot, A., Malik, P., Sharma, R., Obaid, A. J., Kumar, S. Microgrids 4.0: Digitalisation of Microgrid with Iot and Recent Technology Interventions, IET Communications. 2025, 19(1), 1–12. https://doi.org/10.1049/cmu2.12864

Biography

Daniel Kumi Owusu holds a PhD in Electrical and Electronic Engineering from the University of Mines and Technology, Tarkwa, Ghana where he also completed his MSc in Electrical and Electronic Engineering in 2016. He earned a BEng in Electrical and Electronic Engineering from the Accra Institute of Technology in 2012 and a Higher National Diploma in Electrical and Electronic Engineering from Takoradi Technical University in 2006. He is currently a Lecturer in the Department of Electrical Engineering at Takoradi Technical University and serves as an A1 Assessor and Lead Trainer in Electrical and Instrumentation Engineering at the Jubilee Technical Training Centre of Takoradi Technical University. Dr Owusu is a member of the Ghana Institution of Engineering (GhIE) and the International Association of Engineers (IAENG). His research interests encompass artificial intelligence applications in engineering, electrical machines, instrumentation, and control systems engineering.

Research Field

Daniel Kumi Owusu: application of artificial intelligence in engineering, electrical machines, instrumentation, control systems engineering, power systems analysis