

# **Development of Risk-Informed UAV Pathfinder Based on A\* Algorithm**

**Micko Lesmana**

**11201901010**



**Undergraduated Thesis**

Department of Aviation Engineering  
International University Liaison Indonesia  
BSD City  
7 July 2024

# **Approval Page**

## **Report**

### **Risk Informed UAV Pathfinding, UPX**

**3 December 2022**

**Micko Lesmana**

**11201901010**

**Department of Aviation Engineering**

**Acknowledge by:**

# CONTENTS

<b>Contents</b>	<b>2</b>
<b>List of Figures</b>	<b>6</b>
<b>List of Tables</b>	<b>9</b>
<b>List of Abbreviations</b>	<b>10</b>
<b>1 Introduction</b>	<b>12</b>
1.1 Introduction . . . . .	12
1.2 Research purpose . . . . .	12
1.3 Research scope . . . . .	12
1.4 Research problem . . . . .	13
1.5 Significance of the study . . . . .	13
1.6 Research question . . . . .	13
<b>2 Literature Review</b>	<b>14</b>
2.1 Overview . . . . .	14
2.2 Ground risk model . . . . .	14
2.2.1 Failure model . . . . .	16
2.2.2 Impact location model . . . . .	19
2.2.3 Recovery model . . . . .	21
2.2.4 Stress model . . . . .	22
2.2.5 Exposure model . . . . .	23
2.2.6 Incident stress model . . . . .	24
2.2.7 Harm model . . . . .	24
2.3 Pathfinding Algorithms . . . . .	26

2.3.1	Bread-First Search . . . . .	26
2.3.2	Depth-first search . . . . .	27
2.3.3	Djikstra's algorithm . . . . .	28
2.3.4	Greedy best-first search . . . . .	28
2.3.5	A* . . . . .	29
2.3.6	Rapidly-exploring random tree . . . . .	31
2.3.7	Ant colony optimization . . . . .	34
2.3.8	Generative pathfinding AI . . . . .	36
2.3.9	Justification for choosing risk-based A* . . . . .	37
<b>3</b>	<b>Background Theory and Methodology</b>	<b>38</b>
3.1	Case study . . . . .	38
3.1.1	UAV Specification . . . . .	38
3.2	Objective location . . . . .	39
3.3	General Methodology . . . . .	40
3.4	Data retrieval and processing . . . . .	41
3.4.1	Online geographic data . . . . .	41
3.4.2	Population data . . . . .	41
3.4.3	Geographic Information System . . . . .	42
3.5	Ground risk map calculation . . . . .	42
3.5.1	Geographic feature . . . . .	43
3.5.2	Population density layer . . . . .	44
3.5.3	Obstacle layer . . . . .	45
3.5.4	Sheltering Factor Layer . . . . .	46
3.5.5	No-fly zone layer . . . . .	48
3.6	Ground Risk Assessment . . . . .	48
3.6.1	Failure rate . . . . .	50

3.6.2	Descent Event . . . . .	51
3.6.3	Wind effect . . . . .	55
3.6.4	Probability of Impacting an Individual . . . . .	55
3.6.5	Probability of Fatality . . . . .	56
3.7	Finding optimal path . . . . .	57
3.7.1	Problem statement . . . . .	57
3.7.2	Graph . . . . .	59
3.7.3	Dijkstra's algorithm . . . . .	59
3.7.4	A* algorithm . . . . .	61
3.8	Implementation of risk-based UAV path finder . . . . .	64
3.8.1	Instrumentation . . . . .	64
3.8.2	General working principle . . . . .	65
3.8.3	UPX Python and packages . . . . .	69
<b>4</b>	<b>Results</b>	<b>71</b>
4.1	PDF risk map . . . . .	71
4.1.1	Ballistic descent . . . . .	71
4.1.2	Glide . . . . .	73
4.1.3	Fly Away . . . . .	75
4.1.4	Parachute . . . . .	77
4.1.5	Total Fatality . . . . .	79
4.2	Pathfinder result . . . . .	81
4.2.1	Regular A* . . . . .	82
4.2.2	Risk A* . . . . .	82
4.2.3	Risk result . . . . .	84
4.2.4	Summarized result . . . . .	88
4.3	Comparison with previous research . . . . .	89

<b>5 Conclusion, and Recommendation</b>	<b>90</b>
5.1 Summary . . . . .	90
5.2 Conclusion . . . . .	90
5.3 Recommendation . . . . .	91
5.3.1 For future research . . . . .	91

## LIST OF FIGURES

2.1	Components of ground risk model adopted from Primatesta, Rizzo, and la Cour-Harbo, 2020 . . . . .	15
2.2	Area exposed of casualty area . . . . .	20
2.3	Range covered by parachute landing . . . . .	22
2.4	Generalized logistic fatality function given energy impact . . . . .	25
2.5	BFS graph tree . . . . .	27
2.6	DFS graph tree . . . . .	27
2.7	Differences of cost function between GBFS and Dijkstra's . . . . .	29
2.8	Simplified RRT Graph . . . . .	33
2.9	RRT vs RRT* . . . . .	34
2.10	Graph with ACO to optimize shortest total route . . . . .	35
3.1	Operation objective . . . . .	40
3.2	General overview . . . . .	40
3.3	Grid risk matrix cell representation . . . . .	43
3.4	Talon ground risk map fatality distribution . . . . .	43
3.5	Population density map . . . . .	45
3.6	Height obstacle map . . . . .	46
3.7	Shelter factor map with varying levels of sheltering provided by physical structures and objects across different locations. . . . .	47
3.8	Type of descent event probability being summed . . . . .	50
3.9	Ballistic descent by aircraft type . . . . .	52
3.10	Glide descent by aircraft type . . . . .	53
3.11	Fly away descent by aircraft type . . . . .	54
3.12	Parachute descent by aircraft type . . . . .	55
3.13	Dijkstra's pseudocode algorithm . . . . .	61

3.14 A* pseudocode algoithm . . . . .	63
3.15 UPX architecture design . . . . .	65
3.16 Tesselation process, color does not represent anything . . . . .	66
3.17 Spatial join and attribute aggregation, color does not represent anything	67
3.18 Graph convert, color does not represent anything . . . . .	68
4.1 Talon ballistic PDF . . . . .	71
4.2 Disco ballistic PDF . . . . .	72
4.3 DJI Phantom 4 ballistic PDF . . . . .	72
4.4 DJI Inspire 2 ballistic PDF . . . . .	73
4.5 Talon glide PDF . . . . .	74
4.6 Disco glide PDF . . . . .	74
4.7 DJI Phantom 4 glide PDF . . . . .	75
4.8 DJI Inspire 2 glide PDF . . . . .	75
4.9 Talon fly away PDF . . . . .	76
4.10 Disco fly away PDF . . . . .	76
4.11 DJI Phantom 4 fly away PDF . . . . .	77
4.12 DJI Inspire 2 fly away PDF . . . . .	77
4.13 Talon parachute PDF . . . . .	78
4.14 Disco parachute PDF . . . . .	78
4.15 DJI Phantom 4 parachute PDF . . . . .	79
4.16 DJI Inspire 2 parachute PDF . . . . .	79
4.17 Talon total fatality PDF . . . . .	80
4.18 Disco total fatality PDF . . . . .	80
4.19 DJI Phantom 4 total fatality PDF . . . . .	81
4.20 DJI Inspire 2 total fatality PDF . . . . .	81
4.21 Regular A* flight path . . . . .	82

4.22	Talon risk A* flight path . . . . .	82
4.23	Disco risk A* flight path . . . . .	83
4.24	DJI Phantom 4 risk A* flight path . . . . .	83
4.25	DJI Inspire 2 risk A* flight path . . . . .	84
4.26	Aircraft risk score . . . . .	86

## LIST OF TABLES

3.1	UAV Specifications . . . . .	39
3.2	Sheltering factor values for different types of features. . . . .	47
3.3	MTBF and $P_{fatal}$ for Various UAVs . . . . .	51
3.4	Energy correlate to fatality . . . . .	57
4.1	Total Pathway Risk . . . . .	86
4.2	Pathway Risk Statistic . . . . .	87
4.3	Areal map risk statistic of North Jakarta (%) . . . . .	88

## LIST OF ACRONYMS

<b>A*</b>	A-star
<b>AIS</b>	Abbreviated Injury Scale
<b>API</b>	Application Programming Interface
<b>BBN</b>	Bayesian Belief Networks
<b>BC</b>	Blunt Criteria
<b>CFIT</b>	Controlled Flight Into Terrain
<b>CRS</b>	Coordinate Reference System
<b>DOJC</b>	Dropped or Jettisoned Component
<b>DoD</b>	United States Department of Defense
<b>ETA</b>	Event Tree Analysis
<b>FAA</b>	Federal Aviation Administration
<b>FMECA</b>	Failure Mode, Effects and Criticality Analysis
<b>FTA</b>	Fault Tree Analysis
<b>GHSL</b>	Global Human Settlement Layer
<b>GIS</b>	Geographic Information System
<b>GPS</b>	Global Positioning System
<b>GRM</b>	Ground Risk Model
<b>geoJSON</b>	Geographic JavaScript Object Notation
<b>EPSG</b>	European Petroleum Survey Group
<b>ICAO</b>	International Civil Aviation Organization
<b>KE</b>	Kinetic Energy
<b>LOC</b>	Lost of Control
<b>MTBF</b>	Mean Time Between Failure
<b>OSM</b>	OpenStreetMap
<b>PDF</b>	Probability Density Function
<b>RRT</b>	Rapidly Exploring Random Tree

<b>UDS</b>	Unpremediated Descent Scenario
<b>UAS</b>	Unmanned Aerial System
<b>UAV</b>	Unmanned Aerial Vehicle
<b>UPX</b>	UAV Pathfinding eXtension
<b>CE</b>	Casualty Expectation
<b>AGL</b>	Above Ground Level
<b>BFS</b>	Bread-First Search
<b>DFS</b>	Depth-First Search
<b>ACO</b>	Ant Colony Optimization
<b>AI</b>	Artificial Intelligence
<b>GBFS</b>	Greedy Best-First Search
<b>JRC</b>	Joint Research Centre

# CHAPTER 1

## INTRODUCTION

### 1.1 Introduction

Unmanned Aerial Vehicle (UAV) also known as a drone, is getting more recognition in recent years. The booming of UAV is caused by the component needed to manufacture the UAV, especially microcontrollers getting cheaper, and smaller Austin and Brewster, 2024. Compare to previous UAV iteration, UAV nowadays offers much wider capability, such as enhanced range, the ability to mount hardware, and camera for example, and of course easily deployable without requiring many technicians to maintain them.

UAV are often being used in area where it is dangerous, impractical, and expensive for human. The affordable entry point of using a UAV incentivized emerging industries to exploit the capability that UAV offer. Not only that but the existing industry is also getting interested to use UAV as an augmentation to its existing operational infrastructure. And that industries are, logistics companies, agricultural companies, multimedia companies, and surveyors to name a few DJI, 2023.

Although it provides a great solution, UAV have the potential to pose risks to people, property, and other aircraft. Many countries have regulations in place that require UAV to be operated in a safe and responsible manner FAA, 2016. Risk-based UAV path planner can help to ensure that UAV are operated safely and responsibly. By considering potential risks and hazards, a risk-based approach can help to reduce the likelihood of accidents or incidents occurring during the flight. By considering potential risks and hazards, this approach can help to reduce the likelihood of accidents or incidents occurring during deployment.

Other than its impacts towards public, UAV can also potentially impact the environment, for example by disturbing wildlife or damaging natural habitats Holland, 2015. UAV path planner can help to minimize these impacts by establishing rules around the use of UAV in sensitive areas.

### 1.2 Research purpose

For ensuring the safety and efficiency of UAV operations, this research aims to develop a UAV trajectory planning based on the risk assessment on the ground.

### 1.3 Research scope

- UAV model that being used in this research are fixed wing and multirotor.
- This research only use point mass representation with generalize aerodynamic equation.

- Based on the point above, UAV failure is assumed to be only power failure, without damage to its shape.
- Risk assessed in this paper are combination of human exposed area and area that can cause major accident such as airfields, and power station.
- This research use Jakarta, Indonesia as its geographic location, since it has varying building density, rivers, and random patch of lands.

## 1.4 Research problem

- There are risks associated when of UAV falls on the ground.
- Modeling physical interaction of UAV crash.
- Selection of pathfinding algorithm to get optimum time.

## 1.5 Significance of the study

UAV Pathfinder application could be one of the toolkit for UAV operator to plan their safe operation of flying their UAV, reducing incident and accident. Developer and researcher could use the library itself as an extension or starting point in the development of the Unmanned Aerial System (UAS) Traffic Management application.

## 1.6 Research question

- Which kind of geographic data and data processing model perfectly suited for pathfinding risk assesment.
- How to statistically create the best approximation of crash dynamics.
- What kind of pathfinding method that produces the quasi-optimal path without sacrificing computational time.

## CHAPTER 2

# LITERATURE REVIEW

### 2.1 Overview

In this part of the literature review, we will discuss two main concepts: the Ground Risk Model and the Path Finder. In the first part, we will examine the Ground Risk Model, which involves discussing the concept of probability within a spatial context. Consider this scenario: if we want to calculate the likelihood of someone being hit by a drone in a specific location, how do we model this phenomenon? As you can see, scenarios like this are key to understanding the Ground Risk Model.

Next, we will explore the Path Finder. After identifying a model that answers the question of where a UAV is most likely to hit someone, we move on to the next question: how do we navigate the UAV or provide the safest route to minimize the risk of severe casualties? This is why we will discuss the Risk-Based Path Finder.

### 2.2 Ground risk model

A ground risk model is a type of assessment tool that is used to evaluate the potential risks associated with the operation of UAV. This model takes into account a variety of factors, including the physical environment in which the UAV will be operating, the capabilities, and the limitations of the UAV itself. The goal of a ground risk model is to identify and prioritize potential risks in order to produce informed decision-making and risk management strategies for UAV operations.

We found a total of 22 different research papers that contributed to our study on ground risk models. Various approaches have been taken to develop ground risk models for UAVs. One approach is the use of scenario-based methods, as discussed in the research of Ancel et al., 2017, which involves the development of detailed scenarios that describe potential UAV operations and the associated risks.

To clarify further, in the research done by Ancel et al., 2017, there are two common methods for describing the severity of accidents using scenario-based approaches. These scenarios are typically coupled with safety and hazard levels on a per-operation basis. The first method utilizes a lookup table to determine the severity of an accident. The second method represents the scenario as a directed acyclic graph, a type of tree graph, where we start at the root node and traverse based on yes or no questions. These questions can range from simple ones, such as whether the UAV hit a person, to more detailed ones, such as whether the accident caused harm.

This method is computed using techniques such as risk tree diagrams, Bayesian belief networks Ancel et al., 2017, and hazard risk assessment matrices Barr et al., 2017. Although these methods might produce different analyses, they generally

describe the same phenomena.

From the description above, it appears that scenario-based models do not suit our needs for a spatially based ground risk model. Essentially, they reduce to questions of general progression from the point of impact to the accident caused by the drone. They provide no information on how the position of the UAV affects the casualty risk of failed operations and only serve as a preliminary assessment.

To add a spatial component to our models, we can simply incorporate location information into the yes/no question tree graph. What I mean by this is to propose an idea where, in a certain area, if a UAV somehow falls, does it cause harm, and based on the surrounding environment, does that harm result in fatality?

This brings us to the most widely used method, so called ground probabilistic risk assessment methods, which involve the quantification of risks per area using probability and consequence metrics. Although a greater number of research papers use this method, as compiled by Washington et al., 2017, our concept shown in figure 2.1 is derived from the research conducted by Primatesta, Rizzo, and la Cour-Harbo, 2020.

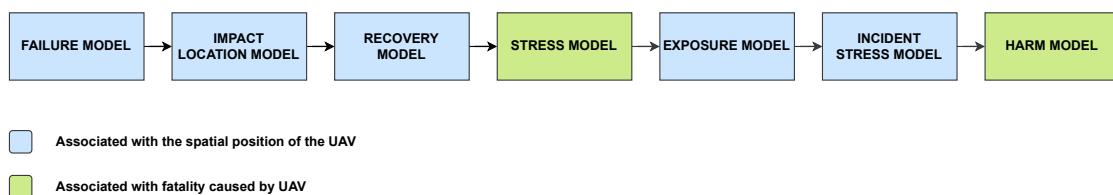


Figure 2.1: Components of ground risk model adopted from Primatesta, Rizzo, and la Cour-Harbo, 2020

A ground probabilistic risk assessment model, shortened to ground risk model, of a UAV typically consists of several subcomponents that are used to assess various factors that can impact the safety and effectiveness of UAV operations. These components in the ground risk model include the failure model, impact location model, recovery model, stress model, exposure model, incident stress model, and harm model, which will be explained further below.

However, since this is the most widely used approach, it is not surprising that this approach is not new, as many earlier researchers have already followed guidelines to model these components. For example, Breunig et al., 2018 provides a general model guideline. These guidelines focus more or less on the probability of said accidents causing fatalities but do not compute the overall area risk.

Washington et al., 2017 offers a meta-analysis of different kinds of research using more or less the same paradigm to describe the ground risk model. Washington et al., 2017 itself uses the method of its analysis based on the guidelines from Pat-Cornell, 1996, which studied the level of analytical detail when modeling risk. Pat-Cornell, 1996 provides us with a degree of certainty ranging from one to six, where one uses

a very simple model and six accounts for every possible situation that can affect the model.

After a thorough discussion, we derived our study based on Primatesta, Rizzo, and la Cour-Harbo, 2020 simply because it shows the most comprehensive way to calculate the ground risk model, which can be summarized into Equation 2.1.

$$P_{fatal}(x) = P_{failure}(x) \cdot P_{impact}(x) \cdot P_{recovery} \cdot P_{stress} \cdot P_{exposure}(x) \cdot P_{incident}(x) \cdot P_{harm}(x) \quad (2.1)$$

All of these models are represented as the probability of the said event-model occurring. Where  $P$  is a probability of an event and  $x$  is the location on a grid map. If we take a closer look at Equation 2.1, some  $P_n$  terms do not contain any location parameter. This is because those probabilities are applied as global variables to every area.

Back to the model, since all models occur consecutively from the failure model to the harm model, each component of the ground risk model is multiplied by the others. Although many papers do not use all of these components in their ground risk models and often use different methodologies to analyze the same component, the fundamental approach remains consistent. For example, Primatesta, Rizzo, and la Cour-Harbo, 2020 only use the event, impact location, and fatality models, while Kim and Bae, 2022 use the event, impact, exposure, incident stress, and harm models. Despite apparent differences in modeling ground risk, researchers sometimes combine multiple subcomponents into one. For instance, in Primatesta, Rizzo, and la Cour-Harbo, 2020, the fatality model component is, in fact, a combination of the incident, stress, and harm models. All the models in ground risk assessment are explained in following sub-chapters.

### 2.2.1 Failure model

A failure model is the first subcomponent in the ground risk model that evaluates the potential for failures in UAV. This potentially leads to a crash impact in a given variety of factors that can contribute to UAV failures, including design and manufacturing defects, operational errors, and environmental factors.

The failure model can help to understand on how the UAV would behave after experiencing the failure, Washington et al., 2017 described four distinct events following the failure. Which are Unpremediated Descent Scenario (UDS), Lost of Control (LOC), Controlled Flight Into Terrain (CFIT), and Dropped or Jettisoned Component (DOJC).

All of these failure events describe how the UAV crashes into the ground. UDS is a mode where the drone automatically enters a safe mode and tries to land as safely as possible. Drone manufacturers often implement this feature when critical components are detected to be defective. UDS can also be triggered when the UAV

loses its signal to its pilot. Most mid- and high-priced drones usually have this feature, such as those from DJI, 2024.

Controlled flight into terrain occurs when the UAV still has flight authority, but its main propulsion is no longer working or its battery is at very low power, allowing it to only move its control surface servos. It is similar to UDS, except that in UDS, the UAV still has propulsion power. Also like UDS, if the UAV is advanced enough Nemire, 2015, it can auto-navigate to a safer area.

Loss of control scenario occurs when the UAV is no longer responsive and has lost its navigation system, causing it to fly away while maintaining its velocity and altitude until it runs out of energy or crashes into terrain. This condition is much more degraded compared to UDS. Lastly, the "dropped or jettisoned" scenario, probably the least common failure stated by Washington et al., 2017, occurs when some components detach from the drone and fall to the ground.

Although the definition of a ways that UAV can fell into the terrain, it does not necessarily help us calculate the likelihoodness of a UAV to fell into the ground. With that being said this definition can help us to formulate the descent model of the ground risk model.

If we want to calculate the probability of the drone starts to fail, there are two main models that we can use. The first one is failure model proposed by Breunig et al., 2018 and Weibel and Hansman, 2012 where it is based on historical data. Data driven historical data are using previous UAV crash, and its UAV specification as a starting point for data interpolation and extrapolation for calculating so called Mean Time Between Failure (MTBF).

Unfortunately since UAV industries is in its infancy and an added fact with many variants of UAV already exist in short period of time. There are not enough historical data to make an interpolation, which is an issue commented by Breunig et al., 2018, many model using an MTBF simply use a standard assumed by ARC, 2015 which is average of 1 failure per 100 hours of flight for small UAV.

Although with its limitation, there are few research that trying to emulate MTBF to solve target level of safety to help certifying authorities to make a proper judgment in regards to ever growing small UAV industry. Such model is developed by Burke et al., 2011, where they make a system-level airworthiness tools that predict how safe a UAV is based on the sizes of the aircraft which is explained the equation 2.2 below.

$$MTBF = \frac{\rho\pi b^2}{2.78} = k(\rho\pi b^2) \quad (2.2)$$

Where  $\rho$  is population density,  $b$  is geometric wingspan of the drone, and  $k$  is population correction factor, where  $k$  is  $\frac{1}{2.78}$  in a setting where the population density is calculated per square miles. This equation is derived from target level

safety algorithm proposed by Burke et al., 2011. Which then derived from Casualty Expectation (CE) equation 2.3

$$CE = P_f \times P_d \times A_l \times P_k \times S \quad (2.3)$$

Where  $CE$  is expected casualty rate per flight hours,  $P_f$  is the probability of failure in failure per flight hour,  $P_d$  is the population density in people per square mile,  $A_l$  is the dimensionless lethal area,  $P_k$  is fatality probability, and  $S$  is the dimensionless shelter factor. We then change the  $P_d$  as a  $\rho$  to signify the average in time for population in a given mission area.  $A_l$  is modeled from using the maximum circular area of the UAV itself which is the maximum wing span  $b$  in feet or in the case of multirotor, maximum diagonal distance. Therefore  $A_l$  can be substitute into  $\pi b^2$ . Pointer given by Burke et al., 2011 simply using  $A_l$ , which calculate in feet, will give an error when being multiplied by  $\rho$  where it is in square miles. Therefore Burke et al., 2011 give an error correction to divide  $\rho$  with the number of  $1 \text{ ft}^2$  in a  $1 \text{ mile}^2$  which roughly is 27.8 million.

Burke et al., 2011 normalized the  $P_k$  to 1 to argue based on the fact provided by ARC, 2015. This normalization came from the calculation of kinetic energy for small UAV with a weight of 64 lb with the speed of 100 ft/s, in which the UAV is still in non-lethal configuration. Move on to the  $S$ , it describe to be in range where 1 is non-covered area and 0 is fully sheltered. Fully sheltered in this case is when an UAV with the weight of 64 lb and the speed of 100 ft/s crash into a populated building, the occupants will be in no harm condition. If we rearrange the equation and set the CE into  $1 \times 10^{-7}$  which also has the same meaning with equivalent level safety, where we will get resulting equation 2.4

$$10^{-7} = P_f \times \frac{\rho}{2.78 \times 10^{-7}} \times \pi b^2 \quad (2.4)$$

$$P_f = \frac{\rho}{2.78} \times \pi b^2 \quad (2.5)$$

$$P_f = \frac{k}{2.78} \times \pi b^2 \quad (2.6)$$

Since  $P_f$  is probability of failure per flight hours which also the inverse of the MTBF, we can change the  $P_f$  into MTBF. And if we rearrange the equation one more time we will get the same equation of 2.2. However since that equation is in US customary unit, we must convert it into metric system, where we divided  $1 \text{ m}^2$  per  $1 \text{ km}^2$ . Fortunately it is quite simple, as we just need to change error correction factor from  $2.78 \times 10^{-7}$  into  $10^{-6}$  which resulted into equation 2.7

$$MTBF = \frac{\rho \pi b^2}{10} = k(\rho \pi b^2) \quad (2.7)$$

Aside from system-level airworthiness tool, there are systematic methodologies used in risk management and safety engineering to assess the failure of UAV. These tools are Fault Tree Analysis (FTA), Event Tree Analysis (ETA) , and Failure Mode, Effects and Criticality Analysis (FMECA) Barr et al., 2017. Where focuses on identifying root causes of specific failures using logical diagrams. ETA evaluates the outcomes of initiating events and their subsequent paths. And FMECEA identifies potential failure modes, their effects, and prioritizes them based on criticality.

Also there is a new development using a machine learning model such as Bayesian Belief Networks (BBN) Ancel et al., 2017 to be used for failure analysis. Where Ancel et al., 2017 provide the most detailed failure risk model incorporating real-time monitoring of UAV component status to determine the likelihoodness of failure. BBN itself is a graphical model used for probabilistic reasoning and decision-making, which represents relationships among variables through directed acyclic graphs. In the context of UAV failure analysis, BBNs can help identify the likelihood of various failure modes by modeling the dependencies and interactions between different system components and environmental factors.

The model such as FTA, ETA, and FMECA sadly only giving us an answer to a question of what is and are the cause of the crash. BBN proposed by Ancel et al., 2017 also is a only a framework for implementing realtime decision making when it is in operation. Some researcher also assume a constant MTBF of 100 hours proposed by the ARC, 2015. We also make use the assumption that the failure's probability of happening is constant. Therefore, variables like phase, length, environment circumstances, or mission profile that might affect the chance of failure are not taken into account.

### 2.2.2 Impact location model

The impact location model describes the position and size of the impact region caused by a UAV crash. Various models are employed to determine this impact area, ranging from single-point impact models to more complex models that account for multiple potential impact points along a UAV's trajectory. The latter strategy establishes a perimeter around which all of the UAV's possible impact points are located, calculated based on the maximum descent trajectory distance. The impact location model can be formulated using empirical or hypothetical approaches.

Empirical models leverage historical aircraft event data as a starting point, using interpolation and extrapolation algorithms to identify potential impact sites. For example, Melnyk et al., 2014 categorizes these models based on the study by Ale and Piers, 2000, further refining impact predictions through weight-based classification.

Where generally speaking, the hypothetical model in research often attempts to establish a relationship between the size of the UAV and the extent of its impacted area. For instance, models proposed by Primatesta, Rizzo, and la Cour-Harbo, 2020 and Burke et al., 2011 utilize aerodynamic principles to simulate impact behavior. This model can be further categorized into ballistic descent, gliding descent, and

planform models. The ballistic descent calculates the trajectory based on a ballistic profile, while the gliding model assumes the UAV will continue descending along its glide path. The planform model, on the other hand, uses the aircraft's wing and layout aerodynamic properties to determine the descent trajectory.

According to Primatesta, Scanavino, et al., 2020, it is very costly to calculate full aerodynamic properties per cell in a large grid, which is why many researchers rely on gliding and ballistic descent calculations for impact modeling. Additionally, Primatesta, Rizzo, and la Cour-Harbo, 2020 introduced scenarios such as parachute descent, where the UAV deploys a parachute during descent, and fly-away descent, where the UAV maintains a stable cruise after the pilot loses control.

Primatesta uses a method from the US Department of Transportation FAA, 2000 for calculating the casualty area of a vertically falling inert piece of debris. According to this method, the casualty area is a circle whose radius is the sum of the radius of a circle with an area equal to the largest cross-sectional area of the debris and the radius of a human being ( $30\text{ cm}$ ). For these calculations, an acceptable dimension for a human is considered to be a  $182.5\text{ cm}$  tall cylinder with a  $30\text{ cm}$  radius. To account for any horizontal velocity component, such as wind or trajectory angles, the basic casualty area can be calculated using the following equation and figures:

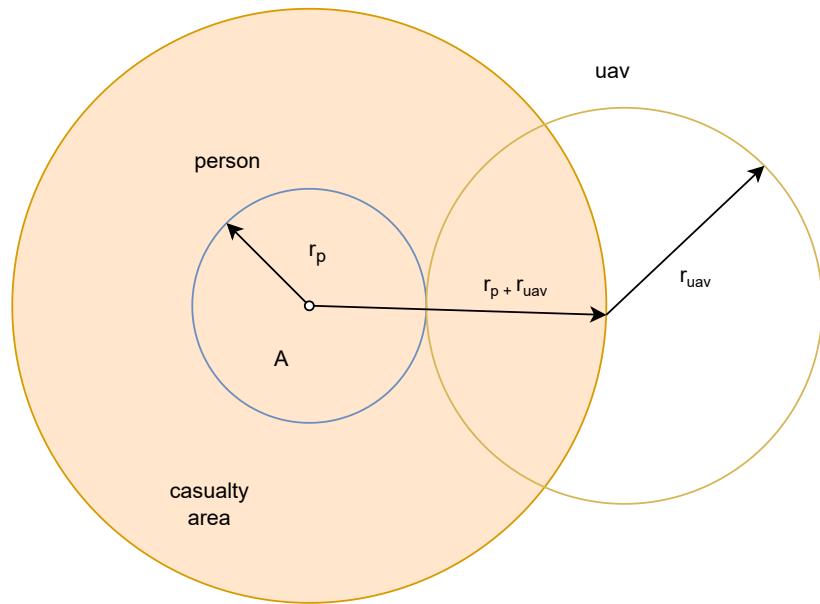


Figure 2.2: Area exposed of casualty area

$$A_{exp}(\theta) = \pi(r_p + r_{uav})^2 \sin(\theta) + (r_p + r_{uav})(h_p + r_{uav}) \cos(\theta) \quad (2.8)$$

Where  $A_{exp}$  is an area exposed in  $\text{m}^2$ ,  $r_p$  and  $h_p$  is the average radius ( $0.3\text{ m}$ ) and average height ( $182.5\text{ m}$ ) of a person respectively in meter,  $r_{uav}$  is the radius of the UAV in meter, and  $\theta$  is the impact angle on the ground. This conservative com-

putation takes into account the maximum radius of the UAV, ensuring a thorough evaluation of the impact area.

### 2.2.3 Recovery model

A recovery risk model for UAV is a model used to assess the likelihood of successful recovery of a UAV to a optimal or degraded operational state following a failure or malfunction. The model takes into account a variety of factors, the nature and severity of the failure, and the environment in which the recovery is taking place. The recovery risk model also takes into account the UAV design, capabilities and limitations. For example some UAV has parachute failsafe built into it. This parachute is then deployed when the control logic of the UAV detected major system breakdown. Some high-end UAV will enter a automatic landing or controlled crash into terrain before it deploys its parachute. Although not always it is a general rule of thumb that the UAV that has already equiped by parachute usually already has an advance flight controller. Please note that at the time of this writing, there are little to none database containing every specification of each highend drone.

Although there is a significant in studying UAV recovery behaviour, only three researches that managed to include recovery model in its ground risk model. Shelley, 2016 calculated the safest highest possible flight altitude of the UAV based on its weight, or to be precised its resulted kinetic impact. It is showed that under any circumstances UAV with the weight over 1.5 kg must not be flown over a dense population. The study also give a benefit to equiping UAV with parachute will enable most UAV from only permitted to flown in tens of feet to 400 ft Above Ground Level (AGL).

The reasoning behind the small numbers of parachute descent being incorporate into GRM is because it assumes if the UAV already deployed its parachute, there are little to no threat to population. Hence, most of the reasearcher just give if else statements, where to just ignore its calculation if parachute is installed into the UAV. However some researcher goes into detail to see the effect of UAV impacting the surrounding area even with parachute, after all parachute can be drag by wind and still conatains an energy.

Primatesta, Rizzo, and la Cour-Harbo, 2020 use parachute descent calculation based on the study from Bleier et al., 2015, where it calculate possible area where the UAV will land. Bleier et al., 2015 use its calculation based on the relation ship of  $s = V \cdot t$ , where  $s$  is distance,  $V$  is velocity and  $t$  is time. The relationship is expanded into include wind effect. Which resulted into equation 2.9 and 2.10.

$$t_d = \frac{h}{V_{sink}} \quad (2.9)$$

$$\begin{pmatrix} x \\ y \end{pmatrix} = (V_{\text{wind}} \pm \Delta V_{\text{wind}}) \frac{h}{V_{\text{sink}}} \cdot \begin{pmatrix} \cos(\alpha_{\text{wind}} \pm \Delta \alpha_{\text{wind}}) \\ \sin(\alpha_{\text{wind}} \pm \Delta \alpha_{\text{wind}}) \end{pmatrix} \quad (2.10)$$

Where  $t_d$  is time from deployed state to the ground,  $v_{\text{sink}}$  is the descent velocity,  $V_{\text{wind}}$  and  $\Delta V_{\text{wind}}$  is wind speed and random change in wind speed respectively,  $\alpha_{\text{wind}}$  and  $\Delta \alpha_{\text{wind}}$  is wind angle and random change in wind angle respectively

We can see how the equation works from figure 2.3 below. Please note that Bleier et al., 2015 only shows the equation for north down east coordinate system, where we use general cartesian coordinate for the figure. Back to the figure, if we view how the UAV is fell from the sky from a top down perspective, where the 0,0 is the point of deployment and has a wind affecting its descent trajectory. The UAV will take its time to be dragged into the ground by the wind to the red area, where the left most side is the furthest possible point, and vice versa for the right most side of the red area.

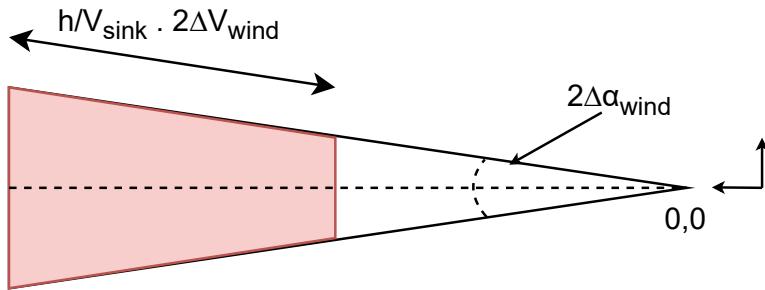


Figure 2.3: Range covered by parachute landing

#### 2.2.4 Stress model

The stress model explains the unpredictability of harmful circumstances occurring at a specific time and place. Or in other term what kind of harm are received by the population. For instance, stresses may be described as kinetic energy if the outcome to be calculated was human physical harm modeled in Kinetic Energy (KE). While secondary effects such as explosions and the release of hazardous chemicals will have distinct related stress qualities that need to be modeled differently. There are model proposed by Ball et al., 2012 to analyze the consequences of heat radiation and explosions which often produces by chemical reaction. 18 of 22 of the papers are using only kinetic energy as an only harm type in their ground risk model. This is not surprising since majority of the damage done into human body is kinetic energy based on the bomb explosion effect into human body research by Harwick et al., 2007.

Depending on the kind of UAV, different stress factors that relate to one or more harmful processes may exist. Additionally, the UAV's characteristics and its trajectory may be changed to alter the stress profile at the site of impact. Through one or

more harm mechanisms, the various stresses may be linked to the undesirable consequence. Examples of physical injury to humans include blunt force, penetration, crushing, blast, burns, lacerations.

It should be mentioned that trauma generated by a blunt force collision is the main cause of injury taken into account by current models. It is possible that blunt force trauma is not the main cause of injury for UAV with tiny mass. Smaller multirotor UAV, for instance, have a higher tendency to lacerate individuals and inflict physical harm. In the stress model, laceration injuries must be taken into account, according to the Small, 2009.

The KE at impact was calculated by the models using a variety of mass and speed values, which is why the results of the stress models can differ from one another. For instance, Ancel et al., 2017 and other models simply assume the maximum velocity for estimating the impact energy for small UAV flying at low altitudes. Although using terminal velocity would be an overly conservative estimate, Dalamagkidis et al., 2008 evaluation of the impact KE employs a similar assumption.

### **2.2.5 Exposure model**

An exposure model is a method for determining how affected individual or property may be to dangers related to the use of UAV.

A uniform exposure model is the model of population exposure that is most frequently utilized. The use of this approach overlooks possible population clustering within a specified geographic region and, as a result, averages the potential exposure factors contribution to the overall risk assessment. Such methodologies tend to overestimate or underestimate the risk value associated with the current setup because they do not provide sufficient information about the spatial and temporal distribution of the population.

Models such as those created by Burke et al., 2011 and Melnyk et al., 2014 aim to add temporal features, acknowledging how individuals spend their days, taking such factors into consideration. An illustration of a thorough exposure model that emphasizes the time dependencies in population exposure is given by Melnyk et al., 2014. Research conducted by Klepeis et al., 2001 found that individuals spent 68.7% of their time at home, 7.6% outside, 5.5% in cars, 5.4% at factories or offices, and another 12.8% in various indoor settings. This data is used by Melnyk et al., 2014 to specify their exposure model along with the information.

As expected population density data gathered for exposure model came from census, geographical data from satellite imagery. Depending on when each study was done and what regions were being analyzed, a different census data source was used. For instance, Burke et al., 2011 used data from the U.S. Census Bureau, and Clothier et al., 2007 used data from the Australian Bureau of Statistics. And to provide non uniform population data, cellular networks were used by Ancel et al., 2017 to create a near-real-time model of the population distribution.

### 2.2.6 Incident stress model

The incident stress model describes how much harm an individual is exposed to. Where in other term, how protected an individual from getting hit by the debris. Some papers such as Primatesta, Rizzo, and la Cour-Harbo, 2020, Cour-Harbo, 2020, and Dalamagkidis et al., 2008 to name a few, take into account an attenuating factor based on the sheltering effect to calculate how much kinetic energy a certain building can hold. In the model described by Ball et al., 2012, the amount of KE that is absorbed by the shelter as a result of its deformation, breaking, and movement is then subtracted from the overall risk posed by the system.

Most of the shelter model in the paper who use it is split into four or five different type of building. From no shelter, sparse tree, small building, medium size building, and large building. Where small building typically describe a residential house or a small shop. Medium sized building where it describe a 2 stories building. And large building is from 3 stories and up. Based on investigations undertaken by the Columbia Accident and Investigation Board and the Department of Defense, Melnyk et al., 2014, where it describe a method for calculating the risk from population distribution, most of the shelter are mostly residential structures and commercial buildings.

As previously mentioned each of these shelter types has its attenuating factor, where it is just a coefficient factors to reduce the ammount of kinetic energy. It can be described as an simple proportional equation 2.11 below

$$S \cdot E_k \quad (2.11)$$

Where  $S$  is a shelter factor ranging from 0 to 1.0, and  $E_k$  is kinetic energy in joule

### 2.2.7 Harm model

The harm model describe the severity of UAV impact towards the population. The harm model usually measured in probability of fatality given UAV impact to a person. The impact harm model can be categorized based on its severity from minor, major, and fatal injury. As previously mentioned many literature paper use kinetic impact energy as a variable when it comes to determining the fatality imposed by the UAV. Of every injury mechanism, the most common studied harm mechanism is a blunt force trauma. According to Shelley, 2016, the most likely impact injuries are injuries to the head, particularly skull fracture.

For various impact locations, the a charateristic of fatality probability model have been developed as as a function of kinetic energy by Harwick et al., 2007. The model describes the harm response of an average male, averaged over varying impact orientations, as a binomial regression function. Binomial distribution is a function that map a given input in this case kinentic energy to a false or true statement in

this case a resulted fatality. Therefore, given a kinetic energy, does that energy can cause fatality to a person.

The binomial distribution is calculated using a sigmoid function, more commonly known as logistic curve, shown in figure 2.4. logistic curve probability of fatality described in Shelley, 2016 is presented in equation 2.12. Where threshold is the impact energy associated with a 50% probability of a fatality (measured in Joules), Energy imapct is the impact energy. A number of advancements over this standard equation have been presented by Dalamagkidis et al., 2008 and Shelley, 2016 to name a few.

$$f(x) = \frac{1}{1 + e^{-\frac{(x-\text{threshold})}{\text{energy impact}}}} \quad (2.12)$$

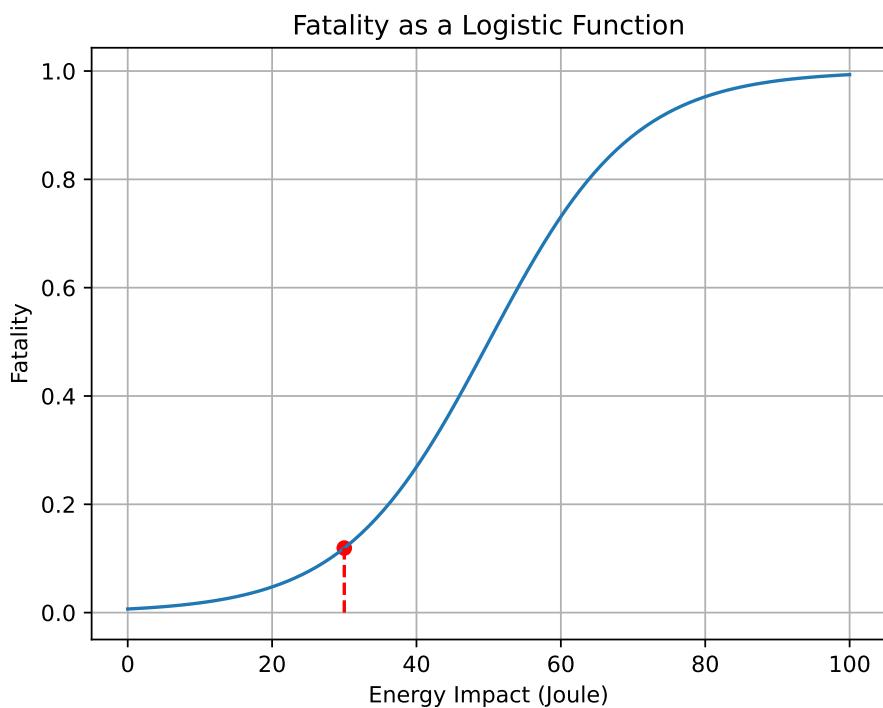


Figure 2.4: Generalized logistic fatality function given energy impact

Magister, 2010 makes use of a Blunt Criteria (BC) that relates the kinetic energy on impact with the body's ability to tolerate the energy on impact, which is expressed using the Abbreviated Injury Scale (AIS). The Blunt Criterion is a parameter used to assess the severity of blunt force trauma, taking into account the energy transferred during the impact and the body's tolerance to that energy. The Abbreviated Injury Scale is a globally accepted scoring system used to classify and describe the severity of injuries. The AIS assigns a numerical value to injuries ranging from 1 (minor) to 6 (fatal), providing a standardized method to quantify the impact of injuries.

Overall, the use of ground risk models has the potential to significantly improve the safety and effectiveness of UAV operations. By identifying and prioritizing potential risks, ground risk models can help operators make informed decisions about how to mitigate or manage those risks. As the use of UAV continues to grow and expand into new applications, the importance of ground risk modeling is likely to increase as well.

As an summary, think the ground risk model as a series of queestion asked in order. Where the questions are, how likely the will UAV fail, if the UAV fail where will it land, what does the descent landing looks like in ballistic, fly away, glide, or in parachute mode. If it land how likely is it to hit people in certain area. Are people in the somekind of shelter? if it does how safe is it if the drone hit them, does that hit caause fatality?

## 2.3 Pathfinding Algorithms

This study not only for generating the ground risk map for the risk assesment but also measures and finds the path taken by the aircraft to find the least amount of risk and distance traveled required. That is why pathfinding algorithms are essential for UAV navigation, particularly for determining safe and efficient routes. There are multiple pathfinders, or in general, a method that is being used in the UAV industry as a navigation algorithm. All of these algorithms have different criteria, operating principles, and goals to complete. Therefore we are trying to reviews several common pathfinding algorithms, highlighting their advantages and limitations, and explains why the A\* algorithm is selected for our study.

### 2.3.1 Bread-First Search

Bread-First Search (BFS) is a fundamental searching algorithm where it works by starting at the root node and exploring all neighboring nodes. Then, it moves to the nearest nodes and explores their unexplored neighbors, continuing this process until the goal is reached. For a more analogous example, you are at a crowded networking event, beginning at the entrance, and you need to locate a colleague. You initially survey all individuals in the immediate vicinity, then progress incrementally, systematically examining each subsequent group. This methodical approach ensures you comprehensively search each "level" before advancing further.

We can see how the algorithm supposed to "walk" on the simple graph tree on the figure 2.5. Where we start at the left-most node, and it spread between its current neighbor first before increment to the right.

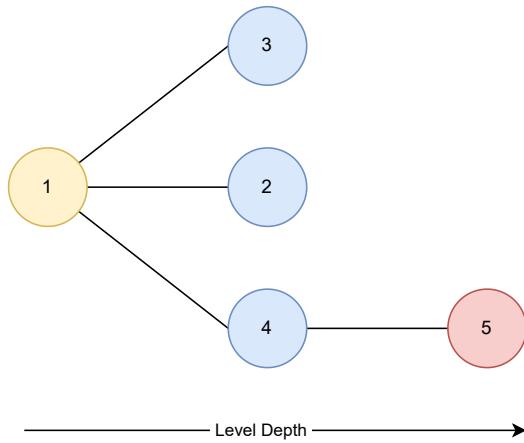


Figure 2.5: BFS graph tree

One of the main advantages of BFS is its guarantee to find the shortest path in terms of the number of edges in unweighted graphs. This algorithm is also straightforward and easy to implement. However, BFS can be very memory-intensive as it needs to store all nodes at the current depth level before proceeding to the next, which can be inefficient for large graphs or when memory resources are limited. Furthermore, BFS does not perform well on weighted graphs where the cost of traversing edges varies significantly Cormen et al., 2022.

### 2.3.2 Depth-first search

Depth-First Search (DFS) explores as far down a branch as possible before backtracking, making it useful for exhaustive searches of all possible paths. DFS starts at the root node and explores each branch to its fullest before moving to the next branch, we can see the algorithm works on the figure 2.6. This approach can be beneficial in situations where the complete path needs to be explored or specific deep paths are of interest.

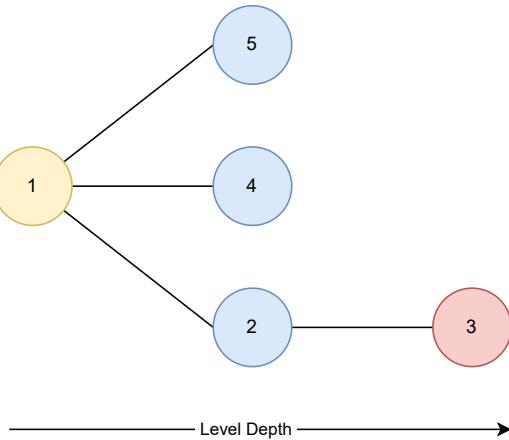


Figure 2.6: DFS graph tree

DFS is less memory-intensive than BFS since it only needs to store the nodes along a single path from the root. It is also simple to implement. However, DFS does not guarantee finding the shortest path. In cases of deep or potentially infinite graphs, DFS can get stuck in long paths, leading to inefficiency. This characteristic makes DFS less suitable for pathfinding where the shortest or optimal path is required Knuth, 1997.

### 2.3.3 Dijkstra's algorithm

Dijkstra's Algorithm is a widely used method for finding the shortest path in graphs with non-negative edge weights. It works by maintaining a set of nodes whose shortest distance from the source is known and repeatedly expanding the shortest known path by one edge. The algorithm uses a priority queue to efficiently select the node with the smallest tentative distance. This algorithm can be represented by equation 2.14 and can be expanded into equation 2.13

$$f(n) = \min\{g(n)\} \quad (2.13)$$

Equation 2.13 is a generalized form of the algorithm, where to find the shortest distance, we must minimize the cost function of  $g(n)$ . The cost function can be a weight between node connection or the distance between the node. Since we mostly interested in calculating optimum distance, we then expand the cost function to calculate euclidean distance  $w(u, v)$  as is cost function. Which resulted in equation 2.14 below

$$d(v) = \min_{u \in V}\{d(u) + w(u, v)\} \quad (2.14)$$

Where  $d(v)$  is the shortest known distance from the start node to node  $v$ .  $u$  is a node in the graph that is adjacent to  $v$ .  $V$  is the set of all nodes in the graph. And  $w(u, v)$  is the weight, or cost, of the edge connecting nodes  $u$  and  $v$ .

Dijkstra's Algorithm guarantees finding the shortest path in weighted graphs without negative weights, making it highly reliable for a variety of applications. It is efficient when implemented with a priority queue, especially for sparse graphs. However, the algorithm can be computationally intensive for very large graphs, which limits its scalability. For graphs with uniform weights, simpler algorithms like BFS may be more efficient Dijkstra, 2022.

### 2.3.4 Greedy best-first search

Greedy Best-First Search (GBFS) is an algorithm that uses a heuristic to prioritize nodes that appear to be closest to the goal. Unlike Dijkstra's Algorithm, which considers the cost to reach the current node and the estimated cost to reach the

goal, Greedy Best-First Search only uses the heuristic estimate, as is in equation 2.15.

$$f(n) = \min\{h(n)\} \quad (2.15)$$

Where instead of cost function  $g(n)$  we use only heuristic function of  $h(n)$ . Heuristic function can be any function that suited the use cases of the algorithm. But when we are discussing about spatial based path finder, we commonly use distances as a cost function.

Even though in a glance that it might seems that this algorithm heuristic function use the same function cost as in the Dijkstra's algorithm. Be aware that in Dijkstra that it use sum of all is traversed node. While in greedy algorithm it only measure the distance between the current node to its final node. Take a look at figure 2.7

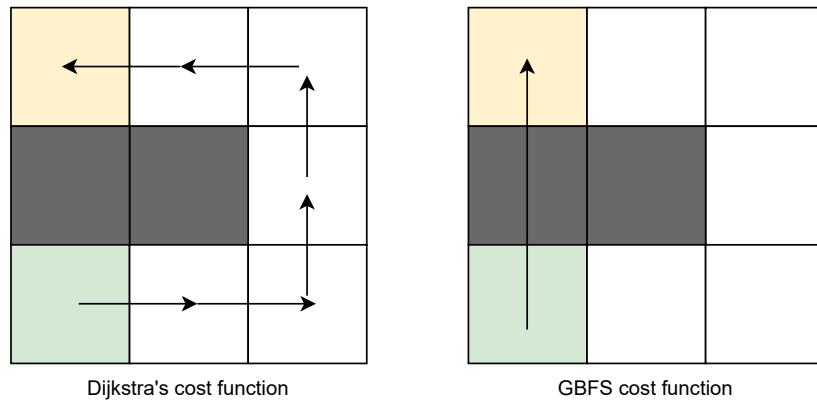


Figure 2.7: Differences of cost function between GBFS and Dijkstra's

As you can see, when there is blocked path between start node colored in green and the end node colored in orange. The cost function of the Dijkstra's algorithm compound its true per step distance, while the GBFS point straight from the start to the end.

GBFS required much lower computation resource and relatively easy to implement. However, this speed comes at the cost of not guaranteeing the shortest path, especially in graphs where the heuristic is not perfectly accurate. This algorithm can be useful when a quick, approximate solution is needed rather than an exact shortest path Russell and Norvig, 2016.

### 2.3.5 A\*

The A\* algorithm essentially combines the distance calculation of Dijkstra's algorithm with an additional heuristic function that estimates the remaining distance to the goal, shown in 2.16. This heuristic guides the search process more efficiently

toward the goal, often reducing the number of nodes that need to be explored. This become apparent when exploring very large ammount of grid cells where dijkstra need to explore every cell to determine least ammount of distance, where A\* just need to move into the direction of the heuristic vector.

$$f(n) = \min\{g(n) + h(n)\} \quad (2.16)$$

The A\* algorithm is a widely used pathfinding and graph traversal algorithm that is known for its efficiency and accuracy. By using a heuristic function, A\* often explores fewer nodes than Dijkstra's algorithm, especially in large graphs. A\* is guaranteed to find the shortest path if the heuristic function is admissible. An admissible heuristic is one that never overestimates the cost to reach the goal; in other words, it is always optimistic. This ensures that A\* will always find the optimal path. However, the performance of A\* heavily depends on the quality of the heuristic function. An inappropriate heuristic can degrade the efficiency Knuth, 1997.

An admissible heuristic is one that never overestimates the cost to reach the goal. The Euclidean distance is a common example of an admissible heuristic for pathfinding in a 2D plane because it represents the shortest possible straight-line distance between two points.

For two points  $A$  and  $B$  with coordinates  $(x_1, y_1)$  and  $(x_2, y_2)$  respectively, the Euclidean distance  $h(A, B)$  is given by:

$$h(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (2.17)$$

This heuristic is admissible because it never overestimates the true shortest path distance, which would be along the actual path in the grid or graph.

In contrast, a non-admissible heuristic overestimates the cost to reach the goal. For example, let's define a heuristic  $h'(A, B)$  as:

$$h'(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} + 5 \quad (2.18)$$

In this case,  $h'(A, B)$  is not admissible because it always overestimates the actual shortest path by an additional 5 cells. This overestimation can lead the A\* algorithm to make incorrect decisions about which nodes to explore, potentially causing it to miss the optimal path.

## Risk-based A\*

Instead of using none weighted graph in a normal A\*, we can insert weighted factor in each node. This weight factor is the risk itself, where now not only it tries to minimized the path length but it also tries to find the path with the least ammount of risk. This new function is The Risk-Based A\* algorithm which is an extension of the standard A\* algorithm. The cost function for the Risk-Based A\* algorithm is modified to include a risk component. The cost function  $f_{\text{risk}}(x)$  is given by on equation 2.19:

$$f_{\text{risk}}(x) = g(x) + h(x) + R(x) \quad (2.19)$$

where:

- $g(x)$  is the actual distance from the start node to the current node.
- $h(x)$  is the heuristic estimate of the distance from the current node to the goal.
- $R(x)$  is the risk associated with the path from the start node to the current node.

In this formula,  $f_{\text{risk}}(x)$  represents the total cost of a path, considering both distance and risk. By minimizing this cost function, the algorithm finds the path that balances the shortest distance with the least risk. Primatesta, Rizzo, and la Cour-Harbo, 2020 and la Cour-Harbo, 2019 use this method to find the overall least risk path with the least possible distance to take to.

### 2.3.6 Rapidly-exploring random tree

The Rapidly-exploring Random Tree (RRT) algorithm is a widely used method for solving motion planning problems, particularly in high-dimensional spaces and environments with obstacles. It is extensively utilized in robotics for pathfinding and motion planning. The algorithm works by incrementally building a space-filling tree that explores the feasible regions of the state space. RRT is especially advantageous in scenarios where the environment is dynamic and complex, making it challenging to find a path using traditional methods.

RRT operates in a state space, where each state represents a configuration of the system, such as the position and orientation of a robot. The algorithm constructs a tree  $T$  that starts at the initial state  $x_{\text{init}}$  and incrementally expands towards unexplored regions of the state space. Each node in the tree represents a state, and each edge represents a feasible transition between states. The tree is expanded by randomly sampling a state  $x_{\text{rand}}$  in the state space, finding the nearest node  $x_{\text{nearest}}$  in the tree to this random state, and then creating a new node  $x_{\text{new}}$  in the direction of  $x_{\text{rand}}$  from  $x_{\text{nearest}}$ .

The RRT algorithm's main steps can be summarized as follows: 1. Initialize the tree  $T$  with the initial state  $x_{\text{init}}$ . 2. Repeat until the goal state is reached or a maximum number of iterations is exceeded:

- Randomly sample a state  $x_{\text{rand}}$  from the state space.
- Find the nearest node  $x_{\text{nearest}}$  in the tree  $T$  to  $x_{\text{rand}}$ .
- Create a new node  $x_{\text{new}}$  by moving from  $x_{\text{nearest}}$  towards  $x_{\text{rand}}$ .
- Add the new node  $x_{\text{new}}$  to the tree  $T$ .

Mathematically, the nearest node  $x_{\text{nearest}}$  is found using the Euclidean distance, shown in equation 2.20:

$$x_{\text{nearest}} = \arg \min_{x \in T} \|x - x_{\text{rand}}\| \quad (2.20)$$

where  $\|\cdot\|$  denotes the Euclidean norm. The new node  $x_{\text{new}}$  is then generated by moving a fixed step size  $\delta$  from  $x_{\text{nearest}}$  towards  $x_{\text{rand}}$ , shown in equation 2.21:

$$x_{\text{new}} = x_{\text{nearest}} + \delta \cdot \frac{x_{\text{rand}} - x_{\text{nearest}}}{\|x_{\text{rand}} - x_{\text{nearest}}\|} \quad (2.21)$$

Here,  $\delta$  is the step size, and the direction vector  $\frac{x_{\text{rand}} - x_{\text{nearest}}}{\|x_{\text{rand}} - x_{\text{nearest}}\|}$  ensures the new node is on the line connecting  $x_{\text{nearest}}$  and  $x_{\text{rand}}$ . To give more clarity, take a look at figure 2.8 below, where the root node 1 try to reach the final node 4 by spawning node  $x_{\text{rand}}$  and extends its connection to the nearest node  $X_{\text{nearest}}$

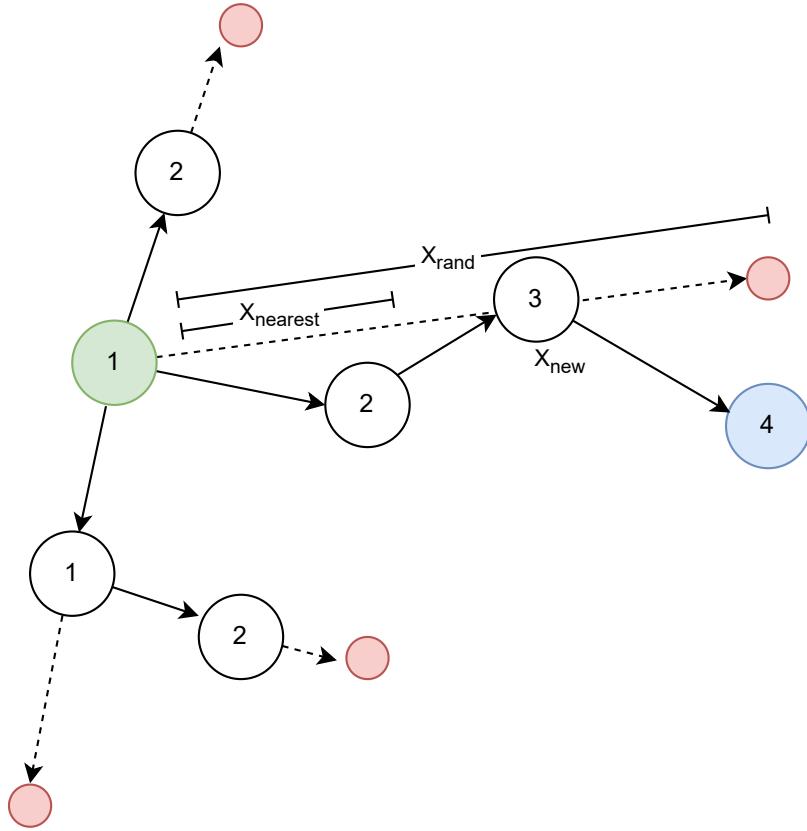


Figure 2.8: Simplified RRT Graph

In layman's terms, the RRT algorithm can be likened to a person exploring a forest. The person starts at a known location (initial state) and randomly picks a direction to move towards (randomly sampled state). They then find the closest point they have already explored (nearest node) and take a step in the direction of the new point (new node). This process is repeated, expanding the area explored until the destination is reached or the exploration is halted.

The RRT algorithm is advantageous because it efficiently explores large state spaces by focusing on unvisited regions. This exploration property makes it suitable for high-dimensional and complex environments. However, RRT may not always find the optimal path, and its performance can be sensitive to the choice of step size  $\delta$  and the distribution of random samples  $x_{rand}$ . Extensions and variations of RRT, such as RRT\* and Informed RRT\*, have been developed to address some of these limitations by improving path optimality and convergence rates. RRT\* works by pruning a longer path that take to the same node, take a look at figure 2.9 where 2 lower node 3 is not pruned in RRT but pruned in RRT\*

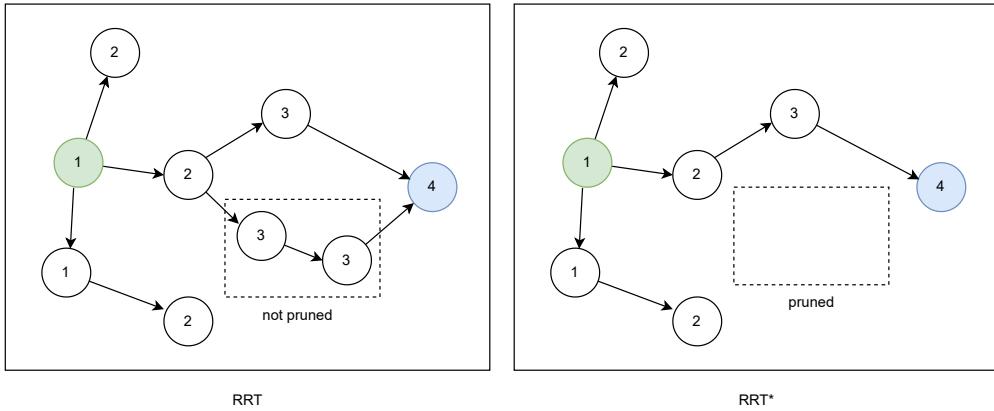


Figure 2.9: RRT vs RRT\*

In conclusion, the RRT algorithm is a powerful tool for motion planning in robotics, capable of handling complex and dynamic environments. Its ability to rapidly explore the state space makes it a popular choice for real-time pathfinding applications LaValle, 2006, however because since it excel to find solution fast, it is not guarantee to be optimal. But we are not going to dive deeper into these extension since RRT itself is for real time on the fly path finder, where our paper is for calculating pre-flight ground risk model.

### 2.3.7 Ant colony optimization

Ant Colony Optimization (ACO) algorithm is a probabilistic technique inspired by the foraging behavior of ants. It is used for finding optimal paths in graphs, and it is particularly effective for solving combinatorial optimization problems such as the travelling salesman problem Dorigo, 2007. ACO simulates the behavior of ants searching for food, where ants deposit pheromones on paths they take, and these pheromones guide subsequent ants towards promising paths.

The algorithm operates in the following steps: 1. Initialize pheromone levels on all edges of the graph. 2. For each iteration, simulate the movement of multiple ants:

- Each ant constructs a solution by moving from one node to another based on the probability influenced by pheromone levels and heuristic information.
  - Update pheromone levels based on the quality of the solutions constructed by the ants.
3. Repeat the iterations until a termination condition is met (e.g., a maximum number of iterations or convergence).

The probability  $p_{ij}(t)$  that an ant moves from node  $i$  to node  $j$  at time  $t$  is given by equation 2.22:

$$p_{ij}(t) = \frac{[\tau_{ij}(t)]^\alpha [\eta_{ij}]^\beta}{\sum_{k \in \mathcal{N}_i} [\tau_{ik}(t)]^\alpha [\eta_{ik}]^\beta} \quad (2.22)$$

where:

- $\tau_{ij}(t)$  is the pheromone level on edge  $(i, j)$  at time  $t$ .
- $\eta_{ij}$  is the heuristic information (e.g., the inverse of the distance between nodes  $i$  and  $j$ ).
- $\alpha$  and  $\beta$  are parameters that control the influence of pheromone and heuristic information, respectively.
- $\mathcal{N}_i$  is the set of neighboring nodes of node  $i$ .

The pheromone update rule is given by:

$$\tau_{ij}(t+1) = (1 - \rho)\tau_{ij}(t) + \Delta\tau_{ij}(t) \quad (2.23)$$

where:

- $\rho$  is the evaporation rate, representing the pheromone evaporation over time.
- $\Delta\tau_{ij}(t)$  is the amount of pheromone deposited by the ants, typically proportional to the quality of the solution.

The ACO algorithm can be likened to a group of ants exploring different paths to find the shortest route to food. As ants travel, they leave a trail (pheromone) that other ants can follow. Over time, the paths with stronger pheromone trails attract more ants, reinforcing good paths and eventually leading to the discovery of the optimal route. Take a look at figure 2.10 below where the first iteration all path has the same amount of pheromone, but after couple of iterations the shorter path has a much more pheromone since ant agent traverse it more often, where the least likely route has its pheromone evaporated.

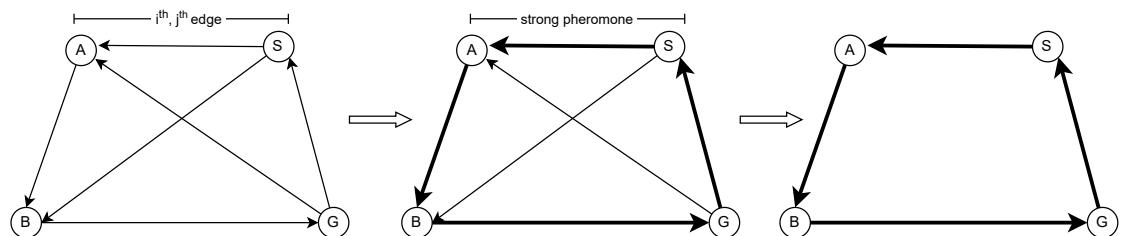


Figure 2.10: Graph with ACO to optimize shortest total route

The Ant Colony Optimization algorithm offers several advantages. It is highly effective for tackling complex optimization problems, making it a versatile tool for

a wide range of applications. The algorithm's ability to find near-optimal solutions within a reasonable timeframe adds to its appeal, especially for time-sensitive tasks. Furthermore, ACO is adaptable to dynamic changes in the environment, allowing it to maintain performance even when conditions shift. However, the algorithm is not without its drawbacks. It can require significant computational resources, particularly for large-scale problems, which may limit its practicality in resource-constrained settings. Additionally like any other complex algorithm with hyperparameters, the performance of ACO is sensitive to parameter settings, necessitating careful tuning to achieve optimal results. There is also a risk of converging to a suboptimal solution if the algorithm is not properly managed, highlighting the importance of adequate oversight and control Dorigo, 2007.

Although the ACO algorithm is a robust and flexible optimization technique that mimics the natural behavior of ants to solve complex problems. The problem it tries to solve is not suitable for finding ground risk model. Remember that ACO is try to spread the ant to explore every possible cells to determine the shortest distance to cover every cell.

### 2.3.8 Generative pathfinding AI

Generative Pathfinding AI algorithms leverage advanced machine learning techniques, particularly generative models, to discover optimal paths in complex environments. These algorithms utilize neural networks, such as Generative adversarial networks (GANs), to generate feasible and efficient paths. By learning from data, these models can generalize to new, unseen environments and provide robust pathfinding solutions Silver et al., 2016.

The algorithm operates in the following steps: 1. Train a generative model on a dataset of known paths. 2. Use the trained model to generate new paths by sampling from the learned distribution. 3. Evaluate the generated paths based on specific criteria (e.g., shortest distance, minimal risk).

Generative Pathfinding AI offers several significant advantages. The most notable is its capability to learn and generalize from vast amounts of data Russell et al., 2010, enabling it to adapt to complex and dynamic environments efficiently. This adaptability makes it suitable for scenarios where the environment is continually changing or unpredictable. Additionally, it can efficiently generate multiple path solutions, providing a variety of options to choose from, which is particularly useful in decision-making processes Russell et al., 2010. However, these advantages come with some challenges. Generative Pathfinding AI requires large amounts of training data to achieve high performance, which may not always be readily available. The training process itself is computationally intensive, demanding substantial computational resources and time. Moreover, the paths generated by the model need to be validated for feasibility, as not all generated paths will be practical or optimal. This validation step is crucial to ensure the reliability of the solutions provided by the AI.

In summary, Generative Pathfinding AI represents a cutting-edge approach to pathfinding, leveraging the power of machine learning to discover efficient paths in complex environments. Its ability to learn and adapt makes it a promising tool for a wide range of applications.

### 2.3.9 Justification for choosing risk-based A\*

Among the algorithms reviewed, A\* stands out due to its balanced approach combining the benefits of Dijkstra's Algorithm and Greedy Best-First Search. A\* is particularly advantageous for pre-operational UAV path planning as it ensures the optimal path in terms of both distance and risk. By incorporating a risk coefficient into the cost function, the Risk A\* variant can balance the need for short paths with the requirement to minimize exposure to hazardous areas. Studies by Primatesta, Rizzo, and la Cour-Harbo, 2020 and Clothier et al., 2007 have demonstrated the effectiveness of the Risk A\* algorithm in developing ground risk management models, further supporting its suitability for this application.

While generative pathfinding AI algorithms offer advanced capabilities by learning from vast amounts of data and adapting to complex environments, they come with significant drawbacks. The training process for these models is prohibitively expensive, requiring substantial computational resources and time. This makes them less practical for applications where resources are limited or quick deployment is necessary.

In conclusion, while various pathfinding algorithms offer different benefits, A\* provides the most balanced and optimal solution for UAV navigation in terms of minimizing both distance and risk, making it the preferred choice for our study.

# CHAPTER 3

## BACKGROUND THEORY AND METHODOLOGY

### 3.1 Case study

#### 3.1.1 UAV Specification

In this study, we are going to analyze the risk factor of multirotor and fixed-wing UAV. In comparison, Talon and Parrot Disco are fixed-wing. While DJI Phantom 4 and DJI Inspire 2 are multirotor. Four of them are chosen since they are generally known as a go-to standard in the commercial drone industry. Each category is also differentiated into two different sizes, where Talon and Inspire 2 are the large size, while Phantom 4 and Disco are much smaller UAV. All properties are obtained from manufacturer design specifications.

As for the aerodynamic properties of these drones, we are using a very simple drag model. This will not include an Induced drag since we cannot predict every possible angle of attack. It is also beneficial to use simplify drag model since it would give a higher impact velocity, which increases the impact energy and results in a much more conservative value.

It is worth mentioning the mechanic of the glide model in this risk analyzer. Since only fixed wings are capable of gliding, and to an extent, a variable-pitch helicopter by autorotation, it is almost impossible to make a glide descent on a fixed-pitch propeller multirotor. Where for the multirotor, ballistic descent is the more likely outcome, even though there is little to no damage in the aircraft. It is then assumed that a glide descent, and to an extent a glide speed in multirotor, as a low powered horizontal descent where the power is barely sustaining a hover maneuver.

Table 3.1: UAV Specifications

Specification	Talon	Parrot Disco	Phantom 4	Inspire 2
Type	Fixed wing	Fixed wing	Quadcopter	Quadcopter
Mass (kg)	3.75	0.75	1.4	4.25
Front area ( $m^2$ )	0.1	0.07	0.02	0.04
Radius (m)	0.88	0.575	0.2	0.4
Maximum flight time	1.25 h	45 min	20 min	25 min
Ballistic drag coefficient	$N(0.9, 0.2)$	$N(0.9, 0.2)$	$N(0.7, 0.2)$	$N(0.7, 0.2)$
Horizontal initial speed (m/s)	$N(18, 2.5)$	$N(15, 2.5)$	$U(0, 15)$	$U(0, 20)$
Vertical initial speed (m/s)	0	0	0	0
Glide speed (m/s)	16	12	7.5	10
Glide ratio	$N(12, 2)$	$N(2.7, 0.8)$	$N(2.7, 0.8)$	$N(2.7, 0.8)$
Parachute drag coefficient	$N(13, 0.2)$	$N(0.9, 0.2)$	$N(0.9, 0.2)$	$N(0.9, 0.2)$
Parachute area	1	0.5	0.5	0.5

$N(\mu, \sigma)$  represents a normal distribution.

$U(start, end)$  represents a uniform distribution.

## 3.2 Objective location

The ground risk would be performed in North Jakarta, Indonesia. The place is chosen for its dense population, an assortment of building heights, large water bodies, and multiple residential and commercial zone areas. Shown in 3.1

And for the operation itself, the UAV would be flying from Pluit Village to Jakarta Sailing Vocational School. This objective is chosen based on the criteria previously mentioned.

We would compare the performance of the Risk A\* Algorithm vs the Normal A\* algorithm. We do not include Djikstra's algorithm in this section since the A\* algorithm provided by the NetworkX library is already admissible and able to obtain optimal pathways.

It is worth mentioning the normal A\* is the same for all types of aircraft since aircraft properties are only important in Risk A\* in generating various descent PDFs.

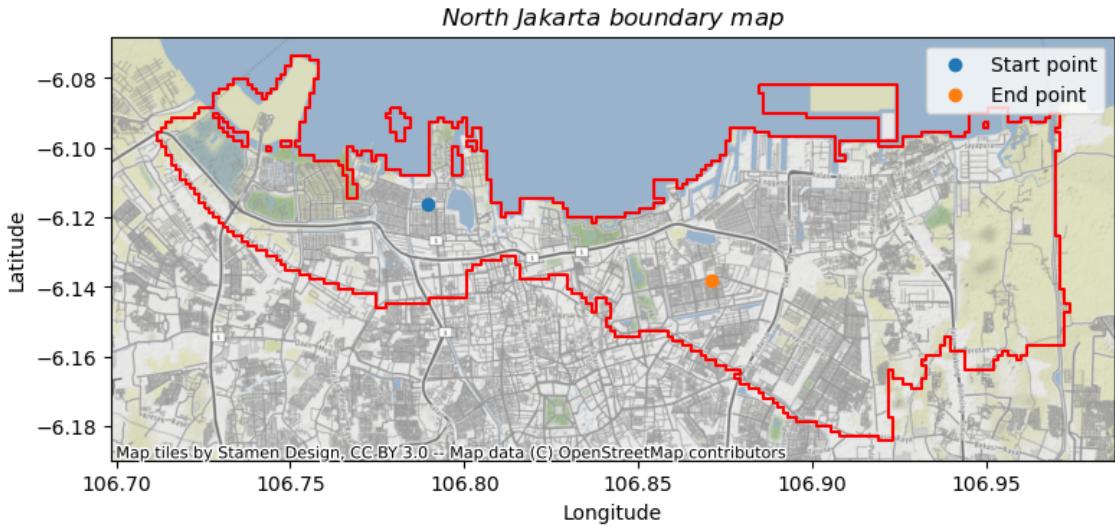


Figure 3.1: Operation objective

### 3.3 General Methodology

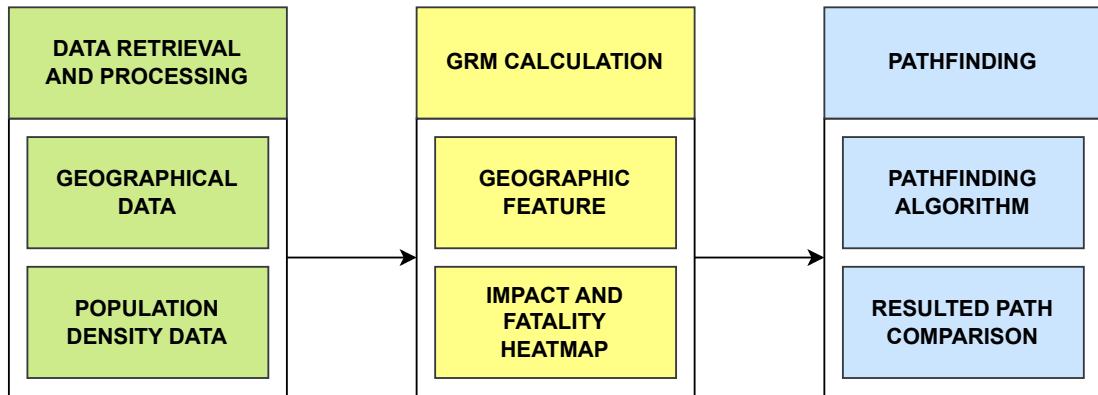


Figure 3.2: General overview

In general, there are three main objectives to create an optimal risk-aware flight path. These are data retrieval and processing, ground risk map calculation, and generating the most optimal path using the pathfinder algorithm.

To collect necessary data, we are using free services provided by OpenStreetMap (OSM), OpenStreetMap contributors, 2017, to retrieve geographic data, and Global Human Settlement Layer (GHSL) , Commission and Centre, 2023, to retrieve population density throughout the earth.

The Ground Risk Model (GRM) is calculated using geographic and population data, and also the type of aircraft that are being used in the calculation. The result would be descent location, impact energy, and fatality probability heat map.

After we compute and discretized the GRM. We find the optimal path with

the least amount of risk using pathfinding algorithms, where the result would be compared.

## 3.4 Data retrieval and processing

### 3.4.1 Online geographic data

OSM is a collaborative project that aims to create a free, editable map of the world. It was founded in 2004 and has since grown to become one of the largest and most popular sources of geospatial data.

OSM is built on the idea that anyone can contribute to the map by adding and updating data. Users can add data using Global Positioning System (GPS) devices, aerial imagery, or by manually tracing over satellite imagery. The data is then stored in a database and made available under an open license, allowing anyone to use and share the data.

There are multiple ways to use the OSM service. To retrieve the geographic data, we can send the query to Overpass Application Programming Interface (API) and receive the defaulted Geographic JavaScript Object Notation (geoJSON) data OpenStreet, 2022. OSM also provides geocoding and reverse geocoding using the Nominatim API.

Geocoding is the process of converting addresses or other human-readable location descriptions into geographic coordinates (latitude and longitude). This is useful for mapping addresses, analyzing spatial data, and integrating location-based services into applications. Reverse geocoding, on the other hand, is the process of converting geographic coordinates into a human-readable address. This is useful for identifying the address of a specific point on a map, providing location information in a user-friendly format, and enhancing the usability of geographic data.

### 3.4.2 Population data

GHSL Commission and Centre, 2023 is a dataset produced by the JRC (Joint Research Centre) of the European Commission that offers details on the distribution and location of human settlements worldwide.

The GHSL classifies and examines patterns of urbanization and human settlement using new spatial data mining technologies and remote sensing data, such as satellite images. The dataset is periodically updated to reflect changes in urbanization and population growth and contains data on built-up areas, population density, and other indices of human presence.

Overall, the GHSL is a valuable tool for comprehending the spatial dynamics and patterns of human settlement and can assist planners in making better-educated choices concerning urbanization and sustainable development.

### 3.4.3 Geographic Information System

Geographic and population data are stored and processed by Geographic Information System (GIS), which are tools used for storing, manipulating, and visualizing geographic data. GIS allows users to analyze and understand spatial relationships, patterns, and trends in geographic data.

GIS consists of several key components:

- Data: GIS uses data that is linked to a geographic location. This data can be in the form of points, lines, or polygons (areas) and can include attributes such as population, land use, or elevation.
- Hardware and software: GIS requires specialized hardware and software to store and analyze the data. This includes computers, servers, and specialized software such as ArcGIS or QGIS.
- Maps: GIS allows users to create interactive maps that can be used to visualize and analyze the data. Maps can be customized with different layers of information, such as roads, land use, and population density.
- Analysis: GIS allows users to perform spatial analysis on the data, such as calculating the distance between two locations or identifying patterns in the data.

There are multiple GIS software suites, ranging from standalone programs such as ArcGIS and QGIS to package-based libraries for Python, such as GeoPandas and Rasterio. In this paper, we are going to use GeoPandas and Rasterio along with their supporting libraries.

## 3.5 Ground risk map calculation

The Risk map is a graphic representation depicting a potential risk associated with the UAV operation. Risk maps are divided equally in distance into cells. Each cell represented a bounded geographical location. The map can be represented by a  $M \times N$  size matrix, which each  $M \times N$  is the number of cells in each row and column respectively. Each  $\mathbf{R}(i, j)$  cell has centroid coordinates which represented the geographic location of  $(x, y)$  in any coordinate reference system Coordinate Reference System (CRS).

If we use a slight notation abuse, which is  $\mathbf{R}(i, j)$ , each cell coordinate  $(i, j)$  is represented as a geographical location in which case shifted by  $x$  and  $y$  amount. Figure 3.3 illustrates the Risk map matrix. If we actually deployed ground risk map using real world map it should have look like in the figure 3.4, where we are going to use North Jakarta as our testing ground and Parrot Talon as our drone. This map that are in CRS of European Petroleum Survey Group (EPSG) 4236 are showing us the fatality if certain UAV failed and fall down.

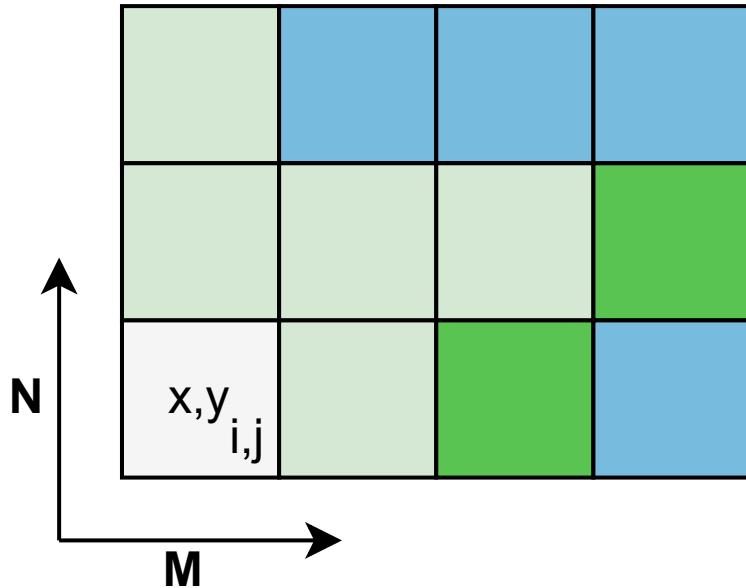


Figure 3.3: Grid risk matrix cell representation

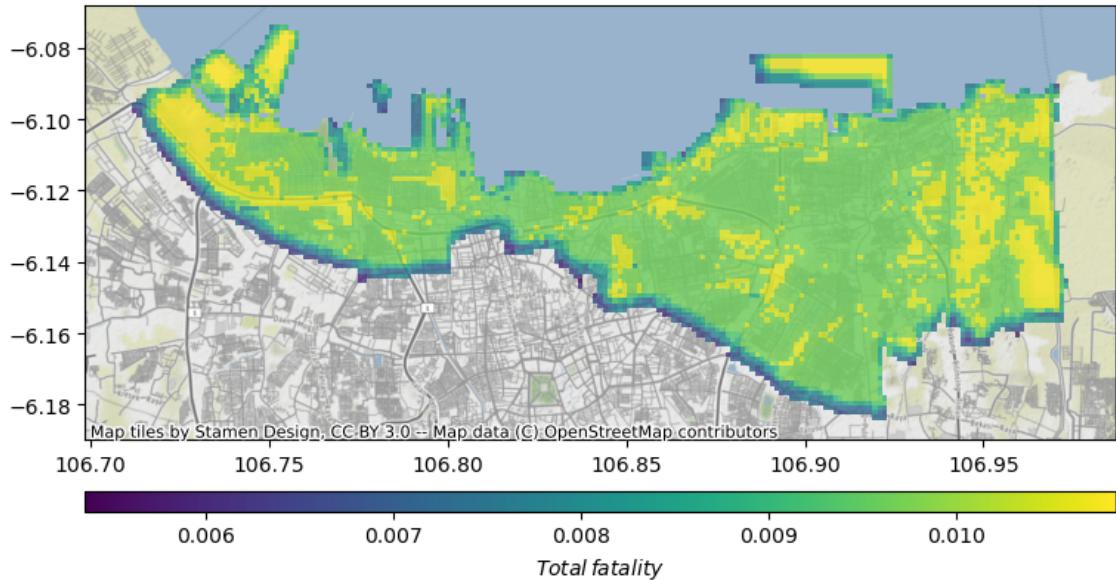


Figure 3.4: Talon ground risk map fatality distribution

### 3.5.1 Geographic feature

In a sense risk map is constructed by stacking layer of features, each layer itself is obtain by analyzing each geographical feature and population density to calculate the ground risk map. In this paper, multilayer frameworks are used based on the work of Primatesta, Rizzo, and la Cour-Harbo, 2020. As the name implied, this framework contains multiple georeference layers. These layers are :

- **Population density layer:** Describes population density in the area.
- **Obstacle layer:** Describes geographic features that can block aircraft flight paths such as building height, or any other obstacle.
- **Sheltering factor layer:** Describes geographic features that provide shelter for humans in the area.
- **No-Fly zone layer:** Describes geographic features in which aircraft cannot enter, such as military base, or airport

### 3.5.2 Population density layer

The population density layer describes the population density and distribution in the area. It is one of the most crucial factors in the risk assessment is population density since it indicates how many individuals can be affected by a vehicle crash and how they are dispersed over the map.

To conduct an accurate risk assessment it is necessary to provide population density data with fine resolution. since the likelihood of colliding with a person on the ground is particularly impacted by population density.

The population data are frequently gathered by the government human settlement institute. However, this approach is not easily ported from country to country since different institutes produce population data differently. As previously mentioned above GHSL would be used in this paper, since it already has the global database.

In this study, the population density layer is represented by a 2D georeference map with each cell corresponding to the value of population per cell area size with cell area in unit  $m^2$ , therefore  $people/m^2$ . Logically, the map would be in matrix form,  $\mathbf{D}$ , with each cell element of  $\mathbf{D}(x, y)$ . Figure 3.5 illustrates the population density of North Jakarta. The map uses a color gradient to depict varying levels of population density. Areas with higher population densities are shown in yellow lighter shades, while areas with lower population densities are represented in darker purple shades. This allows for quick identification of densely populated neighborhoods as well as less crowded areas. The data for this map is derived from Commission and Centre, 2023, ensuring accurate and up-to-date insights into the population distribution.

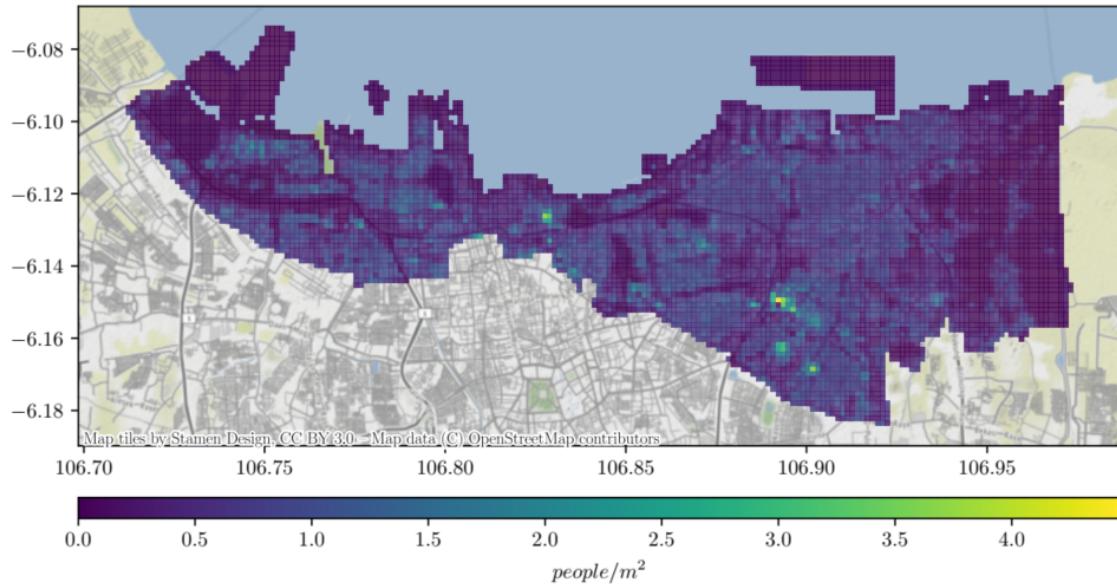


Figure 3.5: Population density map

### 3.5.3 Obstacle layer

Obstacle layer describe which cell has contains a building or geographic feature which higher than flight altitude of UAV. the data can easily be gathered by filtering, and mark it as impasable terrain. since it is just another georeference layer, we can represent the as a matrix  $\mathbf{O}(x, y)$  which describe maximum height of building in a given cells.

Figure 3.6 illustrates the obstacle map for North Jakarta, which highlights the height of buildings in the area in meter. This map is a crucial tool for urban planning and navigation, especially for UAV and other low-flying aircraft. The map uses a color-coded system to indicate the heights of different buildings, providing a clear visual representation of vertical obstacles. Noted that this layer does not explicitly accounted into risk factor analysis, but instead only for determining each cell is impasable for path finder.

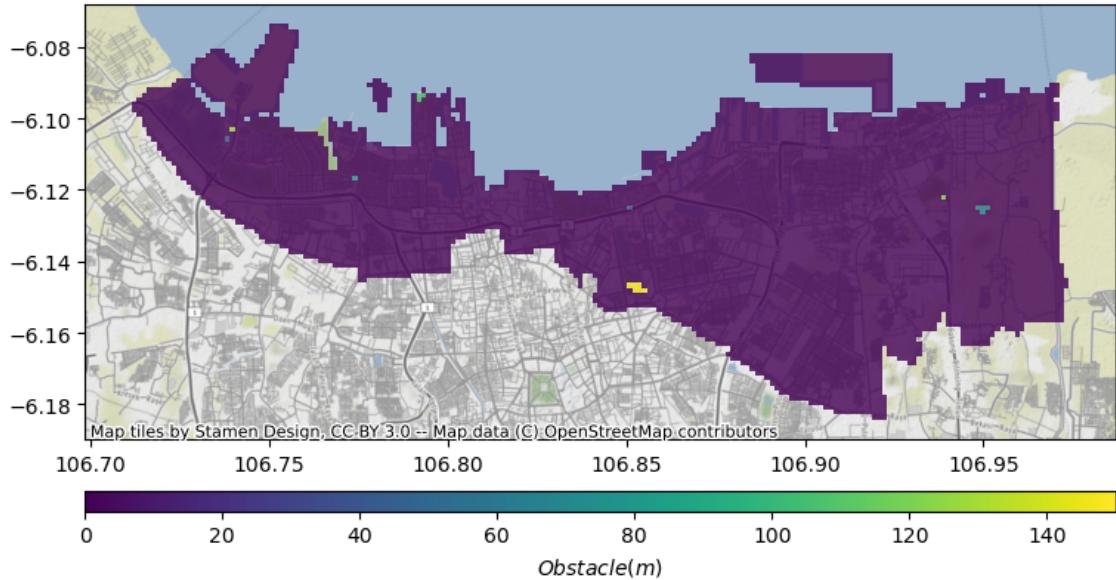


Figure 3.6: Height obstacle map

### 3.5.4 Sheltering Factor Layer

The Sheltering Factor layer focuses on assessing the presence and characteristics of physical structures or objects that can provide shelter. It is represented with a matrix  $\mathbf{S}$ , where each element  $\mathbf{S}(x, y)$  assumes the sheltering factor value of the corresponding location.

The sheltering factor layer takes into account factors such as the size, shape, material, and location of the structures or objects in the drone's operating environment. It aims to assess how these factors can influence the risk of fatality in case of failure.

Unfortunately, it is impractical to evaluate the sheltering element in full detail in a study of this kind because it is a very complex scenario. The dynamics of drone collision are a complication in itself, particularly since detailed physical building features, like material or wall thickness, are recorded differently around the world.

This issue is compounded by the fact that multiple studies have conflicting results when calculating the absorption of shelter since there is no consensus on defining the sheltering factor. Some studies by Dalamagkidis et al., 2008 and Primatesta, Rizzo, and la Cour-Harbo, 2020 use a scale factor from 0 to 1 or 1 to 10, respectively, representing the energy absorption factor of a given building. Meanwhile, Melnyk et al., 2014 utilizes a study conducted by the United States Department of Defense (DoD) Harwick et al., 2007 regarding explosive studies to measure the energy absorption capability.

Even though it is tempting to use the method given by Melnyk et al., 2014 and the DoD since it has the most accurate data, the study conducted by them is more suitable for singular operation-specific evaluation (i.e., best used for generating event

trees) and not suited for generating risk maps. There are also multiple problems when using this method, such as the lack of accurate data on building material geographic maps, especially when OSM does not provide building material data, which would result in inaccurate measurements.

Given these considerations, in this paper, we will use the scale factor proposed by Klepeis et al., 2001 and Primatesta, Rizzo, and la Cour-Harbo, 2020 as a foundation for sheltering factor risk generation since it provides a general method to calculate a risk factor. Table 3.2 shows the sheltering factor for different types of features.

Figure 3.7 shows the sheltering factor map, which visually represents the sheltering factors across different geographic locations. The map uses color coding to indicate various levels of sheltering, providing a clear understanding of areas that offer different degrees of protection.

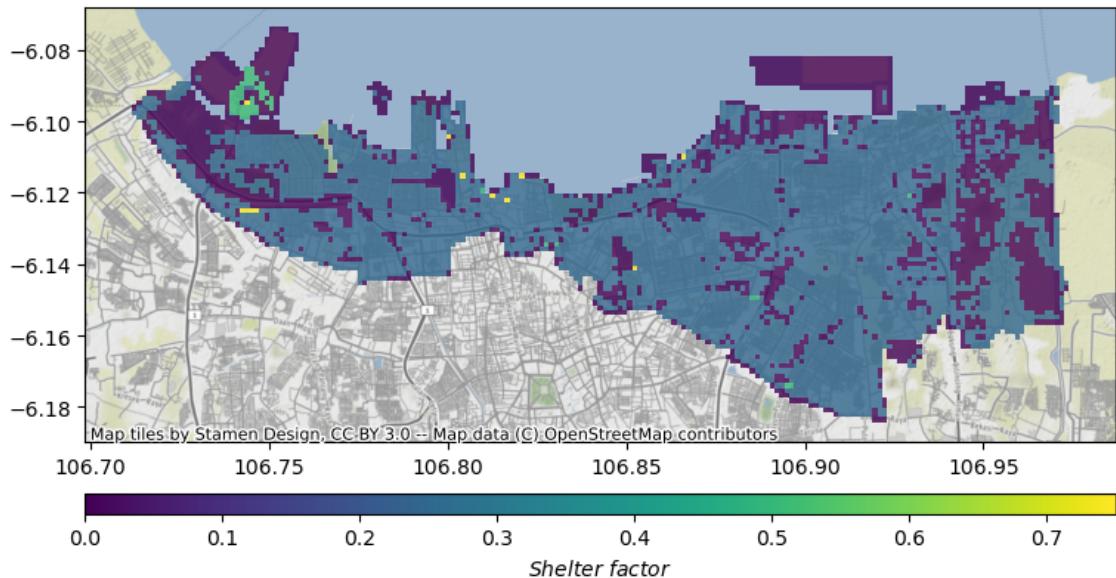


Figure 3.7: Shelter factor map with varying levels of sheltering provided by physical structures and objects across different locations.

Table 3.2: Sheltering factor values for different types of features.

Sheltering Factor	Type
0	No shelters
0.25	Sparse trees
0.5	Low building
0.75	High building
1.0	Industrial building

### 3.5.5 No-fly zone layer

The No-fly zone layer refers to a component of a ground risk map or airspace map that delineates areas where drone flights are prohibited. The restrictions in these zones can include areas such as:

- Regulation agency-defined areas, i.e. Federal Aviation Administration (FAA), and International Civil Aviation Organization (ICAO). They restrict airports and military areas.
- Nature-sensitive areas, such as National Park where flying is restricted.
- Safety and security areas, very crowded public areas, such as the city center, or events.
- Operator-defined areas, where areas are restricted by the operator itself for various reason.

In this paper, the no-fly zone layer is a georeferenced map similar to the obstacle layer. And generally work in the same method to obstacle layers. The No-fly zone layers explicitly prohibit the flight of UAV by using two different values, -1 if flying is forbidden, and 0 if flying is allowed. The no-fly zone layer is represented by matrix  $\mathbf{F}(x, y)$  defined as:

$$\mathbf{F}(x, y) = \begin{cases} -1 & \text{if flight is prohibited.} \\ 0 & \text{if flight is allowed} \end{cases} \quad (3.1)$$

## 3.6 Ground Risk Assessment

In this research, the risk is described as the potential to cause harm of UAV to the population on the ground. quantified as the number of fatalities per hour of flight. The risk model is built upon a series of consecutively occurring events, each with its probability. which can be summarized into Equation 3.2.

$$P_{\text{fatal}}(x) = P_{\text{failure}}(x) \cdot P_{\text{impact}}(x) \cdot P_{\text{recovery}} \cdot P_{\text{stress}} \cdot P_{\text{exposure}}(x) \cdot P_{\text{incident}}(x) \cdot P_{\text{harm}}(x) \quad (3.2)$$

- $P_{\text{failure}}(x)$ : Represents the probability of the UAV losing control and descending to the ground. This probability is calculated using the MTBF, which measures the average time between failures of the UAV.
- $P_{\text{impact}}(x)$ : Describes the likelihood of the UAV impacting a specific location on the ground. This probability is derived from a two-dimensional Probability Density Function that considers various descent scenarios, such as ballistic, unpowered glide, flyaway, and parachute descent. These events are calculated independently, and their probabilities are summed for each cell in the grid to determine the overall impact probability for that cell.

- $P_{recovery}$ : Accounts for the effectiveness of any recovery mechanisms, such as parachutes, that can reduce the severity of the UAV's impact. The probability of successful recovery is factored into the overall risk model to determine the likelihood of a controlled descent versus an uncontrolled crash.
- $P_{stress}$ : Refers to the probability of harm caused by the kinetic energy of the UAV at the time of impact. This probability is calculated using models that relate kinetic energy to injury severity, such as blunt force trauma models.
- $P_{exposure}(x)$ : Represents the probability of individuals being exposed to the UAV impact, based on population density. This probability is determined using population distribution data and the impact exposure area, which indicates the likelihood of the UAV striking individuals in the area.
- $P_{incident}(x)$ : Involves the sheltering factor, which assesses the presence and characteristics of physical structures or objects that can provide shelter. This probability is represented by the sheltering factor matrix,  $\mathbf{S}(x, y)$ , which assigns sheltering factor values to different locations based on the type and characteristics of structures present.

This paper considers four types of events: Ballistic, Unpowered Glide, Flyaway, and Parachute. Mathematical models are utilized to determine the probable impact area for each type of descent. It is important to note that these events are calculated independently of each other. Once the probabilities for each event are computed, they are summed up for each cell, resulting in the probability of the drone impacting that particular cell, shown in the figure 3.8. This probability is described using a 2D Probability Density Function (PDF). For further information on the descent models, please refer to Sections 3.6.2. The two-dimensional PDF is utilized to calculate the probabilities of  $P_{impact}$  and  $P_{fatality}$ , considering factors such as population density, sheltering factor layer, and impact velocities.

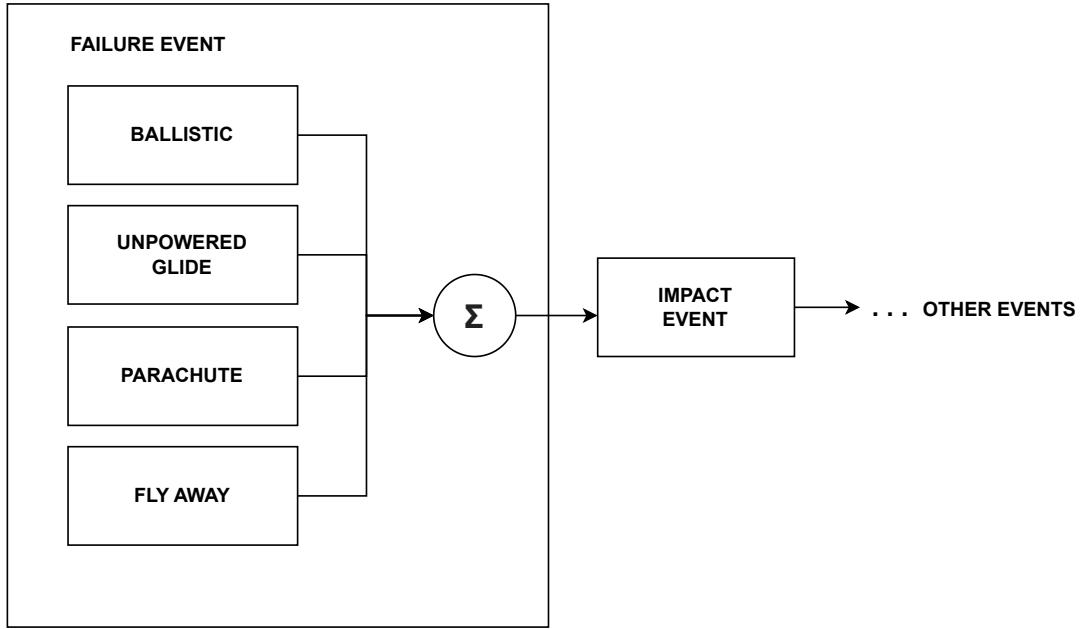


Figure 3.8: Type of descent event probability being summed

After calculating the UAV crash location following an uncontrolled descent, we need to calculate the  $P_{impact}$  which indicates the likelihood of the UAV strikes individuals. Where the population density and the impact exposure area can have its influence. And finally  $P_{fatality}$ , which denotes the probability of fatal injuries occurring after a person has been impacted. It depends on the kinetic energy at the time of impact and the sheltering factor.

The overall probability of fatality,  $P_{fatal}(x)$ , is then computed. The  $\mathbf{R}$  is a matrix containing calculated of said  $P_{fatal}(x)$  in every element of  $\mathbf{R}(x, y)$ . This approach allows for the assessment of risk associated with UAS flights by considering various factors and probabilistic models for different events during the flight.

### 3.6.1 Failure rate

The failure rate is a metric used to measure the reliability or average failure rate of a system or component. In this study, the MTBF (Mean Time Between Failures) represents the failure rate. MTBF is defined as the average amount of time that elapses between consecutive failures of a system or component.

According to Shelley, 2016, failure rate can be calculated using equation shown in equation 3.3 where  $\rho$  is the population density,  $b$  is the wingspan, and  $k$  is a conversion factor. The value of  $k$  is  $1/10$  when the population density data is given by people per square meter or square kilometer, and  $1/2.78$  when using square feet or square miles. Since MTBF is an inverse of  $P_{failure}$  we can rearranged into equation 3.4. If we calculated both MTBF and the  $P_{failure}$  of various UAVs in metrics using North Jakarta population density of 14,464 people per square kilometer. We will

get the result shown in table 3.3

$$MTBF = \frac{\rho\pi b^2}{10} = k(\rho\pi b^2) \quad (3.3)$$

$$P_f = \frac{\rho}{k} \times \pi b^2 \quad (3.4)$$

Table 3.3: MTBF and  $P_{fatal}$  for Various UAVs

UAV Model	Wingspan (meters)	MTBF (hours)	$P_{fatal}$
Talon	0.88	35,965.76	0.000278
Parrot Disco	0.575	13,483.80	0.000104
Phantom 4	0.2	1,818.36	0.000014
Inspire 2	0.4	7,273.44	0.000055

### 3.6.2 Descent Event

Information on how the UAV descended into the ground is required for ground risk management. The impact area and the kinetic energy at impact are determined according to the type of descent, which affects how the UAS falls to the ground.

The shape of a UAV has a great influence on how it falls, with fixed wings, generally glides down. While the multirotor fell following the ballistic trajectory. Both of which will influence the falling probability heat map.

These models below are employed to calculate a two-dimensional PDF that quantifies the likelihood of the UAS impacting the ground. The PDF takes into account uncertainties in the initial velocities (horizontal and vertical), drag coefficients, flight altitude, and flight direction.

Furthermore, the model generates a two-dimensional matrix containing the estimated horizontal and vertical impact velocities. These velocities are instrumental in computing the probabilities of impact ( $P_{impact}$ ) and fatality ( $P_{fatality}$ ).

#### Drone type

Drones come in various types, each with distinct descent characteristics, especially in terms of glide and range. Fixed-wing drones, like the Talon, are designed to mimic the flight of traditional airplanes, with wings that provide lift and enable longer glide distances. These drones are typically more efficient for covering large areas and can maintain flight with minimal power consumption. On the other hand, multirotor

drones, such as the DJI Phantom4, utilize multiple rotors to achieve vertical takeoff and landing. These drones are highly maneuverable and can hover in place, but they generally have a shorter range and less efficient glide capabilities compared to fixed-wing counterparts.

For our testing, we have selected the Talon as a representative fixed-wing drone and the DJI Phantom as a representative multirotor drone. These models are chosen due to their mid-range pricing, making them accessible for a wide range of applications while still offering reliable performance. The Talon's fixed-wing design allows us to assess its glide characteristics over long distances, while the DJI Phantom4's multirotor setup provides insights into its descent behavior. By comparing these two types of drones, we aim to better understand their respective strengths and weaknesses in various scenarios, ultimately guiding users in selecting the appropriate drone type for their specific needs.

## Ballistic descent

The occurrence of ballistic descent arises when the aircraft experiences a catastrophic failure, such as a structural failure that results in the loss of most of its lift. Consequently, the vehicle undergoes a descent solely driven by the forces of gravity and drag.

$$m\dot{v} = mg - c \cdot A|v|v \quad (3.5)$$

The ballistic descent model used in this work is based on the standard second-order drag model. The model takes into account factors such as the mass of the UAV ( $m$ ), a drag constant ( $c$ ), area( $A$ ), air density( $\rho$ ), gravitational acceleration ( $g$ ), and the vehicle's velocity vector ( $v$ ). A study done by Cour-Harbo, 2020 provides greater details of the model that also considers uncertainties and variations in wind conditions.

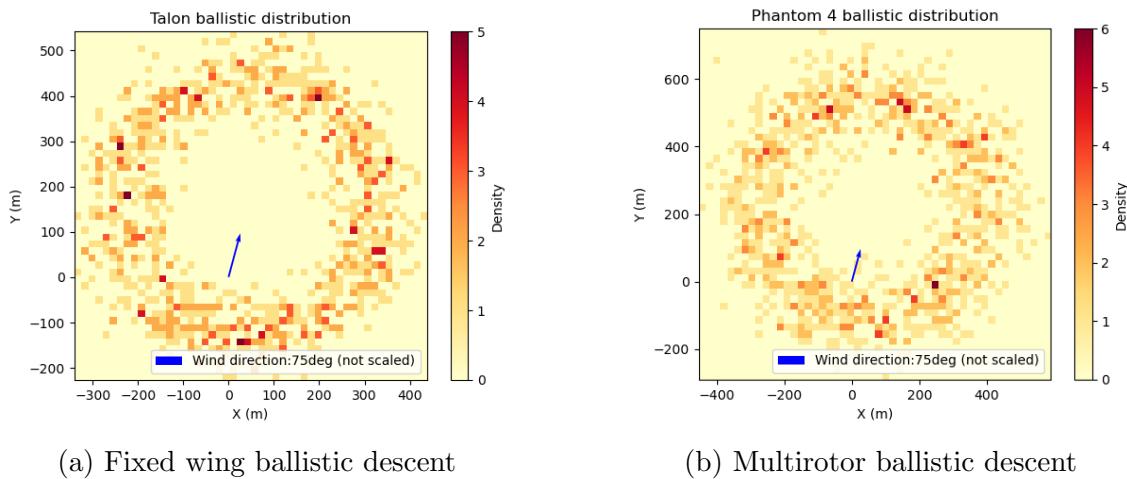


Figure 3.9: Ballistic descent by aircraft type

## Glide descent

Uncontrolled glide may arise depending on the configuration of the aircraft. It is an event when UAV flight profiles are governed solely by the glide ratio or autorotation angle. This event could happen when aircraft loses its power but still retains most of its lift capability.

In the case of a fixed-wing aircraft, this event arises from a loss of thrust or power affecting the flight control surfaces. For a single rotorcraft, it occurs when there is a loss of thrust on the main rotor, causing the aircraft to descend using autopiloted autorotation. However, multi-rotor cannot enter unpowered glide, instead, it enters ballistic descent.

During the uncontrolled glide event, the horizontal distance traveled is computed using the simple formula  $dist(h) = \gamma h$ , where  $h$  represents the flight altitude and  $\gamma$  represents the glide ratio

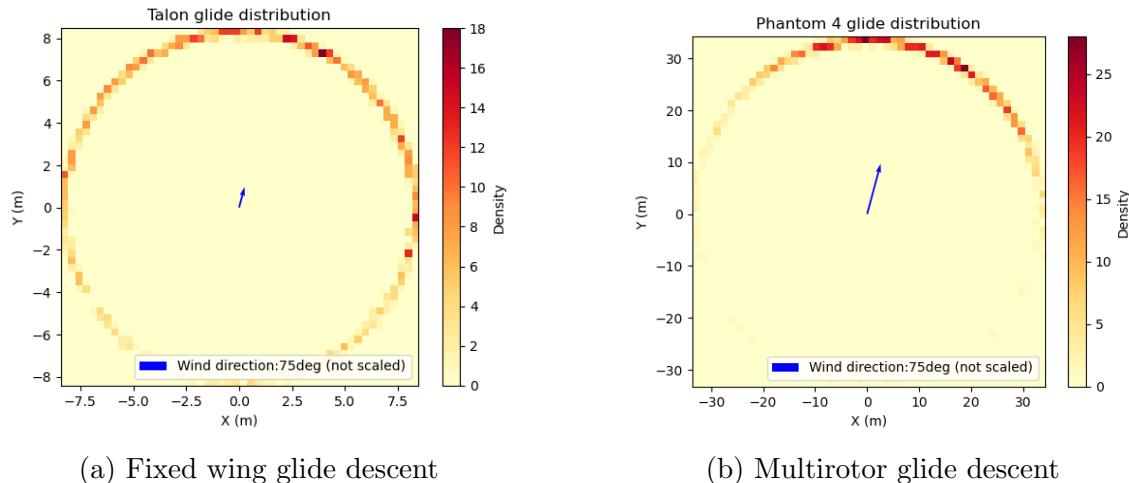


Figure 3.10: Glide descent by aircraft type

## Fly away

The fly-away event occurs when the operator loses complete control over the UAV, while the onboard flight control maintains the stability of the vehicle. As a result, the UAS can fly in any direction until it crashes on the ground. This event is modeled based on a reference paper la Cour-Harbo, 2019, which proposes a model comprising two components. Firstly, the probability of ground impact decreases linearly with the distance from the UAS position until reaching the maximum distance the vehicle can travel. Secondly, the vertical motion of the vehicle is taken into account, with a higher probability of ground impact near the UAS position, modeled as a normal distribution centered around the UAS position. These two contributions are combined linearly using a scaling factor.

In this study, equal probabilities are assumed for both scenarios to be conservative. It is assumed that the fly-away event terminates with an uncontrolled glide descent. Unlike other descent events, the fly-away event can involve motion in all directions. Figure 16 presents the results of the fly-away event.

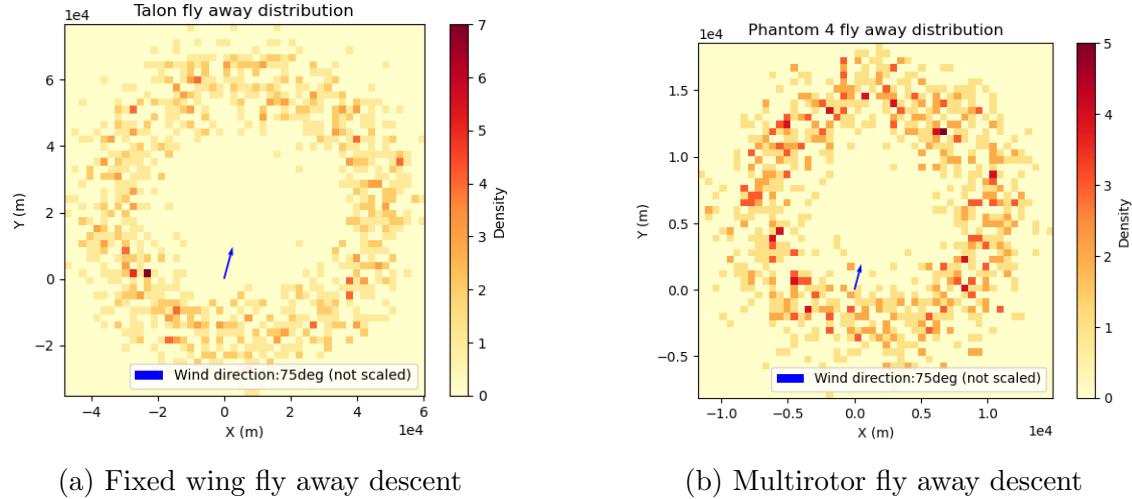


Figure 3.11: Fly away descent by aircraft type

## Parachute

The parachute descent starts when the UAV flies downward while the parachute is fully extended. The power plants are typically shut off and the parachute is opened when a problem occurs. After the emergency detection and the parachute fully opening, there is, however, a small delay. The only factors affecting the drop at this point are the parachute's aerodynamic features, which are intended to slow the vertical speed.

To model this event, a methodology described in a referenced paper la Cour-Harbo, 2019 is employed. The authors of the paper consider factors such as the mass of the vehicle, as well as the physical properties of the parachute, including its area and drag coefficient. Notably, the presence of wind significantly alters the descent dynamics, affecting the direction and velocity of the descent.

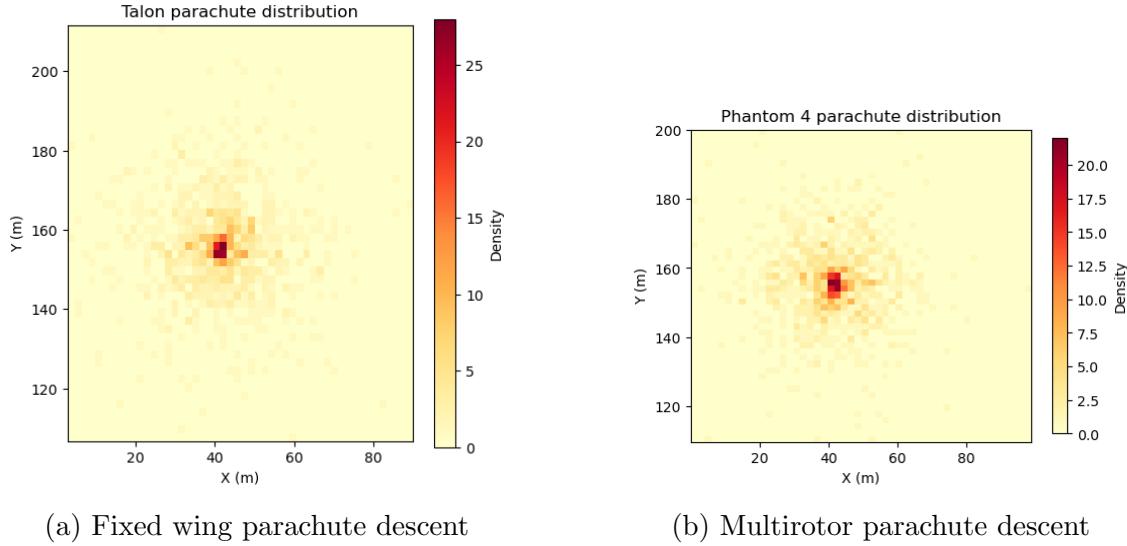


Figure 3.12: Parachute descent by aircraft type

### 3.6.3 Wind effect

The wind is one of the important variables in creating a ground risk map since wind speed and direction is determined the impact location of each descent event. Where the wind data can be received from weather forecast information before the operation.

Noted that this study assumes and applies constant wind speed and direction when simulating the descent event. This assumption arose when the wind effect usually encompasses a large local area which is often greater than medium to micro-sized UAV. Added to that, the wind only affected only the descent event and not when the drone is normally operating.

### 3.6.4 Probability of Impacting an Individual

The  $P_{impact}$  is the probability of a UAV striking individual. This parameter is important because, when used in conjunction with the population density parameter previously discussed, it determines the number of people exposed to risk on the ground in the event of an impact.

$$P_{impact}(x, y) = \rho(x, y) \cdot A_{exp} \quad (3.6)$$

The  $P_{impact}$  equation describe the population density  $\rho$  and exposure area  $A_{exp}$

Similar to sheltering factor, there is no standard approach when calculating the affected area. However, there are two distinct approaches that most researchers find useful, the empirical method and the geometric method.

Empirical methods are the method that uses regression to calculate impact area

based on the size or the weight of an aircraft. While the geometric method uses the general dimension of the aircraft to calculate the impact area.

Research done by Melnyk et al., 2014 uses the regression method from Ale and Piers, 2000, based on the accident data of large aircraft to calculate the impacted area. However, this regression is created using aircraft weight of more than 1500kg. This can result in underestimating the lethality of common UAV operations which weigh less than 500kg.

A study done by Aalmoes et al., 2015 proposes two types of solutions. empirical regression and geometric approach. Which stated that, above 500kg the calculation should use the empirical method, or geometric approach if less than 500kg. The empirical data obtained by the Netherland Aerospace Laboratory, when researching the helicopter crashes behavior, resulted in 3.8 relation. While the geometric approach 3.7 is using human impact model.

$$A_{exp}(m) = 230 \ln \left( \frac{m}{1000} \right) + 330 \quad (3.7)$$

$$A_{exp}(\theta) = 2(r_p + r_{uav}) \frac{h_p}{\tan(\theta)} + \pi(r + r_p)^2 \quad (3.8)$$

Where  $A_{exp}$  is in  $m^2$ , with m is kg,  $\theta$  is impact UAV impact angle. And  $r_{uav}$  is UAV radius ,  $h_p$  is person height, and  $r_p$  is person radius all in m,

### 3.6.5 Probability of Fatality

The probability of fatality, denoted as  $P_{fatality}$ , represents the likelihood that an impact on a person will result in a fatality. Calculating this probability is challenging due to the various complexities involved in an impact event. The human body reacts differently to different types of impacts, and UAV crashes can involve factors such as blunt impact, lacerations, or chemical explosions caused by the battery. As a result, variables like velocity and mass do not have a simple correlation to the severity of injuries, as the human body's response varies depending on the specific body part affected. This lack of consensus on the calculation method for fatality probability arises from the difficulty of accounting for all the intricacies and variations in impact events. However, it is generally accepted to correlate kinetic impact energy to measure the  $P_{fatality}$ .

$$P_{fatality} = \frac{1}{1 + e^{-k(E_{imp} - E_{thres})}} \quad (3.9)$$

3.10 is a generalized equation based on the logistic function, where the function maps a value of impact energy  $E_{imp}$  to the probability of fatality, with  $E_{thres}$  is an energy value that has a probability to cause fatality above 50%. On average, 58 joules of energy can cause a fatality based on the table below.

Table 3.4: Energy correlate to fatality

Information	Value	Unit	Source
Energy for hazardous debris	45	Joule	Cole et al., 1997
Energy required by fragment for 90% probability of fatality	120	Joule	Cole et al., 1997
Energy required by fragment to be hazardous to human	58	Joule	Gonzales and Murray, 2017

This study would be using more comprehensive  $P_{fatality}$  Equation 3.10, with  $\alpha$  as threshold energy to cause 50% probability of fatality in a medium level of shelter  $S \geq 0.6$ , and  $\beta$  is the threshold energy needed to cause fatality when in a non-sheltered area ( $S \leq 0.1$ ).

$$P_{fatality}(x, y) = \frac{1}{1 + \sqrt{\frac{\alpha}{\beta} \left[ \frac{\beta}{E_{imp}(x, y)} \right]^{\frac{1}{4S(x, y)}}}} \quad (3.10)$$

The kinetic energy is computed as:

$$E_{imp}(x, y) = \frac{1}{2} m \cdot v_{imp}(x, y)^2 \quad (3.11)$$

Equation 3.10 is based on Equation 3.10 from a study conducted by Dalamagkidis et al., 2008 and Primatesta, Rizzo, and la Cour-Harbo, 2020. Although it is not accounted for chemical and thermal factors as in Melnyk et al., 2014 and Harwick et al., 2007, we argue that since both Melnyk et al., 2014 and Harwick et al., 2007 only suitable for evaluating per single event and not for generating ground risk map.

## 3.7 Finding optimal path

After creating a risk map, the next objective is to find the optimal flight path with the least amount of risk. From a naive method point of view, computing flight path is quite easy since 99% of the operation is only a straight path from point A to point B since there is little to no obstacle. Instead, the risk factor has a bigger impact on the Pathfinder compared to real physical obstacles.

### 3.7.1 Problem statement

**Given:**

- A grid with dimensions  $M \times N$ , where each cell represents a location in the grid.

- A start cell  $S$  and a goal cell  $G$  in the grid.
- Each cell in the grid has an associated risk factor (a non-negative real value) that represents the level of risk or danger associated with that cell.

**Objective:** Find the shortest path from the start cell  $S$  to the goal cell  $G$  that minimizes the cumulative risk factor along the path.

**Formally, let::**

- Let  $G = (V, E)$  be the graph representation of the grid, where  $V$  represents the set of cells and  $E$  represents the set of edges connecting adjacent cells in the grid.
- Let  $w(u, v)$  be the weight or cost associated with an edge  $(u, v)$  in  $E$ . In this case, the weight  $w(u, v)$  represents the risk factor of moving from cell  $u$  to cell  $v$ .
- Let  $d(u)$  represent the shortest path distance from the start cell  $S$  to cell  $u$ .

The problem is to find the shortest path  $P = (S, v_1, v_2, \dots, G)$  that minimizes the cumulative risk factor along the path, which can be represented by the sum of the weights along the path:

$$\min \sum w(v_i, v_{i+1}) \quad (3.12)$$

subject to:

- $P$  is a valid non negative weighted path from  $S$  to  $G$  in the graph  $G$ .
- $d(G)$  is minimized.

The objective is to find the path  $P$  that minimizes the cumulative risk factor, taking into account the risk factors associated with each cell in the grid. We are going to use Risk A\* as the pathfinder algorithm.

Noted that the specific formulation and constraints of the problem may vary depending on additional considerations such as movement constraints, specific risk factor calculations, and other domain-specific requirements. The above problem statement provides a general framework for finding the shortest path based on risk factor in a grid system.

### 3.7.2 Graph

Before using any pathfinder algorithm, the grid map must be defined in the graph. In computer science and mathematics, a graph is a data structure that represents a collection of interconnected nodes, called vertices, ( $V$ ), and the relationships between them, called edges, ( $E$ ).

A graph consists of two main components: Vertices (Nodes): These are the fundamental units of a graph and represent entities or objects. Each vertex is typically labeled with a unique identifier or value. Edges: These represent the connections or relationships between vertices. An edge connects two vertices and can be either directed or undirected.

A directed edge has an orientation or a specific direction, while an undirected edge does not. Edges can also have weights or costs associated with them to represent the strength of the relationship or distance between vertices.

A grid typically falls into the category of a regular graph or a lattice graph. A regular graph is a graph in which every vertex has the same number of neighbors. In a grid, each cell is typically connected to its adjacent cells, resulting in a regular graph structure. The number of neighbors depends on the dimensionality of the grid.

The grid system, with its regular arrangement of cells and connectivity between adjacent cells, can also be considered a form of lattice graph. A lattice graph is a specific type of regular graph that forms a lattice-like structure. A lattice graph consists of a regular arrangement of vertices in a grid-like pattern, where each vertex is connected to its adjacent vertices.

### 3.7.3 Dijkstra's algorithm

Dijkstra's algorithm is a well-known and widely used algorithm for finding the shortest path between a given source node and all other nodes in a graph. It guarantees optimality, meaning it always finds the shortest path, as long as the graph has non-negative edge weights Dijkstra, 2022.

Dijkstra's algorithm uses a cost function to calculate the shortest path from a given start node to all other nodes in a graph. The distance in Djikstra's case is represented as a cost function of  $g(x)$ .

$$g(x) = \min(g(x_{goal}), g(x_n)) \quad (3.13)$$

The function  $g(x_n)$  represents the shortest distance from the start node to a given node  $x_n$  in the graph. It is initialized to a very large value, often infinity, for all nodes except the start node, which is assigned a distance of 0.

$$g(x_n) = \sum_{i=1}^{n-1} g(x_{i-1}) + c(x_{n-1}, x_n) \quad (3.14)$$

where,  $\sum_{i=1}^{n-1} g(x_{i-1})$  is all combined cost from a start node to the current node, and  $c(x_{n-1}, x_n)$  is a risk-weighted cost function, which describes to

$$c(x_{n-1}, x_n) = \frac{r(x_{n-1}) + r(x_n)}{2} \cdot d(x_{n-1}, x_n) \quad (3.15)$$

where the risk  $r_x$  is the average between two nodes, and  $d(x_{n-1}, x_n)$  is an euclidean distance between two nodes.

The algorithm iteratively selects the node with the minimum distance among the unvisited nodes and updates the distances of its neighboring nodes if a shorter path is found. This process continues until all nodes have been visited or until the goal node is reached, depending on the specific problem requirements.

By continuously updating the distances of nodes based on the weights of the edges, Dijkstra's algorithm ensures that it explores paths with lower total costs, gradually converging towards the shortest path from the start node to each node in the graph.

```
1: function DIJKSTRA(start, goal)
2:   distances  $\leftarrow \{\}$ 
3:   previous  $\leftarrow \{\}$ 
4:   for all node in graph do
5:     distances[node]  $\leftarrow \infty$ 
6:     previous[node]  $\leftarrow$  undefined
7:   end for
8:   distances[start]  $\leftarrow 0$ 
9:   queue  $\leftarrow$  PriorityQueue()
10:  queue.insert(start, 0)
11:  while queue is not empty do
12:    current  $\leftarrow$  queue.pop_min()
13:    if current is goal then
14:      return ReconstructPath(previous, current)
15:    end if
16:    for all neighbor of current do
17:      distance  $\leftarrow$  distances[current] + cost(current, neighbor)
18:      if distance < distances[neighbor] then
19:        distances[neighbor]  $\leftarrow$  distance
20:        previous[neighbor]  $\leftarrow$  current
21:        queue.insert(neighbor, distance)
22:      end if
23:    end for
24:  end while
25:  return Failure
26: end function

27: function RECONSTRUCTPATH(previous, current)
28:   path  $\leftarrow$  [current]
29:   while previous[current] is defined do
30:     current  $\leftarrow$  previous[current]
31:     path.prepend(current)
32:   end while
33:   return path
34: end function
```

Figure 3.13: Dijkstra's pseudocode algorithm

### 3.7.4 A\* algorithm

A\* algorithm is an extension of the Dijkstra algorithm and uses a heuristic function to guide the search. From 3.16 we can see that the addition of heuristic function  $h(x_n)$ . Noted that the implementation of  $g(x)$  is similar to previously mentioned in 3.14.

The cost function can be represented as:

$$f(x) = g(x_n) + h(x_n) \quad (3.16)$$

where:

- $f(n)$  is the total estimated cost of node n,
- $g(n)$  is the actual cost from the start node to node n,
- $h(n)$  is the heuristic estimate of the remaining cost from node n to the goal node.

The cost function in the A\* algorithm typically consists of two components: the actual cost to reach a vertex from the start node, and the heuristic estimate of the remaining cost from the vertex to the goal node. The sum of these two components represents the total estimated cost from the start node to the goal node through a particular vertex.

The heuristic estimate,  $h(x_n)$ , must be an admissible (underestimating) heuristic function that provides an estimate of the remaining cost from node n to the goal node. Admissibility is a requirement for the algorithm to be able to look for the best solution in the graph. If admissibility is not implemented, overestimation might occur in the algorithm, and it may overlook nodes that would lead to the optimal solution.

The solution for admissible heuristic implementation is based upon the paper from Primatesta et al., 2019. Where the author *dialed down* the heuristic value by multiplying it by the risk cost between  $x_{goal-1}$  and the  $x_{goal}$  itself, this is to ensure conservative heuristic admissibility.

$$\begin{aligned} h(x_n) &= \frac{r(x_n) + r_{min}}{2} \cdot \text{dist}(x_n, x_{n+1}) + \text{dist}(x_{n+1}, x_{goal-1})r_{min} \\ &\quad + \frac{r(x_{goal}) + r_{min}}{2} \cdot \text{dist}(x_{goal-1}, x_{goal}) \end{aligned} \quad (3.17)$$

By summing the actual cost and the heuristic estimate, the A\* algorithm prioritizes nodes with lower total costs, resulting in a search that tends to explore paths that are likely to be closer to the goal. This heuristic-guided exploration helps to efficiently find the shortest path from the start node to the goal node.

To summarize, when there is a single unique optimal path, A\* can discover it more efficiently than Dijkstra's algorithm by utilizing an admissible heuristic function. While Dijkstra's algorithm has the capability to eventually find the absolute optimal path given sufficient time.

```
1: function AStar(start, goal)
2:   openSet  $\leftarrow \{start\}$ ; closedSet  $\leftarrow \{\}$ 
3:   cameFrom  $\leftarrow \{\}$ ; gScore  $\leftarrow \{\}$ 
4:   hScore  $\leftarrow \{\}$ ; fScore  $\leftarrow \{\}$ 
5:   gScore[start]  $\leftarrow 0$ 
6:   hScore[start]  $\leftarrow \text{HEURISTIC}(start, goal)$ 
7:   fScore[start]  $\leftarrow gScore[start] + hScore[start]$ 
8:   while openSet is not empty do
9:     current  $\leftarrow$  node in openSet with the lowest fScore
10:    if current is goal then
11:      return RECONSTRUCTPATH(cameFrom, current)
12:    end if
13:    openSet  $\leftarrow openSet \setminus \{current\}$ 
14:    closedSet  $\leftarrow closedSet \cup \{current\}$ 
15:    for all neighbor in NEIGHBORS(current) do
16:      if neighbor in closedSet then
17:        continue
18:      end if
19:      tentativeGScore  $\leftarrow gScore[current] + COST(current, neighbor)$ 
20:      if neighbor not in openSet or tentativeGScore < gScore[neighbor]
21:      then
22:        cameFrom[neighbor]  $\leftarrow current$ 
23:        gScore[neighbor]  $\leftarrow$  tentativeGScore
24:        hScore[neighbor]  $\leftarrow \text{HEURISTIC}(neighbor, goal)$ 
25:        fScore[neighbor]  $\leftarrow gScore[neighbor] + hScore[neighbor]$ 
26:        if neighbor not in openSet then
27:          openSet  $\leftarrow openSet \cup \{neighbor\}$ 
28:        end if
29:      end if
30:    end for
31:  end while
32:  return Failure
33: end function

34: function RECONSTRUCTPATH(cameFrom, current)
35:   path  $\leftarrow [current]$ 
36:   while current in cameFrom do
37:     current  $\leftarrow cameFrom[current]$ 
38:     PREPEND(current, path)
39:   end while
40:   return path
41: end function
```

Figure 3.14: A\* pseudocode algorithm

## 3.8 Implementation of risk-based UAV path finder

### 3.8.1 Instrumentation

UAV Pathfinding eXtension, UPX, is both a python libraries and application to find the optimal risk-informed flight path. In general, UPX works by retrieving geographic data from OSM, receiving geographic data, then being discretized and aggregated. The risk map is then converted into a graph for the pathfinding algorithm. The risk mentioned Risk in UPX is rules to determine the probability of fatality in case the UAV happens to malfunction and falls in that area.

UPX was built based on the concern of negative consequences of increase frequency of UAV flying in urban areas. While UPX is intended to be used to aid UAV operation, its library can be used in any type of aircraft as long they provide aerodynamic parameters and aircraft performance data. However, some caveats would be explained in the UPX limitation below.

Since UPX is strictly for a preliminary assessment only. It does not handle real-world phenomena such as sudden wind gale, rain, storm, or other aircraft that, unfortunately, are present in the path of the UAV. For that reason only, we hope the pilot practice safe flying and the UAV itself is provided with the failsafe feature.

### 3.8.2 General working principle

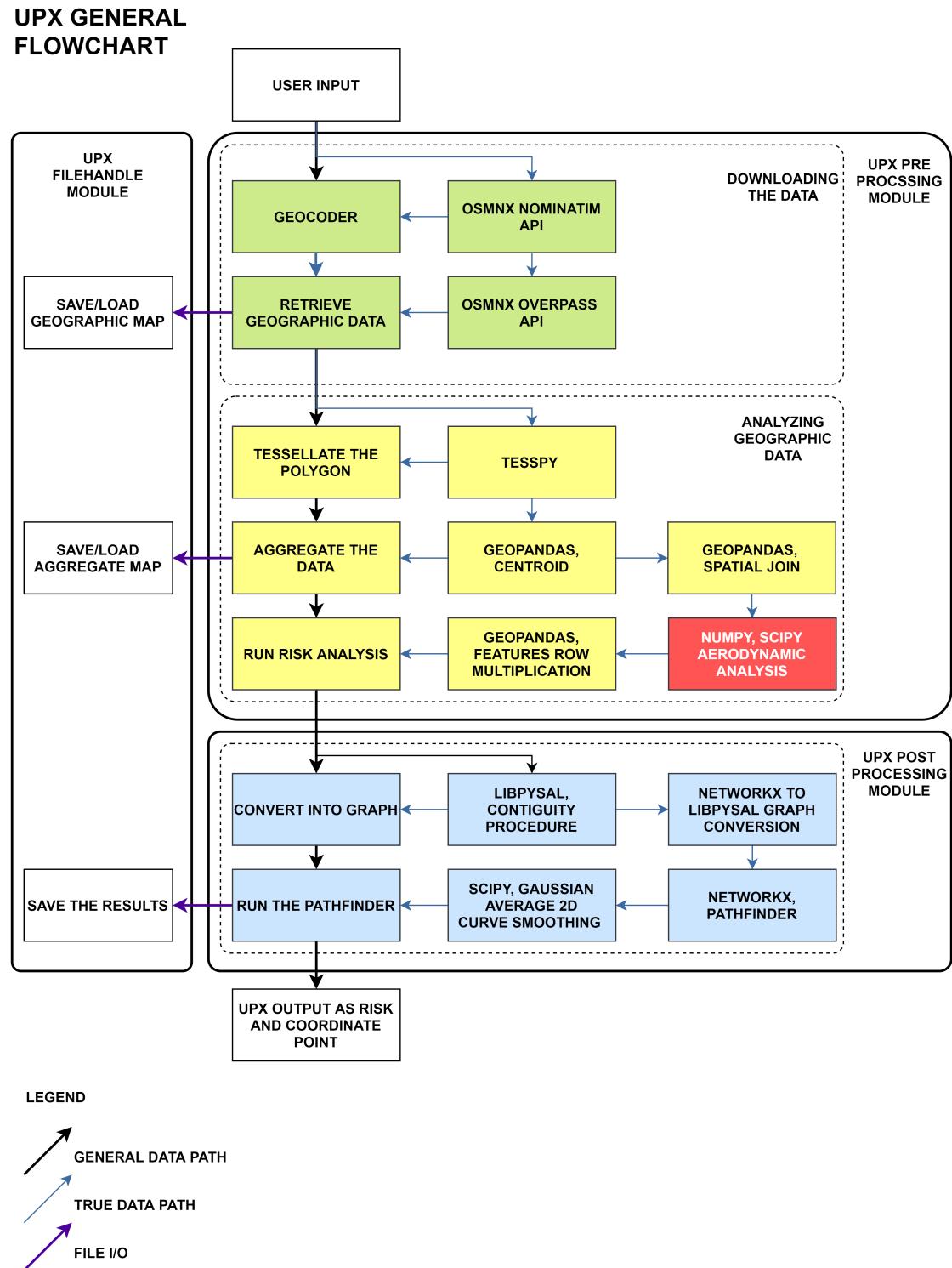


Figure 3.15: UPX architecture design

## Data Collection Using OSM and Nominatim

In the initial stage of the UPX architecture, data collection is performed using OpenStreetMap and the Nominatim API. User input is processed to retrieve relevant geographic data, which includes points of interest, streets, buildings, and other geographic features. The Nominatim API is used to geocode addresses and locations, converting them into geographic coordinates. This collected data forms the foundation for further processing in UPX.

## Tessellation and Geopandas Centroid Calculation

Once the geographic data has been collected, the next step involves tessellating the data into polygons using TessPy. This tessellation process is essential as it converts the continuous geographic area into a discrete grid array, making it easier to manage and analyze. Each cell in the grid represents a polygon, which will be used for detailed analysis. Noted that polygon is a term use by GIS community to indicate any 2D spatial shape, which can include triangle to any n-gon. GeoPandas is then used to find the centroid of each polygon. The centroid acts as an anchor position for the cell, providing a reference point that simplifies further calculations and analyses. To give more clarity refer to figure 3.16 below

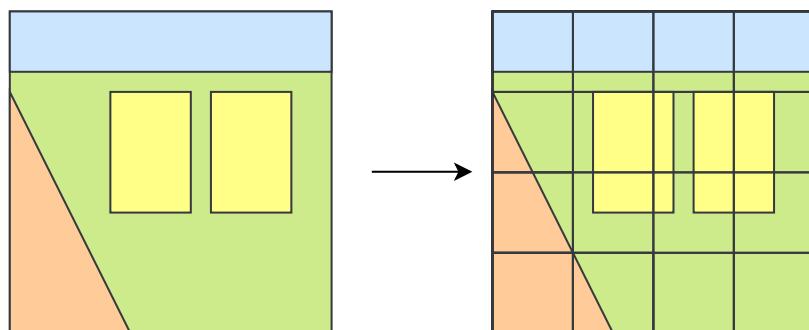


Figure 3.16: Tesselation process, color does not represent anything

## Spatial Join and Attribute Aggregation

After tessellation, the polygons undergo spatial joins using GeoPandas. This process links the geographic cells with their respective attributes, such as population density, land use, and other risk factors. Aggregating these attributes is crucial for the risk analysis, as they help determine the probability of fatality in case of a UAV malfunction. If you take a closer look into the figure 3.17 below, the upper part of the layer map is applied aggregate per row basis, or per field basis. This is because map provided by the OSM is already spatially align with each other, since it uses the same CRS. However we need to combine both the aggregated geographic map and the actual impact PDF. This is because the there are spatial differences we need

to align them using spatial join. Spatial join, or to be precised, spatial inner join essentially stack on top of another

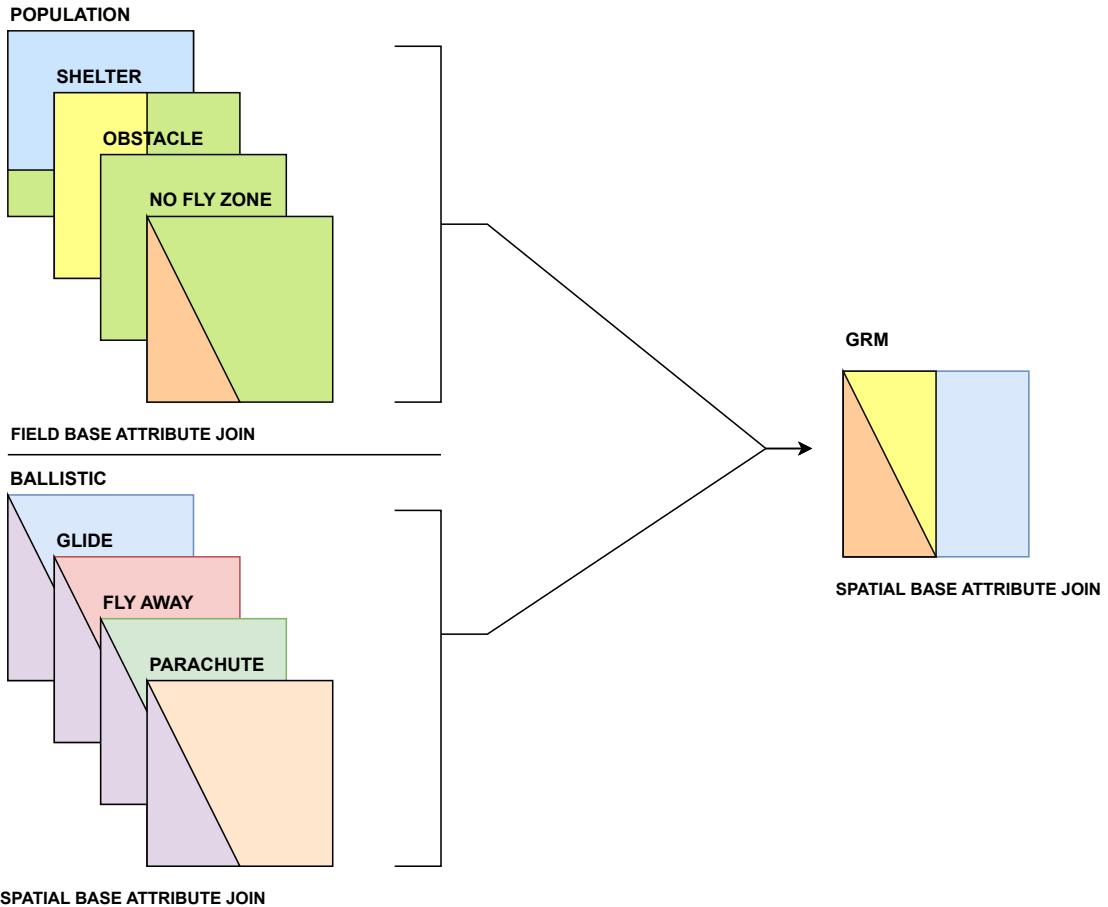


Figure 3.17: Spatial join and attribute aggregation, color does not represent anything

## Aerodynamic Analysis

Each cell undergoes detailed aerodynamic analysis, where calculations for various descent scenarios are performed. These scenarios include ballistic descent, glide descent, fly away, and parachute deployment. The results of these calculations are incorporated into the risk map, providing a comprehensive assessment of risk levels across different geographic areas.

## Risk Analysis and Ground Risk Module Computation

The risk analysis involves computing the probability of descent events, impacts, and fatalities. This is processed using a ground risk module, which evaluates the likelihood of these events occurring within each cell. This step is critical for un-

derstanding the potential consequences of a UAV malfunction and for creating an informed risk map.

### Converting Cells into a Graph

The next stage involves converting the tessellated cells and their associated risk attributes into a graph structure suitable for pathfinding algorithms. This is achieved using the libpysal Queen Contiguity procedure, which connects each node (representing a cell) to its neighboring nodes. This graph structure allows for efficient navigation and analysis. Take a look at figure 3.18 below to see the concept being applied.

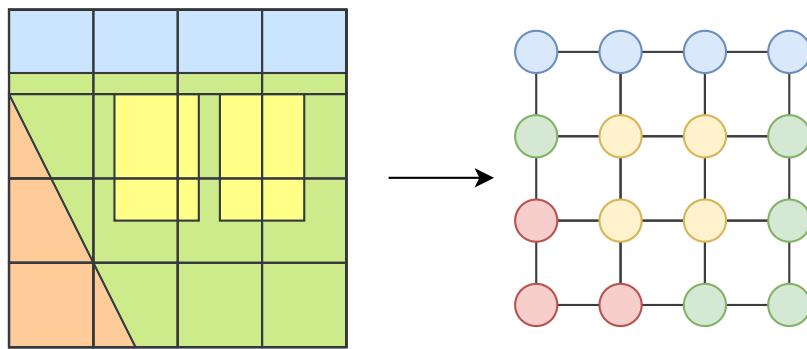


Figure 3.18: Graph convert, color does not represent anything

### Pathfinding Using Risk-A\*

The stage involves running the Pathfinder algorithm. This algorithm uses the risk-informed graph and applies the Risk-A\* method to identify the optimal flight path for the UAV. The Risk-A\* algorithm takes into account the risk values associated with each cell, ensuring the UAV avoids high-risk areas and selects the safest possible route.

### Smoothing the Path

To avoid jagged paths, a Gaussian average is applied for path smoothing. This ensures that the resulting flight path is more realistic and manageable, reducing unnecessary sharp turns and optimizing the route.

### FileHandle Module for Utility and I/O Purposes

The FileHandle module in UPX is dedicated to utility and I/O purposes. It manages the saving and loading of geographic and aggregate maps, ensuring that data can

be efficiently stored and retrieved throughout the various stages of the process.

### 3.8.3 UPX Python and packages

#### Python

Python is an interpreted programming language. Interpreted language has the benefit of being an easier language to compare to a compiled language. Added to this, it has massive scientific-related packages and libraries. Combine with quick deployment and its STEM library, it is a perfect choice for engineering, mathematics, and physics analysis. Based on previous criteria it is a great choice for creating UAV path planner applications and libraries.

While Python indeed has a great advantage. Unfortunately, its comparatively easy language came with a great price. An interpreted language is inherently slow, 8 to 100 times slower compared to C/C++ language. Base Python language also cannot use more than one thread because its GIL, Global Interpreter Lock, prevents the threads from sharing memories and create a parallel workload. Fortunately, most of the heavy calculation is handled by the c/c++ library, where Python uses the caller function to launch the libraries.

#### GeoPandas

GeoPandas is one of the GIS libraries built for tabulating and analyzing geospatial analysis in python. GeoPandas itself can be regarded as the extension of Pandas library, where Pandas is general tabular data processor. GeoPandas has geographic spatial and attribute related function such as overlay, spatial join, clipping to group, aggregate, and measure geospatial data.

GeoPandas exports and imports files in standard GIS-supported files such as geoJSON, the file, and the Shapely object, to name a few. In Python GeoPandas tabular data is called GeoDataFrame, and GeoSeries. What differentiates between GeoDataFrame and regular DataFrame is the addition of a geometry column. The geometry column can contain Multi/Polygon, Multi/Linestring, and Point shapes.

#### Shapely

Shapely is a geometry calculation and manipulation packages available for python. Many of geospatial function in GeoPandas relies on Shapely itself. Overlay, clipping, spatial binary combination in GeoPandas can be found in Shapely. In this paper Shapely is being used to create custom filter for heat map function of aircraft falling probability.

## Tesspy

Tesspy, a tessellation library for Python. Its purpose, as the name implies, is to create a tesselated polygon. Tesspy can be used as long as the polygon object is derived from the Shapely object. Tesspy offers type of tessellated map that commonly use for geospatial processing. Square grid, hexbin, Voronoi city block, and adaptive square is the type of tesselated maps, to name a few.

## OSMnx

OSMnx is a combination of OSM and NetworkX. OSMnx is mainly being uses to create graph representation of the road network. It uses the GeoPandas data frame, to build the network, as long as the roads in question are connected at least with another road. But for this paper, OSMnx is only being used as a geocoder and reverse geocoder to retrieve GeoPandas data.

## NetworkX

To create a graph structure, NetworkX provides the necessary function to do that. NetworkX offers graph creation, graph traversal algorithms, and graph manipulation. The graph, as previously mentioned, in a combination of nodes and edges. The A\* pathfinder for the UAV Path finder also uses the NetworkX built-in A\*.

## CHAPTER 4

# RESULTS

### 4.1 PDF risk map

#### 4.1.1 Ballistic descent

The provided images illustrate ballistic impact probability maps for four different drones: DJI Inspire , Parrot , DJI Phantom, and Talon. Each map is a grid representation where each cell denotes the probability of the drone falling within that specific area. The color gradient, ranging from purple to yellow, indicates the probability scale, with purple representing lower probabilities and yellow representing higher probabilities.

From these maps, we can observe distinct patterns for each type of drone. The DJI Inspire and DJI Phantom maps show a higher concentration of yellow, indicating a higher probability of falling within certain central areas. This might suggest that these drones, likely due to their multirotor design, have a more focused descent pattern. However since we calculate every drone as a point mass and not accounted for the shape of the UAV, this pattern likely came because both DJI Inspire and Parrot has higher velocity compare to Phantom and Talon, which resulted much greater range covered per grid map when falling down.

Which is why Talon and Disco a exhibit a more widespread distribution of probabilities, with less concentration in specific areas. The Talon's map, in particular, demonstrates a broader area of impact, which aligns with the expected behavior of much higher horizontal velocity compare to the others.

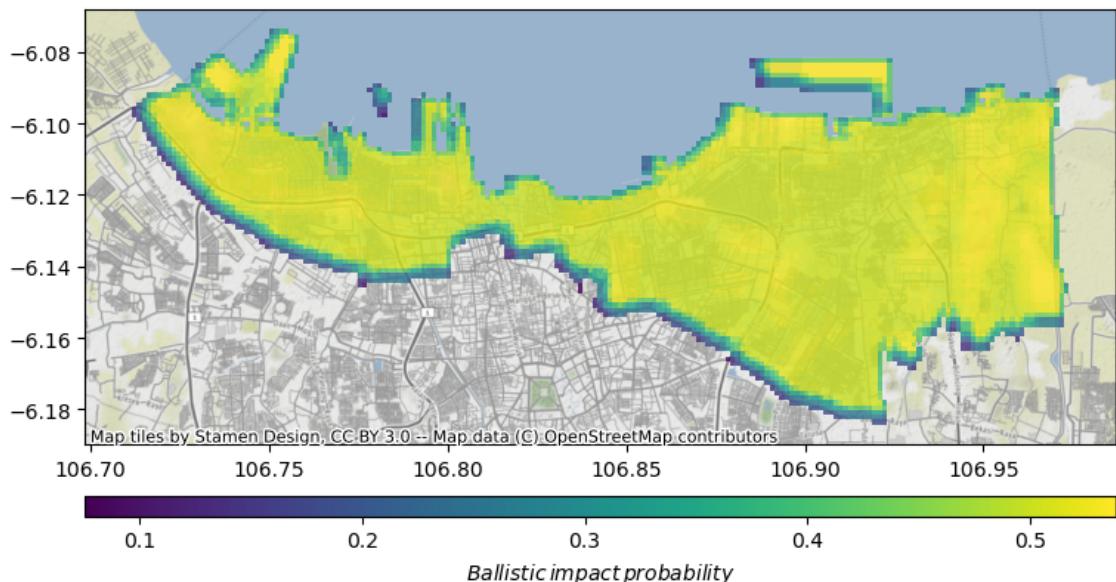


Figure 4.1: Talon ballistic PDF

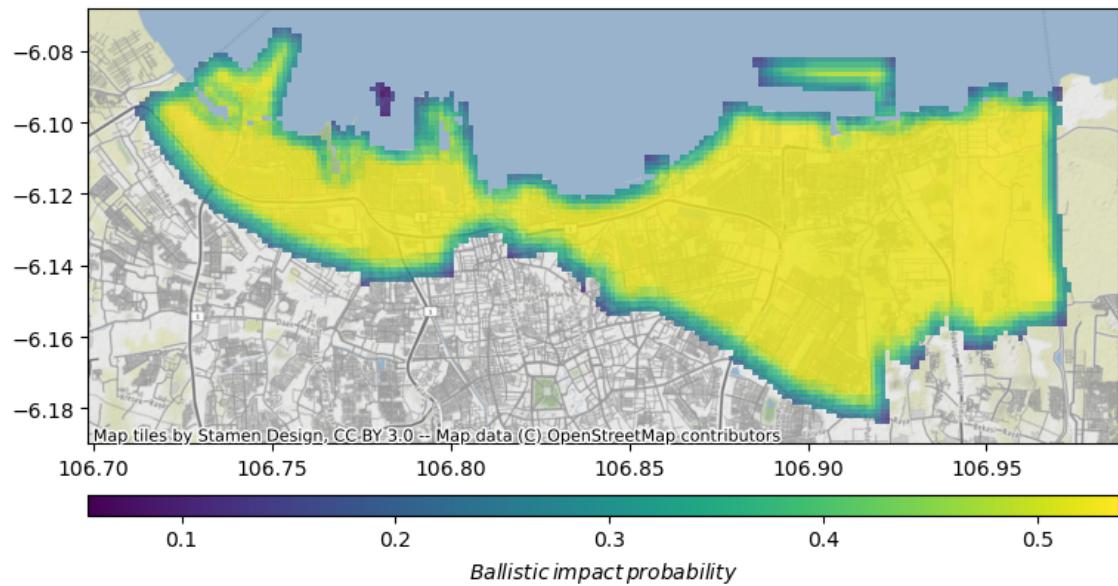


Figure 4.2: Disco ballistic PDF

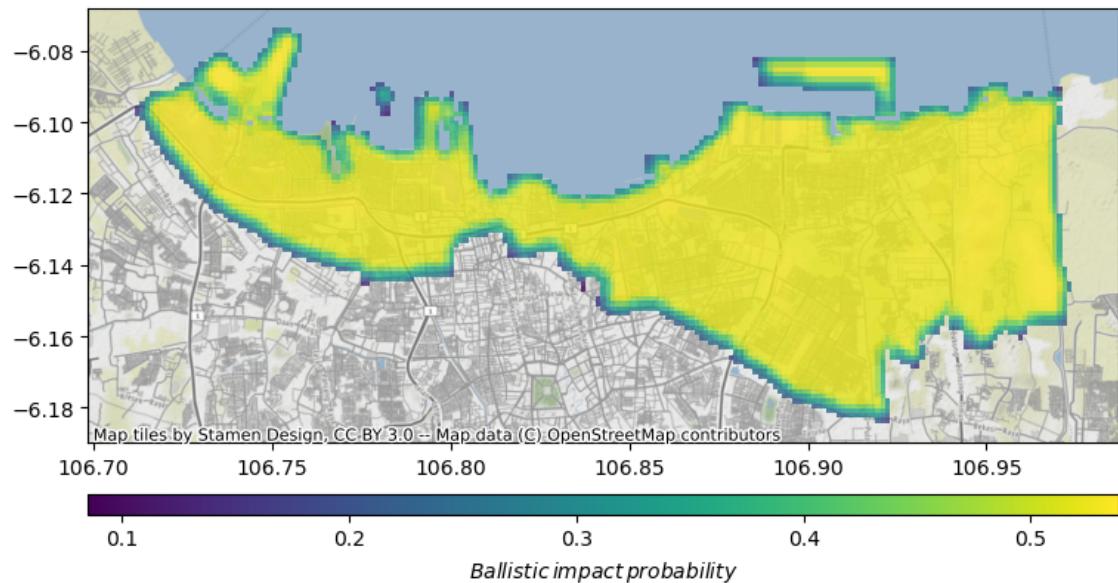


Figure 4.3: DJI Phantom 4 ballistic PDF

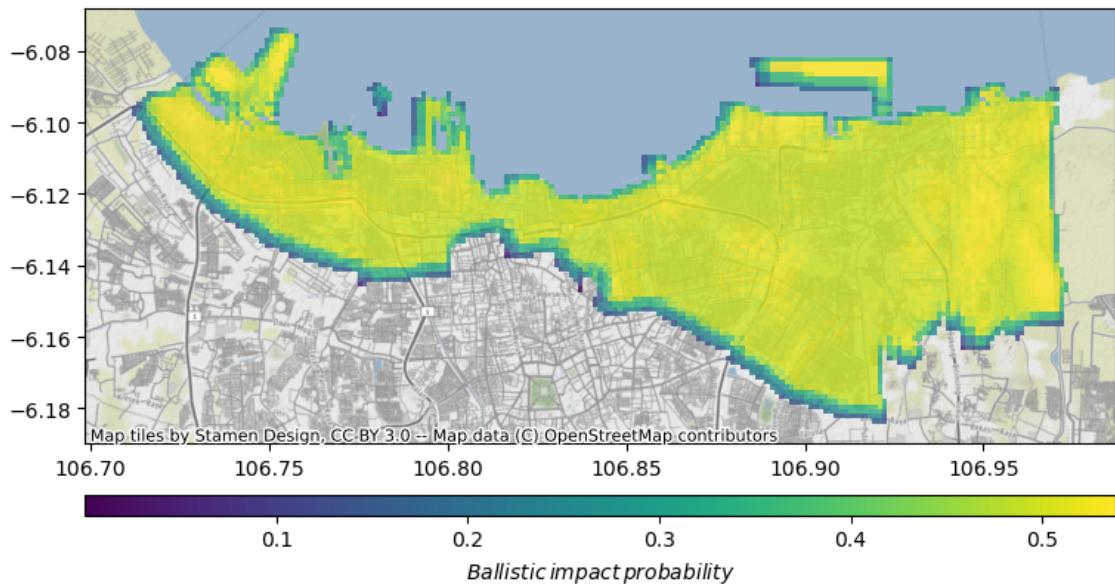


Figure 4.4: DJI Inspire 2 ballistic PDF

#### 4.1.2 Glide

The second set of figure represents glide impact probability maps. These maps display the likelihood of a drone impacting specific areas while gliding, with the color gradient ranging from purple to yellow indicating lower to higher probabilities.

These glide probability maps for these drones does not reflect their notable differences in their descent behaviors. It should be noted from previous chapter that multirotor and fixed wing has a different glide behaviour. Where a multirotor drone, should show a more scattered distribution of probabilities with notable concentrations of yellow in certain areas, suggesting that these drones has lower glide ratio. And a fixed-wing drone, should displays a distinct pattern with broader areas of high probability spread across the map, reflecting its capability to glide over longer distances before descending. However these map show a more or less same behavior. This is because since we using altitude 400 m, where it is the maximum distance allowed by ARC, 2015, even if the fixed wing fall, it will only cover maximum of 4.8 km for the case of talon.

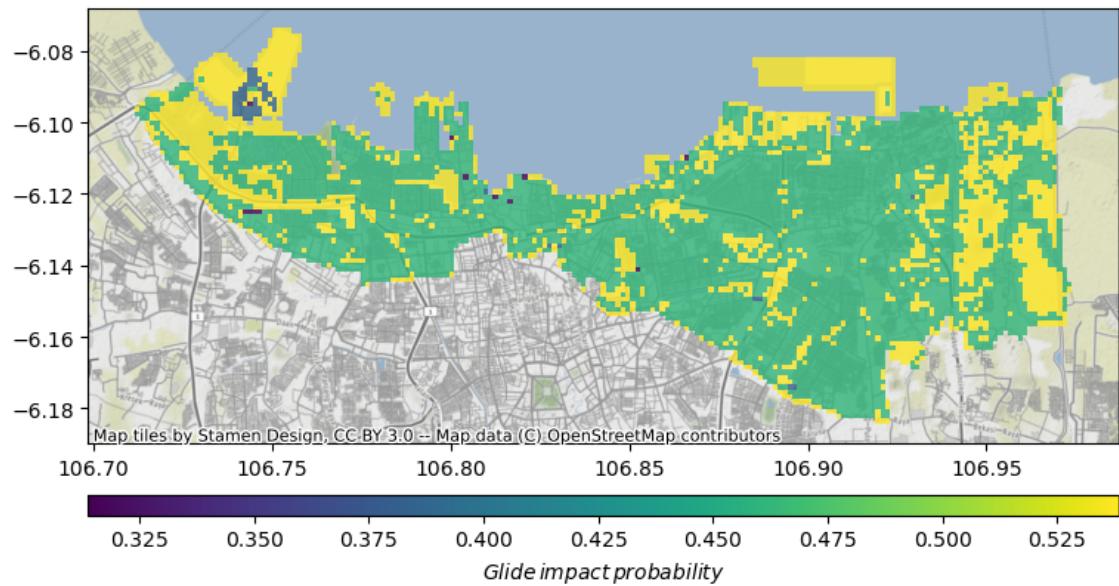


Figure 4.5: Talon glide PDF

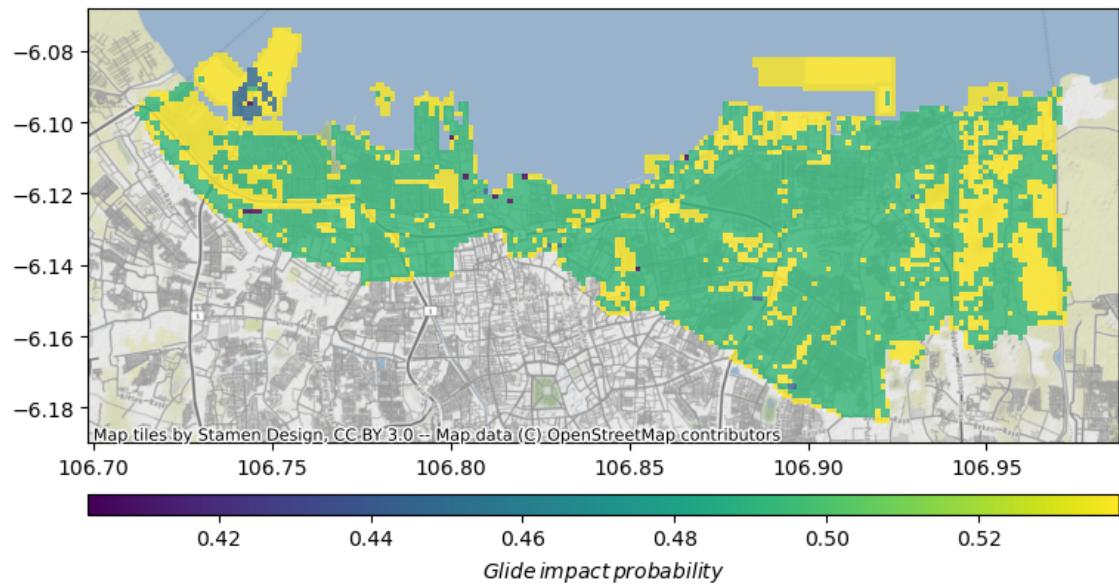


Figure 4.6: Disco glide PDF

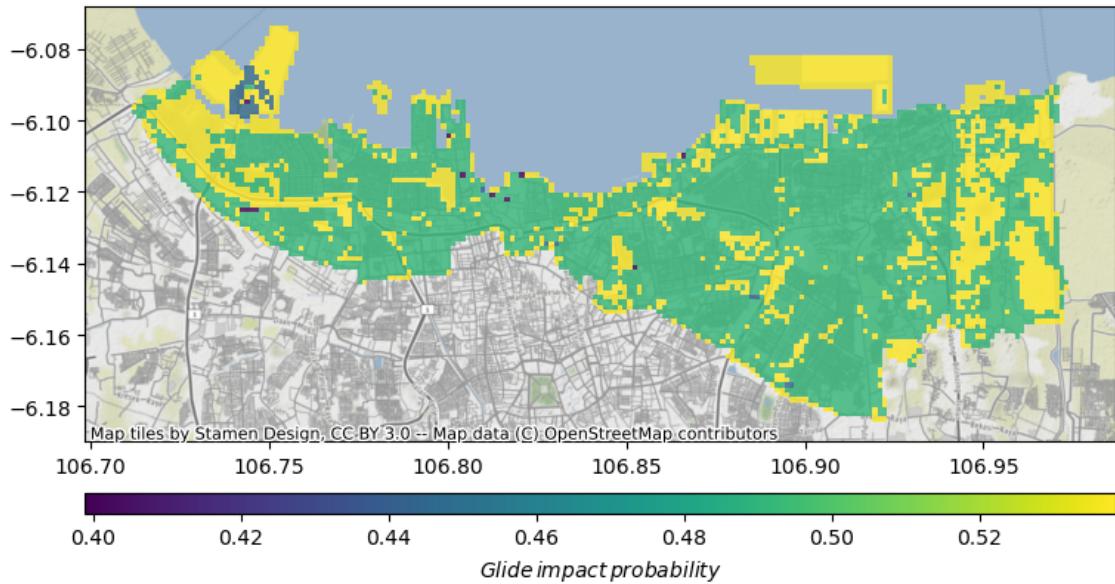


Figure 4.7: DJI Phantom 4 glide PDF

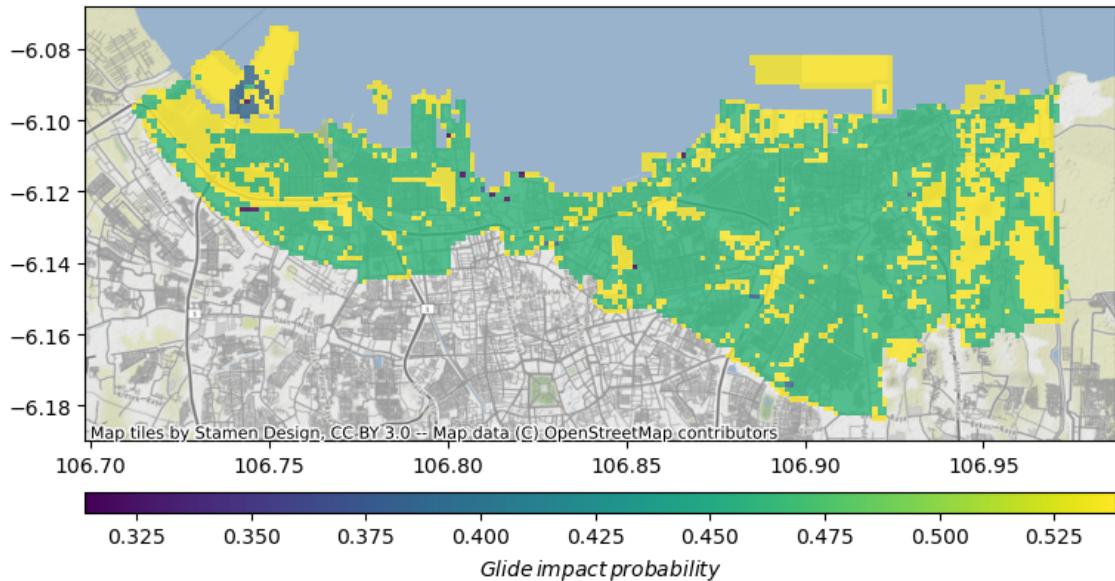


Figure 4.8: DJI Inspire 2 glide PDF

#### 4.1.3 Fly Away

The third set of images showcases flyaway impact probability maps for the same four drones. Examining these maps reveals varied flyaway patterns for each drone type. The DJI Inspire and DJI Phantom show more localized high-probability areas, with significant concentrations of green and yellow, particularly in specific sections of the map. This indicates a higher chance of these drones impacting within confined zones, which may correlate with their operational characteristics and flight control

systems. The DJI Parrot map displays a broader distribution of probabilities, with higher impact probabilities spread more evenly across a larger area, suggesting that in a flyaway event, this drone could potentially impact a wider range of locations. The Talon, consistent with its fixed-wing nature, demonstrates distinct zones with both very low and very high probabilities, reflecting its ability to cover extensive distances during a flyaway scenario.

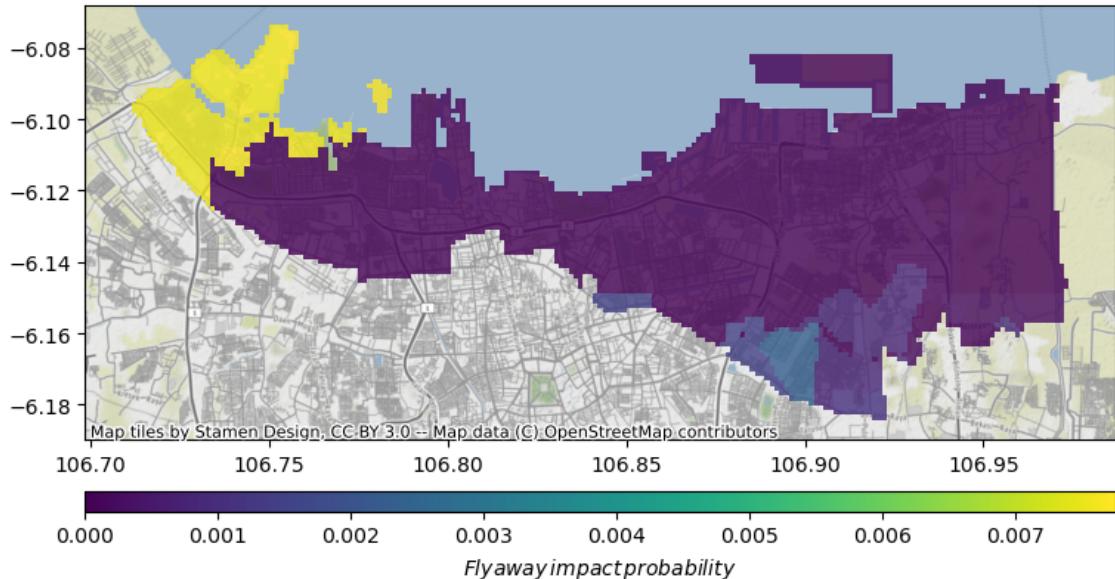


Figure 4.9: Talon fly away PDF

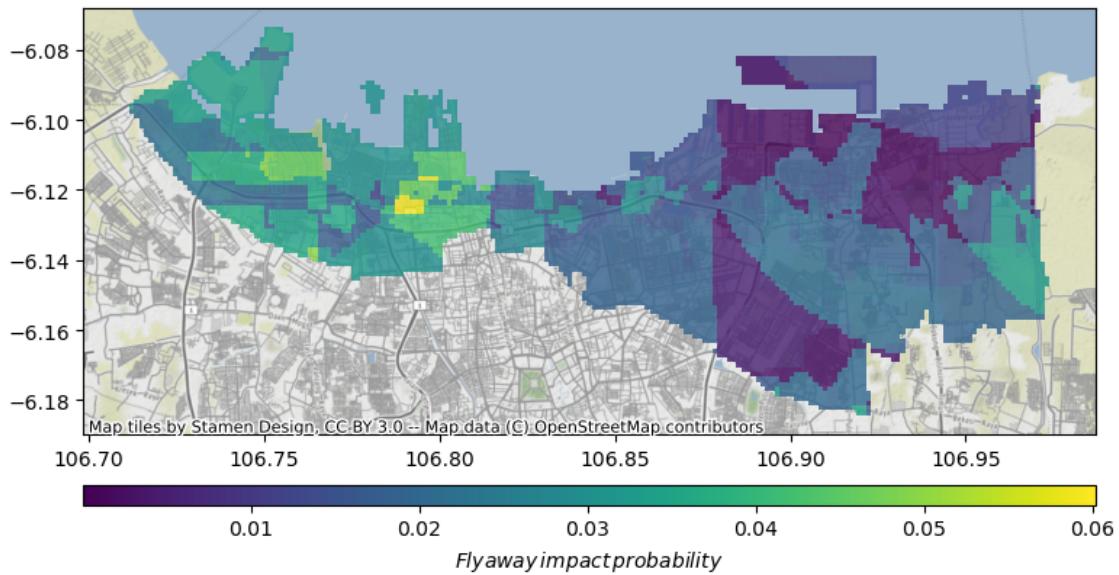


Figure 4.10: Disco fly away PDF

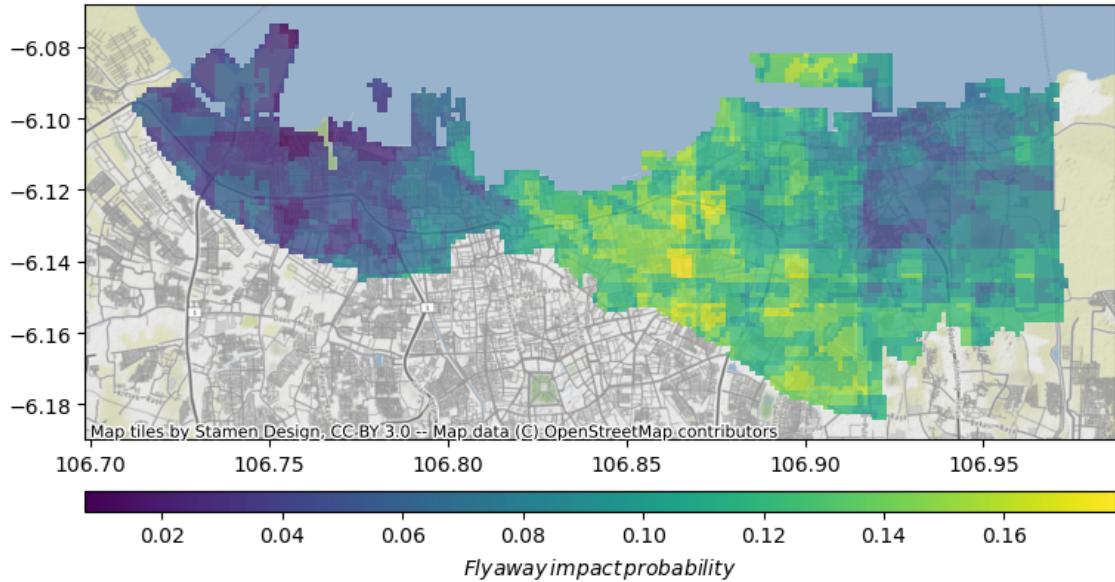


Figure 4.11: DJI Phantom 4 fly away PDF

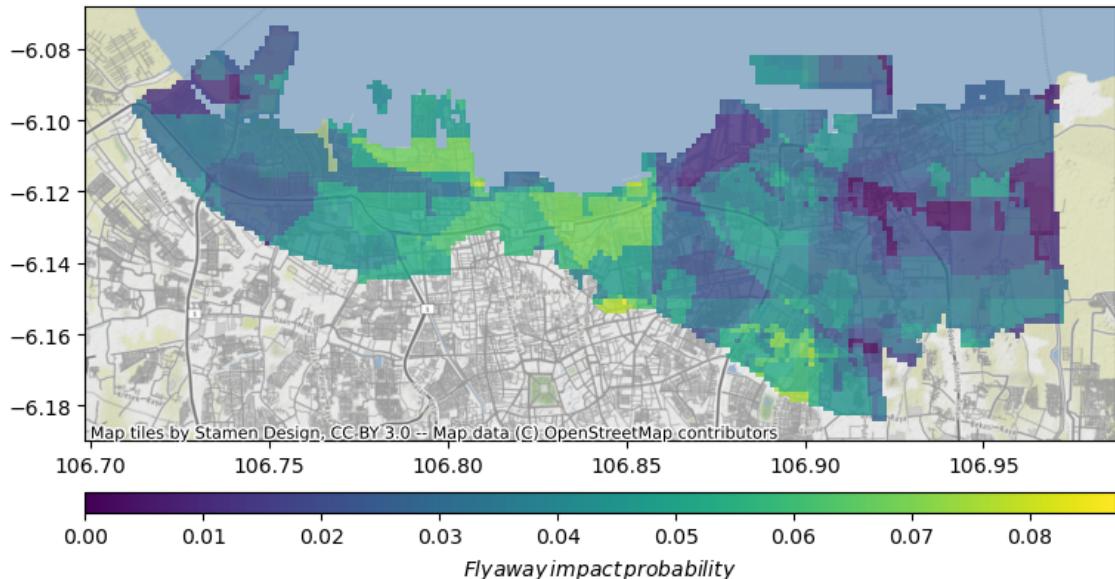


Figure 4.12: DJI Inspire 2 fly away PDF

#### 4.1.4 Parachute

The parachute deployment maps exhibit a consistent and almost uniform distribution across different drones. This uniformity stems from the assumption that the drones will descend solely using parachutes in the event of a malfunction or emergency. Consequently, the descent pattern is less influenced by the specific flight characteristics of the drones, resulting in similar probability distributions regardless of the drone type.

Due to this assumption, the parachute deployment scenarios are characterized by a predictable and controlled descent path. This uniformity in distribution ensures that the impact probabilities are evenly spread across the area, providing a reliable measure for risk assessment and safety planning. The consistent patterns observed in the parachute deployment maps simplify the task of determining safe operational zones and preparing for potential emergency scenarios, as the descent behavior remains largely unaffected by the drone's inherent flight dynamics.

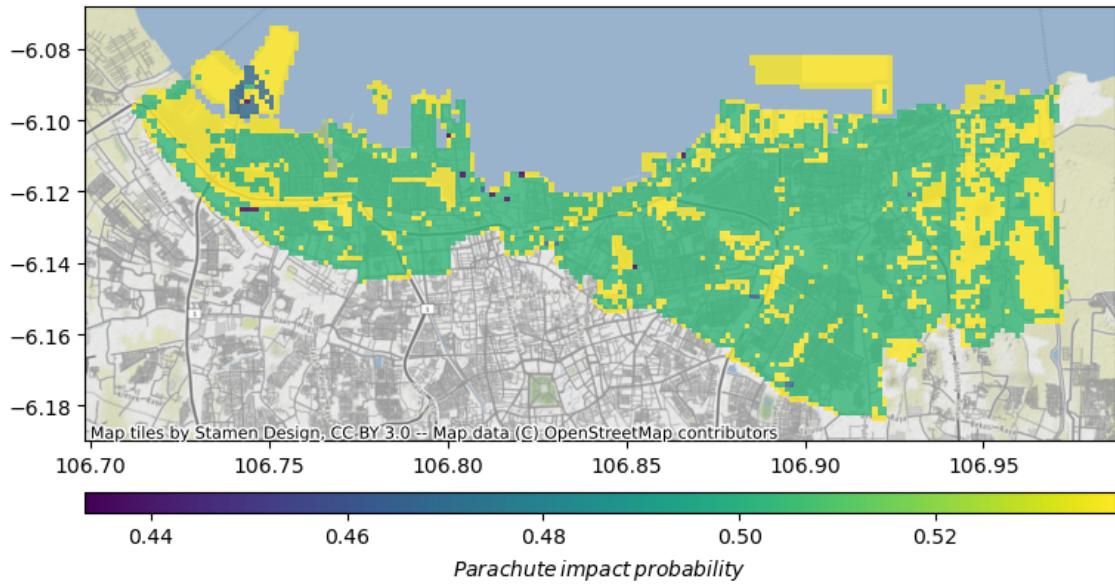


Figure 4.13: Talon parachute PDF

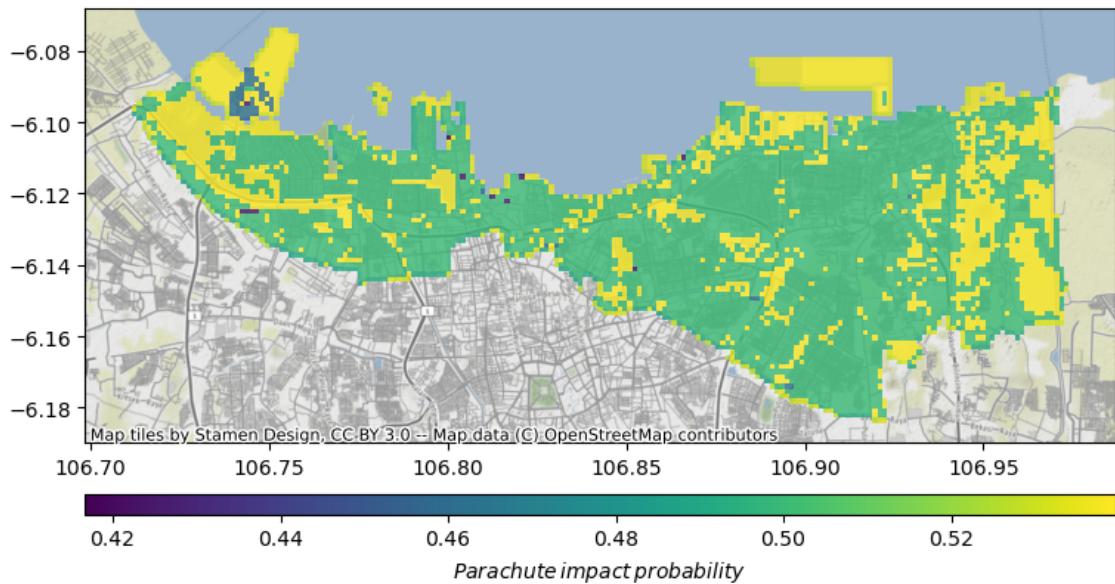


Figure 4.14: Disco parachute PDF

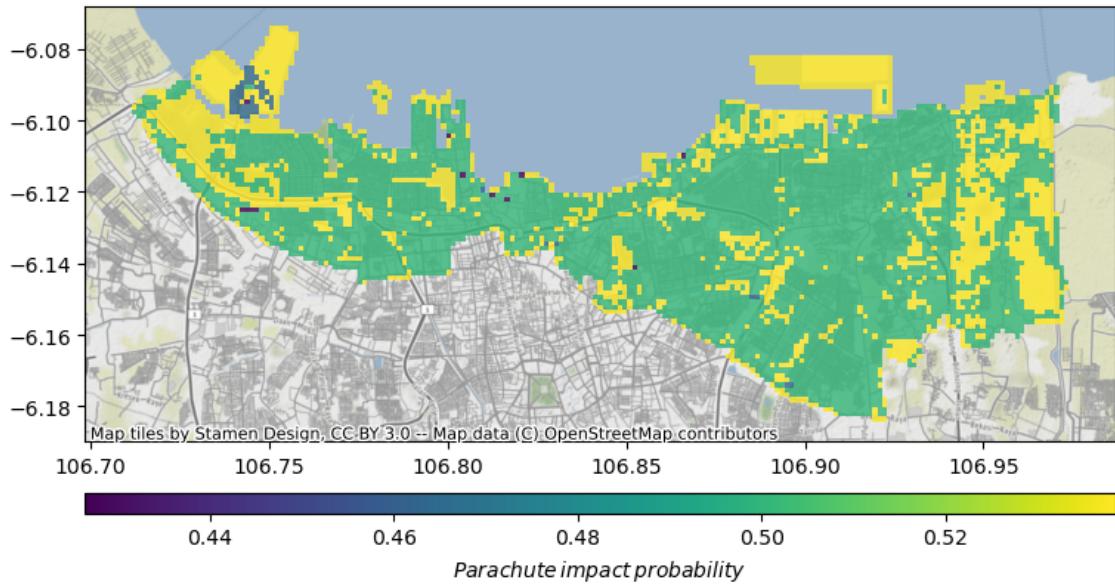


Figure 4.15: DJI Phantom 4 parachute PDF

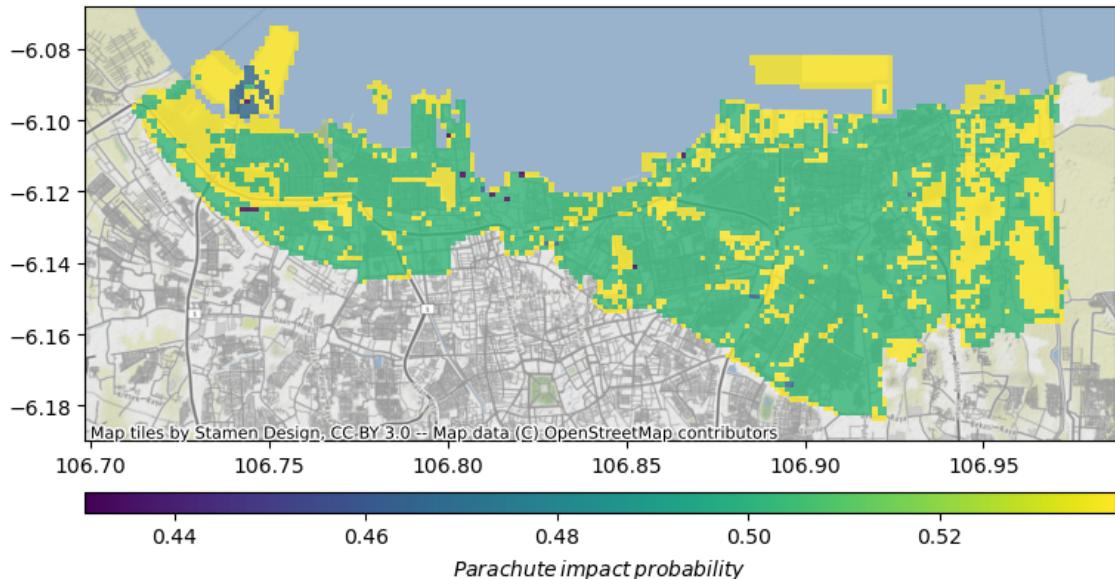


Figure 4.16: DJI Inspire 2 parachute PDF

#### 4.1.5 Total Fatality

A total fatality risk map integrates various hazard models, including ballistic, parachute, glide, and fly-away trajectories, to provide a comprehensive visual representation of potential fatality risks in a given area. This map combines data from multiple descent event to assess the likelihood and impact of different types of hazards that could lead to fatalities. Ballistic models estimate the trajectories and impact points of high-speed projectiles, while parachute models account for objects descending

more slowly under a parachute. Glide models track the paths of objects that can travel significant horizontal distances while descending, and fly-away models consider the risks posed by uncontrolled, drifting objects. By overlaying these models, the total fatality risk map offers a detailed and multifaceted view of potential danger zones, helping to inform safety measures and emergency response planning.

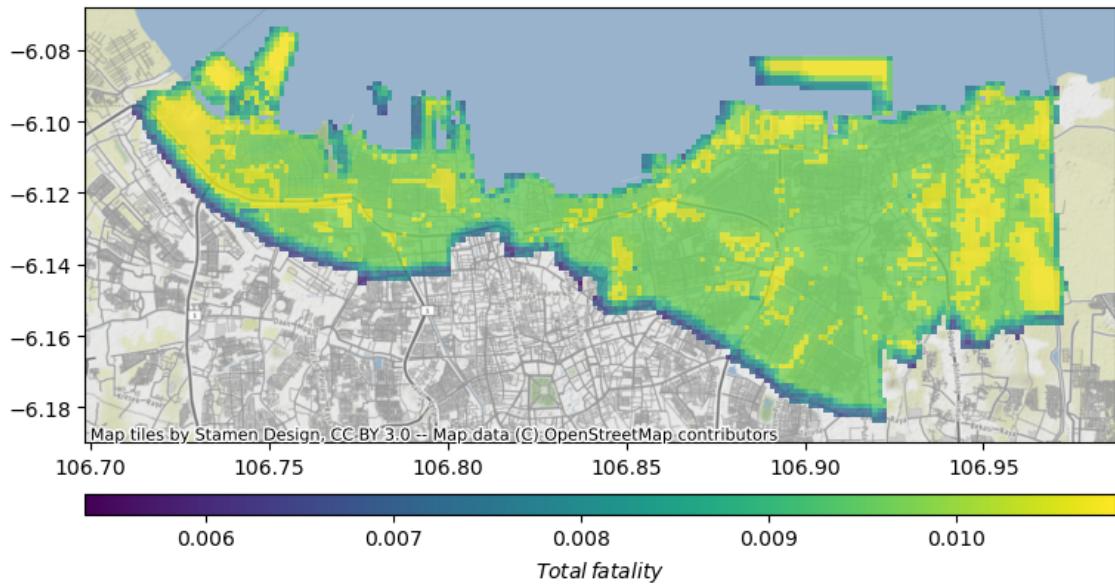


Figure 4.17: Talon total fatality PDF

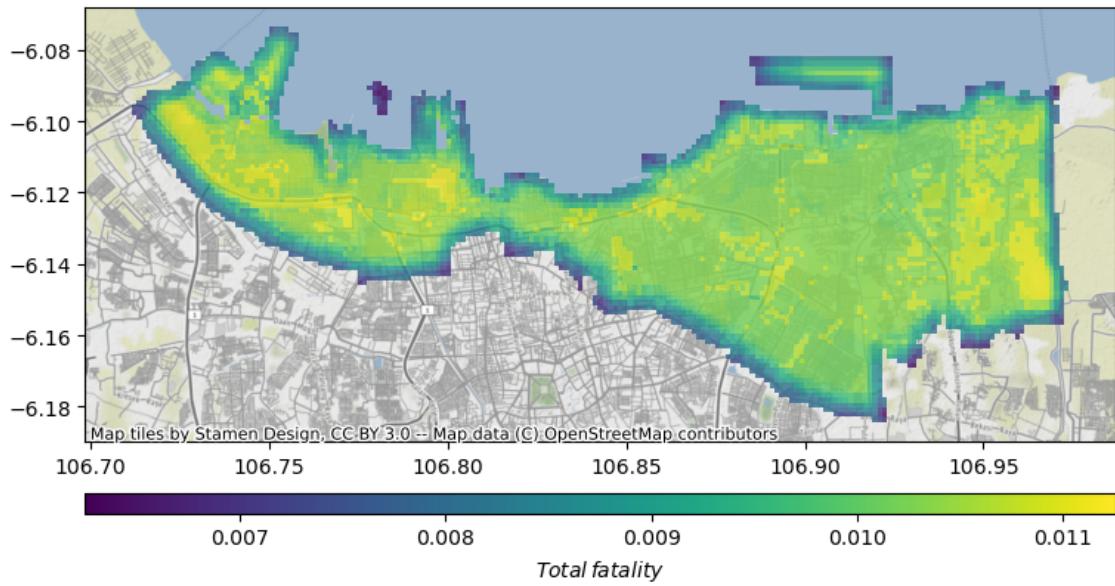


Figure 4.18: Disco total fatality PDF

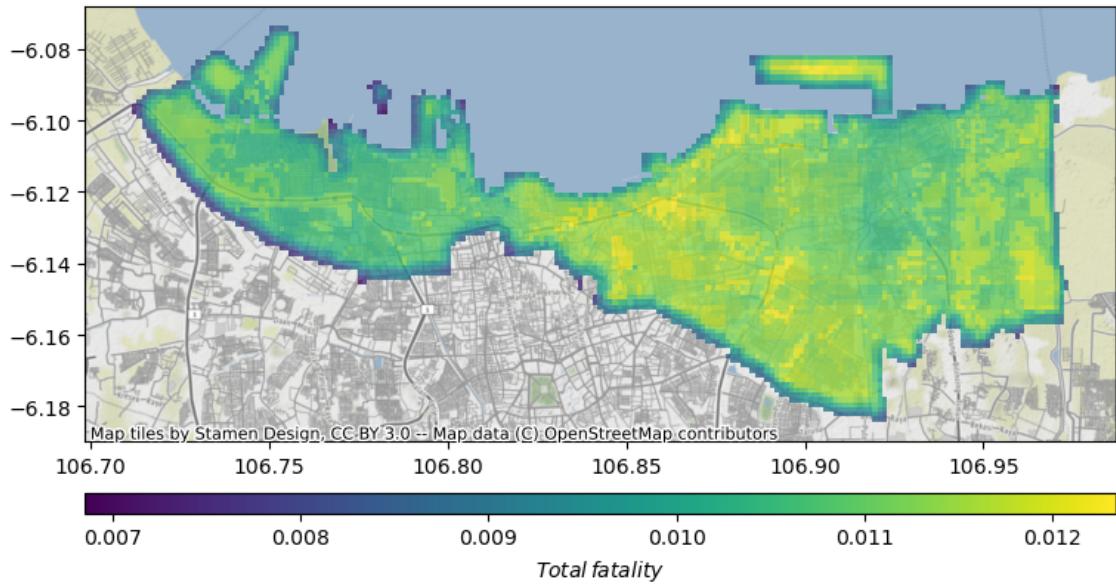


Figure 4.19: DJI Phantom 4 total fatality PDF

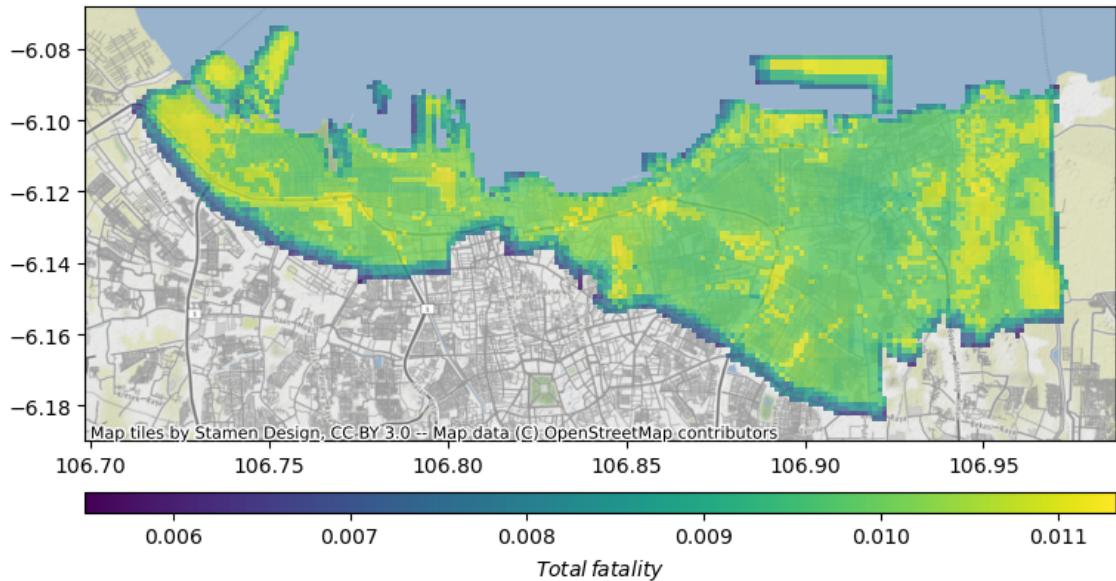


Figure 4.20: DJI Inspire 2 total fatality PDF

## 4.2 Pathfinder result

The implementation of risk-aware A\* algorithms across four drone models shows similarities primarily due to these drones belonging to the same weight class. Although their specific flight characteristics may differ, their classification within the same weight category means they share comparable physical properties. This similarity ensures that, at a glance, the risk-aware A\* algorithms applied to these drones produce similar pathfinding results. However, a closer examination on the next

section reveals distinctions in how each drone's specific risk profiles and operational constraints influence the finer details of route optimization. Despite their shared weight class, we must not ignore the subtle variations in performance, sensor capabilities, and response to environmental factors lead to differences that become apparent in more complex scenarios and longer missions.

#### 4.2.1 Regular A\*

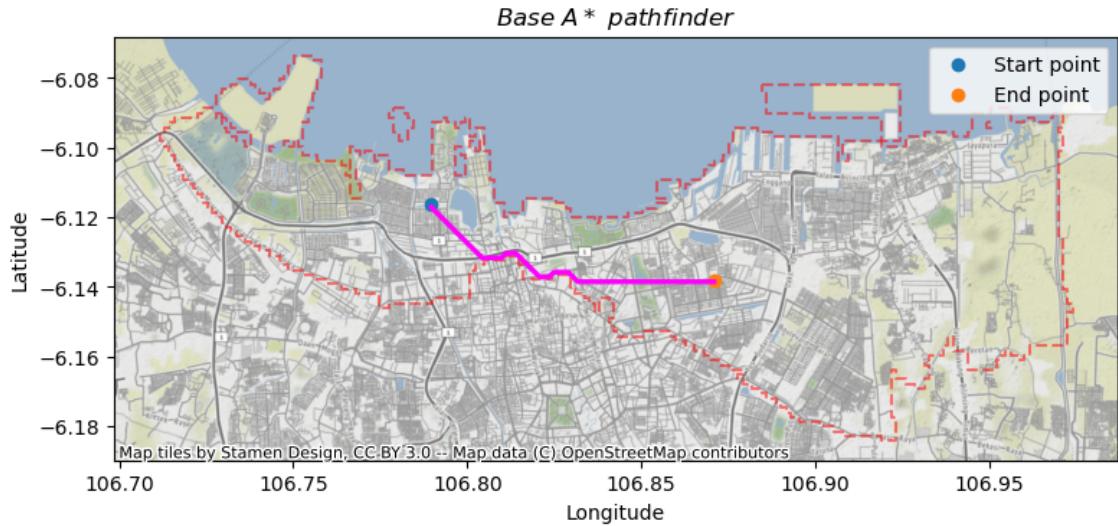


Figure 4.21: Regular A\* flight path

#### 4.2.2 Risk A\*

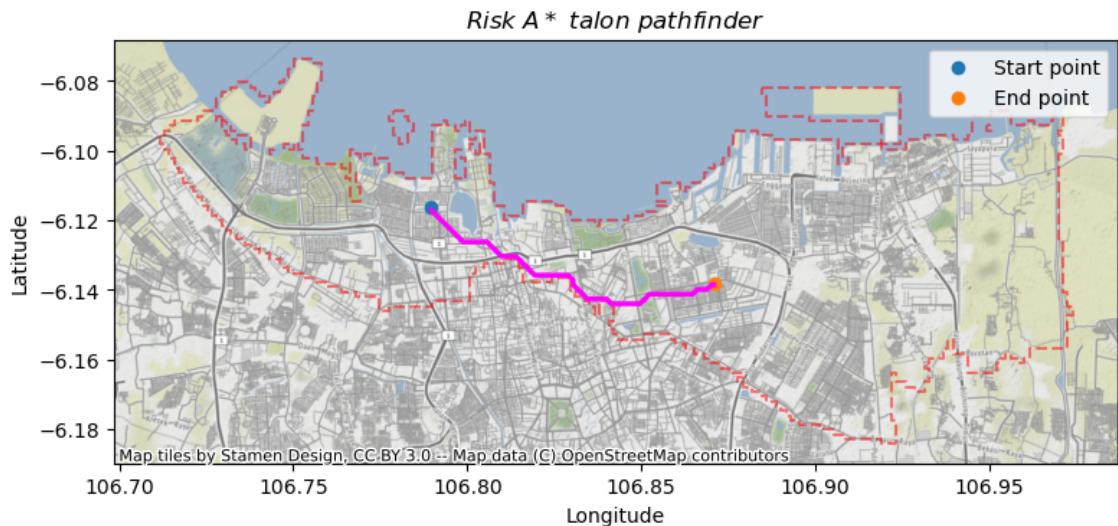


Figure 4.22: Talon risk A\* flight path

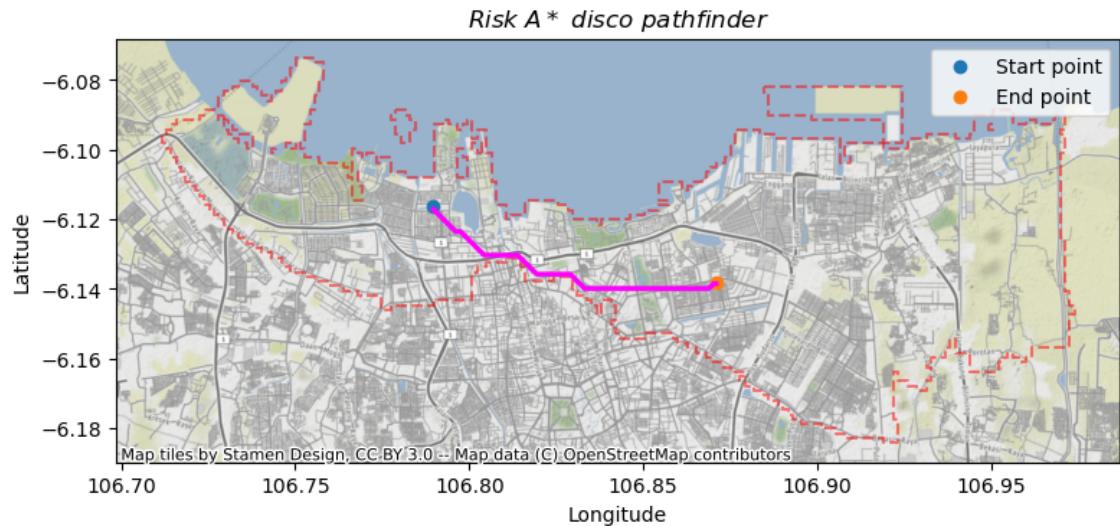


Figure 4.23: Disco risk A\* flight path

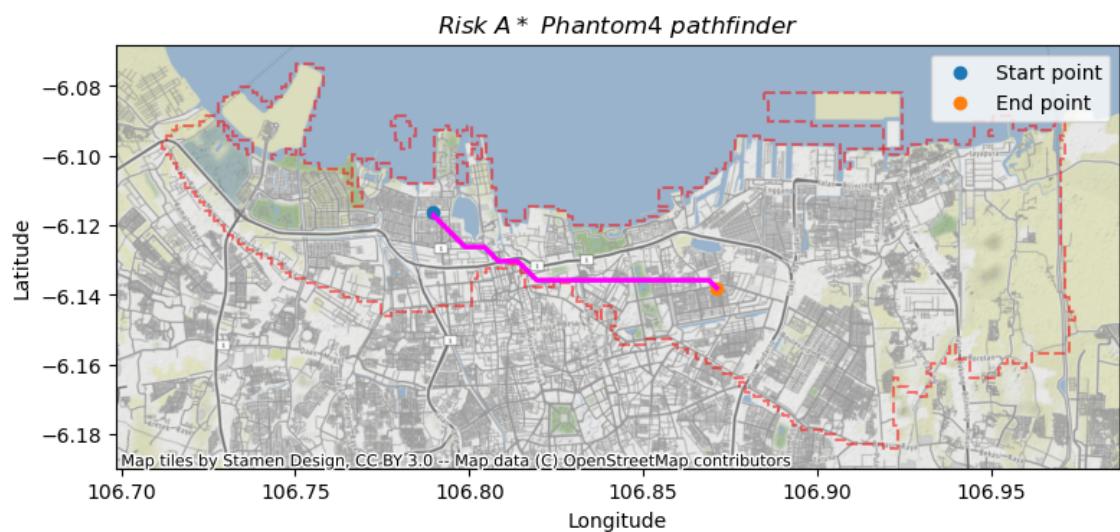


Figure 4.24: DJI Phantom 4 risk A\* flight path

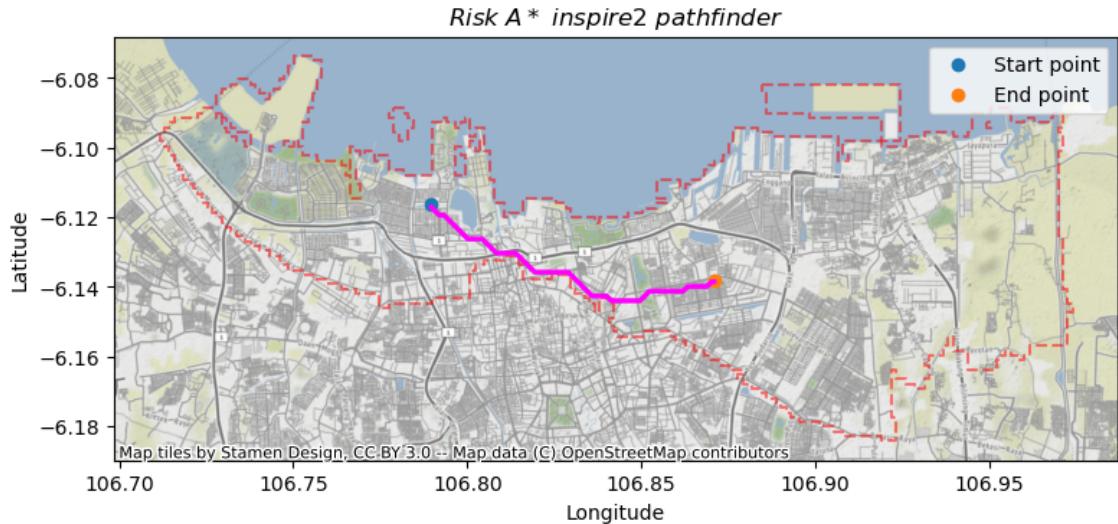


Figure 4.25: DJI Inspire 2 risk A\* flight path

#### 4.2.3 Risk result

This table 4.2 is an expansion of the previous table, providing a detailed explanation of each risk factor associated with the paths taken by various drones. It highlights the specific risks from ballistic, glide, and flyaway incidents, as well as the final risk for each drone. The inclusion of the Normal A\* algorithm showcases a comparative analysis, revealing that while it may incur slightly more risk, it also offers a similar range of risk factors and potentially shorter distances. This analysis helps to better understand the trade-offs between risk and distance for each algorithm.

Normal A\* utilizes the Talon distribution but does not implement the Risk A\* pathfinder. As seen in Figure 4.2, the traveled distance of Normal A\* is the shortest among the risk A\* algorithms. Interestingly, Normal A\* has the highest risk score, which may seem counterintuitive, as one might expect an inverse correlation between Normal A\* and its risk A\* counterpart. While it is not guaranteed for Normal A\* to have the highest risk score, it will consistently produce the shortest path length compared to risk A\*.

To assess different properties, we can measure the risk-to-traveled distance ratio, reflecting how much the algorithm attempts to avoid critical areas by extending the pathway to safer options. This is clearly illustrated in the DJI Phantom 4, which shares almost the same risk as Normal A\* with only a 0.02 difference, yet takes a path that is 163 meters longer.

Upon reviewing the graph, it may be tempting to choose Talon due to its lowest traveled distance and risk score. However, this result stems from the limitations of the program's design, which certainly needs improvement. The problem lies in the flyaway descent of this aircraft. Talon has a long-range flight before crashing into the ground. Since we are only using one cluster of geographic regions, North Jakarta, the region is smaller than the flyaway radius. This results in the outer cells

of the region having a smaller risk compared to the center cells. It is clear that the scores represented by shorter-range UAVs such as Parrot Disco, DJI Phantom 4, and DJI Inspire 2 have more defined scores compared to those produced by Talon.

The table 4.1 and figure 4.26 illustrates the total pathway risks associated with different algorithms, including Talon, Disco, Phantom 4, Inspire 2, and Normal A\*. It outlines the paths taken by these algorithms, the risks involved on those paths, and the average risk per distance traveled. Notably, while the Normal A\* algorithm exhibits a higher total risk involved, it also has the shortest distance traveled, indicating a trade-off between risk and efficiency in pathfinding.

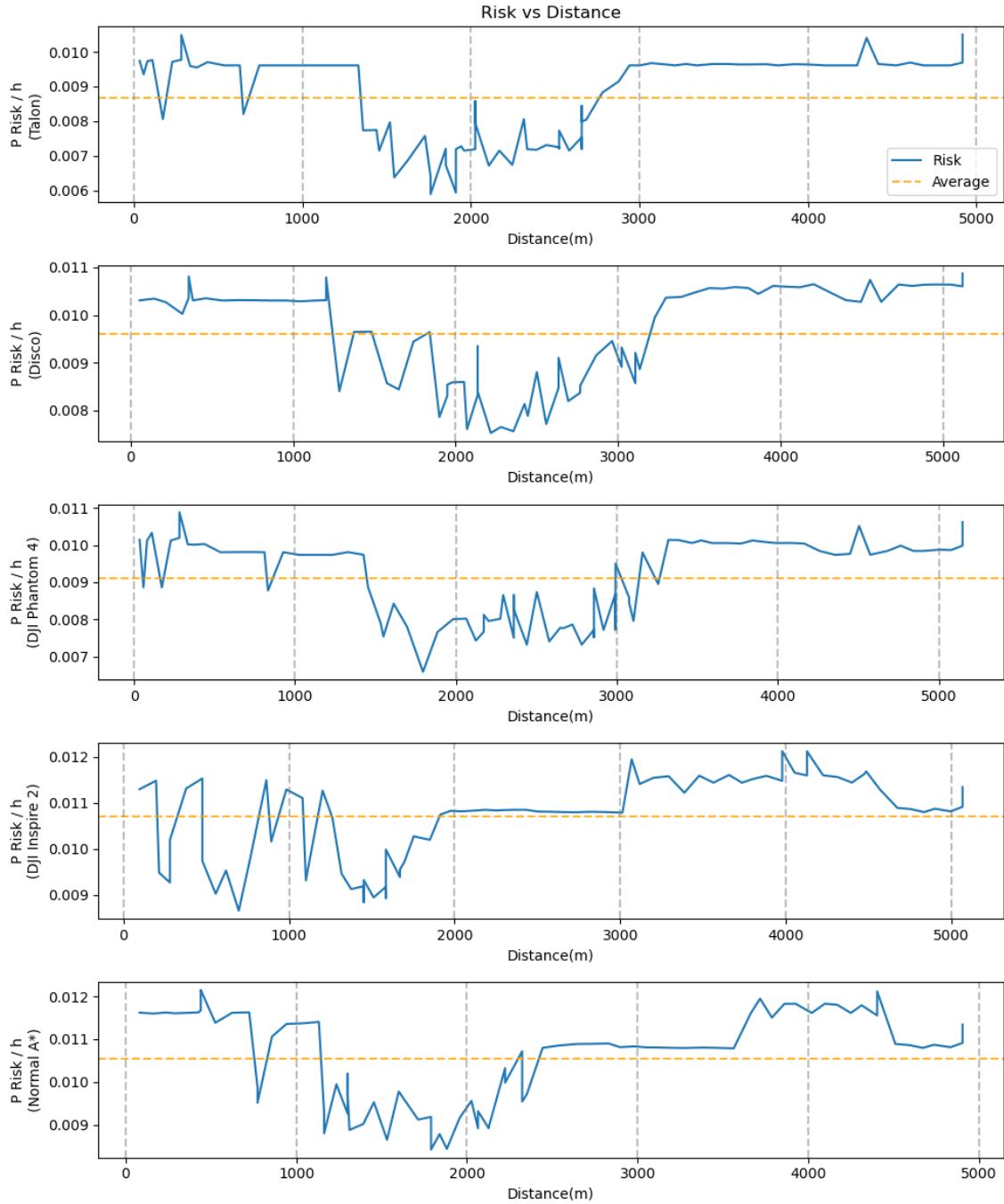


Figure 4.26: Aircraft risk score

Table 4.1: Total Pathway Risk

Parameter	Talon	Disco	Phantom 4	Inspire 2	Normal A*
Total travelled distance (m)	4917.65	5120.85	5066.80	5147.43	4903.0
Total risk involved (%)	0.7277	0.7595	0.8343	0.7654	0.853
Risk vs distance ratio (% fatal/m)	$1.48 \times 10^{-4}$	$1.48 \times 10^{-4}$	$1.65 \times 10^{-4}$	$1.49 \times 10^{-4}$	$1.74 \times 10^{-4}$

This table 4.2 is an expansion of the previous table, providing a detailed explanation of each risk factor associated with the paths taken by various drones. It highlights the specific risks from ballistic, glide, and flyaway incidents, as well as the final risk for each drone. The inclusion of the Normal A\* algorithm showcases a comparative analysis, revealing that while it may incur slightly more risk, it also offers a similar range of risk factors and potentially shorter distances. This analysis helps to better understand the trade-offs between risk and distance for each algorithm.

Table 4.2: Pathway Risk Statistic

Drone	Stat	Ballistic	Glide	Flyaway	Final Risk
<b>Talon</b>	min	$7.55 \times 10^{-2}$	$3.14 \times 10^{-1}$	$0.00 \times 10^0$	$5.36 \times 10^{-3}$
	max	$5.38 \times 10^{-1}$	$5.38 \times 10^{-1}$	$7.76 \times 10^{-3}$	$1.08 \times 10^{-2}$
	avg	$4.74 \times 10^{-1}$	$4.83 \times 10^{-1}$	$8.47 \times 10^{-4}$	$9.57 \times 10^{-3}$
<b>Disco</b>	min	$5.63 \times 10^{-2}$	$4.03 \times 10^{-1}$	$5.38 \times 10^{-5}$	$6.25 \times 10^{-3}$
	max	$5.38 \times 10^{-1}$	$5.38 \times 10^{-1}$	$6.02 \times 10^{-2}$	$1.13 \times 10^{-2}$
	avg	$4.66 \times 10^{-1}$	$5.06 \times 10^{-1}$	$1.93 \times 10^{-2}$	$9.91 \times 10^{-3}$
<b>DJI Phantom 4</b>	min	$8.50 \times 10^{-2}$	$3.99 \times 10^{-1}$	$7.19 \times 10^{-3}$	$6.85 \times 10^{-3}$
	max	$5.39 \times 10^{-1}$	$5.38 \times 10^{-1}$	$1.79 \times 10^{-1}$	$1.23 \times 10^{-2}$
	avg	$4.87 \times 10^{-1}$	$5.05 \times 10^{-1}$	$9.19 \times 10^{-2}$	$1.08 \times 10^{-2}$
<b>DJI Inspire 2</b>	min	$6.03 \times 10^{-5}$	$3.14 \times 10^{-1}$	$0.00 \times 10^0$	$5.50 \times 10^{-3}$
	max	$5.38 \times 10^{-1}$	$5.38 \times 10^{-1}$	$8.72 \times 10^{-2}$	$1.13 \times 10^{-2}$
	avg	$4.63 \times 10^{-1}$	$4.83 \times 10^{-1}$	$3.50 \times 10^{-2}$	$9.81 \times 10^{-3}$
<b>Normal A*</b>	min	$7.57 \times 10^{-2}$	$3.15 \times 10^{-1}$	$1.00 \times 10^{-6}$	$5.37 \times 10^{-3}$
	max	$5.39 \times 10^{-1}$	$5.39 \times 10^{-1}$	$7.78 \times 10^{-3}$	$1.09 \times 10^{-2}$
	avg	$4.75 \times 10^{-1}$	$4.84 \times 10^{-1}$	$8.50 \times 10^{-4}$	$9.58 \times 10^{-3}$

This table 4.3 presents the average risk statistics in North Jakarta. It is important to note that the probabilities for ballistic impact, fly away, parachute, and glide primarily measure the likelihood of falling into a grid. In contrast, the fatality risk encompasses a comprehensive assessment of ground risk mode, providing a more holistic view of the potential hazards in the area.

Table 4.3: Areal map risk statistic of North Jakarta (%)

	Talon	Disco	Phantom 4	Inspire 2
Avg Ballistic impact probability (pct)	$2.28 \times 10^{-3}\%$	$2.25 \times 10^{-3}\%$	$2.35 \times 10^{-3}\%$	$2.24 \times 10^{-3}\%$
Avg Fly away impact probability (pct)	$4.1 \times 10^{-5}\%$	$9.3 \times 10^{-5}\%$	$4.44 \times 10^{-4}\%$	$1.69 \times 10^{-4}\%$
Avg Glide impact probability (pct)	$2.33 \times 10^{-3}\%$	$2.44 \times 10^{-3}\%$	$2.44 \times 10^{-3}\%$	$2.33 \times 10^{-3}\%$
Avg Parachute impact probability (pct)	$2.48 \times 10^{-3}\%$	$2.46 \times 10^{-3}\%$	$2.47 \times 10^{-3}\%$	$2.48 \times 10^{-3}\%$
Avg Total fatality risk probability (pct)	$4.62 \times 10^{-5}\%$	$4.74 \times 10^{-5}\%$	$5.23 \times 10^{-5}\%$	$4.74 \times 10^{-5}\%$

#### 4.2.4 Summarized result

- Talon having  $9.57 \times 10^{-3}$  fatality per flight hour
- Parrot Disco having  $9.91 \times 10^{-3}$  fatality per flight hour
- Phantom 4 having  $1.08 \times 10^{-2}$  fatality per flight hour
- Inspire 2 having  $9.81 \times 10^{-3}$  fatality per flight hour

With a traveled distance of each aircraft being:

- Talon, with 4,918 m
- Parrot Disco, with 5,121 m
- Phantom 4, with 5,066 m
- Inspire 2, with 5,147 m
- Any Aircraft Without Risk involved, with 4,903 m

And average path risk of aircraft:

- Talon, with  $1.48 \times 10^{-4}$  fatality per flight hour
- Parrot Disco, with  $1.48 \times 10^{-4}$  fatality per flight hour
- Phantom 4, with  $1.65 \times 10^{-4}$  fatality per flight hour
- Inspire 2, with  $1.49 \times 10^{-4}$  fatality per flight hour
- Any Aircraft Without Risk involved, with  $1.74 \times 10^{-4}$  fatality per flight hour

### 4.3 Comparison with previous research

When comparing our results to the findings from Shelley's research on Target Level Safety, we find that our average fatality factor aligns closely, maintaining a range of approximately  $1 \times 10^{-7}$  Shelley, 2016. Additionally, the average risk path reported by Primatesta and Courharbo shows values in the range of  $1 \times 10^{-8}$ , which is not surprising given their utilization of significantly shorter paths, leading to the lowest total path risk observed in their study Primatesta, Scanavino, et al., 2020. This illustrates how path optimization can greatly impact the overall risk assessment of UAV operations, highlighting the importance of route planning in achieving safety benchmarks.

# CHAPTER 5

## CONCLUSION, AND RECOMMENDATION

### 5.1 Summary

In the summary ground, a risk map is a discrete PDF in the shape of a grid containing the possible fatality risk of each cell element. It is obtained by multiplying the possible outcome of multiple layers contributing to each risk. These layers are population density, no-fly zone layer, obstacle layer, shelter layer, and descent layer.

- The population layer is obtained by measuring the number of people per area of the cell.
- The no-fly zone layer represents military and aerodrome space and its critical infrastructure that the UAV must not fly into
- The obstacle layer represents a layer where the height of the geographic features becomes an obstacle depending on the flight level.
- The shelter factor represents the capability of the building to absorb the impact energy of a crashing UAV.
- The descent layer is a layer that describes the probability of the drone impacting the area. The descent layer is split into 3 parts that additively contribute to the descent layer. These parts are ballistic descent, glide descent, parachute, and flyaway.

After building the ground risk map. the next step would be finding the optimal Risk A\* flight path where Risk A\* is a modified A\* pathfinding algorithm, which uses risk and distance to measure the most optimal path.

### 5.2 Conclusion

In conclusion, the findings of this study underscore the critical importance of optimizing flight paths in UAV operations to minimize associated risks. The analysis revealed a strong correlation between pathway length and risk factors, emphasizing that shorter, well-planned routes tend to result in lower risk levels. This highlights the need for effective route planning and risk assessment strategies to enhance safety and operational efficiency. As UAV technology continues to advance, ongoing research and development in risk management will be essential to ensure safe and reliable operations in diverse environments.

## 5.3 Recommendation

### 5.3.1 For future research

Create an improvement in the methodology to use a much more elegant solution to smooth out the curve by using the combination of the Dublin curve and spline interpolation instead of using a simple spline method. Create a feedback loop where the risk re-apply itself to the ground risk pdf for each time step to give a much more realistic scenario.

Use a space partitioning algorithm such as a binary tree or RTree for the geographic area. This is to prevent over-boundary risk problems that Talon encountered. This is because the binary tree will expand to every region that overlaps with the furthest flight distance. Not only that, the partitioning algorithm can theoretically increase the performance of calculation since it frees up memory when the partitioned area is not being calculated.

And as for the software design, UPX is currently written purely in Python. This is because it is for proof of concept for the research paper, where the development time is much more critical than the performance. However, if in the future there are similar methods and come across in using UPX. These are some key notes to improving the software.

In order to effectively compute the risk map in great detail for each region, it is highly recommended to utilize distributed systems such as Apache Spark and Dask. These platforms are designed to handle massive amounts of data efficiently by distributing the workload across multiple nodes, allowing for parallel processing and significant reductions in computation time.

**Apache Spark** is a powerful distributed computing framework that can process large datasets quickly through in-memory computing capabilities. It supports a wide range of operations, including SQL queries, streaming data, and complex analytics. By leveraging Spark's robust ecosystem, we can perform extensive risk calculations and data analyses across large geographic areas without the bottlenecks associated with traditional single-machine processing, The Apache Software Foundation, 2024.

**Dask** is another excellent option for distributed computing, particularly for Python-based workflows. It enables parallel computing by extending Python's native data structures, such as NumPy arrays and Pandas DataFrames, to operate in a distributed manner. Dask is highly flexible and can seamlessly integrate with existing Python code, making it an ideal choice for scaling up data processing tasks and risk map computations, Dask Development Team, 2016.

Implementing these distributed systems offers several advantages:

- **Scalability:** Both Spark and Dask can scale horizontally by adding more nodes to the cluster, accommodating growing data volumes and computational demands.

- **Efficiency:** Distributed computing significantly reduces the time required to process large datasets, enabling faster and more detailed risk assessments.
- **Flexibility:** These platforms support a wide range of data processing and analysis tasks, making them versatile tools for comprehensive risk mapping and other analytical needs.
- **Resilience:** Distributed systems are designed to handle node failures gracefully, ensuring reliable and continuous operation even in the face of hardware or software issues.

By integrating Spark or Dask into the software design for risk map computation, we can achieve a higher level of detail and accuracy in risk assessments, ultimately leading to more informed decision-making and better management of potential hazards.

## BIBLIOGRAPHY

- Aalmoes, R., Cheung, Y., Sunil, E., Hoekstra, J., & Bussink, F. (2015). A conceptual third party risk model for personal and unmanned aerial vehicles. *2015 International Conference on Unmanned Aircraft Systems (ICUAS)*, 1301–1309. <https://doi.org/10.1109/ICUAS.2015.7152424>
- Ale, B., & Piers, M. (2000). The assessment and management of third party risk around a major airport. *Journal of Hazardous Materials*, 71(1), 1–16. [https://doi.org/10.1016/S0304-3894\(99\)00069-2](https://doi.org/10.1016/S0304-3894(99)00069-2)
- Ancel, E., Capristan, F. M., Foster, J. V., & Condotta, R. C. (2017). Real-time risk assessment framework for unmanned aircraft system (UAS) traffic management (UTM). *17th AIAA Aviation Technology, Integration, and Operations Conference*. <https://doi.org/10.2514/6.2017-3273>
- ARC. (2015). Unmanned aircraft systems (UAS) registration task force (RTF) aviation rulemaking committee (ARC).
- Austin, J., & Brewster, S. (2024). *The 4 best drones for photos and video of 2024 — reviews by wirecutter* [The best drones for photos and video]. Retrieved July 12, 2024, from <https://www.nytimes.com/wirecutter/reviews/best-drones/>
- Ball, J. A., Knott, M., & Burke, D. D. (2012). CRASH LETHALITY MODEL.
- Barr, L. C., Newman, R., Ancel, E., Belcastro, C. M., Foster, J. V., Evans, J., & Klyde, D. H. (2017). Preliminary risk assessment for small unmanned aircraft systems. *17th AIAA Aviation Technology, Integration, and Operations Conference*. <https://doi.org/10.2514/6.2017-3272>
- Bleier, M., Settele, F., Krauss, M., Knoll, A., & Schilling, K. (2015). Risk assessment of flight paths for automatic emergency parachute deployment in UAVs. *IFAC-PapersOnLine*, 48(9), 180–185. <https://doi.org/10.1016/j.ifacol.2015.08.080>
- Breunig, J., Forman, J., Sayed, S., & Audenaerd, L. (2018). Modeling risk-based approach for small unmanned aircraft systems.
- Burke, D. A., Hall, C. E., & Cook, S. P. (2011). System-level airworthiness tool. *Journal of Aircraft*, 48(3), 777–785. <https://doi.org/10.2514/1.C031022>
- Clothier, R., Walker, R., Fulton, N., & Campbell, D. (2007). A CASUALTY RISK ANALYSIS FOR UNMANNED AERIAL SYSTEM (UAS) OPERATIONS OVER INHABITED AREAS.
- Cole, J. K., Young, L. W., & Jordan-Culler, T. (1997, April 1). *Hazards of falling debris to people, aircraft, and watercraft* (SAND-97-0805). Sandia National Lab. (SNL-NM), Albuquerque, NM (United States). <https://doi.org/10.2172/468556>
- Commission, E., & Centre, J. R. (2023). *GHSL data package 2023*. Publications Office of the European Union. <https://doi.org/doi/10.2760/098587>
- Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2022). *Introduction to algorithms*. MIT press.
- Cour-Harbo, A. I. (2020). Ground impact probability distribution for small unmanned aircraft in ballistic descent. *2020 International Conference on Un-*

- manned Aircraft Systems (ICUAS)*, 1442–1451. <https://doi.org/10.1109/ICUAS48674.2020.9213990>
- Dalamagkidis, K., Valavanis, K. P., & Piegl, L. A. (2008). Evaluating the risk of unmanned aircraft ground impacts. *2008 16th Mediterranean Conference on Control and Automation*, 709–716. <https://doi.org/10.1109/MED.2008.4602249>
- Dask Development Team. (2016). *Dask: Library for dynamic task scheduling*. <http://dask.pydata.org>
- Dijkstra, E. W. (2022). A note on two problems in connexion with graphs. In *Edsger Wybe Dijkstra: His life, work, and legacy* (pp. 287–290).
- DJI. (2023). *DJI agriculture - drones better growth, better life*. Retrieved July 12, 2024, from <https://ag.dji.com/>
- DJI. (2024). *Consumer drones comparison - DJI* [Consumer drones comparison]. Retrieved July 12, 2024, from <https://www.dji.com/id/products/comparison-consumer-drones>
- Dorigo, M. (2007). Ant colony optimization. *Scholarpedia*, 2(3), 1461.
- FAA. (2000, August 20). *EXPECTED CASUALTY CALCULATIONS FOR COMMERCIAL SPACE LAUNCH AND REENTRY MISSIONS*. U.S Department of Transportation. [https://www.faa.gov/about/office\\_org/headquarters-offices/ast/licenses\\_permits/media/Ac4311fn.pdf](https://www.faa.gov/about/office_org/headquarters-offices/ast/licenses_permits/media/Ac4311fn.pdf)
- FAA. (2016). eCFR :: 14 CFR part 107 – small unmanned aircraft systems (FAR part 107). Retrieved July 12, 2024, from <https://www.ecfr.gov/current/title-14/chapter-I/subchapter-F/part-107>
- Gonzales, E., & Murray, D. (2017). *FAA's approaches to ground and NAS separation distances for commercial rocket launches* [Aerospace sciences meetings] [Archive Location: world]. <https://doi.org/10.2514/6.2010-1540>
- Harwick, M. J., Hall, J., Tatom, J. W., & Baker, R. G. (2007, February 2). *Approved methods and algorithms for DoD risk-based explosives siting*: Defense Technical Information Center. Fort Belvoir, VA. <https://doi.org/10.21236/ADA463571>
- Holland, J. (2015). *How drones are affecting wildlife in surprising ways*. Retrieved July 12, 2024, from <https://www.nationalgeographic.com/animals/article/150825-drones-animals-wildlife-bears-science-technology>
- Kim, Y., & Bae, J. (2022). Risk-based UAV corridor capacity analysis above a populated area. *Drones*, 6(9), 221. <https://doi.org/10.3390/drones6090221>
- Klepeis, N. E., Nelson, W. C., Ott, W. R., Robinson, J. P., Tsang, A. M., Switzer, P., Behar, J. V., Hern, S. C., & Engelmann, W. H. (2001). The national human activity pattern survey (NHAPS): A resource for assessing exposure to environmental pollutants. *Journal of Exposure Science & Environmental Epidemiology*, 11(3), 231–252. <https://doi.org/10.1038/sj.jea.7500165>
- Knuth, D. E. (1997). *The art of computer programming, volume 1 fundamental algorithms*. Addison Wesley Longman Publishing Co., Inc.
- la Cour-Harbo, A. (2019). Quantifying risk of ground impact fatalities for small unmanned aircraft. *Journal of Intelligent & Robotic Systems*, 93(1), 367–384. <https://doi.org/10.1007/s10846-018-0853-1>
- LaValle, S. M. (2006). *Planning algorithms*. Cambridge university press.

- Magister, T. (2010). The small unmanned aircraft blunt criterion based injury potential estimation. *Safety Science*, 48(10), 1313–1320. <https://doi.org/10.1016/j.ssci.2010.04.012>
- Melnyk, R., Schrage, D., Volovoi, V., & Jimenez, H. (2014). A third-party casualty risk model for unmanned aircraft system operations. *Reliability Engineering & System Safety*, 124, 105–116. <https://doi.org/10.1016/j.ress.2013.11.016>
- Nemire, B. (2015). *DJI launches GPU-based high performance embedded computer for drones — NVIDIA technical blog*. Retrieved July 20, 2024, from <https://developer.nvidia.com/blog/dji-launches-gpu-based-high-performance-embedded-computer-for-drones-2/>
- OpenStreet. (2022). *Overpass API - OpenStreetMap wiki*. Retrieved September 21, 2022, from [https://wiki.openstreetmap.org/wiki/Overpass\\_API](https://wiki.openstreetmap.org/wiki/Overpass_API)
- OpenStreetMap contributors. (2017). Planet dump retrieved from <https://planet.osm.org> [Published: <https://www.openstreetmap.org>].
- Pat-Cornell, M. E. (1996). Uncertainties in risk analysis: Six levels of treatment.
- Primatesta, S., Guglieri, G., & Rizzo, A. (2019). A risk-aware path planning strategy for UAVs in urban environments. *Journal of Intelligent & Robotic Systems*, 95(2), 629–643. <https://doi.org/10.1007/s10846-018-0924-3>
- Primatesta, S., Rizzo, A., & la Cour-Harbo, A. (2020). Ground risk map for unmanned aircraft in urban environments. *Journal of Intelligent & Robotic Systems*, 97(3), 489–509. <https://doi.org/10.1007/s10846-019-01015-z>
- Primatesta, S., Scanavino, M., Guglieri, G., & Rizzo, A. (2020). A risk-based path planning strategy to compute optimum risk path for unmanned aircraft systems over populated areas. *2020 International Conference on Unmanned Aircraft Systems (ICUAS)*, 641–650. <https://doi.org/10.1109/ICUAS48674.2020.9213982>
- Russell, S. J., & Norvig, P. (2016). *Artificial intelligence: A modern approach*. Pearson.
- Russell, S. J., Norvig, P., & Davis, E. (2010). *Artificial intelligence: A modern approach* (3rd ed). Prentice Hall.
- Shelley, A. (2016). A model of human harm from a falling unmanned aircraft: Implications for UAS regulation. *International Journal of Aviation, Aeronautics, and Aerospace*. <https://doi.org/10.15394/ijaaa.2016.1120>
- Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., et al. (2016). Mastering the game of go with deep neural networks and tree search [Publisher: Nature Publishing Group]. *nature*, 529(7587), 484–489.
- Small, U. A. S. (2009). Aviation rulemaking committee (ARC). *Comprehensive Set of Recommendations for sUAS Regulatory Development*, 1.
- The Apache Software Foundation. (2024). *SparkR: R front end for 'apache spark'*. <https://www.apache.org%20https://spark.apache.org>
- Washington, A., Clothier, R. A., & Silva, J. (2017). A review of unmanned aircraft system ground risk models. *Progress in Aerospace Sciences*, 95, 24–44. <https://doi.org/10.1016/j.paerosci.2017.10.001>
- Weibel, R. E., & Hansman, R. J. (2012). Safety considerations for operation of different classes of UAVs in the NAS.