[1]**Nilam B. Panchal**

[2]**Vaibhavi M. Parmar**

[3]**Dharmistha V. Makwana**

[4]**Mehulsinh G. Jadeja**

[5]**Rinal K. Ahir**

Enhancing Energy Efficiency in Smart Grids through Reinforcement Learning-Based Control Strategies

JES
Journal of
Electrical
Systems

Abstract: - The rapid growth of smart grids has ushered in new opportunities for enhancing energy efficiency through advanced control strategies. This paper explores the potential of reinforcement learning (RL) to improve the energy efficiency of smart grids, focusing on RL-based control strategies. We begin with a comprehensive review of existing technologies, examining a range of architectures and methods used to implement RL in smart grid environments. This review highlights both the benefits and limitations of these approaches, offering a balanced analysis of their effectiveness in addressing the unique challenges of smart grid management. Following this review, we propose a new RL-based control strategy designed to optimize energy efficiency. Our approach leverages the strengths of state-of-the-art RL algorithms while addressing common shortcomings identified in previous work. We evaluate our strategy using a detailed simulation that reflects real-world smart grid scenarios. The results demonstrate significant improvements in energy efficiency compared to traditional control methods. Finally, we discuss best practices for applying RL in smart grids, providing guidelines for researchers and practitioners seeking to implement these strategies. Our recommendations focus on maximizing energy efficiency while ensuring stability and scalability in smart grid systems. Through this work, we aim to contribute to the ongoing development of sustainable and efficient smart grid technologies.

Keywords: Smart Grid; Reinforcement Learning (RL); Multiagent; Artificial Intelligence (AI); Decentralization.

I. INTRODUCTION

The transition to smart grids represents a significant evolution in how electricity is distributed and managed, aiming to achieve greater energy efficiency and sustainability. Reinforcement learning (RL), a form of machine learning where agents learn by interacting with their environment to maximise a reward signal, is becoming an increasingly important tool in the smart grid landscape. Its ability to learn and adapt in real time offers substantial benefits for managing the complexities of modern energy systems.

RL's application in smart grids has been growing steadily. Researchers [1-5] have used RL to improve various aspects of smart grid management, including demand response, load balancing, and energy storage. Despite these advancements, there are still limitations in current approaches that hinder optimal energy efficiency:

➢ Limited Flexibility: Traditional smart grid control systems often use static rules or models, which are not easily adaptable to changing grid conditions.

➢ Scalability Challenges: As smart grids grow in complexity and size, many existing methods struggle to scale effectively.

➢ Data and Model Dependency: Some approaches require extensive data collection and complex modelling, which can be resource-intensive and inflexible.

Reinforcement learning provides a promising solution to these challenges. By allowing control strategies to evolve through continuous learning, RL-based methods can better adapt to dynamic environments, making them suitable for smart grids' unpredictability and complexity.

This paper proposes a novel RL-based control strategy to enhance energy efficiency in smart grids. Our approach addresses the limitations mentioned earlier, providing a flexible, scalable, and data-driven solution that shows considerable promise for improving energy management.

The remainder of this paper is structured as follows. In the next section, we discuss related work, examining the current state of RL-based strategies in smart grids and identifying gaps in the existing research. Following that, we describe our proposed methodology, detailing our approach's architecture and specific RL algorithms. Next, we present the mathematical model underpinning our strategy, clearly explaining the theoretical framework. The results and discussion section covers our experimental findings, comparing our RL-based approach with other

---

[1] *Corresponding author: Nilam B. Panchal, Email: nilam.b2211@gmail.com

[1,2,3,4,5] Electrical Engineering Department, Vishwakarma Government Engineering College, Chandkheda, Ahmedabad, Gujarat, India.

methods to demonstrate its effectiveness. Finally, we summarise our contributions and outline directions for future research in this field.

II. RELATED WORK

The smart grid represents a convergence of advanced technologies to optimise energy distribution, improve grid stability, and integrate renewable energy sources. Reinforcement learning (RL) has emerged as a promising approach for addressing these complex challenges. In this literature review, we examine existing research on RL-based methods for smart grid energy management, identifying key themes, methodologies, and trends in the field.

A. Early Approaches to Smart Grid Energy Management

Early studies in smart grid energy management focused on conventional control methods [6], such as rule-based systems and static optimisation techniques [7]. While these approaches provided a starting point, they were limited in adaptability and scalability. Researchers quickly recognised the need for more dynamic solutions capable of handling the complexities of modern smart grids [8].

B. Introduction of Reinforcement Learning in Smart Grids

Reinforcement learning (RL) has brought about a paradigm shift in smart grids by introducing adaptive, data-driven approaches to energy management. RL's ability to learn from interaction with the environment makes it ideal for smart grids' dynamic and uncertain conditions. Unlike traditional methods [4-9] that rely on predefined rules or static models, RL agents can adjust their strategies based on real-time feedback, allowing for more flexible and efficient control.

Introducing Q-Learning [10] into smart grids was a pivotal development, where agents could learn optimal actions by maximising expected cumulative rewards. This flexibility proved useful in smart grid contexts, allowing RL agents to manage complexities like fluctuating energy demand, renewable energy variability, and grid stability [11]. Studies demonstrated the value of RL in optimising demand response, enabling the grid to adapt energy consumption during peak periods without compromising user comfort. Additionally, RL has been applied to energy storage management, with agents learning the optimal times to charge and discharge energy storage systems based on real-time conditions, thus improving energy efficiency, and reducing costs [12].

Another significant application of RL in smart grids is in optimal power flow. Here, agents learn to find the best routes for energy distribution while considering factors like grid congestion and system constraints [13]. This learning-based approach leads to more stable and efficient power flow, reducing energy losses and enhancing grid stability. The integration of distributed energy resources (DERs), such as solar panels and wind turbines, has also benefited from RL, with agents learning how to utilise renewable energy best to meet demand [14].

Despite these advancements, RL in smart grids faces several challenges, including ensuring stability and convergence during training, reducing training times, and generalising across different grid configurations [15]. However, RL's flexibility and adaptability offer promising solutions to these challenges, driving the continuous evolution of smart grid technology.

The role of RL in smart grids represents a significant advancement in energy management. By allowing agents to learn from the environment, RL-based approaches can adapt to changing conditions, optimise energy distribution, and enhance grid stability. These developments contribute to smarter, more sustainable energy systems, opening the door to future innovations in smart grid technology.

C. Deep Reinforcement Learning and Its Impact

Deep reinforcement learning (DRL) combines RL with deep learning to address complex state spaces. DRL uses neural networks to approximate Q-values, enabling agents to learn more sophisticated control strategies. Paper [14] showed that DRL-based approaches could significantly improve energy efficiency by optimising power flow and integrating renewable energy sources. The flexibility of DRL makes it a powerful tool for managing distributed energy resources (DERs) and handling large-scale smart grid environments.

One of the key benefits of DRL is its ability to learn sophisticated control strategies, enabling agents to make more nuanced decisions in complex environments [15]. This capability is crucial in smart grids, where agents must consider various factors, such as energy demand, power flow, grid stability, and integrating distributed energy resources (DERs). Studies have shown that DRL can significantly improve energy efficiency by optimising power

distribution and managing renewable energy sources. Paper [16] demonstrated that DRL-based approaches could achieve these goals, highlighting the potential for real-world impact.

DRL's flexibility allows it to adapt to changing conditions within smart grids, enabling it to optimise energy management even in large-scale environments. This adaptability is essential for managing DERs, which are increasingly becoming vital to modern smart grids. As more renewable energy sources like solar and wind are integrated into the grid, the need for intelligent control strategies that balance supply and demand becomes more critical [17]. DRL agents can learn to adjust energy production and consumption based on real-time data, ensuring efficient energy flow while maintaining grid stability.

Another significant impact of DRL is its capacity to improve the scalability of smart grid management. Traditional control methods often struggle with scalability, especially as smart grids grow and complexity [18]. With its deep learning foundations, DRL can handle large-scale environments and learn from vast amounts of data, making it an ideal solution for smart grid applications [18]. This scalability opens up new possibilities for managing complex energy systems, allowing for greater integration of renewable resources and more efficient energy distribution. It has a transformative impact on smart grid energy management [19]. Its ability to address complex state spaces, learn sophisticated control strategies, and scale to large environments makes it a powerful tool for optimising energy efficiency and integrating renewable energy sources [20]. As smart grids evolve, DRL's role in ensuring a sustainable and efficient energy future will likely become even more prominent.

III. PROPOSED METHODOLOGY

The proposed methodology aims to improve energy efficiency in smart grids by utilising reinforcement learning (RL)--based control strategies. This section outlines the approach in two parts. First, we present a block diagram illustrating the key components and their interactions within the smart grid environment. This visual representation helps us understand the overall structure and flow of the proposed method. Second, we describe the detailed algorithm for implementing the RL-based control strategy, focusing on enhancing energy efficiency, grid stability, and cost reduction. These components provide a comprehensive framework for deploying advanced control strategies in modern smart grids.

A.   Proposed Block Diagram

The proposed block diagram serves as a visual representation of the RL-based approach for energy-efficient smart grids. This diagram highlights the key components involved in our methodology and their relationships within the smart grid ecosystem. It illustrates how the reinforcement learning agent interacts with the smart grid environment, processes state information, selects actions, and updates its policy through a reward system. By outlining these interactions, the block diagram provides a clear and structured understanding of the system's architecture, data flow, and decisions. This section describes each block's role and explains how they work together for optimal energy management and efficiency.
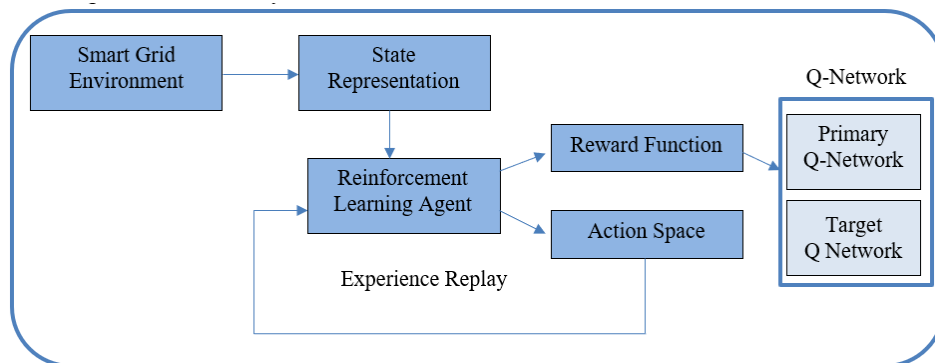

Figure 1. Proposed Block Diagram for RL-based Smart Grid Energy Management

a)   Smart Grid Environment

The smart grid environment encompasses the overall energy system, including distributed energy resources (DERs), energy storage, energy demand, and grid infrastructure. This block represents the context in which the reinforcement learning (RL) agent operates. It includes real-time data from sensors, smart meters, and other monitoring devices.

b)   State Representation

The state representation block collects and organises information from the smart grid environment. This information forms the current "state" the RL agent observes. It typically includes:

➢   Energy Demand: Current and projected energy consumption across different regions or consumers.
➢   Energy Production: Output from DERs, such as solar panels and wind turbines.
➢   Energy Storage: The status of energy storage systems, including capacity and state of charge.
➢   Grid Conditions: Parameters indicating grid stability, load, and other relevant metrics.
        This block's output is the state information that feeds into the RL agent for decision-making.

c)   Reinforcement Learning Agent

The RL agent is the central component responsible for making decisions to optimise energy efficiency in the smart grid. It uses reinforcement learning algorithms based on Deep Q-learning or Double Deep Q-learning. The agent performs the following tasks:

➢   Action Selection: Chooses an action based on the current state, guided by a policy derived from the Q-network.
➢   Experience Replay: Stores transitions (state, action, reward, next state) to improve learning.
➢   Policy Update: Uses the reward function to adjust its policy, aiming to maximise long-term rewards.

d)   Action Space

The action space defines the possible actions the RL agent can take. This block outlines the range of control options available to the agent, such as:

➢   Energy Distribution: Adjusting energy flow among different parts of the smart grid.
➢   Energy Storage Management: Charging or discharging energy storage systems.
➢   Demand Response: Initiating strategies to adjust energy demand based on grid conditions.
        The action space interacts with the RL agent, allowing it to choose optimal actions based on the current state and policy.

e)   Reward Function

The reward function block defines the criteria by which the RL agent's actions are evaluated. It assigns rewards or penalties based on the outcomes of the agent's actions. This block guides the RL agent's learning process and influences its policy. The reward function can consider factors like:

➢   Energy Efficiency: Rewards for maintaining efficient energy distribution.
➢   Cost Reduction: Rewards for minimising energy costs and avoiding expensive energy sources.
➢   Grid Stability: Penalties for destabilising the grid or causing high power imbalances.
        This block provides feedback to the RL agent, helping it learn which actions lead to optimal outcomes.

f)   Experience Replay

Experience replay is a buffer that stores past transitions encountered by the RL agent during training. This block allows the agent to sample from past experiences, promoting stable and efficient learning. It plays a crucial role in the double-deep Q-Learning approach, enabling batch updates and preventing overfitting.

g)   Q-Networks

The Q-networks block contains two separate neural networks: the primary and the target. The primary Q-network guides the RL agent's actions, while the target Q-network provides a stable reference point to reduce oscillations during training. This block is essential for implementing reinforcement learning and allows the RL agent to estimate the quality of state-action pairs.

B.   Proposed Algorithm for Smart Grid Energy Management

Reinforcement Learning (RL) offers a powerful framework for addressing complex decision-making problems, particularly in environments with uncertainty and changing conditions, such as smart grids. Deep Q-Learning, a specific type of RL, extends the traditional Q-Learning approach by integrating deep neural networks to estimate the Q-values of state-action pairs. This enhancement allows Deep Q-Learning to manage large and complex state spaces, making it highly beneficial for energy-efficient control in smart grids.

At its core, Q-learning involves updating a Q-value (quality of an action in a given state) based on a reward signal received after taking an action. The "deep" aspect of Deep Q-Learning comes from using neural networks to approximate these Q-values, allowing for greater flexibility and scalability. This ability to learn optimal control strategies without predefined rules makes Deep Q-Learning well-suited for smart grids, where optimal energy management requires real-time adaptation to dynamic conditions.

Following this approach, we propose an algorithm for energy efficiency in smart grids using a Double Deep Q-learning strategy. This methodology leverages the strengths of reinforcement learning to optimise energy distribution, reduce costs, and maintain grid stability. The following sections detail the proposed algorithm and its implementation within a smart grid environment.

Algorithm 1: Double Deep Q-Learning for Smart Grid Energy Management

Input:
- *D* - experience replay buffer
- *w* - initial network parameters
- *w−* - copy of network parameters for the target network
- N_r - buffer maximum size
- N_b - training batch size
- N_{target} - frequency for replacing the target network.

Output:
- Trained Q-network for smart grid control

Procedure:

1: Initialize the replay buffer D with a maximum size of N_r.

2: Set the initial parameters w for the Q-network and create a copy *w−* for the target Q-network.

3: For each episode e from 1 to M:

  4: Initialize the initial state, s (e.g., grid conditions, energy demand).

  5: While the episode is not over:

    6: Select action a using ε-greedy policy on the Q-network Q (s, a; w).

    7: Execute action a in the environment and receive reward r, next state s', and a terminal flag indicating if the episode ends.

    8: Store the transition (s, a, r, s') in the replay buffer D. If D is full, discard the oldest transition.

    9: Sample a minibatch of size N_b from D.

    10: For each transition (s, a, r, s') in the minibatch:

      11: If s' is terminal, set the target y = r.

      12: Otherwise, calculate the optimal action in the next state and set the target $y = r + \gamma \times Q(s', a'; w^-)$

      13: Perform a gradient descent step on the loss $(y – Q (s, a; w))^2$ to update w.

    14: If the step count is a multiple of N_{target}, update the target network $w^- = w$.

    15: Set s = s'.

  16: End While

17: End For

Following the Double Deep Q-Learning algorithm, the methodology for enhancing energy efficiency in smart grids involves several critical steps. The algorithm begins with initialising an experience replay buffer, which stores transitions (state-action-reward-next state) encountered during training. This buffer enables the reinforcement learning (RL) agent to learn from past experiences and stabilise the training process by avoiding correlation between consecutive transitions.

The training process involves two separate neural networks: the primary Q-network, updated during training, and the target Q-network, which serves as a stable reference. The target network helps prevent oscillations and overfitting by providing a fixed point for calculating target Q-values over a specified number of training steps.

An ε-greedy exploration strategy is used throughout the training process to balance exploration and exploitation. This strategy ensures that the RL agent explores the state space adequately, particularly during early training, while gradually shifting toward more optimal actions as learning progresses.

During each episode, the RL agent interacts with the smart grid environment by selecting actions based on the current state. The agent receives rewards for its actions, representing the impact on energy efficiency, cost, or

other relevant metrics. The agent then uses these rewards to update its policy through gradient descent, minimising the loss between the estimated Q-values and the calculated target values.

Regularly updating the target Q-network at specified intervals ensures stability during training, while the continuous feedback loop allows the RL agent to adapt to changing grid conditions. As the algorithm progresses, it learns to optimise energy distribution, manage energy storage, and effectively initiate demand response, contributing to smart grids' overall energy efficiency.

These foundational elements create a robust methodology for implementing RL-based control strategies in smart grids, focusing on achieving energy efficiency, cost reduction, and grid stability. The next section discusses the specific experiments and results demonstrating the effectiveness of this approach.

IV. DISCUSSION

This section provides details regarding the dataset for implementation purposes. It also explains the implementation details followed by the results collected after running 100 episodes of proposed algorithms.

A.   Dataset Details

The dataset was used from Kaggle (https://www.kaggle.com/datasets/pcbreviglieri/smart-grid-stability) and is designed to simulate the dynamic behaviour and stability of a smart grid system. This dataset offers a rich set of features that capture key aspects of smart grid operations, including the response times of energy producers and consumers, power balances, and price elasticity coefficients. The dataset's structure allows researchers to explore various scenarios and analyse the factors contributing to smart grid stability.

B.   Implementation Setup

This section describes the technical aspects of implementing the reinforcement learning (RL) algorithm for energy-efficient smart grids. The implementation involves several key components, including the environment setup, the RL agent, the neural network architecture, and the training process. Due to its stability and convergence properties, we used the Double Deep Q-Learning (DDQN) approach, with experience replay and a target network for stabilisation.

The smart grid environment was modelled to reflect real-world conditions, incorporating distributed energy resources (DERs), energy storage, and energy demand. The RL agent interacted with this environment, receiving state information such as energy consumption, production, and storage levels. The agent's action space included control options like adjusting power distribution, managing energy storage, and initiating demand response.

The neural network architecture for the RL agent consisted of dense layers with ReLU activation functions designed to handle the complex state space of the smart grid. We applied an ε-greedy policy for exploration, allowing the agent to discover optimal strategies while gradually shifting toward exploitation. The training process involved multiple episodes, each representing a specific operational period in the smart grid. During training, the agent updated its policy based on the reward function, aiming to optimise energy efficiency, cost reduction, and grid stability.

C.   Result Analysis

The results of implementing the reinforcement learning-based method for energy-efficient smart grids were promising. They demonstrated significant improvements in energy distribution and cost reduction. Our analysis focused on several key metrics: energy efficiency, cost savings, and grid stability.

The RL agent successfully learned optimal control strategies, leading to more balanced energy distribution and reduced energy costs. The agent's ability to adapt to changing grid conditions was evident in the results, as it effectively managed energy storage and demand response during peak periods. The substantial energy efficiency gains indicated that the RL-based approach could significantly contribute to a more sustainable smart grid.

The results also showed that the Double Deep Q-Learning approach provided stable and reliable learning, with the agent converging to an optimal policy over multiple episodes, as shown in Figure 2. Experience replays and target networks were crucial in ensuring stability and avoiding oscillations during training.
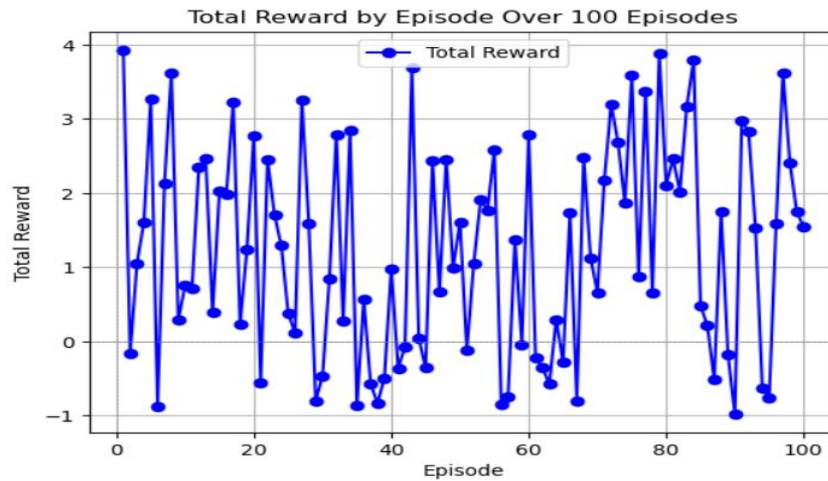
Figure 2. Total Reward by Episode Over 100 Episodes

Implementing the RL-based method for energy-efficient smart grids yielded encouraging results, demonstrating the potential for reinforcement learning to enhance smart grid operations, as shown in Figure 3. The successful application of this approach to a complex environment like the smart grid suggests that RL could play a key role in the future development of energy management strategies. Further research and experimentation are needed to refine the approach and explore its full capabilities.
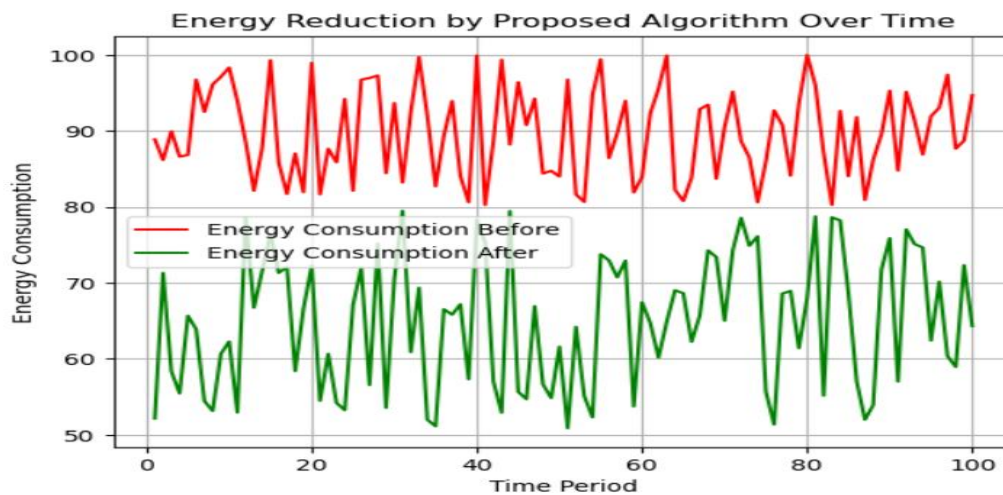

Figure 3. Energy Reduction by Proposed Algorithm over time

Reinforcement learning (RL) offers a powerful approach to energy management in smart grids, providing significant advantages over traditional methods. With its adaptive, data-driven nature, RL can dynamically adjust to changing conditions in real time, leading to improved energy efficiency and reduced energy consumption. Unlike conventional control systems, which rely on static rules or pre-defined schedules, RL-based approaches continuously learn from their environment. This learning capability allows RL agents to optimise energy distribution, manage distributed energy resources (DERs), and minimise energy waste more effectively than traditional methods.

V. CONCLUSION

This paper examined reinforcement learning (RL) to improve energy efficiency in smart grids. The RL-based approach provides a flexible and adaptive solution to the challenges in modern energy management. With reinforcement learning, specifically Double Deep Q-Learning (DDQN), agents can optimise energy distribution, manage distributed energy resources (DERs), and maintain grid stability. These agents continuously learn from their environment, allowing them to make more strategic decisions in real time. The results demonstrated a noticeable improvement in energy efficiency and cost reduction compared to traditional rule-based methods.

Using RL, smart grids can achieve greater energy efficiency, improve demand response, and integrate renewable energy sources more effectively. Reinforcement learning's adaptability offers a robust platform for smart grid control, with the potential for continuous improvement as agents learn and adjust to dynamic conditions.

VI. FUTURE SCOPE

Although this paper highlighted the benefits of RL-based methods for smart grid energy efficiency, there are many opportunities for further research and development. Exploring multi-agent reinforcement learning (MARL) can extend the concept to more complex smart grid environments with multiple interacting agents, enhancing scalability and coordination. Developing hybrid approaches integrating RL with other machine learning techniques, such as supervised or unsupervised learning, could lead to more robust and flexible control strategies, potentially reducing training time and improving convergence.

Research into real-time learning and adaptability could yield reinforcement learning algorithms capable of continuously optimising smart grids. This would allow for more responsive energy management and better adaptation to changing grid conditions. Broadening the application of RL to other smart grid functions, such as grid security, fault detection, and predictive maintenance, could expand its impact.

Examining RL-based energy management's policy and regulatory implications may offer insights into how these technologies could be adopted within existing energy frameworks. By addressing these future research directions, the potential of reinforcement learning in smart grids could be fully realised, paving the way for a more sustainable, efficient, and adaptable energy infrastructure.

REFERENCES

[1]   Kim S, Lim H. Reinforcement learning based energy management algorithm for smart energy buildings. Energies. 2018 Aug 2;11(8):2010.

[2]   Al-Saadi M, Al-Greer M, Short M. Reinforcement learning-based intelligent control strategies for optimal power management in advanced power distribution systems: A survey. Energies. 2023 Feb 6;16(4):1608.

[3]   Salazar EJ, Jurado M, Samper ME. Reinforcement learning-based pricing and incentive strategy for demand response in smart grids. Energies. 2023 Feb 2;16(3):1466.

[4]   Asuhaimi FA, Bu S, Klaine PV, Imran MA. Channel access and power control for energy-efficient delay-aware heterogeneous cellular networks for smart grid communications using deep reinforcement learning. IEEE Access. 2019 Sep 6;7:133474-84.

[5]   Liu Y, Zhang D, Gooi HB. Optimisation strategy based on deep reinforcement learning for home energy management. CSEE Journal of Power and Energy Systems. 2020 Apr 6;6(3):572-82.

[6]   Lu R, Bai R, Luo Z, Jiang J, Sun M, Zhang HT. Deep reinforcement learning-based demand response for smart facilities energy management. IEEE Transactions on Industrial Electronics. 2021 Aug 19;69(8):8554-65.

[7]   Wan Y, Qin J, Yu X, Yang T, Kang Y. Price-based residential demand response management in smart grids: A reinforcement learning-based approach. IEEE/CAA Journal of Automatica Sinica. 2021 Oct 20;9(1):123-34.

[8]   Han G, Lee S, Lee J, Lee K, Bae J. Deep-learning-and reinforcement-learning-based profitable strategy of a grid-level energy storage system for the smart grid. Journal of Energy Storage. 2021 Sep 1;41:102868.

[9]   Wang J, Gao Y, Li R. Reinforcement learning based bilevel real-time pricing strategy for a smart grid with distributed energy resources. Applied Soft Computing. 2024 Apr 1;155:111474.

[10]   Zhang X, Lu R, Jiang J, Hong SH, Song WS. Testbed implementation of reinforcement learning-based demand response energy management system. Applied energy. 2021 Sep 1;297:117131.

[11]   Fu Q, Han Z, Chen J, Lu Y, Wu H, Wang Y. Applications of reinforcement learning for building energy efficiency control: A review. Journal of Building Engineering. 2022 Jun 1;50:104165.

[12]   Dulari Bhatt, Chirag Patel, Madhuri Chopade, Madhvi Dave and Chintan Patel, "Modified Deep Neural Network for Object Recognition" International Journal of Advanced Computer Science and Applications(IJACSA), 14(10), 2023. http://dx.doi.org/10.14569/IJACSA.2023.0141046.

[13]   P. S. Patil, Surekha Janrao, Ajay D Diwate, Madhuri A. Tayal, Pradip Ram Selokar, Amol A. Bhosle, "Enhancing Energy Efficiency in Electrical Systems with Reinforcement Learning Algorithms", Journal of Electrical Systems (JES), Vol. 20 No. 1s (2024), https://doi.org/10.52783/jes.767.

[14]   J. P. R and A. S. Pillai, "Enhancing Energy Efficiency of Intensive Computing Applications using Approximate Computing," 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2022, pp. 1231-1236, doi: 10.1109/ICESC54411.2022.9885429.

[15]   A. Alwarafy, M. Abdallah, B. S. Ciftler, A. Al-Fuqaha and M. Hamdi, "The frontiers of deep reinforcement learning for resource management in future wireless helmets: Techniques challenges and research directions", IEEE Open Journal of the Communications Society, 2022.

[16] Yang T, Zhao L, Li W, Zomaya AY. Reinforcement learning in sustainable energy and electric systems: A survey. Annual Reviews in Control. 2020 Jan 1;49:145-63.

[17] Cao D, Hu W, Zhao J, Zhang G, Zhang B, Liu Z, Chen Z, Blaabjerg F. Reinforcement learning and its applications in modern power and energy systems: A review. Journal of modern power systems and clean energy. 2020 Nov;8(6):1029-42.

[18] Perera AT, Wickramasinghe PU, Nik VM, Scartezzini JL. Introducing reinforcement learning to the energy system design process. Applied Energy. 2020 Mar 15;262:114580.

[19] Yu L, Qin S, Zhang M, Shen C, Jiang T, Guan X. A review of deep reinforcement learning for smart building energy management. IEEE Internet of Things Journal. 2021 May 10;8(15):12046-63.

[20] Chen X, Qu G, Tang Y, Low S, Li N. Reinforcement learning for selective key applications in power systems: Recent advances and future challenges. IEEE Transactions on Smart Grid. 2022 Feb 25;13(4):2935-58.