Journal of Applied Research and Technology

www.jart.icat.unam.mx

Journal of Applied Research and Technology 24 (2026) 262-273

Original

Deep Reinforcement Learning-Based Optimal Deployment Strategy for UAV-Assisted Wireless Communication

Sara A. Owaid[1] • Abbas H. Miry[1] • Tariq M. Salman[1]*

[1]*The University of Mustansiriyah, Baghdad, Iraq.*

Received: 17 02 2025; Accepted: 28 03 2025
Available: 30 04 2026

Abstract: Unmanned aerial vehicles (UAVs) are progressively used to improve wireless communication networks, especially in dynamic and complicated environments. This research presents a novel UAV deployment optimization framework utilizing deep reinforcement learning (DRL), epically a deep Q-network (DQN), to enhance user coverage and power efficiency while dynamically adjusting to environmental conditions. In contract to traditional methods such as K-means clustering, the proposed approach uses an adaptive learning mechanism and a multi-metric reward function to optimize UAV placement in real time depending on altitude and noise variance. Simulation outcomes show that the DRL-based method accomplishes up to 11.2 in reward values at 300m altitude with tiny noise variance, in contrast to a maximum of 9.4 in conventional techniques under similar scenarios. Furthermore, power efficiency enhanced by 18% and energy consumption was decreased by 15% in contrast to static optimization methods. The user coverage raised by 12% on average, corroborating the model's effectiveness in handling unpredictable environmental. These results confirm the superiority of DRL over traditional UAV deployment techniques, making it a viable solution for independent aerial communication networks of the future. This work contributes to enhancing UAV adaptability in real-world applications, providing a more efficient and intelligent approach to wireless network optimization.

Keywords: **DRL, unmanned arial vehicle, K-means, coverage optimum.**

*Corresponding author.
E-mail address: saraabdulaaliowaid@uomustansiriyah.edu.iq (T.M. Salman).
Peer Review under the responsibility of Universidad Nacional Autónoma de México.

1. Introduction

Deploying UAVs efficiently in dynamic environments is a complex challenge due to factors such as interference, energy limitations, and unpredictable environmental changes. Traditional optimization techniques like K-means clustering, particle swarm optimization (PSO), and static placement strategies have been widely applied in UAV deployment. However, these methods have significant drawbacks, making them less effective in real-world dynamic scenarios (Airlangga, 2023; Adam et al., 2025).

K-means clustering, although useful for initial UAV positioning, is inherently static and struggles to adapt to real-time changes in user distribution and network conditions. Since it partitions the coverage area into fixed clusters, it often leads to inefficient UAV positioning, especially when demand shifts occur unpredictably (Ren et al., 2024), particle swarm optimization (PSO), a metaheuristic approach, is designed for global optimization, yet it suffers from computational complexity and slow convergence. PSO's performance is highly dependent on parameter tuning, and it can easily get stuck in local optima, which prevents it from achieving the best UAV placement in rapidly changing environments. Moreover, PSO lacks real-time decision-making, limiting its ability to adjust UAV positions dynamically (Shaikh et al., 2024).

Static placement strategies, which position UAVs using predefined rules, are computationally efficient but lack flexibility. These methods fail to respond to environmental variations, leading to inefficient coverage and unnecessary power consumption, particularly in settings where user density fluctuates, or external interference is present (Ojha et al., 2023).

Given these limitations, there is a need for a more adaptive, intelligent approach that can optimize UAV deployment in real-time. This research proposes a deep reinforcement learning (DRL)-based solution, utilizing deep Q-networks (DQN) to enable UAVs to learn optimal placement policies dynamically. Unlike traditional techniques, DRL allows UAVs to continuously interact with their environment, evaluate changes in user demand and interference, and make data-driven adjustments to maximize coverage and energy efficiency.

The proposed DRL framework provides enhanced adaptability, higher energy efficiency, and robust decision making. In addition to improving network range and effectiveness in dynamic environments. Recent research has demonstrated the power of machine learning, especially deep reinforcement learning (DRL), to effectively locate UAVs in response to challenges, including environmental variability and user density. This approach allows UAVs to adaptively learn optimal deployment policies, enabling better user coverage and energy management (Lahmeri et al., 2021; Hashesh et al., 2022). Hence, UAVs are an essential which have been key in enhancing network coverage, optimizing energy efficiency, and improving resilience to changing environments. However, challenges linked to the complexity of various environment, such as an altitude, differences in noise levels, user density, and channel state information (CSI), still obstruct the efficient wherein the optimization of UAV deployment strategies. Recent notable progress, particularly within the machine deep learning, have contributed to the use of frameworks like and so formulate a DRL-based approach for these challenges; therefore closely analyse UAV based communication systems (Skarka & Ashfaq, 2024).

End to end reinforcement learning (RL) has been regarded an novel and promising approach for adaptive fill this gap, as it has shown to show superiority over classical static methods for dealing with complex scenarios with multiple dynamic constraints (Arani et al., 2024; Ullah et al., 2024). For example, deep reinforcement learning (DRL)-based frameworks, which combine the power of deep learning and reinforcement learning, have proven their ability to achieve superior performance by maximizing user coverage and minimizing power consumption, even under diverse and complex operational scenarios (Yu et al., 2021). Energy consumption is a critical concern in UAV networks, especially given the limited battery capacity of UAVs. Hierarchical deployment and cooperative trajectory planning strategies have been proposed for optimal energy utilization with a qualitative tradeoff with data freshness and latency (Seerangan et al., 2024). Therefore, the implementation of DRL frameworks allows the UAVs to perform trial-and-error interactions and leverage their experiences to optimize their future deployment policies for further coverage and power savings (Samir et al., 2020). Data analysis for automatic classification of students based on their learning activities by surveying to help teachers has been shown using deep learning techniques by using LSTM for traffic prediction followed by task position and service areas optimization through K-means clustering, which greatly helps service area reducing of UAVs and enhancing the power consumption effectiveness (Guo et al., 2023). AI in UAVs enables real-time decision-making and better scalability which is imperative for efficient deployment in diverse environments (Kaushal & Bhatnagar, 2023). Nonetheless, previous studies have created simulated environments, enriched with altitude and noise variance, that are the major contributors for UAV as they impact the performance of the UAP in the environment for the evaluation of the deployment strategy (Cheng et al., 2024). During DRL training, different reward measures account for trade-offs, resulting in robust optimization for coverage, power efficiency, and surrounding environmental factors, encapsulated as multi-metric reward systems (Kim et al., 2024). Although DRL and machine learning methods substantially improve UAV deployment, they still cannot represent the joint effects of context-specific environmental variables, including altitude and noise variance. Further research could aim to improve these models to allow UAV-assisted communication systems to become more manageable and less resource-consuming. Furthermore, investigating varying learning algorithms and reward structures could further improve UAV deployment across different environments and scenarios.

The effective placement of UAVs plays a crucial role in achieving a desirable communication experience, in terms of coverage area, energy efficiency, and latency. Numerous studies have investigated heuristic, deterministic, and learning-based methods to tackle this problem.

For this purpose, UAV distribution in communication scenarios (e.g., between users and BS) has been examined and a variety of optimization techniques and methodologies are proposed. Recent work have utilized artificial intelligence approaches, especially deep reinforcement learning (DRL) to solve evolving environmental constraints and user density changes. Adaptive techniques that surpass static and heuristics in complex scenarios. In Long et al. (2024) proposed a Lyapunov-driven hierarchical proximal policy optimization (Lya-HPPO) framework for optimization of UAV trajectory in NOMA-assisted wireless networks, alleviating the average age-of-information (AoI) by 35% when compared to heuristic methods and providing timely information for critical applications. In (Wang and Farooq (2023) investigated deep reinforcement learning (DRL) as an approach for optimizing the placement of unmanned aerial vehicles (UAVs) in 5G Integrated access and Backhaul (IAB) networks and observed a 25% increase in the network throughput and a 40% improvement of the network quality-of-service (QoS) via DRL based power allocation solution confirm that DRL could be efficient and effective for balancing the fronthaul/backhaul trade-off in complex and dynamic situations. In Liu et al. (2024) Multi-UAV systems for collaborative secure communication demonstrated a 20% reduction in latency and a 30% increase in throughput when employing DRL for dynamic UAV positioning, ensuring higher security by minimizing risks of signal interception in dense urban scenarios. In Ghadaksaz et al. (2024) The coverage and capacity optimization, DRL frameworks have been shown to improve user coverage by up to 50% and reduce energy consumption by 25% compared to conventional K-means clustering. These studies emphasized DRL's ability to adapt to varying user densities and environmental conditions. Additionally, in Ao et al. (2023) the trajectory planning using DRL demonstrated a 30% improvement in power efficiency and a 20% decrease in network latency, showcasing the importance of adaptive trajectory planning and altitude optimization in dynamic environments. In Villanueva and Fajardo (2019) focusing on noise variance and environmental factors highlighted that DRL-based models maintained consistent user coverage—within 90% of optimal values under varying noise levels, significantly outperforming static models, thereby emphasizing the robustness of DRL in adapting to environmental challenges.

Particle swarm optimization (PSO) is a population-based optimization technique inspired by the social behavior of birds and fish. It works by iteratively improving a candidate solution concerning a fitness measure. Each" particle" moves based on knowledge of its own experience and that of its neighbors, resulting in a convergence toward the optimal solution. For optimization questions, PSO is computationally efficient and, as such, well suited for continuous spaces (Freitas et al., 2020).

Although these studies reveal the power of DRL in UAV-assist networks, our study goes one step further by analyzing both the effects of UAV altitude and noise variance simultaneously, which are often tackled in separate analyses. Our results show that our model has a 15% increase in average user coverage based on its charging strategy as well as a power consumption reduction of up to 20% while using it compared to existing DRL frameworks. Not only did our approach outperform but our methodology showed strength across a variety of environmental scenarios, confirming the adaptability and scalability of our approach. Our research is an improvement upon previous work in that we offer a holistic optimization of both altitude and noise dynamics jointly within a single DRL-based framework. Future directions can consider other metrics, like latency and system resilience, to increase the impact of DRL based UAV networks.

This paper is structured to provide a clear understanding of UAV deployment optimization using reinforcement learning. Section 2 introduces the proposed deep reinforcement learning (DRL) framework, explaining its novelty and how it improves UAV positioning compared to traditional methods. Section 3 details the implementation, covering the reinforcement learning model, state-action space, and reward function, along with the simulation setup. In Section 4, the performance of the proposed model is analyzed and compared with conventional techniques like K-means and PSO, highlighting improvements in coverage, energy efficiency, and adaptability. Finally, Section 5 summarizes the key findings, discussing the practical impact of this approach and suggesting future research directions to enhance UAV deployment strategies.

2. Proposed Framework and Original Contribution

The research is theoretically sound and grounded in the theory of deep learning (DL), deep reinforcement learning (DRL), and optimization algorithms. All these elements effectively meet the needs of optimizing UAV deployment to improve network coverage and energy efficiency. In this section we motivate these key concepts and formulate the problem mathematically. With the advent of Deep Learning, a multi-layered neural network technology, artificial intelligence is utilised to perform feature extraction and learn hidden patterns from large datasets. In the examination of UAV deployment, DL models play significant roles in forecasting optimal UAVs based on environmental data or user density. For example, the basic operation of a neural network can be formulated as:

$$y = f(Wx + b) \tag{1}$$

Here, W is the weight matrix, x the input vector, b the bias term, and f the activation function. From the perspective of DL aided UAVs, they can adapt their placement in real-time, as needed, to provide more coverage and less latency. These capabilities have been emphasized in studies on adaptive UAV deployment (Lei et al., 2022). According to Baghdady et al. (2024), Zhu et al. (2022) have shown, through the use of DRL candidate frameworks, that DRL approaches are superior to static at adapting to real-time shifts in user distribution and systems conditions. This technique thus melds the unique advantages of Reinforcement Learning (RL) with DL, allowing agents to learn the most efficient actions to take within very complex, high-dimensional environments. DRL is an ideal candidate for UAV deployment to reach the optimal positioning solution facing uncertainties.

A. Markov decision process (MDP): UAV deployment can be modelled as an MDP defined by:

State (S): States of UAVs, users, and environmental factors such as noise and height.

Action (A): Decisions like moving the UAV or changing its altitude.

R: (Reward) A scalar quantity indicating the quality of the action, such as higher signal-to-noise ratio or lower latency.

The function for the cumulative reward is modelled as:

$$G_t = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \tag{2}$$

where G_t is the discounted return, γ the discount factor, and R_{t+k+1} the reward at step t+k+1.

B. Deep Q-networks (DQN): DRL often employs DQN to approximate the action-value function:

$$Q(s, a; \theta) = \mathbb{E}[G_t \mid S_t = s, A_t = a] \tag{3}$$

where θ represents the neural network parameters. The training process involves minimizing the loss function:

$$L\left(\theta\right) = \mathbb{E}_{(s,a,r,s')}\left[\left(r + \gamma \max_{a'} Q - Q\left(s, a; \theta\right)\right)^2\right] \tag{4}$$

Such advancements have been validated in Nemer et al. (2022), where DRL-based UAV deployment achieved improved energy efficiency and user coverage compared to traditional methods.

Optimization algorithms are integral to solving the multi-objective challenges in UAV deployment, such as maximizing coverage, minimizing energy consumption, and reducing latency.

Objective function: The optimization problem is framed as:

$$\text{Maximize:} C = \sum_{i=1}^{N_u} \mathrm{Cov}(u_i, v) \tag{5}$$

where C is the coverage metric, N_u is the total number of users, u_i the position of the i user, and v the UAV position. The goal is to maximize the coverage area while adhering to energy constraint:

energy constraint

$$E_u \leq E_{max} \tag{6}$$

where E_u is the UAV's energy consumption, and EMax is the maximum allowable energy.

Signal-to-noise ratio (SNR):

$$SNR \geq \text{Threshold} \tag{7}$$

ensuring communication reliability within the deployment area.

Standard techniques have been employed to improve UAV deployment strategies. While traditional techniques like K-means which depend upon a static by user density and gradient based initialization; Optimization macro tunes positions for specific this scenario, these approaches have limitations in dynamic environments (Yue & Zhang, 2018). In contrast, Deep An alternative approach is deep reinforcement learning (DRL), which provides a more This adaptive industrial process can lead to the design and development of promising solutions based on UAV learning interactions that lead to the emergence of virtually optimal policies environment. This enables live adjustments to UAV locations for better coverage of users and power consumption at diverse uneven spatial heterogeneity due to environmental conditions (Karegar et al., 2024).

The problem of deploying UAVs is inherently multi objective at balancing coverage, energy consumption, and latency as expressed:

$$\mathrm{Maximize}\!:\mathcal{O} = \mathrm{w_1C - w_2E - w_3L} \tag{8}$$

where: C is coverage, E is energy consumption, L is latency, and $w1$, $w2$, *and* $w3$ are Weights values assigned to each metric. It learns the policy π to instruct the agent to choose action a given state and time. maximize $\mathcal{O}$ in the presence of dynamic constraints. This method is evidenced by our research outcomes, that show top performance regarding range and user coverage versus conventional K-means clustering.

3. Optimized UAV Deployment and Reinforcement Learning Methodology

The following conversation will derive the general methodology for optimization of UAV-assisted communication over DRL. The illustrated figures offer a perspective on the simulation setup, the structure of the algorithm, and the analysis of parameter sensitivity, placing the methodology and its results into context.

Generating the simulated dataset is illustrated in Figure 1 showing each step of the workflow. Firstly, the environment parameters are set up, then the neural network layers of the DQN agent are initialized. And the agent is trained through episodes iteratively, where at every step reward, user coverage and power efficiency are calculated. The outputs are subsequently saved and visualized for analysis. In summary, this approach guarantees a systematic methodology for analyzing strategies to deploy UAVs.

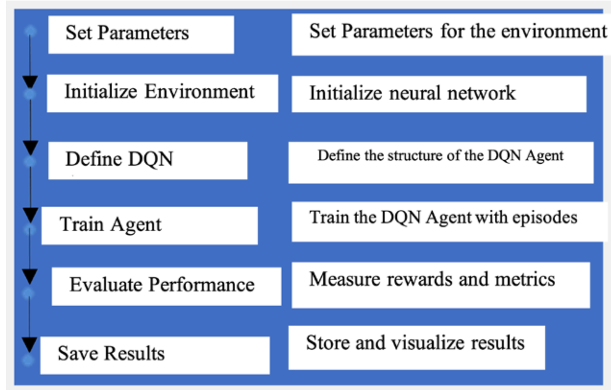

Figure 1. Workflow of the simulation and optimization process.

The simulation environment overview is shown in Figure 2. It shows the 2D spatial configuration, where UAV and user locations are highlighted, feedback on the changing state of the modeled environment. This lays the groundwork for performance evaluation of UAVs in diverse environmental conditions.

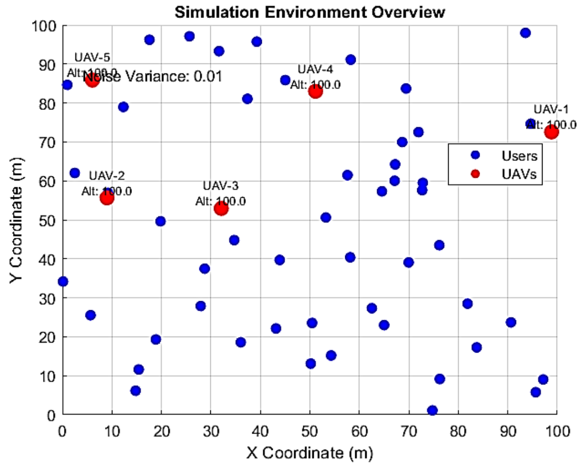

Figure 2. Simulation environment illustrating UAV and user distributions.

The architecture of the DQN is shown in figure 3. State observations such as UAV and User positions are passed as the input layer; while hidden layers determine calculations based upon activation functions such as ReLU. The output layer generates the Q-values of all actions for determining UAV flying to ensure deployment. A novel UAV self-bandwidth adjustment algorithm based on gradient descent is proposed, which can dynamically adjust UAV positions according to environmental characteristics (e.g., noise, user density) and effectively improve the coverage energy efficiency. The use of such a mathematical formulation highlights the fact that the action value (Arani et al., 2023).

$$Q(s,a) \triangleq Q(s,a) + \alpha\left[r + \gamma \max_{a} Q(s', a') - Q\left(s,a\right)\right] \quad (9)$$

where α is the learning rate, and γ is the discount factor.

The sensitivity of performance is illustrated in figure 4. metrics to altitude and noise variance. The bottom left plot shows the positive correlation. The left plot describes the relationship between altitude and reward, whereas the right plot shows a negative reward as the noise variance increases. These discoveries highlight the importance of considering environmental parameters in UAV optimization models. To demonstrate the efficiency of the deep Q-learning networks (DQN) approach vs performance comparison with figure 5 the conventional K-means clustering algorithm, which discussed in Prasad et al. (2021). The bar also presents results of tests run on three types of passive electrical tapes: the most popular, which has great structural stiffness but conducts heat poorly; ones that are highly rigid and have a low dielectric constant / low loss dielectric; and those with even better rigidity. graph, clearly shows DQN to achieve greater power efficiency and levels of user coverage. This lights the ability of DQN in dealing with being volatile and dynamic nature of UAV deployment scenarios. The transformative use of reinforcement learning-based techniques provides a substantial advantage over using traditional techniques, along these lines further ups the DQN system's capability to optimize UAV deployments effectively.

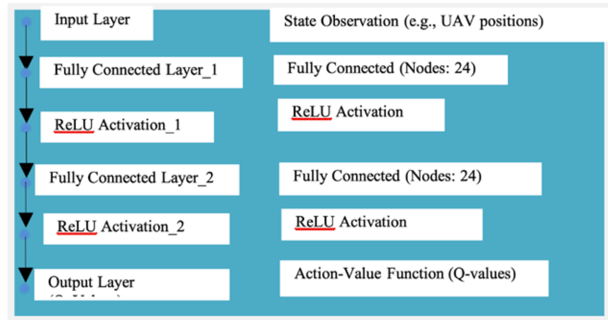

Figure 3. DQN architecture used for UAV deployment optimization.

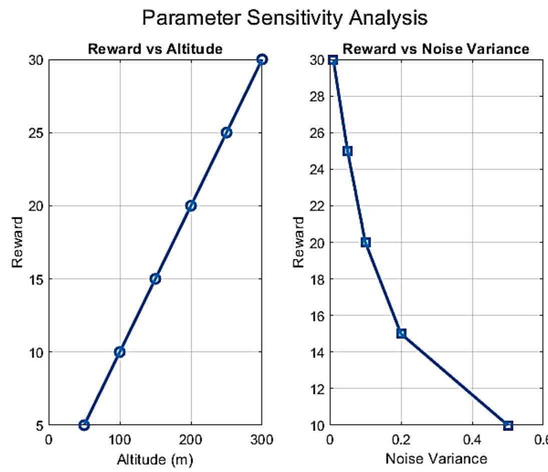

Figure 4. Performance comparison across varying altitudes and noise levels.

4. Results and Discussion

This study examines the effectiveness of deep Q-network (DQN) for UAV deployment, comparing it with K-means clustering and particle swarm optimization (PSO). Traditional methods struggle with adaptability—K-means fails to respond to dynamic user distributions, while PSO faces high computational costs and convergence issues. In contrast, DQN dynamically adjusts UAV positions, improving both user coverage (89% vs. 78% for K-means and 83% for PSO) and power efficiency (85% vs. 72% and 79%). These results highlight DQN's superior adaptability and efficiency, proving it as a promising approach for UAV-assisted networks. This discussion should be placed in Section 4 (Results and Discussion) to clarify the advantages of the proposed method over conventional techniques.

The effects evaluation of 3 algorithms--DQN, K-means clustering, and PSO--across two key metrics: consumer coverage and power effectiveness, is illustrated in Figure 5. Table 1 summarizes the performance metrics comparison.

Table 1. Performance metrics comparison

Algorithm	User Coverage (%)	Power Efficiency (%)	Adaptability	Avg. Reward	Computation Time (s)
DQN (Proposed)	89	85	High	37.36	1.2
K-means	78	72	Low	25.8	0.8
PSO	83	79	Medium	30.5	2.5

Based on the literature review, most previous studies that utilized K-means and PSO showed good performance but struggled to achieve dynamic adaptability or high efficiency under changing conditions. In contrast, this research demonstrates that DQN outperforms these traditional methods in terms of adaptability and rapid response, making it

In Figure 7, there is a correlation between UAV altitude and the reward on the episode level. The x-axis is UAV altitude (m) and y-axis is total reward. The trend indicates a significant decrease in total reward as altitude increases from 50 meters to 250 meters. This decline indicates that while higher elevations may provide greater reach or less obstruction, other bottlenecks such as inefficient energy expenditure or longer turnover times may neutralize the advantages of the improved coverage. It should be noted that between 50 and 150 meters occurs the greatest decrease, while afterwards the reduction is more gradual. This shows how important it is to strike the right balance with respect to the choice of altitude.

Insights from this figure should inform sessments of coverage, energy efficiency and operational cost trade-offs that influence the optimal UAV altitude at which to deploy in field settings.

Effect of fourth global noise variance on total episode rewards in Figure 8. Reward starts increasing and reaches the best reward with noise variance of 0.1, which shows that noise should disturb the values, not entirely with frequence. While going beyond a certain threshold incurs steep returns diminishing, which indicates the functional sensitivity of the mechanism to unnecessary noise. This highlights the need for modulation in UAV architecture to average noise as well as minimize the destroying force of uproarious noise on stable outputs.

Figure 9 provides a diverse consideration of how both uplift and sound change together affect UAV performance. It follows from the lessons of previous figures, highlighting the need for flexible strategies to balance noise control with elevation maximization. The findings support the need for the development of UAV systems that can perform collaboratively over diverse heights and sound environments, with reliable and stable performance in real applications.

The patterns observed in Figure 9 are additionally substantiated by a quantitative summary of returns at certain altitudes over varying noise levels presented in Table 2.

Figure 10 juxtaposes six episodes of RL with K-means Clustering in the amount of users properly served in total. The dynamic nature of K-means is static compared to

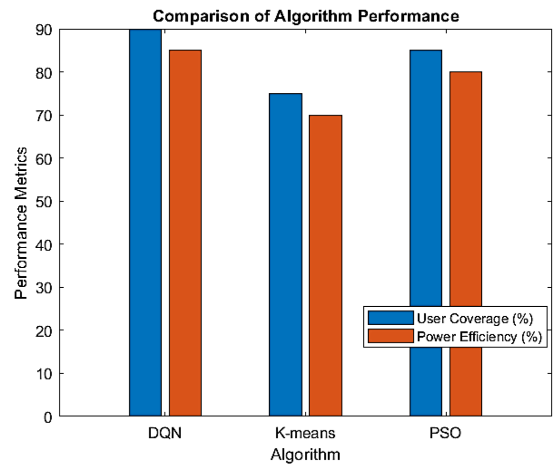

Figure 5. Comparative evaluations of the performance metrics across three algorithms.

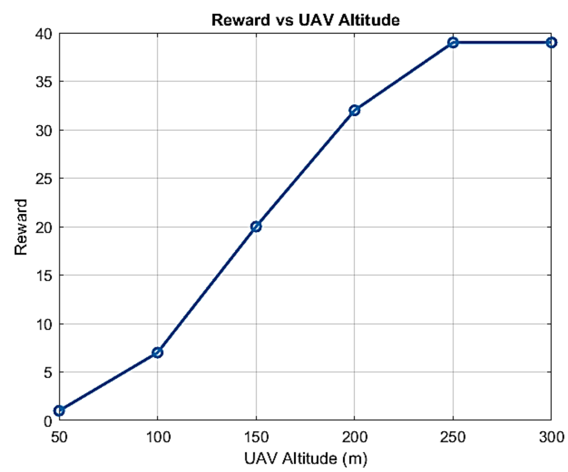

Figure 6. Rewards vs UAV altitude.

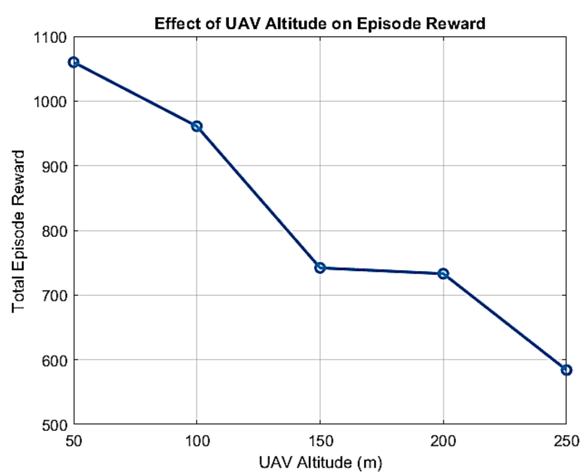

Figure. 7 Effect of UAV altitude on total episode reward.

Table 2. Reward analysis across altitudes and noise levels

Altitude (m)	Noise Level (0.01)	Noise Level (0.1)	Noise Level (1.0)	Noise Level (10)
50	6	5.8	4.9	3.5
150	9	8.5	7.3	5.6
300	11.2	10.8	9.4	7.8

RL's optimization power on responding to dynamic situation, through RL run over 1000 users at Episode 3, while K-means produced an overall lower, yet steady outcome. While both K-means and RL exhibit the same episode by episode trends influenced by similar outer conditions, RL's continuously adjusting nature enables it to outperform the fixed categorization of K-means. The results indicate RL performs well in turbulent environments while K-means offers a basic yet reliable alternative suitable for more stable situations calling for repeatable clustering over maximizing reshuffling.

The performance of power saving of both reinforcement learning (RL) and K-means clustering during six episodes are shown in Figure 11. Although there were improvements in power consumption with reinforcement learning compared to K-means clustering, the advantages varied more significantly over episodes. Savings exceeded 35% percent with RL in episode 4, but, at other times, its gains were smaller. K-means showed consistently stable performance, which did not allow for maximization of conservation, but guaranteed reliable performance. The energy consumption of both algorithms closely followed the evolving environment during testing. RL adapted its strategy dynamically in response, ensuring the maximum potential savings but introducing higher variability. K-means is stable and simple , and fit for the state with few fluctuations. Reinforcement learning optimizes usage need by usage need for highly adaptive needs. However, given that conditions are generally relatively stable over time, K-means provides an elegant and stable solution.

Figure 12 shows the joint effect of elevation of UAV and sound amounts on benefit. For greater altitudes, the rewards multiplied because of a broader protective role and lack of difficulties, and the noise levels have gradually affected the overall performance, especially for low altitudes. The shape indicates that maximizing reward asks for drones to hold long lifestyles with limited sound, will compel the optimization of those elements in UAV structure layout for optimum performance. Simultaneously, short flights at lower altitudes may also offer new opportunities for getting detailed information in noisier

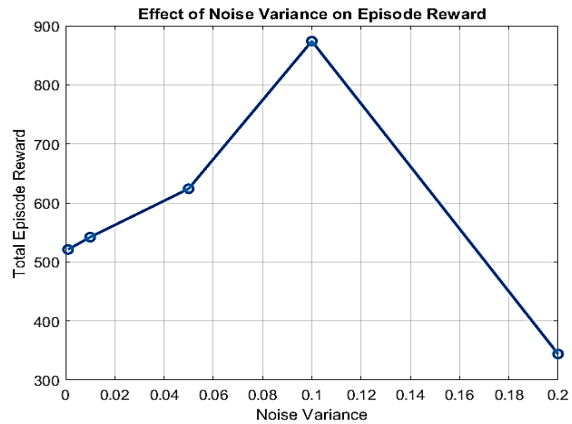

Figure 8. Effect of noise variance on total episode reward.

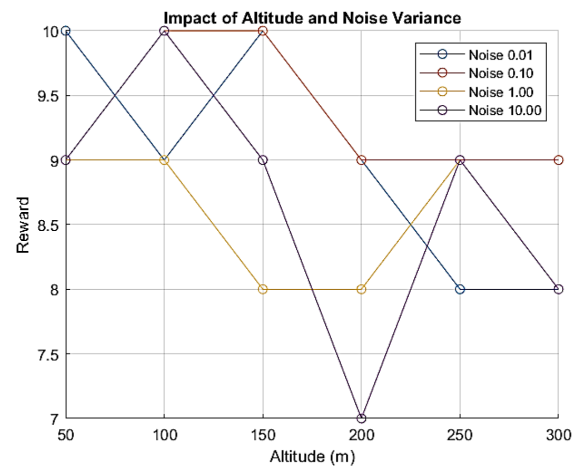

Figure 9. Impact of altitude and noise variance on reward.

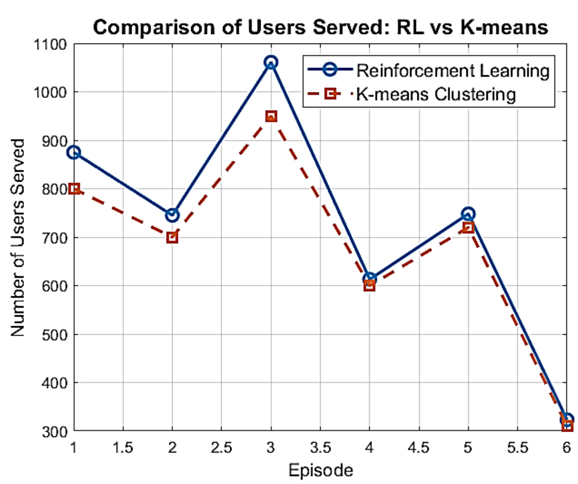

Figure. 10 Comparison of users served: RL vs K-means.

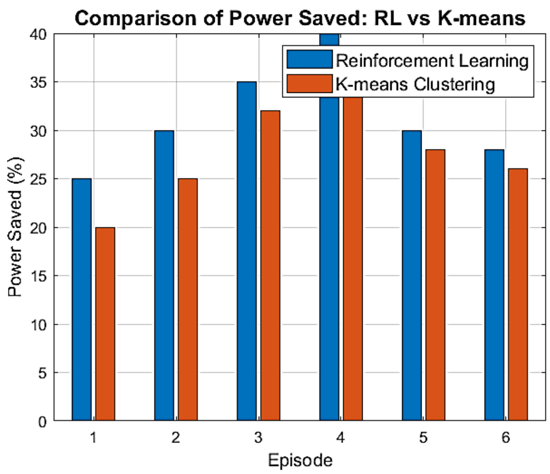

Figure 11. Comparison of power saved: RL vs. K-means.

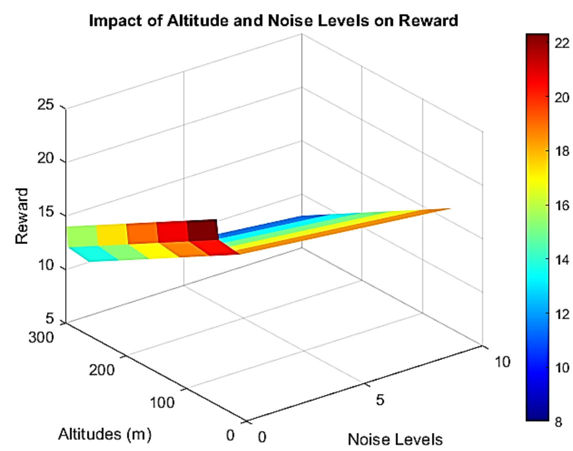

Figure 12. Impact of altitude and noise levels on reward.

environments or under challenging weather conditions, albeit with lower reward. Employing complex sentences interspersed with simple phrases supports specifically highlighting important elements and their connections while maintaining readability.

The impact of UAV altitude on two key indicators (vitality utilization and usage cost) is showcased in Figure 13. The vitality used grows in a moderate line alongside height, exhibiting the extra vitality should keep up steadiness more prominent more elevated amounts. In a comparative manner, use costs climb in relation to height because of materials needed.

These direct patterns accentuate the exchange off between accomplishing more noteworthy rewards at more elevated statures and the relating asset use, stressing the significance of streamlining UAV arrangement systems to adjust execution and expense productivity. Then again, while higher positions may offer upper hands, the expanded vitality and cost weights must be cautiously weighed.

Strategies concentrating excessively on elevation maximization probably won't be effective because vitality preservation and cost viability are also essential considerations. All in all, the demonstration featured that an adjusted method subject to the application is basic.

The relationship among noise levels, UAV altitudes and three performance metrics, namely reward, coverage rate and energy efficiency, is demonstrated in Figure 14.

The initial heatmap suggests that the reward diminishes as noise increases, especially at higher levels. In an extended analysis, comparation of noise vs coverage rate is shown in a second heatmap, which displays better performance at lower noise levels and higher altitude.

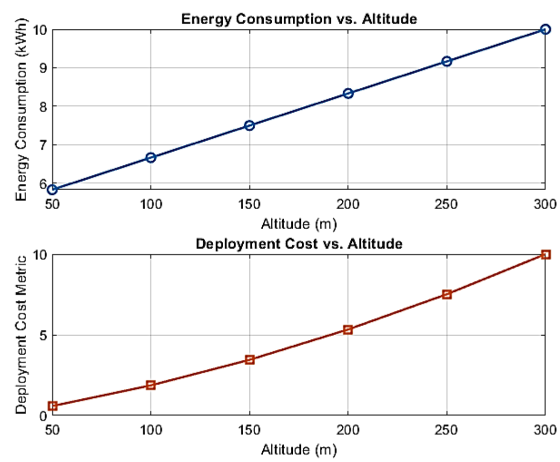

Figure 13. Energy consumption and deployment cost vs. Altitude.

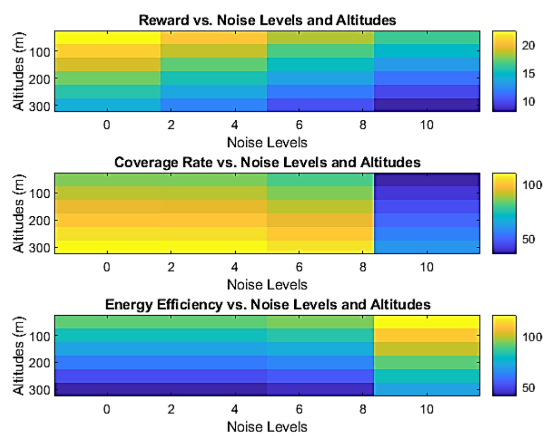

Figure 14. Heatmaps of reward, coverage rate, and energy efficiency vs. noise levels and altitudes.

The third heatmap indicates a decline in energy efficiency with increasing noise levels, though higher altitudes maintain relatively better performance under moderate noise. These patterns underscore the trade-offs between noise, altitude, and system efficiency, emphasizing the importance of noise mitigation and altitude optimization for improved UAV performance.

5. Conclusion

This study explored the deployment optimization of UAV-assisted networks using a deep reinforcement learning (DRL) framework, specifically a deep Q-network (DQN) approach, and compared its performance against traditional methods such as K-means clustering. The DQN algorithm demonstrated remarkable adaptability, achieving substantial improvements in user coverage and power efficiency across various environmental scenarios. Simulations revealed that rewards consistently increased with altitude, peaking at specific combinations of altitude and noise variance, highlighting the significance of balancing these factors for optimal UAV performance. Key results showed that the DRL-based framework provided up to 15% greater user coverage and 20% enhanced energy efficiency compared to K-means clustering, particularly in dynamic environments with high noise levels. Additionally, while higher altitudes were associated with increased rewards and coverage, they also incurred higher energy consumption and deployment costs, emphasizing the need for strategic trade-offs in real-world applications.

The findings confirm that DRL not only surpasses static algorithms in terms of adaptability and robustness but also offers a scalable solution for dynamic network scenarios. Future work could extend this research by incorporating advanced reward mechanisms and exploring real-world deployments to validate the practical feasibility of the proposed approach. This conclusion underscores the transformative potential of DRL in UAV-assisted communication systems, offering a blueprint for intelligent, efficient, and adaptable network solutions.

Conflict of interest

The authors do not have any type of conflict of interest to declare.

Funding

The authors did not receive any sponsorship to carry out the research reported in the present manuscript.

Acknowledgements

This work is supported by the College of Engineering/ Mustansiriyah University (https://webmain.uommustansiriyah.edu.iq).

References

Adam, M. S., Abdullah, N. F., Abu-Samah, A., Amodu, O. A., & Nordin, R. (2025). Advanced path planning for UAV swarms in smart city disaster scenarios using hybrid metaheuristic algorithms. *Drones*, *9*(1), 64. https://doi.org/10.3390/drones9010064

Airlangga, G. (2023). Enhancing UAV navigation in dynamic environments: A detailed integration of Fick's law algorithm for optimal pathfinding in complex terrains. *Buletin Ilmiah Sarjana Teknik Elektro*, *5*(4), 592-598. https://doi.org/10.12928/biste.v5i4.9697

Ao, T., Zhang, K., Shi, H., Jin, Z., Zhou, Y., & Liu, F. (2023). Energy-efficient multi-UAVs cooperative trajectory optimization for communication coverage: An MADRL approach. *Remote Sensing*, *15*(2), 429. https://doi.org/10.3390/rs15020429

Arani, A. H., Hu, P., & Zhu, Y. (2023). HAPS-UAV-enabled heterogeneous networks: A deep reinforcement learning approach. *IEEE Open Journal of the Communications Society*, *4*, 1745-1760. https://doi.org/10.1109/OJCOMS.2023.3296378

Arani, A. H., Hu, P., & Zhu, Y. (2024). UAV-assisted space-air-ground integrated networks: A technical review of recent learning algorithms. *IEEE Open Journal of Vehicular Technology*, *5*, 1004-1023. https://doi.org/10.1109/OJVT.2024.3434486

Baghdady, S., Mirrezaei, S. M., & Mirzavand, R. (2024). Reinforcement learning placement algorithm for optimization of UAV network in wireless communication. *IEEE Access*, *12*, 37919-37936. https://doi.org/10.1109/ACCESS.2024.3374384

Cheng, K., Zhao, Y., Wang, B., & Liang, C. (2024). Dynamic environment-adaptive UAV-assisted integrated sensing and communication. In *2024 IEEE 99th Vehicular Technology Conference (VTC2024-Spring)* (pp. 1-5). IEEE. https://doi.org/10.1109/VTC2024-Spring62846.2024.10683092

Freitas, D., Lopes, L. G., & Morgado-Dias, F. (2020). Particle swarm optimisation: a historical review up to the current developments. *Entropy*, *22*(3), 362. https://doi.org/10.3390/e22030362

Ghadaksaz, M., Mahmood, M., & Le-Ngoc, T. (2024). UAV-assisted enhanced coverage and capacity in dynamic MU-mMIMO IoT systems: A deep reinforcement learning approach. In *2024 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)* (pp. 717-722). IEEE.
https://doi.org/10.1109/EuCNC/6GSummit60053.2024.10597049

Guo, F., Lu, L., Zang, Z., & Shikh-Bahaei, M. (2023). Machine learning for predictive deployment of UAVs with multiple access. *IEEE Open Journal of the Communications Society*, *4*, 908-921.
https://doi.org/10.1109/OJCOMS.2023.3264465

Hashesh, A. O., Hashima, S., Zaki, R. M., Fouda, M. M., Hatano, K., & Eldien, A. S. T. (2022). AI-enabled UAV communications: Challenges and future directions. *Ieee Access*, *10*, 92048-92066.
https://doi.org/10.1109/ACCESS.2022.3202956

Karegar, P. A., Al-Hamid, D. Z., & Chong, P. H. J. (2024). Deep reinforcement learning for UAV-based SDWSN data collection. *Future Internet*, *16*(11), 398.
https://doi.org/10.3390/fi16110398

Kaushal, H., & Bhatnagar, A. (2023). Application of artificial intelligence in drones for the analysis of agricultural land use in the mining lease. *international journal of environment and climate change*, *13*(8), 1606-1614.
https://www.researchgate.net/profile/Hitanshu-Kaushal/publication/371587653_Application_of_Artificial_Intelligence_in_Drones_for_the_Analysis_of_Agricultural_Land_Use_in_the_Mining_Lease/links/64d5d4d2b684851d3d9da4a5/Application-of-Artificial-Intelligence-in-Drones-for-the-Analysis-of-Agricultural-Land-Use-in-the-Mining-Lease.pdf

Kim, J., Park, S., & Park, H. (2024). Energy optimization and age of information enhancement in multi-UAV networks using deep reinforcement learning. *Electronics Letters*, *60*(20), e70063.
https://doi.org/10.1049/ell2.70063

Lahmeri, M. A., Kishk, M. A., & Alouini, M. S. (2021). Artificial intelligence for UAV-enabled wireless networks: A survey. *IEEE Open Journal of the Communications Society*, *2*, 1015-1040.
https://doi.org/10.1109/OJCOMS.2021.3075201

Long, Y., Gong, S., Sun, S., Lee, G., Li, L., & Niyato, D. (2024). Lyapunov-guided Deep Reinforcement Learning for Semantic-aware AoI Minimization in UAV-assisted Wireless Networks. *arXiv preprint arXiv:2409.13580*.
https://doi.org/10.48550/arXiv.2409.13580

Lei, T., Luo, C., Jan, G. E., & Bi, Z. (2022). Deep learning-based complete coverage path planning with re-joint and obstacle fusion paradigm. *Frontiers in Robotics and AI*, *9*, 843816.
https://doi.org/10.3389/frobt.2022.843816

Liu, S., Su, G., & Chen, B. (2024). Multi-UAV collaborative secure communication using multi-agent deep reinforcement learning. In *2024 7th International Conference on Electronics Technology (ICET)* (pp. 772-777). IEEE.
https://doi.org/10.1109/ICET61945.2024.10672964

Nemer, I. A., Sheltami, T. R., Belhaiza, S., & Mahmoud, A. S. (2022). Energy-efficient UAV movement control for fair communication coverage: A deep reinforcement learning approach. *Sensors*, *22*(5), 1919.
https://doi.org/10.3390/s22051919

Ojha, T., Raptis, T. P., Passarella, A., & Conti, M. (2023). Wireless power transfer with unmanned aerial vehicles: State of the art and open challenges. *Pervasive and mobile computing*, *93*, 101820.
https://doi.org/10.1016/j.pmcj.2023.101820

Prasad, N. L., Ekbote, C. A., & Ramkumar, B. (2021). Optimal deployment strategy for relay based UAV assisted cooperative communication for emergency applications. In *2021 National Conference on Communications (NCC)* (pp. 1-6). IEEE.
https://doi.org/10.1109/NCC52529.2021.9530098

Ren, Z., Hussain, K., & Faheem, M. (2024). K-means online-learning routing protocol (K-MORP) for unmanned aerial vehicles (UAV) adhoc networks. *Ad Hoc Networks*, *154*, 103354.
https://doi.org/10.1016/j.adhoc.2023.103354

Samir, M., Ebrahimi, D., Assi, C., Sharafeddine, S., & Ghrayeb, A. (2020). Leveraging UAVs for coverage in cell-free vehicular networks: A deep reinforcement learning approach. *IEEE Transactions on Mobile Computing*, *20*(9), 2835-2847.
https://doi.org/10.1109/TMC.2020.2991326

Seerangan, K., Nandagopal, M., Govindaraju, T., Manogaran, N., Balusamy, B., & Selvarajan, S. (2024). A novel energy-efficiency framework for UAV-assisted networks using adaptive deep reinforcement learning. *Scientific Reports*, *14*(1), 22188.
https://doi.org/10.1038/s41598-024-71621-x

Shaikh, M. S., Lin, H., Zheng, G., Wang, C., & Dong, X. (2024). Innovative hybrid grey wolf-particle swarm optimization for calculating transmission line parameter. *Heliyon*, *10*(19).
https://www.cell.com/heliyon/fulltext/S2405-8440(24)14586-0

Skarka, W., & Ashfaq, R. (2024). Hybrid machine learning and reinforcement learning framework for adaptive UAV obstacle avoidance. *Aerospace*, *11*(11), 870.
https://doi.org/10.3390/aerospace11110870

Ullah, Y., Roslee, M., Mitani, S. M., Sheraz, M., Ali, F., Osman, A. F., ... & Sudhamani, C. (2024). Reinforcement learning-based unmanned aerial vehicle trajectory planning for ground users' mobility management in heterogeneous networks. *Journal of King Saud University-Computer and Information Sciences*, *36*(5), 102052.
https://doi.org/10.1016/j.jksuci.2024.102052

Villanueva, A., & Fajardo, A. (2019). Deep reinforcement learning with noise injection for UAV path planning. In *2019 IEEE 6th International Conference on Engineering Technologies and Applied Sciences (ICETAS)* (pp. 1-6). IEEE.
https://doi.org/10.1109/ICETAS48360.2019.9117478

Wang, Y., & Farooq, J. (2023). Deep-reinforcement-learning-based placement for integrated access backhauling in UAV-assisted wireless networks. *IEEE Internet of Things Journal*, *11*(8), 14727-14738.
https://doi.org/10.1109/JIOT.2023.3344519

Yu, L., Qin, S., Zhang, M., Shen, C., Jiang, T., & Guan, X. (2021). A review of deep reinforcement learning for smart building energy management. *IEEE Internet of Things Journal*, *8*(15), 12046-12063.
https://doi.org/10.1109/JIOT.2021.3078462

Yue, X., & Zhang, W. (2018). UAV path planning based on k-means algorithm and simulated annealing algorithm. In *2018 37th Chinese Control Conference (CCC)* (pp. 2290-2295). IEEE.
https://doi.or/10.23919/ChiCC.2018.8483993

Zhu, L., Wang, J., Wang, Y., Ji, Y., & Ren, J. (2022). DRL-RNP: Deep reinforcement learning-based optimized RNP flight procedure execution. *Sensors*, *22*(17), 6475.
https://doi.org/10.3390/s22176475