FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 10 No. 16, June, 2026, pp 453 – 460
DOI: https://doi.org/10.33003/fjs-2026-1016-5713

Reinforcement Learning-Based Adaptive Data Rate Control in Mobile LoRaWAN with Imperfect CSI

***[1]Simon Nelson Boyi, [1]Umar Suleiman Dauda, [1]Safiu Abiodun Gbadamosi, [1]James Garba Ambafi and [2]Umar Zangina**

[1]Department of Electrical and Electronic Engineering, Federal University of Technology Minna, Niger State, Nigeria.
[2]Sokoto Energy Research Centre, Usmanu Danfodiyo University, Sokoto State, Nigeria.

*Corresponding authors' email: simon.nelson@futminna.edu.ng

Received: 10 Aug. 2026
Accepted: 24 Aug. 2026
Published: 02 Aug. 2026

ABSTRACT
Low Power Wide Area Networks (LPWANs), particularly Long-Range Wide Area Network (LoRaWAN), are increasingly being deployed in mobile Internet of Things (IoT) applications. In these applications, adaptive data rate (ADR) control is essential for maintaining reliable and energy-efficient communication under dynamic wireless conditions. However, ADR performance remains constrained when channel state information (CSI) is imperfect. Although reinforcement learning (RL)-based ADR methods have demonstrated notable improvement over conventional approaches, their effectiveness under imperfect CSI deployment scenarios remains insufficiently explored. In this paper, we investigate the robustness of RL-based ADR approaches with imperfect CSI for mobile LoRaWAN networks using a mobility-aware simulation framework. Standard ADR, heuristic ADR, Deep Q-Network (DQN)-based ADR, and Proximal Policy Optimisation (PPO)-based ADR are evaluated across varying mobility patterns, network layouts, and channel dynamics. Key performance metrics include packet delivery ratio, latency, runtime efficiency, and retention of baseline capability are evaluated. The Results show that PPO achieves the best performance under imperfect CSI conditions, attaining a mean packet delivery ratio of 0.9954, outperforming standard ADR, heuristic ADR, and DQN by 29.49%, 19.12%, and 1.55%, respectively. In addition, PPO reduced latency by 53.12% and runtime by 81.94% compare to DQN while preserving over 99% of baseline communication reliability under scenario variation. These findings demonstrate that policy-gradient learning provides better transferability, robustness, and deployment readiness compared with conventional and value-based ADR approaches. The study highlights the importance of evaluating intelligent wireless controllers not only for optimisation performance, but also for cross-scenario generalisation before real-world deployment in mobile LPWAN systems.

Keywords: Adaptive Data Rate, Generalisation, LoRaWAN, Mobile IoT, Wireless Networks

INTRODUCTION

Low Power Wide Area Networks (LPWANs), particularly LoRaWAN, have become an important communication platform for large-scale Internet of Things (IoT) systems due to their long communication range, low energy consumption, and cost-effective deployment (Adelantado et al., 2017; Alliance, 2015; Centenaro et al., 2016). Their use is expanding beyond static sensing applications toward increasingly mobile scenarios such as fleet monitoring, logistics tracking, wearable sensing, precision agriculture, and intelligent transportation systems (Gbadamosi, 2020; Nouar et al., 2024).

Adaptive data rate (ADR) is a key LoRaWAN mechanism that adjusts transmission parameters such as spreading factor (SF) and transmission power (TP) according to channel quality. Effective ADR operation can improve packet delivery reliability, battery lifetime, and network scalability. However, the conventional LoRaWAN ADR scheme was primarily designed for quasi-static deployments and often responds poorly when channel conditions vary rapidly due to mobility, gateway switching, or fluctuating interference (Bor et al., 2016; Sarmento Neto et al., 2025).

To address these limitations, recent studies have explored machine learning (ML) and reinforcement learning (RL)-based ADR strategies. These approaches enable communication parameters to be learned directly from environmental feedback rather than relying on fixed decision rules. Existing studies have reported gains in packet delivery ratio, energy efficiency, and adaptation capability compared with standard ADR methods (Azizi et al., 2022a; Farhad and Pyun, 2022; Teymuri et al., 2023; Zholamanov et al., 2024).

Despite these advances, an important practical question remains insufficiently addressed. Can learned ADR policies generalise effectively to unseen deployment scenarios? In real mobile IoT systems, devices may encounter speeds, trajectories, gateway layouts, traffic loads, and propagation environments that differ significantly from training conditions. A policy that performs well in one scenario but degrades sharply under deployment uncertainty may have limited real-world value. Similar concerns regarding policy transferability and robustness have been widely recognized in broader RL research (Mnih et al., 2015; Schulman et al., 2017).

This paper investigates the generalisation and robustness of RL-based ADR control in mobile LoRaWAN networks. Using a mobility-aware simulation framework, standard ADR, heuristic ADR, Deep Q-Network (DQN)-based ADR, and Proximal Policy Optimisation (PPO)-based ADR are evaluated across multiple mobility conditions and tested under unseen scenarios involving altered movement patterns and network parameters. Results show that PPO achieved the strongest transferability, preserving over 99% of baseline

communication reliability while reducing latency by 53.12% and runtime by 81.94% relative to DQN under unseen conditions. These findings indicate that policy-gradient learning offers a robust and scalable solution for intelligent ADR control in next-generation mobile LPWAN systems.
The remainder of this paper is organised as follows: Section 2 reviews related literature, Section 3 presents the methodology and evaluation framework, Section 4 discusses experimental results Section 5 concludes the paper.

MATERIALS AND METHODS

Research on adaptive data rate optimisation in LoRaWAN has expanded significantly in recent years due to increasing demand for scalable, reliable, and energy-efficient IoT connectivity. Existing studies can be grouped into four broad categories, conventional ADR mechanisms, heuristic enhancement methods, machine learning approaches, and reinforcement learning-based optimisation.

Conventional LoRaWAN ADR

The standard LoRaWAN ADR mechanism adjusts SF and TP using historical uplink signal statistics collected over time (Alliance, 2015). It is effective in static or slowly varying environments where long-term channel estimates remain representative. Earlier performance studies showed that ADR can improve network capacity and battery lifetime when devices remain stationary and channel conditions evolve gradually (Adelantado et al., 2017; Bor et al., 2016). In mobile environments, previously observed link conditions may quickly become outdated due to user movement, gateway handover, shadowing, or interference changes. As a result, conventional ADR often adapts too slowly to preserve optimal performance.

Heuristic and Rule-Based Enhancements

To improve responsiveness, several researchers proposed heuristic ADR methods based on recent Signal-to-Noise Ratio (SNR), packet loss rate, Received Signal Strength Indicator (RSSI), or threshold-triggered parameter updates. Hybrid approaches combining multiple rules showed faster adaptation than ADR in dynamic scenarios (Farhad and Pyun, 2022). Context-aware schemes have also been proposed to react to changing network conditions with reduced decision delay (Lodhi et al., 2025). Although heuristic approaches are computationally simple and for low-power systems, their performance depends strongly on manually selected thresholds and environment-specific tuning. Consequently, their ability to transfer across unseen scenarios remains uncertain.

Machine Learning-Based ADR

Machine learning methods have been introduced to reduce reliance on manually-crafted decision rules. Multi-armed bandit and contextual learning approaches can adapt transmission settings using online feedback while balancing exploration and exploitation (Azizi et al., 2022; Teymuri et al., 2023). These methods may outperform static heuristics in moderate dynamic conditions. Nevertheless, many ML techniques remain reactive and focus on short-term reward optimisation. They often do not explicitly model long-horizon sequential decision processes or large state transitions caused by mobility.

Reinforcement Learning-Based ADR

Reinforcement learning provides a natural framework for ADR optimisation because transmission decisions are sequential and environment-dependent. Early LoRaWAN studies using Q-learning demonstrated gains in packet delivery ratio and energy efficiency relative to conventional ADR (Yazid et al., 2022). More recent DQN-based methods improved state representation and adaptation capability using neural function approximation (Zholamanov et al., 2024). However**,** value-based RL approaches may struggle under non-stationary conditions because replay-buffer samples can become inconsistent with rapidly changing environments. They may also overfit to training scenarios, reducing deployment robustness. Similar limitations have been discussed in broader RL literature (Mnih et al., 2015; Van Hasselt et al., 2016). Policy-gradient methods such as PPO perform direct policy learning with constrained updates, often improving stability and transferability in continuous or uncertain environments (Schulman et al., 2017). Despite widespread success in robotics and control, their cross-scenario behaviour for LoRaWAN ADR remains underexplored.

Table 1: Summary of LoRaWAN ADR Studies and Identified Research Gaps

Study	Method Type	Mobility	Generalisation	Main Gap
LoRa Alliance (2015)	Standard ADR	No	No	Static design
Bor et al. (2016)	Simulation study	No	No	No ADR learning
Adelantado et al. (2017)	Performance study	No	No	No mobility focus
Farhad and Pyun (2022)	Heuristic ADR	Partial	No	Threshold tuning
Yazid et al. (2022)	Q-learning ADR	Partial	No	Limited transferability
Zholamanov et al. (2024)	DQN ADR	Partial	Limited	Weak robustness
Nouar et al. (2024)	Mobility analysis	Yes	No	No ADR optimisation
This Work	PPO ADR			

*Note**:** "Partial" indicates the factor was present but not systematically central to the study. "Limited" indicates some evidence was reported, but not comprehensive cross-scenario evaluation.*

Research Gap

Based on the reviewed literature, three important gaps remain:

i. Limited generalisation assessment: Most LoRaWAN ADR studies evaluate performance only under training or fixed scenarios.
ii. Insufficient robustness analysis: Existing works rarely quantify how performance changes under mobility, altered gateway layouts, or unseen propagation conditions.
iii. Weak deployment-oriented comparison: Many studies report average metrics without assessing transferability, runtime efficiency, or operational stability.

Mobility-Aware LoRaWAN ADR Evaluation Framework

The framework provides a mobility-aware simulation environment for evaluating the performance, generalisation,

and robustness of ADR strategies in mobile LoRaWAN networks. It models mobile end devices, multiple gateways, and a central network server, while considering device mobility, wireless propagation, channel variability, traffic behaviour, and transmission parameters such as spreading factor (SF7 to SF12) and transmission power. This is illustrated in figure 1.

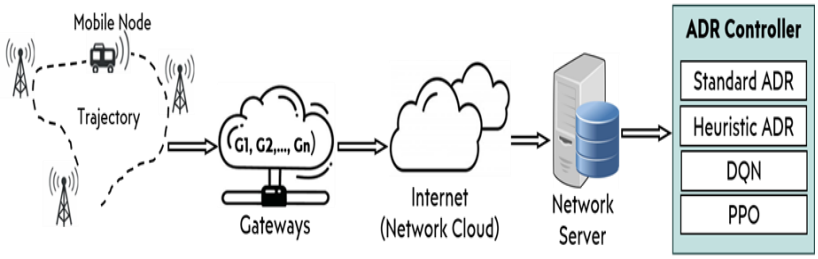

Figure 1: Mobility-Aware Simulation framework for Evaluating ADR Generalisation

As devices move, their changing positions affect received signal strength and SNR, while gateway selection is based on the highest received SNR when multiple gateways can receive a packet, thereby modelling gateway switching. The framework enables a fair comparison of Standard ADR, Heuristic ADR, DQN-based ADR, and PPO-based ADR under identical network and mobility conditions, including both baseline and unseen deployment conditions. Performance is evaluated using reliability, latency, runtime, generalisation, and robustness measures.

Experimental Methodology for ADR Generalisation and Robustness

This section describes the procedures used to evaluate the generalisation capability and robustness of ADR strategies in mobile LoRaWAN networks. A unified mobility-aware simulation environment is employed to ensure fair comparison among benchmark methods. Unlike conventional evaluation procedures that test models only under familiar conditions, this study explicitly separates training scenarios from unseen test scenarios to measure transferability under deployment uncertainty. The procedure adopted in this work is illustrated in Figure 2, which summarises the major stages of the simulation, ADR evaluation, training and testing, performance measurement.

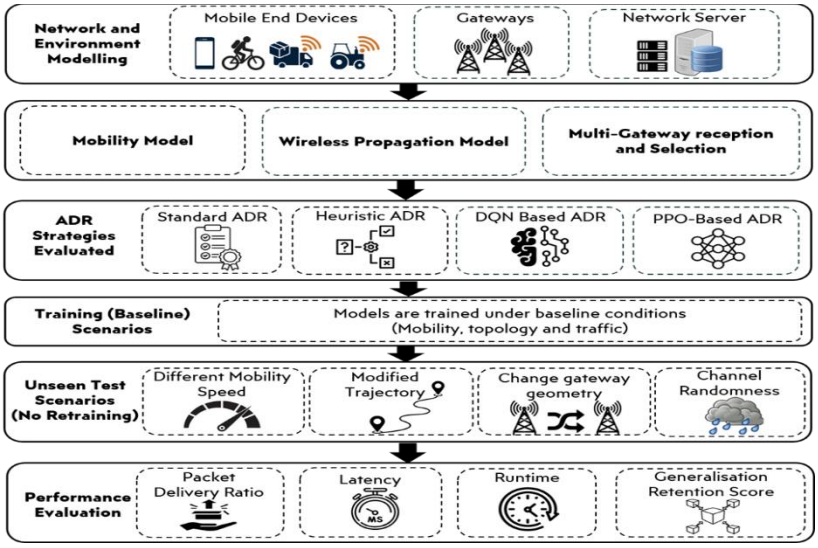

Figure 2. Overview of the Methodology

Network Model

A LoRaWAN network consisting of mobile end devices, multiple gateways, and a central network server is considered. End devices periodically transmit uplink packets while moving within the service region. Depending on instantaneous channel quality, a transmitted packet may be received by one or more gateways. The network server manages packet selection and ADR decisions. Transmission parameters are selected from supported LoRa configurations:

i. Spreading Factor:$SF \in \{7,8,9,10,11,12\}$
ii. Transmission Power: selected from supported radio levels

These parameters directly influence communication range, airtime, reliability, and energy consumption.

Mobility Model

Device movement is generated using the Gauss-Markov mobility model because it captures temporal correlation in speed and direction, making it suitable for realistic mobile IoT behaviour. The velocity update process is defined as:

$$v_t = \alpha v_{t-1} + (1-\alpha)\bar{v} + \sqrt{1-\alpha^2} w_t \quad (1)$$

Where: $v_t =$ Velocity at time step, $v_{t-1} =$ previous velocity, $\alpha =$ Memory coefficient, $\bar{v} =$ Long-term mean direction, $w_t =$ Gaussian random variable

Direction updates follow a similar correlated process. Position is updated by:

$$x_t = x_{t-1} + v_t \cos(\theta_t)\Delta t \quad (2)$$

$y_t = y_{t-1} + v_t \, sin(\theta_t)\Delta t$ (3)

Where:

θ_t =Movement direction, Δt =Time step

This model enables realistic pedestrian, cycling, and vehicular trajectories.

Wireless Propagation Model

Received signal power is estimated using the log-distance path loss model with shadow fading. The received signal power at distance d is gives as:

$$P_r(d) = P_t - \left[PL(d_0) + 10n\, log_{10}\left(\frac{d}{d_0}\right) + X_\sigma\right] \quad (4)$$

Where: $P_r(d) =$ Received power(dBm),$P_t =$ Transmit power(dBm), $PL(d_0) =$ Path loss at reference$d$, $n =$ Path loss exponent, $d =$ Distance between transmitter and receiver,

$d_0 =$ Reference distance, $X_\sigma =$ Shadow fading random variable

Signal-to-noise ratio is computed as:

$SNR_t = P_r(d) - N_0$ (5)

Where: $SNR_t =$ Signal-to-noise ratio, and $N_0 =$ Receiver noise floor in dBm

A packet is considered decodable when the measured SNR exceeds the sensitivity threshold corresponding to the active spreading factor.

Multi-Gateway Packet Selection

Each transmitted packet may be received by multiple gateways. For packet *k*, the selected

Gateway is: $$g_k^* = \arg\max_{g \in G} SNR_{k,g} \quad (6)$$

Where: $\mathcal{G}$ =Set of available gateways, and$SNR_{k,g} =$ Received SNR of packet k, at gateway g.

This models spatial diversity and gateway handover behaviour under mobility.

Evaluated ADR Methods

Four benchmark strategies are considered

i. Standard ADR: Conventional LoRaWAN ADR using historical uplink statistics.
ii. Heuristic ADR: Rule-based parameter updates triggered by recent packet loss trends.
iii. DQN-Based ADR: Value-based RL with discrete state-action mapping.
iv. PPO-Based ADR: Policy gradient RL using clipped policy updates

Training and Unseen Test Scenarios

To evaluate generalisation, experiments are divided into two phases:

i. Training/main scenarios: Learning agents are developed under baseline mobility settings using representative movement speeds and standard gateway placement.
ii. Unseen test scenarios: Trained policies are evaluated under altered conditions such as:
    a. Different mobility speeds
    b. Modified trajectories
    c. Changed gateway geometry
    d. Varied channel randomness
    e. Shifted traffic behaviour

No retraining is performed during unseen test scenarios.

Performance Metrics

i. Packet Delivery Ratio

$$PDR = \frac{N_{succ}}{N_{tx}} \quad (7)$$

Where:

$N_{succ} =$ Number of successfully received packets, $N_{tx} =$ Total number of transmitted packets

ii. Average Latency

$$L_{avg} = \frac{1}{N_{succ}} \sum_{k=1}^{N_{succ}} \left(t_i^{rx} - t_i^{gen}\right) \quad (8)$$

Where: $t_i^{gen} =$ Generation/Transmission time of packet$i$, and $t_i^{rx} =$ Reception time of packet

iii. Runtime

$$T_{run} = T_{end} - T_{start} \quad (9)$$

Where T_{start} and T_{end} denote the start and end times of the computation, respectively.

iv. Generalisation Retention Score

$$G = \frac{M_{unseen}}{M_{baseline}} \times 100 \quad (10)$$

Where: $M_{unseen} =$ Performance under unseen scenario, $M_{baseline} =$ Performance under baseline scenario

Retention score measure how much original performance is preserved after scenario shift. A value close to 100% indicates strong transferability.

Robustness Quantification

Performance consistency across unseen scenarios is measured using:

$$R_m = 1 - \frac{\sigma(m_1, m_2, \ldots, m_s)}{\mu(m_1, m_2, \ldots, m_s)} \quad (11)$$

Where: $\mu(\cdot)$and$\sigma(\cdot)$are the mean and standard deviation of metric m across $s$scenarios.

Robustness score measures the stability across changing conditions. Higher values indicate more stable behaviour.

Statistical Validity

All experiments are repeated using multiple random seeds. Mean values are reported with comparative percentage gains. Cross-method comparisons are further supported through variability analysis and unseen-scenario retention scores.

This methodology enables rigorous evaluation of not only raw performance, but also transferability and deployment readiness under realistic mobile LoRaWAN uncertainty.

RESULTS AND DISCUSSION

This section evaluates the ability of ADR strategies to maintain strong communication performance when deployed under unseen mobility scenarios. Unlike benchmarking under familiar conditions, the emphasis here is on transferability, robustness, and operational stability after scenario shifts. The benchmark methods include standard ADR, heuristic ADR, DQN-based ADR, and PPO-based ADR.

Robustness under Increasing Mobility

To evaluate the stability of each ADR strategy under progressively dynamic operating conditions, packet delivery performance was examined across increasing mobility speeds ranging from low-speed pedestrian motion to higher vehicular movement. Figure 3 presents the variation of PDR with mobility speed for the evaluated methods, providing insight into their robustness and ability to sustain reliable communication as channel conditions become more time-varying.

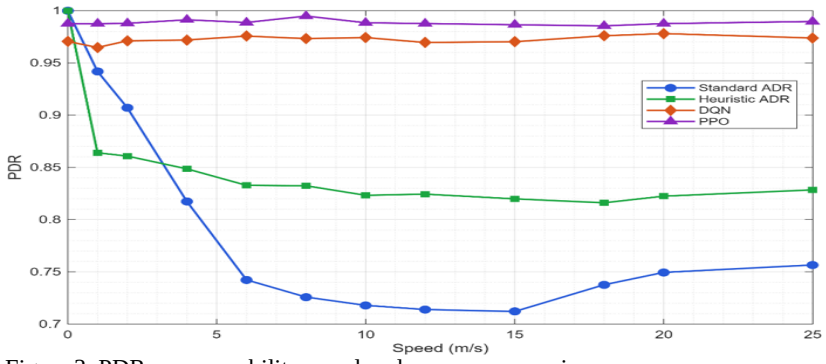

Figure 3: PDR versus mobility speed under unseen scenarios

Standard ADR showed progressive degradation in PDR as speed increased, reflecting delayed adaptation to rapidly changing channel conditions. Heuristic ADR performed better than standard ADR at moderate speed, but became unstable under aggressive mobility transitions. DQN-based ADR preserved stronger reliability than non-learning baselines but exhibited higher variability across scenarios. PPO achieved the most stable cross-scenario behaviour, maintaining consistently high PDR with lower variance in energy and latency metrics. This suggests that policy-gradient learning better captures transferable control behaviour across varying mobility intensities.

Reliability under Unseen Scenarios

Reliable packet delivery remains the most fundamental requirement for mobile IoT applications Figure 4 compares the mean Packet Delivery Ratio (PDR) achieved by each method under unseen deployment scenarios.

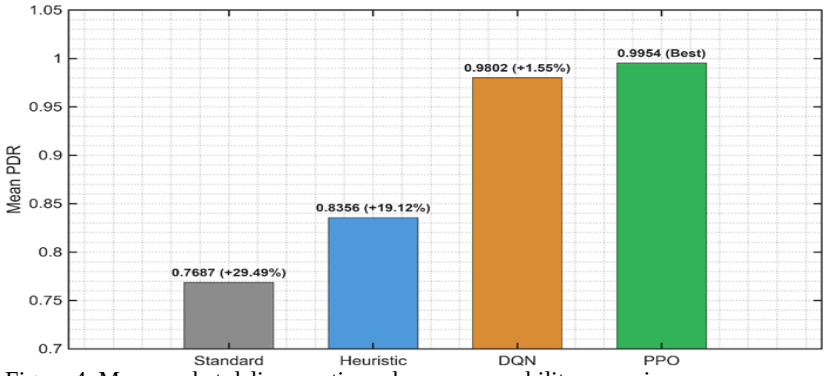

Figure 4: Mean packet delivery ratio under unseen mobility scenarios.

As shown in Figure 4, PPO achieved the highest communication reliability under unseen deployment condition. PPO achieved the highest mean PDR of 0.9954, outperforming standard ADR (0.7687), heuristic ADR (0.8356), and DQN (0.9802). This corresponds to performance gains of 29.49%, 19.12%, and 1.55%, respectively. The results indicate that PPO learned transmission policies that transfer effectively to conditions not encountered during development.

By contrast, standard ADR and heuristic ADR experienced larger reliability losses because their decision processes depend strongly on historical or manually tuned behaviour. Although DQN retained strong PDR, it remained slightly below PPO, suggesting weaker policy transferability under scenario shifts.

Latency Stability under Scenario Shift

Low communication delay is important in logistics tracking, industrial monitoring, and mobile sensing applications. Figure 5 compares average latency under unseen scenarios. Figure 6 compares latency behaviour under scenario shifts and confirms superior responsiveness of PPO. PPO achieved a mean latency of 0.1067 s, significantly lower than standard ADR (0.1350 s), heuristic ADR (0.1113) and DQN (0.2276 s), this represents latency reduction of 20.94%, 4.14% and 53.12% respectively.

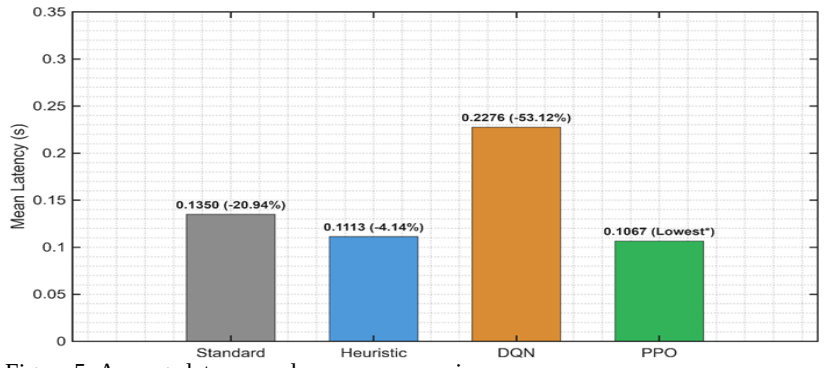

Figure 5: Average latency under unseen scenarios.

The poor latency performance of DQN suggests that policies optimised under baseline conditions may become inefficient when scenario dynamics change, leading to suboptimal parameter choices or increased retransmission behaviour.

Computational Efficiency

Practical deployment of intelligent ADR methods requires not only good communication performance, but also manageable computational overhead. Figure 6 compares runtime across benchmark approaches.

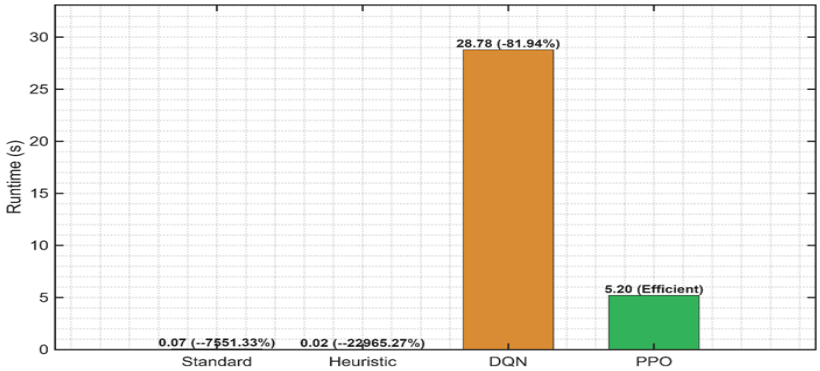

Figure 6: Runtime comparison under unseen scenarios.

The run time comparison in figure 6 shows that PPO remains more efficient than DQN. PPO required 5.20 s average runtime, compared with 28.94 s for DQN, representing an 81.94% reduction. This result is important because lower computational burden improves suitability for real-time adaptation and large-scale simulation or optimisation tasks.

Although standard and heuristic ADR required lower runtime due to simpler decision rules, their reduced reliability under unseen scenarios limits practical effectiveness. PPO therefore provides a favourable balance between intelligence and efficiency.

Consolidated Generalisation Performance

To provide a holistic view, Figure 7 summarises reliability, latency, and runtime behaviour under unseen scenarios. Across all key dimensions, PPO consistently maintained strong operational performance and demonstrated better balance than competing methods. Figure 7 present a consolidated comparison of reliability, latency, and runtime under unseen scenarios.

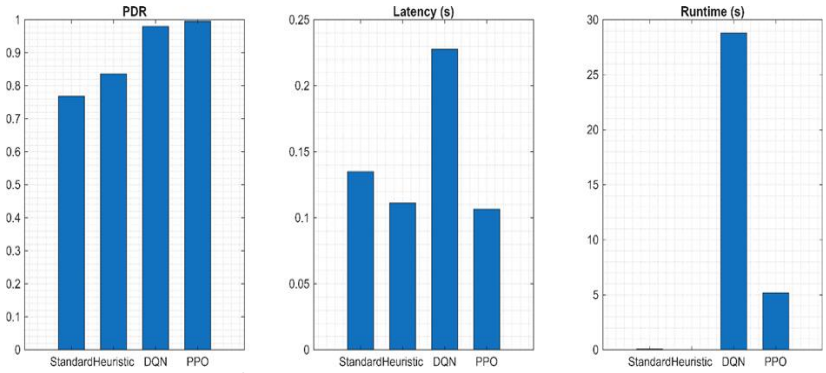

Figure 7: Combined Performance Summary

Figure 7 present a consolidated comparison of reliability, latency, and runtime under unseen scenarios. DQN preserved high reliability but suffered from substantially worse latency and runtime. Standard and heuristic ADR remained computationally lightweight but underperformed in communication reliability. PPO emerged as the most deployment-ready solution because it preserved near-optimal reliability while maintaining low delay and efficient execution.

Generalisation Retention Analysis

Beyond unseen-scenario performance, it is useful to compare how much baseline capability each method retained after scenario shift. Table 2 presents generalisation retention relative to baseline scenarios.

Table 2: Generalisation Retention Summary under Unseen Mobility Scenarios

Method	PDR Retention (%)	Latency Change (%)	Runtime Change (%)	Energy Change (%)	Overall Interpretation (%)	Rank
Standard ADR	103.23	+3.85	+24.30	+6.73	Moderate transferability with slight latency/runtime penalty	4
Heuristic ADR	101.46	-5.12	+30.22	+2.38	Good latency retention but weaker reliability than learning methods	3
DQN-Based ADR	99.21	+31.67	-9.96	+54.04	Strong reliability retention but major latency and energy degradation	2
PPO-Based ADR	99.31	+0.66	-7.39	+21.49	Best balance of reliability retention, latency stability, and efficiency	1

i. PDR Retention $= (Unseen\ Scenario\ PDR\ /\ Baseline\ PDR)\ x\ 100$.
ii. Positive latency runtime/energy change indicates increase under unseen scenarios; negative values indicate reduction.
iii. Ranking reflects balanced generalisation performance across all metrics.

Table 2 shows that PPO achieved the strongest overall generalisation performance under unseen mobility scenarios. Although standard ADR and heuristic ADR recorded retention values above 100%, their absolute reliability remained lower than PPO. PPO preserved 99.31% of baseline PDR performance while maintaining near-constant latency (+0.66%) and reduced runtime (-7.39%). In contrast, DQN retained reliability but experienced substantial latency (+31.67%) and energy (+54.04%) degradation. These results indicate that PPO offers the most balanced and deployment-ready ADR strategy under scenario uncertainty.

These findings suggest that policy-gradient learning captures more transferable control behaviour than value-based learning under dynamic mobile LoRaWAN uncertainty

Discussion

The results demonstrate that high baseline performance alone does not guarantee deployment readiness. A practical ADR controller must remain effective when device movement, propagation conditions, or network layout differ from prior experience.

Among the evaluated methods, PPO exhibited the strongest combination of reliability retention, latency stability, and runtime efficiency. This indicates that direct policy optimisation may be better suited to uncertain mobile LPWAN environments than replay-based value learning methods such as DQN. For real-world systems such as fleet tracking, smart logistics, and mobile industrial sensing, these properties are especially valuable because operating conditions often change unpredictably after deployment.

CONCLUSION

This paper investigated the generalisation capability and robustness of RL-based ADR strategies for mobile LoRaWAN networks under unseen deployment scenarios. Unlike conventional evaluations that focus mainly on average performance under familiar conditions, this study examined whether learned ADR policies remain effective when exposed to altered mobility patterns, changing network conditions, and deployment uncertainty.

Using a mobility-aware simulation framework, standard ADR, heuristic ADR, DQN-based ADR, and PPO-based ADR were comparatively evaluated. Results showed that PPO consistently achieved the strongest overall unseen-scenario performance. PPO obtained the highest mean PDR of 0.9954, outperforming standard ADR, heuristic ADR, and DQN by 29.49%, 19.12%, and 1.55%, respectively. PPO also reduced latency by 53.12% and runtime by 81.94% to DQN while preserving over 99% of baseline communication reliability under scenario shifts.

These findings indicate that policy-gradient learning provides stronger transferability than conventional and value-based alternatives for dynamic Low Power Wide Area Network environments. In practical terms, PPO demonstrated the most favourable balance of reliability, responsiveness, and computational efficiency, making it a strong candidate for intelligent ADR deployment in mobile applications such as logistics tracking, fleet monitoring, smart transportation, and industrial mobility systems.

The present evaluation was conducted in a simulation environment and therefore may not capture all practical effects such as hardware imperfections, regulatory duty-cycle constraints, gateway backhaul delays, antenna orientation, or severe urban blockage. In addition, the unseen scenarios were generated from controlled variations of modelled parameters rather than large- scale field measurements.

Future research may extend this work in several directions

i. Real-world validation using physical LoRaWAN gateways and mobile end devices.
ii. Online continual learning for post-deployment adaptation.

Multi-agent cooperative ADR among gateways and devices.

REFERENCES

Adelantado, F., Vilajosana, X., Tuset-Peiro, P., Martinez, B., Melia-Segui, J., & Watteyne, T. (2017). Understanding the limits of LoRaWAN. IEEE Communications Magazine, 55(9), 34–40.

Alliance, L. (2015). Lorawan specification. LoRa Alliance, 1–82.

Azizi, F., Teymuri, B., Aslani, R., Rasti, M., Tolvaneny, J., & Nardelli, P. H. J. (2022a). MIX-MAB: Reinforcement learning-based resource allocation algorithm for LoRaWAN. 2022 IEEE 95th Vehicular Technology Conference :(VTC2022-Spring), 1–6.

Bor, M. C., Vidler, J. E., & Roedig, U. (2016). LoRa for the Internet of Things. Ewsn, 16, 361–366.

Centenaro, M., Vangelista, L., Zanella, A., & Zorzi, M. (2016). Long-range communications in unlicensed bands: The rising stars in the IoT and smart city scenarios. IEEE Wireless Communications, 23(5), 60–67.

Farhad, A., & Pyun, J.-Y. (2022). HADR: A hybrid adaptive data rate in LoRaWAN for Internet of Things. ICT Express, 8(2), 283–289.

Gbadamosi, A. S. (2020). Cloud-based IoT monitoring system for poultry farming in Nigeria. Arid Zone Journal of Engineering, Technology and Environment, 16(1), 100–108.

Lodhi, M. A., Wang, L., Farhad, A., Qureshi, K. I., Chen, J., Mahmood, K., & Das, A. K. (2025). A Contextual Aware Enhanced LoRaWAN Adaptive Data Rate for mobile IoT applications. Computer Communications, 232, 108042.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., & Ostrovski, G. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529–533.

Nouar, A., Abbes, M. T., Boumerdassi, S., & Chaib, M. (2024). Impact of mobility model on LoRaWAN performance. Journal of Communications, 7–18.

Sarmento Neto, G. A., da Silva, T. A. R., de Abreu, P. F. F., Veloso, A. F. da S., Mendes, L. H. de O., Soares, A. C. B., & Junior, J. V. dos R. (2025). Evaluating a mobility-aware ADR scheme in urban and suburban LoRaWAN environments. Annals of Telecommunications, 1–14.

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms. ArXiv Preprint ArXiv: 1707.06347.

Teymuri, B., Serati, R., Anagnostopoulos, N. A., & Rasti, M. (2023a). Lp-mab: Improving the energy efficiency of lorawan using a reinforcement-learning-based adaptive configuration algorithm. Sensors, 23(4), 2363.

Van Hasselt, H., Guez, A., & Silver, D. (2016). Deep reinforcement learning with double q-learning. Proceedings of the AAAI Conference on Artificial Intelligence, 30(1).

Yazid, Y., Guerrero-González, A., Ez-Zazi, I., El Oualkadi, A., & Arioua, M. (2022). A reinforcement learning based transmission parameter selection and energy management for long range internet of things. Sensors, 22(15), 5662.

Zholamanov, B., Bolatbek, A., Saymbetov, A., Nurgaliyev, M., Yershov, E., Kopbay, K., Orynbassar, S., Dosymbetova, G., Kapparova, A., & Kuttybay, N. (2024). Enhanced Reinforcement Learning Algorithm Based-Transmission Parameter Selection for Optimization of Energy Consumption and Packet Delivery Ratio in LoRa Wireless Networks. Journal of Sensor and Actuator Networks, 13(6), 89.