Integrating swarm intelligence with bayesian networks for continuous UAV-based surveillance in dynamic environments

Master of Science Thesis

J.H.J. Knuyt

TU Delft

Integrating swarm intelligence with bayesian networks for continuous UAV-based surveillance in dynamic environments

Master of Science Thesis

by

J.H.J. Knuyt

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on 21 April 2022.

Student number:	4601963	
Project duration:	August 2020 – April 2022	
Thesis committee:	Dr.ir. B.F. Santos	TU Delft, Chairman
	Dr. O.A. Sharpans'kykh	TU Delft, Supervisor
	Dr.ir. E. van Kampen	TU Delft, Examiner

TU Delft

Preface

Dear reader,

During the summer of 2020 I was fortunate to find a thesis opportunity concerning UAV-based surveillance, which is a topic I personally find particularly interesting for its operational challenges and for the various possibilities of future applications. This thesis was written in partial fulfilment of the requirements for the degree of Master of Science in Aerospace Engineering at the Delft University of Technology.

I would like to express my gratitude to Alexei Sharpans'kykh for being both a passionate, supporting and challenging supervisor. Despite the challenges of working from home and discussing the technical aspects of this thesis in-depth through online meetings, I like to think that we successfully managed these challenges. I thoroughly enjoyed our collaboration and appreciate the feedback and the personal discussions that we had. I would also like to express my sincere gratitude to my family, friends and fellow students, who provided all the support I needed to stay motivated. Specifically, I would like to thank Jari, Sabine and Vincent for having my back during the entire process, your support definitely made working on my thesis more pleasant.

This thesis concludes my time as a student at the Delft University of Technology. I like to think that I have grown during this journey, both on an academic and on a personal level. There are countless beautiful moments to cherish from my time in Delft, thanks to my fellow students.

J.H.J. Knuyt
Delft, 21 April 2022

Contents

List of Figures

List of Tables

List of Abbreviations

ABM	Agent Based Modelling
ABN	Artificial Bayesian Network
ACO	Ant Colony Optimisation
ACOSG	Adaptive Collaborative Opportunistic Security Game
APF	Artificial Potential Field
APFM	Artificial Potential Field Map
ASDI	Average Standard Deviation of Idleness
ASU	Adaptive Subjective Utility
BAPS	Bayesian Ant-based Patrol Strategy
BN	Bayesian Network
BR	Bounded Rationality
BRQR	Bounded Rational Quantal Response
COSG	Collaborative Opportunistic Security Game
CPP	Coverage Path Planning
CPT	Conditional Probability Table
DACLD	Dynamic Ant Colony's Labour Division
DAG	Directed Acyclic Graphs
DBN	Dynamic Bayesian Network
DTCM	Discrete-Time Markov Chain
EM	Expectation Maximisation
FOV	Field of View
FRTM	Fixed Response Threshold Model
GAI	Global Average Idleness
GSA	Global Sensitivity Analysis
GSG	Green Security Game
HAPF-ACO	Hybrid Artificial Potential Field and Ant Colony Optimisation
HAPF-BLACOPS	Hybrid Artificial Potential Field and Bayesian Learning Ant Colony Optimisation Prioritised Surveillance
KDE	Kernel Density Estimation
KPI	Key Performance Indicator
LHS	Latin Hypercube Sampling

LPM	Local Pheromone Map
LRM	Local Rhino Map
LWPR	Locally Weighted Projection Regression
MAL	Multi-Agent Learning
MLE	Maximum Likelihood Estimation
MT	Mission Type
MTD	Mean Time between Threat Detection
NDVI	Normalised Difference Vegetation Index
NP	Non-deterministic Polynomial Time
PD	Probability of Detection
PND	Probability of Non-Detection
QMI	Quadratic Mean of Intervals
QR	Quantal Response
RSF	Resource Selection Function
RTM	Risk Terrain Modelling
SS	Spatial Scalability
SSG	Stackelberg Security Game
ST	Scalability of Team Size
SU	Subjective Utility
SUQR	Subjective Utility Quantal Response
TPM	Target Probability Map
UAV	Unmanned Aerial Vehicle
VNN	Von Neumann Neighbourhood
VTOL	Vertical Take-Off and Landing
WGAI	Weighted Global Average Idleness

Introduction

Illegal animal poaching has been an active trade for decades. It poses serious threats to the survival of iconic animal species and causes economic problems in local communities due to an absence of wildlife. Surveillance has previously been identified to be effective at preventing a loss of wildlife due to poaching. The characteristics and increasing capabilities of unmanned aerial vehicles (UAVs) provide new, innovative possibilities to address the problems of illegal animal poaching.

This thesis proposes a novel multi-UAV surveillance model to prevent the loss of wildlife in national parks. While UAV technology has previously been used to provide surveillance, these models require further development to fully exploit their potential. An autonomous multi-UAV surveillance model that integrates swarm intelligence with bayesian networks to predict poacher activity is therefore proposed to address adaptive poacher behaviour and prevent the loss of wildlife. This research analyses the effects of the bayesian network on surveillance effectiveness and efficiency. In addition, the influences of online bayesian learning and adaptive poacher behaviour on surveillance performance is considered in order to prevent a potential loss in surveillance effectiveness in the long-term. The emergent behaviour is evaluated in a case study concerning Aloegrove Safari Park in Namibia and simulated using the agent-based modelling paradigm.

This thesis report is organised as follows. In Part I, the scientific paper of this research is presented, after which the appendices relevant to this paper are presented in Part II. These appendices provide elaboration on the proposed model, as well as additional simulation results and statistical analyses of the performed simulations. Finally, Part III contains the relevant literature study that supports this research.

I

Scientific Paper

Integrating swarm intelligence with bayesian networks for continuous UAV-based surveillance in dynamic environments

J.H.J. Knuyt

Under supervision of Dr. O.A. Sharpans'kykh

Air Transport and Operations, Faculty of Aerospace Engineering
Delft University of Technology, Delft, The Netherlands, 21 April 2022

Abstract

The loss of wildlife due to illegal poaching activity poses threats on both the survival of iconic animal species and the livelihood of local communities. This research proposes a distributed surveillance model in which a UAV swarm autonomously coordinates continuous surveillance in a dynamic environment. The adaptive behaviour of poachers has the potential to negatively affect surveillance performance and is therefore taken into consideration through the proposed ACOSG model. The novelties of this research are twofold. A mission selection algorithm is proposed that addresses the deficiencies of the existing HAPF-ACO model while improving on surveillance effectiveness. Bayesian learning is applied to dynamically prioritise surveillance efforts of the proposed HAPF-BLACOPS model. Additionally, the learning rate of both poachers and UAVs is analysed to determine whether surveillance remains effective in response to adaptive poacher behaviour. Simulation results show that the proposed model significantly outperforms the current state-of-the-art HAPF-ACO model. Prioritisation of surveillance efforts is achieved through the use of (A)BNs, such that coverage of the target area is reduced by 30%, while maintaining the surveillance effectiveness of the current state-of-the-art. It is found that the interactions between the APF and ACO modules limit the extent to which the (A)BNs' predictions influence the UAVs' spatiotemporal patterns and therefore limit the effects of prioritisation on surveillance performance as well. The adaptive capabilities of the poachers and the learning rate of the UAV swarm do not significantly affect surveillance performance and the loss of wildlife, due to a limited amount of newly gained experience. Future research opportunities are identified that can improve the influence of the (A)BN on surveillance performance and prioritisation.

1 Introduction

Animal poaching has been an active trade for decades [1, 2]. The illegal trade in wildlife is considered to be the fourth largest criminal industry worldwide [3, 4] and also provides financial support to terrorist organisations [5, 6, 7]. Poaching therefore poses risks to national security, along with serious threats to the survival of iconic animal species [8, 9]. Additionally, local communities face economic problems as a result from a decline in tourism activities, caused by the absence of wildlife [10].

As early as 1990 it was concluded that law enforcement and adequate manpower were required to effectively fight poachers. Detection of poaching activities was found to be most effective at preventing a loss of wildlife [1]. Effective and innovative solutions are required to address the lack of available resources and to address these poaching activities in national parks [11, 12]. Unmanned Aerial Vehicles (UAVs) are commonly called upon to address this topic [13]. Their advantage is their capability to carry a large variety of sensors, their increasing affordability [14, 5, 15] and their ability to monitor remote target areas [16]. The disadvantage of traditional surveillance strategies that rely on ground patrols, is the inability to monitor vast national parks in their entirety [17]. Their observations regarding poaching activity are therefore biased [18].

Current state-of-the-art surveillance models provide evolving coordination of distributed and continuous surveillance through learning from successes. These models usually do not consider dynamic and adaptive targets such as poachers. Continuous adjustment of surveillance efforts is needed to address such evolving crime patterns [19], in order to prevent a loss of effectiveness as a result from initial model parameters not remaining applicable under changing poacher behaviour [20, 21, 22]. Such complex systems with adaptive intruders, where continuous and distributed surveillance is needed, have not been explored sufficiently. Neither have the adaptive and coordination mechanisms of poacher behaviour been integrated in a single model. Current state-of-the-art poacher decision models use Game Theory to accurately describe these behavioural aspects of their decision making

process on an individual basis. The poachers' objective has not been simulated as dynamic input for the UAVs either, which results from wildlife movement ecology. Previously proposed surveillance models therefore do not actively consider the driving factors of the intruders to optimise path planning. This two-sided surveillance problem, which is a dynamic environment in which both the intruders and the defenders adapt to each other's behaviour, remains a relatively unexplored area of research [23]. This leaves opportunities for improvement by realising learning UAV-based surveillance that can dynamically prioritise surveillance efforts through distributed coordination in response to intelligent intruders. This addresses the potential influence of adaptive poacher behaviour on long-term surveillance effectiveness, which has been disregarded in previous research.

This research explores the effects of online Bayesian learning in a two-sided surveillance problem on surveillance performance. The proposed HAPF-BLACOPS model utilises swarm intelligence to provide a distributed approach to the coordination of autonomous and continuous multi-UAV surveillance. It builds on the current state-of-the-art HAPF-ACO model by using Bayesian Networks (BNs) for online learning and probabilistic inference regarding poacher activity. These BNs are characterised by a graphical representation (directed a-cyclic graph) of the joint probability distributions (JPDs) to describe poacher behaviour through random variables. Online learning can be achieved by updating the underlying conditional probability tables (CPTs) when new evidence becomes available. These BNs fit the characteristics of the problem at hand, given their ability to identify causal relationships and perform risk estimation through probabilistic inference [24, 25]. This knowledge of poacher behaviour is used to allocate available resources towards locations with relatively high probabilities of criminal activity [26, 27, 28]. The UAV swarm in the proposed model achieves adaptive behaviour through coordination of the UAVs' local knowledge of the environment. Multi-agent learning (MAL) therefore becomes an emergent property of the UAV swarm [29] that depends on the exchange of newly obtained information regarding poacher activity [20, 21, 28]. The agent-based modelling (ABM) paradigm provides a framework to define the behavioural properties and interactions of model elements from a bottom-up approach. It is therefore used to formalise the interactions of the proposed surveillance model, to simulate the presence of adaptive poachers through the proposed ACOSG model and to simulate wildlife movement ecology of a black rhino population through a relatively simple resource selection function (RSF).

The novelties of this research are twofold. First, distributed BNs are integrated with the ant colony optimisation (ACO) module within the proposed HAPF-BLACOPS model in order to consider the poachers' behaviour in improving surveillance effectiveness by enforcing dynamically adjusted prioritised surveillance. Second, this proposal aims to prevent a loss in long-run effectiveness of surveillance performance resulting from initial model parameters not remaining applicable under adaptive poacher behaviour. The effects of online learning on surveillance performance in a two-sided surveillance problem are therefore analysed.

The remainder of this work is structured as follows. Section 2 elaborates on relevant research within the fields of autonomous surveillance, poacher behaviour and rhino behaviour models. Section 3 then introduces the methodology used in this work, after which Section 4 formalises the ABM to simulate the proposed models. Section 5 elaborates on the calibration process of model parameters, validation of model output and on the simulation results. Section 6 discusses these simulation results, after which conclusions are drawn and future research opportunities are identified in Section 7.

2 Preliminary work

This section provides an overview of academic work relevant to this research. It is structured such that Section 2.1 introduces the various concepts within the topic of surveillance, after which active research areas within the field of automated and learning surveillance are discussed. Section 2.2 familiarises the reader with developments within the field of Game Theory to model human decision making, after which Section 2.3 introduces concepts to simulate animal movement ecology.

2.1 Autonomous UAV-based surveillance

This section elaborates on the classification of surveillance problems first, after which the most relevant work regarding surveillance is discussed. We do not provide a comprehensive summary and categorisation of previously performed work within this extensive field of research, For detailed overviews of research within this field, we like to invite the reader to refer to [23, 14, 30, 31, 32].

The surveillance problem at hand is known as the coverage path planning problem and originates from the second world war, when the objective was to find submarines [23]. The topic at hand is not to be confused with the coverage problem, where the objective is to find optimal, static locations for a set of sensors that

are to continuously monitor a target area. Coverage path planning addresses the robotics' subtopic of motion planning, where each robot is assigned to a path such that the entire target area is explored within the capabilities of the robots, while avoiding obstacles. The paths are limited by the robot's capabilities in terms of kinematics, payload, endurance, communication and/or sensor constraints [32, 30, 33]. Another distinction is to be made between discrete and continuous surveillance, as this work focuses on the latter. Discrete surveillance is terminated once the entire target area has been searched or once all intruders within the target area are found [34, 35]. The objective in discrete surveillance is to minimise the total search time required to cover the entire target area, or to find all intruders [36]. Continuous surveillance is a generalisation of the surveillance problem that has not been explored as much. Persistence is required within the context of poachers, since an additional challenge arises in the behaviour of poachers that can enter a target area at different locations at unknown times. This means that the set of intruders to be found evolves over time and locations that were previously covered are to be revisited to achieve effective surveillance [30]. Continuous surveillance focuses on periodically revisiting locations and usually minimises the largest revisit time [23], also referred to as the idleness, refresh time, latency or age. This is defined as the time between two consecutive visits to a location within the target area [30, 37, 32]. Sub-optimal solution techniques are usually applied since the continuous surveillance problem is NP-hard [38]. Finally, a distinction between solution approaches can be made by identifying systems capable of global communication and bio-inspired systems driven by local autonomy [34]. Such distributed approaches do experience the disadvantage of not guaranteeing exhaustive coverage and the need to overcome local minima [35, 39]. Local autonomy translates to each UAV having local knowledge regarding the system and its environment, that is obtained from short range communication. The accuracy of local knowledge varies due to the communication constraints, but it is utilised to make decisions regarding path planning. Such systems can overcome challenges arising due to the absence of digital infrastructure in extreme environments where law enforcement aims to find poachers in vast target areas [40]. Since this research focuses on utilising learning and autonomous surveillance to detect poaching activities in vast, hard-to-reach areas, the focus of this research lies with continuous coverage path planning (surveillance) that utilises local autonomy. Local communication between the swarm members of a distributed model and their interactions with the environment cause emergent behaviour at a macro level, which can be observed from a global point of view [41, 42].

The Green Security Game (GSG) is a concept that provides a Game Theory based approach to continuous surveillance that aims to prevent illegal poaching activity. It is a field of research that has received substantial attention. Optimal patrol routes can be computed through the previously developed PAWS [43], OPERA [44] or CAPTURE [45] frameworks, by simulating the interactions between poachers and rangers as a repeated strategy game. These centralised solution methods render these techniques ineffective for online path planning [46] in scenarios with an absence of digital infrastructure to support the required communication channels [40]. A continuous surveillance approach that is to be implemented within the context of this research is required to be robust and able to operate in the presence of limited communication capabilities [47, 48]. There is also a lack of consideration of ecology movement within these Game Theory based solutions [49].

Most research towards (multi-agent) surveillance addresses this topic with predetermined patrol routes [50, 51, 52, 53, 54, 48, 55] or uses space decomposition to create patrol routes [56, 57, 47]. These approaches also lack online path planning and distributed coordination, which can be achieved through local autonomy utilising artificial potential fields, swarm intelligence and behaviour-based methods [58, 59]. Distributed approaches are also advantageous over centralised models in terms of adaptability, simplicity, modularity, reliability, robustness and low communication bandwidth requirements [30, 42].

The common approach in distributed models to achieve continuous surveillance is to toggle the state of grid cells between an explored and unexplored state. The period of time after which this state is to be switched is to be tuned carefully. Toggling the state too early prevents the swarm from covering the entire target area [35], whereas re-exploration is performed too late if the opposite occurs. In distributed systems, elements perform internal decision making without any centralised control point. Ant colony optimisation (ACO) achieves this through evaporating pheromone concentrations. Each swarm member has a certain environmental awareness through its own unique digital pheromone map. The pheromone concentrations on this map are the product of a target location's priority and the location's idleness. Collaborative surveillance is achieved by broadcasting these pheromone maps to nearby swarm members. The local autonomy in turn motivates each swarm member to plan their path according to the pheromone concentrations [34, 35].

Prioritisation of surveillance efforts has proven to be effective in fighting crime [26]. Such a concept is proposed in [37], where continuous surveillance of a target area is performed in a non-uniform manner by specifying the pheromone evaporation rates to be proportional to a grid cell's relative importance [12]. The proposed BAPS model achieved stochastic path planning through a greedy algorithm that computes transition probabilities from the pheromone concentrations and the distance from a target location. In [60] the swarm members use

gained experience regarding successfully completed tasks to adjust response thresholds for specific tasks. The transition probabilities are therefore adjusted using learning and forgetting rates, depending on the agent's experience. This is a different approach to gradually improve the effectiveness and success rate of the overall surveillance mission objective through implicit coordination.

The limitation of these proposals is that the objective of the intruder that is to be found is not considered during path planning. Neither do these approaches incorporate operational and/or kinematic constraints that apply to the swarm members. These aspects can influence surveillance performance negatively [61]. The HAPF-ACO model addresses a number of these limitations by combining kinematic constraints, artificial potential fields (APF) and an ACO module. The magnitude of the artificial forces is used to switch between the APF and the ACO modules and to incorporate collision avoidance. The ACO module utilises pheromone maps to establish distributed coordination of continuous surveillance. Prioritised surveillance in consideration of the intruders' objective is not considered however. In addition, a range constraint is introduced that forces swarm members to occasionally return to their base of operations to recharge [40].

2.2 Modelling adaptive and rational poachers

The review of surveillance models in the previous section shows that current state-of-the-art surveillance models mainly consider static intruders that do not respond to the searching party. Humans are intelligent rational beings however, who are capable of expressing behaviour that includes evasive and cooperative action. This section discusses models that have gained significant support to simulate poacher decision making in order address such two-sided surveillance problems [23].

Game Theory-based models describing decision making of rational entities have established accurate models to simulate bounded human rationality [43]. The Quantal Response (QR) model proposed in [62] models human rationality as a probabilistic decision making process. It assumes that better choices are more likely to occur relative to worse choices, although the best choice is not made with certainty. This QR model has been extended to incorporate bounded rationality. This extension considers that a rational entity may not have full knowledge of its surroundings and that decisions are influenced by the interplay of emotion and cognition. The proposed Subjective Utility Quantal Response (SUQR) model was found to predict human decisions more accurately compared to the QR model [63]. The Collaborative Opportunistic Security Game (COSG) model is another extension that considers the advantages of collaborative behaviour for poachers through adjustment of the reward factor in the QR function [64]. The SUQR model has also been extended to incorporate the effects of gained experience on decision making processes, which resulted in the Adaptive Subjective Utility (ASU) function [10].

To date, these proposals modelling the influence of bounded rationality and gained experience on decision making, have not been integrated. This research proposes to integrate the COSG and the ASU models to analyse the two-sided surveillance problem at hand. The emergent behaviour of this proposed poacher model will be used as input for the proposed surveillance model in order to address the identified limitations of current state-of-the-art surveillance models.

2.3 Modelling wildlife movement ecology

Both the UAVs and the poachers discussed so far are interested in wildlife, albeit that their objectives conflict. Where the ultimate objective of GSGs is to prevent wildlife loss due to illegal poaching, poachers aim to kill wildlife for other gains. It is therefore crucial to incorporate animal movement ecology in order to prevent the loss of animal species [65, 66]. Wildlife movement ecology models the spatiotemporal patterns of animals by depicting the mechanistic components of animal movement in response to changes in the environment [67].

Simplistic models that can simulate animal movement include random walk theory, where equal probabilities are assumed for movement in each possible direction. More advanced models allow for simulating correlated random walks to simulate a correlation between successive steps using a directional bias towards preferred locations [68, 69]. Topographical and environmental attributes influence these biases [66, 70]. Levy walks aim to simulate movement where local (tactical) movement is mixed with less frequent (strategic) far-ranging commutes, also allowing to return to previous locations. These methods require dense positional data for empiric evaluation to verify that animals indeed perform Levy walks [68] and calibration of such models is important to avoid inappropriate management strategies [71, 66].

Such historic data is not available for this research. It is therefore proposed to model wildlife movement ecology using calibrated resource selection functions (RSFs). These provide a solid foundation to develop an empirical

link between habitat use and population distribution. It assesses which habitat characteristics, both biotic and abiotic factors, are important to a specific animal species to quantify a probability of an animal utilising a certain resource [72].

Modelling these animal decision processes allows for quantification of intruder rewards and penalties based on the animal population distribution within the target area [63]. In addition, animal movement dynamics can be provided as input for the proposed surveillance model such that the intruders' objective can be taken into account to adjust path planning [12, 73].

3 Methodology

The methodology visualised in Figure 1 is applied in this research. The steps comprise four phases: defining the purpose of this research, defining the ABM model, collecting training data to construct BNs and testing the hypotheses through simulation.

The purpose of this research is to quantify the effects of online Bayesian learning in a two-sided surveillance problem on surveillance performance. Swarm intelligence is integrated with (A)BNs and simulated using the ABM paradigm to obtain quantitative data describing model performance. The research question has been formulated as follows:

> *"How does a learning multi-UAV approach to coordination of autonomous continuous surveillance influence effectiveness of detecting coordinating and adapting intruders in vast target areas?"*

It is hypothesised that prioritised surveillance efforts using (A)BNs as input for swarm intelligence results in improved surveillance effectiveness. Additionally, online learning of the UAV swarm is expected to prevent a long-term reduction in surveillance performance as a result from adaptive poacher behaviour.

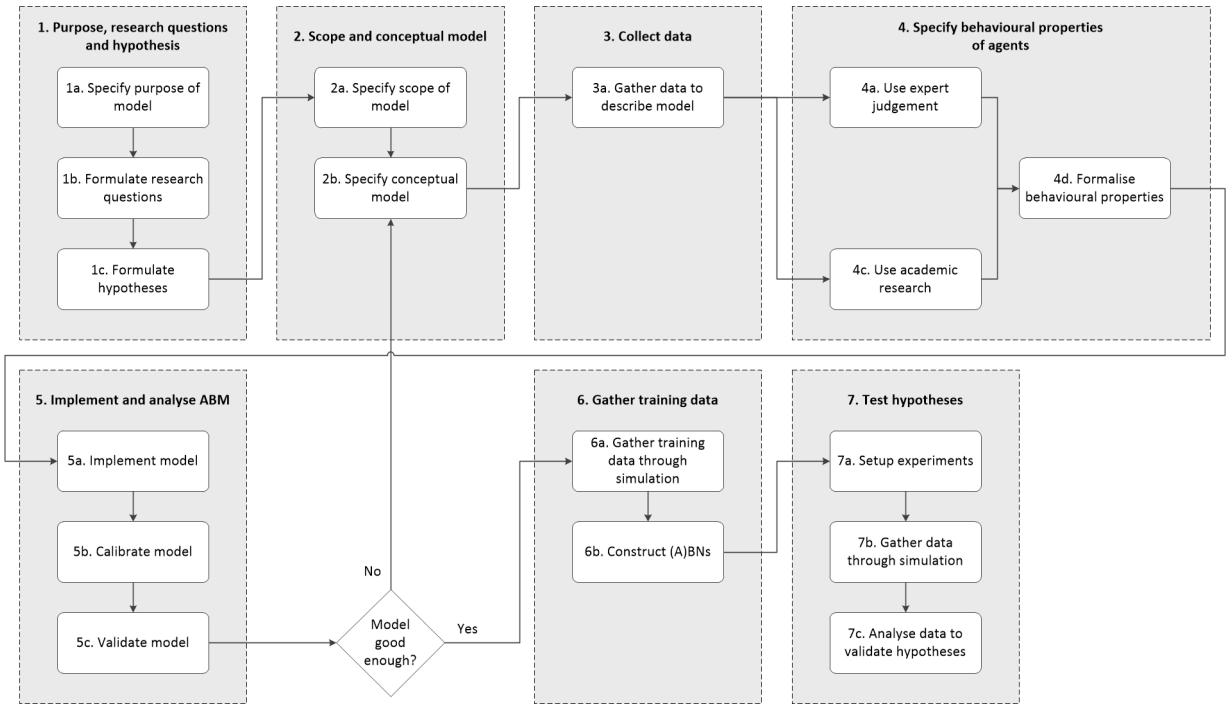


Figure 1: The methodology used in this research.

The ABM paradigm lies at the basis of the proposed model that is used to answer the research question. This paradigm has shown to be effective at overcoming limitations of equation-based and game theoretical models. It has also shown to be a promising method for simulating anti-poaching models [74, 75, 65, 49]. ABMs can incorporate detailed behavioural and ecological processes, while also being able to simulate adaptive responses of agents. This is achieved by defining the local static and dynamic properties of the environment, the agents and the interactions between these model elements from a bottom-up approach. It therefore supports modelling complex spatiotemporal interactions and analysis of the local interactions that affect overall system performance and emergence [49, 65]. An ABM model is proposed in Section 4.

The ABM paradigm provides a framework for defining the scope and the behavioural properties of the model elements from a bottom-up approach. The proposed ABM model[1] comprises UAV, poacher and rhino agents to simulate the emergent behaviour of an anti-poaching surveillance model. The presence of park rangers is excluded to prevent additional complex model interactions. Provided that a target area in Namibia is considered [76] and the availability of relevant wildlife ecology movement research, it was chosen to consider the black rhino (or Diceros bicornis) in the ABM model. Data is collected from literature and satellite images to quantify the environment of Aloegrove Safari Park and the rhino population. Expert knowledge provided by Eyeplane B.V. is taken into consideration to describe poacher behaviour in the target area. Parameters of the poacher and rhino models are then calibrated through sensitivity analyses to obtain realistic results, given the unavailability of historic data regarding poacher and rhino activity. Calibration of the UAV swarm is particularly important, since the difficulty with swarm intelligence lies in creating autonomous models that are precise enough to timely adapt to the dynamics of the environment while dealing with operational constraints [59]. This is performed by systematically calibrating the individual modules of the surveillance model and integrating these one at a time. UAV kinematics, battery re- and discharging are simplified in order to facilitate the overall calibration process and to simplify analysis of emergent behaviour of the ABM model. In addition, poacher evasive behaviour upon detection by a UAV is excluded as well. ABM model specification is concluded once the emergent behaviour of the ABM model has been validated and results in a desired level of accuracy to simulate realistic (anti-)poaching activities.

A BN relies on historic data regarding poaching behaviour to predict criminal activity. This training data is collected through simulation of the validated ABM model due to an absence of historic data from the real world. The amount of training data gathered is based on required computation time and available hardware capacity. The collected data is used as input for structure and parameter learning to define several BNs[2]. This is needed to enlighten our knowledge with an accurate understanding of the relation between the random variables in order to describe poacher behaviour with a BN. In addition, we propose to exploit expert knowledge to define the directed a-cyclic graphs (DAGs) and conditional probability tables of a number of artificial BNs (ABNs). These ABNs are used to design controlled experiments and to compare simulation results of the experiments with the performance of BNs inferred from the training data. A variety of (A)BNs is compared to determine the impact of structure and parameter learning on surveillance effectiveness.

Several experiments are designed to test the hypotheses and the influence of model parameters on observed emergent behaviour. These experiments are discussed in further detail in Section 5. Both sample size and simulation length is determined from analysis of the coefficient of variation. Performance of the proposed HAPF-ACO model is analysed through sensitivity analysis and quantified using key performance indicators (KPIs) obtained from literature. The results are assessed through correlation factors, analysis of spatiotemporal patterns and statistical tests. This includes testing for normality using the Shapiro-Wilk test and measuring significance using the Wilcoxon signed rank test and the paired t-test for normally distributed KPI data. Paired tests are used since model output is obtained with the same population of agents. A significance level of 5% is used for these statistical tests.

4 ABM Specification

This section formulates the ABM and the interactions of its elements, which are visualised in Figure 2. The local properties and capabilities of the individual elements define the emergent behaviour of the ABM [77, 78]. The proposed UAV-based surveillance model, HAPF-BLACOPS, is an improved version of the HAPF-ACO model that is integrated with an (A)BN. The UAV agents rely on swarm intelligence to prevent wildlife loss by detecting illegal poaching activity. UAV behaviour is therefore based on the exchange of locally available information among swarm members to facilitate local autonomy [59]. Poacher agents are simulated through a novel decision model, ACOSG, that combines the COSG and the ASU models. This allows for simulation of poachers that can coordinate with each other and adapt their strategy in response to surveillance efforts. A wildlife population is simulated as well, utilising RSFs to simulate black rhino movement ecology.

[1]Implemented using the Python based Mesa Framework: `https://mesa.readthedocs.io/en/master/`

[2]Implemented using the Python based pgmpy Framework: `https://pgmpy.org/`

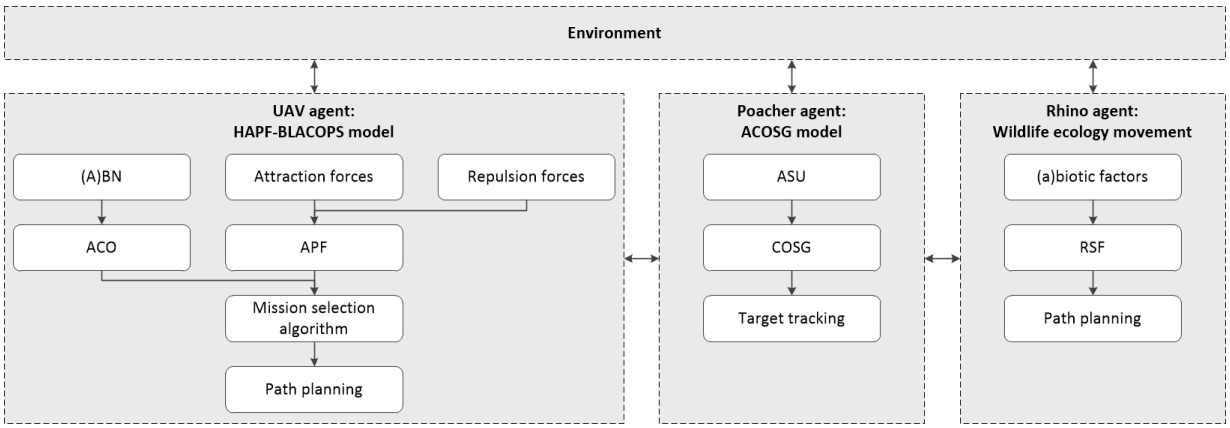


Figure 2: Overview of interactions between ABM elements and the models describing the agents' behaviour.

The following subsections are structured in accordance with the ABM paradigm and specify the models and interactions shown in Figure 2 in further detail. The ABM's environment is therefore elaborated upon first in subsection 4.1. Subsection 4.2 then provides an overview of the interactions among the agents. After which subsections 4.3, 4.4 and 4.5 specify the models describing the behavioural properties and interactions of the black rhino, poacher and UAV agents respectively. Finally, subsection 4.6 proposes a number of (A)BNs that are integrated with the HAPF-BLACOPS model.

4.1 Environment

A 10 by 10km area within the Aloegrove Safari Park in Namibia is used to describe the environment of this ABM. This particular target area is selected due to available expert knowledge from Eyeplane B.V. and previous research regarding poaching activity in the area [76]. This target area E is discretised for simulation purposes into N_x x N_y equally sized grid cells σ (from hereon referred to as patches) of size L_x x L_y. All patches represent an area of 500 by 500m, such that a unit distance Δd is defined as 500m and a balance is found between computation costs and preventing flocking of UAVs for larger swarm sizes as a result of an insufficient size of the target area. Each patch's static properties are defined from abiotic factors. Abiotic factors relevant to rhinos and poachers include the presence of water sources [45, 3], roads [79, 3], vegetation (density) [80, 45, 71] and the terrain slope [81]. SRTM V3 data [82] is used to quantify the target area's elevation and vegetation density is quantified using NDVI data provided by Landsat 8 satellite data [83]. Appendix A elaborates further on the definition of the environment and quantification of these abiotic factors. These environmental attributes influence the decisions made by the set of agents, S_{agents}, which comprises all types of agents in the ABM model. The target area is therefore described by a graph $G(E, A)$ where $E = \{(x, y)|x = 1, 2, \cdots, N_x, y = 1, 2, \cdots, N_y\}$ are the positions of the patches and $A = \{(x_k, y_k)| \forall\ k \in S_{agents}\}$ are the positions of the black rhino, poacher and UAV agents. A base of operations for the UAV swarm is located at the (x, y) position $(8, 14)$ [76].

4.2 Agent interactions

The environmental attributes provide input for various ABM model elements. This is visualised in figure 3, which indicates how agents consider both environmental attributes and observations regarding other agents in their cognitive decision processes. It summarises how environmental attributes are utilised by both rhino, poacher and UAV agents, while the presence of individual agents provides input for other agents. The following sections specify the behavioural properties of the agents in further detail by elaborating on the actions defined in figure 3.

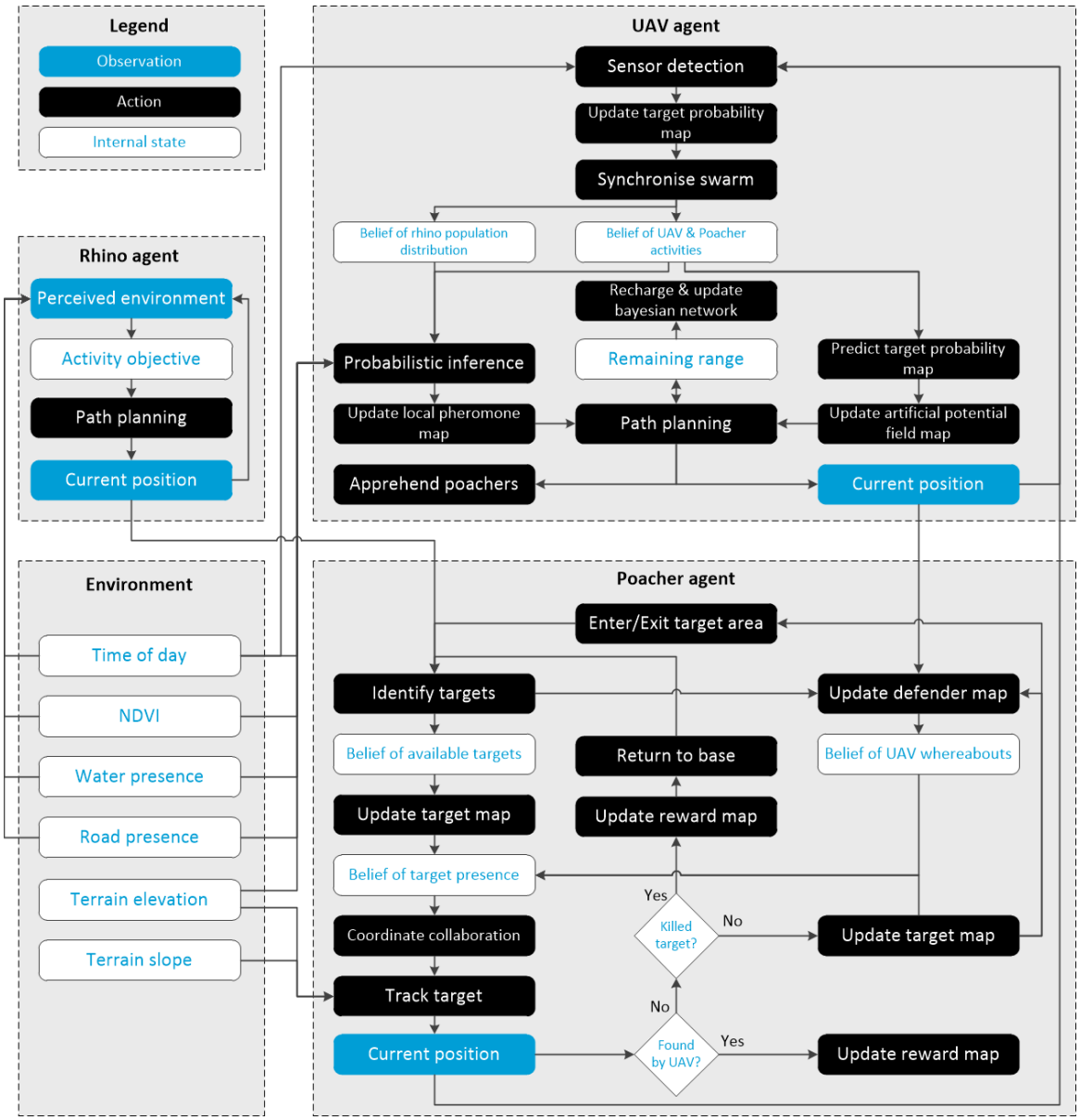


Figure 3: Visualisation of agent interactions and their internal decision processes.

4.3 Rhino agents

Black rhinos are a ubiquitously solitary species and crepuscular animal, formation of groups is unlikely and ephemeral if it does indeed occur [80]. Rhino agents therefore have internal belief properties regarding their environment that allows them to consider the presence of other rhinos in their decision making. The rhino population being simulated is defined as $S_{rhinos} = \{a_1^{rhino}, \cdots, a_n^{rhino}\}$, $1 \leq n \leq N_{rhinos}$ where N_{rhinos} is the size of the rhino population. These rhinos are also eager to be nearby water sources during midday and dusk [3, 84]. Previous research has also found a correlation between vegetation density, the distance from roads (and therefore tourism) and the presence of black rhinos [71, 85]. Their longest period of rest is during the mid-day heat [80]. Stochastic modelling of animal movement remains a state-of-the-art basis to gain a deeper understanding of an animal's choices regarding habitat utilisation [68]. Black rhino behaviour is therefore proposed simulated through a Markov Decision Process utilising RSFs to quantify transition probabilities [86]. A number of time periods is distinguished to account for the different behavioural activities of black rhinos throughout the day. The RSF therefore utilises a different set of (a)biotic factors as input to differentiate between these activities. The number of input variables, $\vec{x}$, for the RSF is limited to reduce added model complexity that can result from multiple input variables [78]. These input variables are normalised in order to simplify the calibration and validation process of the RSF model. Their weights, $\vec{w}$, are calibrated through sensitivity analysis in the absence of historical data regarding black rhino movement. Quantification of these (a)biotic factors and the weights is elaborated upon in further detail in Appendix A.

4.3.1 Path planning

Black rhino movement behaviour is simulated through a stochastic model utilising a RSF (see Equation 1) to compute transition probabilities from patch σ_i to an adjacent patch σ_j. These transition probabilities from σ_i are quantified for each adjacent patch σ_j in the Von Neumann Neighbourhood (VNN), which is defined as the set of possible options, $S_i = \{\sigma_1, \cdots, \sigma_k\}$, $k \in \text{VNN}_i$, towards which a rhino agent can move. The RSF uses a linear combination of weighted factors to simulate black rhino behaviour [87, 86]. The output of the path planning action from Figure 3 is an updated patch towards which the rhino agent moves.

$$p(\sigma_i, \sigma_j, t) = \frac{e^{\vec{w} \cdot \vec{x}_{\sigma_i, \sigma_j, t}}}{\sum_{k \in S_i} e^{\vec{w} \cdot \vec{x}_{\sigma_i, \sigma_k, t}}} \tag{1}$$

The RSF considers different input variables and their respective weights for each defined time period, shown in Table 1. The RSF considers the euclidean distance, x_{road,σ_j}, of patch σ_j to the closest road, as well as the number of black rhinos present on the patch in question, $x_{\text{rhino},\sigma_j}$, for each time period.

Table 1: Input variables for RSF describing black rhino behaviour.

Time Period	**Input 1**	**Input 2**	**Input 3**
Forage (00h00 - 04h30)	x_{road,σ_j}	$x_{\text{rhino},\sigma_j}$	$x_{\text{vegetation},\sigma_j}$
Drink (04h30 - 09h30)	x_{road,σ_j}	$x_{\text{rhino},\sigma_j}$	$x_{\text{water},\sigma_j}$
Forage (09h30 - 12h00)	x_{road,σ_j}	$x_{\text{rhino},\sigma_j}$	$x_{\text{vegetation},\sigma_j}$
Rest (12h00 - 20h30)	x_{road,σ_j}	$x_{\text{rhino},\sigma_j}$	x_{rest,σ_j}
Forage (20h30 - 24h00)	x_{road,σ_j}	$x_{\text{rhino},\sigma_j}$	$x_{\text{vegetation},\sigma_j}$

During periods of foraging, the rhino agents prioritises for vegetation, which is quantified from NDVI data [83]. Black rhinos tend to revisit patches with NDVI values of, on average, 0.23, with a standard deviation of 0.02 [80]. Equation 2 therefore quantifies $x_{\text{vegetation},\sigma_j}$ this preference with a normal distribution to stimulate visits to patches with an NDVI value that is similar to these findings.

$$x_{\text{vegetation},\sigma_j} = \frac{1}{1 + (\frac{NDVI(\sigma_j) - 0.23}{0.02})^2} \tag{2}$$

During sunrise, rhino agents are motivated to drink water by favouring patches with a relatively small euclidean distances $x_{\text{water},\sigma_j}$ towards such water sources. Lastly, resting is preferred during the heat of the day. This implies a preference for no movement, which is quantified by defining x_{rest,σ_j} according to Equation 3. It motivates black rhino agents to remain stationary during this time period.

$$x_{\text{rest},\sigma_i} = \begin{cases} 1, & \text{if } \sigma_j = \sigma_i \\ 0, & \text{else} \end{cases} \tag{3}$$

4.4 Poacher agents

The main objective of each poacher agent, $a^{poacher}$, is to find and kill the rhino agents. The poachers move through the target area towards their target, χ, by performing the actions visualised in Figure 3. These agents therefore rely on observations regarding the presence of rhino and UAV agents to determine their strategy. The proposed ACOSG model combines the COSG and ASU models. The COSG model utilises the SUQR function to simulate the interplay between emotion and cognition influencing rational decision making and coordination among poachers [64]. The ASU model utilises gained experience to adjust the preference for a given strategy [10]. The following elaborates on the behavioural properties of the poacher agents visualised in Figure 3 and thereby introduces the proposed ACOSG model.

4.4.1 Enter target area

Poacher agents enter the target area by road for its easy access. The exact entry point is chosen randomly from the roads specified in Appendix A. Initiation of these poaching activities occurs daily at 20h30, emphasising that these activities occur during night time (Eyeplane B.V., personal communication, 1 September 2020). Considering the possibility that poachers can scout the environment and the rhino population prior to this, the poacher agents are assumed to know the current locations of all rhino agents at the time of entry. This information is stored in their target map $S_{targets} = \{\chi_1, \cdots, \chi_n\}$, $\forall\ n \in S_{rhinos}$.

Considering the collaborative aspects of the ACOSG model, two poacher agents enter the target area at this point in time. The choice for two poachers was on the one hand made to increase the rate at which observation for the UAVs' BN are gathered and on the other hand to prevent an exponential increase in collaboration possibilities and therefore model complexity. Each of these sets of poachers (one set for each simulated day) is from hereon referred to as a pair of poacher agents, denoted by $S_{pair} = \{a_{\Psi_1}^{poacher}, a_{\Psi_2}^{poacher}\}$. Collaboration among poacher agents occurs only within one such a set of poacher agents.

4.4.2 Identify targets

Upon initialisation of a poacher agent at its entry point, the agent collects the current whereabouts of all unique targets that are present within the environment in $S_{targets}$. The poacher then removes targets from this set, if the target in question was killed at an earlier time by the poacher. The belief of available targets therefore shrinks as the poacher progresses through the targets in its target map.

4.4.3 Update target map

Additionally, the poacher scans for rhinos within a radius R_{poach} of its current position σ_i at time t to update $S_{targets}$ with the most current rhino whereabouts. The updated whereabouts are also passed on to the other poacher that is part of S_{pair}, if these poachers have agreed to collaborate.

4.4.4 Update defender map

The poacher simultaneously searches for UAV agents within a radius of R_{detect} from the poacher's current position s_i at time t. The location and time of detection of the UAV agent is then stored by the poacher and shared with the other poacher agent within S_{pair}. Sharing this information among S_{pair} only occurs in case of collaboration. The defender map is an empty set upon entry of the target area and is updated as new information regarding UAV whereabouts comes available.

4.4.5 Coordinate collaboration

The poacher agents $a_{\Psi_1}^{poacher}$ and $a_{\Psi_2}^{poacher}$ within a pair S_{pair} then each choose a target from their unique set $S_{targets}$ and coordinate their collaboration strategy. When a poacher successfully tracks a target χ, it kills the rhino and earns a reward R_χ. Collaboration increases this reward with a factor ϵ. Collaboration is advantageous for two reasons. First, the poachers can cover more ground and therefore collect more information of UAV and agent whereabouts, which increases their chances of avoiding UAVs and successfully tracking targets [10]. Second, poachers receive a larger reward for collaborating. The disadvantage of collaboration is that the penalty for being apprehended increases, since the penalty, P_χ, is proportional to the reward as discussed in Appendix A. Collaboration between poachers only occurs in case of mutual agreement.

Algorithm 1 describes how the poachers can reach agreement regarding collaboration and formulates how targets are chosen from a poacher's own unique target map $S_{targets}$. First, each poacher computes the transition probabilities for the two possible strategies of a given target $\chi \in S_{targets}$. These two strategies refer to the possibility to collaborate, or not. The transition probabilities are computed according to Equation 4. It is based on three input variables; the terrain slope (Eyeplane B.V., personal communication, 1 September 2020), the reward and the penalty. It combines the ASU model to consider the poachers' experience, which is the combined set $S_{rewards}$, which contains rewards and penalties obtained through killing rhinos and poachers being apprehended by UAVs. This experience adjusts the weights of the input variables through the inclination I. The weight of the experience increases over time, which is quantified in Appendix A [10].

$$p(\chi_i, \chi_j, t) = \frac{e^{\left(1-I_{\chi_j}(t)\right)\cdot\omega_{\text{slope}}x_{\chi_j}+\left(1+I_{\chi_j}(t)\right)\cdot\omega_{\text{reward}}R_{\chi_j}+\left(1+I_{\chi_j}(t)\right)\cdot\omega_{\text{penalty}}P_{\chi_j}}}{\sum_{\chi\in S_{targets}} e^{(1-I_\chi(t))\cdot\omega_{\text{slope}}x_\chi+(1+I_\chi(t))\cdot\omega_{\text{reward}}R_\chi+(1+I_\chi(t))\cdot\omega_{\text{penalty}}P_\chi}} \tag{4}$$

The inclination I is based on the average of collected payoffs (both rewards and penalties) within the time interval $[t_0, t]$. The inclination increases as the average payoff at a patch becomes relatively large in comparison to the average payoff collected over all other positions, up to the current time t. It therefore encourages poachers to prefer tacking targets in areas with a higher success rate. The relative importance of the inclination can be adjusted with the C_{ASU} constant [10]. Appendix A provides details regarding quantification of the inclination, calibration of the weights in Equation 4 and the C_{ASU} constant.

The attractiveness of a target, described in Algorithm 1, is defined as the strategy with the highest transition probability (see equation 5). This results in S_{att} for each poacher, containing the attractiveness for each target.

Each poacher then chooses its optimal target χ^{optimal} from S_{att}. Each poacher also chooses a sub optimal target $\chi^{\text{sub optimal}}$ from the strategies without collaboration.

$$Att(\chi_j, t) = \max\left\{p^{\text{not collaborate}}(\chi_i, \chi_j, t), p^{\text{collaborate}}(\chi_i, \chi_j, t)\right\} \tag{5}$$

Algorithm 1: ACOSG algorithm for choosing a poaching strategy and target.

Input : S_{pair}, $S_{targets}$
Output: New targets χ for each poacher in S_{pair}

;
for $a^{poacher} \in S_{pair}$ **do**
  for $\chi_j \in S_{targets}$ **do**
    Compute transition probability $p(\chi_i, \chi_j, t)^{\text{not collaborate}}$ according to Equation 4 ;
    Compute transition probability $p(\chi_i, \chi_j, t)^{\text{collaborate}}$ according to Equation 4 ;
    Compute attractiveness $Att(\chi_j, t)$ according to Equation 5 ;
;
for $a^{poacher} \in S_{pair}$ **do**
  $\chi^{\text{optimal}} = \arg\max_j S_{att}(\chi_j, t)$ $\chi^{\text{sub optimal}} = \arg\max_j p(\chi_i, \chi_j, t)^{\text{not collaborate}}$
;
`// Coordinate collaboration strategy`
;
if $a^{poacher}_{\Psi_1}$ *and* $a^{poacher}_{\Psi_2}$ *prefer the same type of collaboration strategy* **then**
  Each poacher chooses its optimal target χ^{optimal} as their active target χ_Ψ
else
  Each poacher chooses its sub optimal target $\chi^{\text{sub optimal}}$as their active target χ_Ψ
;
return χ_{Ψ_1}, χ_{Ψ_2}

Once each poacher within S_{pair} has determined its preferred target. If both poachers indeed prefer choosing a target with a collaborative strategy, then the two poachers collaborate until either poacher needs to choose a new target. If either poachers does not prefer a collaborative strategy, then both poachers are forced to choose their sub optimal target [64].

4.4.6 Track target

Moving towards the active target is achieved through Equation 4 as well, except now patches σ_j are considered rather than targets. Stochastic poacher movement is simulated based on these transition probabilities as a discrete-time Markov Chain (DTMC) with a discrete state-space [88]. The active target χ, the presence of UAVs and the terrain slope provide input to quantify the reward and penalty that motivate a poacher to move to an adjacent patch. Refer to Appendix A for further details regarding the quantification of these transition probabilities.

4.4.7 Update reward map

While tracking an active target χ, a number of situations can occur. First, the poacher can find its active target, after which the poacher will update the reward map. It also updates its target map and re-initiates Algorithm 1. Second, the poacher can find a non-active target that is present in its target map. In this case, the poacher updates the reward map and its target map, but Algorithm 1 is not re-initiated. Third, the poacher can encounter UAV agent, in which case the poacher is apprehended.

Once a poacher agent has found a rhino agent that is within the poacher's target map, the position σ_j, time t and reward R_χ of the kill are stored in the reward map $S_{rewards}$ [3]. The information in $S_{rewards}$ is at all times available to all poacher agents.

In case a poacher is found by a UAV agent, the poacher is apprehended and removed from the simulation. The penalty defined in Algorithm 1 is stored in $S_{rewards}$, together with the position σ_j and time t.

[3] A rhino agent that is found by a poacher agent is only killed for the pair of poachers to which the poacher in question belongs. The death of a rhino agent is therefore only virtual, such that it does not affect the perception of the rhino population for a different pair of poachers. This method is used to prevent simulating reproduction factors that affect population dynamics.

4.4.8 Return to base

Upon entry, poachers choose a base within the target area. This choice is biased towards elevated locations, as quantified in Appendix A. Once a poacher successfully kills its active target, the poacher returns to its base in accordance with the algorithm described in subsection 4.4.6. Once the poacher has reached its base, the poacher starts tracking a newly identified target obtained from Algorithm 1.

4.4.9 Leave target area

Once a poacher's target map is empty, the poacher performs path planning back to its original entry point in accordance with the algorithm described in subsection 4.4.6. Once a poacher successfully reaches this target, the poacher is removed from simulation and from its S_{pair}.

4.4.10 Poacher types

Two poacher types are distinguished and incorporated in the ABM model, based on expert knowledge from Eyeplane (Eyeplane B.V., personal communication, 1 September 2020). Type 2 poacher agents identify a base as discussed in subsection 4.4.8. Type 1 poachers do not choose such a base however. Type 1 poacher agents choose to track a single target from their initial target map, which is chosen from an initial deployment of Algorithm 1. Type 1 poachers therefore return to their original point of entry upon killing their target. The behaviour described in subsection 4.4.8 is therefore only applicable to type 2 poacher agents.

4.5 UAV agents

A swarm of UAV agents, $S_{UAVs} = \{a_1^{UAV}, \cdots, a_n^{UAV}\}$, $1 \leq n \leq N_{UAVs}$, performs continuous surveillance in the target area, E, in order to find poacher agents and prevent wildlife loss. The proposed surveillance model is based on the HAPF-ACO model [40] and is integrated with (A)BNs to achieve online learning and prioritise surveillance efforts. The HAPF-ACO model integrates an ACO module with an APF module to achieve swarm coordination, target tracking, collision avoidance and continuous surveillance. The proposed Hybrid Artificial Potential Field Bayesian Learning Ant Colony Optimisation Prioritised Surveillance (HAPF-BLACOPS) model is introduced in the following subsections.

Each individual UAV in the swarm has a local belief regarding the state of the swarm, regarding the presence of intruders and regarding the distribution of wildlife in the target area. This belief is based on a UAV's local target probability map (TPM), artificial potential field map (APFM), its local pheromone map (LPM) and its local rhino map (LRM). These maps are updated through local communication among swarm members and through observing the environment. Communication among swarm members is possible if the euclidean distance between UAVs is less than the maximum communication range, R_{sen} [40]. The available knowledge stored in these maps influences a UAV's path planning, as visualised in Figure 3.

The TPM of the k-th UAV agent, $TPM_k(t) = \{p^k_{(x,y)}(t) \mid x = 1, 2, \cdots, N_x, y = 1, 2, \cdots, N_y\}$, describes the distribution of poachers in the target area as believed at time t. It assigns a probability of poacher existence $p^k_{(x,y)}(t) \in [0, 1]$ to each patch. The existence probabilities are computed from Equation 6, by integrating over some distribution $f(x, y)$ that describes the known intruders' presence in the target area, as proposed in [40].

$$p^k_{(x,y)} = \frac{\int_{(x-1)L_y}^{xL_y} \int_{(y-1)L_x}^{yL_x} f(m,n) dmdn}{\sum_{m=1}^{N_x} \sum_{n=1}^{N_y} p^k_{(m,n)}} \tag{6}$$

The presence of poachers induces an attraction force, $\hat{F}^k_{att}$, at the k-th UAV's current position (σ_i). In addition, UAVs sense the presence of obstacles. This research only considers other UAVs as obstacles. The presence of a UAV induces a repulsion force, $\hat{F}^k_{rep}$ on the k-th UAV, which prevents the UAV in question from moving closer to nearby UAV agents. The resulting APFM, $APFM_k(t) = \{\hat{F}^k_{res}(\sigma_i)\}$, $\forall\ \sigma_i \in E$, describes the resultant force obtained after summation over all present attraction and repulsion forces at a given patch. These attraction and repulsion forces are computed from Equations 7 and 8 respectively. The parameters b and k_{att} influence the magnitude of these forces. The c parameter influences the rate of change of the magnitude of the repulsion forces as the euclidean distance, $\|\hat{x}_{\sigma_i,\sigma_j}\|$, increases [40]. In Equation 8, a^{UAV}_{nearby} is the UAV agent which is closest to the patch σ_i in consideration. Appendix A elaborates on the calibration process of these parameters.

$$\hat{F}^k_{att}(\sigma_i) = \sum_{\sigma_j \in E} k_{att} \cdot \frac{\hat{x}_{\sigma_i,\sigma_j}}{\|\hat{x}_{\sigma_i,\sigma_j}\|^2} \cdot (p^k_{\sigma_j} - p^k_{\sigma_i}) \tag{7}$$

$$\hat{F}^k_{rep}(\sigma_i) = \sum_{a^{UAV}_j \in N^c_k} \frac{b}{c} \cdot \frac{e^{\frac{1}{c}\cdot\|\hat{x}_{\sigma_i,a^{UAV}_j}\|} \cdot \hat{x}_{\sigma_i,a^{UAV}_j}}{(e^{\frac{1}{c}\cdot\|\hat{x}_{\sigma_i,a^{UAV}_j}\|} - e^{\frac{1}{c}\cdot\|\hat{x}_{\sigma_i,a^{UAV}_{nearby}}\|})^2} \quad \text{, if} \quad d_{min} < \|\hat{x}_{\sigma_i,a^{UAV}_j}\| \leq d_{max} \tag{8}$$

The LPM contains the pheromone concentrations, τ, for each patch in E as believed by the k-th UAV, such that $LPM_k(t) = \{\tau^k_{(x,y)}(t) \mid x = 1, 2, \cdots, N_x, y = 1, 2, \cdots, N_y\}$. These pheromone concentrations are bound to a minimum, τ_{min}, and maximum value, τ_{max}. These boundaries and the pheromone increments used to update the LPM are to be determined carefully in order to achieve effective coverage of the target area [35]. Appendix A therefore addresses the calibration of these parameters.

Finally, the LRM contains the known whereabouts of rhino agents, as perceived by the k-th UAV. This map is continuously being updated from sensor detections and synchronisation among swarm members. This LRM is used to provide input for the (A)BN, as discussed in subsection 4.6.

Upon initialisation of the ABM, each UAV is given a random heading ϕ and position (x, y). Each UAV agent initialises its TPM, APFM, LPM and its LRM without initial knowledge regarding poacher presence. A uniform target existence distribution is therefore applied to quantify the TPM [40]. Similarly, the LPM is initialised such that all pheromone concentrations are equal to the midpoint between the boundaries on the concentrations. The magnitude of the resultant forces in the APFM determine whether the UAV relies on its APFM for its path planning. These behavioural characteristics are specified in further detail in the following sections.

4.5.1 Sensor detection

The UAV is assumed to be equipped with a sensor that can detect poachers within each unique patch within its vicinity, R_{det}. The influence of environmental factors on sensor performance [89, 90, 91, 92] is accounted for by considering the probability of a true detection, P_D and a failure to successfully detect the presence of poachers, P_F [40]. In addition, these sensor parameters are adjusted between day- and nighttime to address the varying lighting conditions under which poachers operate and its affect on detection performance [93].

The LRM is also updated by detecting the presence of rhino agents. A distinction with poacher detection is made by neglecting the above sensor parameters for detecting rhinos.

4.5.2 Update target probability map

These poacher sightings are stored and shared among the UAV swarm. The most current poacher whereabouts are used to update and quantify the poacher existence probabilities in the TPM. This is described by the Wiener random process described in [40], given that the UAV does not know the speed of movement of the poacher.

4.5.3 Synchronise swarm

Communication is possible among UAV agents within coordination distance, r^{coor}_k, of each other. This radius is dependent on the number of UAV agents within the communication range R_{sen}, such that the number of UAVs the k-th UAV communicates with, is limited [40]. The k-th UAV can therefore only communicate with a subset, $N^c_k \subset S_{UAVs}$, of the swarm. These UAVs exchange poacher sightings, rhino sightings and their travelled paths with each other.

4.5.4 Update local pheromone map

The obtained information from synchronisation is used to update the pheromone concentrations in the LPM at each time step according to Equation 9. This mechanism enables autonomous coordination of continuous and prioritised swarm surveillance [50, 40]. The pheromone concentration at a patch $\tau^k_{(x,y)}(t)$ is increased by a factor $\Delta\tau_{g_0}$ at each time step to motivate UAVs to re-explore the target area, whereas visits from a UAV agent to a patch reduce these pheromone concentrations.

$$\tau^k_{(x,y)}(t + \Delta t) = \tau^k_{(x,y)}(t) + p^k_{BN}(x, y, t) \cdot \Delta\tau_{g_0} - \sum_{a^{UAV} \in N^c_k} \Delta\tau^{(k,a^{UAV})}_{l(x,y)}(t) \tag{9}$$

It is proposed to utilise the posterior probability distribution from an (A)BN to adjust the increments of these pheromone concentrations. The rate at which the pheromone concentrations increase is therefore proportional to the posterior probability, $p^k_{BN}(x, y, t)$, describing poacher activity. The (A)BNs used to perform probabilistic inference are discussed in subsection 4.6.

The k-th UAV accounts for the presence of other UAVs based on exchanged UAV positions $(x_{a^{UAV}}(t), y_{a^{UAV}}(t))$. The magnitude by which the pheromone concentrations are reduced in response to visits from these UAV agents is dependent on the local update constant, $\Delta\tau_{l_0}$ and the distance $\|\hat{x}\|$ between patches, as seen in Equation 10 [40]. Appendix A elaborates on the calibration method of these model parameters.

$$\Delta\tau_{l(x,y)}^{(k,a^{UAV})}(t) = \begin{cases} \Delta\tau_{l_0} \cdot \frac{1}{R_{det}^4} \cdot \left(R_{det}^4 - \|\hat{x}_{(x,y),(x_{a^{UAV}}(t),y_{a^{UAV}}(t))}\|^4\right), & \text{if} \quad \|\hat{x}_{(x,y),(x_{a^{UAV}}(t),y_{a^{UAV}}(t))}\| \leq R_{det} \\ 0, & \text{else} \end{cases} \tag{10}$$

4.5.5 Predict target probability map

The UAVs adjust the TPM prior to computing the APF to incorporate future movement of the poachers. Equation 11 therefore combines the target existence probabilities currently in the TPM with the target detection results to adjust for movement of the dynamic targets, where $\tau \in [0, 1]$ is used to characterise the dynamic environment. If the sensor detected a poacher at a given patch and the local knowledge regarding the poacher whereabouts obtained through synchronisation is in agreement with said observation, then $b(t)$ is defined to be equal to 1 [40].

$$p_{mn}^i(t + \Delta t) = \begin{cases} \tau \cdot p_{mn}^i(t) & \text{(not detected)} \\ \frac{P_D \cdot p_{mn}^i(t)}{P_F + (P_D - P_F) \cdot p_{mn}^i(t)} & \text{(detected and } b(t) = 1\text{)} \\ \frac{(1-P_D) \cdot p_{mn}^i(t)}{1 - P_F + (P_F - P_D) \cdot p_{mn}^i(t)} & \text{(detected and } b(t) = 0\text{)} \end{cases} \tag{11}$$

4.5.6 Update artificial potential field map

The k-th UAV then quantifies the resultant force for each patch in the APFM from the TPM through summation over the artificial forces obtained with Equations 7 and 8.

4.5.7 Recharge and update BN

Once a UAV reaches its base of operations for recharging, it exchanges newly gained observations regarding poacher activity with the database at the base. The exchanged data is used to update the conditional probability tables of the UAV's BN in order to achieve online learning. Updating the BN is performed during recharging since it is computationally expensive [25]. A simple linear recharge rate is assumed for the recharging purpose, as to limit the complex interactions with the emergent behaviour of the UAV swarm.

4.5.8 Path planning

Each UAV decides how path planning is performed based on Algorithm 2. It requires a UAV's unique APFM, LPM, LRM, current position σ_i and the UAV's current heading ϕ_i. Additionally, each UAV monitors its travelled distance since leaving the base of operations, L_{past}, to determine when the UAV needs to return for recharging. The UAV chooses one of five path planning mission types (MTs) in accordance with the proposed algorithm. By choosing for one of these mission types, the UAV either returns to base or it relies on its APF or its ACO module for path planning. These five mission types are discussed below. The output of Algorithm 2 is a new patch, σ_j, towards which the UAV moves. This patch is chosen from a set of adjacent patches, Ω_i, that can be reached without violating the turn angle constraint induced from the maximum turn angle ϕ_{max}. The proposed algorithm differentiates from the HAPF-ACO model by prioritising the UAV's remaining range and considering the maximum turn angle in the range constraint.

Algorithm 2: HAPF-BLACOPS mission selection algorithm

Input : APFM, LRM, LPM, $L_{past}, \sigma_i, \phi_i$
Output: σ_j

```
;
q ~ U(0,1) ;
;
if L_past(t + Δt) > (1/2) · L_max − L_Δφ then
    // Path planning according to mission type 1
else if q ≤ 1 − λ_APF · F_res(σ_i) / F_res,max(Ω_i) then
    // Path planning according to mission type 2
else
    q_1 ~ U(0,1) ;
    if q_1 < q_2 then
        // Path planning according to mission type 3
    else
        // Path planning according to mission type 4
;
return σ_j
```

Mission type 1: return to base of operations
The UAV moves towards and adjacent patch within Ω_i that minimises the difference, $\Delta\phi$, between the UAV's future heading ϕ_j and the heading towards the base of operations. This method of path planning is performed in case the travelled distance at $t + \Delta t$ becomes larger than half the UAV's maximum range L_{max}. The UAV also adjusts its choice of mission type by considering the distance required to align its current heading with the heading towards the base of operations. The difference between these two headings is defined as $L_{\Delta\phi}$.

Mission type 2: path planning using the APF module
The UAV draws a random variable q from a uniform distribution and compares this with the relative magnitude of the resultant forces in the APFM to determine whether the magnitude of the resultant forces within Ω_i is such that obstacle avoidance or target tracking is to be prioritised. The λ_{APF} parameter influences the importance of the APF module and is calibrated in Appendix A. A new patch is chosen by minimising for the difference between the resultant force's direction at the current patch and the heading ϕ_j obtained from moving towards adjacent patch $\sigma_j \in \Omega_i$.

Mission type 3: path planning using the ACO module
The ACO module utilises the pheromone concentrations and heuristic information $\eta_{\sigma j}$ to determine a path that maximises the weighted product of these factors according to Equation 12. The heuristic information describes the frequency of visits to a patch σ_j as perceived by a UAV's communication with nearby swarm members. It is quantified using Equation 13, where $N_{\Delta t}(t - 24\text{hours}, t)$ is the number of time steps in a 24 hour time period and where $\nu_{\sigma_j}(t - 24\text{hours}, t)$ is the number of visits to a patch during this time period. The exponents α and β characterise the optimised path and are included in a Global Sensitivity Analysis (GSA), discussed in subsection 5.4.

$$\sigma_j = \underset{k \in \Omega_i}{\arg\max} \left\{ \tau_{\sigma_k}(t)^{\alpha} \cdot \eta_{\sigma_k}(t)^{\beta} \right\} \tag{12}$$

$$\eta_{\sigma_j}(t) = 1 - \frac{\nu_{\sigma_j}(t - 24\text{hours}, t)}{N_{\Delta t}(t - 24\text{hours}, t)} \tag{13}$$

Mission type 4: stochastic path planning using the ACO module
The parameter q_2 is set to 0.5 to stimulate unpredictable UAV path planning through stochastic movement. This is needed to limit the ability of poachers to accurately learn, predict and evade the surveillance efforts of UAVs through the ASU model [94]. Stochastic path planning is therefore simulated through a DTMC, for which the transition probabilities are computed from Equation 14.

$$p(\sigma_i, \sigma_j, t) = \begin{cases} \dfrac{\tau_{\sigma_j}(t)^{\alpha} \cdot \eta_{\sigma_j}(t)^{\beta}}{\sum\limits_{\sigma_k \in \Omega_i} \tau_{\sigma_k}(t)^{\alpha} \cdot \eta_{\sigma_k}(t)^{\beta}}, & \sigma_j \in \Omega_i \\ 0, & \sigma_j \notin \Omega_i \end{cases} \tag{14}$$

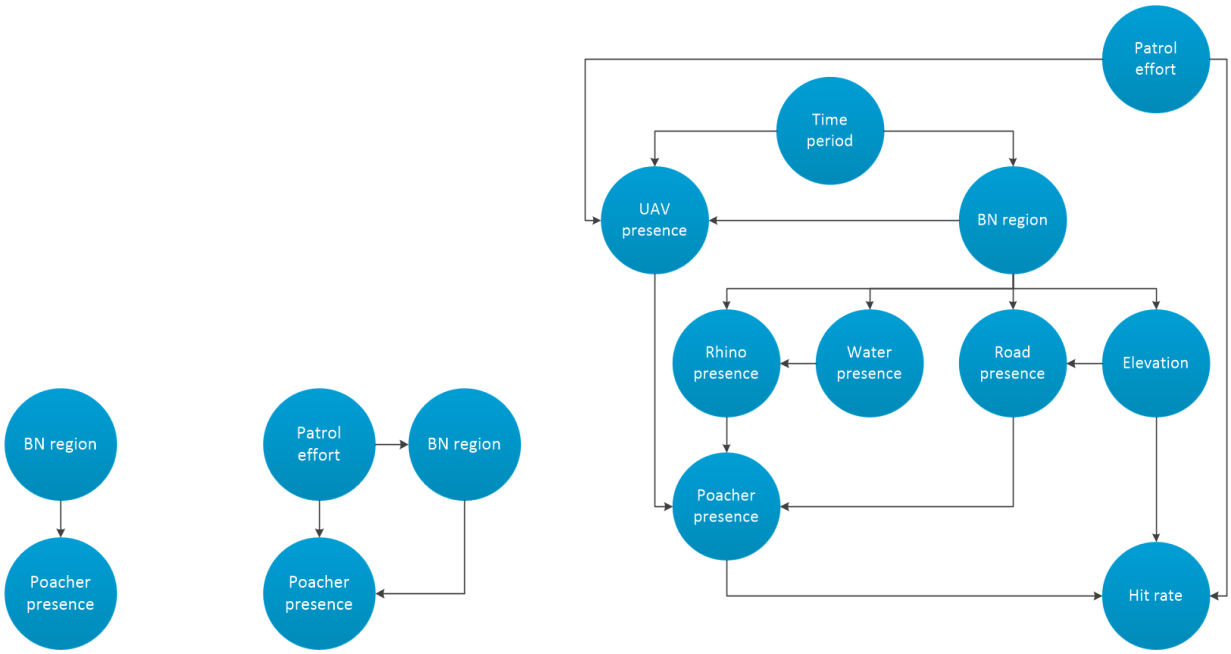
4.5.9 Apprehend poachers

When a UAV agent moves to a patch at which it successfully detects a poacher within a distance of $R_{apprehend}$, the UAV apprehends the poacher. The poacher is then removed from simulation, as discussed in subsection 4.4.7, after which the UAV continuous its surveillance task.

4.6 Integration of Bayesian Networks

The proposed HAPF-BLACOPS utilises (A)BNs to achieve prioritised surveillance by integrating the ACO with an (A)BN, as defined by Equation 9. (A)BNs are characterised by a graphical representation describing the joint probability distributions (JPDs) that can be used for probabilistic inference and risk prediction. These can be characterised by reoccurring behavioural features and are therefore predictable by considering both geographical and temporal features to forecast criminal activity [95, 96]. The underlying conditional probability tables can be updated when new evidence becomes available [25]. A distinction is made between artificially defined BNs (ABNs) and BNs obtained from training data. The DAG and conditional probability tables of an ABN are defined from expert judgement. The (A)BNs are used to perform probabilistic inference of poacher activity and to update the pheromone concentrations of the ACO module accordingly.

A DAG consists of a set of nodes (random variables) and edges, indicating the conditional (in-)dependencies between the discrete random variables. The random variables used to define the DAGs of the (A)BNs considered in this research are selected from expert knowledge (Eyeplane B.V., personal communication, 1 September 2020) and [97]. The various combinations of random variables can result in a large set op (sub-)optimal DAGs to describe the training data. The optimal DAG to describe a training dataset can be obtained from score-based structure learning algorithms such as PC and Hill Climb Search [98, 99] or expert knowledge [25]. The accuracy of a DAG describes how well it fits a given dataset and is quantified with the BIC, BDeu and K2 scoring functions [100, 101]. Appendix B therefore defines the DAGs that are obtained through structure learning and the conditional probability tables for a number of (A)BNs being analysed in this research. This selection of (A)BNs results from a qualitative analysis and from comparison of the dependencies in DAGs to expert judgement. Figure 4 visualises three DAGs that are analysed. The first DAG being defined from expert judgement as ABN4, the second (BN9) being learned from the training data and the third (BN expert) being a combination of structure learning and expert judgement.



(a) A DAG defined from expert judgement (ABN4).

(b) A DAG defined from structure learning (BN9).

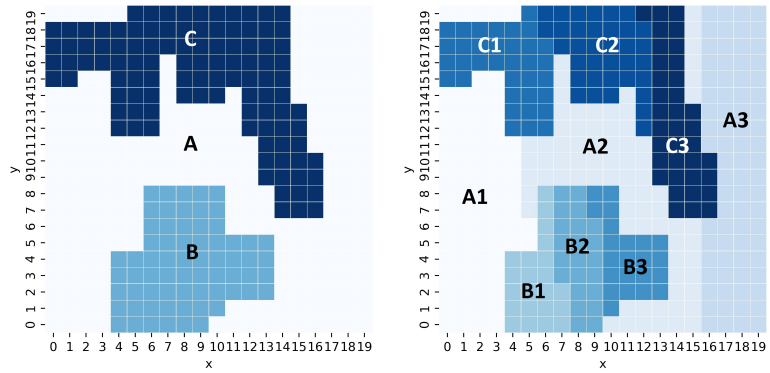
(c) A DAG defined from combining structure learning results and expert judgement (BN expert).

Figure 4: Visualisation of three DAGs being analysed.

The random variables providing the structure of the DAGs in Figure 4 are discretised from the (a)biotic factors characterising the target area (discussed in Appendix A). In addition, the target area is discretised into

a number of BN regions among which surveillance efforts are to be prioritised. These BN regions, shown in Figure 4, are defined to reduce computation costs and such that each region is characterised by a unique set of features. Since the discretisation methods to define the random variables and the BN regions can influence this goodness of fit [102], the influence of these parameters on surveillance performance is analysed in Section 5. Both supervised and unsupervised data discretisation techniques are considered to define uniform bin widths for this purpose.

The conditional probability tables are then estimated through a parameter learning algorithms such as Maximum Likelihood Estimation (MLE), Bayesian Estimation or Expectation Maximisation (EM) [100]. Finally, the UAV agents perform probabilistic inference to predict adaptive poacher activity in a given region within the target area and perform path planning.



(a) Discretisation into 3 BN regions. (b) Discretisation into 9 BN regions.

Figure 5: Discretisation of the target area into predefined BN regions.

5 Simulation Results

This section elaborates on steps 5 through to 7 from the methodology presented in Figure 1. First, a description and a flowchart is provided to elaborate on the approach used to calibrate model parameters, validate model output and to analyse the emergent behaviour of the HAPF-BLACOPS model and its performance in comparison to the current state-of-the-art. Second, the results of the experiments are presented and elaborated upon. The required number of simulation runs and simulation time for these experiments was quantified with the coefficient of variation of model output, for which the results are presented in Appendix A.

5.1 Simulation Approach

Prior to performing the experiments from Table 2, calibration of the parameters of the discussed rhino, poacher and UAV models was performed through sensitivity analyses. The model output was validated against expert knowledge, research describing rhino behaviour and against historic poaching data. A systematic approach was used to calibrate the model parameters and the individual HAPF-BLACOPS modules to obtain sufficiently realistic model output. This was achieved by simulating the ABM model with the individual models that define the rhino, poacher and UAV agents' behaviour respectively. Calibration of the weights of the RSF in Equation 1 simulating black rhino behaviour was performed first. After which the ACOSG's weights from Equation 4 were calibrated. Then the parameters of the APF and the ACO were calibrated individually, before building the hybrid APF and ACO model. Compositional analysis of the individual models and their modules, prior to combining these, prevents interactions between these features (shown in Figure 3) from influencing model output. By gradually combining each of these features, their joint influence on model output was analysed [103]. Appendix A elaborates upon these steps in greater detail and discusses the validation of model output. Appendix C provides an overview of the model parameters that were used in the experiments, unless stated otherwise.

The validated ABM is used in experiments E1 through to E8[4] to perform simulations that test the hypotheses of this research and answer the research question. Table 2 provides a summarising description of each of these experiments and their objective. The results of the performed simulations are quantified according to a

[4]Experiments E5 and E6 are presented and discussed in Appendix F.

number of KPIs, which are elaborated upon in Appendix D. ABNs are used in several experiments[5] to provide a controlled simulation environment, in which the quality of the training data does not influence the emergent behaviour of the ABM.

Table 2: Description of performed experiments.

Experiment	**Description**	**Hypothesis tested with KPI**
E1	The effect of the proposed mission selection algorithm on model output regarding collision avoidance and violation of the range constraint is analysed by simulating the HAPF-ACO and the HAPF-BLACOPS models without a (A)BN.	Range violations Collisions
E2	The effect of prioritising surveillance efforts through ABNs on surveillance effectiveness and efficiency is analysed by simulating the HAPF-BLACOPS and HAPF-ACO models.	Apprehensions Catch per unit effort Coverage fairness
E3	A global sensitivity analysis of the HAPF-BLACOPS input parameters (N_{UAVs}, α, β, R_{sen}, $\Delta\tau_{l0}$, Acknowledgement of the APF) is performed to quantify their influence on the emergent behaviour of the UAV swarm and determine a set of parameters that improves surveillance performance and the influence of the (A)BN on model output.	Spearman correlation coefficients
E4	The effects of prioritising surveillance efforts through various (A)BNs on surveillance effectiveness and efficiency are analysed by simulating the HAPF-BLACOPS with the adjusted input parameters obtained in E3 and comparing to model performance obtained in E2.	BN scoring functions Apprehensions Catch per unit effort Coverage fairness
E5	The influence of the dimensions of the BN regions, specified in Figure 5, on surveillance effectiveness is analysed by simulating the HAPF-BLACOPS with the adjusted input parameters obtained in E3.	Loss of wildlife GAI Spatiotemporal patterns
E6	The potential bias resulting from considering solely apprehended poacher data to predict poacher activity is analysed by extending the training data with observations of unapprehended poachers and simulating the HAPF-BLACOPS model with the adjusted input parameters obtained in E3.	Apprehensions Catch per unit effort Coverage fairness Spatiotemporal patterns
E7	The influence of the ASU function on the emergent adaptive behaviour of poacher agents is analysed through adjustment of the ACOSG input parameter C_{ASU} in an environment where the adjusted HAPF-BLACOPS input parameters obtained in E3 are used.	Coefficient of variation of spatiotemporal patterns Apprehensions Loss of wildlife
E8	Compositional analysis of the HAPF-BLACOPS modules is performed to determine how internal interactions influence the effects of updating the ACO module with (A)BN predictions on surveillance performance. The proposed surveillance model is simulated with the adjusted input parameters obtained in E3 and with ABM parameters that are updated according to the results obtained in E4, E5, E6 and E7.	Coverage Coverage fairness Apprehensions Loss of wildlife Catch per unit effort Collisions

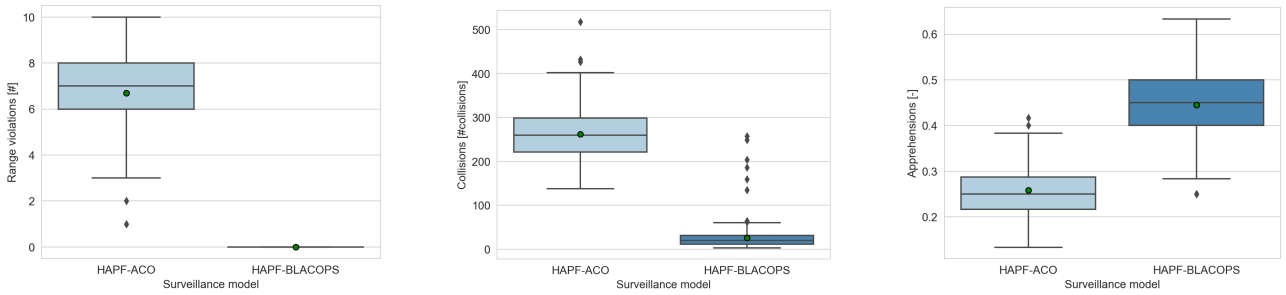
The experiments discussed in Table 2 are also shown in Figure 6. This elaborates on step 7 of the methodology in Figure 1 by providing an overview of the interconnection between the successive simulation experiments. The hypotheses of this research are listed as well.

[5]Simulations in E3, E5 and E6 are performed solely using ABN4, shown in Figure 4(a).

Figure 6: Description of causal relations between the experiments and their respective hypotheses.

5.2 E1: Analysis of proposed mission selector

Experiment 1 tests the hypothesis that the mission selection algorithm (see Algorithm 2) from the proposed HAPF-BLACOPS model addresses the identified shortcomings of the HAPF-ACO model's mission selector. The proposed mission selector was designed with the objective of preventing UAVs from violating their range constraint, since it was found that the range constraint in the HAPF-ACO model did not consider the maximum turning angle in determining the remaining range. In addition, the HAPF-ACO did not prioritise a violated range constraint over other mission types. Simulations for this experiment were performed with t_{max} set at 10 days. Both the HAPF-ACO and the HAPF-BLACOPS models were simulated with the same model parameters and without a (A)BN. The HAPF-ACO model shows that several UAVs returned to their base of operation with a negative remaining range, as seen in Figure 7(a). The proposed mission selector in the HAPF-BLACOPS model prioritises the range constraint over other mission types and considers the maximum turn angle combined with the UAV's current heading to determine when the operational range will be violated, as discussed in subection 4.5.8. This prevents any UAV from violating its range constraint, as seen in Figure 7(a).



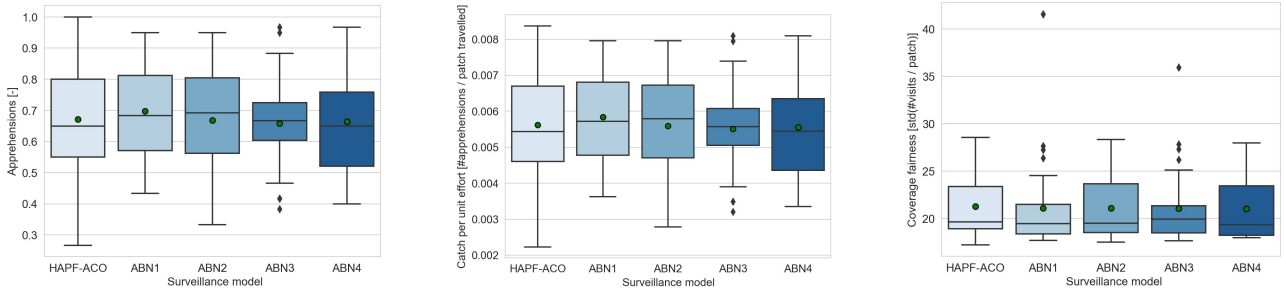
(a) Range violations by UAVs. (b) Collisions among UAVs. (c) Poachers apprehended by UAVs.

Figure 7: Performance comparison of proposed mission selector in HAPF-BLACOPS with HAPF-ACO.

In addition, due to simplifications in the mission selector achieved from the prioritisation of the range constraint, the UAV is more consistent in choosing a mission type. This results in more frequent use of the APF, resulting in a reduction of, on average, 84% mid-air collisions between UAVs, as seen in Figure 7(b). Another secondary effect visualised in Figure 7(c) shows that the use of the ACO and APF becomes such that the surveillance effectiveness, in terms of number of apprehensions, is, on average, doubled, such that 50% of poachers are being apprehended. The variation in surveillance effectiveness remains unaffected however. Statistical analysis of these results, presented in Table E.1 in Appendix E, show that the improved performance of the HAPF-BLACOPS model is indeed significant, since the p-values are less then 0.05. The hypothesis for this experiment, stated in Figure 6, is therefore accepted.

5.3 E2: Baseline performance of HAPF-BLACOPS with ABNs

Experiment 2 analyses how the use of various ABNs influences both surveillance effectiveness and efficiency of the proposed HAPF-BLACOPS model in comparison to the current state-of-the-art. Both the DAG and the conditional probability tables of these ABNs were designed based on the spatiotemporal patterns found during calibration. Appendix B provides the specification of these DAGs and their conditional probability tables. Figure 8 quantifies how these ABNs influence surveillance performance.



(a) Poachers apprehended by UAVs. (b) Catch per unit effort. (c) Coverage fairness.

Figure 8: Performance comparison of HAPF-BLACOPS utilising ABNs to predict poaching activity.

Due to the increased simulation duration in comparison to E1 in subsection 5.2, the HAPF-ACO model is able perform significantly better on average at apprehending poachers. However, the HAPF-ACO model does show significant variation in model performance, due to the randomness in the mission selector. The performance of the ABNs in terms of apprehending poachers, the catch per unit effort and the coverage fairness shows, based on the statistical analysis in Table E.2, no significant improvement or degradation in comparison to the HAPF-ACO model. This is caused by the underlying interactions in the mission selector defined in Algorithm 2. These, combined with the presence of the swarm, limit the amount of path planning being performed according to the BN and the ACO. This is also supported by the spatiotemporal patterns of the UAVs, visualised in Appendix F. These indicate no different preference for certain BN regions in comparison to their preferences in the HAPF-ACO model. The magnitude of the variance in Figure 8 also indicates that HAPF-BLACOPS performance is unstable, which is similar to the HAPF-ACO model. The hypothesis in Figure 6, stating that the use of ABNs improves surveillance effectiveness and efficiency, is rejected based on this experiment. A GSA was therefore performed to determine the influence of model parameters on model output and improve performance.

5.4 E3: Global sensitivity analysis of HAPF-BLACOPS

A global sensitivity analysis of the model parameters in Table 3 was performed with simulations of the HAPF-BLACOPS model utilising ABN4 (specified in Appendix B) in order to determine their effects on model output and test the hypothesis stated in Figure 6. Parameter screening was performed to determine this set of parameters for which their influence on the emergent behaviour is both complex to predict and beneficial to understand. Latin Hypercube Sampling (LHS), utilising discrete uniform distributions to sample parameter combinations, was used to reduce the required number of simulations. The Spearman rank correlation coefficients were computed, since the model input-output relationship was not expected to be linear [75]. A 3^k factorial design was implemented to be able to capture both the strength and direction of nonlinear effects on model output [104, 105], except for the *Acknowledgement of the APF*. This parameter enforces whether a UAV agent omits path planning according to mission type 2 (defined in Algorithm 2) by setting the magnitude of q to 0 when the magnitude of *Acknowledgement of the APF* is set to 0.

Table 3: HAPF-BLACOPS parameters analysed during GSA.

Model parameter	N_{UAVs}	α	β	R_{sen}	$\Delta\tau_{l0}$	Acknowledgement of the APF
Considered values	5, 10, 20	1, 3, 5	1, 3, 5	1, 3, 5	0.1, 10, 100	0, 1

The obtained Spearman correlation coefficients are presented in the lower-triangle of Figure 9. Their respective p-values are presented in the upper-triangle of the this figure. Most of these p-values meet the significance level of 5%, except for, mainly, the p-values related to correlations between the β parameter and model output. This could be related to the quantification of the heuristic information, which is dependent on path planning of the UAV swarm, the mission selection algorithm and therefore the complex and partially stochastic interactions within the ABM. The hypothesis for E3 is only partially accepted, since no significant effect on model output is observed for the parameters α and β.

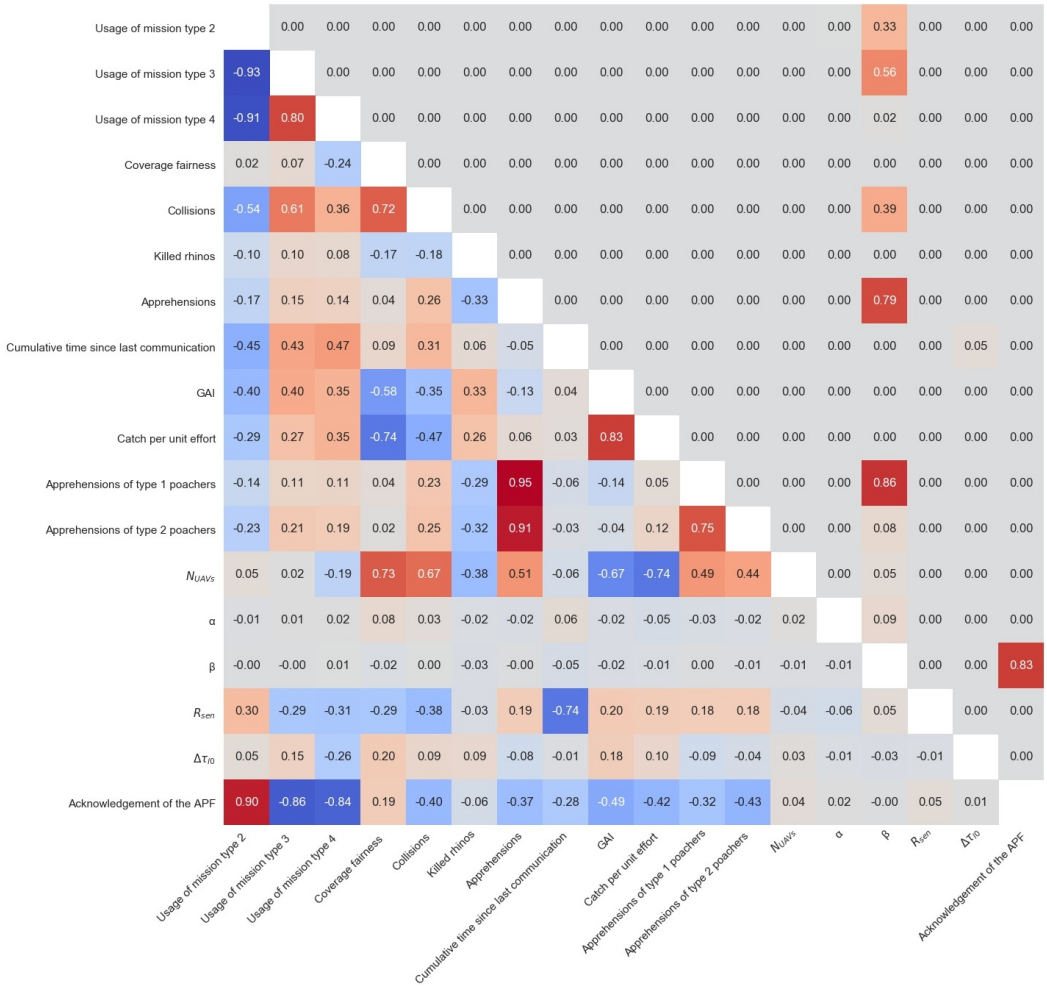


Figure 9: GSA Correlation matrix showing the Spearman correlation coefficients.

The size of the UAV swarm, N_{UAVs}, has the strongest effect on the number of apprehensions being made by the UAVs and results in a significant reduction of loss of wildlife. Albeit that this correlation with wildlife loss is weaker than the correlation with swarm size. This indicates that scaling of the HAPF-BLACOPS model may be the most effective way to reduce the loss of wildlife, but there is no linear relation between these two aspects. Additionally, it shows that surveillance efficiency reduces as the swarm size is increased. This was expected since the travelled distance grows exponentially, whereas the amount of poachers that needed to be apprehended did not change. In addition, Figure 8 indicated an apprehension rate of well above 70%. The spawn rate of the poacher agents therefore limits these results to some extent, as supported by Figure F.7 in Appendix F. This figure indicates an apprehension rate exceeding 90% for swarm sizes of 10 UAVs, which approaches 100% as the swarm size is increased further.

The increasing swarm size also shows strong correlation with prioritisation of surveillance efforts, as quantified by the coverage fairness. The amount of collisions also increases significantly however. This combination of correlation factors indicates that, although most patches in the target area are visited more frequently as quantified by the reduction in GAI, the UAVs are performing surveillance relatively close to each other. The UAVs therefore collaborate on patrolling certain regions such that these regions are prioritised. The downside is that the UAVs can no longer prevent mid-air collisions through the APF due to the relatively small distance between the swarm members and the solutions provided by the APF. The APF is characterised by defining a short-term optimal solution, therefore not necessarily preventing future collisions.

As expected, increasing the R_{sen} parameter, the maximum communication distance, results in UAVs communicating more frequently, which is also quantified in Figure F.9 in Appendix F. The main result of this is an increased reliance on mission type 2 however. This results from a better situational awareness regarding the swarm's spatial distribution, which means that the individual UAVs utilise the APF more frequently to prevent collisions with nearby UAVs. The secondary effect is therefore the observed reduction in collisions seen in Figure 9. The correlation factors also indicate that the improved synchronisation frequency among swarm members results in improved efficiency in terms of the catch per unit effort. It appears that surveillance efficiency can either be improved by adjusting the maximum communication range, which strengthens the effects of UAV presence in the LPM update mechanism, or by directly adjusting the $\Delta\tau_{l0}$ parameter. An increase of this ACO parameter results in more surveillance prioritisation, a slight reduction in apprehensions and also a slight increase in loss of wildlife (see Appendix F). Surprisingly, it also influences the use of mission type 3 and 4, for which no evident cause has been identified. The data presented in Figure F.7 in Appendix F also shows that the improvement regarding surveillance effectiveness as a results from adjusting R_{sen} reaches a stagnation point, which is reached when communications are no longer limiting coordination.

Since the results presented so far indicate a strong influence of the APF module on the emergent behaviour, analyses were performed by disregarding the APF module. A positive correlation between *Acknowledgement of the APF* and model output means that the use of the APF module, and therefore the use of mission type 2, also results in an increase in model output. The correlation matrix in Figure 9 shows that use of the APF has a smaller impact on wildlife loss in comparison to use of the ACO module (mission type 3 and 4), although these correlations are relatively weak. The correlation with the amount of collisions is fairly strong and the data indicates that the avoidance forces dominate the attraction forces, such that the APF module has a relatively small affect on the apprehension rate. Disregarding the APF therefore significantly influences how the (A)BN affects the swarm's spatiotemporal patterns, as supported by Figure F.10 in Appendix F.

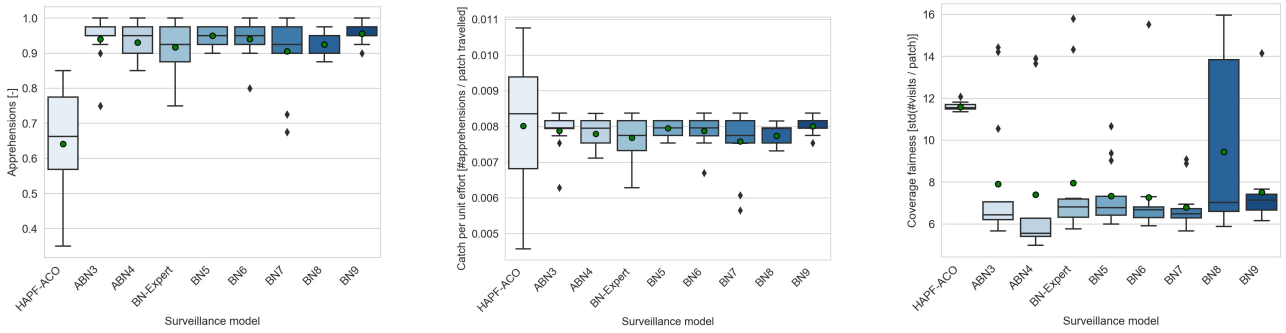
Finally, the α and β parameters do not influence surveillance effectiveness and/or surveillance efficiency. Albeit unexpected, this behaviour results from the different order of magnitude of the pheromone concentrations and the heuristic information. Where the heuristic information is bound to a maximum magnitude of 1, the pheromone concentrations can reach a magnitude of 4000 (refer to Appendix A). The influence of the pheromones is therefore dominant in the choice for the ACO module. These unequal orders of magnitude therefore limit the influence of these weight parameters on model output.

Given these additional insights, the remaining experiments from Figure 6 are performed with an updated set of model parameters. The updated parameter value for R_{sen} is 5, since this results in better surveillance effectiveness and efficiency compared to adjusting the local update constant of the ACO module, $\Delta\tau_{l0}$. Additionally, the APF module is disregarded by setting the magnitude of *Acknowledgement of the APF* to 0 in the remaining experiments, unless explicitly stated otherwise.

5.5 E4: Performance of HAPF-BLACOPS with (A)BNs and GSA parameters

This experiment simulates various (A)BNs with the adjusted parameters from E3 to test the hypothesis that adjusted HAPF-BLACOPS parameters can improve influence of (A)BNs on surveillance effectiveness and efficiency. The selection of simulated (A)BNs was determined from analysis of data discretisation, structure learning and various combinations of random variables to achieve an optimal DAG (refer to Appendix B for further elaboration). This analysis did not determine a single (A)BN that outperformed the others, based on scoring functions. Both ABNs and BNs were therefore simulated with the HAPF-BLACOPS model and compared the model behaviour obtained in E2, discussed in subsection 5.3.

The simulation results in Figure 10(a) indicate an improvement in surveillance effectiveness for ABN3 and ABN4 of approximately 40% in comparison to E2 (see Figure 8). Despite the small magnitude of the catch per unit, Figure 10(b) also shown an improvement of 40% in surveillance efficiency. This is the combined result from the correlation factors observed in Figure 9, which indicates that both *acknowledgement of the APF* and the communication distance R_{sen} have the strongest influence on these aspects. Additionally, adjustment of the local pheromone update rule also aids in achieving this performance improvement, albeit to a smaller extent. The statistical significance of these results in Table E.3 indicates that surveillance effectiveness and prioritisation is indeed significantly altered, whereas the improvement in surveillance efficiency does not meet the significance level of 5%. The hypothesis for E4, stated in Figure 6, is therefore partially accepted. These results are elaborated upon further below.



(a) Poachers apprehended by UAVs. (b) Catch per unit effort. (c) Coverage fairness.

Figure 10: Performance of HAPF-BLACOPS utilising (A)BNs and updated parameters obtained from GSA.

Model performance of the HAPF-ACO model without the adjusted parameter from the GSA is also presented in Figure 10. It is seen that surveillance effectiveness can be improved significantly by optimising model parameters, whereas prioritisation efforts are, in accordance with Figure 9, significantly reduced for the HAPF-BLACOPS model. Increased use of mission types 3 and 4 by disregarding the APF module shows that path planning is more severely influenced by the heuristic information, which emphasises the need for uniform surveillance of the entire target area.

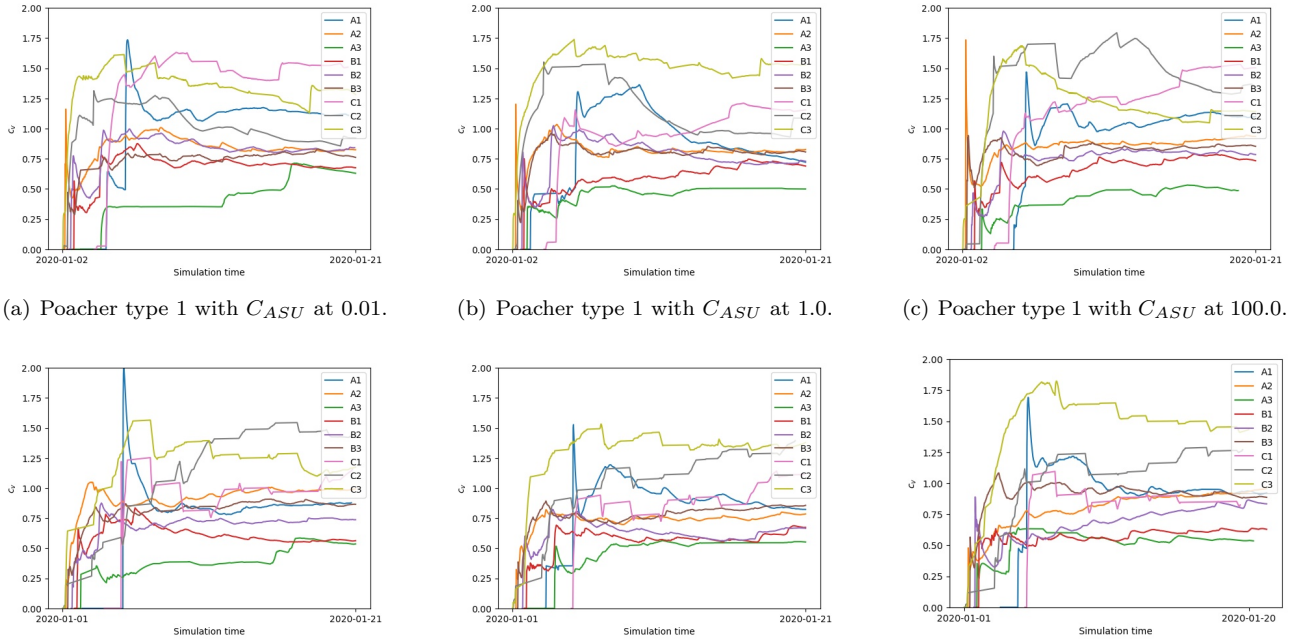
Additionally, using different (A)BNs does not significantly affect performance. The spatiotemporal patterns shown in the swarm plots in Appendix F do not show significant changes among these various (A)BNs either. The adjusted parameters do show changing patterns in comparison to results obtained in subsection 5.3, which indeed converge to some extent to the regions where poachers and rhino agents are active. This indicates that Algorithm 2 influences the effects of (A)BNs on surveillance efforts. Their prioritisation either does not alter path planning significantly or the (A)BNs do not capture the behaviour of the poachers correctly. Finally, the differences between the DAGs and their respective conditional probability tables do not provide closure on a distinct explanation for the observed variance in BN8 coverage fairness.

5.6 E7: Influence of adaptive poachers on surveillance effectiveness

Provided the knowledge gained through the previous experiments, this experiment analyses to what extent adaptive poacher behaviour can be observed and its potential effect on surveillance effectiveness. The ACOSG model simulates adaptive poacher behaviour by using gained experience regarding apprehensions and successful rhino kills to determine new strategies and perform path planning. The rate at which this experience results in adaptive behaviour depends on the amount of collected experience and the parameter C_{ASU}. These simulations analyse the influence of this parameter on the emergent poacher behaviour and the secondary effects on

surveillance effectiveness. It is hypothesised that an increased reliance on this experience through adjustment of the ASU function parameter destabilises the coefficient of variation of the poachers' spatiotemporal patterns. In addition, it is expected that surveillance effectiveness is reduced.

The coefficient of variation used to quantify whether adaptive behaviour can be observed in the spatiotemporal patterns of the poachers is shown in Figure 11. Provided that the magnitude of the inclination increases in response to a reduction in the magnitude of C_{ASU} (see Equation A.4), it is expected that the coefficient of variation becomes less stable as simulations with relatively small values of C_{ASU} progress. Specifically, this is expected for those regions where poachers find rhinos and are being apprehended by UAVs, which is within the BN regions A2, B1 and B2 (see Appendix A). The coefficient of variation for the remaining regions is therefore also expected to fluctuate as a result from the interchanged preference regarding the BN regions from Figure 5(b). Figure 11 does not show such trends in the poachers' spatiotemporal patterns, since the order of magnitude of the coefficient of variation remains relatively identical when the magnitude of C_{ASU} is altered. No consistent increase in the coefficient of variation and the spatiotemporal patterns (visualised in Appendix F) is observed and the hypothesis is therefore rejected.

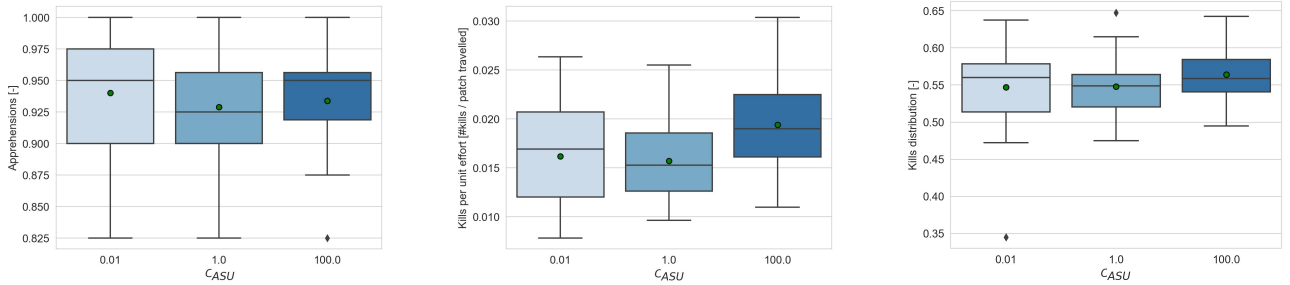


(a) Poacher type 1 with C_{ASU} at 0.01. (b) Poacher type 1 with C_{ASU} at 1.0. (c) Poacher type 1 with C_{ASU} at 100.0.

(d) Poacher type 2 with C_{ASU} at 0.01. (e) Poacher type 2 with C_{ASU} at 1.0. (f) Poacher type 2 with C_{ASU} at 100.0.

Figure 11: Coefficient of variation of BN region utilisation by poacher types 1 and 2 during simulation.

Figure 12 and Table E.6 elaborate on these observations by showing that surveillance effectiveness is not significantly altered as a result from adjustment of the parameter C_{ASU}. It is also confirmed by Figure 12(c) that the poachers do not utilise the target area in a different manner, which is either caused by the limited simulation time or a lack of experience. The calibration process, discussed in Appendix A, elaborates on the limited amount of experience that can be gained from killing rhinos and being apprehended.



(a) Apprehensions. (b) Loss of wildlife per unit effort. (c) Fraction of target area where rhinos are killed.

Figure 12: Influence of poacher experience on poacher and UAV behaviour.

5.7 E8: Compositional analysis to quantify BN impact

This experiment combines the gained knowledge from the previous experiments and analyses the impact of the (A)BNs on surveillance performance through compositional analysis. The model parameters for E8 were not updated in accordance with Figure 6, since the hypotheses in E5, E6 and E7 were rejected. E8 is performed to determine how internal interactions of the HAPF-BLACOPS model influence the effects of updating the ACO module with (A)BN predictions on surveillance performance. It tests the hypothesis that the complex interactions of the model's modules limit the (A)BNs influence on model output and that optimising model parameters can further improve surveillance performance in comparison to the HAPF-ACO model.

Simplifications of the internal interactions of the proposed surveillance model's mission types (MTs) are analysed by performing an experiment according to the model parameters in Table 4. The results from the GSA in subsection 5.4 are combined with the results found in subsection 5.5 and Appendix F. Both ABN4 and BN9 are simulated with the HAPF-BLACOPS model (see Table 4). The HAPF-ACO model is simulated as well, with the model parameters from E1, which is referred to as the reference HAPF-ACO model. The HAPF-BLACOPS model is also simulated with the model parameters from E1, which is referred to as the reference HAPF-BLACOPS model. The HAPF-ACO (Reference) and the HAPF-BLACOPS (Reference) models update the pheromone concentrations as specified in [40]. The HAPF-BLACOPS model is also simulated with ABN4 and the parameters of E2, which is referred to as the reference ABN4 model.

Additionally, the HAPF-BLACOPS model is simulated with both ABN4 and BN9 and the parameters obtained from the GSA. A number of alternatives are simulated as well, which include an increased swarm size and an alternative where β is set to zero. Setting β to zero essentially reduces Algorithm 2 to rely solely on the ACO module for path planning, since the APF and the heuristic information are disregarded. It was also chosen to disregard mission type 4, which simulates stochastic path planning based on the ACO module, when setting β to zero. This allows for analysis of the prioritisation that is expected to result from the (A)BNs in further detail, since the results in Figures 8 and 10 were unable to fully quantify the added value of the proposed (A)BNs. The simulation results are elaborated upon below and their statistical significance is summarised in Table E.8.

Table 4: Model parameters for E8.

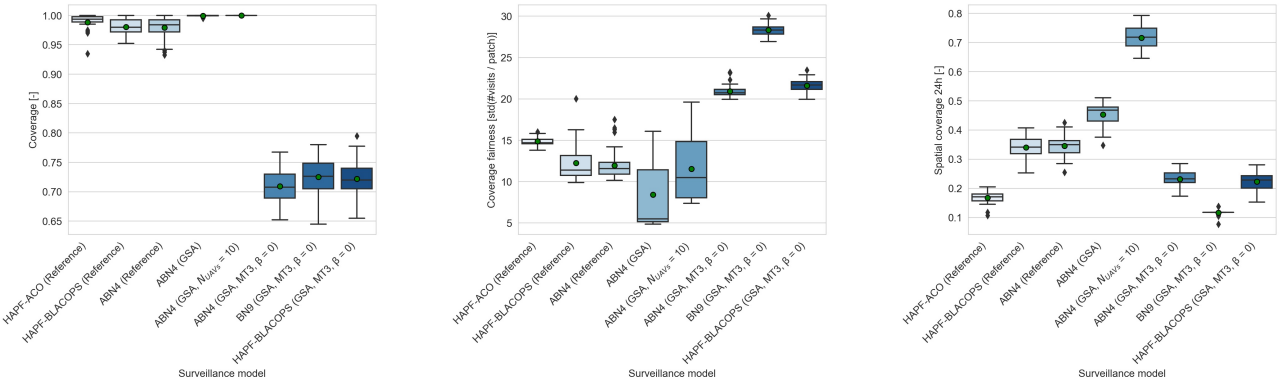
Surveillance model	**(A)BN**	N_{UAVs}	$\delta\tau_{l0}$	C_{ASU}	α	β	**use of MT2**
HAPF-ACO (Reference)	-	5	0.1	1.0	1.0	3.0	True
HAPF-BLACOPS (Reference)	-	5	0.1	1.0	1.0	3.0	True
ABN4 (Reference)	ABN4	5	0.1	1.0	1.0	3.0	True
ABN4 (GSA)	ABN4	5	0.1	1.0	1.0	3.0	False
ABN4 (GSA, $N_{UAVs} = 10$)	ABN4	10	0.1	1.0	1.0	3.0	False
ABN4 (GSA, MT3, $\beta = 0$)	ABN4	5	0.1	1.0	1.0	0.0	False
BN9 (GSA, MT3, $\beta = 0$)	BN9	5	0.1	1.0	1.0	0.0	False
HAPF-BLACOPS (GSA, MT3, $\beta = 0$)	-	5	0.1	1.0	1.0	0.0	False

Figures 13(a), 13(b) and 13(c) indicate how well the target area is being monitored by the UAV swarm for these various surveillance models. From comparison of the HAPF-ACO (Reference) and HAPF-BLACOPS (Reference) it is seen that the proposed mission selector (refer to Algorithm 2) results in a slight (insignificant) reduction of the total coverage of the target area. This is mainly a result from a significant increase in use of the APF module through mission type 2. This prevents UAV collisions (see Figure 14(a)) and draws the UAVs apart and therefore indirectly causes the UAVs to explore the target area at a higher rate (seen in Figure 13(c)). Integrating ABN4 does not result in significantly different model behaviour, as previously found in subsection 5.3. The interactions between the various mission types prevent the UAVs from consistently relying on the (A)BN's predictions to perform path planning.

The effects of incorporating GSA parameters is discussed in subsection 5.5. It influences the emergent behaviour of the UAV swarm significantly. Although significant changes in prioritisation of surveillance efforts are not achieved (see Figure 13(b)). Increasing the size of the swarm results in significant improvement regarding the loss of wildlife (see Figure 14(c)) and therefore surveillance effectiveness, but this merely results from an improved coverage rate (see Figure 13(c)).

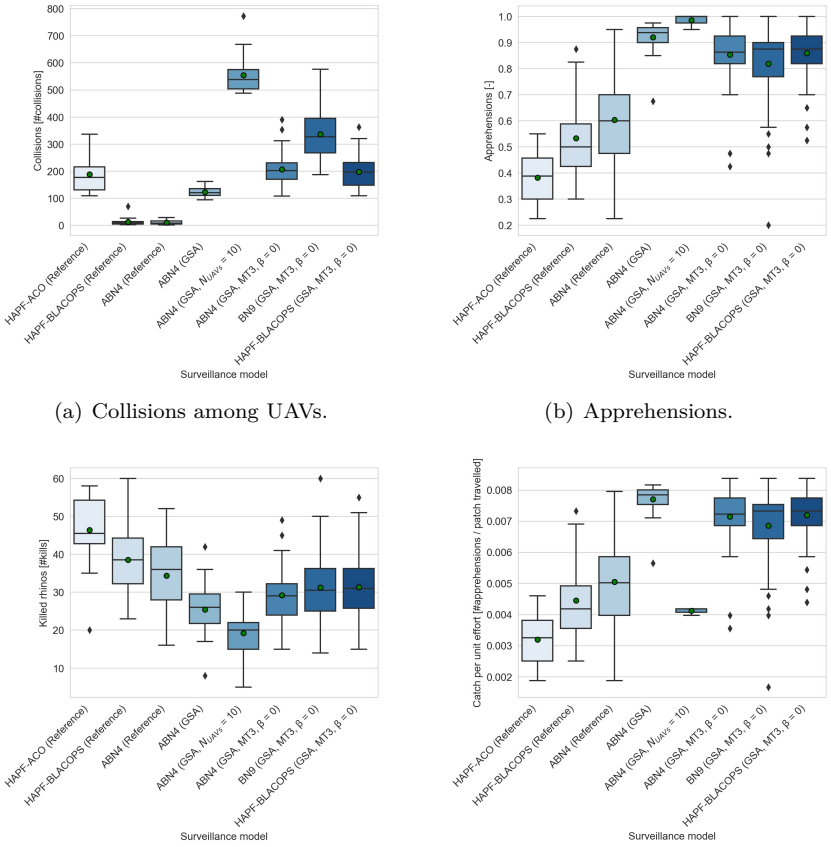
Setting the weight factor of the heuristic information β to zero, while simultaneously disregarding mission type 4, results in significant prioritisation of UAV surveillance efforts, see Figure 13(b) and the spatiotemporal patterns in Appendix F. By solely relying on the ACO and the BN for path planning, coverage of the target area and the rate at which coverage is performed drop significantly (see Figure 13) as a direct result from prioritised

surveillance. This comes at a cost of a severe increase in the number of mid-air UAV collisions (see Figure 14(a)) due to the absence of the APF module. It also negatively affects the loss of wildlife to some extent. Surveillance efficiency is not improved (see Figure 14(d)) from these prioritisation efforts, since the UAVs cannot hover above patches.



(a) Coverage of target area. (b) Coverage fairness. (c) Coverage in 1 day.

Figure 13: Performance comparison of utilising adjusted parameters obtained from GSA.



(a) Collisions among UAVs. (b) Apprehensions.

(c) Loss of wildlife. (d) Catch per unit effort.

Figure 14: Performance comparison of utilising adjusted parameters obtained from GSA.

The HAPF-BLACOPS model was also simulated without a (A)BN, with the GSA parameters, β set to zero and with sole use of mission type 3. Figures 13 and 14 show that the added value in terms of surveillance effectiveness and efficiency achieved from using (A)BNs is diminished. Despite the spatiotemporal patterns of the UAVs (presented in Appendix F) prioritising their efforts to the regions where rhinos and poachers are expected, the limitation of the turn angle and not having the option to hover limit this performance improvement. However, it is seen that the (A)BN's probabilistic inference (when using these GSA parameters, with β set to zero and

with only MT3 results in significantly different spatiotemporal patterns in comparison to this HAPF-BLACOPS model, that relies on the global update rule proposed in [40]. The complex interactions between the various mission types, combined with pheromone update rules, limits the influence of (A)BNs' probabilistic inference on path planning. The hypothesis stated previously is therefore accepted. Another limitation of the (A)BNs is that a time delay exists between prioritising a patch and arriving at the patch in question, such that poachers have the possibility to abandon the patch in question.

6 Discussion and future work

The performed experiments focused on three aspects. First, the hypothesis regarding the mission selector algorithm of the proposed HAPF-BLACOPS model was tested through simulation. Second, the learning rate of the poachers and its influence on surveillance effectiveness was analysed through adjustment of the ASU function in the proposed ACOSG model. Third, analysis of the effects of the (A)BNs on surveillance performance was performed through various steps in order to answer the research question stated in Section 3. The findings of these elements and their root cause are discussed in greater detail below.

6.1 Mission selector of HAPF-BLACOPS model

The proposed HAPF-BLACOPS surveillance model relies on a mission selection algorithm that addresses the deficiency of the current state-of-the-art HAPF-ACO model regarding the range constraint. This deficiency involves a lack of consideration of the maximum turn angle of a UAV in computing the distance required to return to their base and the possibility to neglect the priority of returning to their base for recharging. These factors condone a UAV to violate their range constraint. The proposed mission selector defines the interactions between the mission types that define UAV path planning and the emergent behaviour of the UAV swarm. It successfully re-prioritises the range constraint such that violations of the range constraint are prevented, thereby providing support for the hypothesis regarding the effectiveness of the proposed mission selector.

The frequency at which the APF module is used for path planning also increases significantly, thereby significantly reducing the amoutn of UAV collisions as a result from these re-prioritisations. These collisions are avoided by the swarm members through heading adjustments under influence of the APF module, which do not necessarily align with the optimal path resulting from the ACO module. This limitation of the APF module to a 2-dimensional surveillance space limits the possibilities for the swarm members to coordinate their surveillance efforts in accordance with the ACO's pheromone concentrations. Alternatively, speed adjustments, as proposed in [106], or altitude adjustments could be utilised to prevent collisions, without the APF interfering with the preferred path as described by the ACO module. Despite these potential conflicts between the APF and ACO modules, the proposed mission selection algorithm also achieves an increase of 50% in terms of the amount of apprehensions made. This results from the increased consistency of using the APF module for path planning, which motivates the swarm members to stay clear from each other. It does not result from the attraction forces, since the avoidance forces appear to be dominant in the APF module as a result from the presence of multiple swarm members. Additionally, the attraction forces do not actively steer swarm members towards future locations of a poacher. This effectively results in a larger spread of the swarm throughout the target area, which increases the apprehension rate as a secondary effect. Careful calibration of such modules is therefore recommended, as well as a third dimension such that interference of optimal solutions from the different modules can be minimised.

6.2 Effects of adaptive behaviour on surveillance performance

Both the adaptive behaviour of poachers and online learning of the UAV swarm were simulated. The hypotheses stated that adaptive poacher behaviour that results from gained experience would negatively affect surveillance effectiveness and destabilise the spatiotemporal patterns of the poachers.

Simulations of the proposed ACOSG and HAPF-BLACOPS model did not show these expected patterns however. The poachers' learning rate depends on the amount of apprehensions and killed wildlife, such that differentiation between regions within the target area depends on the amount of obtained rewards and the variance among the average of these rewards. Since the simulation time and the observed spatiotemporal patterns of the poachers did not allow for the collection of rewards in the entire target area, differentiation between various regions within the target area resulting from the ASU function is limited. The consideration of a homogeneous rhino population also limits differentiation, since collected rewards are similar in magnitude, therefore rendering different regions within the target area of equal interest as a result from the definition of the ASU function and the limited amount of gained experience. Analysis of the spatiotemporal patterns and success rate of the

intruders therefore did not indicate changing behaviour and the potential effect on surveillance effectiveness was not identified.

Analysis of the learning rate of the UAV swarm also showed that the conditional probability tables experience minor adjustments from updating the BNs with newly gained poacher data. This observation is in accordance with the absence of adaptive poacher behaviour. The low learning rate of the UAV swarm can be accredited to a number of factors, including a limited communication rate of new observations among swarm members, a limited amount of simulation time and the amount of data describing the conditional probability tables of the BN. A method where more recent observations are weighted relatively severe, could be able to overcome these limitations to some extent. Such a proposal is, in essence, used in [60]. Finally, the learning rate of the UAV swarm and the poachers are co-dependent and could therefore be analysed to a better extent in future research by considering these factors as well as a heterogeneous wildlife population.

6.3 Effects of (A)BNs on HAPF-BLACOPS performance

Both artificially created BNs and BNs trained from synthetic data were used to adjust the emergent behaviour of the ACO module within the proposed HAPF-BLACOPS model. A global sensitivity analysis of HAPF-BLACOPS model parameters and a compositional analysis of its modules showed that model parameters significantly affect internal model interactions and limit the influence of both the ACO and the (A)BN on the emergent behaviour. In fact, the APF was required to be disregarded in order for the (A)BNs to be able to adjust the spatiotemporal patterns of the UAV swarm. In addition, the difference in order of magnitude of the pheromone concentrations and the heuristic information used by the ACO module causes an unequal contribution of these aspects to ACO based path planning. Further simplifications made to the mission selector, by disregarding the heuristic information and relying solely on the ACO module within mission type 3 for path planning, showed promising results.

Despite difference in scores of the DAGs, surveillance performance is similar for the various (A)BNs that were considered. This indicates that the core model component required to predict simulated poacher behaviour is captured by a limited number of random variables. Indeed, incorporating various (a)biotic factors in the (A)BN results in different prioritisation of surveillance efforts as seen in the spatiotemporal patterns of the UAV swarm, without affecting the loss of wildlife however. The level of prioritisation achieved by the (A)BNs does not alter the main focus area where surveillance is being performed and therefore does not improve surveillance effectiveness in comparison to the HAPF-BLACOPS model utilising the pheromone update mechanism of the existing HAPF-ACO model. The probabilistic inference therefore needs more control over ACO based path planning in order to significantly adjust these spatiotemporal patterns. The main effect of relying solely on the ACO module and its (A)BN in mission type 3 is a reduction of 30% in terms of target area coverage without a loss in surveillance effectiveness, which results from prioritisation of surveillance efforts. This improvement comes at the cost of a significant increase of UAV collisions however, since the APF module was required to be disregarded for this behaviour to emerge.

This shows that dynamic prioritisation of surveillance efforts using the (A)BNs does not improve surveillance effectiveness on the long term, since the subtle prioritisation efforts that were achieved, are diminished by the continuous interactions between the (A)BN, the pheromone readjustments from swarm presence and by the synchronisation among swarm members. In addition, surveillance efficiency is not improved either, since the UAVs are forced to move to an adjacent patch at each time step and cannot hover above a patch of interest. The turn angle also limits a UAV from moving within small BN regions and therefore requires larger distances to be covered before being able to return. Research opportunities therefore exist in addressing the limitations of these module interactions. Additionally, this research was limited to expert judgement to define the BN regions, whereas machine learning could result in different spatiotemporal patterns and prioritisation of the UAV swarm.

7 Conclusions

This research focused on distributed coordination of continuous surveillance efforts in accordance with predictions of criminal activity based on (A)BNs. The ABM paradigm was used to explore the effects of learning and prioritised surveillance on performance of the proposed HAPF-BLACOPS model in a two-sided and dynamic environment.

Analysis of adaptive intruder behaviour, simulated with the ACOSG model, indicates that the learning rate of the intruders does not result in significantly changing spatiotemporal patterns. This is accredited to the limited

amount of rewards gained from the simulated wildlife population and the homogeneity of these rewards, which limits differentiation between regions within the target area. A learning rate in the intruders' behaviour is therefore not identified. The anticipated negative influence of the poachers' adaptability on surveillance effectiveness and the loss of wildlife was therefore not observed. Although the swarm achieves online learning by adjusting the conditional probability tables of the BNs in accordance with new observations regarding poacher behaviour, these adjustments do not result in significantly different model predictions regarding poacher activity. The spatiotemporal patterns do not evolve, since these new observations do not describe changed intruder behaviour. In addition, newly gained observations are not prioritised by the BN over historic data. The learning rate of the swarm therefore results in insignificant changes of the conditional probability tables, which do not alter the spatiotemporal patterns of the swarm within the simulation time.

The proposed HAPF-BLACOPS model does outperform the current state-of-the-art HAPF-ACO model on various aspects. The adjustments to the mission selection algorithm readjust the priorities between the mission objectives. This results in a 50% improvement regarding surveillance effectiveness due to increased reliance on the APF module, while simultaneously significantly reducing the amount of collisions and preventing range constraint violations.

Further analysis through global sensitivity analysis and compositional analysis shows that model parameters significantly influence the emergent behaviour of the swarm and limit the influence of the ACO module and the (A)BN on path planning. Adjustment of model input parameters can significantly reduce the loss of wildlife, which results mainly from a reduction in use of the model's APF module. Simulation results show that path planning based on the APF and the ACO modules conflict, since the APF is limited to heading adjustments to avoid collisions. Further simplification of the HAPF-BLACOPS model by neglecting the heuristic information and analysis of the spatiotemporal patterns of the swarm indicates that prioritisation through the (A)BN's probabilistic inference is indeed achieved, albeit that the spatiotemporal patterns are only slightly adjusted. The main focus of surveillance efforts are not altered by the (A)BNs. In addition, the various (A)BNs achieve similar levels of performance, which indicates that only a small subset of the random variables under consideration is required to predict poacher behaviour. Despite the influence of the (A)BNs on the spatiotemporal patterns, their influence on the ACO module is diminished by the complex interactions within the UAV swarm. This results from synchronisation and from accounting for the presence of swarm members. Nevertheless, adjustment of the HAPF-BLACOPS model parameters results in a level of surveillance effectiveness the is similar to the current state-of-the-art, at a reduction of 30% in coverage of the target area. These improvements come at the cost of a significant increase in the number of mid-air collisions however.

These simulation results and the accompanying spatiotemporal patterns indicate that prioritised surveillance shows to be promising for addressing illegal activity. Future research opportunities therefore lie in analysing the effects of three-dimensional path planning to reduce the interference between the APF and ACO modules. Additionally, further analysis of model parameters could shed light on the complex interactions within the HAPF-BLACOPS model that limit the effects of updating pheromone concentrations through probabilistic inference on the prioritisation efforts and the spatiotemporal patterns of the UAV swarm.

References

[1] E. J. Milner-Gulland and N. Leader-Williams, "A Model of Incentives for the Illegal Exploitation of Black Rhinos and Elephants: Poaching Pays in Luangwa Valley, Zambia," *The Journal of Applied Ecology*, vol. 29, no. 2, p. 401, 1992.

[2] H. Hearne, J. Swart, and P. Goodman, "A conservation model for black rhino," *ORiON*, vol. 7, dec 2003.

[3] E. Reuter and L. Bisschop, "Keeping the Horn on the Rhino: A Study of Balule Nature Reserve," in *The Geography of Environmental Crime*, pp. 149–185, Palgrave Macmillan UK, 2016.

[4] H. Thambi Prem, "What are the biggest threats to Wildlife and Why?," feb 2020.

[5] J. Bergenas, R. Stohl, and A. Georgieff, "The other side of drones: Saving wildlife in Africa and managing global crime," tech. rep., 2013.

[6] Bureau of Oceans and International Environmental and Scientific Affairs, "Eliminate, Neutralize, and Disrupt Wildlife Trafficking Act of 2016 PL 114-231, Sec. 301(d) 2019 Strategic Review," nov 2019.

[7] J. Worland, "Drones Are Helping Catch Poachers Operating Under Cover of Darkness," 2018.

[8] United Nations, "World Wildlife Report," 2020.

[9] F. Fang, P. Stone, and M. Tambe, "When Security Games Go Green: Designing Defender Strategies to Prevent Poaching and Illegal Fishing," tech. rep., 2015.

[10] B. Wang, Y. Zhang, Z. H. Zhou, and S. Zhong, "On repeated stackelberg security game with the cooperative human behavior model for wildlife protection," *Applied Intelligence*, vol. 49, pp. 1002–1015, mar 2019.

[11] Etosha National Park, "Can Drones Save Etosha's Wildlife?," aug 2014.

[12] H. Chen, T. Cheng, and S. Wise, "Developing an online cooperative police patrol routing strategy," *Computers, Environment and Urban Systems*, vol. 62, pp. 19–29, mar 2017.

[13] K. E. Doull, C. Chalmers, P. Fergus, S. Longmore, A. K. Piel, and S. A. Wich, "An evaluation of the factors affecting poacher' detection with drones and the efficacy of machine-learning for detection," *Sensors*, vol. 21, jun 2021.

[14] J. J. López and M. Mulero-Pázmány, "Drones for conservation in protected areas: Present and future," mar 2019.

[15] R. Kaufman, "Hunter Becomes the Hunted," dec 2015.

[16] J. C. Hodgson, S. M. Baylis, R. Mott, A. Herrod, and R. H. Clarke, "Precision wildlife monitoring using unmanned aerial vehicles," *Scientific Reports*, vol. 6, mar 2016.

[17] M. A. Olivares-Mendez, C. Fu, P. Ludivig, T. F. Bissyandé, S. Kannan, M. Zurad, A. Annaiyan, H. Voos, and P. Campoy, "Towards an autonomous vision-based unmanned aerial system againstwildlife poachers," *Sensors (Switzerland)*, vol. 15, pp. 31362–31391, dec 2015.

[18] S. Gholami, S. McCarthy, B. Dilkina, A. Plumptre, M. Tambe, M. Driciru, F. Wanyama, A. Rwetsiba, M. Nsubaga, J. Mabonga, T. Okello, and E. Enyel, "Adversary models account for imperfect crime data: Forecasting and planning against real-world poachers," in *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS*, vol. 2, pp. 823–831, International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS), 2018.

[19] J. M. Caplan, L. W. Kennedy, and J. Miller, "Risk Terrain Modeling: Brokering Criminological Theory and GIS Methods for Crime Forecasting," *Justice Quarterly*, vol. 28, no. 2, pp. 360–381, 2011.

[20] B. Śniezyński, "Agent strategy generation by rule induction in predator-prey problem," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. PART 2, pp. 895–903, 2009.

[21] B. Sadel and B. Śniezyński, "Online supervised learning approach for machine scheduling," *Schedae Informaticae*, vol. 25, pp. 165–176, 2016.

[22] R. Critchlow, A. J. Plumptre, B. Alidria, M. Nsubuga, M. Driciru, A. Rwetsiba, F. Wanyama, and C. M. Beale, "Improving Law-Enforcement Effectiveness and Efficiency in Protected Areas Using Ranger-collected Monitoring Data," sep 2017.

[23] M. Raap, M. Preuß, and S. Meyer-Nieberg, "Moving target search optimization A literature review," may 2019.

[24] F. Zong, H. Xu, and H. Zhang, "Prediction for traffic accident severity: Comparing the bayesian network and regression models1," *Mathematical Problems in Engineering*, vol. 2013, 2013.

[25] P. Arora, D. Boyne, J. J. Slater, A. Gupta, D. R. Brenner, and M. J. Druzdzel, "Bayesian Networks for Risk Prediction Using Real-World Data: A Tool for Precision Medicine," *Value in Health*, vol. 22, pp. 439–445, apr 2019.

[26] J. M. Caplan, L. W. Kennedy, E. L. Piza, and J. D. Barnum, "Using Vulnerability and Exposure to Improve Robbery Prediction and Target Area Selection," *Applied Spatial Analysis and Policy*, vol. 13, pp. 113–136, mar 2020.

[27] N. H. M. Shamsuddin, N. A. Ali, and R. Alwee, "An overview on crime prediction methods," in *6th ICT International Student Project Conference: Elevating Community Through ICT, ICT-ISPC 2017*, vol. 2017-Janua, pp. 1–5, Institute of Electrical and Electronics Engineers Inc., oct 2017.

[28] C. Zhang, S. Gholami, D. Kar, A. Sinha, M. Jain, R. Goyal, and M. Tambe, "Keeping Pace with Criminals: An Extended Study of Designing Patrol Allocation against Adaptive Opportunistic Criminals," *Games*, vol. 7, p. 15, jun 2016.

[29] K. M. Khalil, M. Abdel-Aziz, T. T. Nazmy, and A. B. M. Salem, "MLIMAS: A Framework for Machine Learning in Interactive Multi-agent Systems," in *Procedia Computer Science*, vol. 65, pp. 827–835, Elsevier, 2015.

[30] N. Nigam, "The multiple unmanned Air Vehicle persistent surveillance problem: A review," *Machines*, vol. 2, pp. 13–72, mar 2014.

[31] J. Kamminga, E. Ayele, N. Meratnia, and P. Havinga, "Poaching detection technologies-A survey," *Sensors (Switzerland)*, vol. 18, may 2018.

[32] T. M. Cabreira, L. B. Brisolara, and R. Ferreira Paulo, "Survey on coverage path planning with unmanned aerial vehicles," *Drones*, vol. 3, pp. 1–38, mar 2019.

[33] W. Meng, Z. He, R. Su, P. K. Yadav, R. Teo, and L. Xie, "Decentralized Multi-UAV Flight Autonomy for Moving Convoys Search and Track," *IEEE Transactions on Control Systems Technology*, vol. 25, pp. 1480–1487, jul 2017.

[34] D. Howden, "Continuous swarm surveillance via distributed priority maps," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 5865 LNAI, pp. 221–231, 2009.

[35] D. J. Howden, "Fire tracking with collective intelligence using dynamic priority maps," in *2013 IEEE Congress on Evolutionary Computation, CEC 2013*, pp. 2610–2617, 2013.

[36] M. Silvagni, A. Tonoli, E. Zenerino, and M. Chiaberge, "Multipurpose UAV for search and rescue operations in mountain avalanche events," *http://www.tandfonline.com/action/journalInformation?show=aimsScope&journalCode=tgnh20#.VsXodSCLRhE*, vol. 8, pp. 18–33, jan 2016.

[37] H. Chen, T. Cheng, and S. Wise, "Designing daily patrol routes for policing based on ant colony algorithm," in *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 2, pp. 103–109, Copernicus GmbH, jul 2015.

[38] M. Ramasamy and D. Ghose, "A Heuristic Learning Algorithm for Preferential Area Surveillance by Unmanned Aerial Vehicles," *Journal of Intelligent and Robotic Systems: Theory and Applications*, vol. 88, pp. 655–681, dec 2017.

[39] B. Wiandt, V. Simon, and A. Kokuti, "Self-organized graph partitioning approach for multi-agent patrolling in generic graphs," in *17th IEEE International Conference on Smart Technologies, EUROCON 2017 - Conference Proceedings*, pp. 605–610, Institute of Electrical and Electronics Engineers Inc., aug 2017.

[40] Z. Zhen, Y. Chen, L. Wen, and B. Han, "An intelligent cooperative mission planning scheme of UAV swarm in uncertain dynamic environment," *Aerospace Science and Technology*, vol. 100, may 2020.

[41] S. Kalantari, E. Nazemi, and B. Masoumi, "Emergence-based self-advising in strong self-organizing systems: A case study in NASA ANTS mission," *Expert Systems with Applications*, vol. 182, p. 115187, nov 2021.

[42] L. A. Márquez-Vega, M. Aguilera-Ruiz, and L. M. Torres-Treviño, "Multi-objective optimization of a quadrotor flock performing target zone search," *Swarm and Evolutionary Computation*, vol. 60, feb 2021.

[43] F. Fang and T. H. Nguyen, "Green security games: apply game theory to addressing green security challenges," *ACM SIGecom Exchanges*, vol. 15, pp. 78–83, sep 2016.

[44] H. Xu, B. Ford, F. Fang, B. Dilkina, A. Plumptre, M. Tambe, M. Driciru, F. Wanyama, A. Rwetsiba, M. Nsubaga, and J. Mabonga, "Optimal Patrol Planning for Green Security Games with Black-Box Attackers," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10575 LNCS, pp. 458–477, Springer Verlag, 2017.

[45] T. H. Nguyen, A. Sinha, S. Gholami, A. Plumptre, L. Joppa, M. Tambe, M. Driciru, F. Wanyama, A. Rwetsiba, R. Critchlow, and C. M. Beale, "CAPTURE: A New Predictive Anti-Poaching Tool for Wildlife Protection," *AAMAS '16: Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, pp. 767–775, may 2016.

[46] F. Fang, A. X. Jiang, and M. Tambe, "Optimal patrol strategy for protecting moving targets with multiple mobile resources," in *12th International Conference on Autonomous Agents and Multiagent Systems 2013, AAMAS 2013*, vol. 2, pp. 957–964, International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS), 2013.

[47] J. R. Peters, S. J. Wang, and F. Bullo, "Coverage control with anytime updates for persistent surveillance missions," in *Proceedings of the American Control Conference*, pp. 265–270, Institute of Electrical and Electronics Engineers Inc., jun 2017.

[48] R. Jain, R. Tiwari, P. Jain, and P. B. Sujit, "Distributed Fault Tolerant and Balanced Multi-Robot Area Partitioning for Coverage Applications," in *2018 International Conference on Unmanned Aircraft Systems, ICUAS 2018*, pp. 293–299, Institute of Electrical and Electronics Engineers Inc., aug 2018.

[49] E. Neil, J. K. Madsen, E. Carrella, N. Payette, and R. Bailey, "Agent-based modelling as a tool for elephant poaching mitigation," *Ecological Modelling*, vol. 427, p. 109054, jul 2020.

[50] F. Pasqualetti, J. W. Durham, and F. Bullo, "Cooperative patrolling via weighted tours: Performance analysis and distributed algorithms," *IEEE Transactions on Robotics*, vol. 28, no. 5, pp. 1181–1188, 2012.

[51] J. J. Acevedo, B. C. Arrue, J. M. Diaz-Bañez, I. Ventura, I. Maza, and A. Ollero, "One-to-one coordination algorithm for decentralized area partition in surveillance missions with a team of aerial robots," *Journal of Intelligent and Robotic Systems: Theory and Applications*, vol. 74, pp. 269–285, apr 2014.

[52] Y. Altshuler, A. Pentland, and A. M. Bruckstein, "The cooperative hunters Efficient and scalable drones swarm for multiple targets detection," in *Studies in Computational Intelligence*, vol. 729, pp. 187–205, Springer Verlag, 2018.

[53] A. Jahn, R. J. Alitappeh, D. Saldana, L. C. Pimenta, A. G. Santos, and M. F. Campos, "Distributed multi-robot coordination for dynamic perimeter surveillance in uncertain environments," in *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 273–278, Institute of Electrical and Electronics Engineers Inc., jul 2017.

[54] J. Scherer and B. Rinner, "Multi-Robot Persistent Surveillance with Connectivity Constraints," *IEEE Access*, vol. 8, pp. 15093–15109, 2020.

[55] A. V. Borkar, A. Sinha, L. Vachhani, and H. Arya, "Application of Lissajous curves in trajectory planning of multiple agents," *Autonomous Robots*, vol. 44, pp. 233–250, jan 2020.

[56] T. M. Kratzke, L. D. Stone, and J. R. Frost, "Search and Rescue Optimal Planning System," in *13th Conference on Information Fusion, Fusion 2010*, IEEE Computer Society, 2010.

[57] B. Yang, Y. Ding, and K. Hao, "Area coverage searching for swarm robots using dynamic Voronoi-based method," in *Chinese Control Conference, CCC*, vol. 2015-Septe, pp. 6090–6094, IEEE Computer Society, sep 2015.

[58] P. Gaudiano, B. Shargel, E. Bonabeau, and B. T. Clough, "Control of UAV Swarms: What the bugs can teach us," in *2nd AIAA "Unmanned Unlimited" Conference and Workshop and Exhibit*, American Institute of Aeronautics and Astronautics Inc., 2003.

[59] M. Schranz, G. A. Di Caro, T. Schmickl, W. Elmenreich, F. Arvin, A. ekerciolu, and M. Sende, "Swarm Intelligence and cyber-physical systems: Concepts, challenges and future trends," *Swarm and Evolutionary Computation*, vol. 60, feb 2021.

[60] H. Wu, H. Li, R. Xiao, and J. Liu, "Modeling and simulation of dynamic ant colony's labor division for task allocation of UAV swarm," *Physica A: Statistical Mechanics and its Applications*, vol. 491, pp. 127–141, feb 2018.

[61] Y. Jin, Y. Wu, and N. Fan, "Research on distributed cooperative control of swarm UAVs for persistent coverage," in *Proceedings of the 33rd Chinese Control Conference, CCC 2014*, pp. 1162–1167, IEEE Computer Society, sep 2014.

[62] R. D. McKelvey and T. R. Palfrey, "Quantal response equilibria for normal form games," *Games and Economic Behavior*, vol. 10, no. 1, pp. 6–38, 1995.

[63] T. H. Nguyen, R. Yang, A. Azaria, S. Kraus, and M. Tambe, "Analyzing the effectiveness of adversary modeling in security games," in *Proceedings of the 27th AAAI Conference on Artificial Intelligence, AAAI 2013*, pp. 718–724, 2013.

[64] Y. Zhao, M. Li, and C. Guo, "Developing Patrol Strategies for the Cooperative Opportunistic Criminals," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11944 LNCS, pp. 454–467, Springer, 2020.

[65] A. J. McLane, C. Semeniuk, G. J. McDermid, and D. J. Marceau, "The role of agent-based models in wildlife ecology and management," *Ecological Modelling*, vol. 222, pp. 1544–1556, apr 2011.

[66] S. R. Lele and J. L. Keim, "Weighted distributions and estimation of resource selection probability functions," *Ecology*, vol. 87, pp. 3021–3028, dec 2006.

[67] R. Nathan, W. M. Getz, E. Revilla, M. Holyoak, R. Kadmon, D. Saltz, and P. E. Smouse, "A movement ecology paradigm for unifying organismal movement research," dec 2008.

[68] P. E. Smouse, S. Focardi, P. R. Moorcroft, J. G. Kie, J. D. Forester, and J. M. Morales, "Stochastic modelling of animal movement," jul 2010.

[69] E. A. Codling, M. J. Plank, and S. Benhamou, "Random walk models in biology," aug 2008.

[70] G. Aarts, J. Fieberg, S. Brasseur, and J. Matthiopoulos, "Quantifying the effect of habitat availability on species distributions," *Journal of Animal Ecology*, vol. 82, no. 6, 2013.

[71] L. Lush, M. Mulama, and M. Jones, "Predicting the habitat usage of African black rhinoceros (Diceros bicornis) using random forest models," *African Journal of Ecology*, vol. 53, pp. 346–354, sep 2015.

[72] J. Matthiopoulos, J. Fieberg, G. Aarts, H. L. Beyer, J. M. Morales, and D. T. Haydon, "Establishing the link between habitat selection and animal population dynamics," *Ecological Monographs*, vol. 85, no. 3, 2015.

[73] K. Kuchar, E. Holasova, L. Hrboticky, M. Rajnoha, and R. Burget, "Supervised learning in multi-agent environments using inverse point of view," in *2019 42nd International Conference on Telecommunications and Signal Processing, TSP 2019*, pp. 625–628, Institute of Electrical and Electronics Engineers Inc., jul 2019.

[74] N. van Doormaal, "Exploring anti-poaching strategies for wildlife crime with a simple and general agent-based model," in *18th EPIA Conference on Artificial Intelligence*, vol. 10423 LNAI, pp. 51–62, Springer Verlag, 2017.

[75] S. Marino, I. B. Hogue, C. J. Ray, and D. E. Kirschner, "A methodology for performing global uncertainty and sensitivity analysis in systems biology," sep 2008.

[76] K. Dhoore, "Online Agent-Based Aerial Patrol Planning for Wildlife Surveillance," tech. rep., Delft University of Technology, oct 2020.

[77] G. Weiss, *Multiagent Systems*. Cambridge, Massachusetts: MIT Press, 2 ed., 2013.

[78] E. R. Smith and F. R. Conrey, "Agent-based modeling: A new approach for theory building in social psychology," *Personality and Social Psychology Review*, vol. 11, pp. 87–104, feb 2007.

[79] USFWS, "Standard Measures of Effectiveness and Threats for Wildlife Conservation in Central Africa: Guidance for USFWS Applicants," tech. rep., oct 2014.

[80] D. P. Seidel, W. L. Linklater, W. Kilian, P. D. Preez, and W. M. Getz, "Mesoscale movement and recursion behaviors of Namibian black rhinos," *Movement Ecology*, vol. 7, p. 34, nov 2019.

[81] F. Fang, T. H. Nguyen, A. Sinha, S. Gholami, A. Plumptre, L. Joppa, M. Tambe, M. Driciru, F. Wanyama, A. Rwetsiba, R. Critchlow, and C. M. Beale, "Predicting poaching for wildlife Protection," *IBM Journal of Research and Development*, vol. 61, nov 2017.

[82] NASA JPL, "NASA Shuttle Radar Topography Mission Global 1 arc second [Data set]," 2013.

[83] U.S. Geological Survey, "Landsat Surface Reflectance Data (ver. 1.1, March 27, 2019): U.S. Geological Survey Fact Sheet 2015-3034," tech. rep., 2015.

[84] J. R. Muntifering, W. L. Linklater, R. Naidoo, S. Khob, P. Du Preez, P. Beytell, S. Jacobs, and A. T. Knight, "Black rhinoceros avoidance of tourist infrastructure and activity: Planning and managing for coexistence," *ORYX*, vol. 55, no. 1, 2021.

[85] S. A. Oginah, P. O. Ang'ienda, and P. O. Onyango, "Evaluation of habitat use and ecological carrying capacity for the reintroduced Eastern black rhinoceros (Diceros bicornis michaeli) in Ruma National Park, Kenya," *African Journal of Ecology*, vol. 58, pp. 34–45, mar 2020.

[86] H. Thurfjell, S. Ciuti, and M. S. Boyce, "Applications of step-selection functions in ecology and conservation," feb 2014.

[87] M. E. Roberts and R. L. Goldstone, "EPICURE: Spatial and Knowledge Limitations in Group Foraging," *Adaptive Behavior*, vol. 14, pp. 291–313, dec 2006.

[88] L. M. Leemis and S. Park, *Discrete-event simulation : a first course.* New Jersey :: Pearson/Prentice Hall„ 2006.

[89] E. Corcoran, S. Denman, J. Hanger, B. Wilson, and G. Hamilton, "Automated detection of koalas using low-level aerial surveillance and machine learning," *Scientific Reports*, vol. 9, pp. 1–9, dec 2019.

[90] T. Tang, S. Zhou, Z. Deng, H. Zou, and L. Lei, "Vehicle Detection in Aerial Images Based on Region Convolutional Neural Networks and Hard Negative Example Mining," *Sensors*, vol. 17, p. 336, feb 2017.

[91] R. Kays, J. Sheppard, K. Mclean, C. Welch, C. Paunescu, V. Wang, G. Kravit, and M. Crofoot, "Hot monkey, cold reality: surveying rainforest canopy mammals using drone-mounted thermal infrared sensors," *International Journal of Remote Sensing*, vol. 40, pp. 407–419, jan 2019.

[92] J. Li, Y. Dai, C. Li, J. Shu, D. Li, T. Yang, and Z. Lu, "Visual Detail Augmented Mapping for Small Aerial Target Detection," *Remote Sensing*, vol. 11, p. 14, dec 2018.

[93] A. González, Z. Fang, Y. Socarras, J. Serrat, D. Vázquez, J. Xu, and A. López, "Pedestrian Detection at Day/Night Time with Visible and FIR Cameras: A Comparison," *Sensors*, vol. 16, p. 820, jun 2016.

[94] M. Rosalie, G. Danoy, S. Chaumette, and P. Bouvry, "From random process to chaotic behavior in swarms of UAVs," in *DIVANet 2016 - Proceedings of the 6th ACM Symposium on Development and Analysis of Intelligent Vehicular Networks and Applications, co-located with MSWiM 2016*, pp. 9–15, Association for Computing Machinery, Inc, nov 2016.

[95] Y. J. Lee, O. SooHyun, and J. E. Eck, "A Theory-Driven Algorithm for Real-Time Crime Hot Spot Forecasting," *Police Quarterly*, vol. 23, pp. 174–201, jun 2020.

[96] K. J. Bowers, "Prospective Hot-Spotting: The Future of Crime Mapping?," *British Journal of Criminology*, vol. 44, pp. 641–658, may 2004.

[97] T. Kuiper, B. Kavhu, N. A. Ngwenya, R. Mandisodza-Chikerema, and E. J. Milner-Gulland, "Rangers and modellers collaborate to build and evaluate spatial models of African elephant poaching," *Biological Conservation*, vol. 243, p. 108486, mar 2020.

[98] L. E. Sucar, *Probabilistic graphical models : principles and applications.* Cham: Springer„ 2021.

[99] M. Michiels, P. Larrañaga, and C. Bielza, "NeuroSuites-BNs: An open web framework for massive Bayesian networks focused on neuroscience," feb 2020.

[100] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*. London, England: Cambridge, Massachusetts: The MIT Press, 2009.

[101] L. M. De Campos, "A Scoring Function for Learning Bayesian Networks based on Mutual Information and Conditional Independence Tests," *Journal of Machine Learning Research*, vol. 7, pp. 2149–2187, 2006.

[102] J. Law, M. Quick, and A. Jadavji, "A Bayesian spatial shared component model for identifying crime-general and crime-specific hotspots," *Annals of GIS*, vol. 26, pp. 65–79, jan 2020.

[103] H. Iba, *Agent-based modeling and simulation with swarm*. CRC Press, jan 2013.

[104] I. Lorscheid, B. O. Heine, and M. Meyer, "Opening the 'Black Box' of Simulations: Increased Transparency and Effective Communication Through the Systematic Design of Experiments," *Computational and Mathematical Organization Theory*, vol. 18, pp. 22–62, mar 2012.

[105] D. C. Montgomery, *Design and Analysis of Experiments*. John Wiley & Sons, Inc., 9 ed., 2017.

[106] T. M. Lam, H. W. Boschloo, M. Mulder, and M. M. Van Paassen, "Artificial force field for haptic feedback in UAV teleoperation," *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans*, vol. 39, pp. 1316–1330, nov 2009.

[107] D. Masad and J. Kazil, "Mesa: An Agent-Based Modeling Framework," in *Proceedings of the 14th Python in Science Conference*, pp. 51–58, SciPy, 2015.

[108] Google, "[Satellite image provided by Google Maps of Aloegrove Safari Park in Namibia]. Retrieved March 2, 2021."

[109] B. Okita-Ouma, R. Amin, F. van Langevelde, and N. Leader-Williams, "Density dependence and population dynamics of black rhinos (Diceros bicornis michaeli) in Kenya's rhino sanctuaries," *African Journal of Ecology*, vol. 48, pp. 791–799, oct 2009.

[110] Poaching Facts, "Rhino Poaching Statistics," 2020.

[111] S. Kotsiantis and D. Kanellopoulos, "Discretization Techniques: A recent survey," *undefined*, 2006.

[112] D. Freedman and P. Diaconis, "On the histogram as a density estimator:L2 theory," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 57, pp. 453–476, dec 1981.

[113] M. Rosalie, M. R. Brust, G. Danoy, S. Chaumette, and P. Bouvry, "Coverage Optimization with Connectivity Preservation for UAV Swarms Applying Chaotic Dynamics," in *Proceedings - 2017 IEEE International Conference on Autonomic Computing, ICAC 2017*, pp. 113–118, Institute of Electrical and Electronics Engineers Inc., aug 2017.

[114] J. Dentler, M. Rosalie, G. Danoy, P. Bouvry, S. Kannan, M. A. Olivares-Mendez, and H. Voos, "Collision Avoidance Effects on the Mobility of a UAV Swarm Using Chaotic Ant Colony with Model Predictive Control," *Journal of Intelligent and Robotic Systems: Theory and Applications*, vol. 93, pp. 227–243, feb 2019.

[115] P. T. Millet, D. W. Casbeer, T. Mercker, and J. L. Bishop, "Multi-agent decentralized search of a probability map with communication constraints," in *AIAA Guidance, Navigation, and Control Conference*, 2010.

II

Supporting work

A

Model elaboration

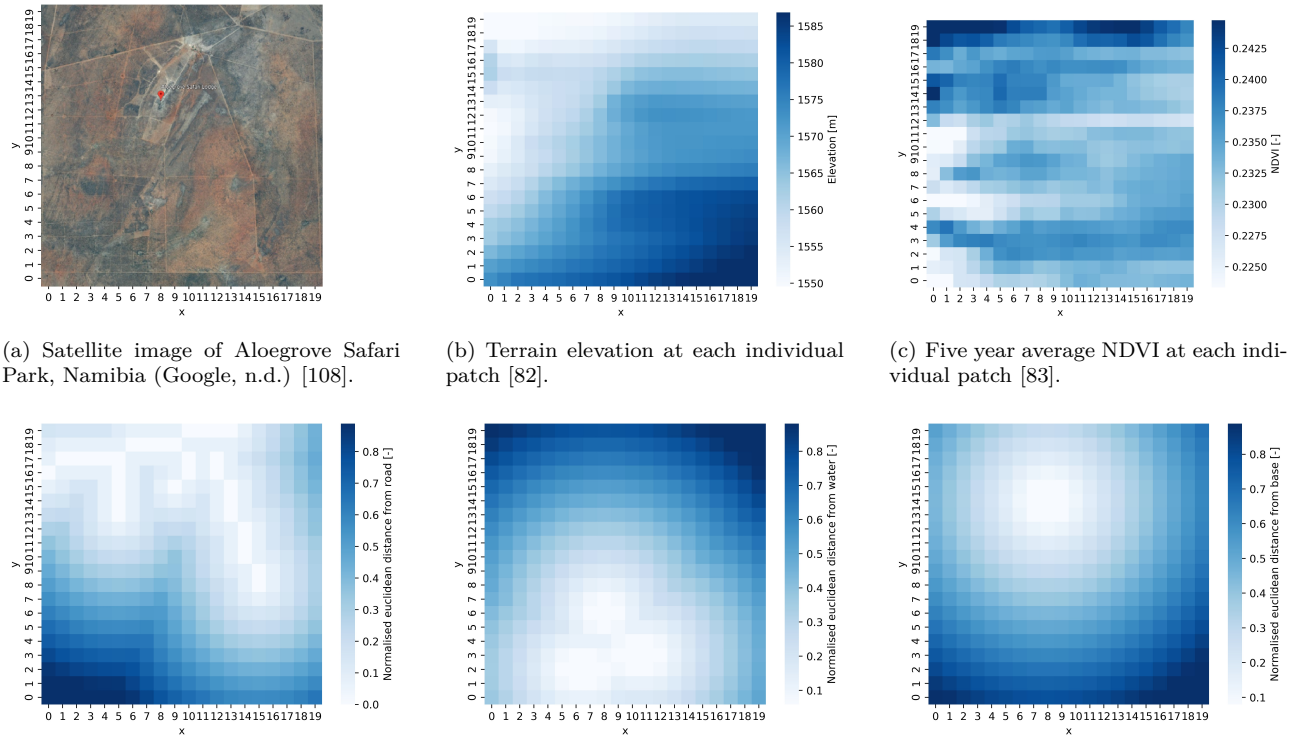
This appendix provides a detailed elaboration of the ABM specification. The ABM is built on the Mesa framework [107] and provides a bottom-up approach to simulating the interactions of the ABM elements [103]. An ABM consists of multiple elements, referred to as agents, that form a multi-agent system. The behaviour of each agent is determined by rules that describe its response to events in the environment. The ABM in consideration consists of three agents; the rhino, poacher and UAV agents. Their interactions result in higher-order system behaviour, referred to as emergent behaviour [103]. The remaining model element is the environment, in which the agents aim to achieve their objectives. This appendix is structured such that details regarding the environment are provided first in Section A.1. The following sections elaborate on the behavioural properties of the individual agents. Sections A.2, A.3, A.4 elaborate on the rhino, poacher and UAV agents respectively. Finally, section A.5 analyses the variability of model output through the coefficient of variation in order to determine the required sample size and simulation time for the experiments.

A.1 Environment

Aloegrove Safari Park in Namibida provides the basis for the environment in the ABM. This particular target area is selected due to available expert knowledge from Eyeplane regarding poaching activity in the area [76]. The 10 by 10km target area shown in Figure A.1(a) is discretised into 400 equally sized patches. Each patch is characterised by a number of abiotic factors that are relevant to rhinos and poachers. These include the presence of water sources [45, 3], roads [79, 3], vegetation (density) [80, 45, 71] and the terrain slope [81]. The ground elevation in the target area is quantified using SRTM V3 data [82]. The elevation h of an individual patch is calculated from the mean of available elevation data within this patch, for which the result is visualised in Figure A.1(b). In turn, the terrain slope between adjacent patches σ_i and σ_j is computed in accordance with Equation A.1, where Δd is the unit distance.

$$\text{terrain slope}(\sigma_i, \sigma_j) = \frac{h_{\sigma_j} - h_{\sigma_i}}{\Delta d} \tag{A.1}$$

The vegetation density is quantified using NDVI data provided by Landsat 8 satellite data [83]. In order to account for seasonal changes in NDVI, the average NDVI is computed over a five year period starting January 1$^{\text{st}}$ 2015 and ending on January 1$^{\text{st}}$ 2020. This data is transformed to the discretised target area in a similar manner as to the elevation data, see Figure A.1(c). Values of NDVI range from -1 to +1 and indicate vegetation greenness, where larger values represent richer vegetation [83]. No dynamic (biotic) factors have been implemented, since this is outside the scope of this research. The locations of roads and water sources [76] are used to quantify the euclidean distance between individual patches and its closest road/water source, see Figures A.1(d) and A.1(e). In addition, the euclidean distance between individual patches and the base of operations has also been computed in Figure A.1(f). The euclidean distances are scaled based on the relevant maximum euclidean distance such that normalised euclidean distances are obtained, which simplifies calibration of model parameters. The aforementioned properties can be accessed by the agents and provide input for their cognitive behaviour.



(a) Satellite image of Aloegrove Safari Park, Namibia (Google, n.d.) [108].

(b) Terrain elevation at each individual patch [82].

(c) Five year average NDVI at each individual patch [83].

(d) Normalised euclidean distance between each individual patch and its closest road [76].

(e) Normalised euclidean distance between each individual patch and its closest source of water [76].

(f) Normalised euclidean distance between each individual patch and the base of operations [76].

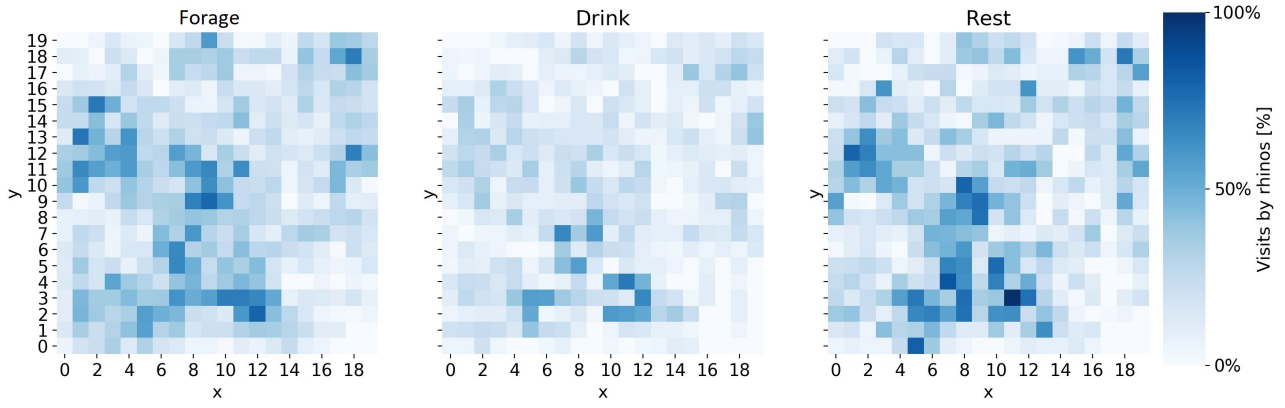
Figure A.1: Visualisation of abiotic factors describing the target area.

A.2 Rhino Agents

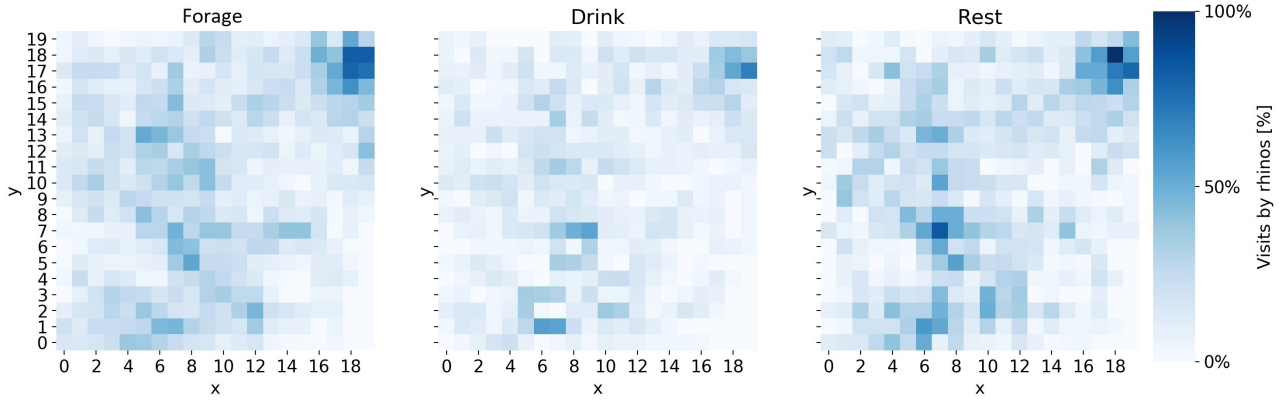
The black rhino agent population consists of 10 rhinos, which is based on estimated black rhino population densities that range from 0.48 to 0.94 rhinos/km^2 [109]. The simulated rhino agents move to an adjacent patch every four timesteps Δt to account for the different speeds of the agents. Black rhino behaviour is described in [80], where distinction is made between male and female rhinos. This research considers homogeneous black rhinos, such that no distinction is made between age, sex or other rhino characteristics. All black rhinos have an equal reward, R_{rhino}, for being killed. Black rhino movement is simulated in the environment of Figure A.1 with a RSF to quantify transition probabilities, see Equation 1. Different behaviour types for a number of time periods throughout the day are specified in Table 1. Calibration of the weights in the RSF model is therefore performed for each of these time periods individually.

A.2.1 Calibration of RSF parameters

Sensitivity analyses were performed for each of sets of the parameters in Table 1 on the interval $w \in [0.0, 1.0]$ with steps of 0.1. Rhino behaviour during the defined time periods was then visualised through heatmaps, as shown in Figure A.2.



(a) Equal RSF parameter values of 0.6.



(b) Calibrated RSF parameter values.

Figure A.2: Heatmaps visualising the relative amount of visits by rhino agents to a patch within the environment.

The objective of this calibration step was to determine a set of RSF weights that results in clustering of rhinos around water sources (see Figure A.1(e)) and areas with relatively high NDVI values (see Figure A.1(c)). A systematic analysis of the influence of various ratios between the weights indicates that too small values for w_{road} and w_{rhino} result in clustering around roads and a relatively high rhino population density. It was found that a value of 0.6 for both these weights results in a more realistic distribution of the rhino population throughout the environment. The analysis then focused on determining weights that characterise the objectives during the individual time periods defined in Table 1. Figure A.2(b) shows how the rhino agents cluster around the three available water sources, while also spreading out over the environment in comparison to Figure A.2(a), where all weights are set at 0.6. Calibration of the rhino RSF model was therefore concluded by setting the weights w_{water}, $w_{vegetation}$ and w_{rest} at 0.8. The emergent behaviour in Figure A.2(b) was found to best represent how black rhinos are a ubiquitously solitary species that show circuitous behaviour around water sources, while remaining relatively stationary during the period of resting and perform habitat selection based on NVDI [80].

A.3 Poacher Agents

The ACOSG model uses the reward for killing a rhino agent to simulate poacher behaviour. The poachers compute this reward, R_χ, from Equation A.2. This also accounts for the amount of normalised euclidean distance, $\|\hat{x}_R\|$, required to travel to the target, by penalising targets that are relatively far away. It also considers the amount of targets at the patch, N_χ, in question, which is determined from a poacher's target map (discussed in subsection 4.4).

$$R_\chi = R_{rhino} \cdot N_\chi \cdot (1 - \|\hat{x}_R\|) \tag{A.2}$$

The penalty, P_χ, is proportional to this reward and adjusted for observed UAV presence to determine a benefit-cost ratio. The most current UAV presence is stored in a poacher's defender map (discussed in subsection 4.4). The objective of the penalty is to stay clear from UAVs and prevent apprehension. It is quantified from Equation A.3, which motivates poachers to increase the normalised euclidean distance $\|\hat{x}_P\|$ to the most nearby UAV.

$$P_\chi = R_\chi \cdot (\|\hat{x}_P(t)\| - \|\hat{x}_P(t + \Delta t)\|) \tag{A.3}$$

The inclination, I, quantifies the experience of poachers gained through (un-)successfully killing rhinos. It is a function of the rewards and penalties collected at certain patches within the target area and is computed from Equation A.4 [10]. The magnitude of the inclination, and therefore the learning rate of poachers, can be influenced with the C_{ASU} parameter. The inclination of a patch σ_j increases as the average of the rewards and penalties obtained at the patch in question, $avg(S^k_{rewards}(\sigma_j))$, increases with respect to the average of rewards and penalties collected in the entire target area E. It is penalised by the variance of these average reward values. The importance of the inclination and therefore its influence on the transition probabilities of the ACOSG model is dependent on elapsed simulation time. This in quantified by considering t as the time that has passed since the simulation started (t_0) and t_{max} as the total simulation time. The weight of the inclination therefore increases as simulation progresses [10].

$$I_{\sigma_j}(t) = \frac{\Delta t^2}{t_{max} \cdot t - t^2} \cdot \frac{1}{C_{ASU}} \cdot \sum_{k \in [t_0, t]} \left(\frac{avg(S^k_{rewards}(\sigma_j)) - avg(S^k_{rewards}(E))}{\mathrm{Var}(avg(S^k_{rewards}(\sigma_j)) - avg(S^k_{rewards}(E)))}\right) \tag{A.4}$$

A.3.1 Choice for a base of operations

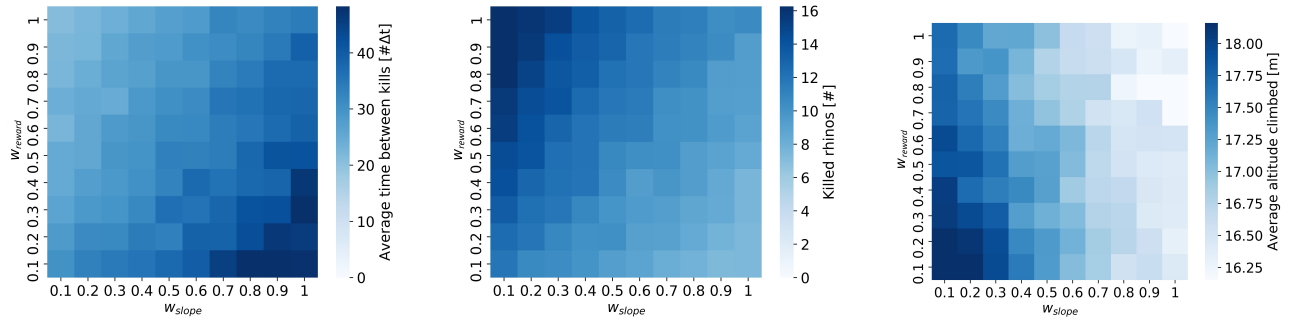
Poacher agents can also choose a base of operation, whereto these agents return at predefined milestones during their missions (discussed in subsection 4.4). This base is chosen upon entry of the target area, with a preference for elevated patches. The probability of choosing a base at a given patch is therefore quantified from Equation A.5 in accordance with the environment's elevation.

$$p(\sigma_j) = \frac{e^{h_{\sigma_j}}}{\sum_{\sigma_k \in E} e^{h_{\sigma_k}}} \tag{A.5}$$

A.3.2 Calibration of ACOSG parameters

The parameters of the ACOSG model are calibrated through sensitivity analyses due to an absence of historic data on poacher activity. The weight parameter C_{ASU} of the inclination is set at 1.0 during these steps. The poaching statistics in [110] are used as guidance to validate whether ACOSG model behaviour is realistic. These statistics indicate that a minor loss of wildlife within the considered target area on an annual basis would be realistic. Albeit realistic, these low rates of wildlife loss do not result in experience needed for the inclination to influence poacher behaviour. The validation process therefore did not correct for the difference in size of the target area in relation to the reported statistics on country-level.

Since the ACOSG model considers the presence of UAV agents, the calibration process of the poacher model parameters is an iterative process. Calibration therefore initially focused on determining parameter values for w_{reward} and w_{slope} by performing a sensitivity analysis and simulating poacher and rhino behaviour for ten simulation days. The performance of the poacher agents is quantified (see Figure A.3) in terms of the average time between kills, number of killed rhinos and the average altitude covered by the poachers. Figures A.3(a) and A.3(b) show that poaching efficiency does changes significantly as a results from varying the w_{reward} and w_{slope} parameters. It also indicates that reducing the w_{slope} parameter has an adverse effect on loss of wildlife and the average time between these kills. This is caused by the location of the elevated areas (see Figure A.1(b)). Figure A.3(c) shows that a low preference for killing rhinos and a low preference to avoid large terrain slopes, results in behaviour that does not succeed at killing rhinos. In addition, this causes poachers to move through the environment, while being ignorant regarding the amount of energy spent on covering rough terrain. These results indicate that the poachers are more effective at killing rhinos when they stay away from large elevation differences and have a large preference for a reward. This is expected since the rhino agents focus on objectives which are not located in elevated locations within the target area. Provided the objective of this research is to analyse poaching activity and provided the relatively small elevation differences in the target area (see Figure A.1(b)), calibration of the poacher parameters was continued with values for w_{reward} and w_{slope} of 1.0 and 0.2 respectively. In addition, these ACOSG parameters result in somewhat similar losses of wildlife as reported for Namibia in 2020 [110], albeit in a shorter time period and within a smaller target area.



(a) Influence on average time between kills.

(b) Influence on number of killed rhinos.

(c) Influence on terrain slope preference.

Figure A.3: Sensitivity analysis of ACOSG parameters w_{reward} and w_{slope}.

The following step of ACOSG model parameter calibration was performed with the presence of 5 UAV agents randomly moving through the target area. The ACOSG parameters found previously resulted in small amount of observations regarding poaching activity. The following therefore analysed poacher behaviour with significantly larger parameter values, which is visualised in Figure A.4. When the weight of the reward is increased significantly, while the weight of the penalty is reduced, the poachers achieve to kill approximately 50% more rhinos in comparison to the results in Figure A.3. A large number of rhino kills is preferred for simulation purposes, since this allows for a larger set of rewards to be collected that in turn stimulates the adaptability of the poachers. Figure A.4 also indicates a nonlinear relation between the reward, the penalty, the weight of the terrain slope and terrain utilisation by the poachers. The objective of this step was to simultaneously stimulate poachers that utilise elevated areas while achieving a loss of relatively large loss of wildlife that is similar in magnitude to the data reported in [110]. Simultaneously, poachers were needed to be apprehended to such an extent that both improvement and deterioration could be shown be implementing the proposed HAPF-BLACOPS model without the ACOSG being a limitation on surveillance performance. Utilisation of elevated areas was required such that type 2 poachers were motivated to go to their base.

The weights for w_{reward}, $w_{penalty}$ and w_{slope} were therefore calibrated at values of 200, 100 and 0.2 respectively. This set of calibrated weights provides a balance between the different objectives, while also limiting the percentage of apprehensions to approximately 0.6.

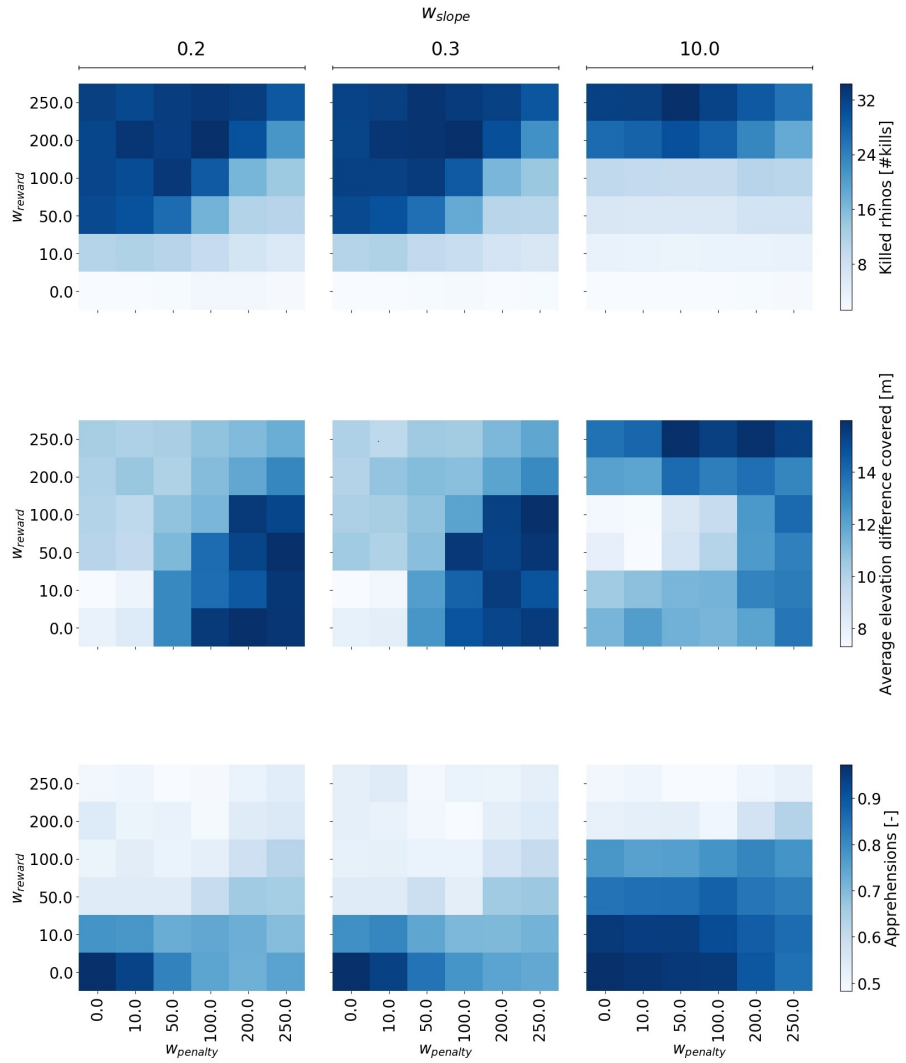
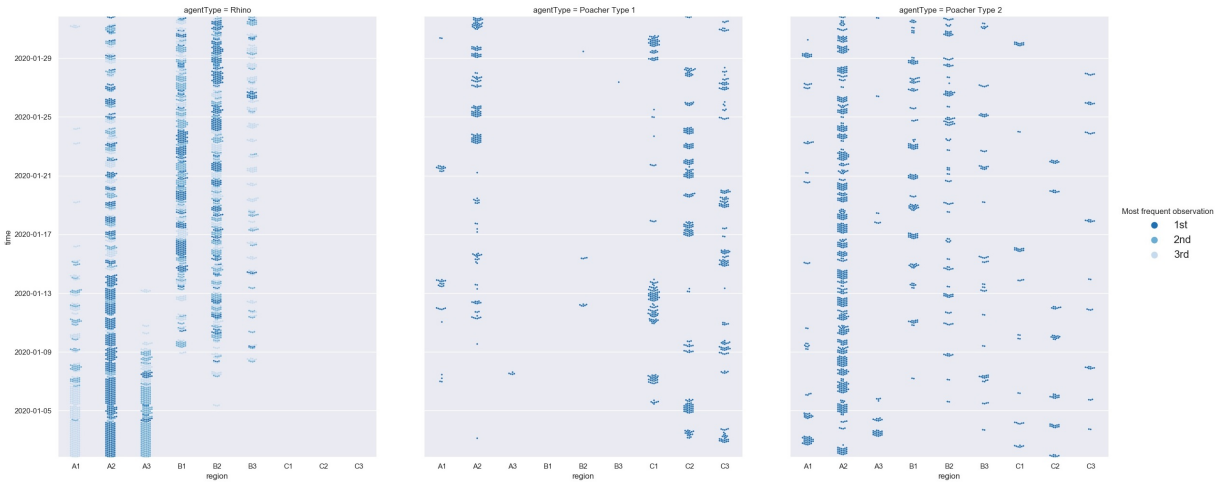
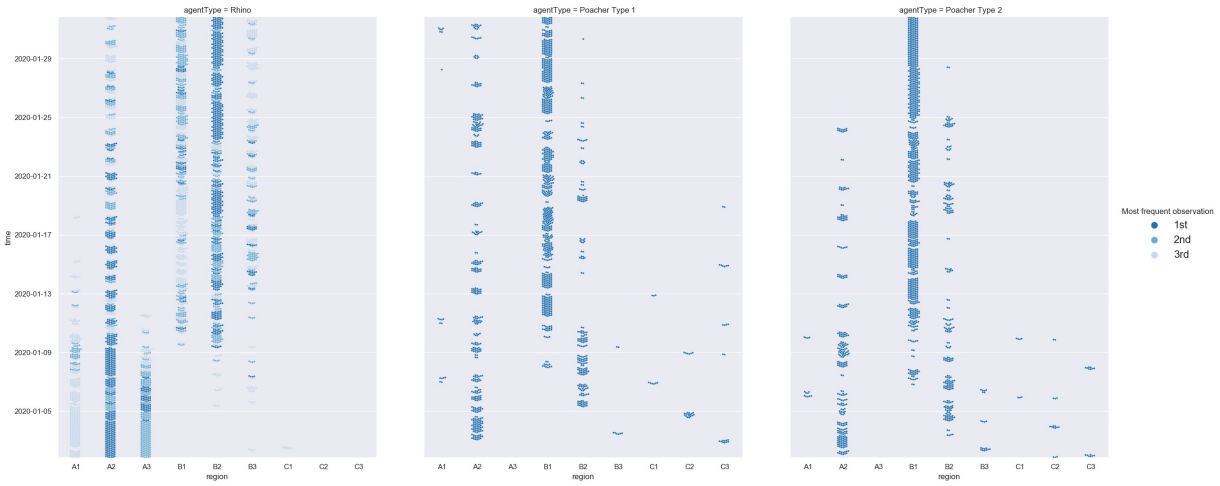


Figure A.4: Sensitivity analysis of ACOSG parameters w_{reward} and w_{slope} and $w_{penalty}$.

Lastly, the spatiotemporal patterns of the poachers were analysed in comparison to those of the rhino agents. All rhino agents are initially given a reward R_{rhino} of 1.0. Figure A.5 shows how the behaviour of both type 1 and 2 poachers is influenced when this parameter is increased by a factor 4. The spatiotemporal patterns are visualised by differentiating between the BN regions in the target area according to Figure 5(b). The data in these swarmplots indicate the number of agents present within a specific BN region at a given time, but only for that region that is observed most frequently when considering all simulation runs. During these calibration steps, 60 simulation runs were performed for each combination. The data further differentiates between the first three n-th most frequently observed regions to provide an indication of the agent distribution throughout both space and time.



(a) Spatiotemporal patterns of agents for R_{rhino} of 1.



(b) Spatiotemporal patterns of agents for R_{rhino} of 4.

Figure A.5: Swarm plots visualising spatiotemporal patterns of most frequently observed poacher positions during simulation.

The observed patterns for both type 1 and type 2 poachers with a rhino reward of 1.0 indicate stochastic movement through the target area, whereas an increase of the reward to a value of 4.0 results in converging spatiotemporal patterns for both poacher types. Type 2 poachers also appear to utilise a larger fraction of the target area, compared to type 1 poachers. This is caused by the different objectives of these poacher types, as discussed in subsection 4.4.10. Poachers of type 2 aim to kill several rhinos, therefore requiring the poacher to move through various regions of the target area. Poachers of type 1 focus on a single rhino instead, which therefore supports the observed spatiotemporal patterns that type 1 poachers are mainly active in the C-regions and the A2-region with a rhino reward of 1.0. This is related to the fact that poachers utilise the roads, which are located in the C-regions (see Figure A.1(d)), as entry points. The A2-region is located in central position in the target area, which indirectly forces the poachers to move through this region on multiple occasions, towards their targets.

The converging patterns converge to the A2- and B-regions, which is where the rhinos are present as well. The increased rhino reward therefore motivates the poachers to become active mainly in those regions that are of interest to the rhinos as well. In addition, the converged spatiotemporal patterns in Figure 5(b) allow for better analysis of possible adaptive behaviour in response to UAV surveillance efforts. We therefore choose to set the R_{rhino} parameter at a value of 4.0, despite the secondary effects this imposes on the emergent behaviour observed in Figure A.4.

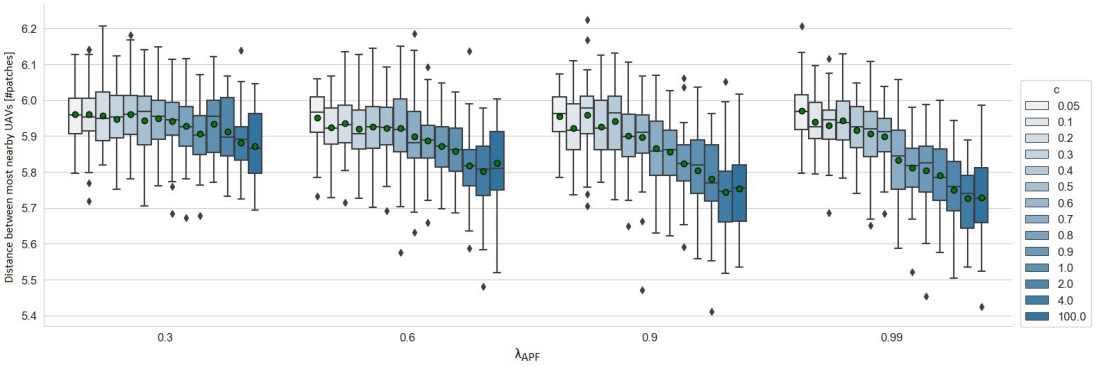
A.4 UAV Agents

The proposed HAPF-BLACOPS surveillance model is built on the current state-of-the-art HAPF-ACO model, which relies on two modules to simulate path planning. The proposed surveillance model inherits the model parameters defined in [40], but requires calibration of several parameters that influence the behaviour of the APF and ACO modules. A systematic approach was taken such that calibration is performed on the individual modules to prevent interactions between these modules. These simulations were performed with a swarm of 5 UAVs performing path planning according to the HAPF-ACO model, while disregarding the range constraint. The simulation time was 10 days and a total of 60 simulation runs were performed for each combination of parameters. The following subsections describe the systematic approach to calibrate the model parameters through sensitivity analysis.

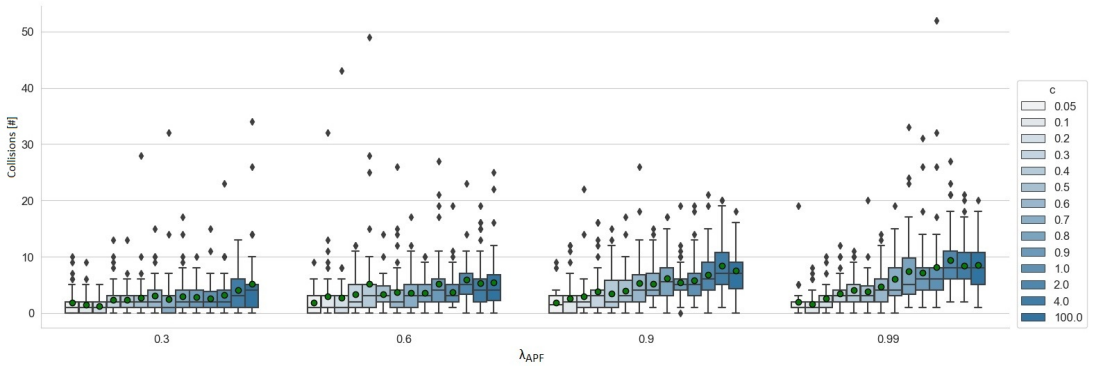
First, the parameters influencing the magnitude of the avoidance forces, computed from Equation 8, are calibrated such that the amount of collisions among UAVs is minimised. No attractions forces are considered at this point and the maximum pheromone concentration is limited to 1.0, such that path planning according to the ACO module is purely stochastic. Second, the parameters influencing the magnitude of the attraction force, computed from Equation 7, are calibrated such that suitable surveillance effectiveness is achieved while disregarding collision avoidance. Finally, the results from the previous two steps were united with the objective of achieving effective surveillance, while preventing a significant increase in the number of collisions. The maximum pheromone concentration of the ACO module was also fine tuned to ensure that the target area is monitored sufficiently. A number of iterations were performed on the steps detailed below, but only the final results are presented.

A.4.1 Calibration of the APF module's avoidance force

The parameters b, c and λ_{APF} influence the path planning based on the APF module, see Equation 8 and Algorithm 2. It was found that the parameter b does influence the collision rate among UAVs, despite its influence on the magnitude of the avoidance forces. This results from Algorithm 2, which considers only the ratio of avoidance force magnitudes to determine whether the UAV adheres to the APF module. A value of 2.0 for b was used during the calibration steps, unless stated otherwise. Figure A.6 therefore shows only the influence of the c and λ_{APF} parameters on the distribution of the UAV swarm and the amount of collisions.



(a) Influence on distance between UAVs.



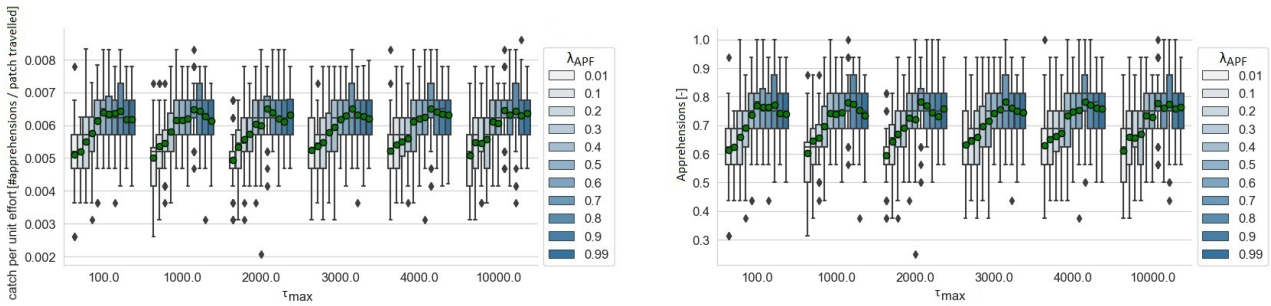
(b) Influence on amount of UAV collisions.

Figure A.6: Sensitivity analysis of APF module parameters c and λ_{APF}.

The c parameter influences the rate of change of the avoidance force's magnitude and therefore affects how soon collision avoidance becomes important. The choice for utilisation of the APF module increases when λ_{APF} is given a small magnitude. A large range of parameter values were therefore analysed. The results in Figure A.6(a) show that c has the strongest effects on the distance between the UAVs. In addition, this results in Figure A.6(b) show that there is an inverse relation between the magnitude of c and the amount of collisions, as expected from a reduction in smaller distances between the UAVs for large values of c. The data shows that the impact of its magnitude does not strongly affect the amount of collisions, although smaller values for c are preferred to avoid collisions.

A.4.2 Calibration of the APF module's attraction force

Calibration of parameters affecting the attraction force and its resulting influence on APF path planning was analysed by disregarding collision avoidance. The magnitude of the b parameter was therefore set to 0.0 during this calibration step. Additionally, the magnitude of the ACOSG parameter $w_{defender}$ was set to 0.0 to circumvent the effects from poacher-UAV interactions on surveillance effectiveness. The magnitude of the attraction force depends on the distance between the UAV and the target, the UAV's TPM and the attraction force coefficient k_{att}. Similarly to the mechanism regarding collisions avoidance, the magnitude of the attraction force does not influence APF behaviour when the attraction and avoidance forces are not united. The data in Figure A.7 therefore only shows the influence of the λ_{APF} and τ_{max} parameters on surveillance performance. The parameter τ_{max} is considered to analyse how the ACO module and therefore exploration of the target area interacts with the attraction forces and surveillance effectiveness.



(a) Influence on surveillance efficiency.

(b) Influence on surveillance effectiveness.

Figure A.7: Sensitivity analysis of APF module parameters τ_{max} and λ_{APF}.

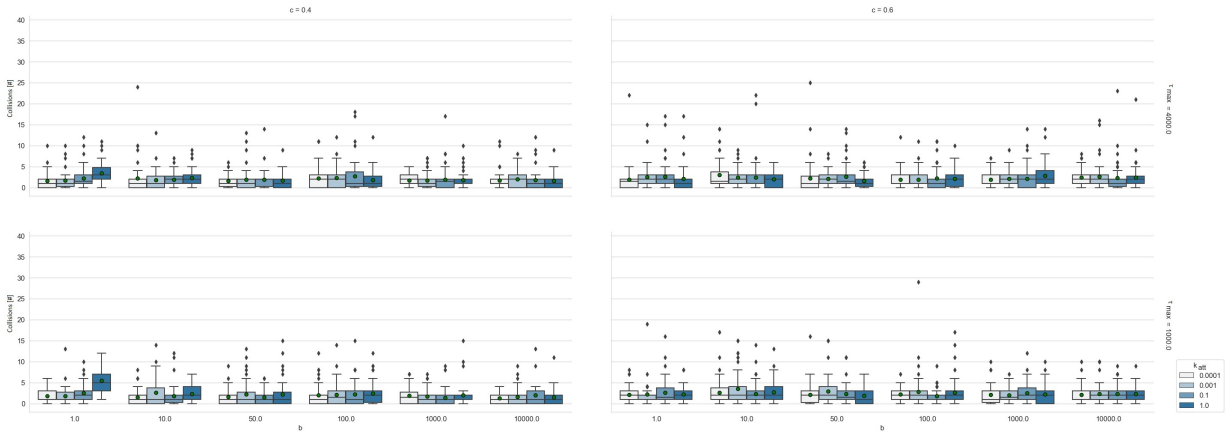
The results in Figures A.7(a) and A.7(b) indicate a fairly parabolic relation between λ_{APF} and surveillance performance, whereas a linear relation was found in Figure A.6. In order to effectively and efficiently apprehend poachers, while also minimising the amount of UAV collisions, it appears an optimum can be deduced from these results. Collision avoidance requires a magnitude for λ_{APF} as small as 0.3, whereas surveillance effectiveness and efficiency is optimal for values between 0.4 and 0.7. However, surveillance effectiveness reduces relatively quickly when the magnitude of λ_{APF} is reduced below 0.5. The magnitude of λ_{APF} was therefore set to 0.6.

The influence of the maximum pheromone concentration, and therefore the influence of exploration through the ACO module, does not significantly affect surveillance performance. Smaller values for τ_{max} do imply that the pheromone concentrations more easily reach their maximum ceiling, which reduces prioritisation of surveillance efforts. In addition, it was found that the pheromone concentrations do not reach the lower bound as a result from the update constants inherited from [40]. The ACO module's update mechanism was therefore found to be balanced.

A.4.3 Calibration of the APF and ACO modules united

Using the calibrated λ_{APF} parameter, the APF's attraction and avoidance forces are united with the ACO module. This step is performed to ensure that neither force fully counteracts the influence of the other. Both rhino and poacher behaviour is simulated in this step as well, although poachers disregard the presence of UAVs. Figure A.8 present the surveillance performance for a range of parameter values of b, c, k_{att} and τ_{max} determined from previous calibration steps.

(a) Influence on surveillance efficiency.



(b) Influence on amount of UAV collisions.

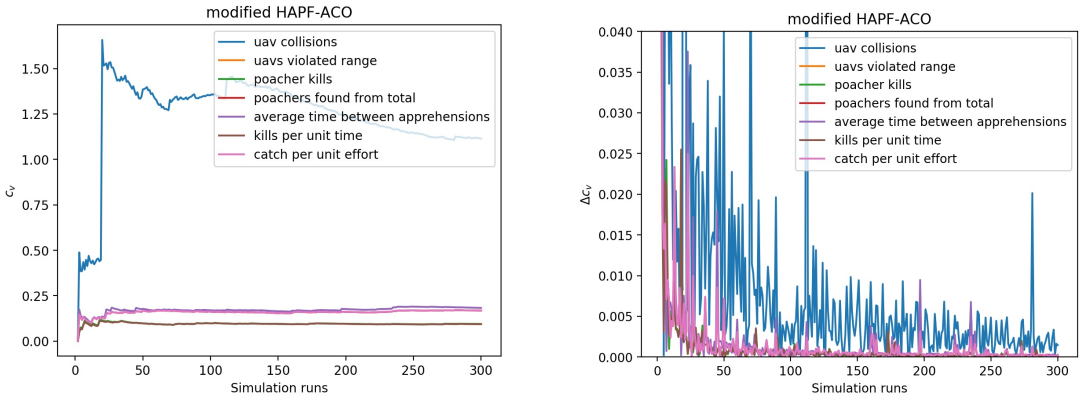
Figure A.8: Sensitivity analysis of APF and ACO module parameters on surveillance performance.

The parameters b and k_{att} influence the prioritisation between avoidance and attraction forces, since these parameters determine the magnitude of these forces. The data in Figure A.8(a) show that these magnitude do not affect surveillance efficiency. Neither is surveillance efficiency affected by the upper bound on the pheromone concentrations. Figure A.8(b) indicates that k_{att} does influence the amount of UAV collisions for relatively small values of c. This indicates that the rate of change of the avoidance force is more important than the magnitude. The ACO module does not appear to affect the rate of collisions. The magnitude of k_{att} was therefore calibrated at 0.001, the magnitude of b was set at 1.0, the magnitude of c at 0,4 and the magnitude of τ_{max} at 4000.0 in order to minimise the amount of UAV collisions, prevent the ACO of limiting prioritisation efforts, without affecting surveillance effectiveness.

A.5 Analysis of ABM output variability

Stochastic model dynamics affect model output. A fixed seed was therefore used for the random number generator and analysis of model output was performed using the coefficient of variation, c_v, to balance output variability with computation costs [104].

The stability of KPIs describing model performance is visualised as a function of the sample size (the number of simulation runs) in Figure A.9. The number of collisions among UAVs is the sole KPI indicating a relatively large coefficient of variation as the sample size increases. Considering the priority of the proposed model that lies with apprehending poachers and considering that the change in magnitude of the coefficient of variation for the remaining KPIs reduces below 0.02 at a sample size of 50, we used a sample size of 50 in the performed experiments.

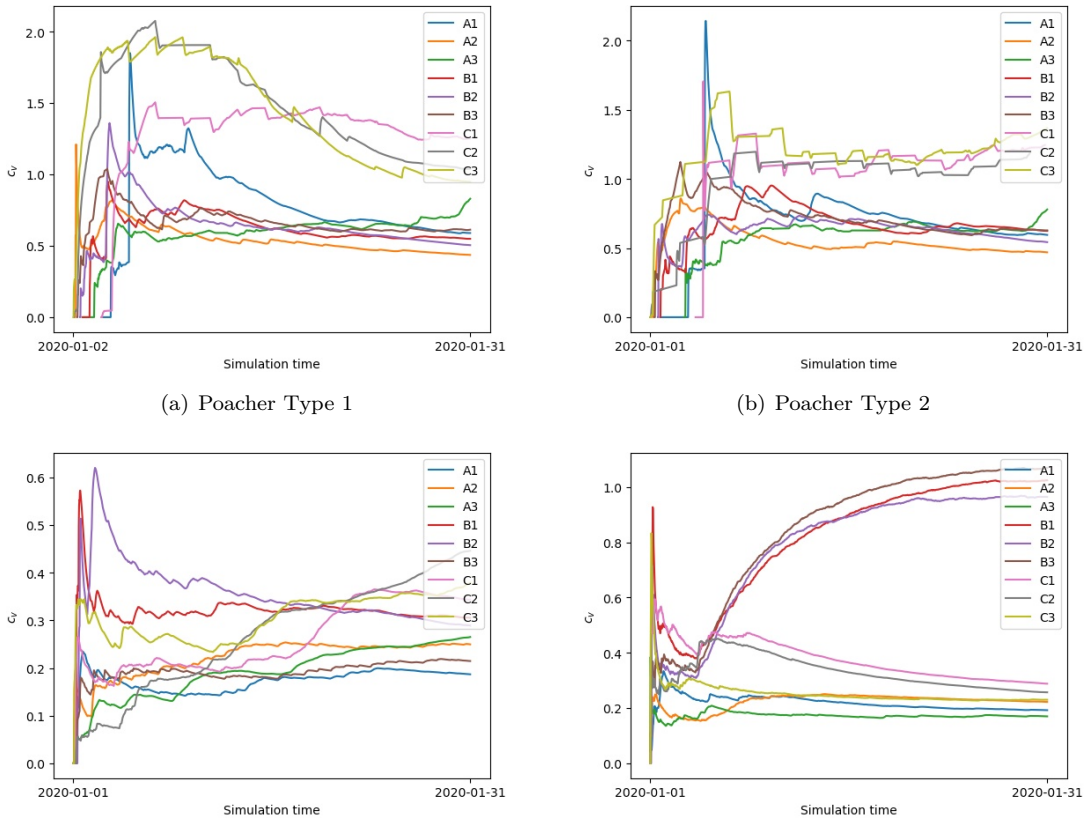


(a) Coefficient of variation.

(b) Change in magnitude of coefficient of variation.

Figure A.9: The coefficient of variation for the HAPF-BLACOPS model as function of sample size.

In addition, Figure A.10 visualises the coefficient of variation of the spatiotemporal patterns of the agents in the ABM. It indicates how the preference of individual agent types to visit the BN regions, defined in Figure 5, fluctuates throughout duration of the simulation runs. The simulation duration, t_{max}, is set at 20 days based on the change in magnitude of the coefficient of variation, which reduces below 0.01 after 14 days. This added simulation time is to allow for adaptive behaviour to emerge in the experiments.



(a) Poacher Type 1

(b) Poacher Type 2

(c) Rhino

(d) UAV

Figure A.10: Coefficient of variation for the ABM based on the spatiotemporal patterns of the agents.

B

Specification of Bayesian Networks

This appendix provides details regarding (A)BNs analysed through simulations. First, the random variables describing the DAGs are discussed, after which the DAGs and their resepctive conditional probability tables of the (A)BNs are specified. The conditional probability tables of the BNs are not provided due to their complexity. This appendix is concluded with an analysis of the scores, quantified by the BDeu, BIC and K2 score functions. These are used to compare how well the different DAGs fit the training data. Not all of the random variables are used for each BN. The score functions were used to determine the set of random variables that represents the training data best.

B.1 Definition of BN nodes

The training data upon which structure and parameter learning is performed is obtained from 100 simulations with the HAPF-BLACOPS model (without utilising a BN), during which observations regarding poacher activity are collected by the UAVs. These observations consist of a number of random variables, represented by nodes in a DAG, that are found in literature [97] and selected based on their data availability. These include the presence of rhinos, poachers and UAVs. These three random variables are discretised and assumed to be dichotomous. Additionally, the BN region (as specified in Figure 5(b)) and the time period defined in 1 are discrete random variables that are used by the UAVs. Abiotic factors including the presence of water sources and roads are considered by computing the fraction of their surface as a function of the surface of a specific BN region. Other abiotic factors that are part of these observations are the NDVI and terrain elevation defined in Appendix A. Finally, based on collected data regarding poacher and UAV presence, the UAVs compute the frequency at which a BN region is monitored (the patrol effort) and the frequency at which rhinos are being killed in a BN region (the hit rate). The following provides de DAGs for various (A)BNs.

B.2 Specification of Artificial Bayesian Networks

This section specifies the DAGs and the conditional probability tables belonging to the ABNs that are defined from expert knowledge regarding UAV, poacher and rhino behaviour.

B.2.1 Specification of ABN1

The DAG and conditional probability tables for ABN1 are defined in Figure B.1 and Tables B.1, B.2 and B.3 respectively.

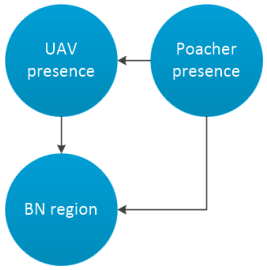


Figure B.1: The DAG of ABN1.

Table B.1: Conditional probability table for ABN1 node Poacher presence.

Poacher presence	**True**	**False**
	0.9	0.1

Table B.2: Conditional probability table for ABN1 node UAV presence.

Poacher presence	**UAV presence (True)**	**UAV presence (False)**
True	0.2	0.8
False	0.4	0.6

Table B.3: Conditional probability table for ABN1 node BN region.

UAV presence	**Poacher presence**	**A1**	**A2**	**A3**	**B1**	**B2**	**B3**	**C1**	**C2**	**C3**
True	True	0.05	0.1	0.05	0.1	0.3	0.2	0.05	0.1	0.05
True	False	0.0125	0.1	0.0125	0.15	0.35	0.25	0.0125	0.1	0.0125
False	True	$\frac{0.4}{6}$	$\frac{0.4}{6}$	$\frac{0.4}{6}$	0.1	0.3	0.2	$\frac{0.4}{6}$	$\frac{0.4}{6}$	$\frac{0.4}{6}$
False	False	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$

B.2.2 Specification of ABN2

The DAG and conditional probability tables for ABN2 are defined in Figure B.2 and Tables B.4, B.5 and B.6 respectively.

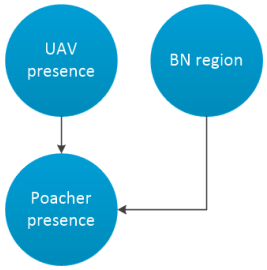


Figure B.2: The DAG of ABN2.

Table B.4: Conditional probability table for ABN2 node BN region.

BN region	A1	A2	A3	B1	B2	B3	C1	C2	C3
	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$

Table B.5: Conditional probability table for ABN2 node UAV presence.

UAV presence	True	False
	0.2	0.8

Table B.6: Conditional probability table for ABN2 node Poacher presence.

BN region	UAV presence	Poacher presence (True)	Poacher presence (False)
A1	True	0.05	0.95
A2	True	0.1	0.9
A3	True	0.1	0.9
B1	True	0.2	0.8
B2	True	0.05	0.95
B3	True	0.05	0.95
C1	True	0.9	0.1
C2	True	0.95	0.05
C3	True	0.6	0.4
A1	False	0.6	0.4
A2	False	0.6	0.4
A3	False	0.6	0.4
B1	False	0.3	0.7
B2	False	0.3	0.7
B3	False	0.3	0.7
C1	False	0.3	0.7
C2	False	0.3	0.7
C3	False	0.3	0.7

B.2.3 Specification of ABN3

The DAG and conditional probability tables for ABN3 are defined in Figure B.3 and Tables B.7, B.8, B.9 and B.10 respectively.

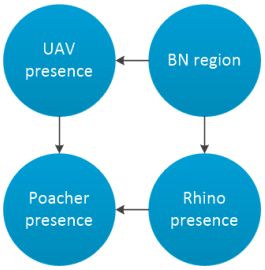


Figure B.3: The DAG of ABN3.

Table B.7: Conditional probability table for ABN3 node BN region.

BN region	A1	A2	A3	B1	B2	B3	C1	C2	C3
	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$

Table B.8: Conditional probability table for ABN3 node UAV presence.

BN region	UAV presence (True)	UAV presence (False)
A1	0.6	0.4
A2	0.4	0.6
A3	0.6	0.4
B1	0.5	0.5
B2	0.5	0.5
B3	0.5	0.5
C1	0.4	0.6
C2	0.6	0.4
C3	0.4	0.6

Table B.9: Conditional probability table for ABN3 node Rhino presence.

BN region	Rhino presence (True)	Rhino presence (False)
A1	0.3	0.7
A2	0.7	0.3
A3	0.2	0.8
B1	0.8	0.2
B2	0.8	0.2
B3	0.6	0.4
C1	0.1	0.9
C2	0.1	0.9
C3	0.1	0.9

Table B.10: Conditional probability table for ABN3 node Poacher presence.

UAV presence	Rhino presence	Poacher presence (True)	Poacher presence (False)
True	True	0.75	0.25
True	False	0.95	0.05
False	True	0.1	0.9
False	False	0.9	0.1

B.2.4 Specification of ABN4

The DAG and conditional probability tables for ABN4 are defined in Figure B.4 and Tables B.11 and B.12 respectively.

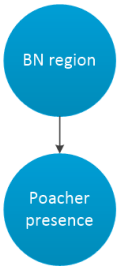


Figure B.4: The DAG of ABN4.

Table B.11: Conditional probability table for ABN4 node BN region.

BN region	A1	A2	A3	B1	B2	B3	C1	C2	C3
	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$

Table B.12: Conditional probability table for ABN4 node Poacher presence.

BN region	Poacher presence (True)	Poacher presence (False)
A1	0.1	0.9
A2	0.8	0.2
A3	0.1	0.9
B1	0.99	0.01
B2	0.9	0.1
B3	0.8	0.2
C1	0.1	0.9
C2	0.1	0.9
C3	0.1	0.9

B.3 Specification of Bayesian Networks obtained from training data

This section specifies the DAGs belonging to the BNs that are obtained from structure and parameter learning. Hill-Climb Search was used to specify these DAGs, since this structure learning method resulted in DAGs that best described the relations between the ABM interactions. Their respective conditional probability tables are not listed due to the complexity of these DAGs. The collected training data was combined with Bayesian Estimation to obtain these conditional probability tables.

B.3.1 Specification of BN5

The DAG of BN5 is defined in Figure B.5.

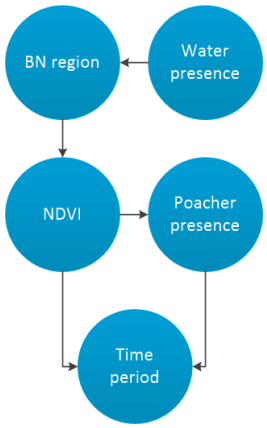


Figure B.5: The DAG of BN5.

B.3.2 Specification of BN6

The DAG of BN6 is defined in Figure B.6.

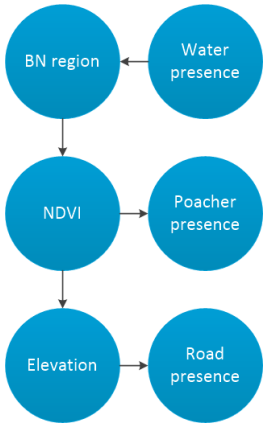


Figure B.6: The DAG of BN6.

B.3.3 Specification of BN7

The DAG of BN7 is defined in Figure B.7.

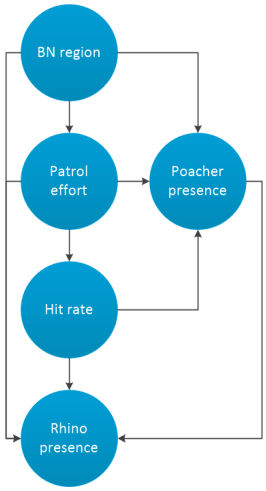


Figure B.7: The DAG of BN7.

B.3.4 Specification of BN8

The DAG of BN8 is defined in Figure B.8.

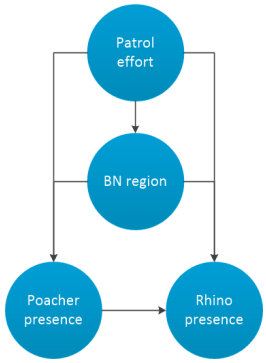


Figure B.8: The DAG of BN8.

B.3.5 Specification of BN9

The DAG of BN9 is defined in Figure B.9.

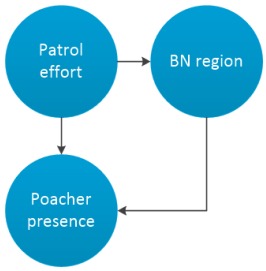


Figure B.9: The DAG of BN9.

B.3.6 Specification of BN-Expert

The DAG of BN-Expert is defined in Figure B.10. This DAG is obtained from a combination of structure learning results and expert judgement in order to determine whether human knowledge regarding the model interactions can aid surveillance performance.

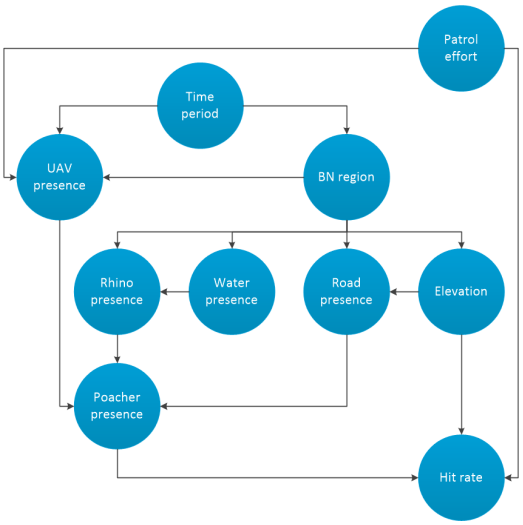


Figure B.10: The DAG of BN-Expert.

B.4 Scores of BN DAGs

Table B.13 summarises how well the previously presented DAGs fit the training data. The BDeu, BIC and K2 scoring functions were used to quantify this goodness of fit and to analyse whether the scoring method influenced these results. The data in Table B.13 shows that these different scoring functions result in the same score for a given DAG.

Some of the random variables considered in the presented DAGs represent continuous data. Discretisation is required, but it can result in a loss of information [111] and it can therefore affect how well the DAGs fit the training data [98]. Several discretisation techniques were therefore analysed and compared with these BDeu, BIC and K2 scoring functions. Both supervised and unsupervised discretisation techniques were analysed. The CAIM, MDLP and Freedman-Diaconis rule were applied to discretise the training data for their performance and efficiency characteristics [111]. The CAIM and MDLP discretisation techniques were unable to determine the optimal set of bins and their scoring results are therefore excluded from Table B.13. The Freedman-Diaconis rule was also applied since it uses uniform bin widths and is relatively simple to implement [112]. This method resulted in over 100 bins per random variable and as a result was excluded from Table B.13 as well, since this level of discretisation requires a more sufficient amount of training data to describe the continuous data without loss of information [111].

Discretisation of the continuous data was therefore performed by predefining the number of uniform bins. The results for 2, 4 and 6 bins are shown in Table B.13. The magnitude of the obtained scores is such that none of the DAGs appears to significantly outperform the others. Interestingly, an increase in the number of bins to discretise the training data only improves the score for BN6, whereas the opposite trend is seen for BN7, BN8 and BN9. The main difference between these BNs is that BN6 mainly considers abiotic factors to predict poacher activity, whereas the remaining BNs do not. This trend does not change when unapprehended poachers data is included in the training data. This suggests that the additional observations from unapprehended poachers do not provide the BN with unaccounted for insights regarding poacher behaviour. Instead, the magnitude of the scores is reduced as a result from an increase in sample size of the training data. The scores also indicate that the most simplistic DAG, which is ABN4, results in the best fit to the training data. These BNs do not differentiate between the poacher types, despite literature suggesting to differentiate for different types of criminal activity in order to accurately predict their behaviour [95]. This choice was made since the calibration results of the ACOSG model, shown in Appendix A, indicate that the spatiotemporal patterns are similar for the two poacher types being considered. Finally, the choice was made to discretise the continuous data in 4 uniform bins in order to be able to differentiate better between various circumstances during simulation of the proposed surveillance model.

Table B.13: Goodness of fit of DAGs, quantified with scoring functions BDeu, BIC and K2.

(A)BN model	Bins (If applicable)	Unapprehended poachers	Scoring function	Score
ABN1	-	-	BDeu	-2360170.9
ABN1	-	-	BIC	-2360226.4
ABN1	-	-	K2	-2360145.6
ABN2	-	-	BDeu	-2528034.0
ABN2	-	-	BIC	-2528082.1
ABN2	-	-	K2	-2528089.1
ABN3	-	-	BDeu	-2815478.7
ABN3	-	-	BIC	-2815491.7
ABN3	-	-	K2	-2815473.8
ABN4	-	-	BDeu	-2035022.6
ABN4	-	-	BIC	-2035039.0
ABN4	-	-	K2	-2035024.3
BN5	2	Exclude	BDeu	-2979641.6
BN5	2	Exclude	BIC	-2979659.4
BN5	2	Exclude	K2	-2979718.7
BN5	4	Exclude	BDeu	-2972398.2
BN5	4	Exclude	BIC	-2972555.8
BN5	4	Exclude	K2	-2972687.6
BN5	6	Exclude	BDeu	-2945593.9
BN5	6	Exclude	BIC	-2945809.5
BN5	6	Exclude	K2	-2945960.4
BN5	2	Include	BDeu	-5867895.1
BN5	2	Include	BIC	-5867912.2
BN5	2	Include	K2	-5867978.6
BN5	4	Include	BDeu	-5777693.3
BN5	4	Include	BIC	-5777755.6
BN5	4	Include	K2	-5777851.0
BN5	6	Include	BDeu	-5777693.3
BN5	6	Include	BIC	-5777755.6
BN5	6	Include	K2	-5777851.0
BN6	2	Exclude	BDeu	-2924434.4
BN6	2	Exclude	BIC	-2924443.6
BN6	2	Exclude	K2	-2924512.0
BN6	4	Exclude	BDeu	-2760588.2
BN6	4	Exclude	BIC	-2760693.9
BN6	4	Exclude	K2	-2760890.9
BN6	6	Exclude	BDeu	-2523178.3
BN6	6	Exclude	BIC	-2523377.8
BN6	6	Exclude	K2	-2523658.7
BN6	2	Include	BDeu	-4140756.6
BN6	2	Include	BIC	-4140669.7
BN6	2	Include	K2	-4140772.0
BN6	4	Include	BDeu	-4051685.3
BN6	4	Include	BIC	-4051649.5
BN6	4	Include	K2	-4051816.5
BN6	6	Include	BDeu	-4051685.3
BN6	6	Include	BIC	-4051649.5
BN6	6	Include	K2	-4051816.5

Table B.13: Goodness of fit of DAGs, quantified with scoring functions BDeu, BIC and K2.

(A)BN model	Bins (If applicable)	Unapprehended poachers	Scoring function	Score
BN7	2	Exclude	BDeu	-2704029.6
BN7	2	Exclude	BIC	-2704452.1
BN7	2	Exclude	K2	-2703963.9
BN7	4	Exclude	BDeu	-3061502.7
BN7	4	Exclude	BIC	-3063862.2
BN7	4	Exclude	K2	-3061424.6
BN7	6	Exclude	BDeu	-3394910.1
BN7	6	Exclude	BIC	-3400644.5
BN7	6	Exclude	K2	-3394824.7
BN7	2	Include	BDeu	-5167838.8
BN7	2	Include	BIC	-5168277.8
BN7	2	Include	K2	-5167797.3
BN7	4	Include	BDeu	-5824401.0
BN7	4	Include	BIC	-5826839.7
BN7	4	Include	K2	-5824365.7
BN7	6	Include	BDeu	-6465229.2
BN7	6	Include	BIC	-6471203.5
BN7	6	Include	K2	-6465214.0
BN8	2	Exclude	BDeu	-2702083.4
BN8	2	Exclude	BIC	-2702202.7
BN8	2	Exclude	K2	-2702054.2
BN8	4	Exclude	BDeu	-3013855.8
BN8	4	Exclude	BIC	-3014153.3
BN8	4	Exclude	K2	-3013777.8
BN8	6	Exclude	BDeu	-3284946.4
BN8	6	Exclude	BIC	-3285406.5
BN8	6	Exclude	K2	-3284791.0
BN8	2	Include	BDeu	-5146304.1
BN8	2	Include	BIC	-5146427.5
BN8	2	Include	K2	-5146277.9
BN8	4	Include	BDeu	-5708132.4
BN8	4	Include	BIC	-5708434.5
BN8	4	Include	K2	-5708055.8
BN8	6	Include	BDeu	-6205615.3
BN8	6	Include	BIC	-6206084.7
BN8	6	Include	K2	-6205458.7
BN9	2	Exclude	BDeu	-2326952.1
BN9	2	Exclude	BIC	-2327001.7
BN9	2	Exclude	K2	-2326956.2
BN9	4	Exclude	BDeu	-2657494.5
BN9	4	Exclude	BIC	-2657632.6
BN9	4	Exclude	K2	-2657508.6
BN9	6	Exclude	BDeu	-2931627.7
BN9	6	Exclude	BIC	-2931850.4
BN9	6	Exclude	K2	-2931637.7
BN9	2	Include	BDeu	-4487483.7
BN9	2	Include	BIC	-4487534.2
BN9	2	Include	K2	-4487489.2
BN9	4	Include	BDeu	-5090356.5
BN9	4	Include	BIC	-5090494.4
BN9	4	Include	K2	-5090368.2
BN9	6	Include	BDeu	-5595164.1
BN9	6	Include	BIC	-5595389.8
BN9	6	Include	K2	-5595173.2

C

Default model parameters

Most of the model input parameters are inherited from the HAPF-ACO model proposed in [40]. Table C.1 summarises the parameters used for the various experiments in this research, unless explicitly specified otherwise.

Table C.1: Default model parameter values.

Parameter	Definition	Default value	Unit
Δt	Each model time step simulates an amount Δt of real world time	30	mins
N_{rhino}	Rhino population size within the environment	10.0	-
w_{road}	Weight of the euclidean distance to the closest road	0.6	-
w_{rhino}	Weight of the presence of other rhino agents	0.6	-
w_{water}	Weight of the euclidean distance to the closest water source	0.8	-
$w_{vegetation}$	Weight of the vegetation index quantified by the NDVI	0.8	-
w_{rest}	Weight of not moving to a different location	0.8	-
$w_{defender}$	Weight of the presence of UAV agents	100.0	-
w_{reward}	Weight of the presence of rhino agents	200.0	-
w_{slope}	Weight of the terrain slope	0.2	-
R_{rhino}	Reward of a rhino agents	4.0	-
ϵ	Collaboration reward coefficient	1.1	-
δ_0	Variance of normal distribution	5.0	-
C_{ASU}	Weight parameter of the Inclination	1.0	-
ϕ_{max}	Maximum turning angle	45.0	°
R	Maximum radius of detection	3.0	km
R_{attack}	Maximum radius of attack	0.0	km
R_{sen}	Maximum radius of communication	3.0	km
k	Adjusting rate	1.0	-
η_{topo}	Expected number of neighbour UAV agents	5.0	-
L_{max}	Maximum range	330.0	km
P_D	Sensor detection probability during day time	0.9	-
P_F	Sensor false detection probability during day time	0.1	-
$P_{D_{night}}$	Sensor detection probability during night time	0.6	-
$P_{F_{night}}$	Sensor false detection probability during night time	0.4	-
Recharge rate	Recharge rate of UAV agent battery	20	$\%/\Delta t$
d_{in}	Maximum radius of detection by sensor without performance degradation	0.5	km
d_{out}	Maximum radius of successful detection by sensor	3	km
τ_{min}	Minimum pheromone concentration	0.0	-
τ_{max}	Maximum pheromone concentration	4000.0	-
F	Environmental uncertainty factor	0.02	-
$\Delta\tau_{g_0}$	Global update rule pheromone concentration	15	-
$\Delta\tau_{l_0}$	Local update rule pheromone concentration	0.1	-
k_{att}	Attraction force coefficient	0.001	-
b	Avoidance force magnitude coefficient	1.0	-
c	Avoidance force rate of change coefficient	0.4	-
τ	Attenuation coefficient	0.5	-
d_{max}	Maximum operating distance of avoidance force	3	km
d_{min}	Minimum operating distance of avoidance force	0.5	km
λ	Environmental perception factor	0.6	-
δ_e	Wiener random process parameter	2	-
α	Importance factor of pheromone information	1	-
β	Importance factor of heuristic information	3	-
q_2	Threshold for mission selector	0.5	-
N_t	Iteration threshold for selecting exception transition rule	10	Δt
recent coverage period	Time period used to compute the heuristic information	24	hours

D

Specification of KPIs

This appendix provides details on the KPIs that are used to quantify the performance of the surveillance model.

Catch per unit effort
The surveillance efficiency of the swarm that coordinates surveillance is computed from the cumulative distance travelled by the swarm and the cumulative number of apprehensions, according to Equation D.1 [22]. The poachers' kills per unit effort is computed in a similar manner from the loss of wildlife and the cumulative distance covered by the poachers.

$$\text{Catch per unit effort}(t) = \frac{N_{\text{apprehensions}}(t)}{\sum_{k \in S_{UAVs}} d_k(t)} \tag{D.1}$$

Coverage fairness
The prioritisation efforts of the UAV swarm are quantified through the coverage fairness. This KPI indicates the amount to which the efforts of monitoring each patch in the target area deviates from a uniform distribution of surveillance efforts, as specified in Equation D.2 [113, 114]. It computes the standard deviation from the set $S_{\text{visits}}(t)$, which describes the number of visits by each UAV to each patch in the target area up to time t.

$$\text{Coverage fairness}(t) = \sigma(S_{visits}(t)) \tag{D.2}$$

Global average idleness (GAI)
Continuous surveillance is commonly optimised by minimising the global average idleness, which is computed from Equation D.3. It quantifies the average time lag between two consecutive visits to the patches σ_j in the target area E during the time period $[t_0, t]$. It is based on a summation of the average idleness $AIdl(\sigma_j, t)$ for all individual patches and dividing by the number of locations to obtain an overall average [12].

$$GAI(t) = \frac{\sum_{\sigma_i \in E} AIdl(\sigma_i, t)}{N_{\text{patches}}} \tag{D.3}$$

Wildlife loss distribution
The spatial distribution of killed rhinos is a value between 0 and 1 and is quantified from Equation D.4. The larger the magnitude of this ratio is, the larger the percentage of the target area is wherein rhinos are being killed by the poachers.

$$\text{Wildlife loss distribution}(t) = \frac{\sum_{\sigma_i \in E} k_{\sigma_i(t)}}{N_{\text{patches}}} \quad , \quad k = 1 \text{ if a rhino has been killed by a poacher at patch } \sigma_i \text{ at time } t \tag{D.4}$$

Reward distribution
The distribution of the rewards is a value between 0 and 1 which quantifies the proportion of the target area where no rewards and penalties have been collected. It is computed from Equation D.5. Larger values indicate that in a relatively small percentage of the target area apprehensions/kills are not recorded.

$$\text{Reward distribution}(t) = 1 - \frac{\sum_{\sigma_i \in E} k_{\sigma_i(t)}}{N_{\text{patches}}} \quad , \quad k = 1 \text{ if a reward/penalty has been collected at patch } \sigma_i \text{ at time } t \tag{D.5}$$

Coverage
The coverage indicates the percentage of the target area E that has been monitored by the UAV swarm during the time period $[t_0, t]$, see Equation D.6 [113, 40, 114].

$$\text{Coverage}(t) = \frac{\sum_{\sigma_i \in E} k_{\sigma_i(t)}}{N_{\text{patches}}} \quad , \quad k = 1 \text{ if patch } \sigma_i \text{ has been visited by a UAV at time } t \tag{D.6}$$

This KPI is also quantified by only considering the visits within a time period such as 24 hours [115].

Cumulative time since last communication
For each UAV in the swarm, S_{UAVs}, the most recent time at which synchronisation with another swarm member, $t_{sync}^{a_{UAV}}$, is performed, is stored. The cumulative time since last communication is quantified from Equation D.7 [115] and is a measurement for how current the local information at each swarm member is.

$$\text{Cumulative time since last communication}(t) = \sum_{a_{UAV} \in S_{UAVs}} t - t_{sync}^{a_{UAV}} \tag{D.7}$$

E

Statistical Elaboration

This appendix presents the statistical results describing the significance of the obtained results in Section 5. The KPI data was first tested for normality with the Shapiro-Wilk test and QQ-plots. The KPI data resulting from analysed models was then tested using either the Wilcoxon signed rank test or using the paired t-test for normally distributed data to compare model performance and determine the significance of achieved results. Paired statistical tests were used since the KPI data is obtained from simulation with the same population of agents. A significance level of 5% was used for all statistical tests.

E.1 Results for E1

The data in Table E.1 shows the significance of model output, resulting from the proposed mission selector in Algorithm 2. The results from the Wilcoxon test indicate significantly different model output for the HAPF-BLACOPS model.

Table E.1: Statistical significance of model performance in E1.

KPI	Reference model	Model	Statistical test	Statistic	P-value
UAVs violated range	HAPF-ACO	HAPF-BLACOPS	Wilcoxon	0.0	1.77E-51
Collisions	HAPF-ACO	HAPF-BLACOPS	Wilcoxon	5.0	6.38E-51
Apprehensions	HAPF-ACO	HAPF-BLACOPS	Wilcoxon	44.0	1.35E-50

E.2 Results for E2

The data in Table E.2 shows the significance of model output, resulting from simulating the proposed HAPF-BLACOPS model with different ABNs defined in Appendix B. The pairwise t-tests for ABN1, ABN2, ABN3 and ABN4 indicate that these models do not result in significantly different model performance since the p-values are larger than 0.05.

Table E.2: Statistical significance of model performance in E2.

KPI	Reference model	Model	Statistical test	Statistic	P-value
Coverage fairness	HAPF-ACO	ABN1	Wilcoxon	170.0	0.1986
Catch per unit effort	HAPF-ACO	ABN1	Wilcoxon	132.0	0.0387
Apprehensions	HAPF-ACO	ABN1	Wilcoxon	131.5	0.0377
Coverage fairness	HAPF-ACO	ABN2	Wilcoxon	224.0	0.8612
Catch per unit effort	HAPF-ACO	ABN2	Wilcoxon	140.0	0.0938
Apprehensions	HAPF-ACO	ABN2	Wilcoxon	110.5	0.0987
Coverage fairness	HAPF-ACO	ABN3	Wilcoxon	192.0	0.4048
Catch per unit effort	HAPF-ACO	ABN3	Wilcoxon	164.0	0.1588
Apprehensions	HAPF-ACO	ABN3	Wilcoxon	163.5	0.1558
Coverage fairness	HAPF-ACO	ABN4	Wilcoxon	231.0	0.9754
Catch per unit effort	HAPF-ACO	ABN4	Wilcoxon	177.0	0.2536
Apprehensions	HAPF-ACO	ABN4	Wilcoxon	165.5	0.2607

E.3 Results for E4

The data in Table E.3 shows the significance of model output, resulting from simulating the proposed HAPF-BLACOPS model the results obtained from GSA.

Table E.3: Statistical significance of model performance in E4.

KPI	Reference model	Model	Statistical test	Statistic	P-value
Coverage fairness	HAPF-ACO	ABN3	Wilcoxon	5.0	0.0046
Catch per unit effort	HAPF-ACO	ABN3	Wilcoxon	38.0	0.6002
Apprehensions	HAPF-ACO	ABN3	Wilcoxon	0.0	0.00147
Coverage fairness	HAPF-ACO	ABN4	Wilcoxon	6.0	0.0058
Catch per unit effort	HAPF-ACO	ABN4	t-test	-0.6	0.5681
Apprehensions	HAPF-ACO	ABN4	t-test	7.5	7.73E-06
Coverage fairness	HAPF-ACO	BN5	Wilcoxon	0.0	0.0015
Catch per unit effort	HAPF-ACO	BN5	Wilcoxon	34.0	0.4212
Apprehensions	HAPF-ACO	BN5	Wilcoxon	0.0	0.00146
Coverage fairness	HAPF-ACO	BN6	Wilcoxon	1.0	0.0019
Catch per unit effort	HAPF-ACO	BN6	Wilcoxon	35.0	0.4631
Apprehensions	HAPF-ACO	BN6	Wilcoxon	0.0	0.00147
Coverage fairness	HAPF-ACO	BN7	Wilcoxon	0.0	0.0015
Catch per unit effort	HAPF-ACO	BN7	Wilcoxon	45.0	0.9721
Apprehensions	HAPF-ACO	BN7	Wilcoxon	1.0	0.00286
Coverage fairness	HAPF-ACO	BN8	Wilcoxon	16.0	0.0392
Catch per unit effort	HAPF-ACO	BN8	Wilcoxon	38.0	0.6002
Apprehensions	HAPF-ACO	BN8	Wilcoxon	0.0	0.00147
Coverage fairness	HAPF-ACO	BN9	Wilcoxon	1.0	0.0019
Catch per unit effort	HAPF-ACO	BN9	t-test	-1.0	0.335
Apprehensions	HAPF-ACO	BN9	t-test	8.2	2.86E-06
Coverage fairness	HAPF-ACO	BN-expert	Wilcoxon	5.0	0.0046
Catch per unit effort	HAPF-ACO	BN-expert	t-test	-0.3	0.7384
Apprehensions	HAPF-ACO	BN-expert	t-test	6.0	6.35E-05

E.4 Results for E5

The data in Table E.4 shows the significance of model output, resulting from simulating the proposed HAPF-BLACOPS model with different BN regions.

Table E.4: Statistical significance of model performance in E5.

KPI	Reference model	Model	Statistical test	Statistic	P-value
GAI	9 BN regions	3 BN regions	Wilcoxon	99.0	0.0793
Kills distribution	9 BN regions	3 BN regions	t-test	-0.1	0.9464
Reward distribution	9 BN regions	3 BN regions	t-test	0.1	0.9464
Killed rhinos	9 BN regions	3 BN regions	t-test	-0,3	0,7584

E.5 Results for E6

The data in Table E.5 shows the significance of model output, resulting from simulating the proposed HAPF-BLACOPS model and including observation from unapprehended poachers in the training data.

Table E.5: Statistical significance of model performance in E6.

KPI	Reference model	Model	Statistical test	Statistic	P-value
Apprehensions	Exclude	Include	Wilcoxon	128.5	0.7717
Catch per unit effort	Exclude	Include	Wilcoxon	176.0	0.7547
Coverage fairness	Exclude	Include	Wilcoxon	204.0	0.5578

E.6 Results for E7

The data in Table E.6 shows the significance of model output, resulting from simulating the ACOSG model for various weights of the inclination in the ASU function.

Table E.6: Statistical significance of model performance in E7.

KPI	Reference model	Model	Statistical test	Statistic	P-value
Apprehensions	$C_{ASU} = 0.01$	$C_{ASU} = 1.0$	Wilcoxon	64.5	0.9517
Kills per unit effort	$C_{ASU} = 0.01$	$C_{ASU} = 1.0$	t-test	0.3	0.778
Coverage fairness	$C_{ASU} = 0.01$	$C_{ASU} = 1.0$	Wilcoxon	81.0	0.3703
Apprehensions	$C_{ASU} = 0.01$	$C_{ASU} = 100.0$	Wilcoxon	93.5	0.5683
Kills per unit effort	$C_{ASU} = 0.01$	$C_{ASU} = 100.0$	t-test	-2.5	0.022
Coverage fairness	$C_{ASU} = 0.01$	$C_{ASU} = 100.0$	Wilcoxon	64.0	0.1259

E.7 Results for E8

The data in Table E.8 shows the significance of model output, resulting from simulating various alternatives for the HAPF-BLACOPS model. Table E.7 defines the models shown in Table E.8.

Table E.7: Statistical significance of model performance in E8.

Model	**Abbreviation**
ABN4 (Reference)	M_{Ref}
ABN4 (GSA)	M1
ABN4 (GSA, $N_{UAVs} = 10$)	M2
ABN4 (GSA, MT3, $\beta = 0$)	M3
BN9 (GSA, MT3, $\beta = 0$)	M4
HAPF-BLACOPS (GSA, MT3, $\beta = 0$)	M5

Table E.8: Statistical significance of model performance in E8.

KPI	**Reference model**	**Model**	**Statistical test**	**Statistic**	**P-value**
Coverage	M_{Ref}	M1	Wilcoxon	0.0	0.0001
Coverage	M_{Ref}	M2	Wilcoxon	0.0	0.0001
Coverage	M_{Ref}	M3	Wilcoxon	0.0	0.0000
Coverage	M_{Ref}	M4	Wilcoxon	0.0	0.0000
Coverage	M_{Ref}	M5	Wilcoxon	0.0	0.0000
Coverage fairness	M_{Ref}	M1	Wilcoxon	24.0	0.0025
Coverage fairness	M_{Ref}	M2	Wilcoxon	96.0	0.7369
Coverage fairness	M_{Ref}	M3	Wilcoxon	0.0	0.0000
Coverage fairness	M_{Ref}	M4	Wilcoxon	0.0	0.0000
Coverage fairness	M_{Ref}	M5	Wilcoxon	0.0	0.0000
Coverage (24h)	M_{Ref}	M1	t-test	-8.4	0.0000
Coverage (24h)	M_{Ref}	M2	t-test	-27.1	0.0000
Coverage (24h)	M_{Ref}	M3	Wilcoxon	0	0.0000
Coverage (24h)	M_{Ref}	M4	Wilcoxon	0	0.0000
Coverage (24h)	M_{Ref}	M5	t-test	21.6	0.0000
Collisions	M_{Ref}	M1	Wilcoxon	0.0	0.0000
Collisions	M_{Ref}	M2	Wilcoxon	0.0	0.0000
Collisions	M_{Ref}	M3	Wilcoxon	0.0	0.0000
Collisions	M_{Ref}	M4	Wilcoxon	0.0	0.0000
Collisions	M_{Ref}	M5	Wilcoxon	0.0	0.0000
Apprehensions	M_{Ref}	M1	Wilcoxon	1.0	0.0001
Apprehensions	M_{Ref}	M2	Wilcoxon	0.0	0.0000
Apprehensions	M_{Ref}	M3	Wilcoxon	74.5	0.0000
Apprehensions	M_{Ref}	M4	Wilcoxon	159.5	0.0000
Apprehensions	M_{Ref}	M5	Wilcoxon	34.0	0.0000
Loss of wildlife	M_{Ref}	M1	Wilcoxon	50.5	0.0418
Loss of wildlife	M_{Ref}	M2	Wilcoxon	19.0	0.0013
Loss of wildlife	M_{Ref}	M3	Wilcoxon	449.0	0.0010
Loss of wildlife	M_{Ref}	M4	Wilcoxon	602.5	0.0500
Loss of wildlife	M_{Ref}	M5	Wilcoxon	566.5	0.0589
Catch per unit effort	M_{Ref}	M1	Wilcoxon	1.0	0.0001
Catch per unit effort	M_{Ref}	M2	Wilcoxon	26.0	0.0032
Catch per unit effort	M_{Ref}	M3	Wilcoxon	73.0	0.0000
Catch per unit effort	M_{Ref}	M4	Wilcoxon	152.0	0.0000
Catch per unit effort	M_{Ref}	M5	Wilcoxon	31.0	0.0000

F

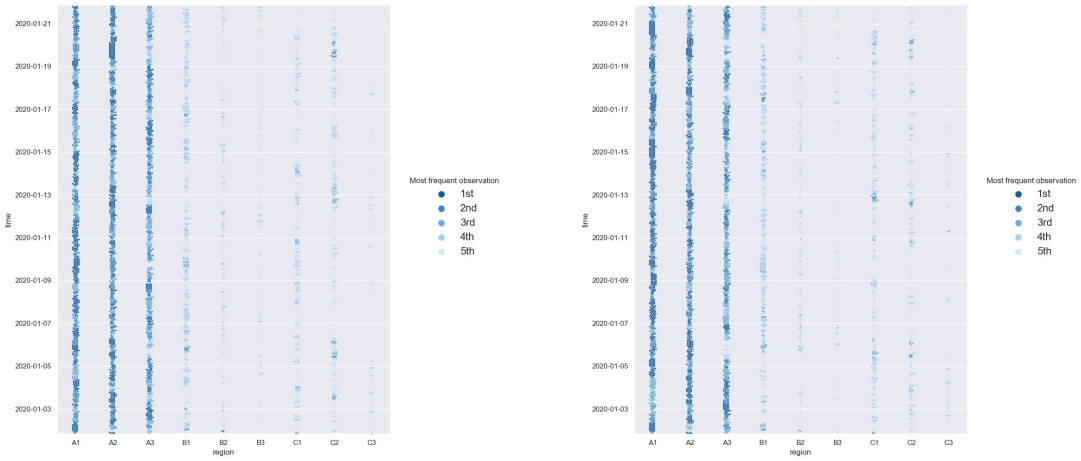
Additional simulation results

This appendix provides the results for experiments E5 and E6 (as defined in Table 2 and Figure 6). In addition, additional results are provided to elaborate on the emergent behaviour described in Section 5.

F.1 E5: Influence of BN regions on surveillance performance

Previous research has focused on analysing the size of hotspots to predict criminal activity and to differentiate between different types of crime [102]. Provided that the ABM model simulates two types of poachers, the number of BN regions is adjusted in accordance with Figure 5. This experiment tests the hypothesis (stated in Figure 6) that smaller BN regions result in prioritisation efforts that better reflect poacher activity. These adjusted dimensions of the BN regions influence the conditional probability tables and therefore the influence of probabilistic inference on UAV path planning. The BN regions are defined such that each region is characterised by its own unique abiotic features. The C-regions represent the entry points for the poachers. The B-regions cover the areas with water sources that are preferred by rhinos during certain time periods of the day as defined in Table 1. The A-regions are mainly characterised by their location relative to the B- and C-regions, the elevation of the environment and the NDVI. In addition, the A-regions are relatively large in comparison to the B- and C-regions.

Figure F.1 indicates that the spatiotemporal patterns of the UAVs in the ABM model are not significantly influenced by varying these BN regions. The spatiotemporal patterns of the UAVs do not converge to the regions where poacher are active. The hypothesis of E5 is therefore rejected.



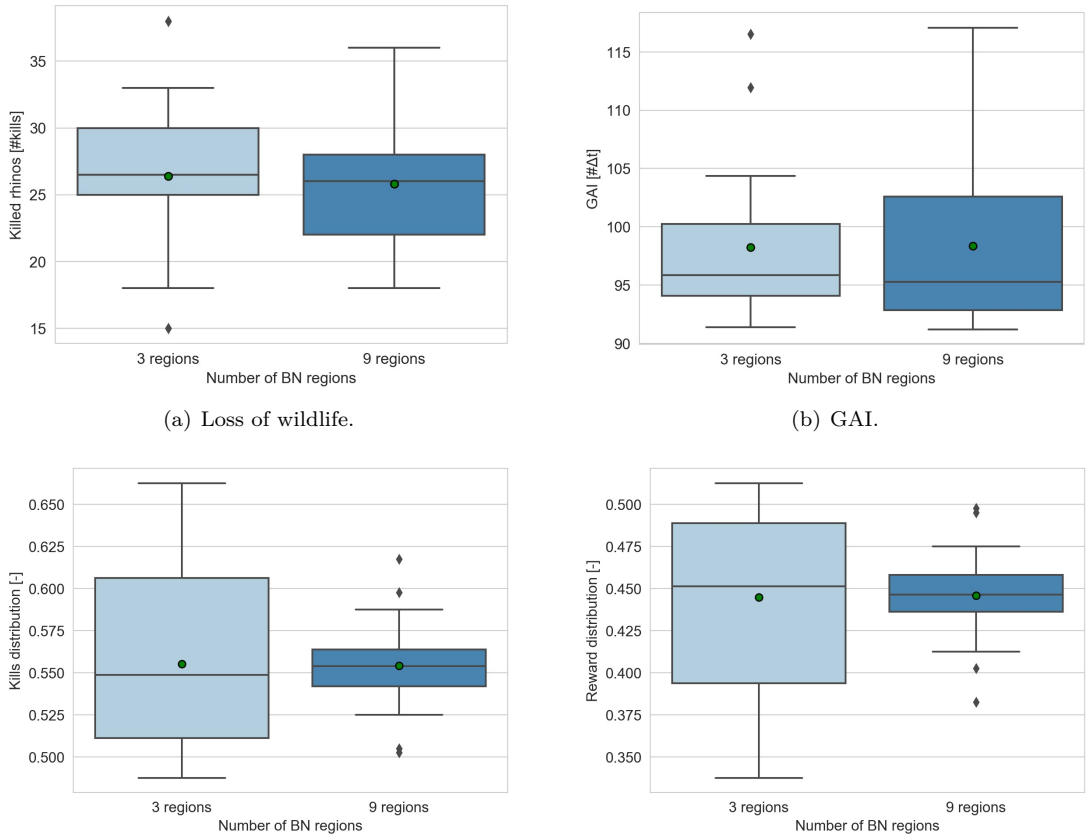
(a) UAVs differentiating between 3 BN regions.

(b) UAVs differentiating 9 BN regions.

Figure F.1: Spatiotemporal patterns of agents for different BN regions.

With an increase in the number of BN regions, the variance reduces of both the surveillance effectiveness (Figure 2(a)), the fraction of the environment in which rewards are collected becomes (Figure 2(d)). This is similar to the trend seen in Figure 2(d), which visualises the distribution of the rhinos being killed over the target area. Accurately differentiating between various hotspots within the target area therefore positively helps in reducing wildlife loss to some extent. More BN regions also results in more prioritisation and therefore causes an increase

in the variance of the GAI. The average model performance does not change significantly however. In addition, the p-values for these KPIs do not meet the significance level of 5% (refer to Table E.4). The spatial patterns in Figure F.1 only visualise the n-th most preferred region of choice, which indicates that the changes in the variance are caused by changes in BN region preference of regions that are of lesser importance for the general emergence of the ABM model. Indeed, the preference for the A-regions does not change significantly, whereas slight changes, that are in accordance with the conditional probability tables, can be observed in the B-regions.



(a) Loss of wildlife.

(b) GAI.

(c) Wildlife loss distribution

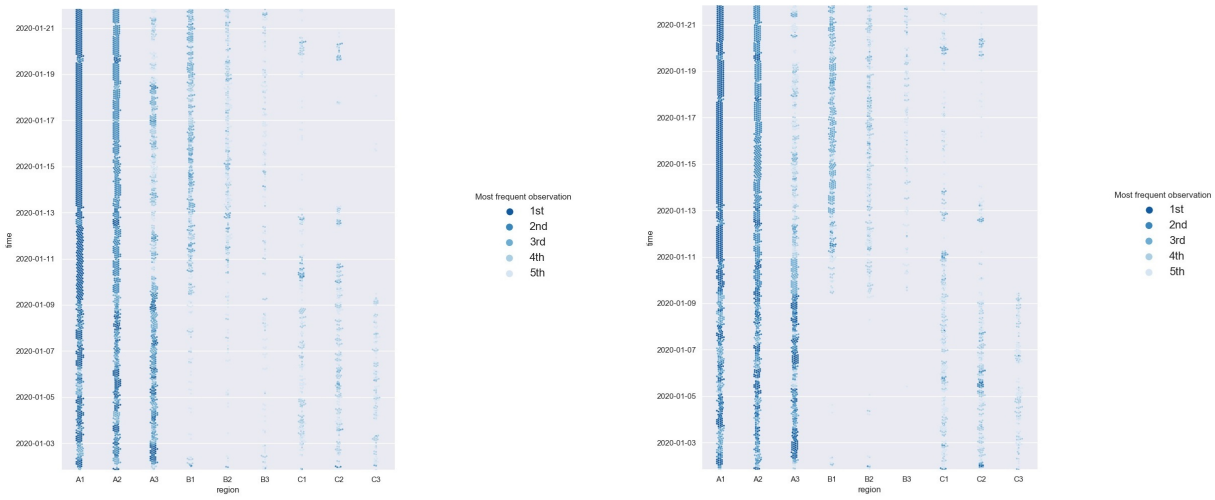
(d) Reward distribution

Figure F.2: Coefficient of variation for the ABM based on the spatiotemporal patterns of the agents.

F.2 E6: Analysis of BN bias through expansion of training data

This experiment analyses whether the training data captures the poacher behaviour without a bias and whether this limits surveillance effectiveness. It addresses the inability to observe the behaviour of all poachers due to the size of the target area, which can result in an incomplete understanding of intruder behaviour and therefore a certain bias [17, 18]. Since the training data used in the experiments consists only of observations from apprehended poachers, this experiment was designed to analyse whether this training data provides a complete understanding of poacher behaviour. The training data is therefore expanded by observations regarding poacher behaviour from poachers that remained unapprehended. The DAG of ABN4 was used in combination with the training data to obtain the conditional probability tables through parameter learning. These conditional probability tables were obtained by discretising the training data into 4 bins and through Bayesian Estimation.

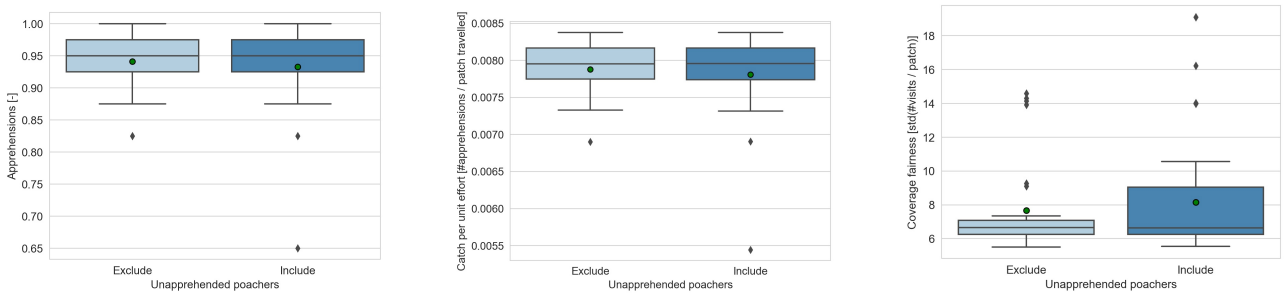
Both the spatiotemporal patterns in Figure F.3 and the KPIs in Figure F.4 show no change in surveillance performance from including unapprehended poacher data in the training data. Analysis of the conditional probability tables shows a marginal increase of preference for the A1- and C1-regions as a resulting from including unapprehended poachers. This directly causes the observed increase in differentiation quantified by the coverage fairness. Table E.5 also shows that no significantly different behaviour is observed in model output. It is therefore determined that the additional training data does not contain significantly different dependencies among the random variables in the BN such that a bias is not detected.



(a) Unapprehended poachers excluded from training data.

(b) Unapprehended poachers included in training data.

Figure F.3: Spatiotemporal patterns of UAVs resulting from ex- and including unapprehended poachers in training data.



(a) Poachers apprehended by UAVs.

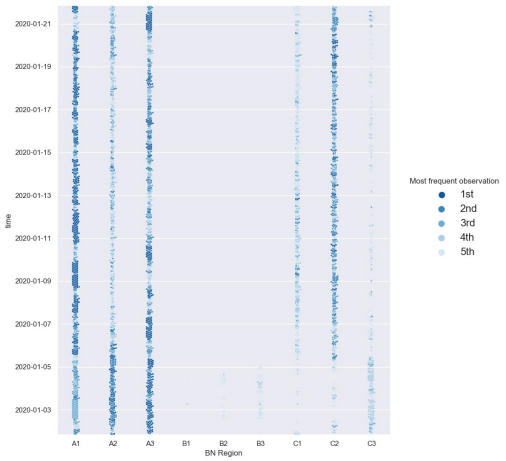
(b) Catch per unit effort.

(c) Coverage fairness.

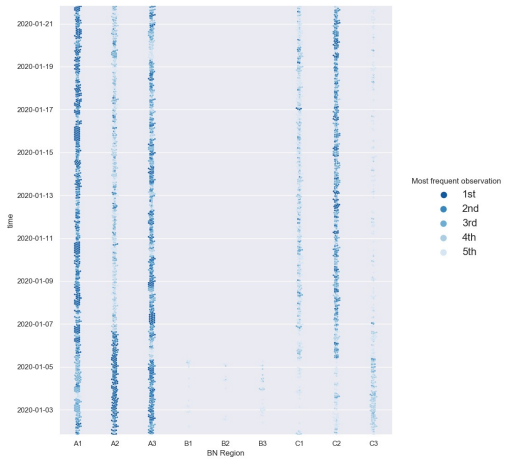
Figure F.4: Performance comparison of BN4 from ex- and including unapprehended poachers in training data.

F.3 Additional results for E2

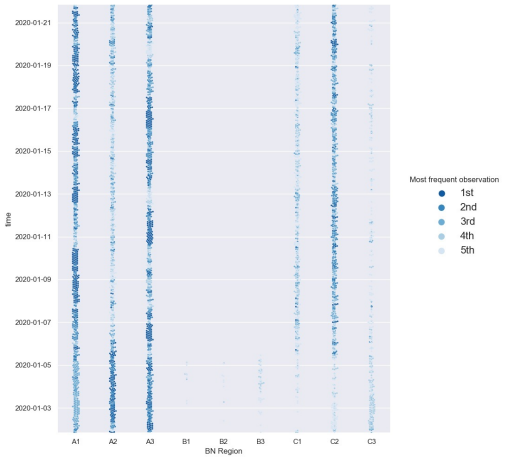
The swarmplots for E2 are visualised in Figures F.5 and F.6. The observed spatiotemporal patterns support the hypothesis that the BN is not utilised by such an extent that it significantly changes UAV behaviour in comparison to the HAPF-ACO model.



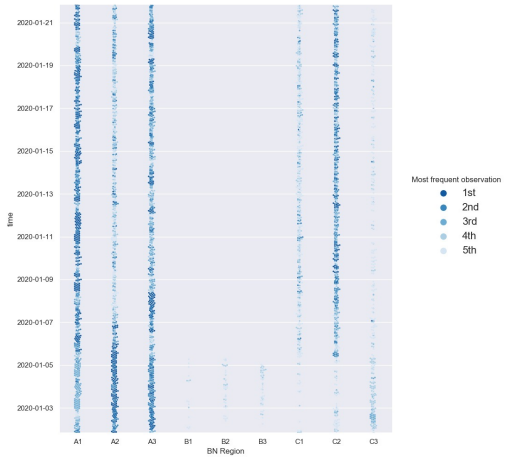
(a) HAPF-BLACOPS model utilising ABN1.



(b) HAPF-BLACOPS model utilising ABN2.



(c) HAPF-BLACOPS model utilising ABN3.



(d) HAPF-BLACOPS model utilising ABN4.

Figure F.5: Swarmplots visualising spatiotemporal patterns of the UAV swarm in E2.

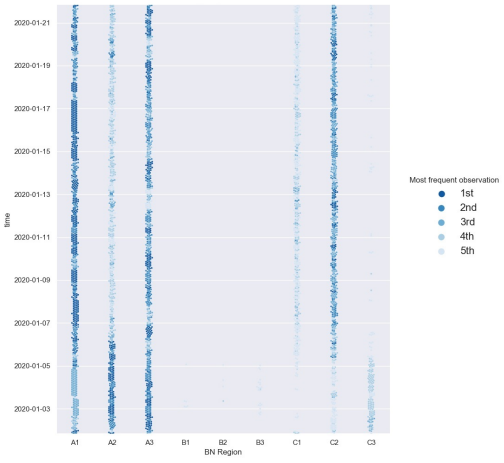
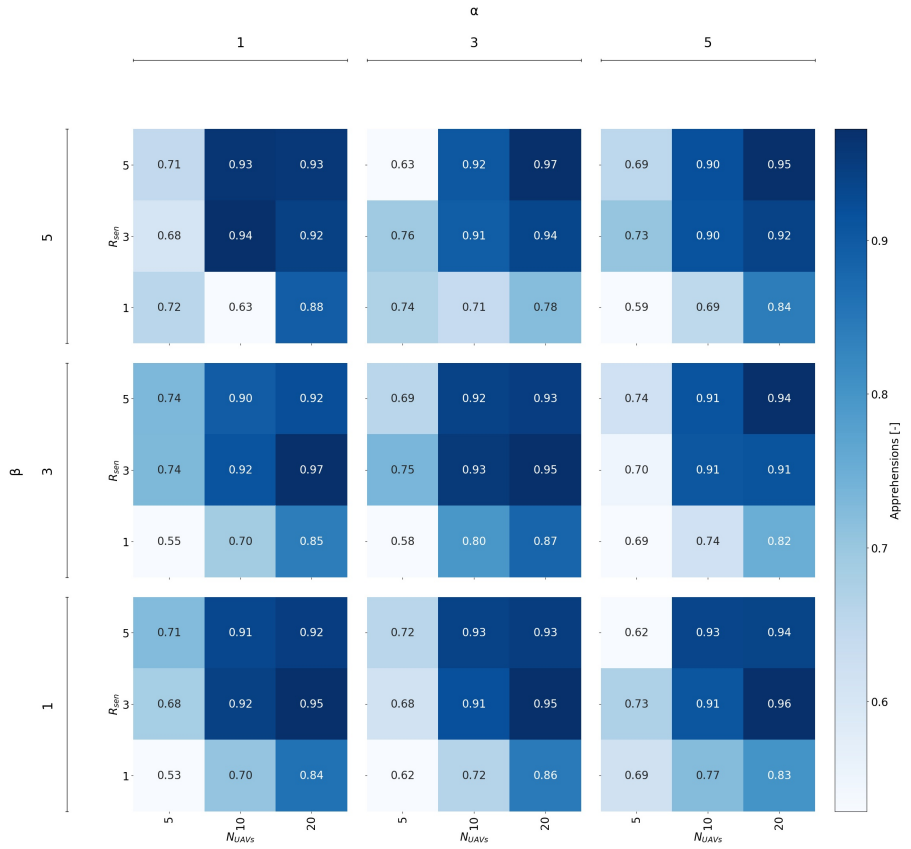


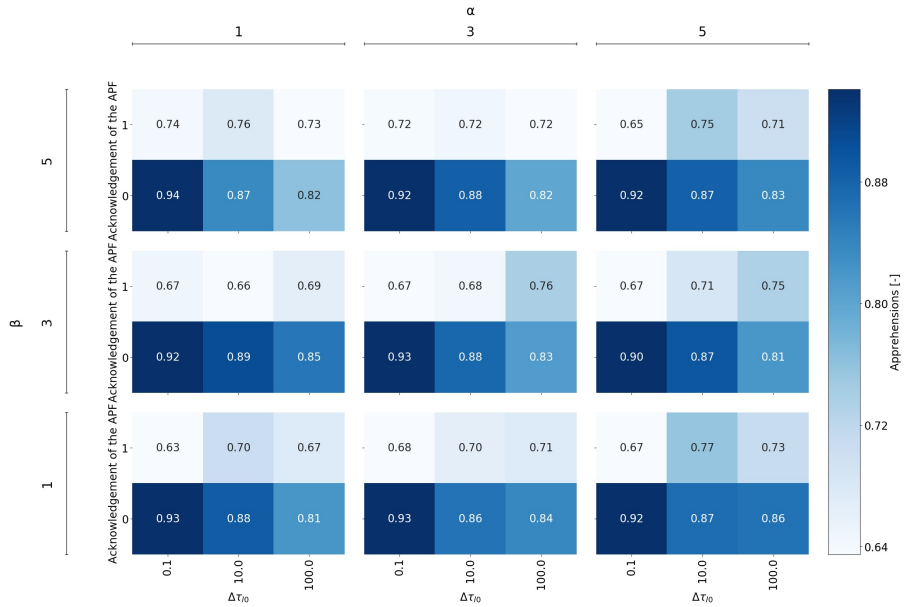
Figure F.6: Swarmplot visualising spatiotemporal patterns of the UAV swarm using the HAPF-ACO model.

F.4 Additional results for E3

The contour plots below indicate the influence of the HAPF-BLACOPS parameters α, β, R_{sen}, $\Delta\tau_{l0}$ and the size of the UAV swarm, N_{UAVs} on model output, as obtained through GSA discussed in subsection 5.4. These results were obtained by simulating ABN4, defined in Appendix B, integrated with the proposed HAPF-BLACOPS surveillance model. A value of 0 for the acknowledgement of the APF indicates that mission type 2 is disregarded by the HAPF-BLACOPS model.

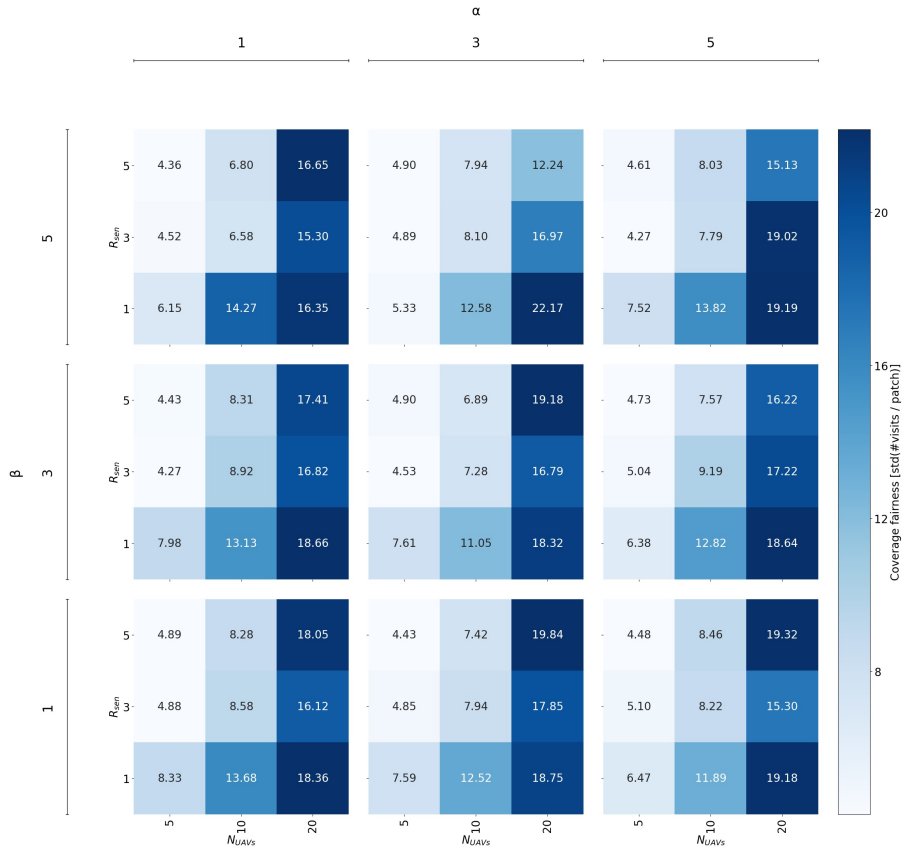


(a) Influence of GSA parameters α, β, R_{sen} and N_{UAVs} on apprehensions.

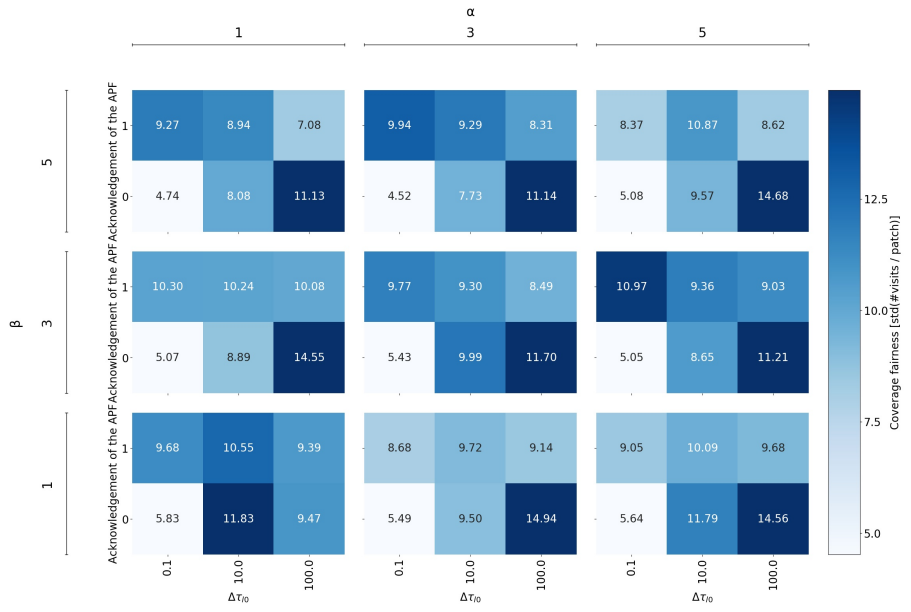


(b) Influence of GSA parameters α, β, APF acknowledgement and $\Delta\tau_{l0}$ on apprehensions.

Figure F.7: Contour plots of the percentage of apprehensions.

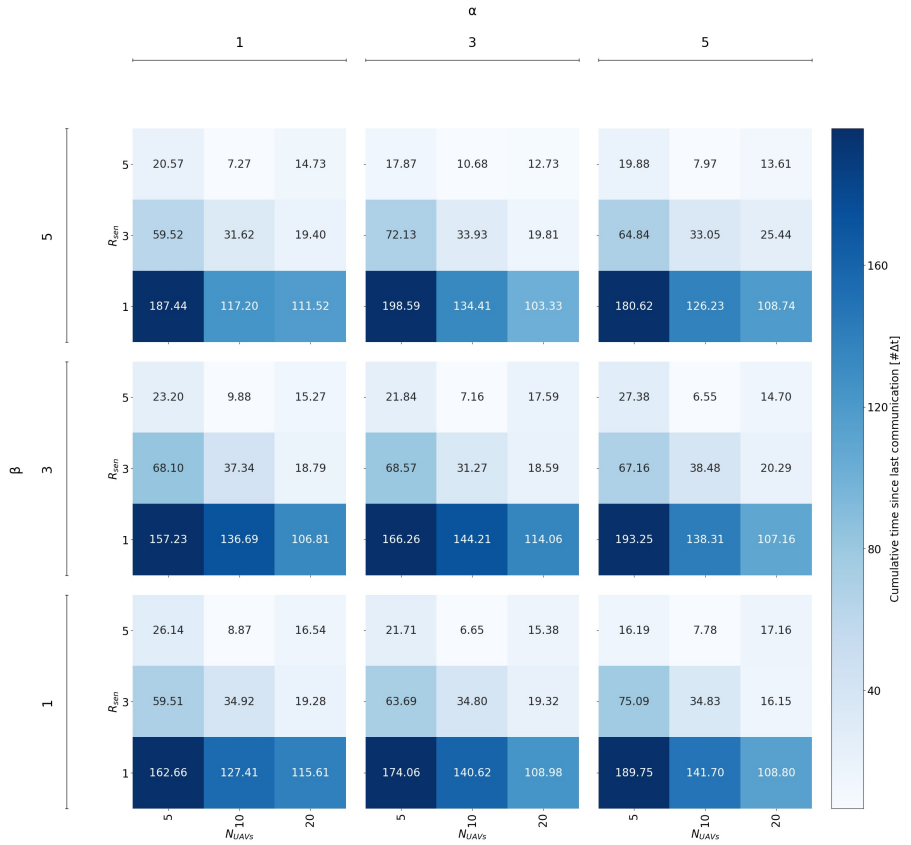


(a) Influence of GSA parameters α, β, R_{sen} and N_{UAVs} on coverage fairness.

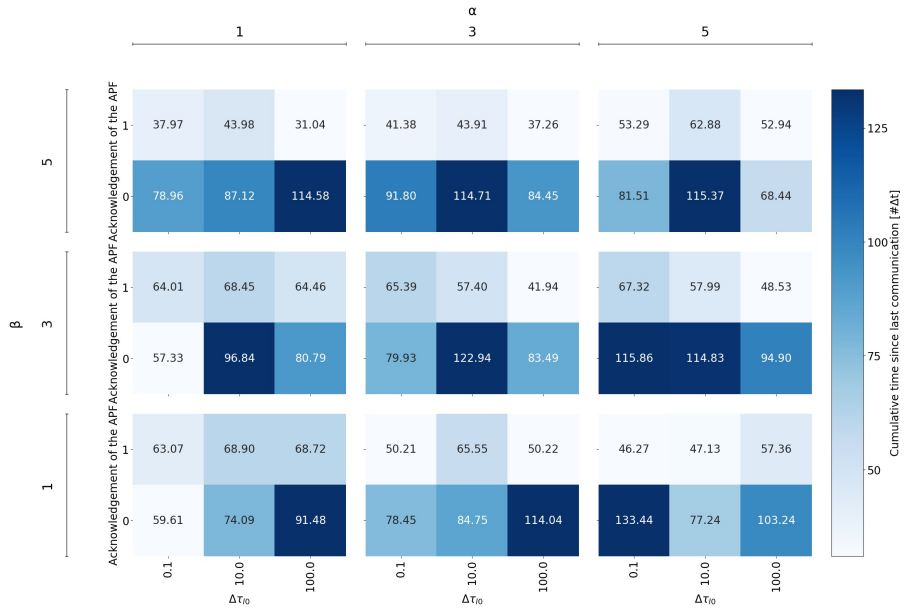


(b) Influence of GSA parameters α, β, APF acknowledgement and $\Delta\tau_{l0}$ on coverage fairness.

Figure F.8: Contour plots of the coverage fairness.

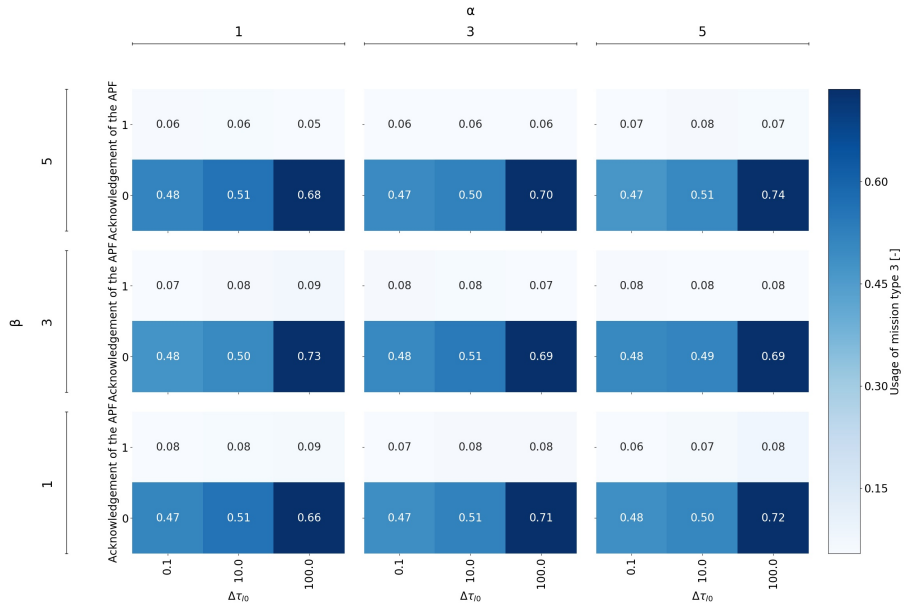


(a) Influence of GSA parameters α, β, R_{sen} and N_{UAVs} on cumulative time since communication.

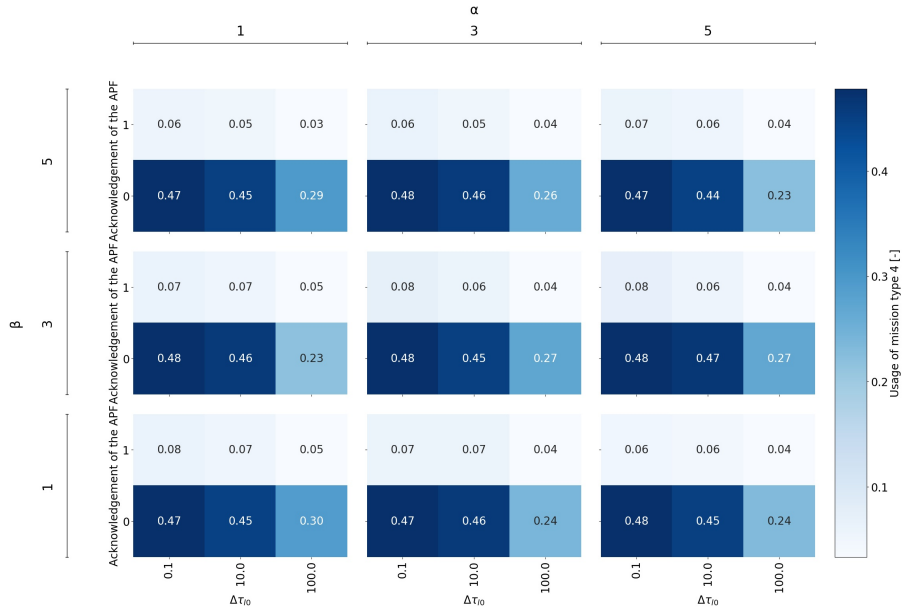


(b) Influence of GSA parameters α, β, APF acknowledgement and $\Delta\tau_{l0}$ on cumulative time since communication.

Figure F.9: Contour plots of the cumulative time since last communication between UAVs.



(a) Influence of GSA parameters α, β, R_{sen} and N_{UAVs} on mission type 3 utilisation.

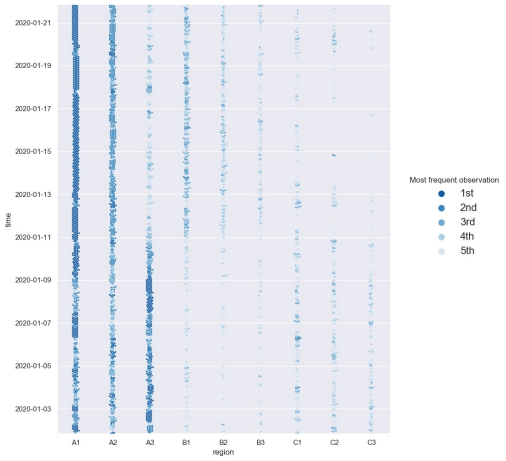


(b) Influence of GSA parameters α, β, APF acknowledgement and $\Delta\tau_{l0}$ on mission type 4 utilisation.

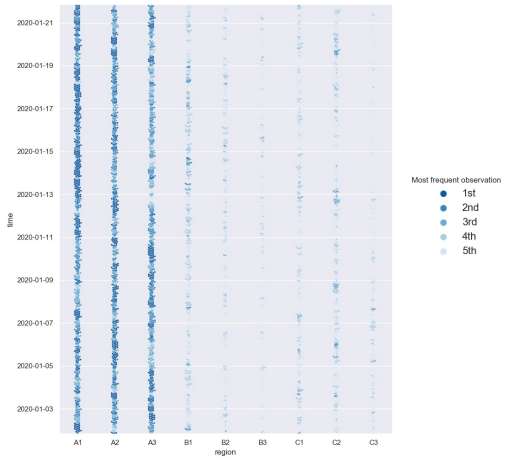
Figure F.10: Contour plots of the utilisation of mission types 3 and 4 for UAV path planning.

F.5 Additional results for E4

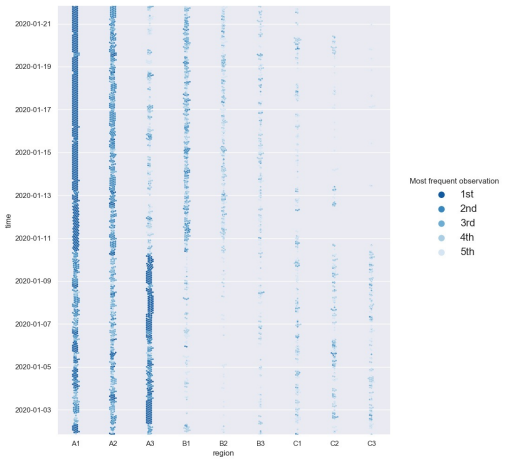
The swarm plots below show the n-th most preferred BN-region, as preferred by the UAV swarm simulating with the HAPF-BLACOPS model utilising various (A)BNs, defined in Appendix B. These swarm plots make use of GSA results to improve surveillance performance, which results in the UAVs targeting regions where both poachers and rhinos are active (as seen in Appendix A). Additionally, it is seen that the different DAGs and their respective conditional probability tables all result in similar behaviour, which indicates that various random variables being considered, do not accurately address poacher behaviour.



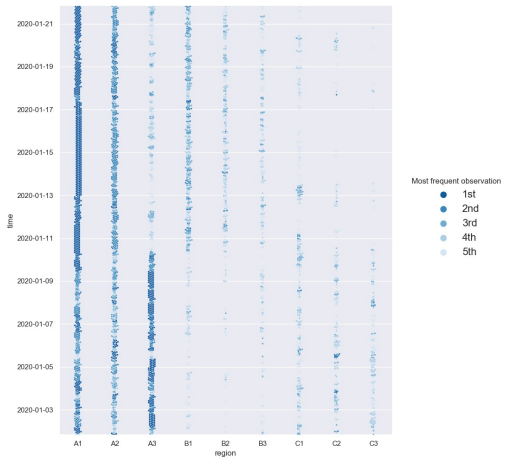
(a) HAPF-BLACOPS model utilising ABN3.



(b) HAPF-BLACOPS model utilising ABN4.

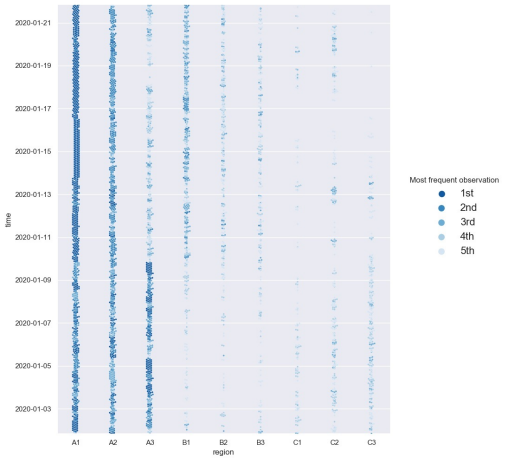


(c) HAPF-BLACOPS model utilising BN5.

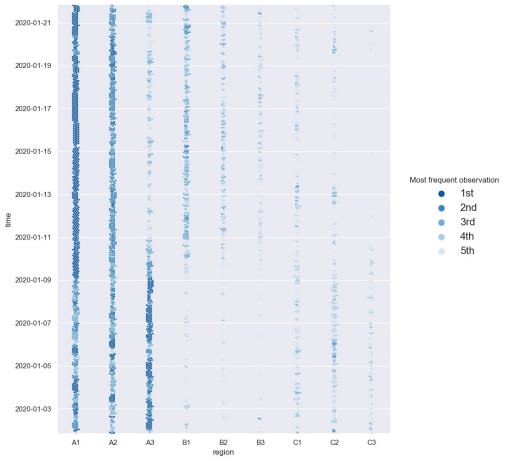


(d) HAPF-BLACOPS model utilising BN6.

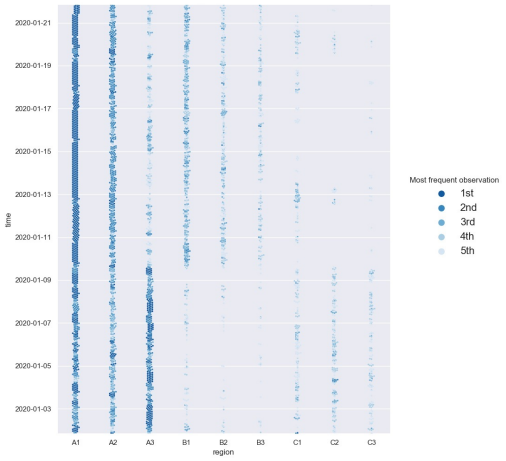
Figure F.10: Swarmplots visualising spatiotemporal patterns of the UAV swarm in E4.



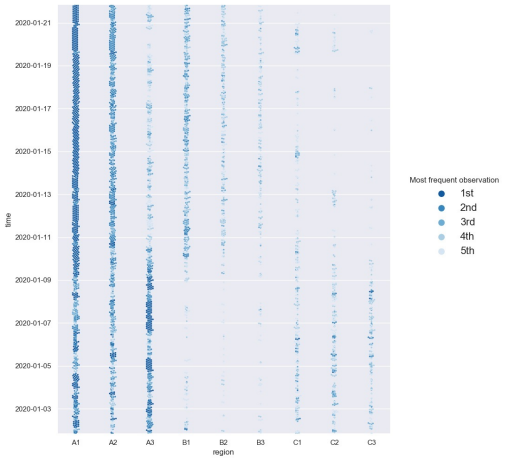
(e) HAPF-BLACOPS model utilising BN7.



(f) HAPF-BLACOPS model utilising BN8.



(g) HAPF-BLACOPS model utilising BN9.

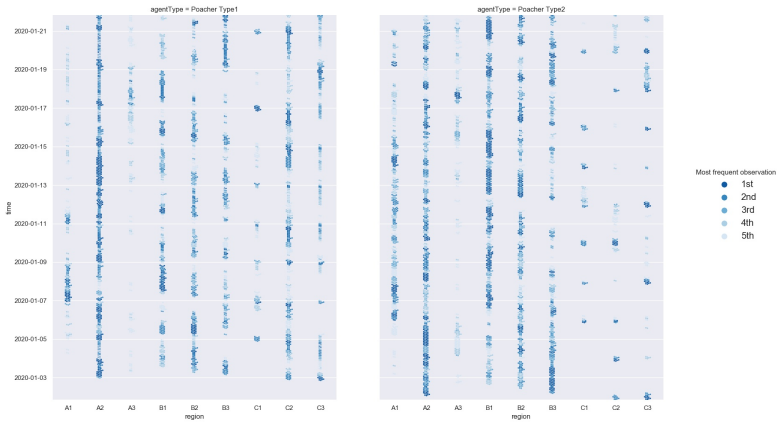


(h) HAPF-BLACOPS model utilising BN-Expert.

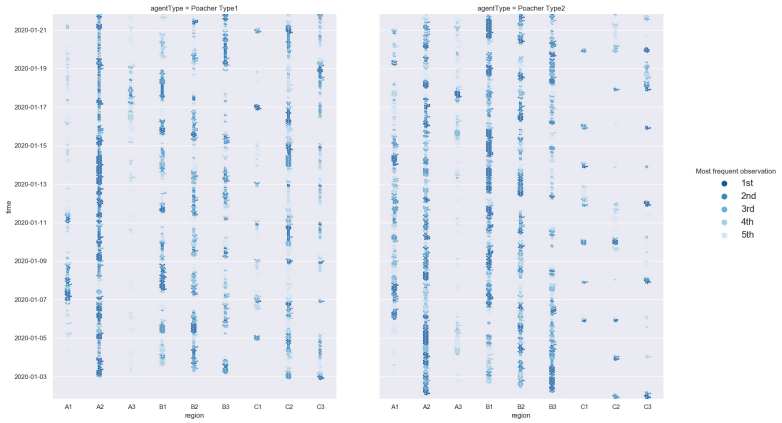
Figure F.11: Swarmplots visualising spatiotemporal patterns of the UAV swarm in E4.

F.6 Additional results for E7

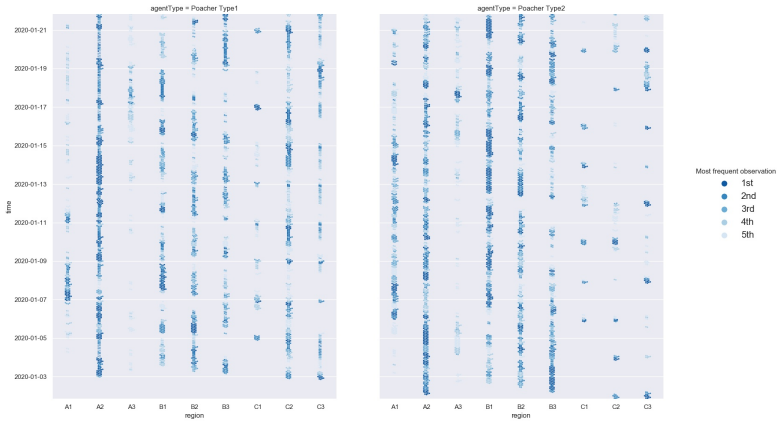
The swarm plots below show the n-th most preferred BN-region, as preferred by both type 1 and type 2 poachers for various levels of consideration of gained experience, which is quantified by the C_{ASU} parameter.



(a) ACOSG model utilising C_{ASU} of 0.01.



(b) ACOSG model utilising C_{ASU} of 1.0.

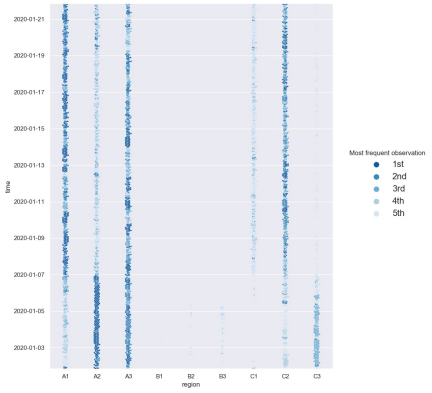


(c) ACOSG model utilising C_{ASU} of 100.0.

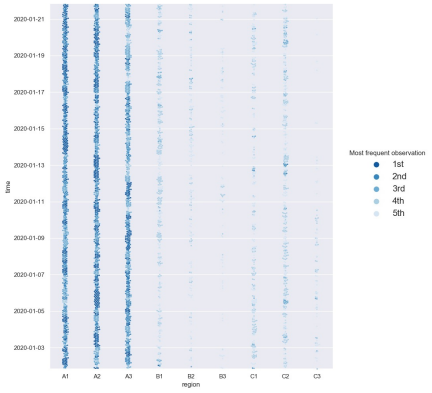
Figure F.12: Swarmplots visualising spatiotemporal patterns of type 1 and type 2 poachers.

F.7 Additional results for E8

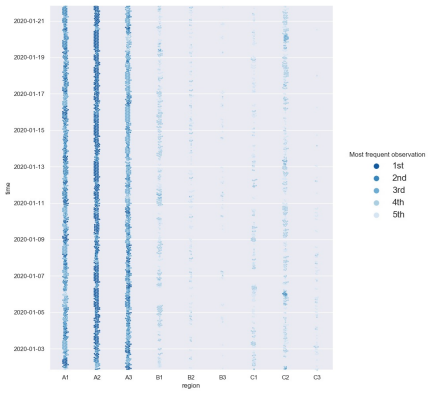
Figure F.13 presents the spatiotemporal patterns of the UAV swarm for various alternatives of the proposed surveillance model.



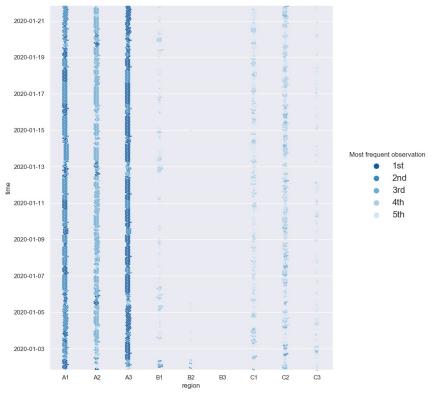
(a) Spatiotemporal pattern of ABN4 (Reference).



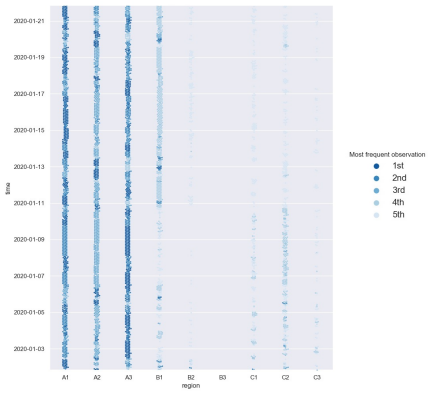
(b) Spatiotemporal pattern of ABN4 (GSA).



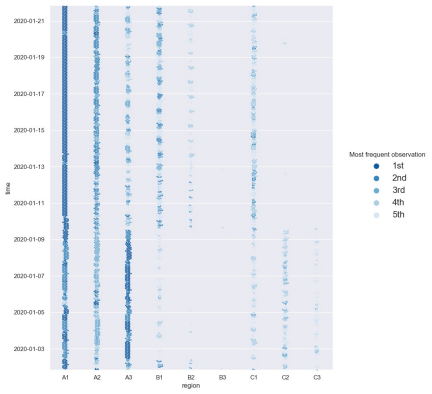
(c) Spatiotemporal pattern of ABN4 (GSA, $N_{UAVs} = 10$).



(d) Spatiotemporal pattern of HAPF-BLACOPS (GSA, MT3, $\beta = 0$).



(e) Spatiotemporal pattern of ABN4 (GSA, MT3, $\beta = 0$).



(f) Spatiotemporal pattern of BN9 (GSA, MT3, $\beta = 0$).

Figure F.13: Swarmplots visualising spatiotemporal patterns of the UAV swarm in E8.

III

Literature Study

previously graded under AE4020

Abstract

The continued loss of wildlife due to poaching poses threats on iconic species and the livelihood of local communities. A search for cost effective and innovative solutions to aid in fighting poachers and save ecosystems and biodiversity in national parks has emerged as consequence. UAVs have been subject to growing interest within green security contexts due to their ability to monitor illegal activities in large target areas. An extensive literature review regarding UAV based anti-poaching efforts revealed interesting research opportunities. In dynamic environments, where intelligent poachers can coordinate to find wildlife and can adapt to law enforcement, there is a need for effective coordination of distributed multi-UAV persistent surveillance. A swarm based surveillance model is therefore adjusted to incorporate prioritised, persistent surveillance. Adaptive surveillance in response to adaptive poachers is achieved through online, distributed machine learning using Bayesian Networks to predict criminal activity. Implementation and simulation in an ABM framework is to be used to compare performance with current state-of-the-art surveillance and perform sensitivity analysis of model input parameters.

1 Introduction

Wildlife is threatened by several forces including illegal wildlife trade, habitat destruction, pollution and climate change. Together, this amounts to more than 30000 animal species being at the risk of extinction [33]. Illegal trade in wildlife is considered to be the fourth largest criminal industry in the world [70, 80] and poses serious risks on the survival of iconic species such as the tiger, rhino and the elephant [25, 81]. Poaching of these animals has been going on for decades [30, 53] and has reduced the elephant population in Africa by 144.000 animals in the time period from 2007 to 2014. This trend has been continuing to affect the elephant population at an annual rate of 8% due to poaching activity, although other animal species also remain under threat [23, 29, 77]. Besides affecting wildlife, poaching also affects the livelihood of local communities that depend on wildlife tourism [85]. It has also been found that it is used to partially finance terrorist organisations, therefore indirectly posing risks to national security as well [8, 12, 70, 87]. Already in 1990 it was concluded that law enforcement and adequate manpower was required to fight poachers, since detection was found to be the most effective strategy [53]. Today, politicians still face challenges when it comes to balancing economic development and the maintenance of healthy environments [51]. Park managers are still calling for cost effective and innovative solutions to handle environmental problems that threaten the ecosystems in national parks [22]. Challenges also arise from the lack of resources to fight illegal activities, therefore requiring efficient and strategic allocation and scheduling of resources [17].

Surveillance drones have been subject to growing interest throughout the last decade due to their flexibility to carry a variety of sensors. Efficient control and surveillance of illegal activities in terrestrial areas has been identified as the most common practical implementation area for Unmanned Aerial Vehicles (UAVs). Although their implementation is restricted by their maximum endurance [8, 40, 48], conservation applications have gained increasing interest in the last decade thanks to the affordability, flexibility and safety characteristics of UAVs. Despite their endurance, their usefulness is particularly of interest when large areas are to be monitored, that are difficult to cover by ground. This is specially an interesting topic when UAV technology is combined with machine learning techniques. These techniques can improve effectiveness of crime prevention through continuously learning from criminals [94]. In turn, this can be used to predict patterns of illegal activities, such as poaching in national parks [8, 22, 59, 74, 77, 87]. These characteristics have resulted in the surveillance problem being studied extensively. The focus from these studies has shifted over the years from single UAV to multi-UAV approaches and a distinction can be made between discrete and persistent surveillance [13, 57]. A single sweep of the target area to detect criminal activity ends once full coverage has been achieved. This is not effective in green security, since poachers can continuously enter national parks. The behaviour of poachers has been studied extensively using Stackelberg Security Games (SSGs) in order to optimise security resource allocation. This has lead to the development of advanced human decision models to model poacher behaviour in multi-stage games [23], although the wildlife dimension has been neglected in most research.

This thesis proposes a multi-agent based approach to tackle the challenges regarding the allocation of limited resources to perform surveillance in complex dynamic environments. The use of a bottom-up approach to understand the collaborative and adaptive behaviour of the poachers and the UAVs, along with local spatial communications between agents indicates that an agent-based method is less complex than Game Theory

approaches [2, 54]. Agent Based Modelling (ABM) is also capable of explicitly representing the environment along with its dynamics, it can accommodate spatial patterns of inter-agent mechanics, while exploring feedback loops and adaptations within these systems [51, 82].

1.1. Problem definition

The various characteristics of green security require an effective multi-UAV approach to perform persistent surveillance such that poacher activity can be detected. The lack of resources at national parks and other contextual aspects pose challenges to be overcome by the to-be-proposed multi-UAV surveillance model.

A fully autonomous system would allow local authorities to experience economic benefits by abandoning the task of searching for poachers. This is usually done on foot in extreme environments, leaving large areas uncovered [29]. A non-deterministic online learning surveillance approach is needed due to these vast areas in which poachers can collaborate, evade local authorities and adapt to surveillance strategies [51, 56]. The surveillance approach also requires a distributed coordination method to account for a lack of digital infrastructure [97], which is also beneficial to the system's robustness.

This research takes a novel multi-agent based modelling approach to implementing and analysing a fully autonomous, distributed and learning multi-UAV persistent surveillance model that can be used to mitigate adaptive poacher activity in dynamic environments, while considering operational constraints. Additionally, surveillance effectiveness is to be improved by incorporating prioritised surveillance such that locations of known criminal activity are targeted more frequently. In doing so, state-of-the-art human decision making models will be considered that incorporate possibilities for collaborations and adaptation of behaviour. The dimension of wildlife motion will also be included to quantify the animal population density as input for the poacher's decision making. From the ABM perspective, this requires modelling the target area and the behaviour and decision making of the poachers, the animals and the UAVs [51, 82]. Finally, the to-be-developed surveillance model used to mitigate poacher activity will be analysed using simulations and sensitivity analyses. The results will be used to compare effectiveness with a state-of-the-art surveillance model to validate the proposed model.

1.2. Report structure

The body of this report is structured in order to elaborate further on the problem definition by first discussing previous research, from which a research proposal is derived. Chapter 2 therefore discusses the state-of-the-art literature in which relevant work is discussed and analysed. From there, a research gap is identified such that a research proposal can be formulated. This is covered in Chapter 3, where the research questions are formulated in accordance with the extensive literature review and the research gap. Chapter 4 presents the research methodology that will be used to address the identified research gap and to provide a scope for the research project. This chapter also elaborates on the to-be-performed simulation setup, which will be used to quantify the effectiveness of the proposed model. Chapter 5 discusses the expected results from this simulation experiment and research project in general. Here, the results of this research proposal are related to previous work in order to justify the relevance of this work. The schedule of this research, containing the various project phases discussed in the previously mentioned chapters, is discussed in Chapter 6. Finally, a conclusion will be drawn in Chapter 7 regarding this research proposal's feasibility and its relevance.

2

Current state-of-the-art research

In order to understand how current state-of-the-art literature is progressing regarding the problem of surveillance in national parks, an extensive literature review is performed. This literature review first covers the challenge of performing surveillance using autonomous robots in target areas. Various solution approaches are discussed, after which the most relevant solution techniques are discussed in further detail. Next, the review focuses on the targets to be detected through surveillance and aims to identify relevant techniques that model human poachers and their decision making in a realistic manner. These two aspects form the basis of the model that is to be developed. This is followed by discussing work regarding machine learning techniques in surveillance contexts that can be used to improve effectiveness of crime detection. Finally, this Chapter concludes with a formulation of the research gap such that further research opportunities can be identified.

2.1. Surveillance using autonomous robots

The problem statement introduced some aspects that need to be considered for the multi-UAV persistent surveillance model. This section elaborates further on these contextual requirements that are based on the relevant state-of-the-art literature discussed below. A number of models are discussed, between which a comparison will be made in the following sections to conclude this aspect of the literature review.

The surveillance problem is not a new concept as it dates back to the second world war, when the objective was to find submarines [66]. Surveillance is also known as the coverage path planning (CPP) problem, which is a non-deterministic polynomial-time-hard (NP-hard) problem. This means that it is unknown whether efficient algorithms exist to find solutions in polynomial time, whereas given solutions can be efficiently verified to be correct solutions to the problem [11]. The solution approaches address a motion planning subtopic in robotics, where each robot needs to be assigned to a created path such that every location in a target area can be explored within the capabilities of the robot, while avoiding obstacles. These capabilities include the robot's limitations w.r.t. kinematics, endurance, communications and sensor footprint [13, 57]. The basic concept of surveillance is visualised in Figure 2.1 where several UAVs are used to monitor the target area. It is not to be confused with the coverage problem, where optimal locations are to be determined for a set of sensors that are to continuously monitor a target area. The exploration problem in which dynamic targets are to be identified is more similar, although exploration does not necessarily result in full area coverage and is a discrete search [57]. Surveillance can be achieved through application of various types of UAVs, usually categorised as fixed-wing and rotary-wing UAVs. Each type has its own challenges and advantages. Fixed-wing UAVs support longer flights due to their aerodynamics and endurance capabilities. Heavier payloads can be carried, although a runway is required for take-off and landing. Rotary-wing UAVs are capable of hovering thanks to the use of rotary blades which support Vertical Take-Off and Landing (VTOL) operations, whereas the fixed-wing design does not allow for hovering above targets. Rotary-wing UAVs have lower payload carrying capabilities and their endurance is limited to about 20-25 minutes. Hybrid UAVs exist to overcome these restrictions [13], although these are not considered in the remainder of this literature review.

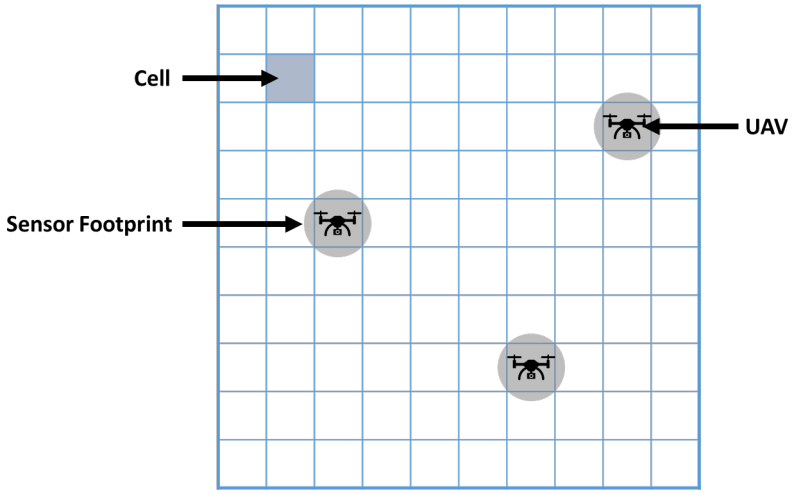


Figure 2.1: Multi-UAV surveillance in a discretised target area [57]

The need for robust and reliable systems that improve the efficiency of search algorithms utilising multiple robots has increased due to the rise of autonomous systems [20, 57, 66]. The surveillance problem has been solved through a variety of approaches since the second world war [9, 13, 31], although limited research has been aimed at the persistent surveillance problem since then [57]. A distinction within literature can be made by recognising that persistent surveillance is a generalisation from surveillance [57], where the search is completed upon finding the targets [31]. Introducing persistence also brings additional challenges that need to be dealt with [37]. Persistence is achieved by revisiting locations within a given area [57]. This is a crucial distinction since the intruders within this thesis' problem definition can enter the area at any moment. Criminals can wait before entering the national park in order to avoid surveillance drones, enabling intruders to move through areas that were patrolled by UAVs at an earlier point in time. It is also possible that once a poacher has been detected in a certain area, which is then considered to be clean from criminals, a different poacher decides to go to this same area. This emphasises to need to continuously monitor the target area in order to achieve effective surveillance. The persistent surveillance problem is also commonly referred to as patrolling or continuous sweep coverage [67]. The objectives of solution algorithms are closely related to periodically revisiting locations in order to minimise the largest revisit time, also referred to as the idleness, refresh time, latency or age [13, 16, 57]. Despite the requirement of achieving this objective being at the core of each of these approaches, contextual differences cause variations in solution approaches. Sub-optimal solution techniques are usually applied since the persistent surveillance problem is NP-hard [67]. Another distinction among solution approaches can be made by identifying systems capable of global communication and bio-inspired systems driven by local autonomy [31]. Such distributed approaches do experience the disadvantage of not guaranteeing exhaustive coverage and the need to overcome local minima [32, 86]. Local autonomy translates to each UAV having local knowledge regarding the system and environment states and to make decisions based on this information combined with information gained from indirect (or short range) communication. Flocks of birds and termites are the source of inspiration for these local autonomy models and define the swarm intelligence algorithms. Similar to pheromone usage is the method utilising social potential fields. These systems use both attraction and repulsion forces which can be adjusted dynamically and determine the motion of UAVs. The law forces are based on distances and influence the emergent behaviour of the system [69]. Communication in such systems is generally over short ranges via a one-to-many protocol and can be the preferred solution method for a number of reasons. The first being scalability, as local communication does not cause significant performance requirements within the system of UAVs when a UAV is added and/or removed from the system. Second, limited awareness regarding the system and environment results in diversity regarding decision making of individual UAVs. In case all UAVs have access to all available information, it is more likely that each UAV will make the same decision. In turn, the system will converge to some optimal state. Introducing diversity also improves the system's resilience against failure. Thirdly, short range communications pose less stringent requirements on UAV hardware and weight restrictions, which translates into a cheaper UAV [32].

As mentioned, a distinction among surveillance methods can be made by considering discrete and contin-

uous searches. Discrete approaches are used when the number of targets is known in advance, such that the mission is considered complete either once the target area has been fully searched or once all targets have been found [32]. Such approaches are useful in context of search and rescue where online methods that continuously update the knowledge of the environment are preferred [13]. Considering the need to identify poacher activity and adapt to the dynamic environment and evasive manoeuvres, online methods are required to be applied within this research as well. An example of the discrete surveillance approach is the distributed Mixed Integer Linear Program (MILP) approach considering communication shortcomings in [26]. It uses greedy heuristics to reduce the solution space and limit computation times. Robustness is introduced through the assumption that all robots have similar and current knowledge of the system such that individually computing the MILP solution results in similar solutions [26], thus introducing a computationally heavy solution approach relative to swarm based solutions. In [20] the authors emphasised the need to improve computational costs in order to effectively consider dynamic environments. A discrete exploration approach based on reinforced random walk was proposed by Albani, D. et al [3] that is computationally cheaper compared to the method proposed in [26]. Each individual UAV splits the target area into two parts and explores either one based on a utility value such that prioritisation occurs. The UAVs are randomly guided to poorly explored areas to cover the entire target area. Apart from being discrete, the algorithm also prevents UAVs from revisiting previously areas [3].

The common approach to modifying such methods in order to obtain persistent surveillance is to toggle the state of grid cells within the target area from explored to unexplored after some period of time. This forces the UAVs to re-explore the node in question such that persistent surveillance is created. The challenge with such an approach is to determine the period of time after which the state of such a node is to be toggled. Toggling the state too early prevents the swarm from covering the entire target area as the UAVs are being drawn to the same locations over and over [32]. The effectiveness of detecting criminal activity drops if the state is toggled too late, resulting in the UAVs not re-exploring these nodes in time. In the extreme, this can cause the UAVs to become stuck at certain locations upon full coverage of the target area without having the objective to re-explore.

Within game theory, models regarding conservation are referred to as green security games. Examples of such continuous surveillance models include PAWS [23], OPERA [91] and CAPTURE [56]. These models incorporate spatial aspects as well as adaptive poachers and imperfect law enforcement behaviour. The focus is mainly on the interactions between poachers and rangers in order to find optimal surveillance strategies [23] and ignore the behavioural and ecological aspects of wildlife that drive the poachers [54]. Such solution methods use centralised optimisation approaches rendering such surveillance techniques inefficient in online, distributed environments. Similar to the proposed method in [24], the model of the target is integrated in the solution models, such that these models cannot be generalised to different targets. Computing these solutions requires extensive computation times as well, especially for an increasing number of UAVs [24] and scaling up to larger environments [23]. In [51] the authors point out that wildlife's adaptive behaviour of habitat selection, movement ecology, and its responses to dynamic environments are key to successful conservation. Wildlife research has therefore mainly focused on understanding wildlife use of habitats and how resource availability affect their avoidance of certain areas, since it can have profound effects on movement patterns. Research to green security is therefore a cross-disciplinary field [51], as highlighted by Game Theory based models that heavily rely on behavioural models describing the poachers [56]. Despite the urgency of incorporating wildlife dynamics, artificially created environments are advantageous when the focus of research does not lie with the dynamics of a particular ecosystem. The environments can be modeled using biotic and abiotic factors. Abiotic factors used to model environment in ABM models include land cover, elevation, temperature and water. These factors are considered essential in finding emergent properties from local interactions in realistic ecosystems [51].

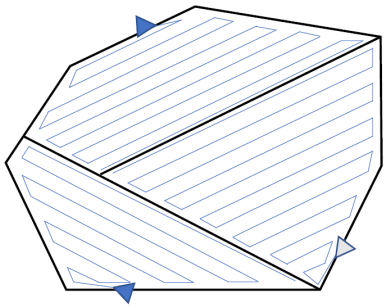


Figure 2.2: Sweep surveillance approach on decomposed target area, blue lines indicate the search paths of the individual UAVs [13]

Other continuous surveillance approaches use some form of grid decomposition to define partitions that are to be monitored by individual robots, visualised in Figure 2.2. Such approaches were presented in [43, 63, 92], where the target area was subdivided into rectangular subareas by minimising the overlap. A different grid decomposition method is to use the Voronoi algorithm [92]. Decentralised graph partitioning approaches have been studied as well that rely on predefined rules [86]. Most of these works do not incorporate path planning since these methods are considered to be modular. Decentralised decomposition of the environment introduces additional challenges such as the need to detect failed UAVs with short-range communication capabilities. Such challenges become larger as the partitioned areas grow larger than the footprint of the UAV. This requires periodic meetings among neighbouring UAVs or with a central node [68]. Another solution is fixed formation surveillance where UAVs have a fixed relative position in order to be able to constantly (indirectly) communicate with each other [35, 75]. Lissajous curves have been applied to solve the persistent surveillance problem as well, although its central characteristics limit its robustness [10].

A simple search path is proposed in [66], where robots coordinate by following equally spaced and parallel search paths resulting in a sweeping search. Royset, J. et al. proposed an exact solution using a MINLP approach to address the surveillance problem with the objective of minimising the largest non-detection probability [72]. Such exact solutions are found to be computationally expensive [46, 69], such that search approaches were introduced that use short-sighted algorithms to determine optimal solutions only for a limited set of future time points [66]. An approach combining grid decomposition and evaporation of certainty levels was proposed in [47], where UAVs define their path using the A* algorithm. Rather than assigning robots to various areas of the target area, it is also possible to compute a search path along the border [34] or through all locations that require monitoring [1, 62]. From there it is easily adjusted to include multiple robots by segmenting the search path as shown in Figure 2.3. Each patrolling robot inverts its direction once the end of its segment is reached, such that continuous surveillance is obtained. The approach is based on one-to-one coordination that aims to minimise the observation intervals and the time between detecting and sharing new information. Their work also considered heterogeneous robots such that larger segments are assigned to robots with a higher velocity, although its robustness to hardware failure is not considered [1]. In [6] individual robot capabilities are used to assign each individual to a partition of the target area, after which a zamboni flight pattern is implemented. Such search patterns are effective and the approach can easily be adjusted to incorporate different search patterns, but its unpredictability is limited. In order to account for the unpredictable search path, a sweeping approach was proposed that ensured detection of moving targets based on the target's maximum relative velocity. This requires several UAVs to fly in fixed formation such that the width of the formation is based on the amount of distance the target can cover before it moves past the last UAV [5]. Despite the 100% probability of detection of moving targets, the approach does not allow for hardware failure and is not robust. It also uses sweep coverage techniques such that targets can plan their visits in between surveillance moments. Despite being able to detect the targets, Li, M. et al. proposed to consider uncertain apprehension. This allows the targets to escape, depending on the available manpower to apprehend them [46]. However, the escape and pursue mechanisms are not within the scope of this research project (see Chapter 4 for further elaboration).

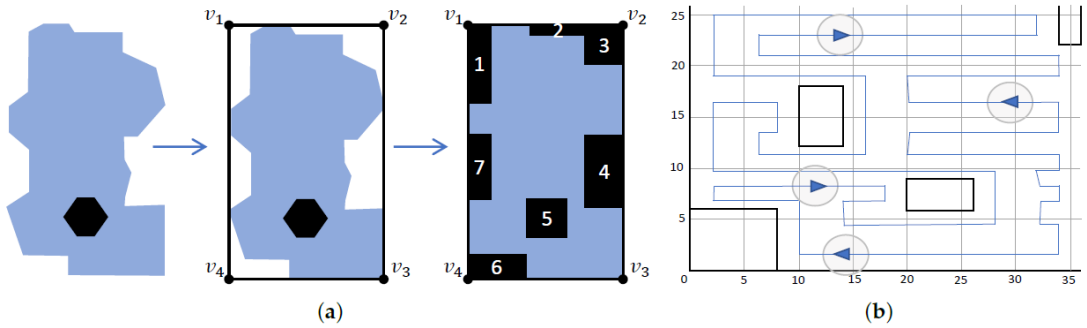


Figure 2.3: Persistent surveillance using a single search path through all locations of interest while avoiding obstacles (in black) [13]

As previously mentioned, some approaches implement heterogeneous UAV capabilities by varying their velocities to assign them to tasks [1]. Another aspect is to constrain surveillance to the operational limits of the UAVs by considering endurance, manoeuvrability and payload restrictions [13, 52].
Swarm intelligence based methods are bio-inspired and build upon natural intelligence fundamentals such as behavioural autonomy, social interaction, and learning. These approaches are pheromone-based in order to guide UAVs through grid-discretised target areas. These pheromones can be left at a given location, can evaporate over time and can be propagated to the neighbourhood [27]. Low computational cost is an important advantage of swarm intelligence, since this allows energy to be used for movement rather than for decision-making [13]. Different types of pheromones can be used to differentiate between information types and was proposed in [61]. It was used to assign individual UAVs to partitions of the target area and to implement collision avoidance. Their method involved a discrete search through implementation of a pheromone to prevent revisits. A similar approach using a central coordinator was implemented in [71] aiming to minimise communications among UAVs. Albani, D. et al. also focused on discrete surveillance using swarm intelligence, but added the possibility to prioritise locations within the target area [4]. In [21], an ant colony optimisation based algorithm to cover discrete searches in a fully distributed manner was proposed. Distributed pheromone maps were stored locally such that inter UAV communications were required to synchronise these maps when within communication range. Nigam, N. et al. extended upon this by considering the persistent surveillance problem and incorporating UAV dynamics to limit the turn radius [58]. Such a distributed approach was shown to be useful in identifying and following wildfires utilising deep reinforcement learning to help UAVs autonomously decide on their next action based on observations from the environment [38]. The need to incorporate UAV kinematics was acknowledged in [36], since this may influence surveillance performance as well. The constraint of limited communication capabilities in a swarm based approach was introduced by Howden, D. such that distributed pheromone maps needed to be synchronised. This model was also capable of prioritising locations within the target area as explained in further detail in section 2.1.1 [31].

The vast target areas considered in this research, combined with the need for robust and reliable systems [57], call for distributed persistent surveillance algorithms. Implicit or locally explicit communication, rather than explicit coordination, is preferred since data transfer over vast areas and/or rough terrain may be difficult. Realistic persistent surveillance approaches must be robust and able to operate in the presence of limited communication capabilities [35, 63]. Such decentralised approaches are also advantageous over centralised systems in terms of adaptability, simplicity, modularity, reliability, robustness and low communication bandwidth requirements. Approaches based on multi-agent planning, MILP optimisation, space decomposition and evolutionary algorithms are generally based on a centralised coordination technique. Algorithms utilising potential fields, swarm intelligence and behaviour-based methods often use a distributed approach [57]. Therefore, the most relevant distributed approaches to persistent surveillance are discussed in greater detail.

2.1.1. Prioritised surveillance using distributed pheromone maps

The need for scalability, robustness and the limitations w.r.t. communications have shifted studies from central systems to autonomous robot swarms. The approach proposed by Howden, D. [31] differentiates from other surveillance solutions by considering the persistent surveillance problem from a pheromone control

based approach. The basis lies with limited communications and the use of digital pheromone maps to model environmental awareness [31]. The approach builds on the work of Erignac, C. where surveillance was limited to a single sweep of the target area and did not prioritise certain areas [21]. A challenge of pheromone based models was to overcome agents getting stuck in local minima. This was overcome by introducing euclidean distances into the heuristic motion planning, which has similar properties to using a greedy hill descent approach [31].

The proposed algorithm utilises distributed time-priority product based pheromone maps to guide the UAVs. Pheromone concentrations are the product of a grid cell's priority priority_C and the time passed since its last visit Δt from a UAV as computed from Equation 2.1 [31].

$$p_C = \text{priority}_C \cdot \Delta t \tag{2.1}$$

Each UAV at location P has access to a local pheromone map where each grid cell C contains two values, the first being the time since a grid cell was last visited and the second being a priority level. These priority levels are to be obtained from knowledge regarding the environment, such as historic data, experience or from expert knowledge. UAVs prefer to move to grid cells with high pheromone concentrations p_C. After visitation by a UAV, the pheromone concentration is reset to zero such that persistent surveillance is incorporated [31].

$$h = \frac{p_C^2}{d(C,P) + d(C,P+\overline{r})} \tag{2.2}$$

The path planning is performed by the A* algorithm, where the pheromone levels are used as weights. The target location for the A* algorithm is obtained from the heuristic h in Equation 2.2, which is computed for all cells in the pheromone map. The distance d in the denominator is used to overcome local minima such that UAVs are motivated to explore the target area. This heuristic h also considers collision avoidance through repulsive forces $\overline{r}$, as this poses an additional penalty to the heuristic h. Synchronisation of the pheromone maps among UAVs is performed when within a pre-specified communication range. This enables global environmental awareness for each UAV as synchronising the pheromone maps indirectly specifies the paths flown by other UAVs. The model outperforms previous work w.r.t. computational load and also improves in the area of revisiting times, which makes it an interesting model within this context [31]. Especially when considering that the UAVs are required to perform computations online, thus asking for quick decision making and motion control. Despite the advances, the presented solution method does not consider dynamically adapting the priority levels. Due to its deterministic motion planning the approach's effectivity is expected to be limited when being applied in real world poacher scenarios that can coordinate and adapt to surveillance strategies. Besides, the approach does not account for operational constraints introduced through the problem characteristics of persistent surveillance.

Howden, D. et al. expanded upon their previous work in [31] by considering the need to detect and track fires [32]. The model builds on Equations 2.1 and 2.2 and each UAV increases the pheromone levels of its own unique pheromone map automatically by an amount proportional to the urgency of surveillance. These internal maps are shared for a given duration at fixed intervals, such that neighbouring UAVs can synchronise their own maps. Synchronisation of the pheromone maps is achieved through comparing and adapting the lowest of local and received pheromone concentrations at a given grid cell C. Convergence of the UAVs is achieved through the introduction of repulsive forces, resulting in spreading of the UAVs. This accomplishes that multiple UAVs do not simultaneously visit the same grid cells and prevents collisions. Again, the pheromone level at a grid cell's location is reset to zero once a node is visited by a UAV, such that other locations become the preferred new objective [32].

In addition to the work in [31], each UAV in [32] carries a sensor to detect phenomena. The authors mention the need to scale grid cells and match their dimensions to the sensor's footprint to minimise overlapping UAV trajectories. Howden, D. et al. propose to dynamically adjust priority levels based on detected phenomena [32]. Grid cells nearby detected phenomena are given higher priority priority_C. The resultant behaviour is that the UAVs moving in the surrounding area adjust pheromone levels such that the fire front is drawn. Due to the increased priority level, other UAVs also tend to converge to these locations, effectively resulting in collaborative fire front tracking. The weights of this high level priority is to be tuned very carefully however, because it prevents the UAVs from exploring other areas of the target area to detect other targets. In such a

case, the UAVs discontinue searching for targets until the current target ceases to exist [32]. Such an approach could be used to detect collaborating poachers that may be spread over a subsection of the target area, or to find poachers once a trail is found by the UAVs.

2.1.2. Stochastic surveillance using Ant Colony Optimisation

Similar to the work in [31, 32], Bontzorlos, H. et al. also propose an ant colony optimisation method to solve the persistent area surveillance problem [9]. The need for robust systems in terms of equipment failure is emphasised, thus imposing indirect communication among UAVs in order to coordinate. This is justified through the example of component failure in systems relying on central coordination which results in the whole system failing to adapt. A system based on physical pheromones is proposed such that indirect communications can be achieved in order to design fault tolerant systems. This allows for minimal hardware requirements in terms of memory, since the physical pheromones act as a collective memory. The pheromones contain information regarding the most recent visit of a certain location, as well as what UAV monitored the location, along with the current distance of the previously visiting UAV from the location [9]. Given the context of surveillance within this research, such physical pheromones are prone to errors. In dynamic environments where wildlife and intruders continuously scout the area, the pheromones may be moved which results in inaccurate pheromone concentrations. Such a physical system is specially prone to intruders that aim to sabotage the surveillance system in order to stay clear of local authorities. Besides, the environmental factors may prove it difficult to sense pheromones dense nature. Nevertheless, this can be adapted to digital pheromone maps similar to the proposed model in [6].

The simulation space is described by a graph $G(V,E)$ such that the environment is divided into a set of grid cells $V = v_{ij} \mid i,j = 1,2,...,n$ where a set of autonomous robots $E = e_{ij} \mid i,j = 1,2,...,m$ monitors the space. The target area is subdivided into two groups A and A' such that monitored locations are grouped apart from locations that have not been covered by the UAVs. The objective of the UAVs is to maximise the coverage C of the target space, which is defined as the percentage of locations that have been monitored within the entire target area. The UAVs pursue locations where pheromone concentrations are low and increase the pheromone level after monitoring such a location. Minimum and maximum pheromone concentrations are introduced as well, such that pheromone concentrations $\tau \in [0,1]$. Upon monitoring a location, the pheromone levels are incremented by a concentration of 1 unit, such that the maximum pheromone level is directly reached and revisits are prohibited. This method is relaxed through the introduction of unique pheromone types for each individual UAV, such that location (i,j) can contain multiple types of pheromone. This relaxation also allows for revisiting locations and UAVs are assumed to identify other UAVs based on deposited pheromones [9].

Evaporation of pheromone levels is included through a steady decrease of α at each time step. Its value was chosen to be a percentage of the deposited pheromone concentrations from UAVs, such that the resulting relative change of concentration is much larger than 1 when a location is being monitored. Larger evaporation rates were found to reduce the coverage performance as robots tend to remain in a given area and continuously revisit a select number of locations [9]. This behaviour is also based on the implemented allowable range of the pheromone concentrations and the aim to move to lower pheromone level locations. Since negative pheromone concentrations do not make sense physically, switching the UAV objective around could results in better coverage. Such an alternative objective could be to pursue high pheromone concentrations instead. This would require neglecting the maximum pheromone concentration such that UAVs notice relatively large concentrations and remain from revisiting a select number of locations.

Motion planning of the robots, given the absence of locally accessible pheromone maps, involves detecting pheromone concentrations of neighbouring locations D. Each robot is therefore capable of moving to one such neighbouring location s at each time step, given its current location. Combining this knowledge with the pheromone evaporation rate, the robots are also capable of computing the current distance of other robots from the neighbouring locations based on the residual pheromone concentrations. Due to the limited field of view and the lack of information regarding the other robot's exact route, only the maximum distance can be computed. Motion planning uses the pheromone type $\tau(i)$ that has the largest concentration at a neighbouring location $i \in D$. Stochastic path planning is implemented according to Equation 2.3, where the pheromone type of the highest concentration for each neighbouring location is multiplied with a randomly drawn value $q \in [0.9,1]$. A limitation here is the lack of information if another UAV has not visited a given location. In such

a case it is impossible to calculate the maximum distance to that UAV due to the lack of pheromones, which can result in multiple UAVs moving towards the same location simultaneously. This can be generalised as the computed maximum range to other UAVs is not considered in decision making, thus not ensuring collision avoidance.

$$s = arg\min_{i \in D}(q \cdot \tau(i)) \tag{2.3}$$

2.1.3. Bayesian path planning to perform prioritised, stochastic surveillance

The surveillance problem is a topic within police patrol as well. Approaches to police patrols cover a great variety of contexts and cultures, although one approach appears to be promising when considering effectiveness. Hotspot, or place-based patrolling in which police officers randomly monitor has shown to be effective. Designing patrol routes is also an important aspect besides the need to identify the location of high criminal activity. Challenges in designing optimal patrol routes arise, specially when considering multiple agents in environments with limited resources. These include the need to repetitively and regularly monitor locations within the target area, while responding to emergency situations. This poses requirements w.r.t. the robustness of the surveillance system such that its effectiveness does not degrade significantly when solving such unexpected situations [17]. As identified in [31, 32], Chen, H. et al. also emphasise the importance of monitoring certain locations within the target area with different levels of priority [16]. In doing so, the system is to ensure coverage of the entire target area while introducing stochastic patrol routes to prevent criminals from predicting patrol behaviour. The problem is very similar to contexts where robots are used to monitor target areas [17].

A routing strategy using heuristics and Bayesian techniques, based on the Bayesian Ant-based Patrol Strategy (BAPS) is introduced. It aims to provide solutions in terms of efficiency, flexibility, unpredictability, scalability and robustness. The model uses a simplified version of the real world by considering a road network in an urban area to be the target area. Certain road segments are identified as hotspots through crime mapping, such that an initial crime activity map is constructed. The n hotspots are denoted as the set $H = h_1, h_2, ..., h_n$. Robots each have full knowledge of the environment w.r.t. the priority locations and their pheromone levels, such that a shortest path problem can be solved to construct a patrol route. BAPS employs pheromones at target location and specifies the evaporation rates proportional to the importance of a hotspot. In this approach, robots look to move towards locations with lower pheromone concentrations. When a hotspot h_i is visited by a robot at time t, the pheromone level is increased by an amount of pheromones $Phe_{Dep}(h_i)$ through Equation 2.4 [17].

$$Phe(h_i, t) = Phe(h_i, t-1) + Phe_{Dep}(h_i) \tag{2.4}$$

The pheromone concentrations decay in an exponential manner as time progresses based on the decay rate $\lambda(h_i) \in (0,1)$, see Equation 2.5 [17]. These decay rates were used to prioritise in earlier work [16], where its effectiveness was analysed as well. The simulations indeed showed reduced average idleness times for locations with relatively higher weights, proving that prioritisation using pheromone levels can be an effective means to steer surveillance in certain directions within the target area.

$$Phe(h_i, t) = Phe(h_i, t_0) \cdot \lambda(h_i)^{t-t_0} \tag{2.5}$$

Stochastic patrol routes are implemented using Bayes theorem to compute the posterior probability of patrolling a hotspot $P(patrol(h_i) \mid G(h_i), S(h_i))$ by applying the decision n independent times. The next hotspot to be monitored h_{next} is the one with the highest posterior probability as defined by Equation 2.6, where normalisation is omitted for simplification as it does not influence the results [17].

$$h_{next} = arg\max_{h_i} P(patrol(h_i) \mid G(h_i), S(h_i)) \tag{2.6}$$

Where $G(h_i)$ is defined as the gain of patrolling a hotspot and is computed according to Equation 2.7. It is therefore inversely proportional to the product of the pheromones and the normalised distance from the robot's location p to the hotspot. Since robots want to move towards locations with low pheromone levels, it can be seen that hotspots with a lower priority and a larger distance from the robot have lower gain. Normalised distances are used to prevent local optima resulting in repeatedly patrolling a small set of hotspots. The number of patrollers that is going to monitor a hotspot is denoted by $S(h_i)$. After omitting terms and

ignoring the normalisation factors for simplification without losing generalisation, the posterior probability itself is computed according to Equation 2.8. In Equation 2.8, L and M can be adapted to tune the probability values of zero gain and the gain saturation respectively, m is the number of robots patrolling. The approach favours robots monitoring the same hotspots through the application of $S(h_i)$ [17].

$$G(h_i) = \frac{1}{Phe(h_i,t) \cdot NormDist(p,h_i)} \tag{2.7}$$

$$P(patrol(h_i) \mid G(h_i), S(h_i)) = \frac{1}{M} \cdot \ln(\frac{1}{L}) \cdot e^{\ln(\frac{1}{L}) \cdot \frac{G(h_i)}{M}} \cdot \frac{2^{m-1+S(h_i)}}{2^m - 1} \tag{2.8}$$

As seen in Equation 2.6 and similar to the work in [9], the BAPS model is a greedy strategy as it determines the optimal choice of patrol target at each time step only. The advantage from this approach is that the model is less computationally heavy relative to models searching for the global optimum solution. This also allows for real time updates, thus being able to include the effects of a dynamic environment. The model implements several techniques that allow distributed coordination, although the authors propose a control centre to assign tasks and adapt to emergencies (see Figure 2.4), thus requiring continuous communications. Although not explicitly mentioned, the approach could also include dynamic updating of the decay rates. Such an approach would need some method to dynamically estimate these values however. Currently, the model only uses predefined criminal activity to define the relative weights of hotspots [17].

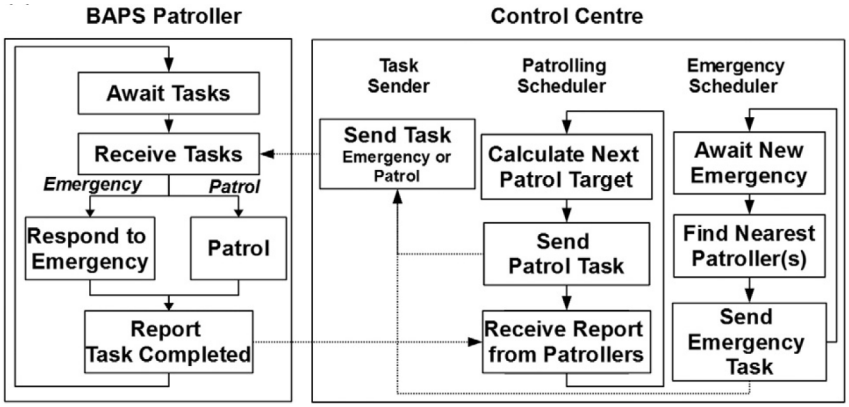


Figure 2.4: BAPS flowchart [17]

2.1.4. Prioritised surveillance using adaptive task thresholds

This paper introduces a dynamic ant colony's labour division (DACLD) model that incorporates self-organisation and flexibility under dynamic environments. The model is based on the classic fixed response threshold model (FRTM) and holds distributed characteristics. The objective is to maximise the total reward obtained by attacking various targets [89].

The DACLD model is built upon the idea that individual agents can perform a variety of tasks, each task having their own payout. Each swarm agent i has a stimulus value corresponding to a certain task, which determines the response of an agent w.r.t that task. The agents also have a response threshold θ_i that corresponds to a given task. Once this threshold is crossed by the stimulus of the respective task, the agent may start the task in question depending on the transition probability. These stimuli of each task degrade and increase over time (see Equation 2.9), therefore controlling what tasks are to be performed. The stimuli are increased at each time step by a fixed amount δ. Simultaneously, the stimuli levels are reduced if agents are performing a given task. This rate of stimuli reduction is proportional to the product of work efficiency φ and the number of agents n_{act} performing given task [89].

$$s(t+1) = s(t) + \delta - \varphi \cdot n_{act} \tag{2.9}$$

The choice of performing a task is obtained from the probability of transitioning from an idle state to an active task as computed from Equation 2.10. This depends on the stimuli level s and the stimuli threshold for the given task [89].

$$P\left(ST_i = 0 \rightarrow ST_i = 1\right) = \frac{s^n}{s^n + \theta_i^n} \tag{2.10}$$

Where n is a constant controlling the curve of the threshold function. The probability of an agent deciding to quit from an active task is a predefined constant [89]. Given the context of this research, the tasks can be translated to grid cells. This would simulate surveillance, although the probability of stopping a task is not useful within such a context. Based on this approach the model does not support collision avoidance however. It does allow for distributed and flexible behaviour through the implementation of distributed stimuli maps as suggested in [31]. Wu, H. et al. further extend upon Equations 2.9 and 2.10 by recognising the need to incorporate more complex environments. This includes the need for differentiation among task types, multiple agent states, operational constraints, heterogeneous agent capabilities and dynamically adjusted thresholds based on gained experience. Thus presenting the Dynamic Ant Colony's Labor Division model [89].

The DACLD model implements a new stimuli, s_j^k, update rule as expressed by Equation 2.11. A task k is either surveillance or attack and j is a target (or grid cell). This stimuli update rule has not changed in essence as it still incorporates a constant increase and a reduction for each agent i proportional to its work efficiency. The stimuli thresholds θ_{ij}^k have become dynamic through implementation of learning and forgetting factors. The learning rate $\zeta_i^k(t)$ of agents combined with the successful completion of tasks reduces the response threshold of tasks according to Equation 2.12, such that the swarm exerts emergent behaviour by gradually improving the success rate of tasks. The learning factor is purely related to the accumulation of experience, N_k, or the number of times task k has been performed in a time interval between t and T. In order to prevent dividing by zero, the initial learning factor $Stu \in (0,1)$ is implemented. A reduction of the learning factor results in a lower task threshold as can be seen in Equation 2.13, where $k = k^*$ indicates that the same task is to be performed sequentially in order to learn. If different tasks are performed sequentially, then the agent can forget about a given task. Thus resulting in a higher threshold due to the multiplication of the threshold with a forgetting factor $\varphi_i^k > 1$. The dynamic threshold is also updated by considering resource consumption in a similar fashion, such that having zero resources available to perform a task results in an infinitely large threshold level [89].

$$s_j^k(t+1) = s_j^k(t) + \delta_j^k - \sum_{i=1}^{N_U}\left[\partial_{ij}^k \cdot x_{i,j}^k(t)\right] \tag{2.11}$$

$$\zeta_i^k(t) = \begin{cases} \frac{1}{N_k(t-T)} \cdot Stu, & N_k(t-T) \neq 0 \\ Stu, & N_k(t-T) = 0 \end{cases} \tag{2.12}$$

$$\theta_{ij}^k(t+1) = \begin{cases} \zeta_i^k \cdot \theta_{ij}^k(t), & k = k^* \\ \varphi_i^k \cdot \theta_{ij}^k(t), & k \neq k^* \end{cases} \tag{2.13}$$

$$\begin{aligned} P\left[ST_i(t) = \alpha \rightarrow ST_i(t+T) = \beta\right] &= P\left[x_{i,j}^k(t) = 1 \rightarrow x_{i,j^*}^{k^*}(t+T) = 1\right] \\ &= \frac{\left[\widetilde{S}_{\alpha\beta}\right]^n}{\left[\widetilde{S}_{\alpha\beta}\right]^n + \left[\widetilde{\theta}_{\alpha\beta}\right]^n + \rho \cdot \left[\Delta\tau_i\left(j \rightarrow j^*\right)\right]^n} \end{aligned} \tag{2.14}$$

State transition probability computations are adapted to account for the possibility of transitioning between various possible states, see Equation 2.14. In doing so, the authors propose the use of a delay penalty coefficient ρ, such that state transition probabilities are reduced if the response time $\Delta\tau_i\left(j \rightarrow j^*\right)$ of a given agent is relatively large. As seen, state transition probabilities are increased if the relative stimuli $\widetilde{S}$ of a state β over state α are large, whereas large relative thresholds $\widetilde{\theta}$ reduce the probability of transitioning. This implies that each UAV only transitions to states that are relatively important to be performed, while taking the likelihood of successfully finishing the task into account. This method ensures efficient task solving by implicit coordination and assigning UAVs to tasks of importance. Such an approach is therefore very useful when heterogeneous UAVs are considered. Collaboration among agents is incorporated through communication of the task thresholds. In case an agent is capable of performing a given task, the stimuli to other agents will rapidly decrease. The inverse holds when the agent is unable to successfully finish a given task, thus drawing other agents to this task such that it can be performed. Despite this mechanism, it does not ensure collision

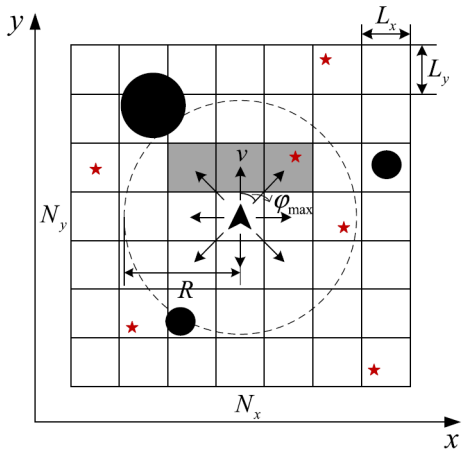


Figure 2.5: HAPF-ACO discretised mission area [97]

avoidance among UAVs. Performance of the DACLD model is compared to a Wolf Pack Algorithm and it is shown that DACLD responds more quickly to unexpected threats. A limitation here is that knowledge of these threats which includes their respective locations and interest values, is known on beforehand [89]. Its performance in dynamic environments where targets can emerge and move through the environment without the UAVs being aware of this, is questionable.

Although the communication protocol not being mentioned explicitly, the approach is relatively easy applied in contexts with limited communication capabilities. Such an approach does introduce latencies since the distance among UAVs may of of such extent that communications must wait until possible. The main contributions of the paper are related to addressing the task allocation problem in dynamic environments, as well as improving on the dynamic response threshold, dynamic environmental stimuli, multi-type tasks and multi-state transition probabilities [89].

2.1.5. Hybrid swarm based surveillance

The proposed distributed surveillance model from Zhen, Z. et al. [97] combines artificial potential fields with an ant colony optimisation method to persistently search for and attack intruders. The Hybrid Artificial Potential Field and Ant Colony Optimisation (HAPF-ACO) algorithm utilises an APF map, as well as pheromone maps. The APF consists of target attraction forces and repulsive forces, also enabling collision avoidance path planning by introducing repulsive fields among UAVs. The target attraction forces are designed according to a Target Probability Map (TPM), as discussed below. The UAV motion planning incorporates collision and obstacle avoidance through the repulsive fields. Manoeuvrability constraints are considered through limitation of the maximum turning angle ϕ_{max} of the UAVs [97]. Figure 2.5 shows that the target area D of size $N_x x N_y$ is discretised into equally sized $L_x x L_y$ grid cells. The UAV is capable of observing its direct environment within a radius R from its own current location. Thus making it possible to detect the targets (visualised in red stars) and threats (visualised by black circles) within the radius R. Considering the UAV's maximum turning angle ϕ_{max} and its speed v such that the UAV can move forward exactly one grid cell, the possible UAV motions are visualised by the grey grid cells. The approach therefore is not limited to rotary or fixed-wings UAVs and is applicable in a broad scala of practical implementations.

The authors also emphasise the need to account for high communication hardware and bandwidth requirements to enable effective coordination among UAVs. The density of the swarm distribution greatly affects the required communication bandwidth as this scales exponentially with the number of communicating neighbours [27]. Besides incorporating a fully distributed approach, the cooperative radius r_i^{coor} is introduced and adjusted accordingly in order to circumvent this hardware requirement. Communication with other UAVs is only possible if the euclidean distance is smaller than this cooperative radius, such that the set of interactive neighbours is defined according to Equation 2.15 [97]:

$$N_i^c = \left\{ j \mid \left\| x_i - x_j \right\| \leq r_i^{coor}, j \in \{1,2,\cdots,N_V\}, j \neq i \right\} \tag{2.15}$$

Where N_V are the number of UAVs in the swarm. Using the set of coordinating neighbours N_i^c of UAV i and the expected number of neighbours N_{topo}, the coordination range is adjusted to achieve a constant number of neighbours (see Equation 2.16) [97].

$$\dot{r}_i^{coor} = k_{adjust} \cdot R_{sen} \cdot \left(1 - \left|N_i^c\right| / N_{topo}\right) \tag{2.16}$$

Where $0 < k_{adjust} < 1$ is the adjusting rate and R_{sen} is the upper bound of the communication range. This equation shows that if more neighbours are present than expected, the coordination range r_i^{coor} is reduced proportionally to k_{adjust}. In doing so, this method does not consider the different information that each neighbours has to offer.

The objective of the UAV swarm is to maximise the sum U^* of attack and search payouts (J_T and J_E respectively), while adhering to manoeuvrability, collision avoidance, obstacle avoidance and range constraints. The distributed swarm system characteristics force the decomposition of these objectives to local indicators on a unique UAV level, see Equation 2.17 [97]:

$$\begin{cases} J_T = \sum_{i=1}^{N_V} \mu_i J_{Ti} \\ J_E = \sum_{i=1}^{N_V} \mu_i J_{Ei} \end{cases} \tag{2.17}$$

Where μ_i can be used to differentiate among UAV importance w.r.t. the mission objective.

Artificial Potential Field

TPMs are used to describe targets and to prioritise areas within the discretised target area D. Each UAV considers its local TPM in its motion planning, based on the target existence probability $p_{mn}^i(k) \in [0,1]$ of grid cell (m,n) as obtained by the i-th UAV at time k. Each UAV therefore has access to a unique TPM of the target area stored in a local probability distribution matrix [97]:

$$\text{TPM}_i(k) = \left\{ p_{mn}^i(k) \mid m = 1,2,\cdots,N_x, n = 1,2,\cdots,N_y \right\} \tag{2.18}$$

Detection of targets is possible within the radius R (see Figure 2.5) and this information is used to dynamically update the TPM. This radius R is based on the sensor detection range. Uncertainty is introduced to model erroneous sensor observations through the implementation of the probability of detection P_D and the false detection probability P_F. The probability of detecting a target, given that the sensor field of view (FOV) covers exactly one grid cell is obtained from Bayesian probability formulas in Equation 2.19 [97].

$$p_{mn}^i(k+1) = \begin{cases} \tau\, p_{mn}^i(k) & \text{(not detected)} \\ \frac{P_D \cdot p_{mn}^i(k)}{P_F + (P_D - P_F) \cdot p_{mn}^i(k)} & \text{(detected and } b(k) = 1\text{)} \\ \frac{(1-P_D) \cdot p_{mn}^i(k)}{1 - P_F + (P_F - P_D) \cdot p_{mn}^i(k)} & \text{(detected and } b(k) = 0\text{)} \end{cases} \tag{2.19}$$

Where b_k indicates the detection result. $b_k = 1$ indicates that a target is present within the sensor detection range. The attenuation coefficient $\tau \in [0,1]$ describes the dynamic environment w.r.t. the targets in grid cells not being covered by the UAV at time k.

The real time TPM of a UAV i is used to generate the target attraction field that stimulates the UAV to move toward the targets. The attraction force $F_{att,i}$ is proportional to the target existence probability and obtained from the gradient of the target attraction field according to Equation 2.20. Similarly, threat repulsive forces $F_{rep,i}^T$ can be used to move UAVs away from no-fly zones or other areas and/or obstacles. This approach is also used to incorporate collision avoidance among neighbouring UAVs (See Equation 2.21) and is valuable due to the distributed method presented. The collision avoidance mechanism can be tuned using the b and c parameters which influence the magnitude and the rate of change of the repulsive field respectively. Due to the limited range of communication, only neighbouring UAVs are considered as possible threats. The magnitude of the repulsive forces mainly depends on position vector x_{ij} of the j-th UAV w.r.t. the i-th UAV, where $\hat{x}_{ij}$ is the unit position vector [97].

$$F_{att,i}(X_i) = \nabla U_{att,i}(X_i) = \nabla k_{att} \cdot TPM_i \tag{2.20}$$

$$F^V_{rep,i}(X_i) = -\nabla U^V_{rep,i}(X_i) = \sum_{j \in N_i^c} \frac{b}{c} \cdot \frac{e^{\frac{\|x_{ij}\|}{c}} \cdot \hat{x}_{ij}}{\left(e^{\frac{\|x_{ij}\|}{c}} - \frac{e^{\|x_{ij}\|_{min}}}{c}\right)^2} \tag{2.21}$$

Ant Colony Optimisation

The model is a hybrid of artificial potential fields and ant colony optimisation. The pheromone maps used by each UAV enable task cooperation among the UAVs in terms of persistent surveillance. The pheromone update mechanisms to be integrated with the discussed artificial forces are provided below.

Each grid cell contains a certain concentration of pheromones, which indicate the attraction of the specific grid cell to the UAV. Higher concentrations attract the UAVs. Together, the concentration values of all grid cells at a certain time k in the target area D constitute the local pheromone structure $\tau_i(k)$ (see Equation 2.22) of a given UAV i. The pheromone concentration levels are limited by an upper and a lower limit.

$$\tau^i(k) = \{\tau^i_{(x,y)}(k) \mid \quad x = 1,2,\cdots,N_x, \quad y = 1,2,\cdots,N_y\} \tag{2.22}$$

The ACO aims to prevent excessive search of certain areas within the target area by implementing a update mechanism that relies on historical state information from other UAVs. The communication range limits the number of UAVs that can share information with each other, thus resulting in regional search coordination. This is achieved through a given UAV i obtaining state information of other UAVs j at time k as stated in Equation 2.23. The historical positions and motion direction data obtained by UAV i from UAV j that is part of the neighbouring UAVs of UAV i is used to predict the state of the swarm at time k. Due to the limited communication range and the target area's dimensions, it is possible that the UAV may not have up to date information of all UAVs. Thus emphasising the importance of predicting the swarm state to obtain accurate pheromone levels. The pheromone levels are updated accordingly by considering the presence of the i-th UAV itself and the neighbours, see Equation 2.24

$$\begin{cases} Info(k) = \{Info_{j,i}(kj,i), V_j \in V\} \\ Info_{j,i}(k_{j,l}) = \{x_{j,i}(k_{j,i}), y_{j,i}(k_{j,i}), \psi_{j,i}(k_{j,i})\} \end{cases}, \quad k_{j,i} \le k \tag{2.23}$$

$$\begin{cases} \tau^i_{(x,y)}(k+1) = \tau^i_{(x,y)}(k) - \Delta\tau^i_{l(x,y)}(k) \\ \Delta\tau^i_{l(x,y)}(k) = \sum_{j \in N_i^c} \Delta\tau^{(i,j)}_{l(x,y)}(k) \end{cases} \tag{2.24}$$

$$\Delta\tau^{(i,j)}_{l(x,y)}(k) = \begin{cases} \Delta\tau_{l_0} \times \frac{R^4 - d^4\left((x,y),\left(x^*_{j,i}(k), y^*_{j,i}(k)\right)\right)}{R^4}, & d\left((x,y),\left(x^*_{j,i}(k), y^*_{j,i}(k)\right)\right) \le R \\ 0, & d\left((x,y),\left(x^*_{j,i}(k), y^*_{j,i}(k)\right)\right) > R \end{cases} \tag{2.25}$$

Persistent surveillance is implemented using the ACO algorithm by including a global pheromone update rule defined in Equation 2.26. This is used to account for the possibilities of new targets appearing in areas that have previously been searched. The rule can be tuned using the environmental uncertainty factor F and also depends on the update step τ_{g_0} [97]. Larger values can be used in more dynamic environments as this encourages UAVs to revisit grid cells. A limitation of the model is the use of a maximum pheromone concentration. Such a maximum is required to be tuned in accordance with the coverage rate of the UAVs, the dimensions of the environment and the pheromone update rate. If not done correctly, this may result in various areas reaching the maximum level as the UAVs are not quick enough in covering the target area. Thus, diminishing any prioritisation from a surveillance perspective through the fact that each area now has the same pheromone concentration.

$$\tau^i_{(x,y)}(k+1) = \tau^i_{(x,y)}(k) + F \times \Delta\tau_{g_0} \tag{2.26}$$

Integration

The APF side of the algorithm mainly focuses on target and obstacle detection in order to find the intruders and avoid collisions (among UAVs). This also incorporates the prioritisation of certain areas through the use of target existence probabilities, although limited. The ACO features introduce the persistence property required within this research's context to account for the persistent attacks by poachers in national parks. Integration of these two different aspects therefore determines the effort put into the various tasks such as persistent surveillance and prioritisation among areas (including collision avoidance). This is achieved by

formulating motion planning transition rules. One downside of integrating the APF with the ACO is the lack of collision avoidance due to the UAV considering various features such that the repulsive force of a threat is neglected. The state transition rule being used at a given location and time depends on the value of $\lambda \cdot F/F_{max}$. Where λ is the environmental perception factor, F and F_{max} are the magnitude of the potential field force of the current grid cell the UAV is located at and the maximum force within the detection range of the UAV. This algorithm therefore adjusts its probability of choosing a transition rule according to the need to act upon nearby threats. A deterministic transition rule that depends on the remaining range of the UAV is used if a random number q is smaller than $\lambda \cdot F/F_{max}$, as shown in Equation 2.27. The transition rule either minimises the change of heading or aims to minimise its movement such that the UAV maximises its endurance, depending on the value of ω_1 and ω_2 and $\omega_1+\omega_2 = 1$. The UAV aims to minimise heading changes only if the travelled distance is less than half its maximum range, i.e. $\omega_1 = 1$ [97].

$$s_j = \omega_1 \times \arg\min_{j\in\Omega}\{\theta_j\} + \omega_2 \times \arg\min_{j\in\Omega}\left(\left|L_{max} - L_{past}(k+1) - D_{left}(k+1)\right|\right), \quad q < \frac{\lambda F}{F_{max}} \tag{2.27}$$

The heuristic transition rule is formulated in Equation 2.28. This decision process either uses a stochastic decision process or a heuristic optimisation step, depending on the random numbers q_0 and q_1 drawn from a standard uniform distribution. If $q_1 \leq q_0$, then the motion transition aims to optimise for optimal search and to maximise endurance. Optimal search is defined as the maximum product of the pheromone concentration and the heuristic information η_{ij} when moving from grid cell s_i to s_j. The heuristic information contains the system performance in terms of coverage of the target area such that unsearched grid cells are chosen with the highest probability. Stochastic decision making is enforced otherwise and according to the transition probabilities obtained from Equation 2.29. The transition probabilities are computed for each grid cell s_j within the set of grid cell options Ω obtained by considering the UAV's position, heading and maximum turning angle [97].

$$s_j = \begin{cases} \omega_1 \times \arg\max_{j\in\Omega}\left\{\left[\tau_{ij}^{\alpha}\right] \times \left[\eta_{ij}^{\beta}\right]\right\} \\ \quad +\omega_2 \times \arg\min_{j\in\Omega}\left(\left|L_{\max} - L_{\text{past}}\,(k+1) - D_{\text{left}}\,(k+1)\right|\right), & q_1 \leq q_0 \\ S, & \text{other} \end{cases} \tag{2.28}$$

$$p_{ij}(k) = \begin{cases} \frac{[\tau_{ij}(k)]^{\alpha} \times [\eta_{ij}(k)]^{\beta}}{\sum_{s_j\in\Omega}[\tau_{ij}(k)]^{\alpha} \times [\eta_{ij}(k)]^{\beta}}, & s_j \in \Omega \\ 0, & s_j \notin \Omega \end{cases} \tag{2.29}$$

It can be noted that these motion planning rules do not include the APF and the TPMs previously discussed. Collision avoidance is therefore applied only when the benefit J_{Ti} at time k outweighs the benefit of search benefit F_{Ei} [97]. Surveillance is therefore not explicitly prioritised, although this is achieved through the APF indirectly.

Finally, the authors consider a variety of possible options regarding available information of dynamic targets. Based on the available information including initial positions, movement directions, magnitude and direction of speeds, the TPM's target existence probabilities are continuously updated according to conditional probability computations [97]. This approach does not consider other attributes about the targets' behavioural characteristics and does not incorporate a feedback loop regarding these predictions. Section 2.3 elaborates on more advanced techniques that can be used to predict target locations.

2.1.6. Surveillance performance indicators

A number of performance indicators have been developed through continued research to persistent surveillance. These are discussed below, since a subset is to be used for the to-be-performed sensitivity analysis as discussed in Chapter 4. Most metrics used in persistent surveillance are based on some time based metric such as age, idleness, latency or refresh time. Another approach is to optimise for visit frequency or to minimise the difference between the largest and lowest visit frequency. This last optimisation method is not of interest in priority based surveillance however [57]. A number of other performance indicators are explained below.

Global average idleness (GAI)

Persistent surveillance efficiency can be measured using the global average idleness performance indicator computed according to Equation 2.30. Efficiency should be optimised such that the time lag between

two consecutive visits to a given location is minimised. It is computed by summing the average idleness $AIdl(h_i,t)$ of all locations and dividing by the number of locations to obtain an overall average [17]. An alternative proposed in [13] is to measure the quadratic mean of the visit intervals (QMI).

$$GAI(t) = \frac{\sum_{i=1}^{n} AIdl(h_i,t)}{n} \tag{2.30}$$

Weighted global average idleness (WGAI)
The flexibility of a surveillance strategy is the ability to prioritise relatively important locations such that their visiting frequency is increased. This can be measured using the WGAI, which is a weighted version of the GAI indicator. The weights $W(h_i)$ assigned to the idleness of a locations correspond to the priority which in turn is dependent on the crime density level [17]. Rather than optimising for the WGAI, one can also optimise for the weighted refresh time. This is simply the longest idleness at a given location, weighted by the location's priority level [62].

$$WGAI(t) = \frac{\sum_{i=1}^{n} W(h_i) \cdot AIdl(h_i,t)}{\sum_{i=1}^{n} W(h_i)} \tag{2.31}$$

Scalability of Team Size (ST)
The ST indicates how performance is influenced due to a change in the number of robots participating in the persistent surveillance method. It is based on Balch's speedup measure which is a ratio of the GAI obtained using a single monitoring robots and a team of R robots. Since it is uncommon for a single patroller to monitor a large area, as would be done using a team of robots, the metric is adjusted to incorporate for this effect. The measure for scalability is therefore multiplied by a target area of size S that would be patrolled by the single patroller, see Equation 2.32 [17].

$$ST(R) = \frac{S \cdot GAI(S)}{R \cdot GAI(R)} \tag{2.32}$$

Spatial Scalability (SS)
Similar to the ST metric for team scalability, a metric for spatial scalability was proposed in [17]. It is based on the relative area covered by crime hotspots of a pre-defined crime density level L divided by the total area to be monitored. Computation is performed according to Equation 2.33 and indicates the relative change in GAI performance due to a change of hotspot density from the baseline level L_B [17].

$$SS(L_i) = \frac{GAI(L_i) - GAI(L_B)}{GAI(L_B)} \tag{2.33}$$

Average standard deviation of idleness (ASDI)
As mentioned, unpredictable monitoring strategies prevent intruders from easily learning and adapting to the implemented surveillance strategies. It should therefore be difficult to predict when a location is going to be monitored by law enforcement. Measuring this unpredictability can be achieved through computing the standard deviation of idleness SDI for each location h_i in the environment, after which an average can be computed from Equation 2.34 [13, 17].

$$ASDI(t) = \frac{\sum_{i=1}^{n} SDI(h_i,t)}{n} \tag{2.34}$$

Robustness to fluctuating team sizes
During persistent surveillance it can occur that robots need to respond to some sort of emergency such as recharging, such that the team size is changed online. The robustness of the surveillance method w.r.t. these online fluctuations can be measured by comparing the GAI when all robots are participating in the surveillance strategy with the GAI when a robot is active in an emergency. A relative idleness (RI) value is then computed according to Equation 2.35 [17].

$$RI_{GAI} = \frac{GAI_{emergency} - GAI}{GAI} \tag{2.35}$$

When the surveillance context becomes multidimensional through the introduction of detecting intruders, additional performance metrics can be defined. Such metrics can be cumulative based and include the target detection probability (PD) or the opposite, the probability of non-detection (PND). The expected, or the mean, time between threat detection (MTD) is an interesting metric to tune the rate of re-visiting locations

in the target area [27, 57, 66]. It is also possible to measure the number of detected events globally [13, 27], although a hit rate per location is used frequently as well [60]. Another efficiency related performance indicator is the catch per unit effort, defined as the number of detections of poaching activity divided by the amount of distance travelled [18].

2.2. Modelling poacher behaviour

The previous section collaborated on the surveillance aspect from various approaches. When monitoring a given target area, the objective is usually to identify unusual activity. Such two-sided search problems must incorporate the evasive behaviour of the targets that are to be found [66], although full knowledge of the dynamic target's decision making is not available. The decision making of rational entities has been tackled through Game Theory in various applications, including the US Coast Guard, Los Angeles Airport and the conservation of national parks [85]. To this end, allocation and optimisation of limited resources to effectively design patrol strategies has been studied through the development of GSGs. Such approaches are based on the repeated SSG in order to model the repeated attacks of intruders [23]. The need to consider multi-UAV approaches has lead to the development of complex optimisation problems, deeming a bottom-up approach to understanding the emergent combined behaviour of the poachers and the UAVs that easily incorporates local spatial communication between agents less complex [2]. Nevertheless, the work performed in the field of Game Theory has established models that effectively model the bounded human rationality in decision making. The following sections therefore elaborate on developed models in order to identify relevant techniques to model coordinating intruders.

2.2.1. Rationally bounded decision making

In 1995 the Quantal Response (QR) model was introduced by McKelvey, R. et al. [50] which uses standard statistical models for decision making of players. This model is an extension of well-developed and commonly used models of choice that are used in biological, pharmacological and social sciences. The statistical approach shifted the process of decision making from being deterministic to being probabilistic by introducing a quantal choice model. Better choices are more likely to occur relative to worse choices, but the best choice is not made with certainty. The probabilities introduce uncertainty in decision making, which translates into players not always deciding on the best strategy. The decision making is based on an individual's relative expected utility from a choice of strategy. These expected payoffs x_{ij} for strategy j are assumed to be obtained in an unbiased way by the individual i themselves. The probability σ of choosing a strategy j can be computed using the logistic quantal response function in Equation 2.36 [50].

$$\sigma_{ij}(x_i) = \frac{e^{\lambda x_{ij}}}{\sum_{k=1}^{J_i} e^{\lambda x_{ik}}} \tag{2.36}$$

Where J^i is the set of pure strategies of individual i. The behaviour of the individual can be tuned using the estimated λ parameter, which defines the magnitude of erroneous decision making. A λ equal to 0 results in continuous error making decisions, whereas an infinitely large λ results in no error making.

The model was tested in the context of game theory in order to model individual human behaviour and also incorporated a learning aspect. Although not elaborated on in detail, it was stated to be based on the gain of experience from observing the real payoffs of chosen strategies after performing them and comparing these values to the expected payoffs [50].

The QR model was adjusted to incorporate bounded rationality, which means that a person may not be aware of all its surroundings or may be influenced in it's decision making by the interplay between emotion and cognition [55]. The work of Nguyen, T. et al. [55] extend upon this and part from the discussed QR model, despite its significant support from literature in modelling human decision making [55, 96]. The assumption used in the QR model which states that human stochastic response is purely based on the expected payoff is relaxed. Instead, a subjective utility function (SU) is proposed defining the SUQR model. This approach uses a linear combination of the reward, penalty and the marginal coverage on target to compute the probability σ of choosing a strategy j using Equation 2.37 [55].

$$\sigma_j = \frac{e^{\lambda \hat{U}_j^a}}{\sum_k e^{\lambda \hat{U}_k^a}} = \frac{e^{\lambda\left(w_1 x_j + w_2 R_j^a + w_3 P_j^a\right)}}{\sum_k e^{\lambda\left(w_1 x_k + w_2 R_k^a + w_3 P_k^a\right)}} \tag{2.37}$$

Where w_1, w_2, w_3 are the weights of the marginal coverage on target x_j, the reward R_j and the penalty P_j respectively. This marginal coverage on target can be translated to the probability of a local authority being on the target's location the moment the intruder expects to be there. The superscript a refers to the attacker, or the poacher in this case.

Compared to the classic utility function (Equation 2.36) as used by the QR model, the authors found that this new approach more accurately predicts human decision making. The downside is that this model requires the estimation of 3 parameters that determine the weights of the reward, penalty and the marginal coverage respectively instead of requiring an estimation for λ only. The λ parameter is set to 1 in the proposed SUQR model, as this does not lose generality. The authors also proposed including the reward and penalty of the defender in the subjective utility function, thus requiring 5 parameters to be estimated. Although this approach did not improve the prediction of human decisions (see Table 2.1) and was further left unexplored in the remainder of their work [55].

It was proposed to estimate the SUQR parameters using Maximum Likelihood Estimation, which was shown to give a unique local maximum to compute the optimal weights. This requires knowledge regarding the defender strategies and data covering human decisions in such situations in order to compute the weights. Rather than computing the weights for each individual independently, Nguyen, T. et al. proposed to compute the weights of the general population of intruding humans due to the lack of available data. This is a relevant topic for this research as well, since knowledge on individual attackers may be marginal or unavailable. It was shown in [55] that the SUQR model outperforms QR in accurately predicting human decision making while using limited data to estimate parameters [55, 85]. A test was performed to measure the prediction accuracy of SUQR by comparing the predicted choice distribution to observations of players' choices. This prediction accuracy of SUQR using three parameters showed the best performance, as can be seen in Table 2.1. A slightly modified Bayesian SUQR model was proposed by Yang, R. et al. that estimated a unique weight vector for each individual intruder in order to account for heterogeneous intruder [93]. However, the lack of knowledge regarding poachers and the lack of available data regarding their activities deems such an approach inappropriate within the scope of this thesis.

Table 2.1: Comparison of prediction accuracy of QR and SUQR [55].

QR	3-parameter SUQR	5-parameter SUQR
8%	51%	44%

2.2.2. Adaptable decision making

A novel human decision making model was proposed by Gholami, S. et al. [28], although this work was limited to fully rational intruders [85]. This work was therefore extended upon by Wang, B. et al. by considering the fact that intruders may not be perfectly rational in making their decisions. This is caused by the complexity of the situation which would prevent humans from computing and comparing every action's payoff, thus using some heuristic approach to choose an action. This is true when criminals take action as well, since many criminals usually act on real-time information only rather than observing the actions and/or strategy of local authorities in order to find opportunity [96]. The proposed model builds on the previously discussed SUQR model by also addressing adaptive decision making based on the intruder's inclination I_i^R in a repeated SSG and allowing intruders to collaborate such that their payoff can be increased.

Intruders receive payoffs upon completion of a task, such as attacking a certain location in the environment independently. However, the intruders also have the option to cooperate. The payoff is shared evenly in such a case, but the total payoff is increased by a bonus reward ϵ that motivates intruders to cooperate. This bonus reward is justified from the perspective that collaborating intruders can share information and smuggling channels in order to commit further destructive crimes associated with poaching. Summarising, the intruders make two decision in this model. The first being what target to attack, and the second being whether to

cooperate with another intruder. Collaboration is only established upon agreement from both parties however. Furthermore, the authors assume that intruders have non-overlapping subsets of targets such that the probability of being caught at the same time is reduced. The payoffs of the attackers R^a therefore depend both on success and on collaboration, as shown in Table 2.2 [85]. Success is achieved when a target is attacked which is uncovered by a UAV, whereas the opposite holds for failure.

Table 2.2: Poacher payoffs are shared when collaborating [85].

Collaboration Status		**Each attacker's payoff**
$a_1^{success}$	$a_2^{success}$	$(R_i^a + R_j^a + 2\epsilon)/2$
$a_1^{success}$	$a_2^{failure}$	$(R_i^a + P_j^a + \epsilon)/2$
$a_1^{failure}$	$a_2^{success}$	$(P_i^a + R_j^a + \epsilon)/2$
$a_1^{failure}$	$a_2^{failure}$	$(P_i^a + P_j^a)/2$

Wang, B. et al. proposed the adaptive subjectivy utility function to incorporate adaptive decision making in Equation 2.38 [85].

$$ASU^R(x_i) = \left(1 - d \cdot I_i^R\right)\omega_1 x_i + \left(1 + d \cdot I_i^R\right)\omega_2 R_i^a + \left(1 + d \cdot I_i^R\right)\omega_3 P_i^a \tag{2.38}$$

Where d is a parameter that addresses situations where not enough information is exposed to the attacker. It is defined according to Equation 2.39, in which N_r is the total number of rounds played in the repeated SSG and r is the current round. Due to the ABM model discretising time, this can be adapted by considering the total simulation time and the time since the simulation started. The parameter therefore increases as time passes, indicating that the intruders gain more information and experience. The inclination in Equation 2.38 is computed according to Equation 2.40, where C is a constant. avg_i^r is the average payoff attackers earn at target i per attack in round r, whereas avg_i^{all} is the overall average payoff of all targets the attackers received in round r. R is the current round, such that the payoffs of all performed rounds are included in the inclination. This definition of the inclination indicates that the inclination increases as more attackers earn a payoff at a given target. It therefore motivates intruders to attack a target i based on historical performance of other intruders [85].

$$d = \frac{1}{N_r - r} \tag{2.39}$$

$$I_i^R = \frac{\sum_{r=1}^{R} \left(\frac{avg_i^r - avg_{all}^r}{C \cdot \mathrm{Var}(avg_i^r - avg_{all}^r)}\right)}{R} \tag{2.40}$$

The probabilities of attacking a target are computed according to this adaptive subjective function in a similar fashion compared to Equation 2.37.

2.2.3. Collaborative and opportunistic decision making

Based upon the model from Wang, B. et al. [85], Zhao, Y. et al. proposed the Collaborative Opportunistic Security Game (COSG) model [96]. This approach divides the target area into cells of the same size (see Figure 2.6) and discretises time such that transition matrices can be defined. Each cell is considered a target and the intruder aims to commit opportunistic crimes in these cells. This opportunistic characteristic comes from the behaviour of criminals where only live information is used to identify crime opportunities rather than observing the strategy of the local authorities for a longer duration of time. The intruder therefore only commits a crime if no local authority is present at the target location. In case intruders agreed to cooperate, their total payoff is shared fifty-fifty, similar to the model proposed in [85]. Transition matrices are introduced to model the motion and decision making of the intruders using a discrete-time discrete-space Markov chain. This model does not incorporate the adaptive behaviour introduced by Wang, B. et al. however. The decisions are purely based on the SU in Equation 2.37 while considering the option to collaborate via the attractiveness of a choice computed from Equation 2.41 [96].

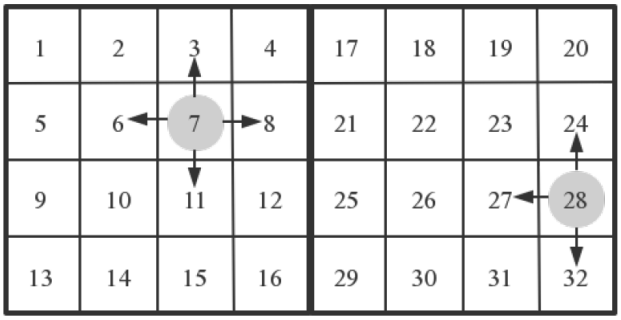


Figure 2.6: Motion modelling using opportunistic adversary decision making limited to bounded rationality [96].

$$Att(j) = \max\left\{p_{\Psi_1}^{\text{not collaborate}}(j), p_{\bar{\Psi}_1}^{\text{collaborate}}(j)\right\} \tag{2.41}$$

The probability of an intruder Ψ_a choosing to move from a cell i to target cell j is denoted by $p_{\Psi_a}(i,j)$ and is computed using Equation 2.42 by considering the attractiveness $Att(j)$ of target j and the probability distribution $\overline{c_{b,t}(j)}$ of the local authorities being present at the target location. The attractiveness is based on the probability of choosing target j that is computed using the subjective utility function of the SUQR model discussed previously [96], see Equation 2.43. This approach does not consider adaptable behaviour of the intruders, as proposed by Wang, B. et al. in [85].

$$p_{\Psi_1}(i,j) = \max\left\{\left(1 - \overline{c_{b,t}(j)}\right) \cdot Att(j)\right\} \tag{2.42}$$

$$p_{\Psi_a}^{\text{not collaborate}}(j) = \frac{e^{SU^{\text{not collaborate}}(j)}}{\sum_{I \in T_a} e^{SU^{\text{not collaborate}}(I)}} \tag{2.43}$$

2.2.4. Quantifying input for decision making models

Quantification of the rewards and penalties has not been discussed so far. This section aims to elaborate on this aspect by introducing the wildlife dimension, since poachers aim to find animals to achieve their objectives. Quantification of the rewards can therefore be achieved by considering the wildlife population distribution, whereas other environmental factors such as slope can be used to quantify the penalties [55]. In turn, the wildlife density depends on the type of animal species under consideration, its habitat, the presence of communities, water sources and animal migration between these water sources [70].

Animal movement in agent-based models can be modelled using different options, depending on the scale of migrations distance. Random walk theory is a basic model that knows a number of extensions, such as correlated random walk and biased random walk. Whereas random walk theory assumes equal probabilities of moving in any direction, correlated random walk involves a correlation between successive steps such that a persistent direction of movement is realised. Biased random walk is similar since it involves a constant bias that directs animal movement towards a preferred direction or location, such as a water source. Levy flights can also be used when animals paths involve large spatial or temporal scales. Another direction is behaviour-based movement models that use multiple stimuli to make multifaceted decision making. This involves considering the internal state of the animal, as well as abiotic factors. Their disadvantage is the lack of available empirical data to validate such models [51]. From a variety of approaches using artificial intelligence and machine learning, it was found that random forests are most accurate at predicting habitat usage by black rhinos [49]. These animals are a critically endangered species due to poaching and live in various countries in Africa, including Namibia where the environment is located that is used in this research [49, 88]. Kidway, Z. et al. [42] show that no single model should be used for management in conservation settings, since different model structures result in different population densities. Inaccurately modelling species presence and behaviour results in inappropriate conservation strategies as well [30, 49]. As such, the research by Lush, L. et al. identified abiotic habitat variables that influence the presence of black rhinos. The most important variables were found to be the density of vegetation and the browse availability [49]. Birth rates and death rates also influence the population density, as shown in [19]. These variables, among others, need

to be incorporated to accurately describe the wildlife population density and how this fluctuates over time under influence of environmental factors and animal behaviour.

2.3. Learning techniques

The previous sections indicate the amount of work that has been put to solving the multi-agent surveillance problem. The various approaches indicate the many challenges within this field of research [66]. A common feature that each of the discussed surveillance model lack is the possibility and flexibility to adapt the surveillance strategy according to the behaviour of dynamic targets. Such learning techniques can in turn be used to predict criminal activity.

Wu, H. et al. proposed an adaptive decision making model (see Section 2.1.4) to select tasks based on experience regarding their payout [89]. The approach does not include other environmental features that could influence these payouts, neither is it an approach focused on surveillance in a conservation context. Nevertheless, the approach connects with prioritised surveillance at a high level. Based on the fact that uniform distributions of crime are far from reality, there is a need for efficient and effective allocation of resources and prioritisation towards locations with high probability of criminal activity [15, 76]. This section connects with prioritised surveillance by discussing work incorporating learning techniques that predict what action robots should take given the circumstances. This is specially beneficial in complex security contexts where intruders search for weaknesses in the defence. The flexibility to adapt to intruder behaviour could therefore improve surveillance effectiveness since initial model parameters may not be and/or remain effective [78].

Game Theory based attacker-defender Stackelberg games have been used extensively to predict intruder behaviour in green security contexts and adapt surveillance strategies accordingly. These approaches are based on substantial knowledge of criminal behaviour and mechanisms in certain areas as identified in [91]. The latest work is CAPTURE, providing anti-poaching strategies using environmental attributes to predict where poachers will go. These attributes include wildlife densities, habitat, slope and various distances such as to the closest river, road or village [56]. The most common metric for crime prediction is the hit rate [60]. Patrol effort at locations can also influence intruder behaviour, meaning that time spent by law enforcement can also be used to predict criminal activity [18]. These centralised solutions explicitly incorporate the intruder's model in the path planning solution, rendering these surveillance techniques difficult to generalise to online, distributed environments [17, 29].

In multi-agent systems the agents are independent and have access to a sub part of the environment surrounding its own current location. Multi-agent learning (MAL) therefore requires that each agent individually incorporates a learning algorithm such that adaptive behaviour can be implemented. This is to prevent initial system configurations from remaining ineffective in dynamic environments. Joint adaptive behaviour is then obtained through coordination among UAVs regarding their learning progress. This is usually achieved through applying machine learning algorithms used in single agent systems to multi-agent systems, such that MAL becomes an emergent property [41]. More specifically, it is achieved by communicating newly learned rules or by exchanging new observations and adding these to the training data to update trained models [73, 78, 94]. In [64], a Local Weighted Projection Regression (LWPR) method is proposed as a computationally efficient and effective supervised algorithm to learn from a number of environmental features to find optimal paths for a single robot. The LWPR is computationally efficient even in large dimensional spaces thus making it possible to consider various environmental features that influence poacher behaviour. The approach is yet to be introduced in a real-time learning environment and could be used to learn from adaptive poaching activity due to its support for incremental learning [64, 84], which is achieved through continuously extending the training data with newly made observations [13, 79]. Heuristics were proposed to increase the probability of target detection in [67] as an approach to improve effectiveness of a persistent surveillance problem in a single UAV setting. Deep reinforcement learning was proposed in [38] to improve surveillance strategies. A neural network is trained using internal beliefs of wildfire locations and reward functions in a persistent surveillance setting. The approach is shown to be useful in domains with real-time trajectory planning as the model learns to adjust the heading based on locally made observations. Reinforcement learning has also been used in discrete surveillance such that the time to cover a target area is minimised [90]. A distributed deep Q-learning method is proposed in [83] that aims to find static targets using a swarm based approach as quick as possible. These methods train the model using a training dataset, after which the robots can make

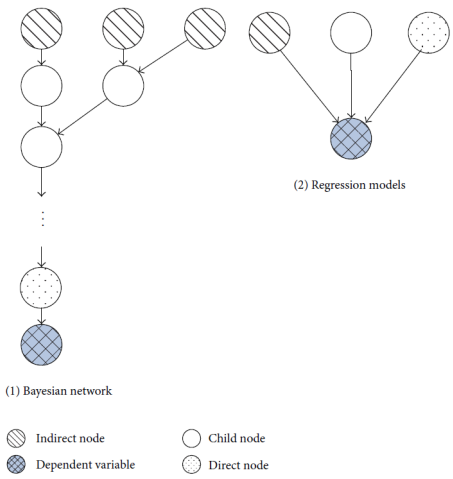


Figure 2.7: Comparison of Bayesian Network and Regression models with respect to the interactions between variables [98]

individual decisions utilising established Q-value tables. A practical implementation utilising crime occurrence maps to predict crime and adjust ranger patrol routes accordingly is proposed in [18]. The authors also emphasise the need to continuously update the patrol routes as the occurrence maps change over time under influence of changed law enforcement strategy [18].

A learning approach based on Dynamic Bayesian Networks (DBN) is proposed by Zhang, C. et al. to account for adaptive intruder behaviour [94]. Patrol areas are assigned from a central perspective by learning incrementally through adding new observations of adaptive opportunistic criminals to the training data and updating the Bayesian Network [7, 94]. It has become popular to represent uncertain systems and due to its ability to identify causal relationships (see Figure 2.7) and perform risk estimation analysis [7, 98]. Bayesian Networks are graphically represented as directed acyclic graphs (DAG) containing interconnected nodes. The nodes represent the random variables and their network defines the hierarchy of conditional independence. The optimal DAG for a given dataset is obtained by optimising the likelihood of the model given the data [7]. Since the number of possible structures scales exponentially with the number of variables, it is computationally expensive to find the most probable network. As such, algorithms have been developed to improve the computational costs [7, 94]. Zong, F. et al. compared Bayesian Networks to Bayesian Regression and concluded that the ability to define dependencies also resulted in better accuracy [98], although not always [7]. Besides, modelling poacher behaviour could benefit from causal relations among variables since many environmental attributes are not independent variables, which collides with assumptions on which regression models are based [98].

Zhao, Y. et al. point out the difficulty of collecting criminal data required to train machine learning methods [96]. Low crime occurrence rates may even disrupt continuity or relevance among crimes such that the observations become independent from each other, thus limiting the prediction accuracy [60]. Risk Terrain modelling (RTM) is a method aiming to predict crime occurrences based solely on environmental attributes rather than using historical data and has proven to be successful [15]. This introduces the need to identify crime specific environmental attributes, since these can be different between countries and between different types of criminal activity [14, 45]. RTM assumes that these features are persistent over time, similar to regression techniques. Risks computed by RTM therefore remain for long time spans, leaving changing relationships unaccounted for. It was shown that RTM did outperform other methods, among which Kernel Density Estimation (KDE) [60]. This method builds on a retrospective approach that utilises past incident locations to predict crime using kernels. Caplan, J. et al. proposed a Boolean integration of RTM and KDE approach to predict crime levels using measures of vulnerability and exposure. The Prediction Accuracy Index was used to show superior performance when RTM and KDE were integrated [15].

2.4. Research gap

Various different techniques have been developed to solve the persistent surveillance problem, slowly evolving from discrete single robot surveillance to distributed swarm based solutions that consider operational constraints. Game Theory based approaches achieve optimal patrol routes through detailed poacher models that rely on extensive knowledge of the poachers. In turn some work has been performed to design surveillance models aiming to detect black-box intruders. Most of the research has focused on 2D problems, therefore leaving work to be done in the field of surveillance in 3D environments as well. Previously developed approaches for surveillance do not necessarily remain effective as poachers adapt to their environment. In fact, most surveillance approaches consider simple, non-adaptive targets. Within the context of this project, the targets are the poachers and have their own unique objectives. The discussed models describing the decision making of poachers rely on rewards and penalties, although these do not explicitly incorporate the presence of wildlife populations to quantify these rewards and/or penalties. Machine learning techniques have been applied to improve patrol routes and to optimise surveillance effectiveness, but without consideration of the target's objective [17, 44]. A set of performance indicators was identified from literature that can measure the effectiveness of surveillance in terms of a number of dimensions. Since initially trained models may not be effective due to the limited size of the training data set, continuously updating the training data with new observations can be used to further optimise surveillance techniques and improve effectiveness [18, 73].

These machine learning techniques have yet to be used for distributed multi-agent persistent surveillance problems with dynamic targets. Reinforcement learning techniques are complex in nature and are difficult to optimise for designing cooperative surveillance in multi-UAV solutions [17]. Instead, mapping crime hotspots has been an effective technique to fight criminal activity. Besides, there is an urgency to respond to dynamic environments, which current approaches lack. This is caused by real world scenarios where crime evolves under the changing behaviour of wildlife in dynamic environments and due to changing surveillance strategies. These two effects influence criminal patterns [14], but none of the surveillance approaches consider both aspects simultaneously. Additionally, since rangers are unable to patrol the entire area of protected parks, their collected records regarding detected illegal activities is limited to about 60% of the total area. Uncertainty in crime data is to be considered to correctly model poacher behaviour due to this inability to patrol the vast areas completely and thoroughly [29].

The identified research gap is that distributed persistent surveillance models do not continuously learn and adapt to intelligent and dynamic black-box intruders to prioritise their surveillance efforts while considering uncertainty in poacher activity observations. In addition, there is a need to incorporate wildlife dynamics and to extensively analyse adaptive surveillance performance in scenarios where intelligent dynamic targets can adapt to changing surveillance strategies.

3

Research proposal

This chapter extends upon the performed literature review and the research gap that was identified in Section 2.4. The research objective and its relevant research questions are defined below in accordance with the problem context that was described in the Introduction, such that the objective of this research proposal is clarified. The scope of this project is limited due to the time constraint of this project and is discussed in this chapter as well.

3.1. Research objective

There is a need to achieve efficient surveillance in green security contexts. Context specific features such as the vast target areas and lack of infrastructure demand for distributed and autonomous surveillance systems. Persistent surveillance is also an important requirement. Moreover, due to the intelligence of poachers, the persistent surveillance system needs to prioritise and adapt to changing poacher behaviour in order to remain effective over time. The research objective of this thesis is formulated accordingly to include all various features of the subject. The objective is:

> *"To improve effectiveness of a distributed multi-UAV approach to coordination of persistent surveillance that autonomously detects coordinating and adapting poachers in vast target areas by implementing online machine learning based on observing poacher behaviour."*

As was identified in Chapter 2, effective allocation of resources is required to solve the problem of criminal activity in the context of green security [15, 76]. The absence of digital infrastructure asks for a distributed approach that depends on local communication to coordinate and prioritise surveillance efforts to areas with high rates of criminal activity [97]. Furthermore, this work considers persistent surveillance since poachers can enter the environment at any moment. Locations therefore need to be monitored continuously since it cannot be assumed that poachers do not target areas that have previously been monitored. The surveillance models described in detail in Section 2.1 meet this requirement through evaporating pheromone concentrations or through task thresholds [9, 17, 31, 32, 89, 97]. Analysing and measuring the effectiveness will be achieved through ABM and according to a sensitivity analysis using a set of performance indicators obtained from literature. By achieving this objective, this thesis will provide an approach to persistent surveillance coordination that is capable of learning from intelligent, adaptive and coordinating poachers. The novelties of this proposal are threefold. First, an approach to distributed multi-UAV coordination of prioritised and persistent surveillance is proposed to detect poachers, while accounting for limitations due to uncertainty, range, communication and kinematic constraints. These poachers can make intelligent decisions based on experience and wildlife dynamics and their population distribution. Second, the surveillance technique incorporates online machine learning through the use of Bayesian Networks such that surveillance effectiveness is improved. Finally, the effectiveness in terms of performance indicators and the secondary effects resulting from poachers adapting to learning surveillance is to be analysed extensively using a sensitivity analysis.

3.2. Research questions

Based on the previously discussed aspects of the multi-UAV surveillance topic and the stated research objective, the following main research question was formulated:

> *"How does a learning multi-UAV approach to coordination of autonomous persistent surveillance influence effectiveness of detecting coordinating and adapting intruders in vast target areas?"*

The effectiveness will be quantified according to performance indicators based on Section 2.1.6. Given the context and the main research question, a set of sub questions can be defined as well. These are used to guide this research and cover the different aspects of the problem. The sub questions are formulated below:

- *What aspects of the environment that was developed in previous research can be incorporated in this approach to agent based surveillance?*
- *What state-of-the-art technique accurately models a poacher's adaptive decision making in which poachers have the ability to coordinate?*
- *How can animal dynamics be modelled in a simplified manner such that it can be used as realistic input for a poacher's decision making?*
- *How to quantify rewards and penalties used for the decision making of poachers?*
- *How do model input parameters describing the poacher's decision making, influence poacher's behaviour?*
- *What set of model input parameters describing the poacher's decision making, results in realistic poaching behaviour?*
- *What state-of-the-art technique accurately models multi-UAV surveillance while accounting for constraints that are relevant in green security?*
- *What are the performance characteristics of the UAV to be used in the multi-agent based model?*
- *What performance indicators can be used to quantitatively analyse and compare effectiveness of multi-UAV persistent surveillance models?*
- *How do model input parameters of the multi-UAV persistent surveillance model influence surveillance effectiveness?*
- *What online machine learning technique can improve effectiveness of the multi-UAV persistent surveillance model while accounting for uncertainty?*
- *How do model input parameters of the machine learning technique influence the effectiveness of a multi-UAV persistent surveillance model?*
- *How does adaptive poacher behaviour influence effectiveness of a learning multi-UAV persistent surveillance model?*

3.3. Scope

This research is limited to some extent by the available project time frame of nine months. The research objective in Chapter 3 has also been narrowed down to perform detailed analysis of the proposed method to identify an answer to the identified research gap. The additional simplifications discussed below reduce the complexity of the model and the analysis, while improving the feasibility of this research project.

In order to perform extensive analysis of the model input parameters on surveillance effectiveness within the nine month time frame and without having to account for interference from various sub-models, pursuit of the intruders is left out. Once intruders are detected by a UAV it is assumed that the intruder is apprehended immediately without being able to execute an escape plan. The reason to exclude local authorities, rangers

within this context, from the model is closely related to the above. Despite the fact that UAVs are not physically capable of apprehending intruders, the presence of rangers could interfere with other model dynamics. The need to account for interference from sub-models is therefore prevented, allowing for more straightforward and extensive analysis of surveillance effectiveness. The discussed HAPF-ACO model does consider range constraints. As consequence, this aspect of UAV operations is not neglected, although the recharging policy will be a simplified version of reality as well. The reason for this is the time frame that is available for the literature review.

Despite the inherent 3-dimensional dynamics of UAV flight, literature has mainly focused on solving the surveillance problem in a 2D context. Although some work towards surveillance has been done in 3-dimensional space, this is not regarded as the next challenge within UAV surveillance [39, 57, 95]. Due to the absence of high obstructions in national parks, this is not considered to influence results significantly. It is therefore assumed that UAVs have built-in capability to adjust their flight altitude in order to prevent collisions if needed.

The targets discussed within the context of this research have the objective to poach wildlife, which will be modelled in a simplified manner such that wildlife presence is considered. This added dimension will improve the realism of the model output by quantifying the input of the decision making model used for the poachers. The movement of wildlife will be considered, although re-population and the wildlife's influence on the habitat's vegetation will be neglected. To further improve realism, this research project will consider day and nighttime influences on detection probability, poaching activity and wildlife movement.

The remaining chapters further elaborate on how these various aspects are combined to provide answers to the research questions.

4

Research methodology

Building upon the stated research objective and research questions in Chapter 3, this Chapter continues by defining the research steps and the research methodology.

ABM is at the basis of the to-be-developed model required to answer the research questions. This framework consists of describing the static and dynamic properties of the environment, the agents, and the interactions among and between the agents and the environment. It uses a bottom-up approach since the behaviour of individual agents is described, from which emergent behaviour is analysed through simulation [51]. This thesis is also an extension of existing work addressing anti-poacher surveillance through ABM, making it a logical choice of framework.

The poachers that need to be detected by the UAVs are to be modelled first. Recent work in green security from a Game Theory approach proposed the COSG model that can accurately model human decision making and incorporates the possibility for collaboration among poachers. It considers bounded rational humans aiming to exploit opportunities and accurately represents poacher decisions. The COSG model is based on the SUQR model, but does not allow for adaptive behaviour [96]. By extending the COSG model through implementation of the ASU function (Equation 2.38) as replacement for the SU function, adaptive poacher behaviour can be simulated. This is considered more realistic since poachers can use their past experience to improve their tactics [56, 85]. These are to be quantified using wildlife population density dynamics and abiotic environmental factors. The ASU function weights these rewards and penalties to compute probabilities of choosing an action. Upon implementation of the ASU function in the COSG model, a set of parameters need to be estimated such that realistic poacher and wildlife behaviour can be modelled. This will be achieved through sensitivity analysis, because there is no historic data available. The results will be analysed using heat maps such that realistic behaviour can be obtained from validation with expert knowledge of the environment.

Surveillance is to be achieved using a team of autonomous UAVs, since this type of law enforcement has shown to be most effective [53]. The context of the problem has resulted in a set of requirements that the surveillance system must adhere to in order to be considered relevant. Persistence of surveillance is required, since the poachers can enter the environment at any moment. Locations therefore need to be monitored continuously because we cannot assume that poachers do not target areas that have previously been monitored. The surveillance models described in detail in Section 2.1 meet this requirement through evaporating pheromone concentrations or through task thresholds [9, 17, 31, 32, 89, 97]. These evaporation rates can be altered to match the importance of a given location, thus realising prioritised surveillance [17]. This is preferred due to the vastness of the target area that is to be monitored effectively and also due to poacher behaviour. Since wildlife is their ultimate target, it can be expected that some areas within the target area are expected to experience more illegal activity. The ability of individual UAVs to store digital maps of the environment also enables distributed coordination [31], where explicit local communication can be used to synchronise and distribute information. Through updating these local maps, online adaptation to dynamic environments is achieved and systems become robust to hardware failure [32]. Such an approach is relevant due to the lack of digital infrastructure in vast target areas such as national parks. In this thesis a multi-UAV persistent surveillance approach is proposed based on HAPF-ACO, since this model also incorporates

stochastic revisit periods and uncertainty in observations, see 2.1.5 for details. Besides, this is the only persistent surveillance model combining distributed coordination with range and kinematic constraints. The approach combines ant colony optimisation and artificial potential fields. Artificial potential fields are used to avoid collisions and other threats, but also to detect nearby targets. The ant colony optimisation aspect of HAPF-ACO introduces persistent surveillance, although not prioritised [97]. This surveillance model is to be extended to include prioritised surveillance based on the proposed methods in Sections 2.1.1 and 2.1.3.

Although the HAPF-ACO model incorporates a target existence probability map that is continuously being updated, the update mechanism does require information regarding the target [97]. The possibility exists that targets can enter the environment at any time, rendering this approach unrealistic. A Bayesian Network learning technique is proposed to replace this update mechanism and to predict poacher activity such that prioritised surveillance can be achieved. Due to its ability to find optimal dependence among variables using DAGs it is more suitable than Bayesian Regression. An optimisation step is required to find the optimal DAG, but this can be achieved through dynamic programming. It is therefore less time consuming than implementing reinforcement learning based on neural networks to perform predictions. Finally, BNs are not fully deterministic as it uses a probabilistic approach to machine learning, making it suitable to perform predictions in uncertain environments [7, 98].

After implementing the models for the various agents within the ABM framework, the developed model is to be used in simulations. These simulations will be used to measure a set of Key Performance Indicators (KPIs) such that performance and effectiveness of the proposed model can be compared to the baseline model, which is build upon the state-of-the-art HAPF-ACO model. Two models utilising BNs will be developed (see Figure 4.1) that can be used to analyse the influence of adaptive poachers on surveillance performance. A sensitivity analysis of the prioritised surveillance model parameters will be executed as well, such that in-depth performance analysis can be performed. The results are to be used to verify the hypotheses stated below and to identify correlations such that model behaviour can be explained. Finally, this will be used to answer the research questions stated in Chapter 3.

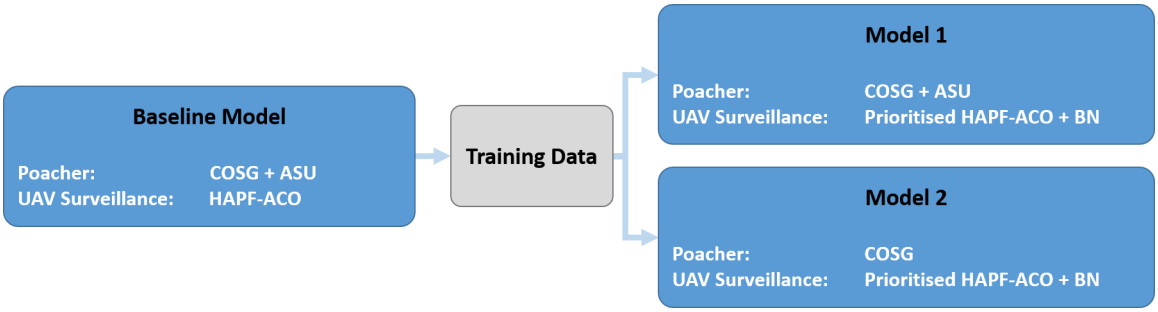


Figure 4.1: Workflow of the experiment setup describing the involved models.

The basis for distributed persistent surveillance is the HAPF-ACO model, implemented in the baseline model in Figure 4.1. The autonomous UAVs are moving according to this algorithm in order to detect poachers, that are being modelled using the Game Theory based COSG model. This model is adjusted to incorporate adaptive poacher behaviour, using the Adaptive Subjective Utility (ASU) function. The baseline model is used to perform sensitivity analysis of the poacher and wildlife models such that realistic behaviour can be achieved through validation with expert knowledge. By performing simulations with this baseline model that incorporates realistic model input parameters, training data is to be obtained from the UAVs that detect poacher activity and store these observations. This data is used to optimise a BN such that poacher activity can be predicted. This BN is then continuously being updated through online learning in Model 1 and 2, which also incorporate the proposed prioritised HAPF-ACO. The difference between Model 1 and 2 is that Model 2 does not include the ASU, such that adaptive poacher behaviour is excluded. This allows for analysis of distributed, online machine learning in adaptive environments without accounting for secondary effects from adaptive poachers w.r.t. learning surveillance. Model 1 is used to simulate and analyse the proposed model that aims to improve effectiveness of online learning in complex dynamic environments compared to the baseline model. It will also be used in a sensitivity analysis to identify correlations between model input parameters and model performance in terms of KPIs based on Section 2.1.6.

4.1. Simulation setup

This section elaborates on the used software frameworks that are the basis of the model.

This research is using an ABM approach that requires software in order to implement models and obtain simulation results to test the hypotheses discussed in the following section. Two software packages offering a framework for agent-based, bottom-up programming that extensively support ABM are NetLogo [1] and Mesa [2]. Both of these frameworks were introduced during the master courses related to ABM. Although NetLogo provides a relatively simple user experience, Mesa is Python based and therefore benefits from the extensive Python community. The Python based Mesa package will be the foundation of the model since Python is commonly used in the industry [65]. The various aspects of the research objective require a variety of modelling techniques to be used. Python also allows the model to be adapted when new, state-of-the-art techniques that are still to be introduced and would be able to extend the scope and overcome limitations of current techniques. Python, combined with the Mesa framework, is also the preferred software framework since most programming experience has been obtained through use of these frameworks.

The available hardware that is to be used to perform the simulations consists of, at the time of writing, a desktop with an i7-3770 CPU @3.4GHz and 8GB RAM. A laptop with lower specifications will be used mainly for development and analysis purposes. This enables simultaneous development and simulation execution. Since large environments combined with small time steps require long simulation runs in order to perform simulations, this setup improves the feasibility of the project.

4.2. Hypotheses

The adaptive behaviour of intelligent poachers allows them to evade areas where the rate of success is lower. This effectively means that poachers adapt to surveillance from law enforcement and try to achieve their goals in a different way. The result is that surveillance systems that are effective initially, will fall short as time progresses. Distributed online machine learning using Bayesian Networks is proposed to overcome this. A Bayesian Network machine learning technique is proposed in an online, distributed learning environment to improve effectiveness by predicting poacher activity. It is expected that incremental learning to predict poacher activity will improve the effectiveness of detecting coordinating and adaptive poachers using a prioritised HAPF-ACO surveillance model. It is therefore hypothesised that the effectiveness of the prioritised HAPF-ACO model that utilises online learning through Bayesian Networks will remain effective, whereas the baseline, the HAPF-ACO model, will degrade over time as the poachers adapt their behaviour. The discussed simulation setup will test these hypotheses by generating data describing the models' effectiveness, which is further discussed in Chapter 5.

[1] NetLogo Documentation: https://ccl.northwestern.edu/netlogo/docs/
[2] Mesa Documentation: https://mesa.readthedocs.io/en/master/

5

Results, outcome and relevance

This chapter briefly summarises what can be expected from this project.

5.1. Results and relevance

The results of this thesis incorporate a number of milestones. First, a multi-agent based model along with documentation regarding its use is expected to be finished by the mid-term meeting. A modular model design allows for simple on/off toggling of various model attributes and performing simulations in different scenarios. This also enhances the adaptability of the model to future extensions and/or improvements such that the model can be used as a platform to explore possible recommendations from this research or elsewhere. Along with the finalised model, a mid-term report and presentation will be provided that summarise the performed work.

The developed model is to be used to simulate the interactions of the agents as discussed in Chapter 4. These simulations generate datasets describing the behaviour and performance of the implemented techniques as result of a set of input parameters. Sensitivity analyses are performed to identify quantitative relations between input and output by considering a set of performance indicators based on the literature in Section 2.1.6. The simulations are also used as a means to compare the baseline model to the performance of the prioritised surveillance coordination approach in model 1 and to identify explanations for behavioural characteristics. Providing answers to the research questions defined in Chapter 3 is achieved through considering the determined relationships and emergent behaviour of the model. The data analysis of these results is to be incorporated in the paper and the thesis report.

The model will be a novel ABM approach to wildlife conservation in national parks. Its modular design aids in designing optimal techniques for various environments in a systematic, cost-effective and controlled manner. It also addresses the uncovered topic of machine learning which can be used to adapt the surveillance strategy and to predict poacher activity in a context where autonomous, distributed and persistent surveillance are modelled to detect intelligent humans in a green security context. The performed research and the findings will be presented during the graduation defence.

5.2. Validation and verification

In order to determine whether the to-be-developed simulation model represents realistic scenarios at an acceptable level, validation steps are required. Although the model environment is based on expert knowledge, assumptions and the absence of data regarding model input and output requires cross-validation with expert knowledge to accurately describe poacher and wildlife dynamics. What-if analyses includes considering model output when subjected to extreme inputs and can be used to verify consistent model performance. Since the model consists of four main aspects, the intruders' behavioural model, the wildlife's motion, the surveillance model and the machine learning model, the validation is required to be performed at each of these aspects individually. This consists of carefully calibrating parameters such that realistic model behaviour can be obtained. The to-be-developed model will also be compared to the state-of-the-art HAPF-ACO model using a set of KPIs to measure the detection effectiveness. As part of the parameter estimation

and sensitivity analysis, the simulations are to be performed in multitude to show that the model is capable of reproducing similar model output. This can also be visualised using the coefficient of variation.

Finally, verification of the computerised multi-agent based model is performed by critically analysing the programming implementation to ensure correctness at the model definition level. This also involves extensive testing and debugging where required.

6

Planning

The contents of this Chapter briefly mention the milestones of this thesis, which have previously been discussed in Chapter 5, and relates these to the time frame. A Gantt Chart is displayed in Appendix 8 visualising the workload and transitions among the various project aspects. The various tasks are subdivided into three phases; literature review, initial phase and the final phase.

The results of the literature review include this report, the research proposal, and is finalised in the kick-off presentation. A three month long initial phase is started afterwards, during which the main effort is put into developing and implementing the model and simulation setup (discussed in Chapter 4) such that simulations and validation with expert knowledge can be performed. The initial phase is concluded in January 2021 with the mid-term meeting. By this time the developed model is to be presented along with an updated project planning, as mentioned in Chapter 5. Although simulations and data analysis are to be initiated prior to the mid-term meeting, most of the remaining work regarding this aspect is performed during the final phase. Writing of the paper and the thesis report is done in accordance with the simulations and the data analysis, from which the results are to be presented in May 2021.

7

Conclusion

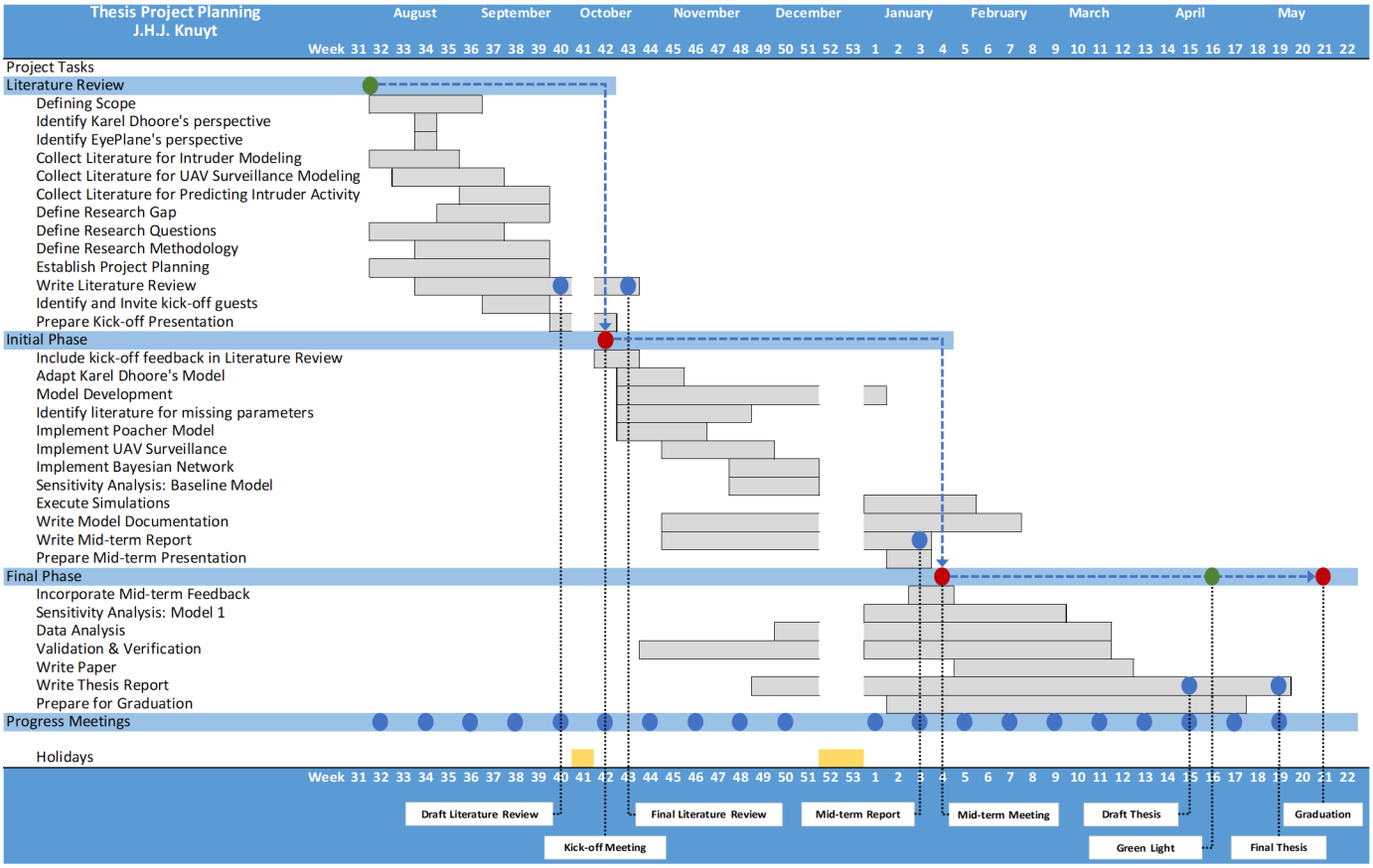
Ongoing poacher activity threatens iconic animal species. UAVs have found many applications in green security, although an extensive literature review identified open research opportunities. The identified research gap is that distributed persistent surveillance models do not continuously learn and adapt to intelligent and dynamic black-box intruders to prioritise their surveillance efforts while considering uncertainty in poacher activity observations. In addition, there is a need to incorporate wildlife dynamics and to extensively analyse adaptive surveillance performance in scenarios where intelligent dynamic targets can adapt to changing surveillance strategies. This research gap was transformed into the following research question:

> *"How does a learning multi-UAV approach to coordination of autonomous persistent surveillance influence effectiveness of detecting coordinating and adapting intruders in vast target areas?"*

The literature review identified several relevant models that will be used to achieve the research objective and provide answers to this question. The proposed model is a novel and relevant approach to anti-poaching strategies and adds to the body of research by filling the research gap. It can also serve as a platform for future research. The proposed model utilises the swarm based HAPF-ACO persistent surveillance model and extends it by incorporating prioritised surveillance through predicting poaching activity. This will be achieved through online machine learning from observing poacher behaviour and building Bayesian Networks in order to improve surveillance effectiveness in response to adapting poachers. These poachers will be modelled according to the COSG model and extending it by incorporating the ASU function, whose inputs will be quantified from wildlife population distributions and environmental factors. Implementation will be achieved in the Python based Mesa framework that is used for ABM. Expert knowledge will be used to validate simulated poacher behaviour, after which a sensitivity analysis is used to identify correlations such that model behaviour can be explained. Thanks to gained experience w.r.t. the mentioned frameworks in master courses, this research project is assumed to be feasible within the available time frame of nine months.

8

Project Gantt Chart



Bibliography

[1] Jose J. Acevedo, Begoña C. Arrue, Jose Miguel Diaz-Bañez, Inmaculada Ventura, Ivan Maza, and Anibal Ollero. One-to-one coordination algorithm for decentralized area partition in surveillance missions with a team of aerial robots. *Journal of Intelligent and Robotic Systems: Theory and Applications*, 74(1-2):269–285, apr 2014. ISSN 09210296. doi: 10.1007/s10846-013-9938-z.

[2] Christoph Adami, Jory Schossau, and Arend Hintze. Evolutionary game theory using agent-based methods, dec 2016. ISSN 15710645.

[3] Dario Albani, Daniele Nardi, and Vito Trianni. Field coverage and weed mapping by UAV swarms. In *IEEE International Conference on Intelligent Robots and Systems*, volume 2017-Septe, pages 4319–4325. Institute of Electrical and Electronics Engineers Inc., dec 2017. ISBN 9781538626825. doi: 10.1109/IROS.2017.8206296.

[4] Dario Albani, Tiziano Manoni, Daniele Nardi, and Vito Trianni. Dynamic UAV swarm deployment for non-uniform coverage: Robotics track. In *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS*, volume 1, pages 523–531. International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS), 2018. ISBN 9781510868083.

[5] Yaniv Altshuler, Alex Pentland, and Alfred M. Bruckstein. The cooperative hunters Efficient and scalable drones swarm for multiple targets detection. In *Studies in Computational Intelligence*, volume 729, pages 187–205. Springer Verlag, 2018. doi: 10.1007/978-3-319-63604-7_7.

[6] J. F. Araujo, P. B. Sujit, and J. B. Sousa. Multiple UAV area decomposition and coverage. In *Proceedings of the 2013 IEEE Symposium on Computational Intelligence for Security and Defense Applications, CISDA 2013 - 2013 IEEE Symposium Series on Computational Intelligence, SSCI 2013*, pages 30–37, 2013. ISBN 9781467359115. doi: 10.1109/CISDA.2013.6595424.

[7] Paul Arora, Devon Boyne, Justin J. Slater, Alind Gupta, Darren R. Brenner, and Marek J. Druzdzel. Bayesian Networks for Risk Prediction Using Real-World Data: A Tool for Precision Medicine. *Value in Health*, 22(4):439–445, apr 2019. ISSN 15244733. doi: 10.1016/j.jval.2019.01.006.

[8] Johan Bergenas, Rachel Stohl, and Alexander Georgieff. The other side of drones: Saving wildlife in Africa and managing global crime. Technical report, 2013.

[9] Tilemachos Bontzorlos and Georgios Ch Sirakoulis. Bioinspired algorithm for area surveillance using autonomous robots. *International Journal of Parallel, Emergent and Distributed Systems*, 32(4):368–385, jul 2017. ISSN 17445779. doi: 10.1080/17445760.2016.1184269.

[10] Aseem Vivek Borkar, Arpita Sinha, Leena Vachhani, and Hemendra Arya. Application of Lissajous curves in trajectory planning of multiple agents. *Autonomous Robots*, 44(2):233–250, jan 2020. ISSN 15737527. doi: 10.1007/s10514-019-09888-7.

[11] Richard P. Braatz, Peter M. Young, John C. Doyle, and Manfred Morari. Computational Complexity of μ Calculation. *IEEE Transactions on Automatic Control*, 39(5):1000–1002, 1994. ISSN 15582523. doi: 10.1109/9.284879.

[12] Bureau of Oceans and International Environmental and Scientific Affairs. Eliminate, Neutralize, and Disrupt Wildlife Trafficking Act of 2016 PL 114-231, Sec. 301(d) 2019 Strategic Review, nov 2019. URL https://www.state.gov/2019-end-wildlife-trafficking-strategic-review/.

[13] Tauã M. Cabreira, Lisane B. Brisolara, and R. Ferreira Paulo. Survey on coverage path planning with unmanned aerial vehicles. *Drones*, 3(1):1–38, mar 2019. ISSN 2504446X. doi: 10.3390/drones3010004.

[14] Joel M. Caplan, Leslie W. Kennedy, and Joel Miller. Risk Terrain Modeling: Brokering Criminological Theory and GIS Methods for Crime Forecasting. *Justice Quarterly*, 28(2):360–381, 2011. ISSN 17459109. doi: 10.1080/07418825.2010.486037.

[15] Joel M. Caplan, Leslie W. Kennedy, Eric L. Piza, and Jeremy D. Barnum. Using Vulnerability and Exposure to Improve Robbery Prediction and Target Area Selection. *Applied Spatial Analysis and Policy*, 13(1):113–136, mar 2020. ISSN 18744621. doi: 10.1007/s12061-019-09294-7.

[16] H. Chen, T. Cheng, and S. Wise. Designing daily patrol routes for policing based on ant colony algorithm. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 2, pages 103–109. Copernicus GmbH, jul 2015. doi: 10.5194/isprsannals-II-4-W2-103-2015.

[17] Huanfa Chen, Tao Cheng, and Sarah Wise. Developing an online cooperative police patrol routing strategy. *Computers, Environment and Urban Systems*, 62:19–29, mar 2017. ISSN 01989715. doi: 10.1016/j.compenvurbsys.2016.10.013.

[18] Rob Critchlow, Andrew J. Plumptre, Bazil Alidria, Mustapha Nsubuga, Margaret Driciru, Aggrey Rwetsiba, F. Wanyama, and Colin M. Beale. Improving Law-Enforcement Effectiveness and Efficiency in Protected Areas Using Ranger-collected Monitoring Data, sep 2017. ISSN 1755263X.

[19] Joris P.G.M. Cromsigt, John Hearne, Ignas M.A. Heitkönig, and Herbert H.T. Prins. Using models in the management of Black rhino populations. In *Ecological Modelling*, volume 149, pages 203–211, mar 2002. doi: 10.1016/S0304-3800(01)00524-5.

[20] Hebah ElGibreen and Kamal Youcef-Toumi. Dynamic task allocation in an uncertain environment with heterogeneous multi-agents. *Autonomous Robots*, 43(7):1639–1664, oct 2019. ISSN 15737527. doi: 10.1007/s10514-018-09820-5.

[21] Charles A. Erignac. An exhaustive swarming search strategy based on distributed pheromone maps. In *Collection of Technical Papers - 2007 AIAA InfoTech at Aerospace Conference*, volume 2, pages 1130–1145. American Institute of Aeronautics and Astronautics Inc., 2007. ISBN 1563478935. doi: 10.2514/6.2007-2822.

[22] Etosha National Park. Can Drones Save Etosha's Wildlife?, aug 2014. URL https://www.etoshanationalpark.org/news/can-drones-save-etoshas-wildlife.

[23] Fei Fang and Thanh H. Nguyen. Green security games: apply game theory to addressing green security challenges. *ACM SIGecom Exchanges*, 15(1):78–83, sep 2016. ISSN 1551-9031. doi: 10.1145/2994501.2994507. URL https://dl.acm.org/doi/10.1145/2994501.2994507.

[24] Fei Fang, Albert Xin Jiang, and Milind Tambe. Optimal patrol strategy for protecting moving targets with multiple mobile resources. In *12th International Conference on Autonomous Agents and Multiagent Systems 2013, AAMAS 2013*, volume 2, pages 957–964. International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS), 2013.

[25] Fei Fang, Peter Stone, and Milind Tambe. When Security Games Go Green: Designing Defender Strategies to Prevent Poaching and Illegal Fishing. Technical report, 2015.

[26] Eduardo Feo Flushing, Luca M. Gambardella, and Gianni A. Di Caro. On decentralized coordination for spatial task allocation and scheduling in heterogeneous teams. In *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS*, pages 988–996. International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS), 2016. ISBN 9781450342391.

[27] Paolo Gaudiano, Benjamin Shargel, Eric Bonabeau, and Bruce T. Clough. Control of UAV Swarms: What the bugs can teach us. In *2nd AIAA "Unmanned Unlimited" Conference and Workshop and Exhibit*. American Institute of Aeronautics and Astronautics Inc., 2003. doi: 10.2514/6.2003-6624.

[28] Shahrzad Gholami, Bryan Wilder, Matthew Brown, Arunesh Sinha, Nicole Sintov, and Milind Tambe. A Game Theoretic Approach on Addressing Cooperation among Human Adversaries. Technical report, 2016. URL www.ifaamas.org.

[29] Shahrzad Gholami, Sara McCarthy, Bistra Dilkina, Andrew Plumptre, Milind Tambe, Margaret Driciru, Fred Wanyama, Aggrey Rwetsiba, Mustapha Nsubaga, Joshua Mabonga, Tom Okello, and Eric Enyel. Adversary models account for imperfect crime data: Forecasting and planning against real-world poachers. In *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS*, volume 2, pages 823–831. International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS), 2018. ISBN 9781510868083.

[30] H.C. Hearne, J. Swart, and P. Goodman. A conservation model for black rhino. *ORiON*, 7(1), dec 2003. ISSN 2224-0004. doi: 10.5784/7-1-473.

[31] David Howden. Continuous swarm surveillance via distributed priority maps. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 5865 LNAI, pages 221–231, 2009. ISBN 3642104266. doi: 10.1007/978-3-642-10427-5_22.

[32] David J. Howden. Fire tracking with collective intelligence using dynamic priority maps. In *2013 IEEE Congress on Evolutionary Computation, CEC 2013*, pages 2610–2617, 2013. ISBN 9781479904549. doi: 10.1109/CEC.2013.6557884.

[33] International Union for Conservation of Nature. Species, 2020. URL https://www.iucn.org/theme/species.

[34] Alexander Jahn, Reza Javanmard Alitappeh, David Saldana, Luciano C.A. Pimenta, Andre G. Santos, and Mario F.M. Campos. Distributed multi-robot coordination for dynamic perimeter surveillance in uncertain environments. In *Proceedings - IEEE International Conference on Robotics and Automation*, pages 273–278. Institute of Electrical and Electronics Engineers Inc., jul 2017. ISBN 9781509046331. doi: 10.1109/ICRA.2017.7989035.

[35] Rishav Jain, Rohan Tiwari, Puneet Jain, and P. B. Sujit. Distributed Fault Tolerant and Balanced Multi-Robot Area Partitioning for Coverage Applications. In *2018 International Conference on Unmanned Aircraft Systems, ICUAS 2018*, pages 293–299. Institute of Electrical and Electronics Engineers Inc., aug 2018. ISBN 9781538613535. doi: 10.1109/ICUAS.2018.8453387.

[36] Yining Jin, Yanxuan Wu, and Ningjun Fan. Research on distributed cooperative control of swarm UAVs for persistent coverage. In *Proceedings of the 33rd Chinese Control Conference, CCC 2014*, pages 1162–1167. IEEE Computer Society, sep 2014. ISBN 9789881563842. doi: 10.1109/ChiCC.2014.6896792.

[37] Tian Jing, Weiping Wang, Tao Wang, Xiaobo Li, and Xin Zhou. Dynamic control scheme of multiswarm persistent surveillance in a changing environment. *Computational Intelligence and Neuroscience*, 2019, 2019. ISSN 16875273. doi: 10.1155/2019/6025657.

[38] Kyle D. Julian and Mykel J. Kochenderfer. Distributed wildfire surveillance with autonomous aircraft using deep reinforcement learning. *Journal of Guidance, Control, and Dynamics*, 42(8):1768–1778, 2019. ISSN 15333884. doi: 10.2514/1.G004106.

[39] Evan Kaufman, Kuya Takami, Zhuming Ai, and Taeyoung Lee. Bayesian Mapping-Based Autonomous Exploration and Patrol of 3D Structured Indoor Environments with Multiple Flying Robots. *Journal of Intelligent and Robotic Systems: Theory and Applications*, 98(2):403–419, may 2020. ISSN 15730409. doi: 10.1007/s10846-019-01066-2.

[40] Rachel Kaufman. Hunter Becomes the Hunted, dec 2015. URL https://insideunmannedsystems.com/hunter-becomes-hunted-drone-wildlife-monitoring/.

[41] Khaled M. Khalil, M. Abdel-Aziz, Taymour T. Nazmy, and Abdel Badeeh M. Salem. MLIMAS: A Framework for Machine Learning in Interactive Multi-agent Systems. In *Procedia Computer Science*, volume 65, pages 827–835. Elsevier, 2015. doi: 10.1016/j.procs.2015.09.035.

[42] Zaara Kidwai, Jose Jimenez, Cornelius J. Louw, H. P. Nel, and Jason P. Marshal. Using N-mixture models to estimate abundance and temporal trends of black rhinoceros (Diceros bicornis L.) populations from aerial counts. *Global Ecology and Conservation*, 19, jul 2019. ISSN 23519894. doi: 10.1016/j.gecco.2019.e00687.

[43] Thomas M. Kratzke, Lawrence D. Stone, and John R. Frost. Search and Rescue Optimal Planning System. In *13th Conference on Information Fusion, Fusion 2010*. IEEE Computer Society, 2010. ISBN 9780982443811. doi: 10.1109/icif.2010.5712114.

[44] Karel Kuchar, Eva Holasova, Lukas Hrboticky, Martin Rajnoha, and Radim Burget. Supervised learning in multi-agent environments using inverse point of view. In *2019 42nd International Conference on Telecommunications and Signal Processing, TSP 2019*, pages 625–628. Institute of Electrical and Electronics Engineers Inc., jul 2019. ISBN 9781728118642. doi: 10.1109/TSP.2019.8768860.

[45] Jane Law, Matthew Quick, and Afraaz Jadavji. A Bayesian spatial shared component model for identifying crime-general and crime-specific hotspots. *Annals of GIS*, 26(1):65–79, jan 2020. ISSN 19475691. doi: 10.1080/19475683.2020.1720290.

[46] Mingchu Li, Yuanpeng Cao, and Tie Qiu. Optimal patrol strategies against attacker's persistent attack with multiple resources. In *2017 IEEE SmartWorld Ubiquitous Intelligence and Computing, Advanced and Trusted Computed, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People and Smart City Innovation, SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI 2017 -*, 2018. ISBN 9781538604342. doi: 10.1109/UIC-ATC.2017.8397517.

[47] Seung-Han Lim and Hyo-Choong Bang. Waypoint Planning Algorithm Using Cost Functions for Surveillance. *International Journal of Aeronautical and Space Sciences*, 11(2):136–144, jun 2010. ISSN 2093-274X. doi: 10.5139/jass.2010.11.2.136.

[48] Jesús Jiménez López and Margarita Mulero-Pázmány. Drones for conservation in protected areas: Present and future, mar 2019. ISSN 2504446X. URL www.mdpi.com/journal/drones.

[49] Lucy Lush, Martin Mulama, and Martin Jones. Predicting the habitat usage of African black rhinoceros (Diceros bicornis) using random forest models. *African Journal of Ecology*, 53(3):346–354, sep 2015. ISSN 13652028. doi: 10.1111/aje.12192.

[50] Richard D. McKelvey and Thomas R. Palfrey. Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10(1):6–38, 1995. ISSN 10902473. doi: 10.1006/game.1995.1023.

[51] Adam J. McLane, Christina Semeniuk, Gregory J. McDermid, and Danielle J. Marceau. The role of agent-based models in wildlife ecology and management. *Ecological Modelling*, 222(8):1544–1556, apr 2011. ISSN 03043800. doi: 10.1016/j.ecolmodel.2011.01.020.

[52] Wei Meng, Zhirong He, Rong Su, Pradeep K. Yadav, Rodney Teo, and Lihua Xie. Decentralized Multi-UAV Flight Autonomy for Moving Convoys Search and Track. *IEEE Transactions on Control Systems Technology*, 25(4):1480–1487, jul 2017. ISSN 10636536. doi: 10.1109/TCST.2016.2601287.

[53] E. J. Milner-Gulland and N. Leader-Williams. A Model of Incentives for the Illegal Exploitation of Black Rhinos and Elephants: Poaching Pays in Luangwa Valley, Zambia. *The Journal of Applied Ecology*, 29(2): 401, 1992. ISSN 00218901. doi: 10.2307/2404508.

[54] Emily Neil, Jens Koed Madsen, Ernesto Carrella, Nicolas Payette, and Richard Bailey. Agent-based modelling as a tool for elephant poaching mitigation. *Ecological Modelling*, 427:109054, jul 2020. ISSN 03043800. doi: 10.1016/j.ecolmodel.2020.109054.

[55] Thanh H. Nguyen, Rong Yang, Amos Azaria, Sarit Kraus, and Milind Tambe. Analyzing the effectiveness of adversary modeling in security games. In *Proceedings of the 27th AAAI Conference on Artificial Intelligence, AAAI 2013*, pages 718–724, 2013. ISBN 9781577356158.

[56] Thanh H Nguyen, Arunesh Sinha, Shahrzad Gholami, Andrew Plumptre, Lucas Joppa, Milind Tambe, Margaret Driciru, Fred Wanyama, Aggrey Rwetsiba, Rob Critchlow, and Colin M Beale. CAPTURE: A New Predictive Anti-Poaching Tool for Wildlife Protection. *AAMAS '16: Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, pages 767–775, may 2016. doi: 10.5555/2936924. URL www.ifaamas.org.

[57] Nikhil Nigam. The multiple unmanned Air Vehicle persistent surveillance problem: A review. *Machines*, 2(1):13–72, mar 2014. ISSN 20751702. doi: 10.3390/machines2010013. URL http://www.mdpi.com/2075-1702/2/1/13.

[58] Nikhil Nigam and Ilan Kroo. Persistent surveillance using multiple unmanned air vehicles. In *IEEE Aerospace Conference Proceedings*, 2008. ISBN 1424414881. doi: 10.1109/AERO.2008.4526242.

[59] Rachel Nuwer. High Above, Drones Keep Watchful Eyes on Wildlife in Africa, 2017. URL https://www.nytimes.com/2017/03/13/science/drones-africa-poachers-wildlife.html.

[60] Tomoya Ohyama and Mamoru Amemiya. Applying Crime Prediction Techniques to Japan: A Comparison Between Risk Terrain Modeling and Other Methods. *European Journal on Criminal Policy and Research*, 24(4):469–487, dec 2018. ISSN 15729869. doi: 10.1007/s10610-018-9378-1.

[61] Matej Paradzik and Gökhan Ince. Multi-agent search stratey based on digital pheromones for UAVs. In *2016 24th Signal Processing and Communication Application Conference, SIU 2016 - Proceedings*, pages 233–236. Institute of Electrical and Electronics Engineers Inc., jun 2016. ISBN 9781509016792. doi: 10.1109/SIU.2016.7495720.

[62] Fabio Pasqualetti, Joseph W. Durham, and Francesco Bullo. Cooperative patrolling via weighted tours: Performance analysis and distributed algorithms. *IEEE Transactions on Robotics*, 28(5):1181–1188, 2012. ISSN 15523098. doi: 10.1109/TRO.2012.2201293.

[63] Jeffrey R. Peters, Sean J. Wang, and Francesco Bullo. Coverage control with anytime updates for persistent surveillance missions. In *Proceedings of the American Control Conference*, pages 265–270. Institute of Electrical and Electronics Engineers Inc., jun 2017. ISBN 9781509059928. doi: 10.23919/ACC.2017.7962964.

[64] Miloš Prágr, Petr Váa, and Jan Faigl. Aerial Reconnaissance and Ground Robot Terrain Learning in Traversal Cost Assessment. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 11995 LNCS, pages 3–10. Springer, 2020. ISBN 9783030438890. doi: 10.1007/978-3-030-43890-6_1.

[65] Python Software Foundation. Python Success Stories, 2020. URL https://www.python.org/about/success/.

[66] Manon Raap, Michael Preuß, and Silja Meyer-Nieberg. Moving target search optimization A literature review, may 2019. ISSN 03050548.

[67] Manickam Ramasamy and Debasish Ghose. A Heuristic Learning Algorithm for Preferential Area Surveillance by Unmanned Aerial Vehicles. *Journal of Intelligent and Robotic Systems: Theory and Applications*, 88(2-4):655–681, dec 2017. ISSN 15730409. doi: 10.1007/s10846-017-0498-5.

[68] Steven Rasmussen, Krishnamoorthy Kalyanam, Satyanarayana Manyam, David Casbeer, and Christopher Olsen. Practical considerations for implementing an autonomous, persistent, intelligence, surveillance, and reconnaissance system. In *1st Annual IEEE Conference on Control Technology and Applications, CCTA 2017*, volume 2017-Janua, pages 1847–1854. Institute of Electrical and Electronics Engineers Inc., oct 2017. ISBN 9781509021826. doi: 10.1109/CCTA.2017.8062725.

[69] John H. Reif and Hongyan Wang. Social potential fields: A distributed behavioral control for autonomous robots. *Robotics and Autonomous Systems*, 27(3):171–194, may 1999. ISSN 09218890. doi: 10.1016/S0921-8890(99)00004-4.

[70] Elisa Reuter and Lieselot Bisschop. Keeping the Horn on the Rhino: A Study of Balule Nature Reserve. In *The Geography of Environmental Crime*, pages 149–185. Palgrave Macmillan UK, 2016. doi: 10.1057/978-1-137-53843-7_7. URL https://link.springer.com/chapter/10.1057/978-1-137-53843-7_7.

[71] Martin Rosalie, Grégoire Danoy, Serge Chaumette, and Pascal Bouvry. From random process to chaotic behavior in swarms of UAVs. In *DIVANet 2016 - Proceedings of the 6th ACM Symposium on Development and Analysis of Intelligent Vehicular Networks and Applications, co-located with MSWiM 2016*, pages 9–15. Association for Computing Machinery, Inc, nov 2016. ISBN 9781450345064. doi: 10.1145/2989275.2989281.

[72] Johannes O. Royset and Hiroyuki Sato. Route optimization for multiple searchers. *Naval Research Logistics*, 57(8):701–717, dec 2010. ISSN 0894069X. doi: 10.1002/nav.20432.

[73] Bartosz Sadel and Bartlomiej Śniezyński. Online supervised learning approach for machine scheduling. *Schedae Informaticae*, 25:165–176, 2016. ISSN 20838476. doi: 10.4467/20838476SI.16.013.6194.

[74] Chris Sandbrook. The social implications of using drones for biodiversity conservation. *Ambio*, 44, 2015. doi: 10.1007/s13280-015-0714-0. URL www.kva.se/en.

[75] Jurgen Scherer and Bernhard Rinner. Multi-Robot Persistent Surveillance with Connectivity Constraints. *IEEE Access*, 8:15093–15109, 2020. ISSN 21693536. doi: 10.1109/aCCESS.2020.2967650.

[76] Nurul Hazwani Mohd Shamsuddin, Nor Azizah Ali, and Razana Alwee. An overview on crime prediction methods. In *6th ICT International Student Project Conference: Elevating Community Through ICT, ICT-ISPC 2017*, volume 2017-Janua, pages 1–5. Institute of Electrical and Electronics Engineers Inc., oct 2017. ISBN 9781538629963. doi: 10.1109/ICT-ISPC.2017.8075335.

[77] Skyeton. Long-range Drone/UAV for Anti-poaching and Wildlife Conservation, 2017. URL https://geo-matching.com/content/long-range-drone/uav-for-anti-poaching-and-wildlife-conservation.

[78] Bartłomiej Śniezyński. Agent strategy generation by rule induction in predator-prey problem. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume PART 2, pages 895–903, 2009. ISBN 3642019722. doi: 10.1007/978-3-642-01973-9_99.

[79] Bartłomiej Śniezyński. Comparison of reinforcement and supervised learning methods in farmer-pest problem with delayed rewards. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 8083 LNAI, pages 399–408, 2013. ISBN 9783642404948. doi: 10.1007/978-3-642-40495-5_40.

[80] Hansen Thambi Prem. What are the biggest threats to Wildlife and Why?, feb 2020. URL https://www.worldanimalprotection.org.in/blogs/what-are-biggest-threats-wildlife-and-why.

[81] United Nations. World Wildlife Report, 2020. URL https://www.unodc.org/unodc/en/data-and-analysis/wildlife.html.

[82] Nick van Doormaal. Exploring anti-poaching strategies for wildlife crime with a simple and general agent-based model. In *18th EPIA Conference on Artificial Intelligence*, volume 10423 LNAI, pages 51–62. Springer Verlag, 2017. ISBN 9783319653396. doi: 10.1007/978-3-319-65340-2_5.

[83] Federico Venturini, Federico Mason, Francesco Pase, Federico Chiariotti, Alberto Testolin, Andrea Zanella, and Michele Zorzi. Distributed reinforcement learning for flexible UAV swarm control with transfer learning capabilities. In *Proceedings of the 6th ACM Workshop on Micro Aerial Vehicle Networks, Systems, and Applications, DroNet 2020*. Association for Computing Machinery, Inc, jun 2020. ISBN 9781450380102. doi: 10.1145/3396864.3399701.

[84] Sethu Vijayakumar, Aaron D'Souza, and Stefan Schaal. Incremental online learning in high dimensions. *Neural Computation*, 17(12):2602–2634, dec 2005. ISSN 08997667. doi: 10.1162/089976605774320557.

[85] Binru Wang, Yuan Zhang, Zhi Hua Zhou, and Sheng Zhong. On repeated stackelberg security game with the cooperative human behavior model for wildlife protection. *Applied Intelligence*, 49(3):1002–1015, mar 2019. ISSN 15737497. doi: 10.1007/s10489-018-1307-y.

[86] Bernat Wiandt, Vilmos Simon, and Andras Kokuti. Self-organized graph partitioning approach for multi-agent patrolling in generic graphs. In *17th IEEE International Conference on Smart Technologies, EUROCON 2017 - Conference Proceedings*, pages 605–610. Institute of Electrical and Electronics Engineers Inc., aug 2017. ISBN 9781509038435. doi: 10.1109/EUROCON.2017.8011183.

[87] Justin Worland. Drones Are Helping Catch Poachers Operating Under Cover of Darkness, 2018. URL https://time.com/5279322/drones-poaching-air-shepherd/.

[88] World Wildlife Fund. Black Rhino, 2020. URL https://www.worldwildlife.org/species/black-rhino.

[89] Husheng Wu, Hao Li, Renbin Xiao, and Jie Liu. Modeling and simulation of dynamic ant colony's labor division for task allocation of UAV swarm. *Physica A: Statistical Mechanics and its Applications*, 491: 127–141, feb 2018. ISSN 03784371. doi: 10.1016/j.physa.2017.08.094.

[90] Jian Xiao, Gang Wang, Ying Zhang, and Lei Cheng. A Distributed Multi-Agent Dynamic Area Coverage Algorithm Based on Reinforcement Learning. *IEEE Access*, 8:33511–33521, 2020. ISSN 21693536. doi: 10.1109/ACCESS.2020.2967225.

[91] Haifeng Xu, Benjamin Ford, Fei Fang, Bistra Dilkina, Andrew Plumptre, Milind Tambe, Margaret Driciru, Fred Wanyama, Aggrey Rwetsiba, Mustapha Nsubaga, and Joshua Mabonga. Optimal Patrol Planning for Green Security Games with Black-Box Attackers. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 10575 LNCS, pages 458–477. Springer Verlag, 2017. ISBN 9783319687100. doi: 10.1007/978-3-319-68711-7_24.

[92] Bin Yang, Yongsheng Ding, and Kuangrong Hao. Area coverage searching for swarm robots using dynamic Voronoi-based method. In *Chinese Control Conference, CCC*, volume 2015-Septe, pages 6090–6094. IEEE Computer Society, sep 2015. ISBN 9789881563897. doi: 10.1109/ChiCC.2015.7260592.

[93] Rong Yang, Benjamin Ford, Milind Tambe, and Andrew Lemieux. Adaptive Resource Allocation for Wildlife Protection against Illegal Poachers. *AAMAS '14*, pages 453–460, may 2014. doi: 10.5555/2615731.2615805. URL www.ifaamas.org.

[94] Chao Zhang, Shahrzad Gholami, Debarun Kar, Arunesh Sinha, Manish Jain, Ripple Goyal, and Milind Tambe. Keeping Pace with Criminals: An Extended Study of Designing Patrol Allocation against Adaptive Opportunistic Criminals. *Games*, 7(3):15, jun 2016. ISSN 2073-4336. doi: 10.3390/g7030015. URL http://www.mdpi.com/2073-4336/7/3/15.

[95] Jian Zhang. Occlusion-aware UAV path planning for reconnaissance and surveillance in complex environments. In *IEEE International Conference on Robotics and Biomimetics, ROBIO 2019*, pages 1435–1440. Institute of Electrical and Electronics Engineers Inc., dec 2019. ISBN 9781728163215. doi: 10.1109/ROBIO49542.2019.8961649.

[96] Yanan Zhao, Mingchu Li, and Cheng Guo. Developing Patrol Strategies for the Cooperative Opportunistic Criminals. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 11944 LNCS, pages 454–467. Springer, 2020. ISBN 9783030389901. doi: 10.1007/978-3-030-38991-8_30.

[97] Ziyang Zhen, Yan Chen, Liangdong Wen, and Bing Han. An intelligent cooperative mission planning scheme of UAV swarm in uncertain dynamic environment. *Aerospace Science and Technology*, 100, may 2020. ISSN 12709638. doi: 10.1016/j.ast.2020.105826.

[98] Fang Zong, Hongguo Xu, and Huiyong Zhang. Prediction for traffic accident severity: Comparing the bayesian network and regression models1. *Mathematical Problems in Engineering*, 2013, 2013. ISSN 1024123X. doi: 10.1155/2013/475194.