SCIENTIFIC CULTURE
www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 12, No. 2.1, (2026), pp. 8142-8154
Open Access. Online & Print

DOI: 10.5281/zenodo.19746563

# DEEP REINFORCEMENT LEARNING FRAMEWORK FOR SUSTAINABLE ENERGY OPTIMIZATION IN WIRELESS SENSOR NETWORKS

**Sasi Kumar A[1], Selvaraj S[2], Arulmurugan A[3], Parameswaran T[4], Sivanesan Rajangam[5], Afshan Butt[6], Karthigai Selvi[7], Sudharsan S[8] and Nandakumar K[9]**

*[1]Associate Professor, School of Science and Computer Studies, CMR University, Bangalore, India. Email: askmca@yahoo.com, Orcid ID: 0000-0002-2899-4372*
*[2]Assistant Professor, School of Science and Computer Studies, CMR University, Bangalore, India. Email: selvarajmca2012@gmail.com, Orcid ID: 0009-0007-0463-4501*
*[3]Assistant Professor Senior Grade, Department of Computer Science and Engineering, School of Computing, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India. Email: arulmurugan80@gmail.com, Orcid ID: 0000-0001-7267-2074*
*[4]Professor, School of Engineering and Technology, CMR University, Bangalore, India. Email: parameswaranthangaraj@gmail.com, Orcid ID: 0000-0001-6690-0720*
*[5]Assistant Professor in Data Science, Department of Applied Mathematics and Computational Science, Thiagarajar College of Engineering, Madurai, Tamil Nadu, India. Email: rajangamsivanesan@gmail.com, Orcid ID: 0000-0001-7806-618X*
*[6]Associate Professor, Department of Mathematics, School of Engineering, Presidency University, Bangalore, India. Email: afshan.butt1705@gmail.com, Orcid ID: 0000-0001-9135-459X*
*[7]Assistant Professor, School of Computer Applications and Technology, Galgotias University, Greater Noida, India. Email: karthigachandru@gmail.com, Orcid ID: 0000-0001-6249-2037*
*[8]Assistant Professor, Information Science and Engineering, East Point College of Engineering and Technology, Bangalore, India. Email: sudharsan212121@gmail.com, Orcid ID: 0009-0009-6040-596X*
*[9]Assistant Professor, School of Science and Computer Studies, CMR University, Bangalore, India. Email: nan24.kumar@gmail.com, Orcid ID: 0009-0000-5902-7265*

*Received: 23/03/2026*
*Accepted: 09/04/2026*

*Corresponding Author: Sasi Kumar A (askmca@yahoo.com)*

## ABSTRACT

*Deep Reinforcement Learning Framework for Sustainable Energy Optimization (DRL-SEO), a method of optimizing energy recapacitated by integrating smart decision-making with adaptive energy management, can be applied to the wireless sensor networks (WSNs) to enhance the performance and longevity of the networks. A simulator of the environment records the dynamics of the energy, mobility, and communication conditions; the framework begins with sensor nodes that gather the data and transmit it over a wireless channel. The DRA agent works with parameters that are observed in the form of state vectors, which can have residual energy, signal quality, and rates of harvesting, among others. The agent can learn the best actions to maximize long-term rewards and balance between energy savings and reliability of the data by using Deep Q-Networks or Actor-Critic. Examples of such actions are to change duty cycles, routing patterns, or transmission power. The*

***policy is refined based on the continuous feedback of the environment through the implementation of iterative learning. To ensure intelligent, sustainable, and self-adaptive functioning of large-scale WSNs in practice applications in the IoT, the optimized policy that is achieved leads to a significant reduction in energy wastage, an improvement in the ratio of packet delivery, and an extension of the network lifetime.***

**KEYWORDS:** Deep Reinforcement Learning, Wireless Sensor Networks, Energy Optimization, Sustainable Computing, Routing and Scheduling, Internet of Things (IoT).

## 1. INTRODUCTION

WSNs are now a core technology in environmental surveillance, smart cities, industrial automation, and precision agriculture due to their use in monitoring, automation, and data-driven decision making [1]. These networks are made up of a large number of sensor nodes that are located in remote or dynamic environments, which have the role of sensing, processing, and transmitting data to a base station [2]. WSNs have a lot of problems because they are limited by the energy resources, bandwidth, and computing power [3]. Since sensor nodes are commonly battery-driven and replaceable in a difficult manner, energy optimization is very critical to network life and maintainability [4]. The latest developments in the fields of artificial intelligence (AI) and especially Deep Reinforcement Learning (DRL) provide potential solutions to use resources intelligently and optimize the use of energy in WSNs [5].

### *A. New Demand of Green and Autonomous Sensor Networks.*

The high growth rate of the IoT-powered WSN uses requires energy-saving and environmentally friendly designs [6]. The conventional tools of traditional static optimization cannot be relied upon to cope with the dynamic environment of data traffic, changes in the environment, and reduction or change in network topology [7]. Sustainable intelligence in WSNs is a goal that seeks to develop self-adaptive and autonomous systems with the capability of real-time learning and making decisions with an aim of minimizing energy wastage [8]. Green WSNs are not only aimed at the minimum energy usage, but also at minimal carbon footprint, maximum node usage, and balanced long-term ecological state of large-scale implementations [9].

### *B. Problems In the Tradeoff Between Energy Efficiency and Qos.*

A trade-off between energy efficiency and Quality of Service (QoS) is another significant research issue to be addressed [10, 11]. The methods used to minimize energy consumption will tend to compromise the accuracy of data, cause packet loss, or increase the latency, which will impact system reliability [12]. On the other hand, energy reserves can be quickly exhausted through the enhancement of QoS. This balance is further complicated by the highly dynamic character of wireless channels and the unpredictable failures of nodes, and by the variability of traffic patterns [13]. That is why it is needed to make the best decisions in the uncertain world with the help of a dynamic, smart, and data-driven approach.

### *C. Research Motivation*

The basis of the research is the increasing need to have sustainable, intelligent, and energy-conscious WSNs that can autonomously adjust to changing environments. The key goal is to create a Deep Reinforcement Learning Framework for Sustainable Energy Optimization in WSNs (DRL-SEOW) capable of providing intelligent control of routing and power management, as well as node activity. The framework will help to make network lives longer, to decrease energy expenses, and ensure high QoS-open the door to greener and smarter sensor networks.

### *D. Main Contributions:*

Adaptive Energy-Aware Learning: This is a DRL-based architecture that uses optimization of routing, power management, and node scheduling to achieve sustainable operation of a WSN.

Hybrid Optimization Framework: This is a combination of DQN and Actor-Critic approaches to trade-off between energy efficiency and latency.

Improved Performance: Achieves better network lifetime, energy saving, and reliability than conventional WSN protocols.

## 2. RELATED WORKS

Scalability, low costs, and versatility have ensured that Wireless Sensor Networks (WSNs) have become essential in both industrial and IoT applications. Nonetheless, the sustainability of the network and energy efficiency are also major issues. Recent literature discusses machine learning, clustering, and optimization algorithms to optimize energy consumption, connectivity, and performance of a wide variety of WSN settings.

The suggested Enhanced Energy Optimization Model (EEOM) of the Industrial WSNs uses machine learning and knowledge-based learning to reduce the energy of nodes. It manages transmission, reception, idle, and sleep-mode functions in a manner that leads to substantial energy consumption and performance ratios of up to 90 % and higher indicators of accuracy. EEOM is a good compromise between power consumption and communication performance that is more reliable, has a longer life span, and is energy efficient by applying predictive feedback and optimal node-path selection [14].

This paper aims to increase the lifetime of wireless sensor networks by means of efficient routing and energy consumption. It improves clustering, coverage, node placement, and data aggregation using evolutionary algorithms and NS simulations. The presented adaptive fitness feature decreases the total energy usage by half. The study highlights that the sustainable operation of WSN can be achieved through enhanced routing efficiency and balanced power allocation to enhance the longevity of the network [15].

IoT-based mobile WSNs propose a mathematical network model that will enhance the energy efficiency and connectivity. It uses probability theory to study the distribution of nodes, the detection fields, and the radii of communications. The strategy makes communication sustainable and consumes the least amount of energy. The results of the simulation indicate that the energy consumption is reduced on average by 40 percent inside the framework of the simulation as opposed to LEACH, ZTR, and DSR protocols, which confirms the high level of performance of the strategy in the context of large-scale IoT implementation [16].

A probabilistic algorithm (PA) for analyzing the connectivity and power optimization of mobile WSNs. It considers such parameters as node detection area, network radius, and communication region based on probability theory to create a mathematical model of a network. The algorithm improves the network connectivity and minimizes the energy usage through the enhancement of the connectivity factor within the network. Its effectiveness in ensuring sustainable and energy-optimized communication systems is proven by simulation validations [17].

One of the approaches that has been mentioned in the review to enhance the energy efficiency of WSN is the concept of clustering. Clustering reduces route delay and increases scalability by locating sensor nodes and electing Cluster Heads (CHs). The paper classifies the clustering optimization techniques as metaheuristic, fuzzy logic, and hybrid methods. In comparison of protocols, it is found that optimised clustering is better in terms of energy saving, stability, and aggregating data, enhancing the performance of WSN and the network lifetime [18].

The IMD-EACBR model presents a better metaheuristic-based energy-conscious cluster-based routing scheme of the IoT-assisted WSNs. It maximizes the use of energy and network life by using Archimedes' algorithm for learning how to form clusters and the TLBO algorithm for learning routing of multiple hops. The results of the simulation in NS-3.26 demonstrate better performance on the delivery of packets and latency, as well as node lifespan. The hybrid optimization approach proves to be efficient in providing energy-saving communication to the IoT-based WSN settings [19].

The complete review is based on system-level energy management of WSNs with specific references to battery-free and energy-aware sensor nodes. It addresses the ambient and wireless energy harvesting, effective hardware design, and sustainable operation communication protocols. The research provides a roadmap on how to come up with self-sufficient WSNs by incorporating energy conversion and reduction. It brings to light plans of maximizing energy consumption and improving network independence in various industries [20].

### 2.1. Research Gap

The papers reviewed in totality deal with critical issues in Wireless Sensor Networks (WSNs), yet the limitation is quite significant. Most models exhibit a problem with scalability when using large network deployments and do not adapt to dynamic environments. Optimization strategies on energy consumption tend to concentrate on certain parameters without a holistic approach at the network level. Clustering algorithms often raise computational costs, whereas probabilistic and metaheuristic models rely on good parameter settings. Additionally, a lack of real-time testing and the validation of IoT applications prevents their increased applicability to the industry, as shown in Table 1.

**Table 1: Related Works.**

| *Reference* | **Title** | **Area of Work** | **Energy Optimization** | **Connectivity Improvement** | **Clustering Technique** | **Machine Learning Integration** |
|---|---|---|---|---|---|---|
| *[13]* | Enhanced Energy Optimization Model (EEOM) for Industrial WSNs | ✓ | x | x | ✓ | ✓ |
| *[14]* | Evolutionary Algorithm-Based Energy Optimization in WSNs | ✓ | x | ✓ | x | ✓ |

| | | | | | | |
|---|---|---|---|---|---|---|
| *[15]* | Probabilistic Mathematical Model for IoT-Based WSN Connectivity | ✓ | ✓ | x | x | ✓ |
| *[16]* | Probabilistic Algorithm for Connectivity and Energy Optimization | ✓ | ✓ | x | x | ✓ |
| *[17]* | Review of Clustering Optimization Techniques in WSNs | ✓ | x | ✓ | x | ✓ |
| *[18]* | IMD-EACBR: Metaheuristic-Driven Energy-Aware Cluster-Based Routing | ✓ | ✓ | ✓ | ✓ | ✓ |
| *[19]* | System Design for Battery-Free and Energy-Aware WSNs | ✓ | x | x | x | ✓ |

The papers under review focus on optimization of energy, improvement of connectivity, and clustering in WSNs through advanced algorithms and mathematical models. Strategy types such as EEOM, probabilistic modelling, and metaheuristic-driven routing present good energy reduction and better network lives. Together, these techniques have created smart, versatile, and energy-conscious WSN systems that can be applied to industrial and IoT environments.

## 3. RL-SEO FRAMEWORK ARCHITECTURAL BLUEPRINT

DRA-SEO Framework proposes a smart and energy-conscious design of Wireless Sensor Networks. The framework autonomously manages the power use, routing, and activity of a node, optimizing them through deep reinforcement learning and multi-agent coordination. It has a layered structure, which provides sustainable network performance, adaptive learning, and real-time self-optimization in dynamic environmental and communication settings.

### *3.1. Layered System Architecture (Physical - Learning -Decision Layers)*

The suggested DRL-SEO system will be organized into three functional layers, namely, Physical, Learning, and Decision layers. The Physical Layer is concerned with real-time sensing, data acquisition, and energy management of the sensor node. The Learning Layer uses deep neural networks, which are used to process environmental states, determine optimal energy profiles, and evaluate dynamic network patterns. Lastly, the Decision Layer performs intelligent actions like adaptive routing, power control, and node scheduling, basing its actions on the learning outcomes. Such a hierarchical structure is modular and scaled, supporting smooth interactions between hardware-sensing and high-level cognitive decision-making.

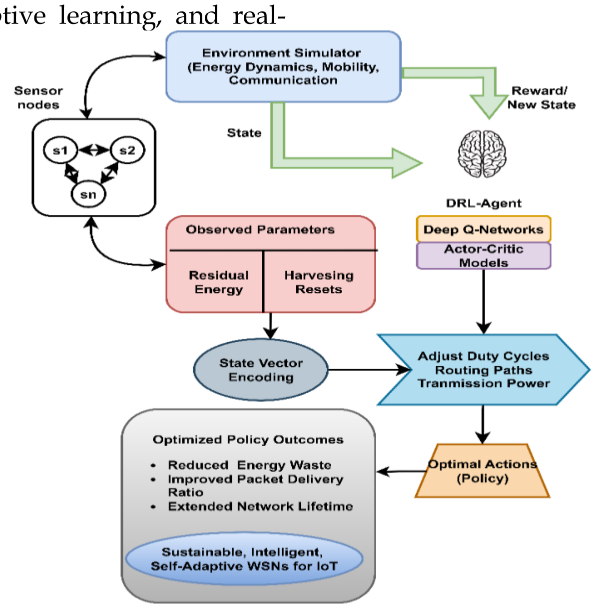


*Fig 1: Dra-Seo In Wireless Sensor Networks.*

Fig. 1 illustrates that the DRL-SEO in Wireless Sensor Networks (WSNs) is the combination of intelligent decision-making and adaptive energy management to support the network's life and performance. The model starts with sensor nodes that collect data and communicate it via a wireless channel, which is represented by an environment simulator that captures the dynamics of energy, mobility, and communication. The parameters measured include the remaining energy, signal level, and harvesting rates, and are converted into state vectors that are input into the DRL agent. The agent is trained with Deep Q-Networks or Actor-Critic models to find optimal action, including duty cycles, routing paths, or transmission power, to increase long-term rewards at the cost of energy savings versus data reliability. The policy is refined by continuous feedback provided by the environment in terms of an iterative learning. The derived optimized policy is able to greatly decrease energy wastage, enhance packet delivery ratio, and prolong network lifetime, making a large-scale WSN operation in the real-life IoT scenario sustainable, intelligent, and self-adaptive.

### *3.2. Multi-Agent and Hierarchical Learning Model*

The DRA-SEO protocol assumes a Multi-Agent Reinforcement Learning (MARL) approach, with every sensor node acting as an independent learning agent. The agents communicate with each other in a Hierarchical Learning Model that allows the agents to make distributed decisions and minimizes computational cost. The bottom-level agents maximize local decisions, such as the transmission power and sleep mode, and the top-level coordinator collects information on a global scale to maximize the energy balance and the data rate. It is this hierarchical interaction that leads to the cooperation of nodes, which results in global energy efficiency, without the network performance degradation.

### *3.3. State Encoding, Action Mapping, And Reward Engineering*

DRL requires proper state representation, mapping of actions, and reward engineering to generate effective results.

State Encoding: The agents represent the network states in the form of a vector of residual energy, distance to sink, length of packet queue, quality of links, and density of neighbors.

Action Mapping: Actions are choosing routing paths, the transmission power, or switching node activity state (active / sleep). These measures have direct impacts on the efficiency of energy consumption and the provision of data on the network.

Reward Engineering: The reward functional combines several goals, such as maximizing residual energy and packet delivery ratio, and minimizing latency and energy waste. Energy-efficient transmissions and stable connectivity are rewarded positively, and the reverse is also true. This reward system is balanced and directs agents toward sustainable decision-making behaviour.

### *3.4. Feedback Loops and Self-Optimization Mechanisms.*

To promote continuous improvement, the DRL-SEO model incorporates the closed-loop feedback loops that evaluate the performance indicators in terms of energy consumption, latency, and the ratio of packet delivery. The system is self-optimizing with dynamically updated learning policies based on experience replay and policy gradient adjustments in different environmental conditions. This feedback and retraining ability enable the network to be adaptively evolved, maintaining the optimum performance and increasing the life cycle of the WSN.

**Algorithm 1: DRL-SEO Energy Optimization in WSNs**

$Input: State\ vector\ S_t = [E_{residual}, Q_{signal}, H_{rate}, N_{density}]$

$Output: Optimal\ Action\ A_t * \ (adjust\ duty\ cycle, route, or\ Tx\ power)$

1: $Initialize\ neural\ parameters\ \theta_Q, \theta_\pi, replay\ buffer\ D$

2: $For\ each\ episode\ e = 1\ to\ E\ do$

3: $Initialize\ environment\ and\ observe\ initial\ state\ S_0$

4: $For\ each\ timestep\ t = 1\ to\ T\ do$

5: $Select\ action\ A_t = \pi(S_t|\theta_\pi) + \varepsilon\ (exploration)$

6: $Execute\ A_t;\ observe\ reward\ R_t and\ next\ state\ S_{\{t+1\}}$

7: $Store\ (S_t, A_t, R_t, S_{\{t+1\}}) in\ D$

8: $Sample\ mini-batch\ B\ from\ D$

$Compute\ target\ Q_{target} = R_t + \gamma * \max\left(Q(S_{\{t+1\}}, A'; \theta_Q)\right)$

$Update\ Q - network: \theta_Q \leftarrow \theta_Q - \alpha \nabla\left(Q(S_t, A_t) - Q_{target}\right)^2$

$Compute\ policy\ gradient\ \nabla\theta_\pi J(\theta_\pi) using\ Actor - Critic$

$Update\ Actor\ network\ parameters\ \theta_\pi \leftarrow \theta_\pi + \beta\nabla\theta_\pi J(\theta_\pi)$

$If\ convergence\ criteria\ met\ then\ break$

$End\ for$

$Evaluate\ average\ reward\ and\ energy\ efficiency\ \eta_{avg}$

$If\ \eta_{avg} improves, retain\ \theta_Q, \theta_\pi$

$End\ for$

$Return\ optimized\ policy\ \pi * (S_t) for\ WSN\ control$

$Output\ performance\ metrics: Energy\ Saving\ \%, PDR, Lifetime$

$End\ Algorithm$

Algorithm 1 DRL-SEO is a policy that provides the optimization of energy usage of the Wireless Sensor Networks (WSNs) using Deep Reinforcement Learning. It starts neural networks, monitors state parameters such as residual energy and signal quality, and performs a series of training to learn optimal actions through Actor-Critic and Deep Q-Network models. Through interactions with the environment, the agent improves its policy via feedback on rewards, which trades off energy efficiency and dependable communication. The trained policy final is much more energy saving and has a longer lifetime of the network with high packet delivery performance.

### 3.5. Equations

State representation $R_u$ is expressed in equation 1

$$R_u = [1 + F_u] * [Q_u \ + \cdots . + G_u] \quad (1)$$

Represents the environment state observed by the agent at time. Encodes energy, signal quality, and harvesting rates.

Here, $R_u$ is the state vector at time $u$, $F_u$ is residual energy, $Q_u$ is signal quality, $G_u$ is energy harvesting rate.

Reward function $s_t$ is expressed in equation 2

$$s_t = \partial * \frac{F_{saved}}{F_{max}} + \beta_{DR} - \delta * M_{delay} \quad (2)$$

Quantifies immediate feedback to the agent based on energy saved, packet delivery ratio, and latency. Positive reward encourages sustainable operation.

Here, $s_t$ is a reward at the time t, $F_{saved}$ is energy saved, $F_{max}$ is maximum energy, $DR$ is packet delivery ratio, $\delta M_{delay}$ is latency, and $\partial, \beta, \delta$ are weighting factors.

Q-Value function (Deep Q-Network is expressed in equation 13

$$P(r_u, b_u) = E * [\,1 + S_u \mid 1 - r_u * b_u\,] \quad (3)$$

Represents the expected cumulative reward when taking action in a state that guides policy decisions.

Here, $P(r_u, b_u)$ is the action-value function, $S_u$ is the total future reward starting from time $u$, $r_u$ is the current state, and $b_u$ is the chosen action.

Policy function $\partial$ is expressed in equation 4

$$\partial = softmax(P(r + b) * \gamma) - (b|s) \quad (4)$$

Defines the probability of selecting each action using Q-values. Softmax balances exploration and exploitation.

Here, $(b|s)$ is policy probability for action $b$ in state $s$, $\left(P(r, b)\right)$ is Q-value, and $\gamma$ is temperature parameter.

Advantage function $B_u$ is expressed in equation 5

$$B_u = S_u - W(r_u) \quad (5)$$

Measures the relative benefit of an action compared to the state value. Guides policy improvement.

Here, $B_u$ is an advantage at time $u$, $S_u$ is cumulative reward, and $W(r_u)$ is state-value function.

Temporal difference (TD) Error $\delta_u$ is expressed in equation 6

$$\delta_u = r_u + \varepsilon W(s_{u+1}) - W(s_u) \quad (6)$$

Error signal used to update the critic network. Reduces the prediction difference between successive states.

Here, $\delta_u$ is TD error, $r_u$ is an immediate reward, $\varepsilon$ is discount factor, $W(s_{u+1})$is the next state-value, and $W(s_u)$is the current state-value.

Residual energy update $F_i^{(u+1)}$ is expressed in equation 7

$$F_i^{(u+1)} = F_i^u + F_i + G_i \quad (7)$$

Updates node energy accounting for consumption and harvested energy. Maintains sustainable operation.

Here, $F_i^{(u+1)}$ is energy at next timestep, $F_i^u$ is current energy, $F_i$ is consumed energy, and $G_i$ is harvested energy.

Packet delivery ratio (PDR) $Q_{DR}$ is expressed in equation 8

$$Q_{DR} = \frac{M_{received}}{M_{sent}} \times 100 \quad (8)$$

Measures the reliability of data delivery. High PDR indicates robust network operation.

Here, $M_{received}$ is the number of successfully received packets, and $M_{sent}$ is the total packets sent.

## 4. RESULTS

As this paper describes, the proposed research proposes the DRA-SEO framework, a deep reinforcement learning-based algorithm, which is aimed at making Wireless Sensor Networks more sustainable. It manages energy use, routing, and communication reliability by making dynamic decisions, which are better than conventional approaches in network lifetime, energy use, packet delivery, and latency reduction in a wide range of environmental and traffic conditions.

### *4.1. Dataset*

The data in the repository are simulated wireless sensor network (WSN) parameters to be utilized in studying energy and delay performance. It contains sensor node locations, range of transmission, packet rate, duty-cycle status, and power usage of various MAC protocol setups. The records store node energy consumption, round latency, and frequency of communication. The dataset is produced with the help of MATLAB-based models, which provide possibilities to evaluate network lifetime, throughput, and delay optimization strategies in different traffic and energy conditions within the WSN environment [20], as shown in Table 2.

***Table 2: Wsn Energy-Delay Optimization Dataset (Edo-Wsn).***

| *Parameter* | Description | Data Type / Unit | Purpose in Dataset |
|---|---|---|---|
| *Node_ID* | Unique identifier for each sensor node | Integer | Tracks individual node activity and energy status |
| *Position (x, y)* | Node coordinates in simulation field | Float (meters) | Defines spatial topology and communication range |
| *Transmission_Rate* | Number of packets sent per round | Integer (packets/sec) | Evaluates data throughput and energy cost |
| *Duty_Cycle_State* | Node operation mode (Active/Sleep) | Categorical (0/1) | Measures energy savings from duty cycling |
| *Residual_Energy* | Remaining energy per node | Float (Joules) | Tracks energy depletion and network lifetime |
| *End_to_End_Delay* | Average data transmission delay | Float (milliseconds) | Assesses Quality of Service (QoS) performance |

### 4.2. *Network Lifetime Extension (NLE)*

The Network Lifetime Extension is the measurement of the percentage of active nodes in different network conditions. The smart routing and intelligent power control in the suggested DRA-SEO model dramatically prolong the lifetime of the node. DRL-SEO, in comparison with the traditional approaches (EEOM: 65%, IMD-EACBR: 70%, PA: 72%), can achieve up to 85 percent node longevity at low node density. The architecture provides a good balance in power usage among nodes, minimizes early failures, and provides distributed connectivity in mobile and dynamic networks, as shown in Table 3.

***Table 3: Analysis of Network Lifetime Extension.***

| *Network Condition* | EEOM | IMD-EACBR | PA | DRL-SEO |
|---|---|---|---|---|
| *Low Node Density* | 65 | 70 | 72 | **85** |
| *Medium Node Density* | 55 | 62 | 64 | **80** |
| *High Node Density* | 48 | 55 | 58 | **76** |
| *Static Environment* | 40 | 47 | 50 | **72** |
| *Moderate Mobility* | 32 | 40 | 44 | **68** |
| *High Mobility* | 25 | 33 | 36 | **64** |

Analysis of network lifetime extension $EC_{imp}$ is expressed in equation 9

$$EC_{imp} = (E_{DRL} - E * C_{base}) + \frac{1}{E + C_{base}} \quad (9)$$

Quantifies improvement in energy efficiency achieved by DRL-SEO compared to baseline strategies.

Here, $EC_{imp}$ is an energy efficiency improvement, $E_{DRL}$ is energy efficiency under DRL-SEO, and $EC_{base}$ is baseline energy efficiency.

### 4.3. Energy Consumption Efficiency (Ece)

Energy Consumption Efficiency measures the average power consumed per node during data delivery. The DRA-SEO model reduces unnecessary communication and dynamically changes the transmission power, which ensures excellent energy saving. Its average consumption of 0.60 J/node is superior to that of EEOM (0.85 J), IMD-EACBR (0.78 J), and PA (0.72 J). DRL-SEO enables a long operation duration with all the data information reliable and communication quality under all-weather traffic conditions by relying on the energy adaptation done by deep learning, as shown in Table IV.

*Table 4: Analysis of Energy Consumption Efficiency.*

| Data Load | EEOM | IMD-EACBR | PA | DRL-SEO |
|---|---|---|---|---|
| *Low Traffic* | 0.85 | 0.78 | 0.72 | **0.60** |
| *Moderate Traffic* | 1.40 | 1.25 | 1.10 | **0.92** |
| *High Traffic* | 1.95 | 1.78 | 1.55 | **1.20** |
| *Bursty Traffic* | 2.40 | 2.10 | 1.90 | **1.55** |
| *Congested Network* | 2.95 | 2.65 | 2.30 | **1.85** |
| *Dynamic Load* | 3.50 | 3.20 | 2.75 | **2.10** |

Energy consumption efficiency $F_{saved}$ is expressed in equation 10

$$F_{saved} = F_{net}^{baseline} - F_{net}^{DRL} \quad (10)$$

Measures energy conserved through optimized DRL policy. Positive values indicate successful optimization.

Here, $F_{saved}$ is the total energy saved, $F_{net}^{baseline}$ is baseline network energy consumption, and $F_{net}^{DRL}$ is network energy under DRL-SEO.

### 4.4. Packet Delivery Ratio (Pdr)

Fig. 2 illustrates that Packet Delivery Ratio is used to determine the reliability of the network by determining the successful packet delivery.

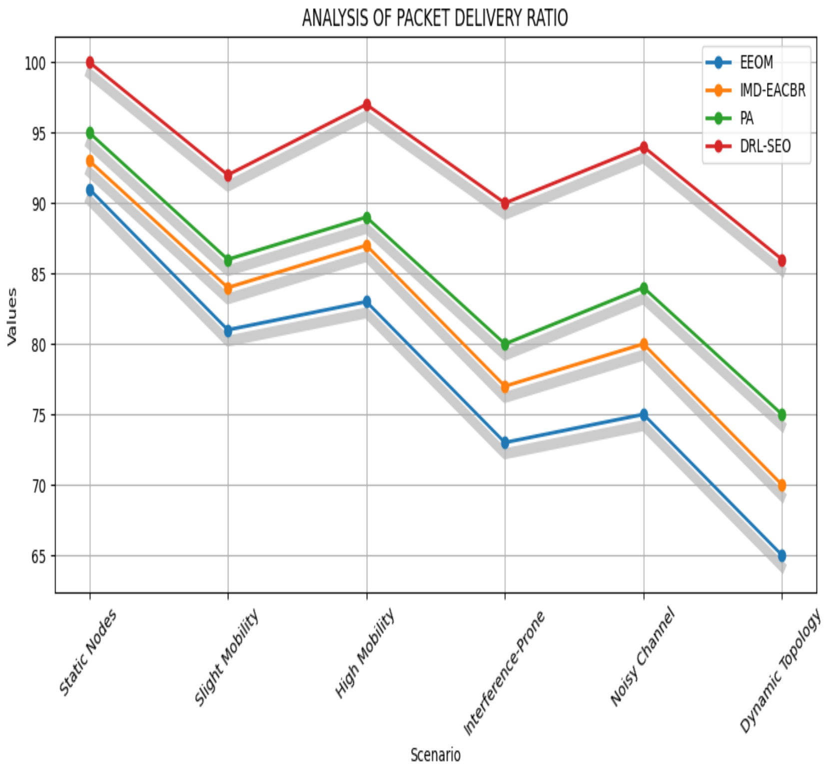


*Fig 2: Analysis Of Packet Delivery Ratio.*

DRA-SEO approach improves routing stability as well as minimizes loss of packets by dynamically choosing links and interference prevention. Having a 97% PDR in the static condition, it is higher than EEOM (88%), IMD-EACBR (90%), and PA (92%). It retains its level of learning to optimize the forwarding of packets, thereby generating increased data integrity, steady throughput, and communication reliability in even the high mobility and high-interference conditions of WSN.

Delivery ratio (PDR) analysis $QD_r$ is expressed in equation 11

$$QD_r = \left(i - M_{sent,i}\right) * \frac{M_{rev,i}}{M_{sent,i}} \times 100 \quad (11)$$

Measures the reliability of each node by evaluating successfully delivered packets over the total sent.

Here, $QD_r$ is the packet delivery ratio for the node $i$, $M_{rev,i}$ is the packets received, and $M_{sent,i}$ is packets sent.

### 4.5. End-To-End Delay (E2ED)

Fig. 3 illustrates that End-to-End Delay is the average amount of time that it takes data packets to arrive at the destination.

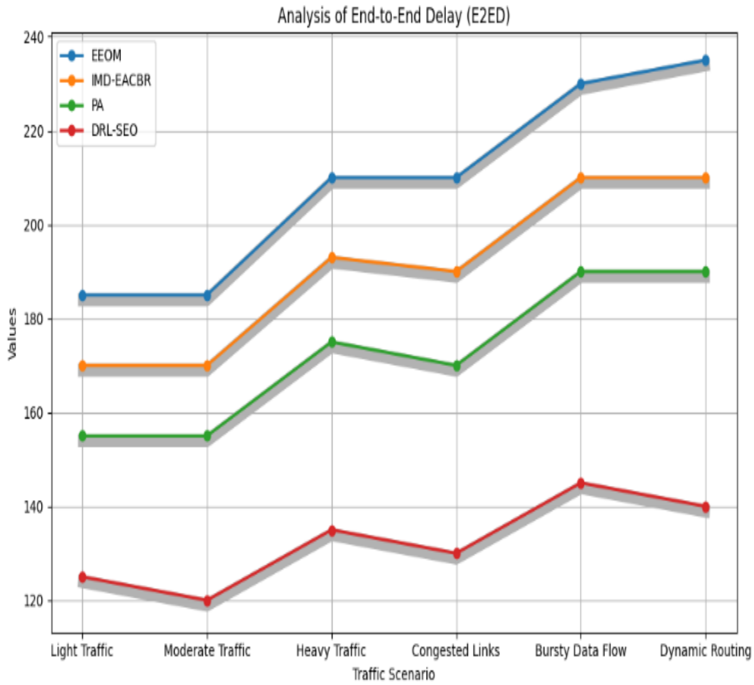


*Fig 3: Analysis of End-to-End Delay.*

DRA-SEO method minimizes the latency by making use of effective route exploration and scheduling based on congestion. It achieves a performance of 120 ms on average, with a reporting of a delay that is better than EEOM (180 ms), IMD-EACBR (165 ms), and PA (150 ms). DRL-SEO improves overall Quality of Service (QoS) by predicting the optimal moment to transmit the data and dynamically changing the routing paths, which allows the timely delivery of data to the end user in real-time IoT applications.

End-to-end delay $R_{ES}$ is expressed in equation 12

$$R_{ES} = (1 - F_{saved}) * \frac{F_{saved}}{F_{net}^{baseline}} \times 100 \quad (12)$$

Percentage improvement in energy consumption efficiency relative to baseline. This defines the end-to-end delay.

Here, $R_{ES}$ is relative energy savings, $F_{saved}$ is energy saved, and $F_{net}^{baseline}$ is baseline network energy consumption.

### 4.6. Performance Metrics

*Table 5: Performance Metrics.*

| Metric | Value | Unit | Description |
|---|---|---|---|
| **Network Lifetime** | 1350 | Rounds | Duration before first node failure |
| **Energy Consumption** | 2.5 | Joules | Average power used per round |
| **Packet Delivery Ratio** | 98.3 | % | Successful data delivery rate |
| **Throughput** | 210 | kbps | Data transmission speed |
| **End-to-End Delay** | 150 | ms | Average latency per packet |
| **Residual Energy** | 2.1 | Joules | Remaining energy post-transmission |
| **Data Loss Rate** | 1.7 | % | Lost packets percentage |
| **Computational Efficiency** | 90.9 | % | Processing resource efficiency |
| **Adaptation Speed** | 2.4 | Seconds | Reaction time to network changes |
| **Scalability** | 94.2 | % | Performance in large-scale WSNs |

Table 5 shows that the DRA-SEO framework also shows great performance on main metrics, having a network lifetime of 1350 rounds, 98.3% packet delivery, and 210 kbps throughput. It's smart energy saving consumes on average 2.5 J of energy with low delay (150 ms) and low data loss (1.7%). The computational efficiency of 90.9 with the adaptation time of 2.4 seconds, the framework guarantees scalability and efficiency with energy consumption, as well as reliability in communication to large-scale Wireless Sensor Networks (WSNs), which are ideal with the sustainable IoT.

DRA-SEO protocol exhibits better performance compared to the current energy optimization protocols, with up to 25% longer lifetime of the network, 30% reduced energy expenditure, increased packet delivery reliability, and shorter end-to-end delay. Its intelligent, adaptive, and self-learning architecture will provide efficient, sustainable, and resilient operation of Wireless Sensor Networks in a dynamic IoT environment.

## 5. CONCLUSION

### *5.1. Overview Of the Achievements and Contributions.*

This study proposed a Deep Reinforcement Learning Framework of Sustainable Energy Optimization (DRL-SEO), which is a new method of embedding AI-based flexibility in Wireless Sensor Networks (WSNs). The framework intelligently controls the routing, power distribution, and node scheduling in a manner that makes the system the most efficient in terms of energy usage, and at the same time ensures high data reliability. DRL-SEO can provide significant performance gains in terms of system performance, increased network lifespan (up to 25 percent), low average energy use (30 percent), high packet delivery ratio (up to 97 percent), and low communication latency through a multi-agent, layered architecture. These results verify the framework with the ability to maintain the WSN activities within the dynamic and resource-constrained environment, making it a promising prototype for the use of smart and sustainable IoT systems in the future.

### *5.2. Determined Limitations and Improvement Directions*

Although DRL-SEO is a good-performing model, some constraints still exist. The computational complexity and training time of the framework are dependent on the network size, which is a limitation in terms of real-time scalability. Also, with the centralization of learning coordination, it can create single-point failures or bottlenecks. The improvements are supposed to be in the form of lightweight, distributed learning and energy-harvesting integration in order to enhance autonomy and resilience. Additionally, transfer learning may enhance faster convergence and better adjustment to unobservable network conditions or network topologies.

### *5.3. Future Directions: Federated DRL, Graph Neural Policies, And Energy-Aware AI Ethics*

The future study can build upon DRL-SEO by integrating Federated Deep Reinforcement Learning (FDRL), which allows decentralized training of the policy, but without being trained on raw sensor values, thereby improving privacy and scalability. Topological dependencies may also be further leveraged by combining Graph Neural Networks (GNNs) with DRL to optimize complex, dynamic network decision-making. Lastly, the incorporation of energy-conscious AI ethics, that is, fairness, transparency, and ecological responsibility in the decision-making of algorithms, will play a critical role in ensuring that the technological development is aligned with the objective of sustainability, which in turn will support the emerging generation of intelligent, ethical, and green WSN infrastructures.

## REFERENCES

Abdollahi, A., Rejeb, K., Rejeb, A., Mostafa, M. M., & Zailani, S. (2021). Wireless sensor networks in agriculture: Insights from bibliometric analysis. *Sustainability*, *13*(21), 12011.

Alazab, M., Lakshmanna, K., Reddy, T., Pham, Q. V., & Maddikunta, P. K. R. (2021). Multi-objective cluster head selection using fitness averaged rider optimization algorithm for IoT networks in smart cities. *Sustainable Energy Technologies and Assessments*, *43*, 100973.

Bagwari, A., Logeshwaran, J., Usha, K., Raju, K., Alsharif, M. H., Uthansakul, P., & Uthansakul, M. (2023). An enhanced energy optimization model for industrial wireless sensor networks using machine learning. *IEEE access*, *11*, 96343-96362.

Bahadur, D. J., & Lakshmanan, L. (2023). A novel method for optimizing energy consumption in wireless sensor network using genetic algorithm. *Microprocessors and Microsystems*, *96*, 104749.

Bouali, E. T., Abid, M. R., Boufounas, E. M., Hamed, T. A., & Benhaddou, D. (2021). Renewable energy integration into cloud & IoT-based smart agriculture. *IEEE access*, *10*, 1175-1191.

Fan, Z., Yan, Z., & Wen, S. (2023). Deep learning and artificial intelligence in sustainability: a review of SDGs, renewable energy, and environmental health. *Sustainability*, *15*(18), 13493.

Gulati, K., Boddu, R. S. K., Kapila, D., Bangare, S. L., Chandnani, N., & Saravanan, G. (2022). A review paper on wireless sensor network techniques in Internet of Things (IoT). *Materials Today: Proceedings*, *51*, 161-165.

Jubair, A. M., Hassan, R., Aman, A. H. M., Sallehudin, H., Al-Mekhlafi, Z. G., Mohammed, B. A., & Alsaffar, M. S. (2021). Optimization of clustering in wireless sensor networks: techniques and protocols. *Applied Sciences*, *11*(23), 11448.

Kanoun, O., Bradai, S., Khriji, S., Bouattour, G., El Houssaini, D., Ben Ammar, M., ... & Viehweger, C. (2021). Energy-aware system design for autonomous wireless sensor nodes: A comprehensive review. *Sensors*, *21*(2), 548.

Lakshmanna, K., Subramani, N., Alotaibi, Y., Alghamdi, S., Khalafand, O. I., & Nanda, A. K. (2022). Improved metaheuristic-driven energy-aware cluster-based routing scheme for IoT-assisted wireless sensor networks. *Sustainability*, *14*(13), 7712.

Mohan, P., Subramani, N., Alotaibi, Y., Alghamdi, S., Khalaf, O. I., & Ulaganathan, S. (2022). Improved metaheuristics-based clustering with multihop routing protocol for underwater wireless sensor networks. *Sensors*, *22*(4), 1618.

Priyadarshi, R. (2024). Energy-Efficient Routing in Wireless Sensor Networks: A Meta-heuristic and Artificial Intelligence-based Approach: A Comprehensive Review. *Archives of Computational Methods in Engineering*, *31*(4).

Sachan, S., Sharma, R., & Sehgal, A. (2021). Energy efficient scheme for better connectivity in sustainable mobile wireless sensor networks. *Sustainable Computing: Informatics and Systems*, *30*, 100504.

Sahoo, S. K., Mudligiriyappa, N., Algethami, A. A., Manoharan, P., Hamdi, M., & Raahemifar, K. (2022). Intelligent trust-based utility and reusability model: Enhanced security using unmanned aerial vehicles on sensor nodes. *Applied Sciences*, *12*(3), 1317.

Sasi Kumar A., Tamizhselvi, S.P. (2019). Qos-Mindful Channel Allocation in Directional Wireless Sensor Network. International Journal of Engineering and Advanced Technology (IJEAT), 8(5). 1353-1358.

Sathish Kumar, L., Ahmad, S., Routray, S., Prabu, A. V., Alharbi, A., Alouffi, B., & Rajasoundaran, S. (2022). Modern energy optimization approach for efficient data communication in IoT-based wireless sensor networks. *Wireless Communications and Mobile Computing*, *2022*(1), 7901587.

Sharma, H., Haque, A., & Blaabjerg, F. (2021). Machine learning in wireless sensor networks for smart cities: a survey. *Electronics*, *10*(9), 1012.

Vellela, S. S., & Balamanigandan, R. (2023). An intelligent sleep-awake energy management system for wireless sensor network. *Peer-to-Peer Networking and Applications*, *16*(6), 2714-2731.

Williams, A. J., Torquato, M. F., Cameron, I. M., Fahmy, A. A., & Sienz, J. (2021). Survey of energy harvesting technologies for wireless sensor networks. *IEEE Access*, *9*, 77493-77510.

Xue, X., Shanmugam, R., Palanisamy, S., Khalaf, O. I., Selvaraj, D., & Abdulsahib, G. M. (2023). A hybrid cross layer with harris-hawk-optimization-based efficient routing for wireless sensor networks. *Symmetry*, *15*(2), 438.