iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: VII Month of publication: July 2025

DOI: https://doi.org/10.22214/ijraset.2025.73295

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VII July 2025- Available at www.ijraset.com

Comparative Analysis of Q-Learning and Rule-Based Approaches for Battery Energy Management in Electricity Markets

Harshi Shrivastava, Seema Pal
Electrical Engineering Department, Jabalpur Engineering College, Jabalpur

Abstract: Modern power grids with the increasing volatility of electricity markets and the growing deployment of battery energy storage systems (BESS) require intelligent energy management strategies to maximize economic benefits while preserving system performance. A Reinforcement Learning (RL) based energy management framework utilizing the Q-learning is presented to optimize BESS operation for peak demand reduction. By formulating the problem as a Markov Decision Process, the Q-learning agent learns an optimal charging and discharging policy through interaction with fluctuating market prices and operational constraints. The proposed approach allows the Battery Energy Storage System (BESS) to carry out energy arbitrage to maximize economic benefits and maintaining demand-supply balance. Simulation results demonstrate that the Q-learning framework achieves superior profitability and battery energy efficiency compared to conventional rule-based optimization method. The results highlight the potential of reinforcement learning techniques, particularly Q-learning, in enabling adaptive and autonomous energy management in complex and uncertain environments.
***Keywords**: Reinforcement Learning, Demand Side Management, Dynamic Electricity Market, Energy Efficiency, Battery Energy Storage System*

I. INTRODUCTION

The expanding renewable energy sources along-with increasing variability in electricity demand and uncertainty energy such as solar and wind due to weather changes, cloud cover and wind speed variations , make power generation non-linear as well as difficult to predict[1]. This results in a huge stress on modern power grids, which results in frequency deviations, instabilities in voltage, and transmission network congestion, causing functional challenges and financial losses as well. There are demand-side uncertainties of irregular consumer behaviour, changing load profiles and also volatility of real-time electricity price in dynamic pricing markets, adding immense complexity and problem in energy management [2]. Battery energy management systems (BEMS) are an excellent solution to ease grid stress through enabling dynamic control over the storage of energy and energy dispatch intelligently by BESS. Traditional battery energy storage control methods that follow rule-based approaches, heuristic algorithms and the scheduling by predefined methodology depend on static assumptions and are not so adaptive to dynamic changes. They are suitable only for ideal and predictable conditions. As a result they are incapable of fully exploiting the economic as well as operational benefits by intelligently performing battery [3].

To combat these limitations different machine learning techniques are being used among which Reinforcement Learning (RL) a very effective strategy is used for energy storage systems designed with highly adaptive control for. RL constitutes a powerful framework for sequential decision-making in complex, uncertain environments. Supervised learning requires labelled data whereas RL does not. RL learns optimal strategies by interacting with the environment and uses feedback in the form of rewards or penalties[4].This work targets the development and implementation of a Q-learning-based strategy that optimally schedules the charging and discharging of battery subjected to fluctuations in renewable generation of energy and real-time pricing of energy. The aim of this research is to outperform static methodology for better energy utilization and enhancement of economic returns in smart grid environments through intelligent data-driven control[5].

The structure of this paper is as follows: Articles related to BEMS and Q learning is reviewed in Section 2. Section 3 shows the methodology of the proposed work. Section 4 represents results and discussion. In Section 5, conclusions and future scopes are provided.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VII July 2025- Available at www.ijraset.com

II. LITERATURE REVIEW

A scheduling framework which is highly adaptable for BESS is presented[6]. They use classical Q-learning approach of a ε-greedy exploration considering electricity prices and demonstrating price adjustment volatility of Q-learning. They have considered battery degradation effects. Economic performance is as compared to heuristic approaches.

Energy Management of a PV micro grid is enhancing the algorithm of Q learning.[7] The model used Time-of-Use (ToU) tariff and did charging and discharging for 24 hour schedule and reached faster convergence. They have also achieved significant reduction in cost and accounted for solar generation and showed the demand's period pattern changes daily.

Markov Decision Process (MDP) brought expansion classical battery energy management [8].They have used solar generation and have improved the charging and discharging actions based on the availability of renewable energy and consumption and brought improvement in BESS's smooth operation.

The batch techniques of reinforcement learning mainly the Fitted Q-iteration(FQI) in microgrid environment and have used historical data for training the offline policies and has optimised BSS and brought 19% improvement in performance over the rule based models conventional using Q-learning[8].

Proposed a framework using reinforcement learning for the electricity markets that are real time by adopting few learning for better battery energy storage system designing The RL agent to better learn the adaptation to market and learning to improve profit and effectively manage battery's SoC [9].

[10] in 2018 presented a work for smart energy buildings using reinforcement learning and modelling by MDP using real world simulation and out-performing the conventional approaches for reduction in the energy cost.

[11] reviewed the techniques of battery modelling and balanced the real time application of battery energy management system by using data driven model and has worked on crucial parameters like state of charge state of health and discussed the state estimation methods in detail.

In [12] it is suggested that a system that well balances the B.E.S.S and monitors and mitigate imbalance of cell the adjustments related to charging and discharging have been taken in account for maintaining the for maintaining for maintaining cell voltages to be uniform across better utilisation of energy and comparison with static strategies is done.

TABLE I
COMPARISON OF RELATED WORKS WITH THIS STUDY

Features	**Research Papers by Authors**						
	Liu et al. (2024)	Arwa & Folly (2021)	Chang et al. (2021)	De Somer et al. (2017)	H. Wang & B. Zhang (2017)	S. Kim & H. Lim (2018)	This work
Algorithm Type	Classical Q-learning with ε-greedy	Enhanced Q-learning for faster convergence	Q-learning used for time-aware cyclic MDP with	FQI used alongwith Batch RL `	Classical Q-learning with price prediction	Standard Q-learning with MDP	Hybrid Q-learning with adaptive ε-decay, prioritized experience replay, and dynamic reward shaping
Adaptation to Price Volatility	Real-time price consideration	ToU tariffs adaptation	Focused on periodic pricing patterns	Used historical prices for offline training	Real-time price learning	Grid import/export cost awareness	Enhanced: Handles dynamic pricing signals with predictive modelling (sinusoidal + random noise)
Exploration Strategy	ε-greedy	Enhanced Q-update	Cyclic time-aware updates	No online exploration (offline batch learning)	ε-greedy	Standard exploration	Improved: Adaptive ε-decay with fine-tuned exploration/exploitation balance for faster convergence
Learning Convergence Speed	Moderate	Faster than vanilla Q-learning	Slower due to cyclic complexity	Offline learning only	Slow convergence (historical data dependent)	Moderate	Superior: Prioritized experience replay + adaptive decay = faster, more stable convergence

Scalability to Multiple Scenarios	Designed for BESS scheduling	Focused on PV micro-grids	Residential BESS	Microgrid only	Real-time market arbitrage only	Smart building energy management	Highly generalized: Supports micro grids, smart buildings, V2G (vehicle-to-grid), and scalable environment
Economic Performance	Outperformed rule-based scheduling	Reduced operational costs	Aligned with dynamic pricing	19% better than model-based optimization	Higher profits than rule-based methods	Lowered energy costs	Achieved higher average profits & efficiency Vs static strategies (shown in plots)
Simulation Visualization	Not reported	No visualization	No visualization	Not reported	No visual analysis	No visual analysis	Provides clear plots for profit & battery efficiency comparisons (supports analysis & debugging)
Key Innovation	Q-learning + degradation-aware reward	Modified Q-update for speed	Cyclic MDP	Batch FQI	Q-learning for market arbitrage	RL for smart building management	Hybrid reward shaping, adaptive exploration & efficient convergence techniques

Collectively, these studies underscore the versatility of Q-learning in addressing the challenges posed by dynamic electricity markets. Despite their contributions, these approaches exhibit certain research gaps, such as limited adaptability to rapidly changing market conditions, insufficient handling of multi-objective optimization. Rather than relying on classical or enhanced Q-learning, our method employs an adaptive exploration strategy with epsilon decay. The inclusion of clear visualization of profits and efficiency further distinguishes our framework, offering better interpretability and practical insights for real-world deployment, enabling faster convergence and more effective exploitation of profitable actions. The proposed Q-learning framework advances beyond existing approaches by integrating multiple enhancements for improved economic and operational performance. This paper addresses these limitations through an improved framework for battery energy management, as detailed in Table 2.1 which compares prior work with the proposed method.

III.PROPOSED METHODOLOGY

In the proposed methodology, a simulation for a 24-hour cycle is done on Google Colab using Python. The libraries imported are numpy, pandas, matplotlib to find the profit, efficiency values and plot graphs. We have used the Machine Learning (ML) technique of Reinforcement Learning (RL) and further Q-learning method In this work, we have implemented a Q-learning-based approach as an alternative to the traditional rule-based control strategy. The proposed method is designed to manage electricity demand and pricing data effectively. Specifically, electricity demand is primarily satisfied from the grid up to a predefined threshold. Beyond this limit, additional demand is met using a Battery Energy Storage System (BESS).

In conventional rule-based systems, the control logic purchases electricity from the grid when prices are low and exports power back to the grid when prices are high, with possible idle states in between. By contrast, in the Q-learning framework, charging and discharging decisions are governed by the ε-greedy exploration strategy, allowing the agent to balance exploitation of known policies and exploration of new actions. The system's dynamics are modelled as a Markov Decision Process (MDP), where the next state depends only on the current state and action. This reinforcement learning formulation enables adaptive energy management without relying on static, pre-defined rules.

Among storage technologies, Li-ion batteries are preferred for effective management [13] [14],especially through Battery Management Systems (BMS).A key role of BMS is of cell balancing to maintain at level the voltage or the state of charge (SoC) of individual cells in[16]. This prevents overcharging or over-discharging, which can reduce performance, shorten battery life, and create safety hazards. Balancing methods include passive and active techniques[15]. In battery management, most ML applications focus on prediction—such as estimating (SoC) [15], State of Health (SoH) [16], load as well as battery safety,. Predictive ML has shown great potential in enhancing battery performance, reliability, and safety [18].Prescriptive ML often uses reinforcement learning (RL)[24].RL learns through trial and error by interacting with the environment, aiming to maximize a reward over time without explicit instructions. Though RL has existed for decades, recent breakthroughs have sparked interest in its applications, including battery management systems. Examples include planning battery use in hybrid systems [17]and optimal battery balancing [12].This study proposes a novel use of reinforcement learning for passive cell balancing, highlighting ML's for better utilization of the potential of BESS.

A. Reinforcement Learning (RL)

Reinforcement Learning (RL) is a type of machine learning involving an agent to interact with an environment to learn an optimal policy by maximizing cumulative rewards. In this case, the battery storage system is the agent, and the environment is defined by time, price, load, and battery status[18]. RL offers:

1) Model-free learning: No prior knowledge of system dynamics required.
2) Online adaptability: Learns directly from experiences.
3) Scalability: Can handle high-dimensional state-action spaces[19].

The Q-learning method of RL aims to learn quality (Q-value) of taking a certain action in a given state. It is particularly suitable for discrete and finite state-action spaces like in this study[5].It is a model-free and off-policy algorithm stands out for its simplicity and effectiveness in a wide range of control tasks. Q-learning is well-suited to problems where the environment is partially unknown or highly dynamic. It enables the agent (in this case, the BESS controller) to learn a policy that maps system states—such as electricity prices, state of charge (SOC), time of day, and predicted demand or generation—to optimal actions, such as when and how much to charge or discharge[20]. Over time, the agent refines its decisions to maximize long-term rewards, which in the energy context can be defined in terms of economic gains, grid support, or system reliability[21].

The proposed methodology is built around Q-learning, where the environment consists of:

States (s): A state represents the condition or situation the agent finds itself in at any given moment within the environment. It serves basis for decision-making, offering the agent relevant information to choose an action.

Actions (a): An action is one of the possible decisions or movements the agent can makes depending on the state the agent is in.

Rewards (r): A reward is a numerical signal sent by the environment after the agent performs an action. It acts as immediate feedback, showing how beneficial or harmful the action was in the given state. The agent uses these signals to learn which actions are preferable in the long run.

In our research work, Markov Decision Process is used to select the following:

a) *States:* Defined by hour of the day, electricity price bin, day type (weekday/weekend), battery state of charge (SoC), and load demand bin.
b) *Actions*: Charge, Discharge, or Idle.
c) *Reward*: Based on the profit or loss from charging/discharging energy.

The agent learns when its most profitable to charge (buy energy) or discharge (sell energy), to maximize total profit over 24 hours. Recommended font sizes are shown in Table 1.

B. Flow Description

The Environment sends the current state or information to the Agent. The Agent decides on an Action, which in this diagram is the "Charge" process. After charging, the system moves to Discharge. The discharge process affects the Environment again, closing the loop. The Environment sends a Reward back to the agent based on the outcome of the discharge. This reward helps the agent learn or adapt future actions.

Key components:

1) State Space: A 5-dimensional space combining temporal, economic, and technical features.
2) Action Space: A discrete set of 3 actions.
3) Training: Over 7500 episodes, the RL agent learns an optimal Q-table by simulating different scenarios.
4) Testing: The learned policy is tested over a 24-hour period and compared against a conventional fixed-threshold strategy.

The block diagram of methodology is shown in Fig. 1

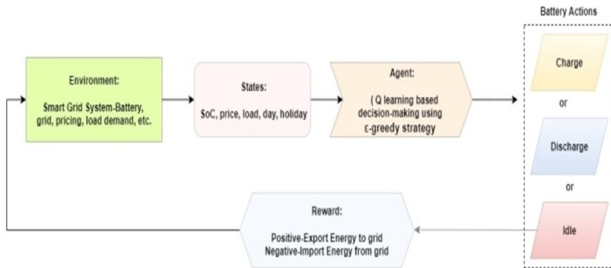


Fig. 1 Block diagram of Methodology

The RL agent chooses actions using the ε-greedy strategy to balance exploration and exploitation[22]. Rewards are calculated based on energy traded and corresponding electricity prices[23]. The agent updates its Q-values using the standard Q-learning update rule:

$$Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma \max Q(s',a') - Q(s,a)] \quad (1)$$

Where,

$Q(s,a)$: The present estimate of the total expected reward for action a taken in state s. Stored in the Q-table.α: Learning rate ($0 < \alpha \leq 1$). It determines how much the new information overrules the previous value. Low α indicates slow learning and high α shows fast but possibility of unstable learning. r: Immediate reward attained after action a is taken in state s. In this case, the profit (positive or negative) depends on energy price and action (charge/discharge).γ (gamma): Discount factor ($0 \leq \gamma \leq 1$). It regulates the future reward amount taken into account. Higher γ makes agent care more about long-term gains. s′: Next state after action a is taken in state s.$\max Q(s',a')$: The best possible Q-value from the next state s′, assuming agent acts optimally in the future.$Q(s,a) \leftarrow$: The updated Q-value is a weighted average of the old Q-value and the new estimate[23].

C. Flowchart of Methodology

The process flow is shown in Fig. 2

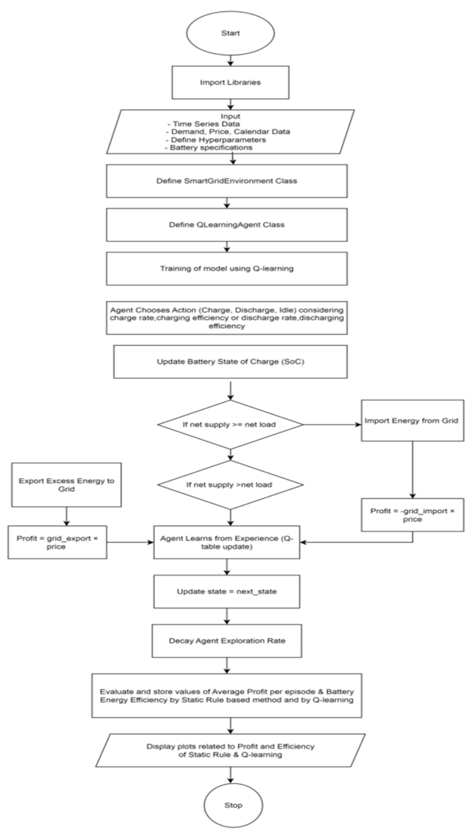


Fig. 2 Flowchart of Methodology

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VII July 2025- Available at www.ijraset.com

IV. RESULTS AND DISCUSSION

To evaluate the performance of the proposed Reinforcement Learning (RL)-based battery energy management system, we conducted a simulation over a 24-hour operation cycle on Google Colab in Python. The RL agent was trained using a Q-learning approach, taking into account dynamic energy prices, load variations, and the battery's state of charge (SoC) . A baseline strategy, following a conventional heuristic without learning, was also implemented for comparative analysis.

The key performance metrics analyzed were monetary profit from energy arbitrage, energy charged/discharged, and overall energy efficiency.

A. Profit Analysis

The goal of our model is to evaluate relative energy management performance considering realistic price fluctuations. We have generated synthetic electricity demand and price to follow periodic patterns and also considered randomness and volatility using the equation [24]:

$$P_t = 10 + 8 \cdot \sin(4\pi t/T) + \sigma \cdot (U_t - 0.5) \qquad (2)$$

Where-

The base price offset is taken as Rs.10 which provides an appropriate scaling for the simulated environment.

The sinusoidal term (amplitude 8, frequency 2/T) is the deterministic sine component.

The noise term, σ (Ut−0.5), is uniform and not Gaussian.

The RL-based strategy achieved a total profit of Rs. 2771.9 Compared to Rs.2612.6 generated by the baseline method. This corresponds to a 6.09% improvement in profit over the conventional approach. This significant increase demonstrates the effectiveness of the RL algorithm in learning optimal charge/discharge decisions in response to time-varying electricity prices and load demands. Profit analysis bar chart is shown below:

Average Profit (Q-learning): 2771.9

Average Profit (Static Strategy): 2612.6

$$\text{Percentage Increase in Profit} = \frac{\text{Avg Profit (Q)} - \text{Average Profit (Static)}}{\text{Average Profit (Static)}} \times 100 \qquad (3)$$

$$\text{Percentage Increase} = \frac{2771.9 - 2612.6}{2612.6} \times 100 = \left(\frac{159.3}{2612.6}\right) \times 100 \approx 6.09\%.$$

Profit Improvement (Q-learning Vs Static): 6.33%

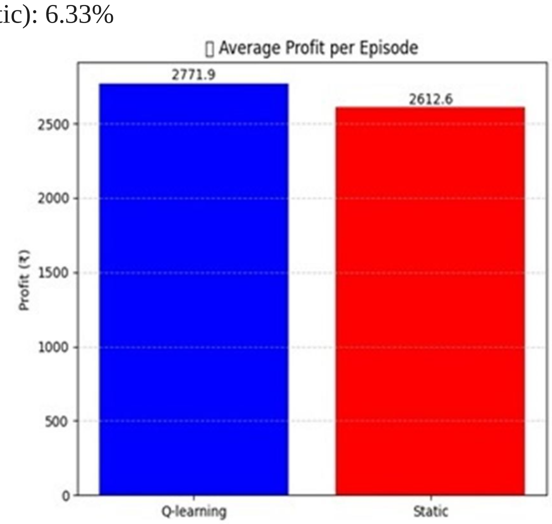


Fig. 3 Bar Graph of Comparison between Average profit per Episode of Q-learning & Rule-based Method

These are found by the RL policy. This high efficiency indicates that the RL agent not only maximizes profit but also utilizes the battery in a relatively loss-minimized manner, maintaining operational integrity and sustainability.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VII July 2025- Available at www.ijraset.com

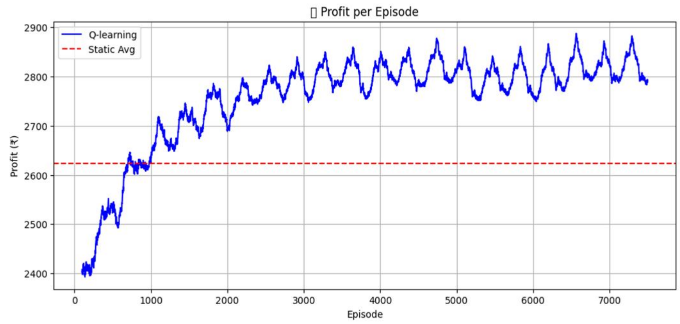


Fig .4 Comparison of Q-learning and Static Rule methods through Profit Vs Episode graph.

B. Energy Utilization

The profit-per-episode graph compares how much net profit the battery scheduling strategy achieves each day. Initially, the Q-learning curve shows significant fluctuations because the agent is still exploring different charge/discharge actions. With time, the smoothing of the curve is obtained and by the agent learning and following an optimal policy. Later episodes are observed to achieve greater daily profits consistently as compared to the static strategy, depicted by a flat horizontal line which shows the average performance. It is the indication that the Q-learning agent have becomes adaptable to dynamic conditions such as fluctuations of electricity price and changing load demands, outperforming the simple static rule.

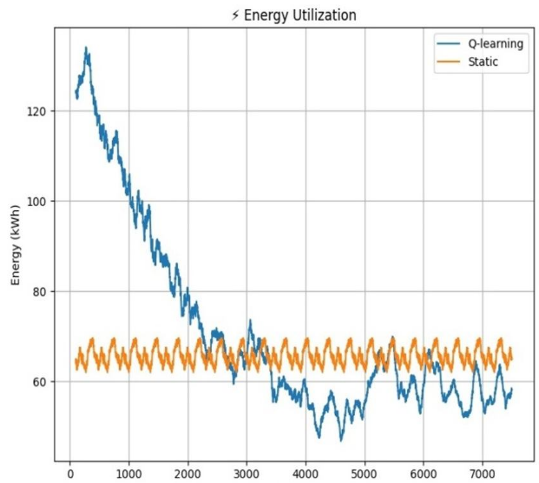


Fig. 5 Comparison of Energy utilization curve of Q-learning Vs. Static Rule Method

C. Energy Efficiency

The Battery Energy Efficiency as the ratio of total energy discharge to total energy charge expressed as a percentage under the dynamic operational conditions which incorporates the physical losses as well as effectiveness of decision making for which the Q-agent was trained to optimise charging.

Battery Management System (BMS) is used and energy efficiency is calculated as[25]:

Energy needed = (Battery capacity * (1 – SoC) / charge efficiency) (4)

Energy to battery is taken as the minimum of the charging rate and the battery needed for the purpose of filling the battery.

The total energy charged is the sum total of amount of energy that is taken from the grid for charging the battery, taking into accounts how well the charging worked, throughout all the episodes.

The total energy discharged is the total amount of energy obtained from the battery and provided to the load also considering discharge efficiency, throughout all the episodes 26].

Energy efficiency = (Discharged energy / Charged energy)*100% (5)

Battery Energy Efficiency is calculated using Static Rule based method and by Q-learning is calculated and their difference is observed.

Battery Energy Efficiency (Q-learning): 80.30%

Battery Energy Efficiency (Static): 53.70%
Absolute Battery % of difference of Battery Energy Efficiencies is found be 26.60%.

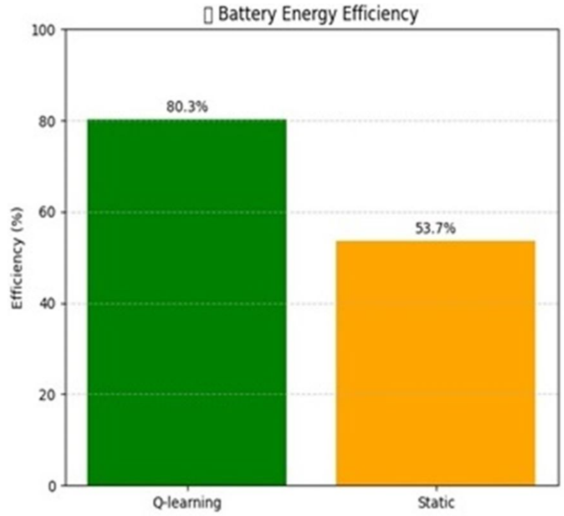


Fig. 6 Comparison of Battery Energy Efficiency of Q-learning and Static Rule methods

V. CONCLUSIONS AND FUTURE SCOPE

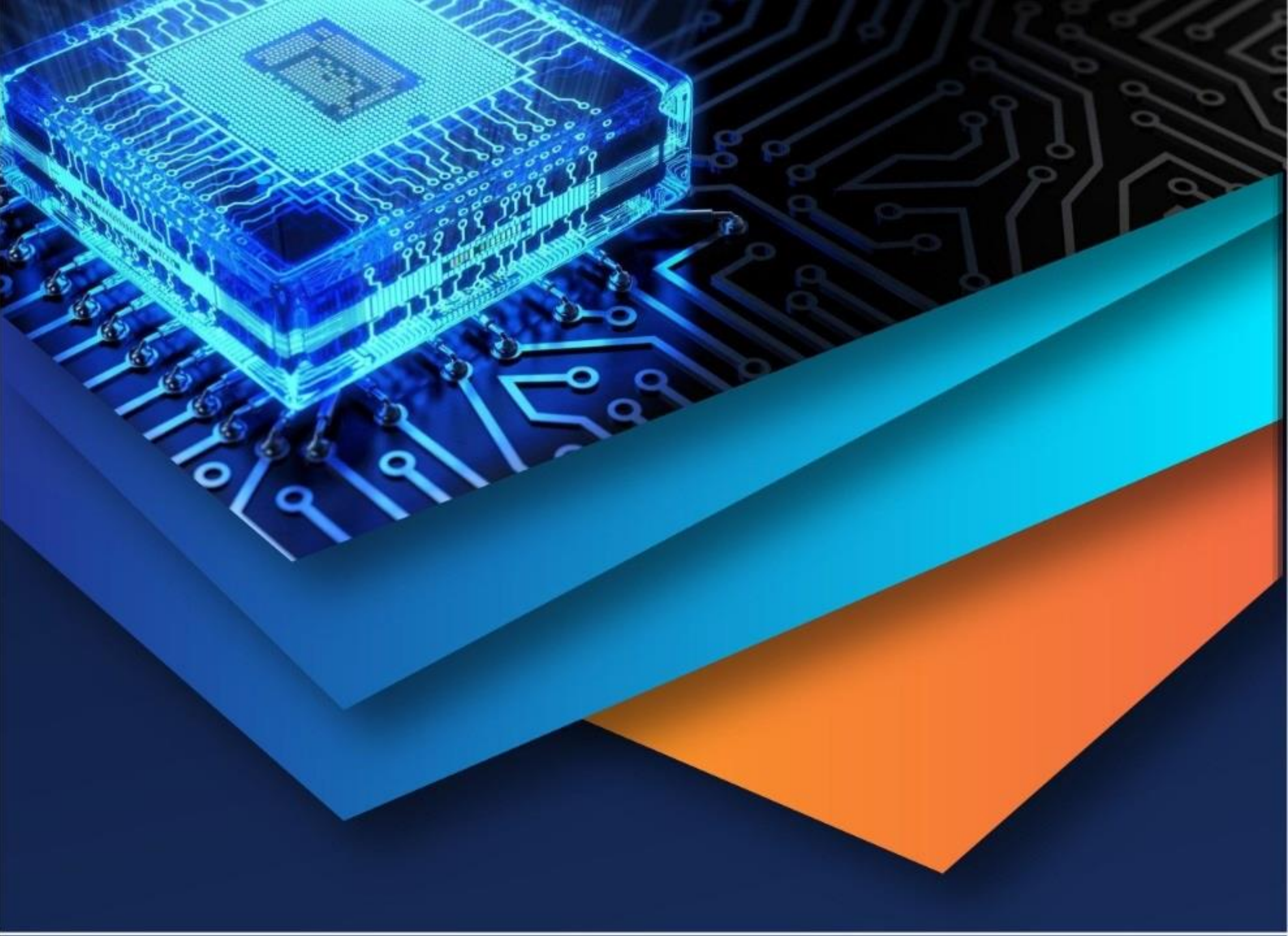
The proposed Reinforcement Learning (RL)-based battery energy management system demonstrated significant improvements over the conventional heuristic baseline. Through the Q-learning approach, the RL agent effectively learned optimal charge and discharge strategies by responding dynamically to time-varying electricity prices, load demands, and the battery's state of charge. The system achieved a 6.09% higher monetary profit, indicating superior economic performance. Additionally, the RL policy maintained a high absolute battery energy efficiency difference of 26.6% showing effective utilization of the battery with minimal losses. The battery energy efficiency may vary considering other constraints. The enhanced energy throughput further highlights the system's ability to capitalize on market fluctuations more proactively than the baseline strategy. Overall, the results confirm that the RL-based approach is a promising solution for intelligent, adaptive battery energy management.

Future work can incorporate focussing on extension of the RL framework like considering temperature-effects, degradation of battery, and even more composite demand response problems to take into account real-world scenarios. Implementation of advanced RL algorithms has potential of improving learning efficiency as well as scalability for wider systems. However, integration of renewable energy sources like wind and solar with BESS would bring highly effective, sustainable and economic energy solutions. The implementation and the testing of these RL based systems in real world scenarios, as in physical micro grids under different operating conditions would justify its practicality. Ultimately, exploration of multiple-agent reinforcement learning would lead to optimization of energy management in interconnected decentralized energy resources.

REFERENCES

[1] "A review of hybrid renewable energy systems: Solar and wind-powered solutions: Challenges, opportunities, and policy implications | Request PDF," ResearchGate, Dec. 2024, doi: 10.1016/j.rineng.2023.101621.
[2] M. S. Bakare, A. Abdulkarim, M. Zeeshan, and A. N. Shuaibu, "A comprehensive overview on demand side energy management towards smart grids: challenges, solutions, and future direction," Energy Informatics, vol. 6, no. 1, p. 4, Mar. 2023, doi: 10.1186/s42162-023-00262-7.
[3] "Battery energy storage control using a reinforcement learning approach with cyclic time-dependent Markov process | Request PDF," ResearchGate, doi: 10.1016/j.ijepes.2021.107368.
[4] Y. Tavakol-Moghaddam, M. Boroushaki, and M. Astaneh, "Reinforcement learning for battery energy management: A new balancing approach for Li-ion battery packs," Results in Engineering, vol. 23, p. 102532, Sep. 2024, doi: 10.1016/j.rineng.2024.102532.
[5] B. Jang, M. Kim, G. Harerimana, and J. W. Kim, "Q-Learning Algorithms: A Comprehensive Classification and Applications," IEEE Access, vol. 7, pp. 133653–133667, 2019, doi: 10.1109/ACCESS.2019.2941229.
[6] Y. Liu, Q. Lu, Z. Yu, Y. Chen, and Y. Yang, "Reinforcement Learning-Enhanced Adaptive Scheduling of Battery Energy Storage Systems in Energy Markets," Energies, vol. 17, no. 21, Art. no. 21, Jan. 2024, doi: 10.3390/en17215425.
[7] E. O. Arwa and K. A. Folly, "Improved Q-learning for Energy Management in a Grid-tied PV Microgrid," SAIEE Africa Research Journal, vol. 112, no. 2, pp. 77–88, Jun. 2021, doi: 10.23919/SAIEE.2021.9432896.

[8] "Battery Energy Management in a Microgrid Using Batch Reinforcement Learning." Accessed: Jul. 12, 2025. [Online]. Available: https://www.mdpi.com/1996-1073/10/11/1846
[9] J. Li, C. Wang, and H. Wang, "Deep reinforcement learning for wind and energy storage coordination in wholesale energy and ancillary service markets," Energy and AI, vol. 14, p. 100280, Oct. 2023, doi: 10.1016/j.egyai.2023.100280.
[10] S. Kim and H. Lim, "Reinforcement Learning Based Energy Management Algorithm for Smart Energy Buildings," Energies, vol. 11, no. 8, Art. no. 8, Aug. 2018, doi: 10.3390/en11082010.
[11] Y. Wang et al., "A comprehensive review of battery modeling and state estimation approaches for advanced battery management systems," Renewable and Sustainable Energy Reviews, vol. 131, no. C, 2020, Accessed: May 24, 2025. [Online]. Available: https://ideas.repec.org//a/eee/rensus/v131y2020ics1364032120303063.html
[12] Y. Yang, J. He, C. Chen, and J. Wei, "Balancing Awareness Fast Charging Control for Lithium-Ion Battery Pack Using Deep Reinforcement Learning," IEEE Transactions on Industrial Electronics, vol. 71, no. 4, pp. 3718–3727, Apr. 2024, doi: 10.1109/TIE.2023.3274853.
[13] W. Li et al., "Deep reinforcement learning-based energy management of hybrid battery systems in electric vehicles," Journal of Energy Storage, vol. 36, p. 102355, Apr. 2021, doi: 10.1016/j.est.2021.102355.
[14] "30 Years of Lithium-Ion Batteries - Li - 2018 - Advanced Materials - Wiley Online Library." Accessed: Jul. 09, 2025. [Online]. Available: https://advanced.onlinelibrary.wiley.com/doi/abs/10.1002/adma.201800561
[15] J. Xu, X. Mei, H. Wang, H. Shi, Z. Sun, and Z. Zou, "A model based balancing system for battery energy storage systems," Journal of Energy Storage, vol. 49, p. 104114, May 2022, doi: 10.1016/j.est.2022.104114.
[16] M. Zhang et al., "A Review of SOH Prediction of Li-Ion Batteries Based on Data-Driven Algorithms," Energies, vol. 16, no. 7, Art. no. 7, Jan. 2023, doi: 10.3390/en16073167.
[17] Q. Hassan, S. Algburi, A. Z. Sameen, H. M. Salman, and M. Jaszczur, "A review of hybrid renewable energy systems: Solar and wind-powered solutions: Challenges, opportunities, and policy implications," Results in Engineering, vol. 20, p. 101621, Dec. 2023, doi: 10.1016/j.rineng.2023.101621.
[18] W. Qiang and Z. Zhongli, "Reinforcement learning model, algorithms and its application," in 2011 International Conference on Mechatronic Science, Electric Engineering and Computer (MEC), Aug. 2011, pp. 1143–1146. doi: 0.1109/MEC.2011.6025669.
[19] O. Al-Ani and S. Das, "Reinforcement Learning: Theory and Applications in HEMS," Energies, vol. 15, no. 17, Art. no. 17, Jan. 2022, doi: 10.3390/en15176392.
[20] Y. Wang et al., "A comprehensive review of battery modeling and state estimation approaches for advanced battery management systems," Renewable and Sustainable Energy Reviews, vol. 131, no. C, 2020, Accessed: Jul. 09, 2025. [Online]. Available: https://ideas.repec.org//a/eee/rensus/v131y2020ics1364032120303063.html
[21] M. Pouyan, A. Mousavi, S. Golzari, and A. Hatam, "Improving the performance of Q-learning using simultanouse Q-values updating," in 2014 International Congress on Technology, Communication and Knowledge (ICTCK), Nov. 2014, pp. 1–6. doi: 10.1109/ICTCK.2014.7033528.
[22] K. Hu et al., "A review of research on reinforcement learning algorithms for multi-agents," Neurocomputing, vol. 599, p. 128068, Sep. 2024, doi: 10.1016/j.neucom.2024.128068.
[23] A. Kumar and D. Singh, "Adaptive epsilon greedy reinforcement learning method in securing IoT devices in edge computing," Discov Internet Things, vol. 4, no. 1, p. 27, Nov. 2024, doi: 10.1007/s43926-024-00080-7.
[24] "bstj27-1-109.pdf." Accessed: Jul. 21, 2025. [Online]. Available: https://vtda.org/pubs/BSTJ/vol27-1948/articles/bstj27-1-109.pdf
[25] Z. Lin, D. Li, and Y. Zou, "Energy efficiency of lithium-ion batteries: Influential factors and long-term degradation," Journal of Energy Storage, vol. 74, p. 109386, Dec. 2023, doi: 10.1016/j.est.2023.109386.
[26] "Green_access_efficiency_paper.pdf." Accessed: Jul. 22, 2025. [Online]. Available: https://pure.manchester.ac.uk/ws/portalfiles/portal/74265359/Green_access_efficiency_paper.pdf



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY