Proceedings of the 35th European Safety and Reliability & the 33rd Society for Risk Analysis Europe Conference
Edited by Eirik Bjorheim Abrahamsen, Terje Aven, Frederic Bouder, Roger Flage, Marja Ylönen

doi: 10.3850/978-981-94-3281-3_ESREL-SRA-E2025-P3688-cd

UAV Swarm Coordination for Flood Area Coverage in Populated Regions using Reinforcement Learning

Maksym Meyer

Munich University of Applied Sciences HM, 80335 Munich. E-mail: meyer.maksym2@hm.edu

Sophia Schorer

Munich University of Applied Sciences HM, 80335 Munich. E-mail: sophia.schorer@hm.edu

Arkadiusz Żurek

Unmanned Air Vehicle (UAV) Section, Center for Advanced Systems Understanding Autnonomous Systems Division, Helmholtz-Zentrum Dresden-Rossendorf e.V. (HZDR), 02825 Görlitz. E-mail: a.zurek@hzdr.de

Marcin Hinz

Munich University of Applied Sciences HM, 80335 Munich. E-mail: marcin.hinz@hm.edu

Flooding is one of the most prevalent natural disasters worldwide and is increasingly recognized as a consequence of climate change. Floods cause substantial economic damage and, moreover, endanger human lives. We present a Deep Reinforcement Learning-based approach, using a centralized Proximal Policy Optimization (PPO)-based agent to coordinate a UAV swarm for the systematic identification of locations with a high likelihood of human endangerment. The agent acts adaptively based on the real-time coverage state, which is crucial for effective inspections of affected areas under a time constraint. We incorporate flood locations and areas of interest—defined by damaged infrastructure—into the decision-making process. We also present a method for extracting relevant data from satellite imagery, based on a previous flood event in the Ahr Valley in Germany in 2021. Our results demonstrate increasing effectiveness in the coverage of diverse flood scenarios. Further advancements are needed before real-world deployment, but the collected data could ultimately be crucial for planning rescue operations and mitigating human risks, especially during the initial disaster response phase. In addition to optimizing coverage efficiency, we highlight key operational risk factors in UAV swarms, such as unpredictable environmental conditions, communication disruptions, and energy constraints, which are essential considerations for ensuring reliable UAV swarm performance in real-world flood scenarios.

Keywords: *Unmanned Aerial Vehicles (UAV), Reinforcement Learning (RL), UAV Swarm Operations, Flood Response, Aerial Coverage, Autonomous Navigation, Emergency Management*

1. Introduction

Flooding is one of the most frequent and destructive natural disasters, exacerbated by climate change and rapid urbanization. The economic damage and human risks associated with floods necessitate efficient and scalable coverage systems. Unmanned Aerial Vehicles (UAVs) have emerged as a viable solution, offering rapid data collection capabilities for disaster response. The deployment of UAV swarms allows for large-scale scanning of flood-affected areas (Tubis et al., 2024). The increasing occurrence of flash floods, such as the catastrophic Ahr Valley flood in Germany 2021, has demonstrated the need for fast and accurate assessments of flooded regions. Traditional flood-assessment methods rely on satellite imagery and ground-based observations, which often fail to provide real-time, high-resolution data. In contrast, UAV-based approaches can deliver timely and detailed information to aid rescue operations and mitigate risks. In recent years, the advancement of UAV technology has significantly enhanced flood coverage, enabling rapid and precise real-time data collection, which is crucial for effective crisis management. However, coordinating UAV swarms in challenging environmental conditions, such as unpredictable weather or communication disruptions, presents significant challenges that require advanced management strategies and adaptive control algorithms. In this paper, we highlight operational risk factors in UAV swarms and possible solutions. Additionally, we provide an example of the optimization of time-limited UAV swarm flights in flood regions using Reinforcement Learning. The agent is trained to act adaptively based on the online coverage state, utilizing data from a historical flood event in the Ahr Valley in 2021.

2. UAV Swarm Deployment for Flood Response: Hardware, Risks, and Case Studies

The effectiveness of UAV swarms in flood response depends on multiple factors, including hardware capabilities, sensor integration, and data processing methods, which will be discussed in the following sub-chapters.

2.1. *Hardware Considerations*

Effective flood response using UAV swarms relies on a diverse set of sensors, each contributing to situational awareness, hazard assessment, and rescue operations. Different sensors can be used in UAV swarms in the context of flood response to collect real-time data. These include:

- LiDAR Sensors – Used often for topographic mapping and water level assessment, enabling precise 3D flood modelling. It is particularly valuable for high-resolution elevation mapping and large-scale flood analysis (Trepekli et al., 2022).
- Thermal Cameras – Allow detection of trapped individuals and warm bodies, improving rescue operations (Thiyagarajan et al., 2024).
- Multispectral and RGB Cameras – Used for classification problems (Zheng et al., 2022).
- Depth Sensors – Help estimate flood depth and flow velocity, supporting real-time hazard assessment. These include various technologies such as sonar (Frias 2023), structured light sensors, and stereocameras, which can complement LiDAR for more localized or underwater depth measurements (Pohl 2020).
- Hemispherical Cameras – Provide 360° situational awareness, enhancing UAV navigation and decision-making in complex urban areas (Rehman 2022).

Power efficiency is a key challenge in UAV deployment. The flight time of UAVs depends on factors such as battery capacity, payload weight, and environmental conditions. In addition to improving efficiency in terms of aerodynamics, propulsion systems, and the use of lightweight materials, a potential (and actively researched) solution is hydrogen fuel cell-powered UAVs, which could offer longer flight times, making them ideal for long-range flood mapping (Saravanakumar et al., 2023).

Hydrogen-powered UAVs have the advantage of higher energy density compared to traditional lithium-ion batteries, allowing for extended flight durations and increased operational range. This makes them particularly useful in large-scale disaster response scenarios, where continuous aerial monitoring is required.

Quadcopters generally have shorter flight times than their fixed-wing counterparts, but they are essential for conducting detailed inspections and scanning, especially when the mission requires hovering.

2.2. *Risk Factors and Mitigation Strategies*

The primary risk factors affecting UAV swarm operations include:

- Uncertainty in Environmental Conditions: Flooded areas present unpredictable conditions, such as sudden changes in water levels, strong wind currents, and debris, which can impact UAV flight stability and navigation (Surmann et al., 2022).
- Obstacle Avoidance and Collision Risk: UAVs operating in swarms must maintain safe distances to prevent collisions. Multi-UAV collision avoidance algorithms could integrate reinforcement learning-based policies to dynamically adjust flight paths (Garg A. and Jha S., 2023).
- Energy Constraints and Operational Failures: The endurance of UAVs is limited by battery capacity, and their operational time must be optimized for maximum area coverage. Strategies such as hybrid UAV deployment (fixed-wing and quadcopters) help balance endurance and manoeuvrability (Sonkar et al., 2022).
- Communication Disruptions: Maintaining reliable communication between UAVs and the ground control station is essential for

coordination. Failures in data transmission can lead to inefficient coverage and redundant flight paths (Chandran 2024).

Possible mitigation strategies include:

- Adaptive Trajectory Planning: UAVs must dynamically adjust their flight paths in unpredictable flood environments to maintain efficiency and avoid obstacles. Various methods, including reinforcement learning (RL), heuristic algorithms, and model-based optimization techniques, can enhance real-time decision-making and improve UAV swarm coordination (Garg 2023). These approaches enable drones to respond autonomously to environmental uncertainties, ensuring optimal area coverage.
- Redundancy and Fault-Tolerant Systems: Implementing backup communication channels and redundant UAV units can prevent mission failure in the event of system malfunctions (Tubis et al., 2024).
- Hybrid UAV Swarms for Enhanced Coverage: Combining fixed-wing UAVs for long-range mapping with quadcopters for detailed inspections improves overall efficiency and mitigates energy limitations (Tubis et al., 2024).
- Risk-Based Mission Planning: Pre-mission simulations using historical flood data and hydrological models help in planning UAV swarm deployment, reducing uncertainty in real-time operations (Chandran 2024).

In this paper, we use Reinforcement Learning as an adaptive method for trajectory optimization under the consideration of area coverage.

2.3. *Meuse and Rhine floods in July 2021*

In July 2021, extreme precipitation amounts resulted in severe flooding across the western European countries Belgium, Luxembourg, Germany and the Netherlands. The two German federal states North Rhine-Westfalia and Rhineland-Palatinate were particularly affected and recorded more than 180 casualties and a large amount of flood damage. The flood disaster resulted from the extreme extent of an underlying meteorological event that caused up to 150 litres of precipitation per square meter within 24 hours. The situation was aggravated by the fact that the soil in the affected regions was not able to absorb the amount of recorded precipitation due to several phases of repeated heavy rainfall in the preceding months. In the Ahr catchment the topography characterized by narrow valleys and a pronounced increase in altitude over a considerable distance contributed to the severe extent of the flood event. Due to the destruction of numerous bridges and infrastructure, access to affected area was rendered difficult and rescue and evacuation measures were assisted from the air (Tradowsky 2023). The difficulty of accessing areas affected by natural disasters raises the question of whether drones can be used in the future to cover flooded areas and are therefore suitable for supporting rescue operations.

3. Problem Description

In order to obtain an overview of the situation on the ground and being able to initiate rescue measures as quickly as possible, the aim of this work is to investigate whether a swarm of drones can be guided to entirely cover an area affected by flooding. Reinforcement learning is applied to find the optimal trajectories that maximize the covered area in a finite amount of time. The benefit of reinforcement learning-based guidance is that the drones are able to adapt to unknown and dynamic environments without the need for information about the situation in advance. A model of the environment simulates the flight dynamics that can be controlled to guide the drones over a region affected by flooding. The drone's location is specified in a grid-world environment that provides information about cells that are either affected by a flood or not. A number of $n = 3$ drones is employed to demonstrate results for the underlying problem.

3.1. *Flood Environment*

The area coverage problem is addressed based on data about the severe floodings that affected parts of Germany in July 2021. During the emerging floods information about the spatial extent as well as a mapping of the damage grade in the affected regions was collected using Copernicus satellites.

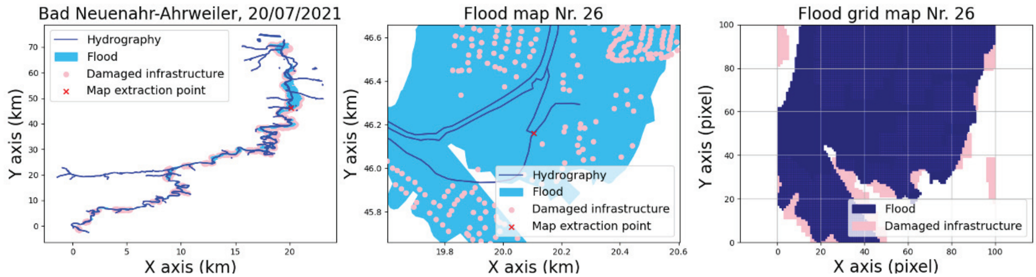

Fig. 1. Flood grid map generation for the flood situation as of 20th of July 2021 in Bad Neuenahr-Ahrweiler, based on data from the European Union's Copernicus Emergency Management Service (Copernicus EMS 2021)

The data is freely accessible and provided by the European Union's Copernicus Emergency Management Service (EMS) (Copernicus EMS 2021). A significant amount of precipitation and subsequent damage incurred due to flooding has been observed in the Ahr catchment area located along the Ahr river in the German federal state Rhineland-Palatinate. The Copernicus data on flooding in the Bad Neuenahr-Ahrweiler area covers the areas along the Ahr river from the district Ahrdorf in the German federal state North Rhine-Westphalia at the southernmost point to the town Sinzig at the northernmost point of the Ahr, as well as the areas along the Ahr tributary Adenauer Bach. The data depicts the static flood scenario at particular days, whereby in this work the data about the flooding situation as of the 20th of July 2021 is employed.

In order to derive maps that depict the flooded areas in a grid world environment, the information about the geographic coordinates of the boundaries of flooded regions is transformed to Universal Transverse Mercator (UTM) coordinates. Maps of the size 1 km × 1km are extracted in 1 km intervals along the Ahr river lines and are subsequently discretized in a grid with resolution of res = 10 m. The discretised maps contain information about the presence of a flooded or non-flooded cell and represent floods of various shape and size.

The maps are superimposed with information on damaged or destroyed buildings in the populated areas of the flooded areas. For every cell containing at least one damaged building, adjacent cells in a surrounding area of 100 meters are also considered to be affected. Cells that indicate either floodings or damaged infrastructure are denoted target cells in the following and are treated equally in their information. The process of the generation of flood grid maps for the area Bad Neuenahr-Ahrweiler is depicted in Fig. 1.

3.2. *Flight Dynamics*

To reduce the complexity of the fixed-wing UAV flight dynamics, the drones are assumed to fly at a constant altitude $h = 200$ m and with constant speed $v = 25$ m/sec with respect to a fixed, inertial reference frame. A low-order model is employed that approximates the flight dynamics using simple kinematic equations and without considering aerodynamic and control forces that act on the vehicle in different flight phases. The relationship between the aircraft's bank angle φ and course angle change $\dot{\chi}$ in the absence of wind or sideslip is given by the equation for the coordinated turn as

$$\dot{\chi} = \frac{g}{v} \tan \varphi \,, \tag{1}$$

where g represents gravitational acceleration (Beard, R. and T. Mclain 2012). The position change of the vehicle along the x and y axis of the inertial reference frame results as

$$\dot{x} = v \cos \chi \,, \quad \dot{y} = v \sin \chi \tag{2}$$

In order to derive the drone's position and course angle under specified initial conditions, the ordinary differential equations can be solved at discrete timesteps using the Euler method.

3.3. *Field of View*

To map the environment and gather information about the emergency situation on the ground, the drones are equipped with cameras. The covered area of the camera on the ground is dependent on the angle of the cameras Field of View FoV and the flight altitude h (see Fig. 2). Assuming that the camera is always facing downwards, and the drone flies with a pitch angle of $\theta = 0$ deg, the coverage area on ground is a square defined by the width w:

$$w = 2h \tan\left(\frac{FoV}{2}\right). \tag{3}$$

In this work the angular Field of View is specified to $FoV = 50$ deg.

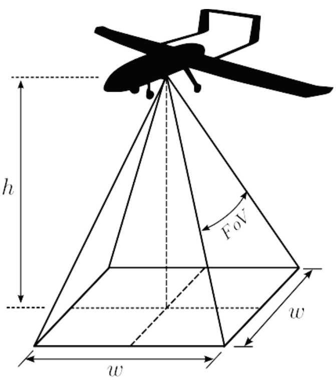

Fig. 2 Field of View and covered area on the ground

4. Partially Observable Markov Decision Problem (POMDP)

Reinforcement Learning Problems in general can be formalized using a Markov Decision Problem (MDP). The basic idea of the MDP is to capture the most important aspects of the problem facing a learning agent that interacts with an environment to achieve a goal (Sutton, R. and A. Barto 2018). An assumption in MDPs is that the agent can fully observe the state of the environment in each timestep. In applications where the agent only has partially knowledge about the environment, the partially observable MDP (POMDP) is a suitable framework. The POMDP can be described according to (Kaelbling 1998) as a tuple $\langle S, A, T, R, \Omega, O\rangle$, where

- S is a finite set of states of the world,
- A is a finite set of actions,
- $T: S \times A \rightarrow \Pi(S)$ is the state transition function that gives to each world state and agent action a probability distribution over world states,
- $R: S \times A \rightarrow \mathbb{R}$ is the reward function that gives the expected immediate reward gained by the agent for taking each action in each state,
- Ω is a finite set of observations the agent can experience of its world,
- $O: S \times A \rightarrow \Pi(\Omega)$ is the observation function, which gives for each action and resulting state a probability distribution over possible observation.

The goal of the POMDP is to maximize the amount of reward R the agent can expect to accumulate over the future depending on the actions it performs. The rule by which the agent selects actions a ϵ A as a function of the belief state b is provided by the policy $\pi(\mathrm{a}|\mathrm{b})$. The computation of the optimal policy is based on the knowledge about the state-transition probabilities provided by T and an estimation of the current belief state based on the received observations o ϵ Ω provided by the observation function O.

4.1. *States*

The state $s_t^{(k)} \in S$ of the k-th drone at time step t, with $k = 1, \ldots, n$, is defined by the position $(x_{g,t}^{(k)}, y_{g,t}^{(k)})$, in the grid world environment

$$x_{g,t}^{(k)} = \frac{x_t^{(k)}}{res}, \quad y_{g,t}^{(k)} = \frac{y_t^{(k)}}{res}, \tag{4}$$

the bank angle $\varphi_t^{(k)}$ and course angle $\chi_t^{(k)}$. The drone is positioned in a flood scenario that is modelled in form of a binary flood grid map $F \in \mathbb{N}^{m_w \times m_h}$ of width and height $m_w = m_h = 100$ pixels:

$$s_t^{(k)}: \left\{x_{g,t}^{(k)}, y_{g,t}^{(k)}, \varphi_t^{(k)}, \chi_t^{(k)}, F\right\}.$$

4.2. *Actions*

In each timestep the agent can choose to perform a discrete action $a_t^{(k)} \in A$ that either increases or decreases the drones bank angle $\varphi_t^{(k)}$ by $\Delta\varphi_t^{(k)} = \pm 20$ degrees or maintains the current bank angle:

$$a_t^{(k)}\colon \left\{+\Delta\varphi_t^{(k)}, 0, -\Delta\varphi_t^{(k)}\right\}.$$

Bank angles are limited to -60 deg $\leq \varphi_t^{(k)} \leq 60$ deg.

4.3. *Observations*

The agent has full knowledge about each drone's position, bank angle and course angle in every timestep. Due to the limited field of view of the camera, pictures taken at each timestep capture only an incomplete state of the environment. Active exploration is required so that the drones get an overall picture of the underlying flood scenario. The observation $o_t^{(k)} \in O$ of the k-th drone at each timestep includes a processed image $P_t \in \mathbb{N}^{m_w \times m_h \times m_c}$ with a number of $m_c = 3$ channels:

$$o_t^{(k)}\colon \left\{x_{g,t}^{(k)}, y_{g,t}^{(k)}, \varphi_t^{(k)}, \chi_t^{(k)}, P_t\right\}.$$

The first channel $P_{1,t}$ encodes binary information about cells that are completely unknown up to the current time step, the second channel $P_{2,t}$ about cells that have already been observed and are target cells and finally the third channel $P_{3,t}$ about cells that have been observed and are no target cells. To refer to a single cell within the image $c_{i,j,l,t}$, for $i = 1,\dots,m_w$, $j = 1,\dots,m_h$, $l = 1,\dots,m_c$, is denoted as the i,j-th cell of the l-th image channel at timestep t in the following.

4.4. *Rewards*

The goal is to maximize the coverage of the target area within a finite time horizon, defined for example by a battery constraint. To fasten convergence, reward shaping is used, where the agent receives an auxiliary, frequent learning signal. We prioritize the coverage of new target cells, while non-target cells receive a reduced reward at half the value of target cells. Boundary violations are penalized in two ways. If an out of boundary event occurs, the episodes are terminated prematurely, stopping the agent from collecting further rewards. On top of that, it receives a penalty for the relevant action. Overall, increased coverage and compliance with the boundaries and therefore the assigned airspace and local obstacle requirements are rewarded. Consequently, the overall reward R_t in each time step is the sum of the following components:

$$R_{1,t} = \sum_{k=1}^{n} R_{1,t}^{(k)} \tag{5}$$

$$R_{2,t} = 0.005 \sum_{i=1}^{m_w}\sum_{j=1}^{m_h}\left(c_{i,j,2,t} - c_{i,j,2,t-1}\right) \tag{6}$$

$$R_{3,t} = 0.0025 \sum_{i=1}^{m_w}\sum_{j=1}^{m_h}\left(c_{i,j,3,t} - c_{i,j,3,t-1}\right), \tag{7}$$

with:

$$R_{1,t}^{(k)} = \begin{cases} -0.1, & \text{if } x_{g,t}^{(k)} \text{ or } y_{g,t}^{(k)} < 0 \\ -0.1, & \text{if } x_{g,t}^{(k)} \text{ or } y_{g,t}^{(k)} > m_w \\ 0, & \text{otherwise}. \end{cases} \tag{8}$$

The used factors were determined heuristically. For simplification, the drones are initialized in fixed positions on a circle, each heading towards the centre of the grid with uniform angular distances. We do not consider collisions in the formulation of the reward function.

5. Model Architecture

The policy π to coordinate the drone swarm in the flood area is obtained using Proximal Policy Optimization (PPO). In this method, data which includes states, actions, and rewards, is sampled from the current version of the policy by interacting with the environment. A surrogate policy gradient objective L_{CLIP} is then used to find new parameters θ of the policy using stochastic optimization. $\widehat{E_t}$ denotes the expectation based on the collected samples of a batch (Schulman 2017). $\widehat{A_t}$ is the advantage function estimate, which predicts the advantage of taking an action a_t in a given state s_t, comparing it to the expected future rewards under its default behavior (Schulman 2016).

$$L_{\text{CLIP}} = \widehat{\mathrm{E}}_{\mathrm{t}}\left[\min\left(\mathrm{r_t}(\theta)\widehat{\mathrm{A}}_{\mathrm{t}}, \mathrm{clip}\left(\mathrm{r_t}(\theta)\right)\widehat{\mathrm{A}}_{\mathrm{t}}\right)\right] \tag{9}$$

PPO is less prone to stability issues from policy updates by clipping the probability ratio $r_t(\theta)$ to

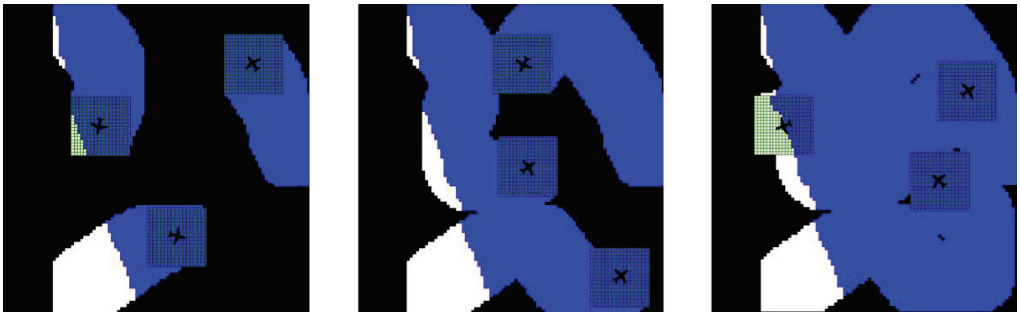

Fig. 3 Trajectory sample during policy execution in the evaluation phase. ($t_1 = 15s$, $t_2 = 30s$, $t_3 = 45s$)

the interval of $[1-\epsilon, 1+\epsilon]$, where ϵ denotes the clipping parameter. (Schulman 2017)

$$r_t(\theta) = \frac{\pi_\theta(a_t \mid s_t)}{\pi_{\theta_{\text{old}}}(a_t \mid s_t)} \qquad (10)$$

We use a centralized agent for the on-line decision making of the swarm, meaning that individual drone actions are derived from a global swarm state. Similar to the approach shown in (Baldazo 2019), we process a two-sided observation space, consisting of the drones positions, bank angles, headings and a coverage map. We scale angular observations by using sine and cosine and divide the position by the grid dimension. The coverage map is categorically encoded into three channels: The unknown cells, the known unflooded or unpopulated cells, and the known cells of interest. During the beginning of the task, all cells are unknown. As the coverage task progresses, the distribution of cells in each channel changes. To leverage spatial features, we process the 3-channel coverage image at every time step using a Convolutional Neural Network (CNN) with three 5×5 convolution layers that yield 8, 16, and 32 feature maps with ReLU activation and max pooling. These feature maps are flattened and passed through two fully connected layers. Meanwhile, drone states are processed separately via two fully connected layers. The outputs from both networks are concatenated into a 256-dimensional vector and then fed through three additional fully connected layers to produce the action probabilities and value function estimate.

6. Training and Results

The parameters used for training are shown in Table 1. 2.5 million steps are run to learn the optimal actions in the given environment. We change the underlying target area map randomly from a selection of 100 target area maps to encourage generalization.

Table 1. Training parameters

Parameter	Value
Learning rate	3e-4
Discount factor (γ)	0.99
Generalized Advantage Estimate	0.95
Clipping range	0.2
Rollout buffer size	512 samples per environment (16 parallel environments)
Batch size	1024
Episode length	90 steps (45s)
Action update interval	0.5s
Optimization algorithm	ADAM (Kingma 2014)

The plot in Fig. 4 shows an upward trend in the moving average of the target area coverage during training, indicating that the swarm is navigated more efficiently within the given time span. Due to changing environments, features between episodes and exploration noise, variations of the raw values are expected.

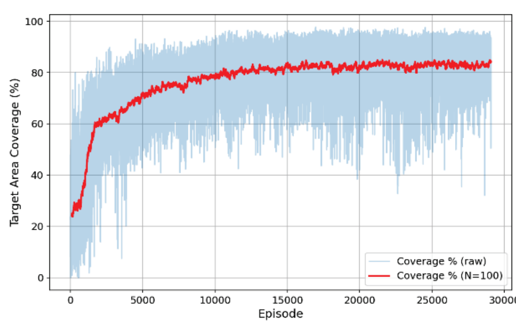

Fig. 4 Moving Average of Target Area Coverage

After training, we deploy the resulting policy to 25 previously unused maps. We execute the policy in a deterministic setting, where the actions of highest probability are chosen for the drones in every step. The swarm shows an average target area coverage of μ=87.02% and a standard deviation of σ=4.51%, indicating that the learned behaviour from the training stage transfers well to other environments. A trajectory sample during policy execution with the learned network weights is shown in Fig. 3.

7. Conclusions and outlook

In this paper, we present a framework to process historic flood data based on satellite imagery for use in a Reinforcement learning coverage optimization environment. We utilize a centralized architecture, combining a MLP (for drone-specific observations) and a CNN (for the coverage image) to process the state information and pass it to a PPO algorithm, to obtain an effective policy. The agent shows improving capabilities at coordinating the drone swarm starting from fixed positions, increasing the covered area of interest within a finite time span and therefore also optimizing the usage of battery resources. Future research includes prioritization of high-risk areas for improved guidance, considering the requirements of an underlying search & rescue task, as well as improved UAV dynamics, sensor models and an energy model, reducing the gap to a real application.

References

Baldazo, D. et al. (2019): Decentralized Multi-Agent Deep Reinforcement Learning in Swarms of Drones for Flood Monitoring. In: 2019 27th European Signal Processing Conference (EUSIPCO). A Coruna, Spain: IEEE.

Beard, R. and T. Mclain (2012): Small Unmanned Aircraft: Theory and Practice, Princeton Univ. Press. DOI: 10.1515/9781400840601.

Chandran, I. and Vipin, K. (2024): Multi-UAV Networks for Disaster Monitoring: Challenges and Opportunities from a Network Perspective. Drone Systems and Applications.

Copernicus EMS (2021): Copernicus Emergency Management Service (© 2021 European Union) [EMSR517], Bad Neuenahr-Ahrweiler: Grading Monitor.

Frias, M. et al. (2025): Combining UAS LiDAR, Sonar, and Radar Altimetry for River Hydraulic Characterization. Drones 9, 31.

Garg, A. and Jha, S. (2023): Deep deterministic policy gradient-based multi-UAV control for moving convoy tracking. Engineering Applications of Artificial Intelligence, 126, 107099.

Kaelbling, L. et al. (1998): Planning and acting in partially observable stochastic domains. In *Artificial Intelligence* 101. DOI: 10.1016/S0004-3702(98)00023-X.

Kingma, D. and J. Ba (2015): Adam: A Method for Stochastic Optimization. In International Conference on Learning Representations.

Pohl, D. and Dorodnicov, S. (2020): Depth Map Improvements for Stereo-based Depth Cameras on Drones. RealSense, Intel Corporation.

Rehman, N. (2022): Hemispherical Photographs: A Review of Acquisition Methods and Applications in the Context of Urban Energy and Environment Assessments. ASME Open Journal of Engineering, 1, 010801.

Saravanakumar, Y. et al. (2023): Power Sources for Unmanned Aerial Vehicles: A State-of-the-Art Review. Applied Sciences, 13, 11932.

Schulman, J. et al. (2016): High-Dimensional Continuous Control Using Generalized Advantage Estimation. In International Conference on Learning Representations.

Schulman, J. et al. (2017): Proximal Policy Optimization Algorithms. Available online at http://arxiv.org/pdf/1707.06347.

Sonkar, S. et al. (2023): Design & Implementation of an Electric Fixed-wing Hybrid VTOL UAV for Asset Monitoring. J. Aerosp. Technol. Manag., v15, e0823.

Surmann, H. et al. (2022): Deployment of Aerial Robots during the Flood Disaster in Erftstadt / Blessem in July 2021.

Sutton, R. and A. Barto (2018): Reinforcement Learning: An Introduction. 2nd ed.: MIT Press Ltd.

Thiyagarajan, S. et al. (2024): Drones in Human Aid and Disaster Relief using Thermal Imaging Technology. International Research Journal on Advanced Engineering and Management.

Trepekli, K. et al. (2022): UAV-borne, LiDAR-based elevation modelling: a method for improving local-scale urban flood risk assessment. Natural Hazards, 113:423–451.

Tradowsky, J. et al. (2023): Attribution of the heavy rainfall events leading to severe flooding in Western Europe during July 2021. In *Climatic Change*. DOI: 10.1007/s10584-023-03502-7.

Tubis, A. et al. (2024): Risks of Drone Use in Light of Literature Studies. Sensors 24, 1205.

Zheng, Z. et al. (2024): Fusion of UAV-Acquired Visible Images and Multispectral Data by Applying Machine- Learning Methods in Crop Classification. Agronomy 14, 2670.