

# An Intelligent Microgrid Energy Management System Using a Deep Reinforcement Approach to Handle the Flexible Demands

Dr. Vijayalaxmi Biradar<sup>1</sup>, Dr. Anu G Pillai<sup>2</sup>

<sup>1</sup>Assistant Professor, Department of Electronics, Kalinga University, Naya Raipur, Chhattisgarh, India, [ku.vijayalaxmibiradar@kalingauniversity.ac.in](mailto:ku.vijayalaxmibiradar@kalingauniversity.ac.in)

<sup>2</sup>Assistant Professor, Department of Electrical, Kalinga University, Naya Raipur, Chhattisgarh, India, [ku.anugpillai@kalingauniversity.ac.in](mailto:ku.anugpillai@kalingauniversity.ac.in)

This research examines Deep Reinforcement Learning (DRL) methods' effectiveness in improving a microgrid (MG) Energy Management System (EMS). The study introduces an intelligent MG EMS that includes a wind turbine generator, an Energy-Storing System (ESS), a collection of thermostatically regulated loads, a group of weights, and a link to the primary grid. The suggested method aims to synchronize the adaptable methods by establishing the hierarchical resources, issuing direct demand, and power price demand. This study implements and compares several DRL methods. The findings indicate the DRL systems' capacity to converge the optimum policies. By including experiences and semi-deterministic learning methods in the well-recognized reward method, the study gets the maximum level of predictive accuracy and achieves convergence towards optimal rules.

**Keywords:** Microgrid, Energy Management System, Deep Reinforcement Approach, Flexible Demands.

## 1. Introduction

The evolution of the power grid revolves around shifting from traditional centralized energy sources to Distributed Energy Resources (DERs) with minimal environmental footprints [1]. This transformation necessitates strategies to address the difficulties that arise from the intermittent nature of Renewable Energy Supplies (RES). Smart grid methods, including upgraded metering facilities, Energy Storage Systems (ESSs) [2], and household Energy Management Systems (EMS) [3], are implemented to facilitate this transition. Microgrid (MG) [4] [19] employs methods in conjunction with DERs to fulfill the local electricity needs and promote the decentralized nature of the MG [16] [17]. MGs consist of low-voltage systems with a restricted local power supply and demand compared to the primary MG. MGs can

function in two ways: directly or indirectly, from the primary MG, where they purchase and sell energy electricity, relying on local production and storage. They provide technical and economic advantages such as improved system dependability, localized energy distribution, and increased opportunities for capital investment in DERs.

The research is fascinated by the EMS management that focuses on maintaining the energy sources, increasing the system's effectiveness, and optimizing the distribution of regional assets. The EMS encounters hurdles because of the characteristics of the MG, like size, fluctuation, unpredictability, and intermittent DERs, as well as the unpredictability of demand and the fluctuating pricing in the dynamic energy market [22]. To address these problems, it is necessary to make more advancements in MG design and control. To address the volatility of DERs, extra sources of flexibility at the design level are required. Novel and intelligent control approaches are needed to enhance energy dispatch and address uncertainties associated with the MG elements [20][24][29].

MG elements consist of DERs, electric weights, and an ESS. The DERs include renewable energy sources, like wind turbines [5] or solar Photovoltaic (PV) [6][21], and are supported by an energy generator that utilizes natural gas or a diesel engine. The growing fascination with DERs arises from their benefits. DERs have a competitive edge in expanded and deregulated energy markets due to their lower costs and faster building timeframes than centralized power facilities. DERs can prevent grid expansion, reduce construction costs, and keep high-voltage electricity lines. Using DERs diminishes the need for central dispatch and minimizes energy losses during long-distance transmission [26]. DERs enhance the MG's resilience by maintaining a local power supply during disruptions in the primary MG. An MG's electrical load elements might consist of residential or corporate loads [7]. The ESSs are constructed using batteries and deployed dispersed inside the MG or in a centralized way [8][18][23].

This research examines a DRL approach to control various energy resources in a practical MG scenario. The method considers the random behavior of multiple components of an MG, such as flexible loads, generators, and electric pricing. It utilizes DRL models to represent various grid elements and considers different power flow limitations.

## **2. Background of Reinforcement Learning and EMS**

This section examines the significant contributions to reinforcement learning and time series prediction. The research concludes by discussing the management strategies studied for MGs.

### **2.1 Reinforcement Learning**

Reinforcement Learning (RL) seeks to improve systems that respond to situations by heuristically determining a set of optimum behaviors for various potential states of the surroundings to maximize a goal [9]. The term "State-Action-Reward" is the overarching designation for this category of techniques. Deep Reinforcement Learning (DRL) ensures the smooth progression of RL when conditions are ongoing or too complex to be efficiently handled using traditional RL methods [10][25]. The authors developed an EMS that is very efficient by using deep deterministic policy gradients [11]. According to the current data, the suggested model offers the most efficient scheduling options for storing and releasing the

*Nanotechnology Perceptions* Vol. 20 No. S4 (2024)

power of the ESS and the input power of the heating, ventilation, and air conditioning systems. A DRL system is created to manage indoor and domestic hot water temperatures [12]. The method aims to minimize energy consumption and maximize renewable energy output. The authors are examining the utilization of RL for response to demand purposes in the smart grid. They showcased a range of RL uses. They have been used to regulate many energy systems, including electric cars, high-voltage systems, smart appliances, and batteries. In deep reinforcement learning, many neural networks have been specifically designed to replicate the weight matrix of the entire system. The Deep Q-Network (DQN) and Policy Gradient (PG) are the primary foundations from which several derivative techniques have been presented [13][27]. Without delving further into a comprehensive examination of these approaches, it is crucial to outline the primary distinctions among all these procedures and propose an innovative solution.

- The approaches provide the most optimum solution. One quickly encounters limitations: 1 - about the duration it takes to resolve, 2 - about addressing issues that are likely to become non-linear.
- The approaches employ a solver for every new dataset. A license accompanies it. Therefore, they need both time and financial resources.
- RL is an efficient heuristic system that rapidly generates solutions to problems, outperforming precise solvers in speed.
- DRL approaches need a substantial volume of data to establish the weights of the system.
- The research has examined these strategies in many studies addressing improving edge placement, optimizing communication drone installation, and EMS.

## 2.2 Energy Management Systems

The Worldwide Electrotechnical Committee defines an EMS as a computerized system comprising an operating system and a set of programs [14]. The EMS is designed to support the efficient functioning of electrical power plants and transmission amenities, ensuring the safety of energy supply at the lowest possible cost. The authors aim to enhance the efficiency of the MG while it operates interconnectedly. The serial linear programming approach suggests and resolves two market approaches. Mobini-Serajy et al. proposed using dynamic software to improve the functioning of a grid-connected MG [15]. The primary goals are to minimize the expenses associated with purchasing power from the source and the costs associated with battery degradation. The techniques were categorized into many groups: Classical methods (including linear and nonlinear coding, dynamic coding, and Rule-Based techniques), meta-heuristic methods (such as genetic and swarming optimization), Artificial intelligence techniques, Stochastic and resilient programming techniques, and modeling predictive management. Lopez-Martin et al. investigated the application of DRL using two approaches [16]. The use of DQN and Deep Policy Gradient algorithms enables the efficient optimization of scheduling for building EMS using online methods. The result demonstrates that Deep Policy Gradient outperforms Q-learning in the context of online scheduling.

### 3. Proposed DRL-based energy management system for microgrid

#### 3.1 Problem formulation

When it comes to constructing the structure of the MG system, there are several alternatives to take into account. Several articles implemented various components of the MG, including restrictions on the usage side and limitations on the generation side. The study considers operational limitations in the MG and the limits imposed by renewable energy production. Additional studies have examined limitations for energy storage, including the rates at which it is charged and discharged, restrictions on power costs, and considerations of carbon emissions. This study examines the primary limitations of MGs, namely the restrictions on exchanging power with the outside grid, the restraints on power production from various distributed generating devices, and the constraints on the flow of electricity. The MG model that has been chosen is linked to the utilities at the Point of Common Coupling (PCC). It includes traditional distributed power plants, energy storage devices, solar PV panels, wind turbines, and variable workloads.

The suggested formulation takes into account traditional generators that are limited by the following calculations:

$$E_{\min}^{\text{MG}} < E_x^{\text{MG}} < E_{\max}^{\text{MG}} \quad (1\text{-a})$$

$$E_x^{\text{MG}} + Q_x^{\text{MG}} < S_{\max}^{\text{MG}} \quad (1\text{-b})$$

$E_x^{\text{MG}}$  and  $Q_x^{\text{MG}}$  denote the active and reactive energy generated by the generator at a particular moment,  $x$ . The generator's score is denoted by  $S_{\max}^{\text{MG}}$ . A quadratic equation represents the expense function for running traditional generators.

$$C_x^{\text{MG}} = \{a_1(E_x^{\text{MG}})^2 + b_1E_x^{\text{MG}} + c_1\} \Delta x \quad (2)$$

$a_1$ ,  $b_1$ , and  $c_1$  are fixed variables. The following formulas limit ESS:

$$0 < E_x^{\text{M}} < E_{\max}^{\text{M}} \quad (3\text{-a})$$

$$P_{\min} < E_x^{\text{M}} < P_{\max} \quad (3\text{-b})$$

$$P_x = P_{x-1} + \alpha_{\text{ch}} v_x E_x^{\text{M}} \Delta x - \frac{(1-v_x) E_x^{\text{M}} \Delta x}{\alpha_{\text{dis}}} \quad (3\text{-c})$$

The power associated with charging or discharging is denoted as  $E_x^{\text{M}}$ , whereas the power level, known as State Of Charge (SOC), is denoted by  $P_x$ . The binary parameter  $v_x$  shows whether the ESS is in the charging state ( $v_x = 1$ ) or draining state ( $v_x = 0$ ). The charged and discharged efficiency are denoted by the terms "ch" and "dis" respectively. The duration of the accused or discharged process is denoted by  $\Delta x$ :

The following expression determines the electricity exchange between the user and the utility:

$$-E_{\max}^{\text{U}} < E_x^{\text{U}} < E_{\max}^{\text{U}} \quad (4\text{-a})$$

$$E_{\max}^{\text{U}2} + Q_x^{\text{U}2} < (S_{\max}^{\text{U}})^2 \quad (4\text{-b})$$

$E_{\max}^{\text{U}}$  and  $Q_x^{\text{U}}$  denote the active and reactive power transactions with the utilities. The most

significant exchange of complicated power is represented by the variable  $S_{\max}^U$ . The expression for the cost of acquiring electricity from the utilities is as follows:  $R$  represents the actual time pricing.

$$C_x^U = E_x^U R_x \Delta_x \quad (5)$$

The regulation of energy resources takes into account power flow restrictions. The electrical flow restrictions at every branch  $yz$  are shown as follows:

$$E_x^{yz2} + Q_x^{yz2} < (S_{\max}^{yz})^2 \quad (6)$$

To maintain voltage levels within acceptable parameters, the voltage is restricted according to the following guidelines:

$$|U^y|_{\min} < |V_x^y| < |U^y|_{\max} \quad (7)$$

The voltage at bus  $y$  at time  $x$ , denoted as  $V_x^y$ , is constrained by specified lowest and upper levels. The primary goal of the reward function is to maximize operating expenses while maintaining compliance with system limitations. The incentive function is designed to be directly related to the operating costs of the MG. These prices involve the costs associated with running the traditional generators and the expenditures for purchasing electricity from the utilities. Equation (8) concisely describes the reward value at every time step.

$$r_x = -\sum_{x=0}^N C_x^{MG} + C_x^U \quad (8)$$

The reward system considers any deviation from the energy flow limits by assigning a negative reward and ending the training cycle. There is a penalty for behaviors that break energy storage limitations, which serves to discourage such actions.

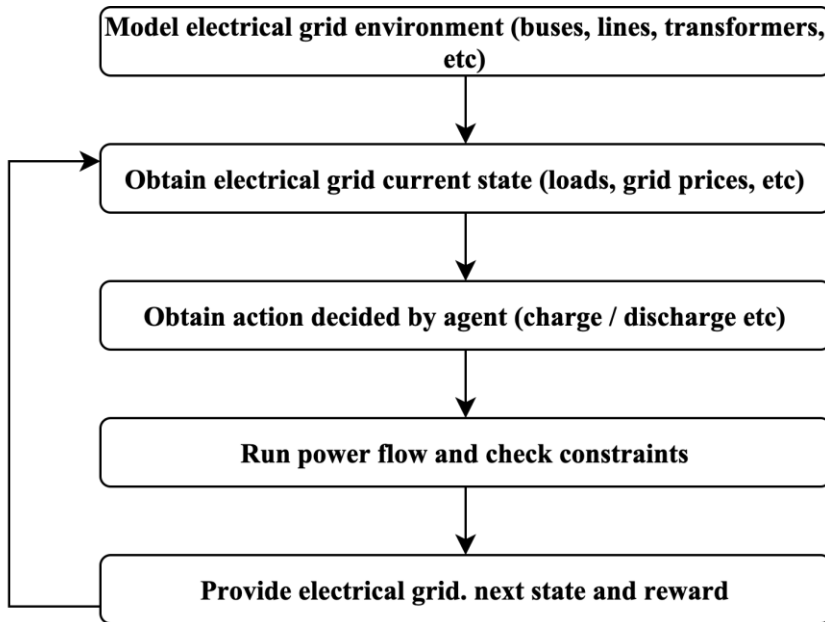


Fig. 1. High-level workflow of the proposed system

The formulation phases are shown in Fig. 1, which illustrates a high-level flowchart depicting the many procedures undertaken to resolve the issue.

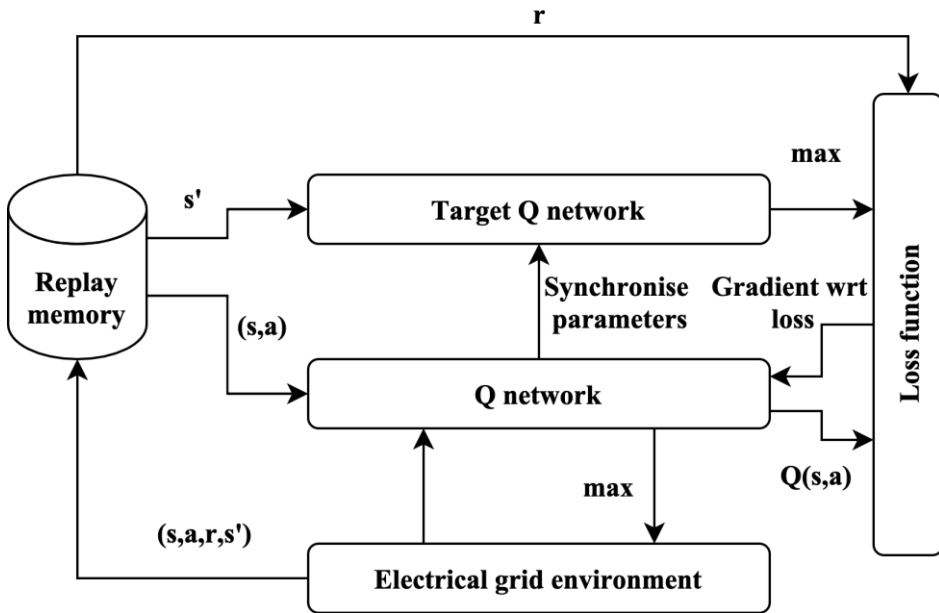


Fig. 2. DRL agents' relationship

Fig. 2 depicts the agent's relationship to the surroundings and the overall progression of the training process. The MG's state parameters are inputted into the Q-network, which then uses a greedy strategy to make decisions that optimize the Q-value. The exchanges between the surroundings and the agents are gathered and stored in a replaying memory, which is then used to train both online and target networks. This procedure continues until cycles conclude or termination requirements are satisfied.

### 3.2 Case study

The research constructs an MG to examine the suggested approach, as seen in Fig. 3. The MG comprises many conventional and non-conventional dispersed power production units, such as PV systems, wind turbines, diesel engines, and energy storage units. It is linked to an outside grid via a PCC.

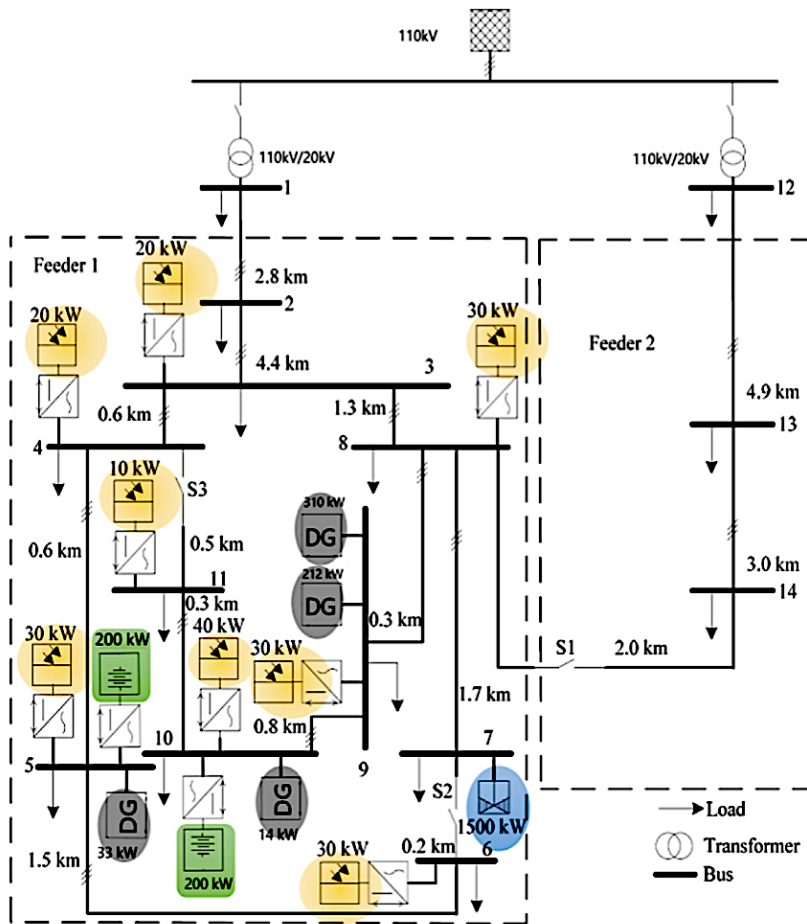


Fig. 3. Sample scenario for the case study

The MG ecosystem is established using the OpenAI Gym library to serve as a structure for the reinforcement learning interface. They use the Pan-depower toolbox to do power flow computations and guarantee that limits are not breached. The DQN agent is implemented using a Stable Baseline to resolve the scenario.

A feedforward neural network is employed to establish a mapping between various states at any given time and their respective state-action function values for every single input. The network receives the data from the previous 24 hours for multiple variables (such as PV, wind, and grid pricing). It currently produces the Q-value for every possible operation (such as charging or discharging). Rectified linear units (ReLU) serve as an activating function in the buried layers.

The desired network is built using a similar architecture to the learning Q-network and is regularly updated. At every stage, a replay memory is created with a capacity 50,000 to hold transitions. At the beginning, the agent is permitted to conduct arbitrary behaviors for the first thousand times before the commencement of the learning process. The efficiency of the

learning network is recorded, and variables are stored at regular intervals. The optimal network's variables are retained for testing purposes.

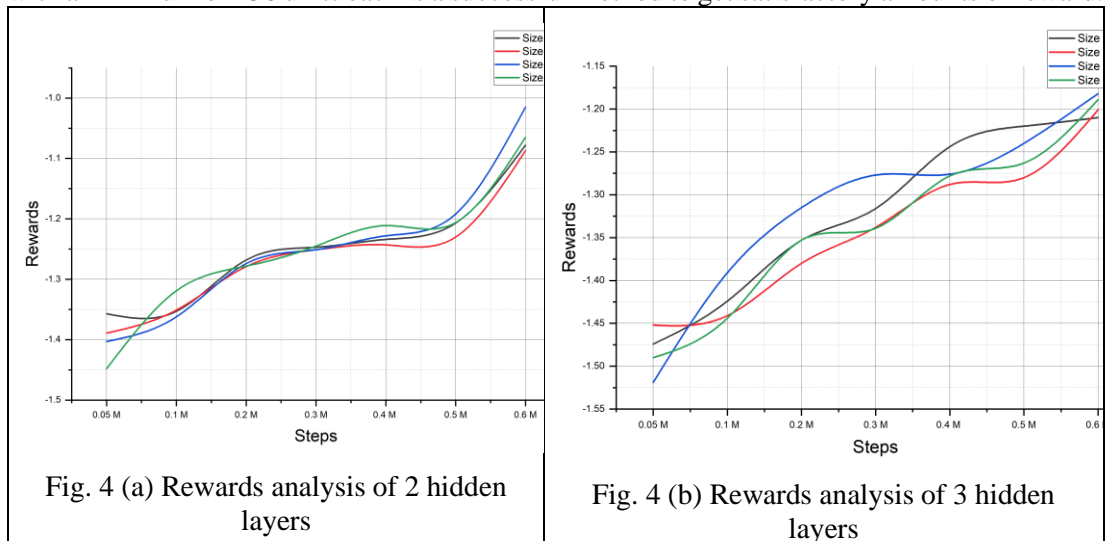
#### 4. Simulation results and outcomes

Multiple tests are carried out to optimize the neural network's various hyperparameters. Due to the lengthy duration of training calculation, which typically takes around 24 hours to complete 1 million steps using an i7 CPU @ 4.1 GHz with 8 GB RAM, the settings are adjusted manually based on cognitive manner and information. This will help actual installations since GPU and cloud computing technology advancements significantly expedite the learning time.

The data is partitioned into two subsets: training (70%) and testing (30%). The agent is permitted to undergo training in the established architecture for 500,000 time steps in every network arrangement. The outcomes are compared to those achieved in the ideal scenario. During the simulation of the test information, every single step lasts about 0.01 seconds. The 2200-hour simulation requires around 25 seconds to complete. This illustrates this strategy's efficacy in real-time scheduling for power supplies.

##### 4.1. Hyperparameter tuning

To choose an appropriate neural network architecture, simulations are performed with varying numbers of layers and nodes per concealed layer. It has been noticed that using two layers with a minimum of 256 units each is a successful method to get satisfactory amounts of reward.



The learning process for different numbers of neural networks with two and three concealed layers is shown in Figures 4(a) and 4(b). There is no discernible benefit to using three layers compared to 2 layers, and in some instances, it performs worse than having two concealed levels. Increasing the size of the community will lead to a decrease in weight sharing, which, in turn, enhances the stability of the learning process. Optimal results were obtained by setting the learning rate from 0.001 to 0.0001. The research analyzes the impact of the discount rate (c) used in revising the value of the desired network. The amount of discount directly impacts



the agent's success aim. Given that a significant portion of the incentive function relies on the stochastic nature of power pricing, this increases the difficulty of accurately forecasting future outcomes. Reducing the price reduction factor is more advantageous in maximizing the accrual of rewards.

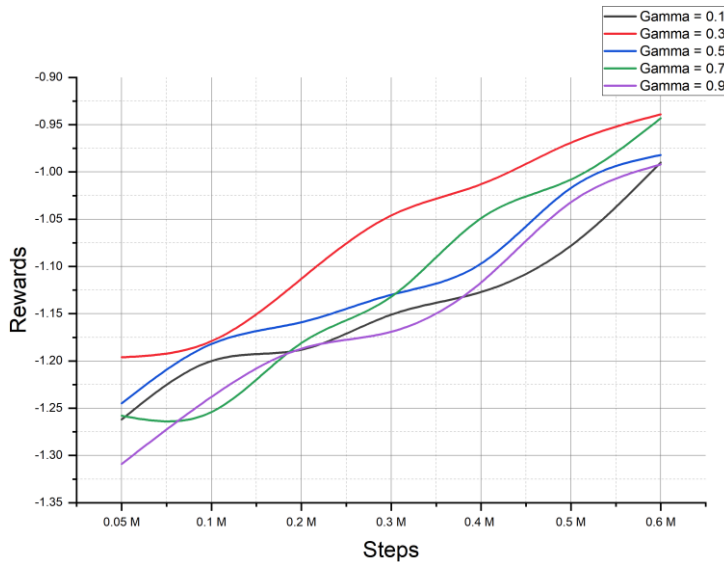


Fig. 5. Rewards analysis over different gamma values

Fig. 5 displays the learning curves of several agents with varying discount rates. Gamma values below 0.5 outperformed higher values. This implies that the present activities only impact relatively brief time intervals in the future. Regarding the impact of batch size on the application of gradient descent for modifying the weighting of the learned networks, it has been shown that bigger batch sizes result in improved learning efficiency. However, batch sizes beyond 256 did not show any significant benefits.

#### 4.2. Energy schedules

To evaluate the efficiency of the created agent, the research simulates the planning of energy resources in the MG. The simulation is conducted in a controlled test setting with past information over three months.

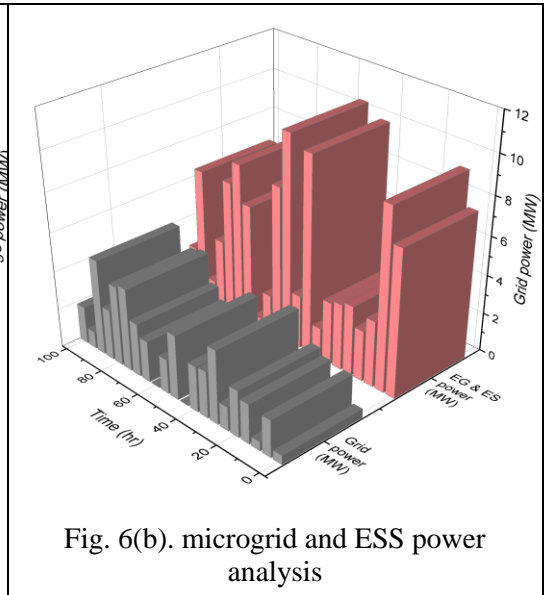
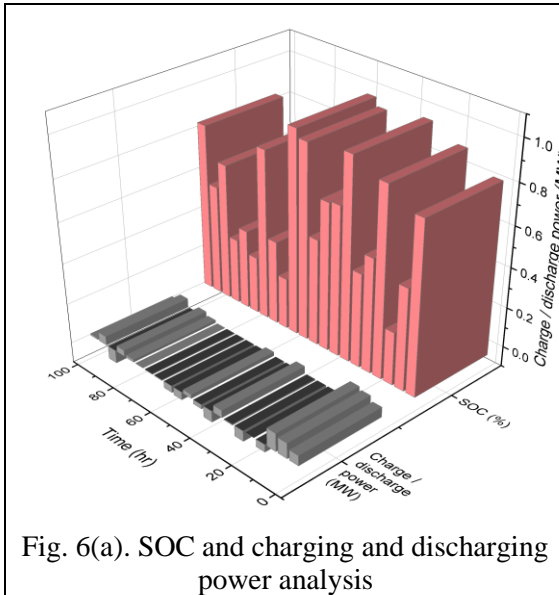


Fig. 6(a) displays the complete charging and discharging schedules and the SOC for the gadgets that store energy. The agent has permission to charge or discharge the ESS so long as it maintains the defined dynamic limitations of the energy storage. The agent aims to optimize the advantages of managing energy storage by considering external grid electricity costs and the assets available inside the MG.

Fig. 6(b) displays the generating schedules for the diesel engines, ESS, and external grid systems. Like ESS, the agent aims to optimize the DG units' functioning by considering the units' operational expenses and other relevant state parameters. To evaluate the efficiency of the created agent, the research contrasts the daily expenses resulting from following the agent's activities with the ideal solution found by a mixed integer linear programming approach. Unlike the suggested method, the solver can anticipate future values for variables over the time range because it does not consider uncertainties.

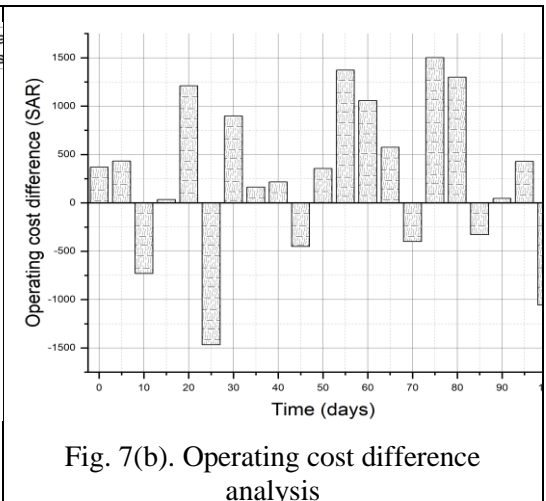
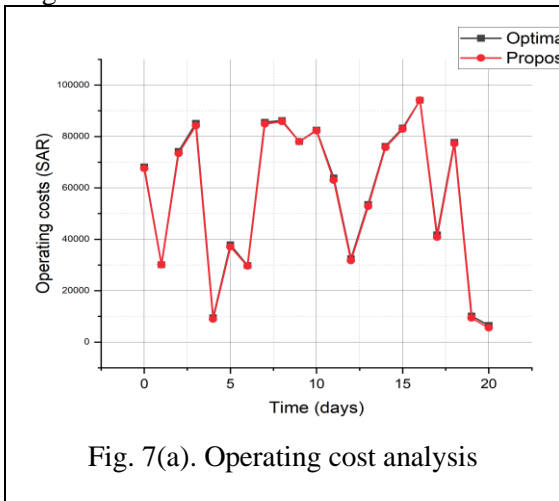


Fig. 7(a) displays the daily expenses accumulated by both strategies during the testing period. Fig. 7(b) illustrates the disparity in daily expenses between the ideal and DQN solution, providing more evidence of the distinctions.

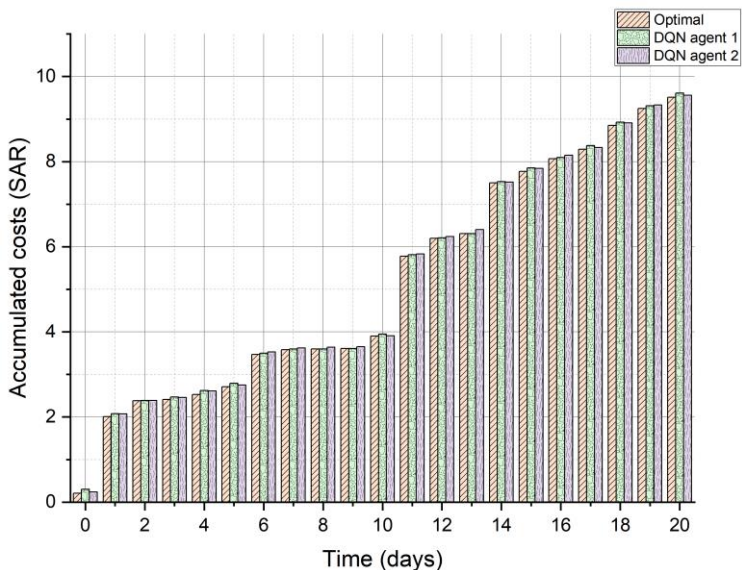


Fig. 8. Accumulated cost analysis

To demonstrate the effectiveness of the proposed DQN approach, the research has graphed the total expenses across the testing period in Fig. 8. The research offers two DQN agents besides the best solution, one with unoptimized hyperparameters and the other with optimized ones. DQN demonstrated its capacity to perform energy administration, yielding outcomes similar to those achieved by the optimum solution. The DQN agent took 0.01 seconds for each step to compute the actions in the experimental set, while the MILP solution took 2.4 seconds.

## 5. Conclusion and findings

This research examines a multi-task EMS designed for a residential MG that incorporates several sources of flexibility. The suggested MG system considers the possible demand flexibility provided by price-responsive loading and loads. The proposed EMS facilitates the coordination of the ESS, the primary power grid, the Loads, and the price-responsive loading to manage local assets efficiently. The inherent unpredictability of the MG elements and the complex structure of their variables make it advantageous to employ intelligent learning-based approaches, such as deep reinforcement learning methods, in the EMS. This work thoroughly and empirically evaluates the most advanced DRL algorithms.

The research has introduced enhanced versions of the methods that have shown superior performance compared to the existing techniques. The numerical findings indicate that the DRL methods achieved varying degrees of convergence. The findings demonstrated that including experience replaying and semi-deterministic learning in the algorithm enhanced convergence and yielded superior rules. The findings were compared to those of a theoretical

optimum controller with complete knowledge of the method's parameters and dynamics throughout the day.

The results were compared to those of an electrical retailer that purchases power from the day-ahead marketplace and provides the same level of demand without using an MG. The findings demonstrate a notable superiority of the suggested MG's architecture and coordination methods over an electrical retailer regarding economic viability and capacity to withstand challenging situations. The technique has accomplished 52% of the conceptual optimum's profit and has outperformed the seller by 25%. Establishing and carrying out an effective EMS for future MGs is challenging due to the high dimensions and unpredictability of the MG elements. While DRL approaches have shown effectiveness in gaming, they have flaws. These methods encounter challenges when applied to real-world situations because of their inefficiency in handling information, unpredictability, and poor convergence. Presently, there is a focused endeavor to raise the efficiency of DRL strategies and broaden their usefulness in practical situations.

## References

1. Strezoski, L. (2023). Distributed energy resource management systems—DERMS: State of the art and how to move forward. *Wiley Interdisciplinary Reviews: Energy and Environment*, 12(1), e460.
2. Stevovic, I., Hadrović, S., & Jovanović, J. (2023). Environmental, social and other non-profit impacts of mountain streams usage as Renewable energy resources. *Archives for Technical Sciences*, 2(29), 57-64.
3. Raya-Armenta, J. M., Bazmohammadi, N., Avina-Cervantes, J. G., Sáez, D., Vasquez, J. C., & Guerrero, J. M. (2021). Energy management system optimization in islanded microgrids: An overview and future trends. *Renewable and Sustainable Energy Reviews*, 149, 111327.
4. Johnson, C., Khadka, B., Basnet, R.B., & Doleck, T. (2020). Towards Detecting and Classifying Malicious URLs Using Deep Learning. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, 11(4), 31-48.
5. Jiang, Z. (2021). Installation of offshore wind turbines: A technical review. *Renewable and Sustainable Energy Reviews*, 139, 110576.
6. Samyuktha, S.S., & Subaji, M. (2024). Variational Autoencoder Diffusion Model (VAEDM) and Divergence Asynchronous Reinforcement Learning (DARL) for Rail Surface Defect Detection. *Journal of Internet Services and Information Security*, 14(2), 298-317.
7. Tang, H., Wang, S., & Li, H. (2021). Flexibility categorization, sources, capabilities and technologies for energy-flexible and grid-responsive buildings: State-of-the-art and future perspective. *Energy*, 219, 119598.
8. Mishra, D., & Kumar, R. (2023). Institutional Repository: A Green Access for Research Information. *Indian Journal of Information Sources and Services*, 13(1), 55–58.
9. Gorostiza, F. S., & Gonzalez-Longatt, F. M. (2020). Deep reinforcement learning-based controller for SOC management of multi-electrical energy storage system. *IEEE Transactions on Smart Grid*, 11(6), 5039-5050.
10. Arora, G. (2024). Desing of VLSI Architecture for a flexible testbed of Artificial Neural Network for training and testing on FPGA. *Journal of VLSI Circuits and Systems*, 6(1), 30-35.
11. Lissa, P., Deane, C., Schukat, M., Seri, F., Keane, M., & Barrett, E. (2021). Deep reinforcement learning for home energy management system control. *Energy and AI*, 3,

- 100043.
12. Zhang, D., Li, S., Deng, Z., & Tang, X. (2023). Lithium-Plating Suppressed and Deep Deterministic Policy Gradient-Based Energy Management Strategy. *IEEE Transactions on Transportation Electrification*, 10(1), 1863-1875.
13. Gökhan, N. U. R., BARIŞ, B. N., Levent, B., SAZAKLIOĞLU, B. S., & Elvan, A. K. (2023). BUSER Transcutaneous Electric Nerve Stimulator Device Design. *Natural and Engineering Sciences*, 8(1), 18-30.
14. Mobini-Serajy, Z., Radmehr, M., & Ghorbani, A. (2022). An improved adaptive controller design for enhancing dynamic stability of grid-connected MGs. *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, 46(1), 91-108.
15. Lopez-Martin, M., Carro, B., & Sanchez-Esguevillas, A. (2020). Application of deep reinforcement learning to intrusion detection for supervised problems. *Expert Systems with Applications*, 141, 112963.
16. Jelena, T., & Srđan, K. (2023). Smart Mining: Joint Model for Parametrization of Coal Excavation Process Based on Artificial Neural Networks. *Archives for Technical Sciences*, 2(29), 11-22.
17. Rana, M. M., Uddin, M., Sarkar, M. R., Meraj, S. T., Shafiullah, G. M., Muyeen, S. M., ... & Jamal, T. (2023). Applications of energy storage systems in power grids with and without renewable energy integration—A comprehensive review. *Journal of energy storage*, 68, 107811.
18. Shadadi, E., & Alamer, L. (2022). Hierarchical Parallel Processing for Data Clustering in GPU Using Deep Nearest Neighbor Searching. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, 13(4), 155-168.
19. Ahmad, S., Shafiullah, M., Ahmed, C. B., & Alowaifeer, M. (2023). A review of microgrid energy management and control strategies. *IEEE Access*, 11, 21729-21757.
20. Sivayazi, K., & Giriraj, M. (2024). Dynamic Path Planning Algorithm for Mobile Robots: Leveraging Reinforcement Learning for Efficient Navigation. *Journal of Internet Services and Information Security*, 14(2), 226-236.
21. Marques Lameirinhas, R. A., Torres, J. P. N., & de Melo Cunha, J. P. (2022). A photovoltaic technology review: History, fundamentals and applications. *Energies*, 15(5), 1823.
22. Rath, S., Mirajkar, O., Shukla, S., Deshmukh, L., & Dangare, L. (2024). Advancing Crack Detection Using Deep Learning Solutions for Automated Inspection of Metallic Surfaces. *Indian Journal of Information Sources and Services*, 14(1), 93–100.
23. Rajagopalan, A., Swaminathan, D., Alharbi, M., Sengan, S., Montoya, O. D., El-Shafai, W., ... & Aly, M. H. (2022). Modernized planning of smart grid based on distributed power generations and energy storage systems using soft computing methods. *Energies*, 15(23), 8889.
24. Dr. Waleed S. Alnumay. (2024). Use of machine learning for the detection, identification, and mitigation of cyber-attacks. *International Journal of Communication and Computer Technologies*, 12(1), 38–44.
25. Xu, G., Shi, J., Wu, J., Lu, C., Wu, C., Wang, D., & Han, Z. (2024). An optimal solutions-guided deep reinforcement learning approach for online energy storage control. *Applied Energy*, 361, 122915.
26. Chowdhury, U. (2022). Millimeter-Wave Technology Based Applications Across Industries. *National Journal of Antennas and Propagation*, 4(1), 33-40.
27. Alhasnawi, B. N., Jasim, B. H., Sedhom, B. E., & Guerrero, J. M. (2023). A new communication platform for smart EMS using a mixed-integer-linear-programming. *Energy Systems*, 1-18.
28. Ozyilmaz, A. T. (2023). Conducting Polymer Films on Zn Deposited Carbon Electrode. *Natural and Engineering Sciences*, 8(2), 129-139.