NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

A PROBABILISTIC MODEL OF ILLEGAL DRUG TRAFFICKING OPERATIONS IN THE EASTERN PACIFIC AND CARIBBEAN SEA

by

Matthew S. Mooshegian

September 2013

Thesis Advisor:	Michael Atkinson
Second Reader:	Johannes O. Royset

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704–0188

1. REPORT DATE *(DD–MM–YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From — To)*
11–9–2013	Master's Thesis	2011-06-24—2013-09-27

4. TITLE AND SUBTITLE

A PROBABILISTIC MODEL OF ILLEGAL DRUG TRAFFICKING OPERATIONS IN THE EASTERN PACIFIC AND CARIBBEAN SEA

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Matthew S. Mooshegian

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Postgraduate School
Monterey, CA 93943

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Department of the Navy

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited

13. SUPPLEMENTARY NOTES

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

14. ABSTRACT

Illicit drug-trafficking is a major concern of the United States and is a primary pillar of President Barack Obama's Strategy to Combat Transnational Organized Crime. In the eastern Pacific and Caribbean Sea, drug-trafficking organizations operate a variety of vessels to transit drugs from South America to the United States. Joint Interagency Task Force (JIATF) South, in cooperation with partner agencies and nations, detects, tracks, and interdicts illegal drug-trafficking in this region. In this thesis, we develop a probability model based on intelligence inputs to generate a spatial temporal heat map specifying the likely location of targets over time. We also formulate a path-finding model that takes the heat map as input and determines route characteristics through transit, such as departure times, waypoints, and speed. We link the results of our models to a separate effort that seeks to provide JIATF South with an optimal search plan to maximize the expected amount of drugs seized. We show that our path-finding model accurately reproduces the target's track even when the target transits along a complex route. Furthermore, we show that the optimal search plan based on our path-finding model is nearly identical to the search plan using the known parameters.

15. SUBJECT TERMS

Multivariate Adaptive Regression Splines, MARS, linear regression, heat map, Path-finding model, Probability model, Drug-trafficking, Route estimation

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
Unclassified	Unclassified	Unclassified	UU	91	

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

A PROBABILISTIC MODEL OF ILLEGAL DRUG TRAFFICKING OPERATIONS IN THE EASTERN PACIFIC AND CARIBBEAN SEA

Matthew S. Mooshegian
Lieutenant Commander, United States Navy
B.S., U.S. Naval Academy, 2002
M.A., Webster University, 2011

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 2013

Author: Matthew S. Mooshegian

Approved by: Michael Atkinson
Thesis Advisor

Johannes O. Royset
Second Reader

Robert Dell
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Illicit drug-trafficking is a major concern of the United States and is a primary pillar of President Barack Obama's Strategy to Combat Transnational Organized Crime. In the eastern Pacific and Caribbean Sea, drug-trafficking organizations operate a variety of vessels to transit drugs from South America to the United States. Joint Interagency Task Force (JIATF) South, in cooperation with partner agencies and nations, detects, tracks, and interdicts illegal drug-trafficking in this region. In this thesis, we develop a probability model based on intelligence inputs to generate a spatial temporal heat map specifying the likely location of targets over time. We also formulate a path-finding model that takes the heat map as input and determines route characteristics through transit, such as departure times, waypoints, and speed. We link the results of our models to a separate effort that seeks to provide JIATF South with an optimal search plan to maximize the expected amount of drugs seized. We show that our path-finding model accurately reproduces the target's track even when the target transits along a complex route. Furthermore, we show that the optimal search plan based on our path-finding model is nearly identical to the search plan using the known parameters.

THIS PAGE INTENTIONALLY LEFT BLANK

Table of Contents

List of Figures

THIS PAGE INTENTIONALLY LEFT BLANK

List of Tables

THIS PAGE INTENTIONALLY LEFT BLANK

List of Acronyms and Abbreviations

ACI	Andean Counterdrug Initiative
AOR	Area of Responsibility
AOU	Area of Uncertainty
AUC	United Defense Forces of Colombia
CARSI	Central America Regional Security Initiative
CBSI	Caribbean Basin Security Initiative
CCDB	Consolidated Counterdrug Database
DoD	Department of Defense
DTO	Drug Trafficking Organization
ELN	National Liberation Army
FARC	Revolutionary Armed Forces of Colombia
FTO	Foreign Terrorist Organization
JIATF	Joint Interagency Task Force
MARS	Multivariate Adaptive Regression Splines
NORTHCOM	U.S. Northern Command
NRL	Naval Research Laboratory
OEG	Operations Evaluation Group
OCDETF	Organized Crime Drug Enforcement Task Forces
SOUTHCOM	U.S. Southern Command
SPSS	Self-propelled Semi-submersible
TOC	Transnational Organized Crime

THIS PAGE INTENTIONALLY LEFT BLANK

Executive Summary

Illicit drug-trafficking originating from South America is a major concern of the United States, and continued expansion of drug-trafficking significantly affects the national security interests of the United States. Organizations tasked with combatting drug traffickers are faced with several problems. They must detect, track, and interdict illegal drug traffickers en route to various follow-on transit zones in Central America, Mexico, and the Caribbean. They utilize incoming intelligence from various sources to determine where and what assets to employ. Their efforts, while productive, are not able to stop a large majority of the drugs from reaching the United States. They battle Mexican and Colombian Drug Trafficking Organizations (DTOs) spread throughout Latin America. The DTOs have extensive capabilities and access to large amounts of money. Combined with the close proximity to the United States, the DTOs pose an extreme threat to the nation's security.

Operators at Joint Interagency Task Force (JIATF) South obtain various types of intelligence information and utilize that information when creating search and interdiction plans. They may have at their disposal information concerning departure times, departure and arrival locations, type of vessel, speed, drug load, and waypoints. At all times, there is a level of uncertainty associated with the intelligence, which can be represented by a probability distribution for each of the aforementioned pieces of information. Given this information and surrounding uncertainty, JIATF South operators can benefit from a visual heat map displaying the likely route of a target or targets over time. In this thesis, we address the problem of creating a heat map of likely drug trafficker locations based on intelligence inputs.

One approach to generating a heat map is through simulation. Simulations have the benefit of being able to account for many real world complexities like weather. However, they also may be computationally expensive and difficult to validate and verify. We desire to complement and vet such complicated simulation by developing more analytically tractable models. We develop probability models to generate a heat map to provide operators with better situational awareness about the likely locations of targets, both in the current period and future periods. We first build a simple probability model that accounts for uncertainty only in the departure and arrival locations. Using this model as a baseline, we further enhance its functionality to account for uncertain speed and time of departure, as well as consideration for multiple legs, multiple targets in the same region, and the ability to output expected amount of drugs. Given the probability distributions for the intelligence inputs (e.g., speeds, locations, and times), the

model produces the spatial temporal probability map of the trafficker's path. Using the heat map as input, we then address the problem of determining a drug trafficker's transit route based on the collected intelligence information.

After developing a model to transform intelligence inputs into a heat map, we next turn to the inverse problem: Given a heat map, what are the most likely intelligence inputs used to create it. We adopt this approach to compare the path-finding model's results with some known inputs via the search vehicle optimization model. Additionally, it is possible that the only data available for JIATF South planners will be a heat map, thus providing a means to estimate intelligence parameters is beneficial for the planners.

We produce the necessary intelligence parameters through the use of various regression techniques and consideration for the input requirements of the optimization model. The path-finding model approximates the path taken by the target as a sequence of piecewise linear legs, where the collection of all legs produces the target's entire track from departure location to arrival location. Standard regression techniques are incapable of automatically identifying where a new leg begins, which occurs when speed or heading changes. As a result, we employ more advanced regression techniques, namely Multivariate Adaptive Regression Splines (MARS). In accordance with the search vehicle optimization model, we define a leg as a segment of the target's track that has a constant heading and speed. This creates a piecewise linear path for the target. Specifically, we determine the target's departure time range, expected departure time, arrival time range, expected arrival time, speed, departure and arrival locations, and waypoint locations.

The final aspect of the path-finding model is specifying the Area of Uncertainty (AOU) of a target as it traverses along a path leg. Roughly speaking, the AOU is the size of the target's "blob" on the heat map. The AOU along a target's track represents the uncertainty in both departure speed and departure time. With these uncertainties, an area exists in which it is possible for the target to be present at any point in the associated area. The search plan algorithm requires that the AOU be constant on each leg of the path. Thus, if the AOU changes significantly, we break one leg of the path into multiple legs even if the speed and heading do not change. We develop some heuristics to accomplish this.

We develop two case studies and the corresponding intelligence information from which we generate a heat map according to the probability model. Next, we use our path-finding algorithm to estimate the input parameters and compare how well they match to the true values.

We show the capability to identify speed changes, changes in heading or course, and changes in the AOU surrounding the target throughout the path. We illustrate the models' functionality with two case studies of varying complexity. First, a simple straight path between the origin and destination shows the path-finding model's ability to identify the target's starting location, destination, speed, and associated travel times. Combining these route characteristics, we are able to reproduce the target's likely track from start to finish. Supplying our path-finding results into the search vehicle allocation model shows similar search plan results when compared with the original case parameters used in the probability model. Next, we demonstrate a more complex scenario with a target traveling to the destination with a course change during the journey. Our model accurately identifies this course change and reproduces the target's track. These two case studies are representative of actual drug trafficker routes but additional analysis is needed to validate the model's ability. Finally, we provide the search vehicle allocation model with the path-finding model results for a scenario with two targets. The resulting search plan is nearly identical to the one generated when the true intelligence inputs are known. This provides strong evidence that we have provided a key link between the probability, path-finding, and optimization models.

THIS PAGE INTENTIONALLY LEFT BLANK

Acknowledgements

I would like to express my sincerest thanks to my thesis advisor, Dr. Michael Atkinson. Your patience, encouragement, and understanding guided me successfully through this process. Your time is truly valuable to me and you have taught me so much more than I ever imagined as I began this journey. Thank you.

To my second reader, Dr. Johannes Royset, you introduced me to this interesting topic and allowed me to freely explore the possibilities. I appreciate your advice and input during this process. It has been a pleasure working with you.

To my wife, Breanna, sons, Mason, Aiden, and Logan, your support and sacrifice enable me to pursue these opportunities. I am forever grateful for your endless love and care.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 1:
Introduction

Illicit drug-trafficking is a major concern of the United States and is a primary pillar of President Barack Obama's Strategy to Combat Transnational Organized Crime. Transnational Organized Crime (TOC) encompasses many areas of global crime to include human smuggling, trafficking in persons, weapons trafficking, intellectual property theft, cybercrime, and drug-trafficking (National Security Staff, 2011). Drug-trafficking has expanded in recent years, and organizations have become increasingly more violent and less centralized. The U.S. government lists Drug Trafficking Organizations (DTOs) as the "greatest organized crime threat" to the United States (National Drug Intelligence Council [NDIC], 2008). Murder rates in known drug-trafficking countries have increased, and continued expansion of drug-trafficking significantly affects the national security interests of the United States. This thesis contributes to the ongoing research to improve the effectiveness of the agencies tasked with combatting illicit drug-trafficking. Based on incoming intelligence, the analysis models movements of drug traffickers.

1.1 Background

Latin America is a prominent player in the global drug market, with involvement in nearly all areas of the illicit drug supply chain. Activities in the region include drug crop farming, drug production, drug-trafficking, and drug consumption. The South American countries in the Andean region, particularly Colombia, Peru, and Bolivia, are the sole producers of cocaine for the global market (United States Department of State, 2012). United States Department of State (2012) estimates total cocoa production in these three countries at 187,500 hectares in 2010. This is down from a peak in 2007 of 232,500 hectares. Combining the Andean region's ability to produce cocaine and their close proximity to the United States, makes Latin America the central hub of the illicit drug market (Seelke, Wyler, Beittel, & Sullivan, 2011). Additional factors that contribute to the region's role in the drug market include poverty, inequality, the absence of other fiscal opportunities, and a weak governing force (Seelke et al., 2011). These factors have left DTOs nearly free to act as they wish in the supply of illicit drugs.

A variety of DTOs operate in Latin America, and they range from small family-based operations to global networks that are well connected throughout the world to include ties with terrorist

organizations. The more advanced DTOs employ considerable paramilitary and counterintelligence assets to contend with legitimate police forces and conduct operations with minimal risk via corrupt governments (Seelke et al., 2011). The family-based operations typically specialize in a few areas of the drug supply chain.

Mexican and Colombian DTOs are the most prominent in the region. They currently have operations worldwide to include Europe, West Africa, and the Asia-Pacific regions as well as connections to organized crime and terrorists (Seelke et al., 2011). The NDIC (2008) reports that Mexican and Colombaian DTOs generate, remove, and launder between $18 billion and $39 billion in wholesale drug proceeds annually. Currently, Mexican DTOs are the dominant groups supplying most illicit drugs to the United States. Seven Mexican DTOs are currently operating in all nine of the Organized Crime Drug Enforcement Task Forces (OCDETF) regions. Table 1.1 and Figure 1.1 show the Mexican DTOs and their primary drugs and regions of operations.

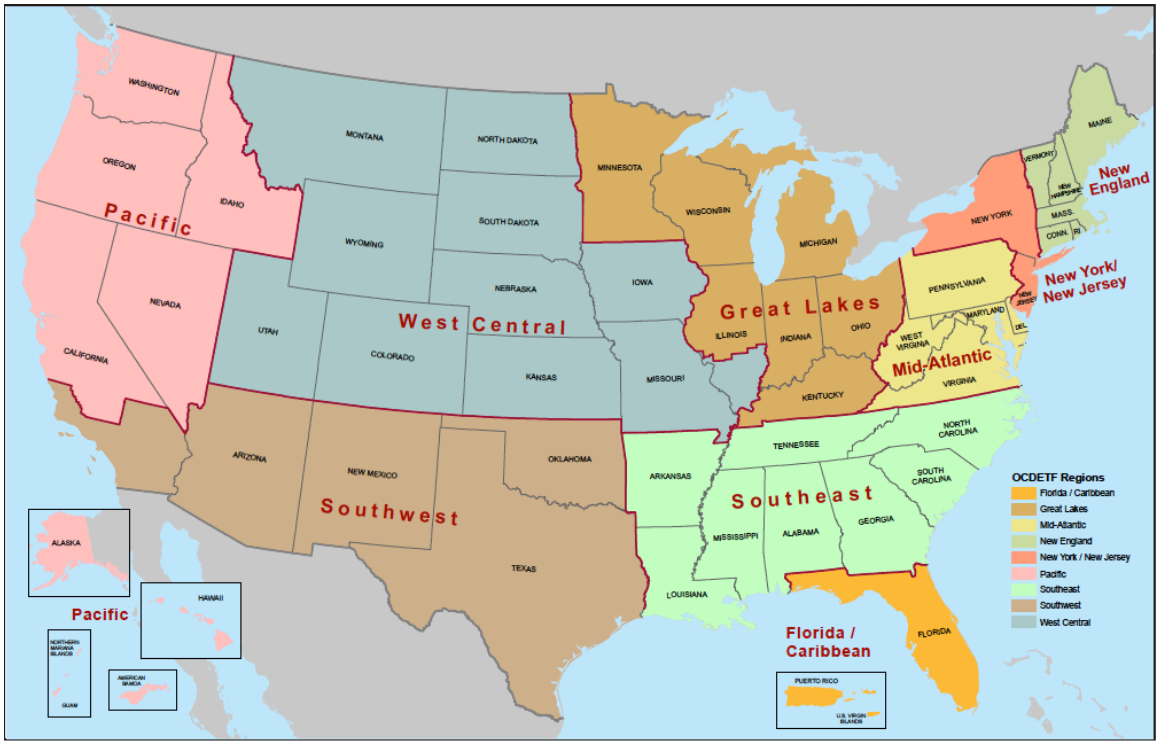


Figure 1.1: Nine regions of the Organized Crime Drug Enforcement Task Force. All seven Mexican-based DTOs operate in all of the OCDETF regions. (From NDIC, 2011)

Over the last ten years, the Mexican DTOs have replaced many of the Colombian DTOs that controlled the market in the 1980s and 1990s. The Mexican organizations have a geographical advantage over the Colombian drug traffickers. They have access and control of the smuggling

DTO	Primary Drugs	Primary Regions
Sinaloa Cartel	Cocaine Heroin Marijuana MDMA Methamphetamine	Florida/Caribbean Great Lakes Mid–Atlantic New England New York/New Jersey Pacific Southeast Southwest West Central
Los Zetas	Cocain Marijuana	Florida/Caribbean Great Lakes Southeast Southwest
Gulf Cartel	Cocaine Marijuana	Florida/Caribbean Mid–Atlantic New England New York/New Jersey Southeast Southwest
Juárez Cartel	Cocaine Marijuana	Great Lakes New York/New Jersey Pacific Southeast Southwest West Central
BLO	Cocaine Heroin Marijuana	Southeast Southwest
LFM	Cocaine Heroin Marijuana Methamphetamine	Southeast Southwest
Tijuana Cartel	Cocaine Heroin Marijuana Methamphetamine	Great Lakes Pacific Southwest

Table 1.1: Mexican-based DTOs operate in all nine of the OCDETF regions and supply most illicit drugs to the United States. (From NDIC, 2011)

routes along the U.S. southwest border, and these routes or plazas are fiercely contested among the Mexican DTOs (NDIC, 2011). Additionally, the ability to obtain, transport, and distribute the illicit drugs in demand in the United States allows them to dominate the market. The Sinaloa Cartel, with operations in all nine OCDETF regions, is the principal Mexican DTO. It boasts an extensive transportation and distribution network, a vast network of personnel to aid in trafficking operations, and control of heroin, marijuana, and methamphetamine production (NDIC, 2011). According to NDIC (2011), the Mexican DTOs will remain the dominant U.S. illicit drug suppliers for the near future.

The Colombian DTOs, the Cali cartel and Medellín cartel that dominated the trade in the 1980s and 1990s, have been replaced by Norte del Valle DTO. Rather than being one central cartel such as their predecessors, the Norte del Valle consists of many smaller cells and is less hierarchical and more decentralized (Seelke et al., 2011). State Department designated Foreign Terrorist Organizations (FTOs), the Revolutionary Armed Forces of Colombia (FARC), the National Liberation Army (ELN), and the United Defense Forces of Colombia (AUC), control and finance Colombian illicit drug-trafficking operations. The FTOs then use the drug money to attack Colombian police and military (United States Department of State, 2012). These organizations are becoming more involved in drug-trafficking operations. They tax cocoa farmers and buyers and have controlling interests in various operations of the drug supply chain (Seelke et al., 2011). When the AUC demobilized, smaller paramilitary groups formed and increased their involvement in the drug market, proving to be a major challenge for law enforcement. Their increased presence has caused a rise in violence in both Colombian cities and rural villages (United States Department of State, 2012).

Many of the drugs, specifically cocaine, that reach U.S. soil transit from South America through the Central America–Mexico transit routes or via the Caribbean transit routes en route to the United States. The Consolidated Counterdrug Database (CCDB) documented 1,494, 1,228, and 804 metric tons of cocaine movement through the Western Hemisphere Transit Zone in fiscal years 2008, 2009, and 2010, respectively (Office of National Drug Control Policy, 2013). The primary transportation route is through the Central America–Mexico corridor to include territorial waters in both the eastern Pacific and the Caribbean Sea. This drug flow accounts for 95% of the cocaine going to the United States. More than half of these drugs transit through Central America first (Seelke et al., 2011). This is a shift from the operations of the 1980s and early 1990s, when drugs were mainly smuggled through the Caribbean to Florida (Seelke et al., 2011). In 2010, 5% of the cocaine transiting to the United States was via the Caribbean,

whereas in 2001, 26% was transiting via the Caribbean (United States Senate Caucus on International Narcotics Control, 2012). In September 2011, President Obama identified 22 countries as major drug transit or major illicit drug producing countries. Of the 22 identified, 19 of the countries are in Latin America or the Caribbean (United States Department of State, 2012). In the Caribbean region specifically, The Bahamas, the Dominican Republic, Haiti, and Jamaica are major countries used in the illegal trafficking of cocaine (United States Senate Caucus on International Narcotics Control, 2012). Approximately one half of the cocaine that flows through the Caribbean transits through the Dominican Republic (United States Senate Caucus on International Narcotics Control, 2012). Puerto Rico presents another major transit hub. Since the country is within the U.S. customs zone, regulations to get into the United States are less stringent, thus 70% to 80% of the cocaine moving through Puerto Rico is bound for the United States while the remaining cocaine is consumed within Puerto Rico (United States Senate Caucus on International Narcotics Control, 2012). The Caribbean presents three main corridors in Figure 1.2 for the illicit transport of drugs.

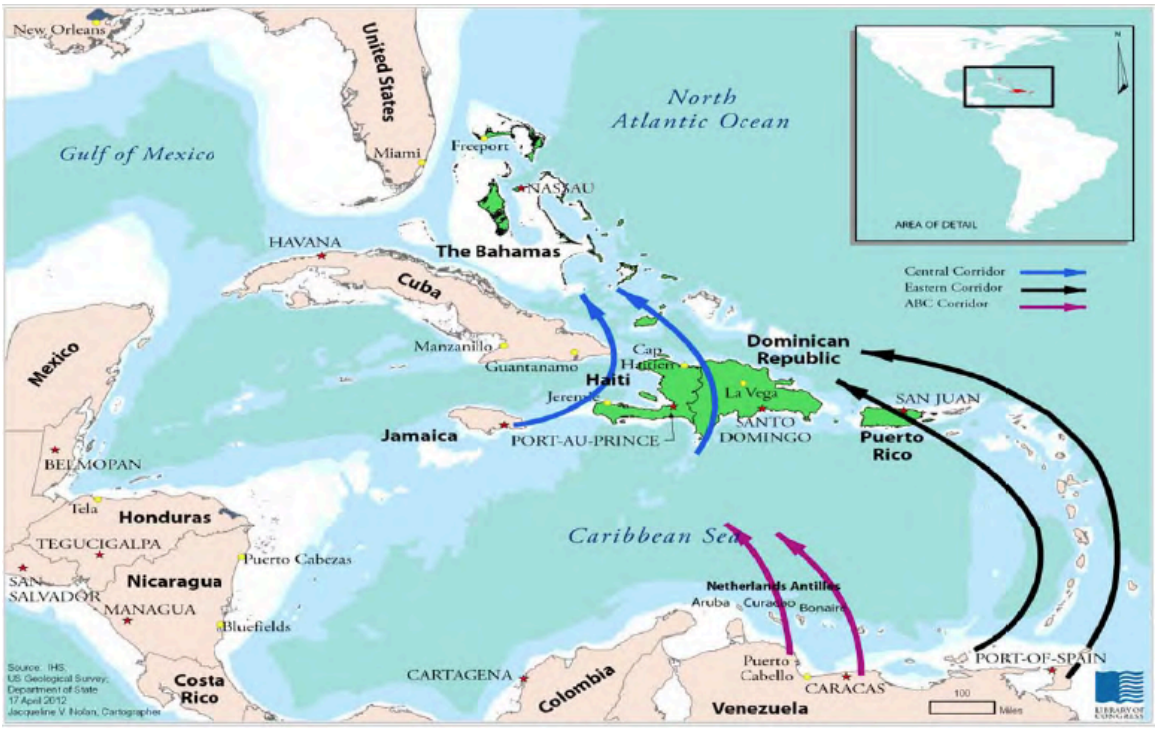


Figure 1.2: Illicit Caribbean drug routes are divided into three main corridors: Central, Eastern, and ABC. (From United States Senate Caucus on International Narcotics Control, 2012)

The central corridor is a primary avenue for moving cocaine from South America through Jamaica, Haiti, the Dominican Republic, or The Bahamas into the United States. The eastern

corridor moves heroin and cocaine through eastern Caribbean nations, Trinidad, and Tobago to Puerto Rico and ultimately into the United States. The ABC corridor of Aruba, Bonaire, and Curacao is prevalent among cocaine traffickers as well (United States Senate Caucus on International Narcotics Control, 2012).

In 2011, 80% of drugs destined for the United States moved via the waterways and the remaining 20% via aircraft (United States Senate Caucus on International Narcotics Control, 2012). Maritime and air transit routes from Joint Interagency Task Force (JIATF) South are shown in Figures 1.3 and 1.4. DTOs using maritime Central America–Mexico transit routes utilize both Caribbean and eastern Pacific routes as evident in Figure 1.3. The air routes leaving the Colombia–Venezuela border avoid Colombian air space in an effort to evade surveillance by the U.S. Government and Colombian Government as part of the Air Bridge Denial Agreement. Colombian officials cite that the surveillance and interdiction efforts force smugglers to remain clear of Colombian air space (GAO, 2009).

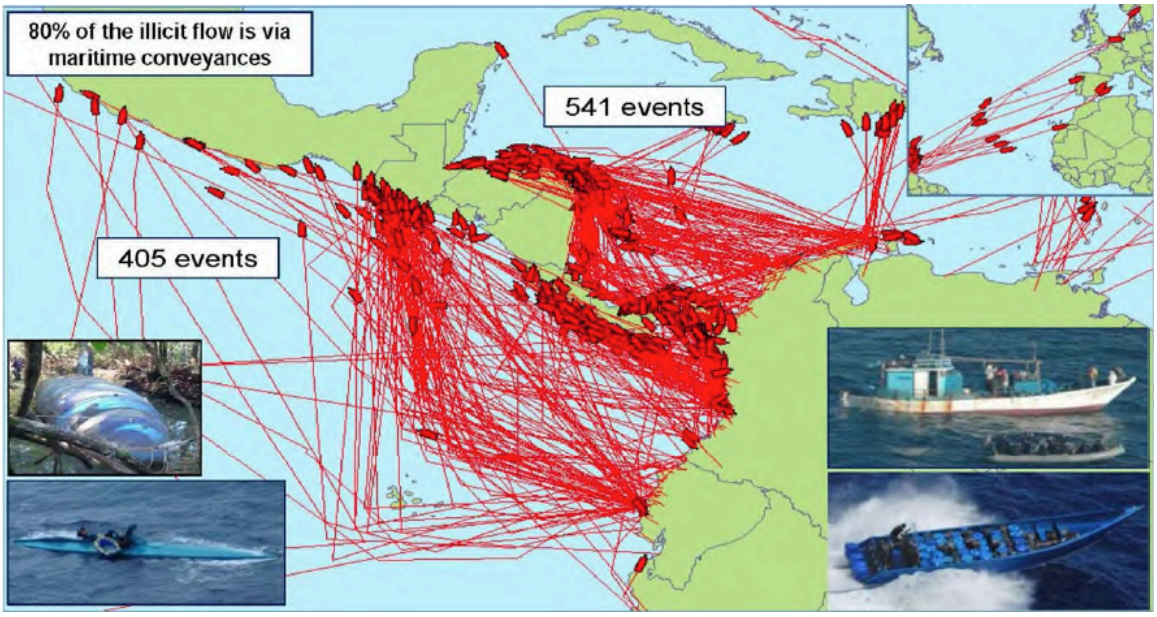


Figure 1.3: Maritime drug transit routes in 2011. 80% of the drugs destined for the United States transit via waterways. (From United States Senate Caucus on International Narcotics Control, 2012)

The primary means of transport is via maritime vessels, which include go-fast boats, pleasure craft, fishing vessels, and Self-propelled Semi-submersible (SPSS). The Caribbean has seen a significant increase in the number of go–fast boats used to transport illicit drugs. These vessels travel at high speeds, are small, display nearly no radar cross–section, and are hard to see

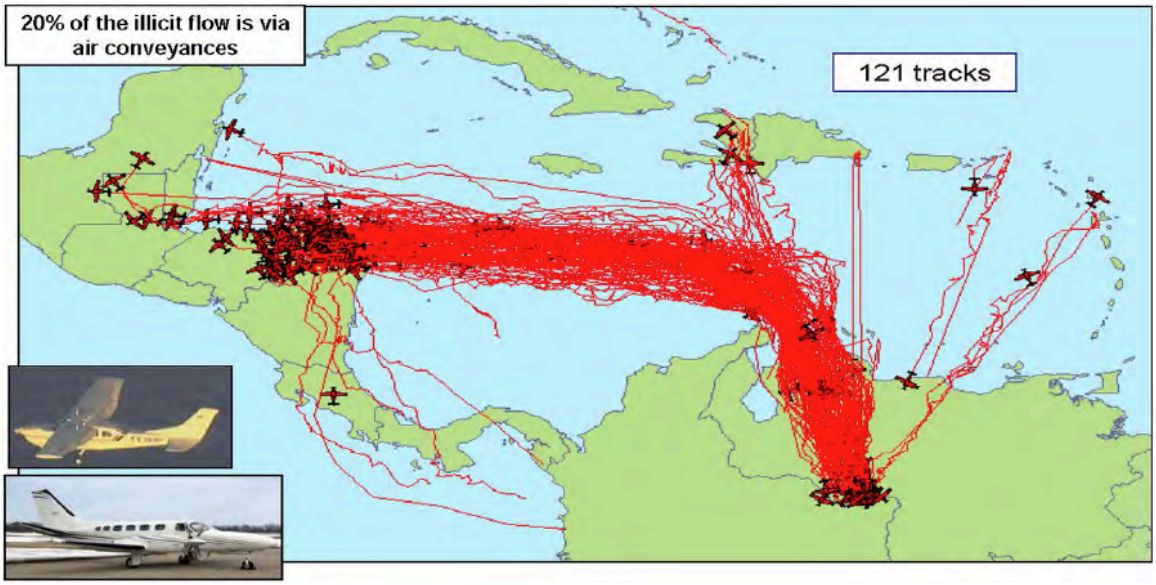


Figure 1.4: Air drug transit routes in 2011. (From United States Senate Caucus on International Narcotics Control, 2012)

during the day, all of which makes them difficult to find, track, and interdict. They can carry as much as two metric tons of cocaine and their transits usually last one or two days (Office of National Drug Control Policy, 2013). In the eastern Pacific, fishing vessels typically depart from Colombian and Ecuadorian ports carrying several tons of cocaine to ports in Central America or Mexico. These operations can last up to six weeks and inject substantial amounts of cocaine into the Central America–Mexico transit route. Drug traffickers have recently been employing, to a greater extent, more subtle maritime transportation in the form of SPSSs.

The use of SPSSs, such as the one in Figure 1.5, to transport cocaine is on the rise in the Caribbean and eastern Pacific. Between 2001 and 2007, 23 events involving SPSSs were documented, compared to 62 SPSS events in the first three quarters of fiscal year 2008 (USCG, 2009). SPSSs are built mainly from wood and fiberglass in the jungles of Colombia. They are capable of carrying up to 10 metric tons of cocaine with a range of 2,500 nautical miles. They range in length from 25 to 80 feet and travel up to 13 knots. A crew of 4 to 5 operates the vessel, which is designed to scuttle upon detection by law enforcement (USCG, 2009). Recently, DTOs have been improving their capabilities with the construction of fully-submersible vessels. In July of 2010, the Ecuador police seized a drug submarine while under construction in the jungles of Ecuador that is 100 feet long and capable of diving to 65 feet (Kraul, 2010). The switch to fully submersible submarines poses a significantly harder threat to combat.

Figure 1.5: SPSS as seen by U.S. Coast Guard prior to being boarded. (From USCG, 2009)

Antidrug efforts have been a mainstay of U.S. drug policy in Latin America and the Caribbean for the last 30 years, with the United States spending $14.6 billion to fight illegal drugs. The strategy is to stop it at the source by destroying drug producing crops, interrupting shipments, and improving socio-economic factors in countries prone to DTOs. The programs are funded by the U.S. Department of State, U.S. Agency for International Development, and the Department of Defense (DoD). The DoD works with foreign countries to train, equip, and improve counter drug efforts. National Security Directive 18, signed by President Bush on August 21, 1989, and the Andean Initiative put counter narcotics on the DoD's list of core priorities. It is now the lead federal agency responsible for detection and monitoring of illegal drugs from foreign countries into the United States. The National Defense Authorization Acts of 1991 and 1998 as amended through FY2011 allows the DoD to provide support to certain countries, mainly Colombia. More recently, the Andean Counterdrug Initiative (ACI) proposed by President Clinton in 2000 and passed through Congress increased funding for counter drug operations in South America. The primary countries of the ACI include Colombia, Peru, and Bolivia (Seelke et al., 2011). In fiscal year 2008, the Merida Initiative increased funding for counter drug operations in Mexico and since then efforts have shifted from Colombia to Mexico, Central America, and the Caribbean. Two recent initiatives, the Central America Regional Security Initiative (CARSI)

and the Caribbean Basin Security Initiative (CBSI) have brought more focus to the region and DoD operations specifically support these initiatives (Seelke et al., 2011).

1.2 U.S. Southern Command

DoD counternarcotics policy is established by the Office of the Deputy Assistant Secretary of Defense for Counternarcotics and Global Threats. Two combatant commands in the western hemisphere execute the mission of providing counter drug support, U.S. Southern Command (SOUTHCOM) and U.S. Northern Command (NORTHCOM). NORTHCOM's Area of Responsibility (AOR) covers Mexico and SOUTHCOM's AOR covers all of Latin America south of Mexico. SOUTHCOM's efforts to counter illegal drug-trafficking directly supports the first objective of their Command Strategy 2020, defend the United States (Fraser, 2010). SOUTHCOM is composed of five component commands, three joint task forces, and one direct reporting unit. SOUTHCOM collaborates with partner nations to combat TOC in their AOR. Illicit drug-trafficking is their primary focus as this is a means for TOCs to strengthen their influence in the region. SOUTHCOM provides support to partner nations enabling them to provide better security in the region. This support is directed to interrupt the flow of illegal drugs, remove TOC networks, and provide assistance to remove criminal influences in lawless areas. As counter narcotics are a core priority of the DoD, SOUTHCOM is the lead agency in the AOR in carrying out this mission. Their efforts are in line with the objectives of the Strategy to Combat Transnational Organized Crime, the CARSI, and the CBSI (United States Southern Command, 2013).

1.3 Joint Interagency Task Force South

JIATF South is a task force of SOUTHCOM that directly performs SOUTHCOM's mission to combat illicit drug-trafficking. Located in Key West, FL, JIATF South detects, tracks, and interdicts illegal drug-trafficking activity in the eastern Pacific, Gulf of Mexico, and the Caribbean. They work with partner nations, military organizations, and U.S. law enforcement agencies in carrying out their mission. U.S. military and partner nations' maritime and air assets are employed in the detection, monitoring, and interdiction of illicit drug-trafficking in the region (United States Southern Command, 2013). The assets available to JIATF South can vary over time and have recently been in decline. Typically, U.S. Navy, U.S. Coast Guard, and partner nation maritime vessels patrol the Caribbean Sea, Gulf of Mexico, and the eastern Pacific. A U.S. Coast Guard Law Enforcement Detachment embarks on a U.S. vessel to take control during boarding operations to seize illicit drugs and apprehend suspects (United States Southern

Command, 2013). In the air, U.S. military, interagency, and partner nation aircraft execute detection and monitoring missions and direct interdiction assets to suspicious surface vessels for further investigation.

In 2011, JIATF South interdicted 119 metric tons of cocaine, arrested 355 drug traffickers, and seized 70 aircraft and vessels, which resulted in a $7.1 billion loss of revenue for DTOs (Posture Statement of General Douglas M. Fraser, 2012). They projected 775 – 930 metric tons of cocaine to move towards the United States in 2012 and in January, JIATF South began Operation Martillo with the goal of disrupting drug-trafficking along Central American coastal waters (Posture Statement of General Douglas M. Fraser, 2012). U.S. Naval Forces Southern Command, in Mayport, FL, supported Operation Martillo with seven frigates, one replenishment ship, and four fixed–wing aviation squadrons (Posture Statement of General John F. Kelly, 2013). In the first year of Operation Martillo, JIATF South and partner nations have disrupted 152 metric tons of cocaine, confiscated $7.2 million in bulk cash, and seized 4 aircraft, 8 SPSSs, and 89 vessels (Posture Statement of General John F. Kelly, 2013). This success is closely dependent on support from partner nations who played a role in 67% of the disruptions associated with Operation Martillo (Posture Statement of General John F. Kelly, 2013). JIATF South has played a major role in the planing and execution of this operation. Based on the amount interdicted in 2011, approximately 85% of the cocaine is not stopped. They oversee a large operational area and effective and efficient employment of available assets is paramount to their success.

1.4 Mathematical Modeling Approach

This thesis addresses some of the uncertainty involved with detection and monitoring of illicit drug-trafficking in the Caribbean and eastern Pacific. JIATF South is the primary agency that is most interested in this work. Given the number of assets they have and the vast amount of area under their watch, their success is heavily dependent on employing available assets as effectively as possible. The main purpose of this overall endeavor is to provide better situational awareness to operators, analysts, and targeteers about the likely location of targets in the present and future. To meet this goal, we formulate two models.

JIATF South faces two primary problems in this context. First, the locations of the targets are unknown. A probability model creates a heat map from various intelligence inputs, which include departure location, arrival location, and departure time. This provides a better picture of the overall AOR. Next, applying a path–finding algorithm provides JIATF South with a

probable track of the target, which will aid in the asset allocation process. Employment of their assets represents the second problem. After developing an estimate for the likely location of a target over time, where to send their assets is the next complexity to overcome. Pietz and Royset (2013) develop an optimization model that takes as input the output from the path–finding algorithm and creates a search plan for JIATF South.

Based on incoming intelligence, the probability model creates a time-layered map indicating the probability associated with a drug trafficker being in a specific area at a specific time. As additional intelligence is received, the map is updated with the new information. Following the creation of the probability map, an algorithm determines the target's course, speed, waypoints (if applicable), time of departure, and time of arrival. The optimization model uses the algorithm output to maximize the effectiveness of JIATF South assets with respect to the potential amount of drugs seized. With the connection of our probability model and path–finding algorithm to the optimization model, JIATF South receives an all-inclusive tactical decision aid to aid in their efforts to counter illicit drug-trafficking.

1.5 Thesis Structure

This thesis develops a method to determine a drug trafficker's course and speed through time given some initial intelligence reports and links additional work being performed in support of JIATF South's mission. Chapter 2 details a literature review to support the efforts of this study. Chapter 3 describes the probability model, key uncertainties, and results of the probability model. Chapter 4 details the target path-finding model and key inputs to the search vehicle optimization model. Chapter 5 presents the conclusions and recommendations of this thesis.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 2:
Literature Review

This thesis covers a range of topics that have been extensively researched in the past and leverages techniques from probability, regression, and optimization. Topics such as search theory can be traced back to World War II and the work of Koopman and colleagues on the Operations Evaluation Group (OEG). More recently, researchers at the Naval Research Laboratory (NRL) have developed a probability model of likely pirate locations based on both intelligence and historical weather data. A major contribution of this thesis links our path-finding model to an JIATF South asset optimization model. This research and other applicable work is reviewed to give context to this thesis.

2.1 Search Theory

The discipline of search theory originates from the need to locate enemy submarines in World War II. Koopman (1946) presents initial findings in this area of research focusing on various types of optimal search and detection techniques with the constraint to maximize use of limited resources. Washburn (2002) provides many mathematical techniques to determine the probability of detecting a target. Rozen (2009) develops a stochastic dynamic programming model to optimize the operational policy of an interdiction force performing a maritime interdiction mission. These works consider cases of varying intelligence about a target, where best to search, and when to perform interdiction, while this thesis develops a grid of probabilities overlaid on the AOR and then determines the target's track. Our focus develops tools to generate the inputs for search models such as Pietz and Royset (2013).

2.2 Piracy Models

Hansen et al. (2011) develop a probability model to aid in the location of pirates. Their model dynamically combines real–time intelligence and weather information to provide a visual picture of likely pirate locations. Intelligence such as homeports, vessel information, performance characteristics, pirate activity, pirate operations, and observations of likely pirate vessels are inputs in the model. Additionally, wind, wave, and current data is supplied creating a "dynamically coupled" model (Hansen et al., 2011). The model produces a prediction of the pirate's track over time with the constraints of intelligence and weather. After thousands of runs are

performed, the model produces a probability map or "heat map" of pirate locations. This map is updated as additional intelligence is received.

Building on the Hansen et al. (2011) piracy model, An et al. (2012) proposes an extension to the model that dynamically allocates search and interdiction assets to minimize the probability of a successful pirate attack. Their research focuses on allocating multiple search and interdiction assets against multiple targets or pirates. An et al. (2012) divide the asset allocation problem into two phases. First, they allocate the interdiction assets. Second, they divide the area not covered by interdiction assets into rectangular areas for search asset allocation.

A pirate's use of a small vessel is significantly impacted by winds and sea state leading to the development of this model. Similarly, drug traffickers operate small vessels during their transit from South America to the major drug transit countries making them susceptible to the same types of weather. The Hansen et al. (2011) model entails a high degree of complexity through the simultaneous analysis of weather, intelligence, and pirate behavior. This model has been extended to the drug trafficker problem but retains its complexity. Our goal is to develop a much simpler model. We analyze straight line paths between an origin and destination without the addition of weather information. Expanding this approach we apply similar techniques to target paths that involve multiple courses and non–constant velocity. Our approach focuses on determining a single target's path through the AOR.

2.3 Probability Model – Optimization Model Link

Pfeiff (2009), Gift (2010), and Bessman (2010) develop attacker – defender type interdiction models to maximize the searcher's probability of detecting an illicit drug trafficker. These theses make use of intelligent smugglers who are able to learn and adapt their voyage to the searcher's actions. This thesis, while retaining some elements of search and detection, is primarily concerned with determining the path of the target to include start time, origin, speed, course, waypoints (if applicable), end time, and destination.

Pietz and Royset (2013) and Pietz (2013) develop an optimization model that routes JIATF South detection and interdiction assets to maximize the amount of drugs interdicted. This model does not take a probability map directly as input, but takes piecewise linear paths to describe the target track and an Area of Uncertainty (AOU) surrounding the path. JIATF South planners may not have access to the necessary parameters to run the optimization model and if weather significantly impacts the target's course, the inputs will be insufficient. Our focus on transform-

ing the probability map into a path links the probabilities generated from gathered intelligence to the optimization model of JIATF South assets. Furthermore, we develop an algorithm to determine the optimal AOU surrounding the target. These efforts directly support the inputs needed for the optimization model.

2.4 Statistics Problem

The probability model produces a set of probabilities indexed over time and identified by the latitude and longitude of the respective grid cell center point. To determine the target's track during the time period, several statistical methods are considered. Principle curves are smooth curves that pass through the middle of the data cloud minimizing the distance from the data (Hastie, 1984). These smooth curves give a general idea of the target path but lack information needed on specific waypoints. Smoothing splines provide similar descriptions of the target track as in principle curves but also lack the needed granularity on when a target changes course (Hastie, Tibshirani, & Friedman, 2009). Hastie et al. (2009) discusses the complexity of determining the appropriate location of knots, and suggests the use of Multivariate Adaptive Regression Splines (MARS).

MARS employs a greedy algorithm and various approximations to reduce the complexity of knot selection (Hastie et al., 2009). MARS has been employed in areas of research concerning survival analysis, time series, ecology, and economics since its introduction by Friedman (1991). MARS is a non–parametric regression technique that automatically models non–linearities and selects variables based on their contribution to overall model fit. This results in a piecewise linear model (Hastie et al., 2009). For this thesis, we desire to define a target's track as a set of piecewise linear paths which is what MARS produces automatically. MARS allows paths with multiple courses or speeds to be modeled accurately by automatically determining locations of potential course or speed changes. Automatic selection of these points makes this technique preferred over other regression techniques. Once MARS determines the points of course or speed changes, all points are checked for validity and any spurious points are removed. MARS supplies the points of speed and heading change and the algorithm builds an ordinary least squares regression model utilizing the selected points of course or speed change.

2.5 Area of Uncertainty

When working with probabilities, there is an inherent uncertainty involved. In our case, this uncertainty resides in the location of the target. At each time step, a set of probabilities are

created over a particular area. An AOU is calculated for each time period using several methods, which include minimum bounding box, convex hull, and a greedy algorithm.

Minimum bounding boxes and convex hulls are used in a variety of applications, including collision avoidance, hidden object determination, and shape analysis. Johnson and Cohen (1998), Lempitsky, Kohli, Rother, and Sharp (2009), and Huebner, Ruthotto, and Kragic (2008) discuss these applications in detail. For our purposes, we present a two–dimensional set of points that we wish to enclose with the minimum bounding box. The area of the minimum bounding box then represents the maximum AOU of the target at a particular time and encloses the entire probability mass. The convex hull further reduces the area while still capturing all of the probability mass. When the maximum AOU or entire probability mass is not necessary, we propose the use of a greedy algorithm to determine the AOU for a user specified level of probability. This approach allows the user to select the minimum amount of probability the AOU should contain.

CHAPTER 3:
Probability Model

JIATF South processes various amounts of intelligence information and makes decisions on where and when to employ available assets. Given this intel, we develop probability models to generate a heat map to provide the operators with better situational awareness about the likely locations of targets, both in the current period and future periods. We first build a simple probability model that accounts for uncertainty only in the departure and arrival locations. From this model, we further enhance its functionality to account for uncertain speed and time of departure as well as consideration for multiple legs. We first present the simplest model and then describe the embellished model.

3.1 Fundamental Probability Model

The fundamental one target analytic model accounts for uncertainty only in the departure and arrival locations on a rectangular area of operation. This setup does not consider the geography of realistic shorelines and we assume the target travels from the bottom of the rectangle to the top of the rectangle in a straight line as shown in Figure 3.1. We further assume the target's departure time and velocity are known. We set up the AOR by defining an area from 0 to X_{max} units in width and 0 to Y_{max} units in height. Within the X_{max} by Y_{max} area, we discretize the AOR into a set of square grid cells for future probability calculations. We further assume that the target's y coordinate at the origin is 0 and the y coordinate at the destination is Y_{max}. Figure 3.1 illustrates the model setup.

The random variables in this case apply to the departure and arrival locations in the x direction and are defined as D and A respectively. The target velocity and departure time remain constant and are deterministic in this simplified model. Additionally, given the departure and arrival locations, we assume the target moves in a straight line between the locations. Let v represent the target velocity and t_0 represent the target departure time. The random variable D follows a probability distribution F_D on the interval $[D_l, D_u]$ and the random variable A follows a probability distribution F_A on the interval $[A_l, A_u]$. Let a represent a realized value of A and d represent a realized value of D. We desire to find the probability that the target is in a specific box at a specific time. We define the box by upper and lower coordinates in both the x and y directions. Let x_l and x_u define the lower and upper edges of the box in the x direction and y_l

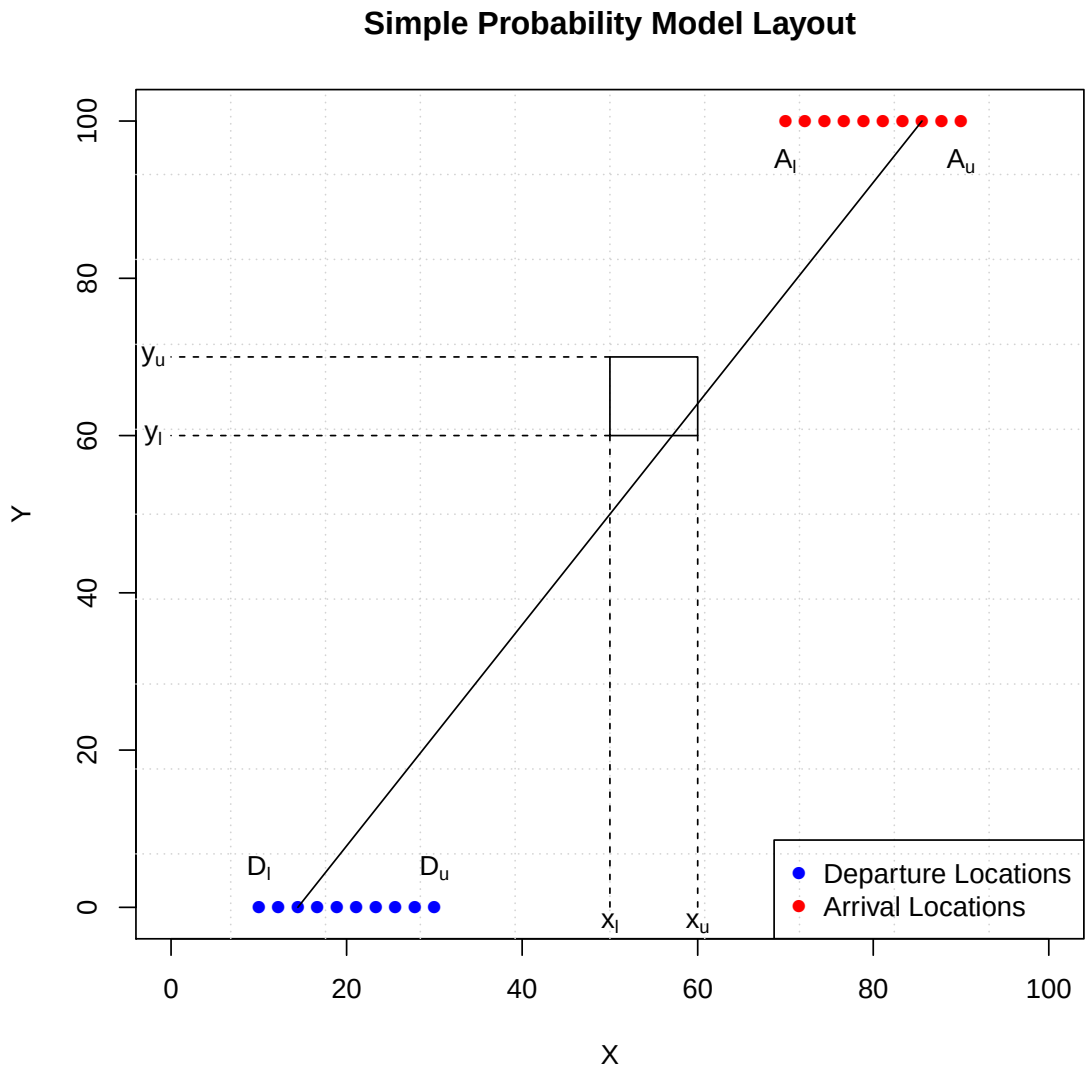


Figure 3.1: Initial layout of the fundamental probability model, which shows the various departure and arrival locations, target track, and box of interest. The target travels from the lower left corner of the area of interest to the upper right corner with constant velocity.

and y_u define the lower and upper edges of the box in the y direction. Let (x_t, y_t) be the coordinates of the target at time t. The target position, (x_t, y_t), is dependent on a, d, v, t, and t_0 and is defined in Equations 3.1 and 3.2. We establish an indicator variable equal to 1 if the target is in the box and 0 otherwise. Equation 3.3 defines the indicator variable. For simplicity, we define square grid cells or boxes for probability calculations but any arbitrary shaped area can be used by modifying Equation 3.3.

$$x_t(d,a,v,t,t_0) = d + \frac{(a-d)v(t-t_0)}{\sqrt{(a-d)^2+Y_{max}^2}} \tag{3.1}$$

$$y_t(d,a,v,t,t_0) = \frac{Y_{max}v(t-t_0)}{\sqrt{(a-d)^2+Y_{max}^2}} \tag{3.2}$$

$$I(d,a,v,t,t_0) = \begin{cases} 1 & \text{if } x_l \le x_t \le x_u \text{ and } y_l \le y_t \le y_u \\ 0 & \text{otherwise} \end{cases} \tag{3.3}$$

Given a departure location, d in $[D_l, D_u]$ and an arrival location, a in $[A_l, A_u]$, we know precisely whether the target is in the box at a specific time from Equation 3.3. The probability based on an uncertain d and a results from integrating over the distributions, F_D and F_A, which yields Equation 3.4.

$$P[x_l \le x_t \le x_u, y_l \le y_t \le y_u] = \int_{D_l}^{D_u}\int_{A_l}^{A_u} I(d,a,v,t,t_0) f_D(d) f_A(a) \mathrm{d}d\mathrm{d}a \tag{3.4}$$

3.2 Numerical Illustrations

After deriving a method to determine the probability a target is in one particular box at a specific time, we expand this to each time period and all of the boxes comprising the operating area. Specifically, for each time period, we want to compute the probability a target is in any grid cell making up the operating area. We present a numerical example to illustrate the application of the model over a multiple hour time period covering the entire area of operations.

The simple one target analytic model as implemented in R, uses discrete departure and arrival distributions. We divide the departure location into n_d bins and the arrival location into n_a bins. For example, the uniform departure distribution takes the value

$$d_i = D_l + (D_u - D_l)\frac{i-1}{n_d - 1} \text{ with probability } \frac{1}{n_d} \text{ for } 1 \le i \le n_d \tag{3.5}$$

and the arrival random variable, A, takes on similar quantities defined as a_j. Thus, the integrals in Equation 3.4 become sums as shown in Equation 3.6.

$$P[x_l \leq x_t \leq x_u, y_l \leq y_t \leq y_u] = \sum_{i=1}^{n_d} \sum_{j=1}^{n_a} I(d_i, a_j, v, t, x_l, y_l, x_u, y_u) f_D(d_i) f_A(a_j) \tag{3.6}$$

For example, let D and A be random variables with uniform probability distributions on the intervals $[D_l, D_u]$ and $[A_l, A_u]$ respectively. Then Equation 3.6 becomes

$$P[x_l \leq x_t \leq x_u, y_l \leq y_t \leq y_u] = \frac{1}{n_d} \frac{1}{n_a} \sum_{i=1}^{n_d} \sum_{j=1}^{n_a} I(d_i, a_j, v, t, x_l, y_l, x_u, y_u) \tag{3.7}$$

The probability model produces target tracks from each d_i to each a_j and plots the associated tracks. When the indicator variables are summed and multiplied by the associated density functions, we produce the probability that the target is in a specific grid cell for each time period. By repeating the calculations for each time period over the entire length of the journey, we produce a matrix of probabilities. This matrix of probabilities is then used to generate a spatial temporal probability map for each time period. An example of the heat map is shown in Figure 3.2 where the departure location follows a uniform distribution, F_D, on the interval $[10, 30]$ and the arrival location follows a uniform distribution, F_A, on the interval $[70, 90]$.

3.3 Enhanced Probability Model

The addition of uncertainty in velocity and departure time further enhances the fundamental probability model. The additional uncertainty introduces two more random variables. We define these as V for velocity and T for departure time. The random variable V follows a probability distribution F_V on the interval $[V_l, V_u]$ and the random variable T follows a probability distribution F_T on the interval $[T_l, T_u]$. Equation 3.4 gains two additional integrals to account for the additional uncertainty as shown in Equation 3.8.

$$\begin{aligned} &P[x_l \leq x_t \leq x_u, y_l \leq y_t \leq y_u] = \\ &\int_{T_l}^{T_u} \int_{V_l}^{V_u} \int_{D_l}^{D_u} \int_{A_l}^{A_u} I(d_i, a_j, v, t, x_l, y_l, x_u, y_u) f_T(t) f_V(v) f_D(d) f_A(a) \mathrm{d}t \mathrm{d}v \mathrm{d}d \mathrm{d}a \end{aligned} \tag{3.8}$$

Furthermore, the enhanced probability model can be used to generate probabilities for targets with multiple legs. The legs are divided by waypoints or geographic positions where the target

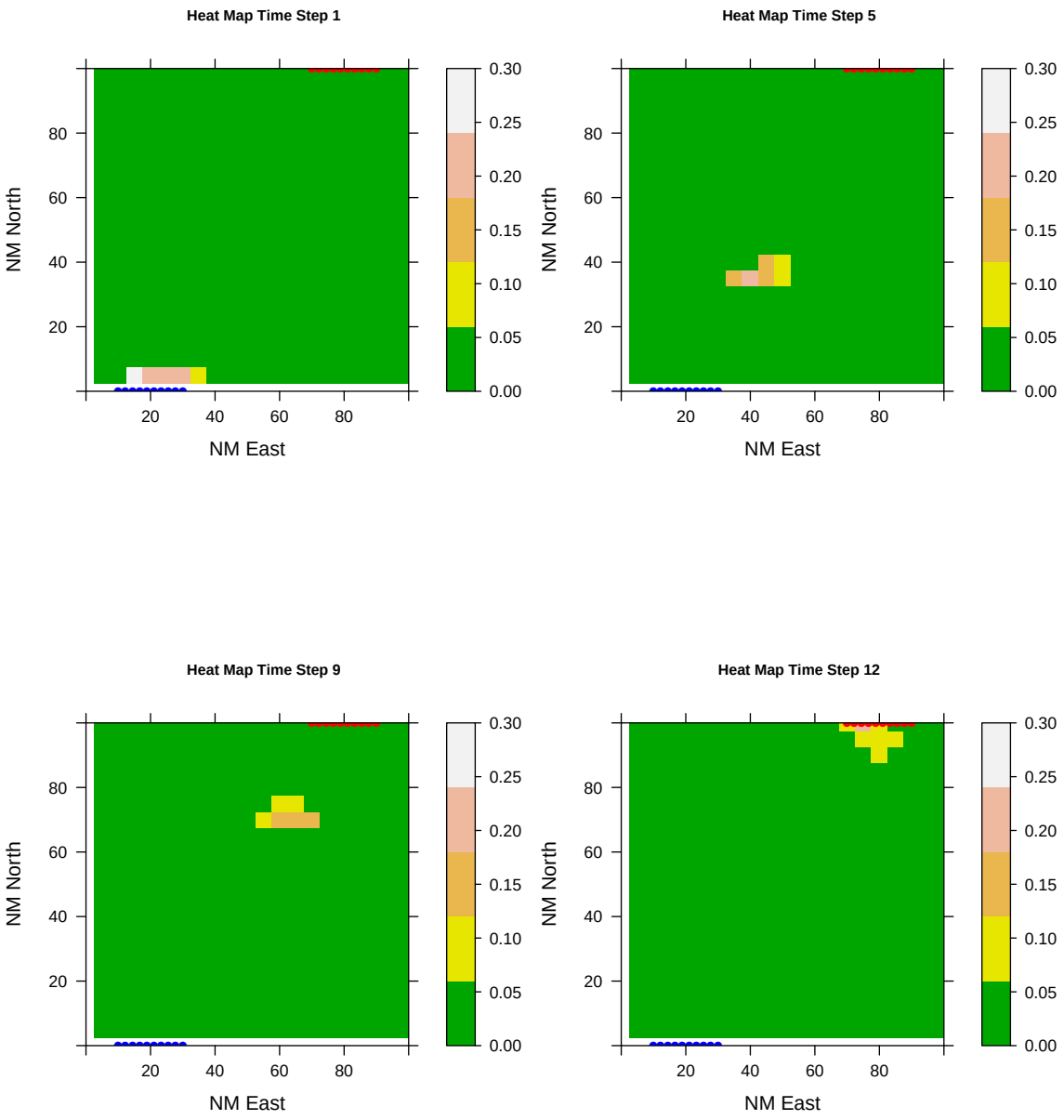


Figure 3.2: Heat maps from several periods of the one target simple analytic model utilizing uniform departure and arrival distributions on the intervals $[10, 30]$ and $[70, 90]$, respectively.

changes course, speed, or both. This functionality of the enhanced model is an extension of the previous model where rather than having just two points on the path, departure and arrival,

there can can be an arbitrary number of waypoints. If there are M total waypoints in addition to the departure and arrival points, then there will be M additional integrals.

As in the fundamental probability model, the basic unit considered by the enhanced model is probability and if the mass of the target is all on the water, the sum of the mass is one. The enhanced model still utilizes square grid cells for probability calculations but the area of interest considers the actual geography of the eastern Pacific and Caribbean. Figure 3.3 shows the enhanced model's area of interest. By considering realistic geographies Equation 3.5 becomes more complex in that the current 1–dimensional equations become 2–dimensional equations.

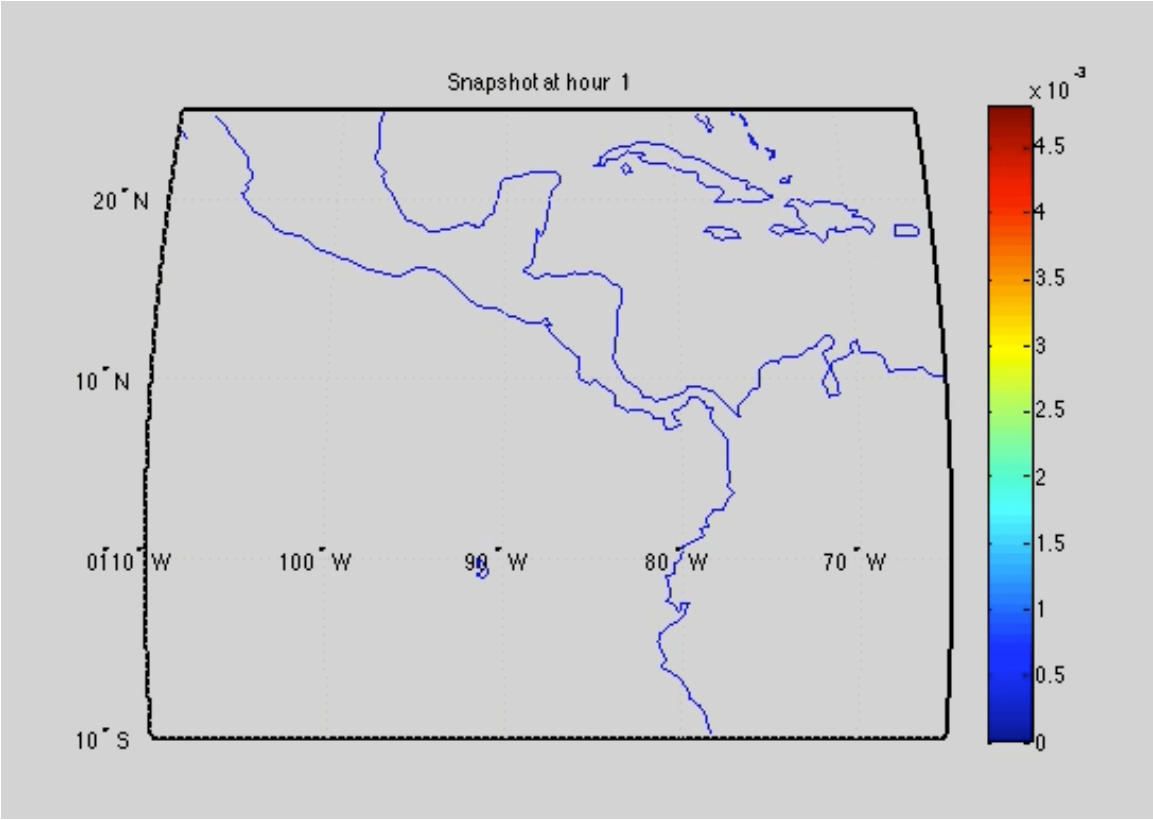
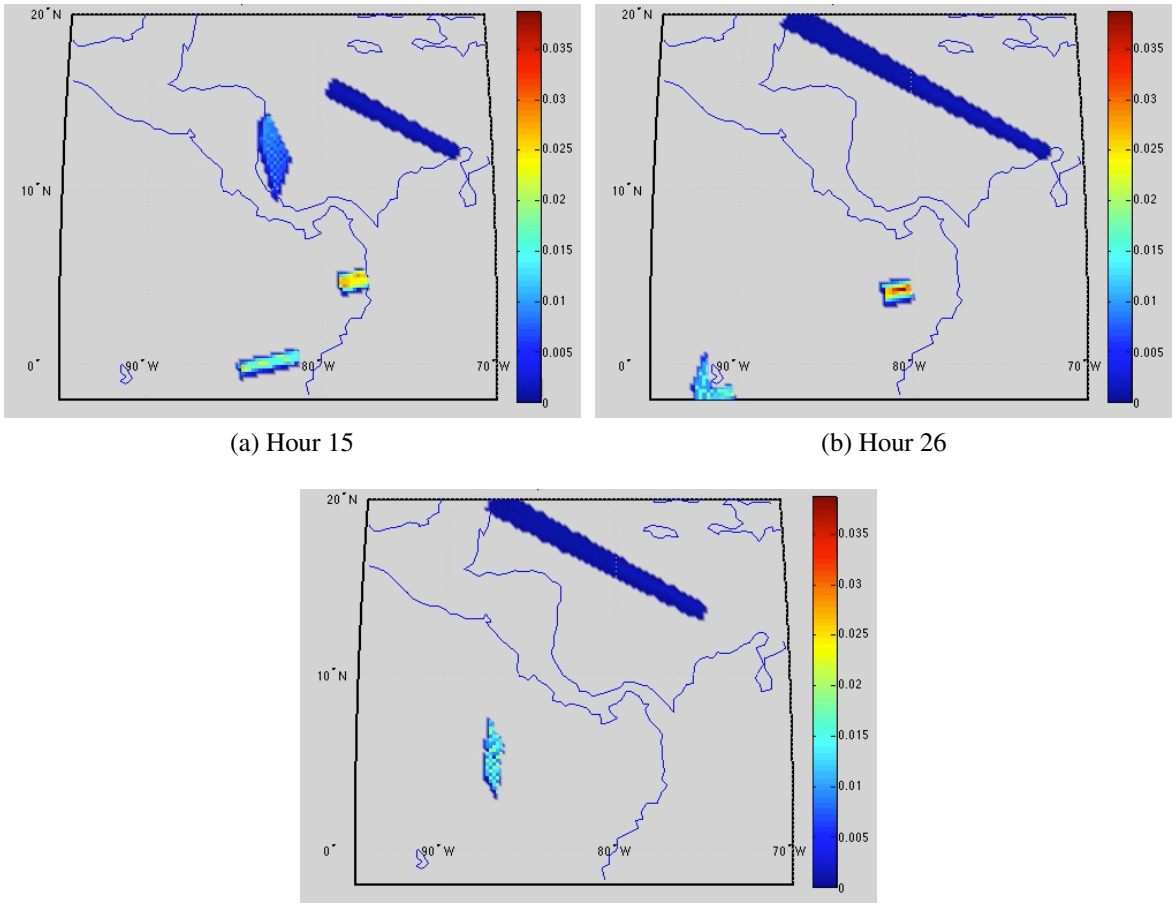


Figure 3.3: Depiction of the area of interest in the enhanced probability model. The enhanced probability model considers the realistic geography of the eastern Pacific and Caribbean.

We construct a model that generates target location probabilities over the course of the target's voyage. The model takes as inputs: a starting location, a destination, target speed, waypoints, expected drug load, and intel certainty. The model applies a predetermined probability distribution to the locations, speed, and departure time inputs. The user can choose either a deterministic distribution, uniform distribution, or triangular distribution. The additional considerations in

the enhanced model introduce more complexity than the basic model in Section 3.1 but the core framework of the model is the same. Numerically, implementation of the enhanced model only requires additional for loops and retains its simplicity when compared to the simulation models such as the piracy model discussed in Section 2.2

The enhanced model provides several generalizations to produce a more realistic heat map. The uncertainty in velocity and departure time simply add two integrals to Equation 3.4 and the more complex area of interest requires a more careful bookkeeping of departure and arrival locations. By including the expected drug load as an input, the model can produce a heat map in terms of pounds of drugs vice probability by simply multiplying the various probabilities by the expected drug load. Considerations for target existence can be made by multiplying the probabilities by the intel certainty. We can model one target or several targets simultaneously on different tracks displaying all targets on one heat map, which would allow JIATF South planners to analyze more complex target scenarios. Figure 3.4 shows a scenario with four targets on the water en route to various destinations in Mexico and Central America. Figure 3.4a depicts the four targets 15 hours into the scenario and Figure 3.4b shows three targets, as one of the four has reached it's destination. The introduction of departure time uncertainty is evident in the long blue channel in Figure 3.4b. This particular target has a departure window of 48 hours thereby spreading the probability mass across the entire voyage. We refer to these cases as "channel" cases. The optimization model handles these cases differently and requires additional inputs. We currently exclude these cases from analysis. Following the construction of a more general probability model, we begin the process of analyzing the probability maps to reproduce the target voyage characteristics.



(a) Hour 15

(b) Hour 26

(c) Hour 57

Figure 3.4: A multiple target scenario produced in the enhanced probability model. The more general probability model is able to consider several targets simultaneously, allowing JIATF South planners to analyze more complex target scenarios. The long "channel" is a special case when the departure window is large. The optimization model handles these differently and we exclude these types of cases from analysis.

CHAPTER 4:
Path-finding Model

A strong focus of this research analyzes how to create the target's track from the probability map generated by the enhanced probability model. Specifically, we determine the target's earliest and latest departure times, earliest and latest arrival times, speed on each leg, course on each leg, location of waypoints, time of change in course or speed, Area of Uncertainty (AOU), departure location, and arrival location. This data is then fed into the optimization model described in Section 2.3 to determine optimum employment of JIATF South assets to maximize the expected amount of drugs seized. Additionally, the analytic probability model provides the ability to verify more complex simulation models such as An et al. (2012)'s extension of piracy model to counter drug operations. The path-finding model assumes that the probability data is the only data available where $p(x,y,t)$ is the probability that a target is located in a grid cell centered at (x,y) at time t. From this data, we derive all other parameters for the optimization model.

4.1 Optimization Model

The Pietz and Royset (2013) optimization model dictates the output required from our path-finding model. This optimization model is a path-constrained optimal search problem in continuous space and time (Pietz & Royset, 2013). The model determines the optimal routing of search assets in the area of operations to maximize the expected drug load. The area of operations is modeled as a transportation network, in which the nodes are search regions and the arcs are the searcher's path between moving search regions. These search regions are areas of uncertainty concerning the target's location, thus knowledge of target parameters is necessary to run the optimization model. The inputs associated with the targets necessary to run the optimization model are listed in Table 4.1 and they correspond to those defined in Pietz and Royset (2013).

The optimization model assumes that the targets travel in a piecewise linear motion at a known constant speed (Pietz & Royset, 2013). The path-finding model derives the uncertainty data listed in Table 4.1 for input into the optimization model. Initially we identify any waypoints, which correspond to changes in a target's course or speed and divide the path into the appropriate number of legs. Subsequent analysis determines velocity, initial position, final position, expected departure time, departure time uncertainty, departure/arrival position uncertainty, and

Expected departure time of target j	τ_j
Time uncertainty range of target j	$\tilde{\tau}_j$
Expected departure location of target j	ρ_j
Expected arrival location of target j	$\bar{\rho}_j$
Departure/arrival location uncertainty range of target j	$\tilde{\rho}_j$
Expected value of detecting target j	q_j
Speed of target j	U_j

Table 4.1: The inputs required to run the optimization model creating a search vehicle allocation plan. (From Pietz and Royset, 2013).

AOU for each leg of the target's path. It is possible some of the uncertainty data in Table 4.1 is directly available from JIATF South databases or intelligence. However, in a worst case scenario, planners have only the probability map and must determine the aforementioned inputs solely from this map to produce a search vehicle allocation plan via the optimization model.

4.2 General Path-finding Model Approach

The path-finding model approximates the path taken by the target as a sequence of legs where the collection of all legs produces the target's entire track from departure location to arrival location. We define a leg as a segment of the target's track that has a constant heading, speed, and AOU. This creates a piecewise linear path for the target. Points in time in which the target's speed, heading, or AOU changes signal the ending of one leg and the beginning of the next. We address new legs based on changes in the AOU once all the constant speed-heading legs are identified. Figure 4.1 shows a potential path from departure location to arrival location for one target. In this case, the target's path consists of two legs in which the target changes heading at the waypoint. The mechanics of the path-finding model automatically identifies where and when a target changes heading or speed. This allows us to generate the piecewise linear path required in the optimization model.

On each leg we must produce linear tracks with a constant speed, heading, and AOU. We determine these inputs for each target separately and compile the inputs by leg for the optimization model. We apply regression techniques to generate piecewise linear paths through the probability mass. Specifically, the path-finding model produces two linear regression equations of the target's location, one for longitude and one for latitude. The coefficients associated with the two linear regression equations are then used to calculate the required quantities. The associated mathematical equations are discussed further in Section 4.4 and numerical illustrations are

Figure 4.1: Illustration of one target traveling along two legs from departure to arrival location. This track consists of two legs with a heading change at the waypoint.

provided in Sections 4.5 and 4.6.

From the probability model, we analyze a set of non-zero probabilities that are spread over a cellular grid in the AOR. For each time period, the probabilities correspond to the center of a grid cell by their associated latitude and longitude. We define n as the number of grid cells in the AOR, $p_i(t)$ as the probability the target is in cell i at time t, and (x_i, y_i) as the center point of grid cell i. Figure 3.2 illustrates four different hours of the probability mass in the AOR with each grid cell colored based on the target's probability of being in that cell at the associated time. This probability mass changes size and shape and moves over time as the target travels from the departure location to the arrival location. At each time step, we calculate the expected value of the target's longitude or x component where $E[X] = \sum_{i=1}^{n} p_i(t)x_i$ and the expected value of the target's latitude or y component where $E[Y] = \sum_{i=1}^{n} p_i(t)y_i$.

Since the probability mass is spread over several grid cells during each time step, weighted target tracks could be produced but for simplicity and limitations of the techniques employed, we use expected values.

The resultant expected x and y coordinates are then plotted in Figure 4.2 for a visual picture of

the target's expected track. Figure 4.2 illustrates the same track from Figure 4.1 but uses the expected values of X and Y. As in Figure 4.1, Figure 4.2 shows two legs with a heading change near 86°W 3°N $(274°, 3°)$.

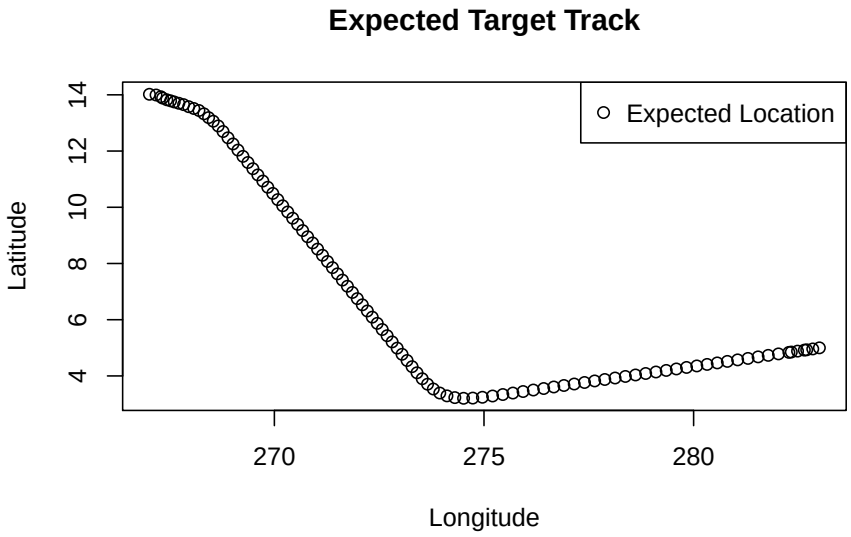


Figure 4.2: Plot of one target's expected track with two legs in the AOR.

Based on the plot of latitude versus longitude, we can easily see a change in heading but discerning departure and arrival locations or speed changes is not possible. The necessary time element is not depicted in Figure 4.2. This leads us to plot longitude versus time and latitude versus time as shown in Figures 4.3 and 4.4 respectively. These plots become the primary means of determining the necessary target parameters in Table 4.1.

Our regression analysis utilizes the probability data between the latest departure time and the earliest arrival time, which corresponds to the time that the entire probability mass is on the water. These points are specifically highlighted in Figures 4.3 and 4.4. Any change in slope between these points signifies a change in the target's speed, heading, or possibly both. Changes in slope outside this interval may produce spurious changes in course or speed due to endpoint effects and are not important for optimization. Figures 4.3 and 4.4 both show a change in slope around hour 45, signifying a change in the target's motion. The points where a change in slope exists in either the x direction or y direction indicates where one leg ends and the next leg begins.

Detecting two legs is obvious in Figure 4.1 because there is a considerable change in course.

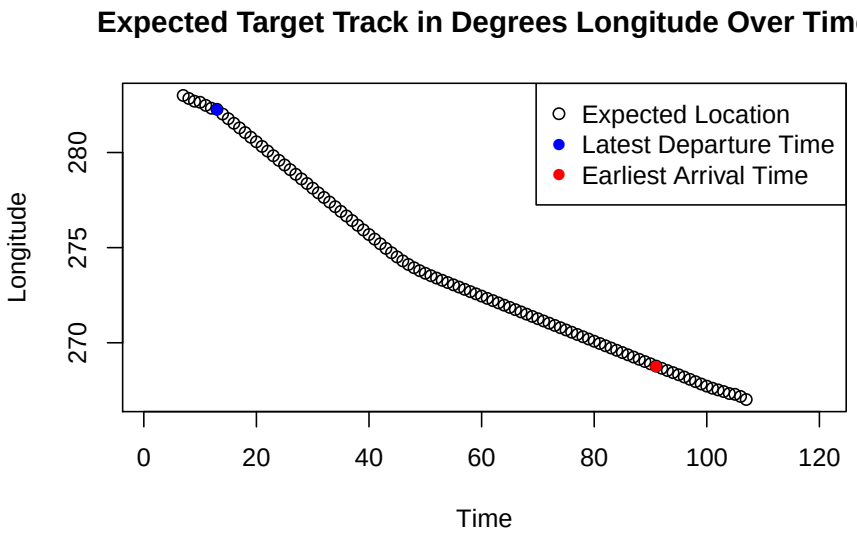


Figure 4.3: Plot of one target's expected track with two legs in degrees longitude over time.

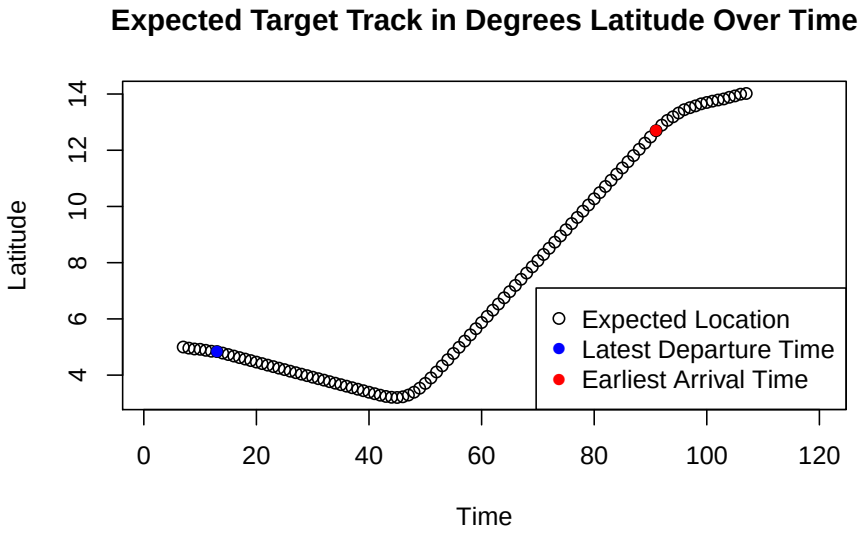


Figure 4.4: Plot of one target's expected track with two legs in degrees latitude over time.

However, a change in speed with a constant course can not be determined with a static figure. Consider the example in Figure 4.5 where the target is traveling in a straight line from the departure location to the arrival location.

Based on Figure 4.5, it appears one leg exists in this target's path. Looking at the expected track

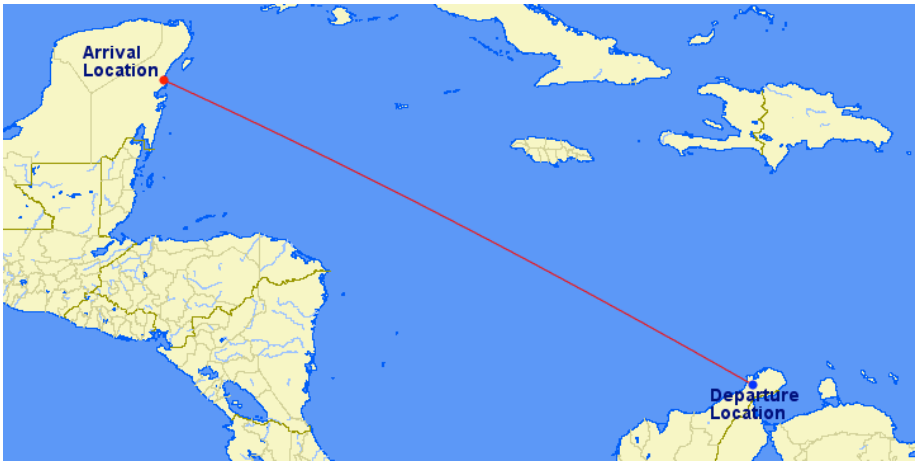


Figure 4.5: Illustration of a potential target traveling in a straight line path from the departure location to the arrival location.

in Figure 4.6 we are inclined to the same conclusion.

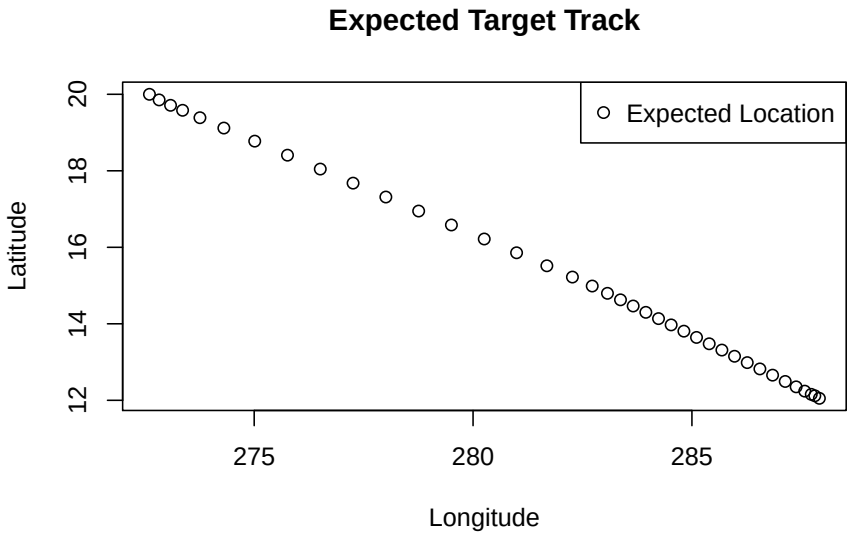


Figure 4.6: Plot of the one target straight line path in degrees latitude versus degrees longitude.

A closer look at this figure shows the spacing between expected locations increasing. This increase in space between expected locations signifies the target has changed speed. We are still missing the time element and cannot tell whether the target is traveling northwest or southeast but we can tell speed has either increased or decreased during the journey. When we consider the plots in Figures 4.7 and 4.8, it is evident the target is traveling northwest and increases speed during the journey based on the increase in slope at hour 25.

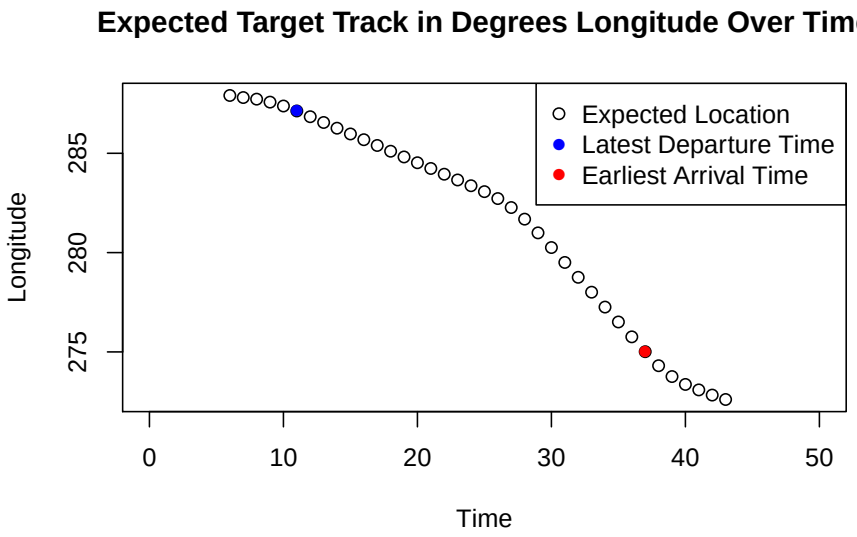


Figure 4.7: Plot of the one target straight line path in degrees longitude over time.

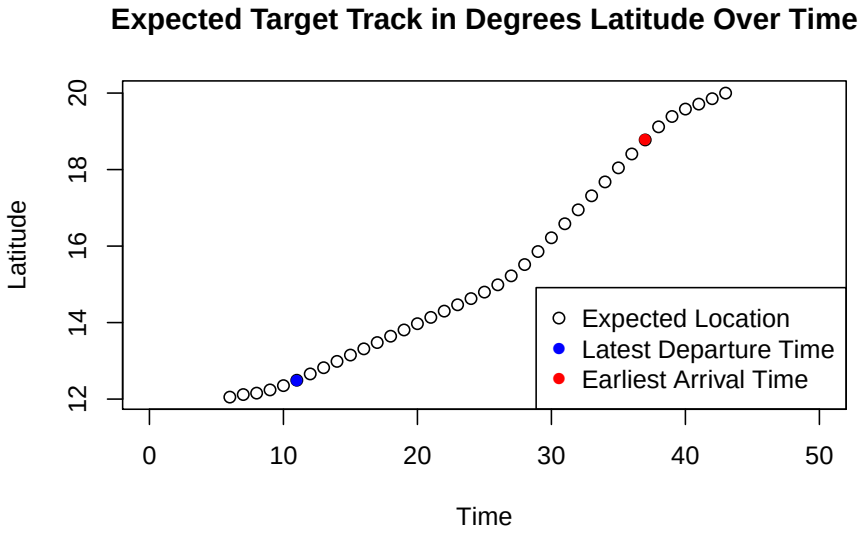


Figure 4.8: Plot of the one target straight line path in degrees latitude over time.

This case clearly illustrates the necessity of analyzing the x and y components of the target's motion separately and the need to employ a technique capable of producing piecewise linear equations.

4.3 Path-finding Model

We desire to provide JIATF South with a path-finding model capable of performing this analysis automatically. To meet this need we employ a technique known as MARS. MARS, as described in Section 2.4 automatically fits a piecewise linear regression model to the probability data by automatically selecting "knots" or points of change in course or speed.

MARS uses a forward model building and backward term deletion procedure to produce the final model (Hastie et al., 2009). During the forward model building procedure, MARS uses a set of basis functions as shown in Equations 4.1 and 4.2 as input.

$$(x-t)_+ = \begin{cases} x-t, & \text{if } x > t, \\ 0, & \text{otherwise} \end{cases} \tag{4.1}$$

$$(t-x)_+ = \begin{cases} t-x, & \text{if } x < t, \\ 0, & \text{otherwise.} \end{cases} \tag{4.2}$$

For each input, MARS forms a set of reflected pairs, $(X_j - t)_+$ and $(t - X_j)_+$, for each value of that input (Hastie et al., 2009). The resulting model is of the form

$$f(X) = \beta_0 + \sum_{m=1}^{M} \beta_m (X - t_m)_+ + \sum_{n=1}^{N} \beta_n (t_n - X)_+ \tag{4.3}$$

where t_m and t_n are the candidate times of course or speed changes. Additionally, the model form can include a product of two or more basis functions (Hastie et al., 2009). For a particular basis function, MARS estimates the associated coefficient, β, by minimizing the residual sum-of-squares (Hastie et al., 2009). The core functionality of MARS lies in the selection of the basis functions. Each step through the forward model building process a new basis function pair and its products of all functions in the model are evaluated for inclusion in the model. The term that produces the largest decrease in error is then added to the model. For a full description of MARS see Friedman (1991) and Hastie et al. (2009).

Consider the target scenarios presented in Figures 4.1 and 4.5. As previously discussed, these targets follow tracks that consists of two legs identified by a course change in the expected track

and a speed change respectively. Utilizing MARS, the path-finding model automatically locates changes in slope as seen in Figures 4.9 through 4.11 and produces final fitted models of the form

$$x = \beta_0 + \beta_1(t-45)_+ + \beta_2(45-t)_+ \tag{4.4}$$

$$y = \beta_0 + \beta_1(t-37)_+ + \beta_2(t-49)_+ + \beta_3(49-t)_+. \tag{4.5}$$

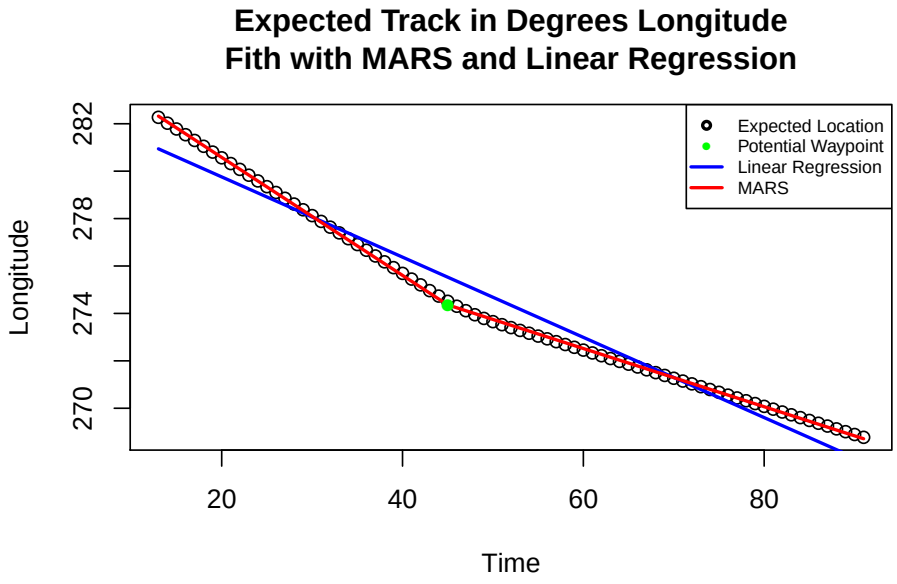


Figure 4.9: Plot of one target's track with two legs in degrees longitude over time. The resulting MARS model fits the path extremely well, whereas ordinary least squares regression is unable to automatically identify the course change.

Figures 4.9 through 4.11 show the resulting MARS model fit and illustrate the poor resulting fit by using standard linear regression. Standard regression does not adequately identify the slope changes, thus it does not produce the piecewise linear track that MARS does. Note, Figure 4.10 displays two slope changes and Equation 4.5 includes different basis functions than Equation 4.4. In fitting models in the x and y directions separately, MARS produces a best fit model irregardless of the number of basis functions in the other direction. A few reasons exist for the additional knot and are discussed further in Section 4.6 as well as the case parameters and associated model output.

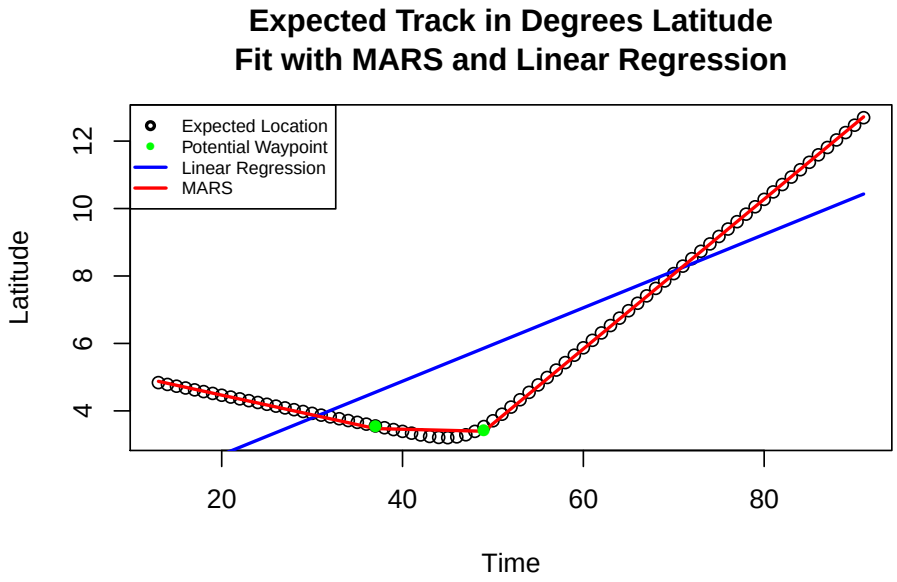


Figure 4.10: Plot of one target's track with two legs in degrees latitude over time. The resulting MARS model fits the path extremely well, whereas ordinary least squares regression is unable to automatically identify the course change.

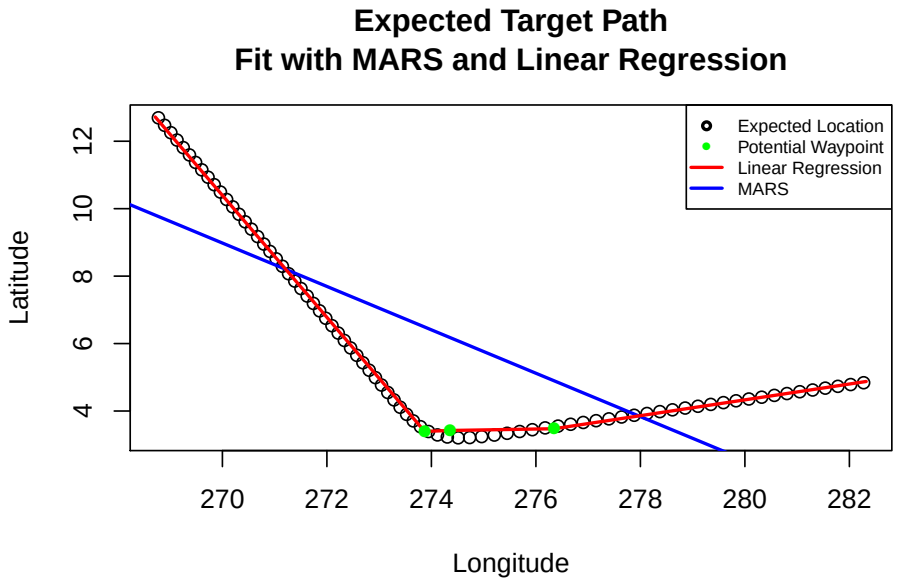


Figure 4.11: Plot of one target's track with two legs in degrees latitude versus degrees longitude. The resulting MARS model fits the path extremely well, whereas ordinary least squares regression is unable to automatically identify the course change.

At this point, the basis functions represent potential waypoints along the target's route. Next, the path-finding model checks the validity of these waypoints by first evaluating the change in slope of reflected pairs, and then evaluating the time between waypoints as compared to the time required for the probability mass to complete the associated turn. These concepts are demonstrated in greater detail in Sections 4.5 and 4.6

Once we know which waypoints are valid, we construct a new data set to account for the waypoints. We add a new data field, wp_i, for each valid waypoint where wp_i equals the value of the associated basis function. If t_{wp} is the time of the respective waypoint, then we define $wp = (t - t_{wp})_+$.

The path-finding model uses this data set and ordinary least squares regression to produce the final model, which is of the form

$$x = \beta_0 + \beta_1 t + \sum_{i=1}^{j} \beta_i wp_i \tag{4.6}$$

and

$$y = \beta_0 + \beta_1 t + \sum_{i=1}^{j} \beta_i wp_i \tag{4.7}$$

where j represents the number of waypoints. The additional data fields, identified by wp_i, enable ordinary least squares regression to produce a piecewise linear fit.

4.4 Input Generation

The probability data and the path-finding model produce the uncertainty parameters listed in Table 4.1 to generate a search vehicle allocation plan. Recall we have uncertainty in departure location, arrival location, departure time, and velocity and in a worst case scenario JIATF South planners will only have the probability map to generate the needed parameters for optimal allocation of their search assets. The following section describes how each of the inputs are generated from either the probabilities or the model output. The desired end-state equates to a list of parameters for each target and leg of the associated target's path.

4.4.1 Times

The introduction of T, the target's random departure time, in the enhanced probability model produces uncertainties in the departure time and subsequently the arrival time for a target j. The uncertainty results in a range of possible departure times between $[T_l, T_u]$ and a corresponding range of arrival times. The specific times for each leg required in the Pietz and Royset (2013) optimization model are listed in Table 4.2. Additionally, we define the earliest departure time of target j from the origin as $\tau_j^{earliestdep}$ and the latest arrival time at the destination as $\tau_j^{latestarr}$ for computational purposes.

Expected departure time of target j	τ_j
Time uncertainty range of target j	$\tilde{\tau}_j$
Latest departure time of target j	τ_j^{min}
Earliest arrival time of target j	τ_j^{max}

Table 4.2: Times required on each leg for the optimization model to generate a search vehicle allocation plan. (From Pietz and Royset, 2013)

Pietz and Royset (2013) discuss calculating τ_j^{min} and τ_j^{max} from the uncertainty data, τ_j and $\tilde{\tau}_j$. Our path-finding model first determines $\tau_j^{earliestdep}$, τ_j^{min}, τ_j^{max}, and $\tau_j^{latestarr}$ for the first and last leg respectively and then we calculate τ_j and $\tilde{\tau}_j$. It is important to note that these initial time determinations are for the first leg and the last leg and are determined directly from the probability data. Each leg of the target's path requires the identification of the times listed in Table 4.2. MARS generates these times for the intermediate waypoints, whereas the probability data generates the times associated with the origin and destination. Figure 4.12 illustrates the associated times. In this case more than one leg exists, thus the intermediate departure and arrival times are determined through regression analysis.

The earliest departure time is the first time period that any non-zero probability exists whereas the latest departure time is the first time period that the sum of the probabilities is equal to 1. These are expressed in Equations 4.8 and 4.9 respectively.

$$\tau_j^{earliestdep} = \min\left\{ t | \sum_{i=1}^{n} p_i(t) > 0 \right\} \tag{4.8}$$

$$\tau_j^{min} = \min\left\{ t | \sum_{i=1}^{n} p_i(t) = 1 \right\} \tag{4.9}$$

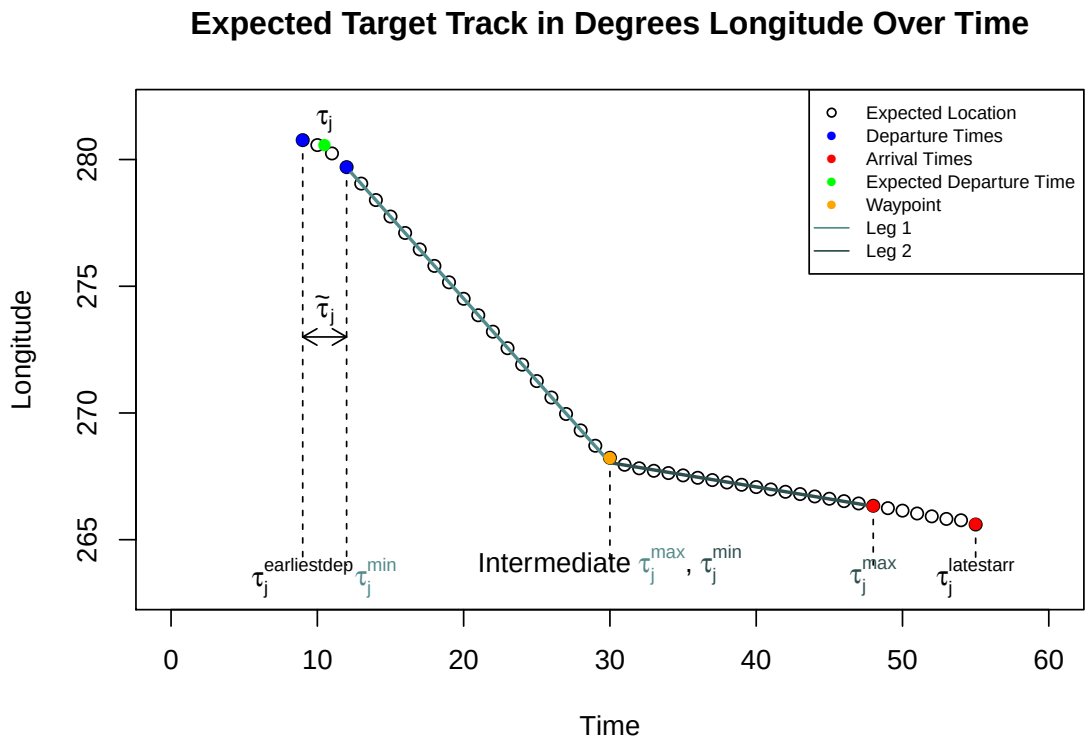


Figure 4.12: The initial time determinations in the path-finding model represent the earliest and latest departure times on the first leg and the earliest and latest arrival times on the last leg.

Similarly for the arrival times, the earliest arrival time is the maximum time that the probabilities sum to 1 and the latest arrival time is the maximum time the probabilities sum to greater than zero. These are expressed in Equations 4.10 and 4.11 respectively.

$$\tau_j^{max} = \max\left\{ t | \sum_{i=1}^{n} p_i(t) = 1 \right\} \tag{4.10}$$

$$\tau_j^{latestarr} = \max\left\{ t | \sum_{i=1}^{n} p_i(t) > 0 \right\} \tag{4.11}$$

It is helpful to view the origin as a source and the destination as a sink. The probability mass begins to grow from the source and eventually 100% of the probability mass has left the source, thus representing the earliest and latest departure times respectively. As the target travels across

the AOR, the probability mass begins to disappear into the sink until all of the mass has finally collected in the sink, thus representing the earliest and latest arrival times. When only one leg exists for a target, all times are determined directly from the probability data. The departure time uncertainty range of target j, $\tilde{\tau}_j$, is simply the difference between the earliest departure time and latest departure time. Based on $\tau_j^{earliestdep}$ and τ_j^{min}, we determine $\tilde{\tau}_j$ and τ_j using Equations 4.12 and 4.13 respectively.

$$\tilde{\tau}_j = \tau_j^{min} - \tau_j^{earliestdep} \tag{4.12}$$

$$\tau_j = \frac{1}{2}(\tau_j^{min} - \tau_j^{earliestdep}) \tag{4.13}$$

If the departure window is large enough, the probability mass can grow from the source to the sink without ever fully leaving the source. These cases are referred to as "channel" cases and the optimization model handles them slightly different. An example is shown in Figure 3.4b. At this point, we only focus on cases where the probability mass develops entirely on the water before beginning to absorb into the sink. For purposes of the path-finding model, we are mainly concerned with the latest departure time and the earliest arrival time. In other words, the path-finding model only uses the probability data when the probabilities sum to 1, which is the probability data between the first τ_j^{min} and the last τ_j^{max}.

4.4.2 Position

Each position is composed of a horizontal component and a vertical component and the random variable D introduces uncertainty in the departure location on the interval $[D_l, D_u]$. The regression models, when run between the first τ_j^{min} and the last τ_j^{max}, result in a set of fitted values between these times. Essentially, they are an estimate of the target's horizontal and vertical position in the AOR for each time step between departure and arrival. The optimization model requires position inputs for each leg, specifically a starting position and an ending position. Pietz and Royset (2013) define the origin of the AOR as 0° N (0°) and 100° W (260°) where 1° in any direction equals 60 nm. All positions are measured from this origin.

The fitted values from the regression model directly correspond to the position of target j. Knowing τ_j^{min} and τ_j^{max} for each leg allows us to determine the target's position by substituting τ_j^{min} and τ_j^{max} into the final models. The resulting $(x_{\tau_j^{min}}, y_{\tau_j^{min}})$ and $(x_{\tau_j^{max}}, y_{\tau j^{max}})$ are

then converted to nautical miles from the AOR origin where $\rho_j(x) = (x_t - x_0) * 60$ and $\rho_j(y) = (y_t - y_0) * 60$.

Pietz and Royset (2013) define ρ_j and $\bar{\rho}_j$ as the expected departure location and the expected arrival location of target j. In other words, ρ_j and $\bar{\rho}_j$ are the target's position at τ_j and the last τ_j^{max} respectively. An accurate estimate of the expected departure location is required for the optimization model, but an accurate estimate of the expected arrival location is not required as long as its along the same track. Since we regress on the probability data between the first τ_j^{min} and the last τ_j^{max}, our estimate, based solely on regression, of ρ_j is always further from the true origin. This fact does not impact the path-finding model, however; to correct this, we use the vertical and horizontal speeds to move the position at the first τ_j^{min} the distance of one-half the departure time uncertainty times the respective speed. For example, with a departure time uncertainty of 4 hours and a vertical speed of 10 knots and a horizontal speed of 12 knots, the ρ_j for the first leg moves 20 nm in the vertical direction and 24 nm in the horizontal direction towards the true origin.

4.4.3 Speed

The optimization model assumes the searcher knows the speed of target j on each leg of the path and defines this speed as U_j (Pietz & Royset, 2013). This assumption then applies to the entire leg without considering the possibility of speed changes during the leg. Recall, a change in speed or heading is cause for the introduction of a new leg since the optimization assumes a constant speed on each leg. Considering the worst case scenario, where only the probability data exists, we desire to determine the target's speed on each leg. The coefficients of the regression model in Equations 4.6 and 4.7 are used to calculate speed. The coefficients in the linear models, excluding the intercept (β_0), represent the speed in the x and y directions. Overall target speed is then defined as $U_j = \sqrt{\beta_x^2 + \beta_y^2}$.

In a multiple leg case, the regression techniques produce a piecewise linear model for the entire journey where, depending on the time, certain coefficients are active. Recall the MARS discussion in Section 4.3, the basis functions equal zero when they evaluate to a negative quantity. Therefore, for times when a basis function equals zero the respective coefficient does not contribute to the target's speed. The coefficients only contribute to the speed calculation when the basis function evaluates to a non-negative number. Knowing the τ_j^{min} and τ_j^{max} for each leg, the path-finding model determines which coefficients are active and based on the active coefficients computes the speed for the respective leg.

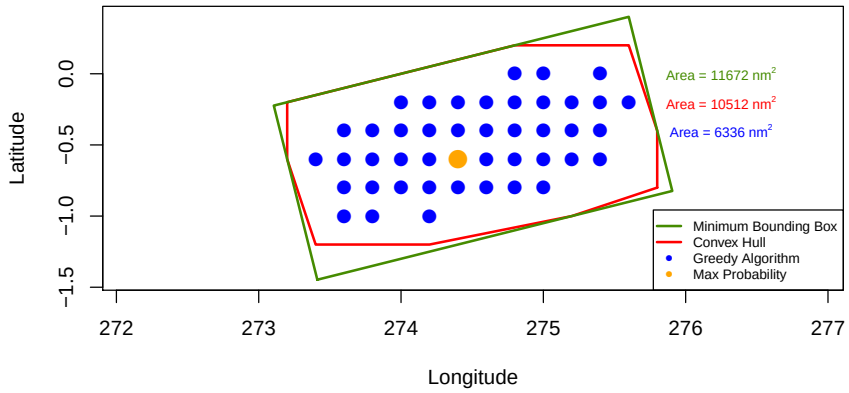
4.4.4 Area of Uncertainty

An AOU results from uncertainty in both the departure time and the departure location. This area is essentially the probability mass moving across the water. Just as a constant speed on each leg is required, so is a constant area along each leg. Changes in the AOU also would indicate the need for additional legs. Determining the area of the probability mass as it moves across the water is rather straightforward. We can simply count the number of grid cells that contain probabilities greater than 0. Knowing the size of each grid cell and the number of grid cells with a non-zero probability, we simply multiply the two and determine the associated AOU. This method while straight forward does not result in a constant area over time. In general, the area changes every time step and the user may not require 100% coverage of the AOU. We explore several methods to determine the AOU.

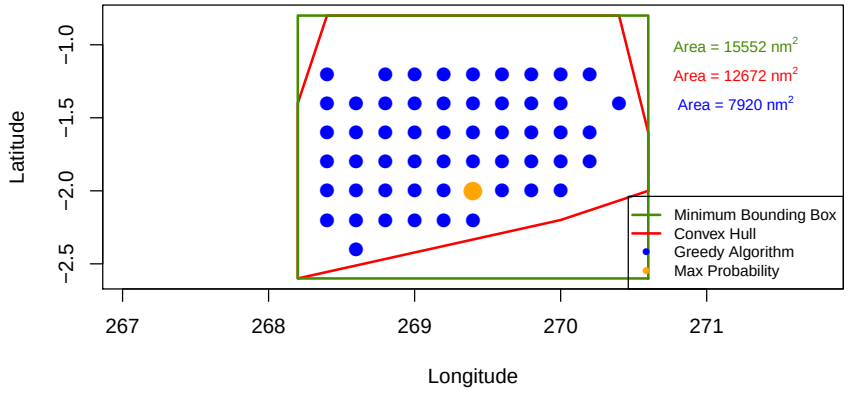
The first approach determines the area of the minimum bounding box at each time step of the target's path. This method captures the entire probability mass and results in the largest AOU. Additionally, it is possible for the minimum bounding box to capture locations with zero probability, as evident in Figure 4.13. Therefore, the second approach determines the convex hull of the probability mass. Again, the convex hull captures all of the probability mass but with minimal area. Figure 4.13 depicts the minimum bounding box and convex hull at several time steps. These methods successfully bound the entire probability mass and capture 100% of the area.

The third approach uses a greedy algorithm to capture a user defined amount of the probability. Throughout this thesis we assume a probability threshold of 90%. For each time between the first τ_j^{min} and the last τ_j^{max}, the greedy algorithm selects the node or cell with the highest probability and then creates a set of cells that neighbor the initial cell. The algorithm proceeds by examining the set of neighbors and selects the neighbor with the highest probability and adds it to the AOU set. Next it calculates the cumulative probability of the AOU set and proceeds to look at the set of neighbors including any new neighbors, again adding the neighbor with the highest probability. This procedure continues until the cumulative probability is greater than or equal to the selected threshold. Once we reach the specified threshold, we simply determine the amount of cells in the set and multiply by the area of a single cell to calculate the total AOU. Figure 4.13 shows the grid cells that are included in the greedy algorithm's AOU set.

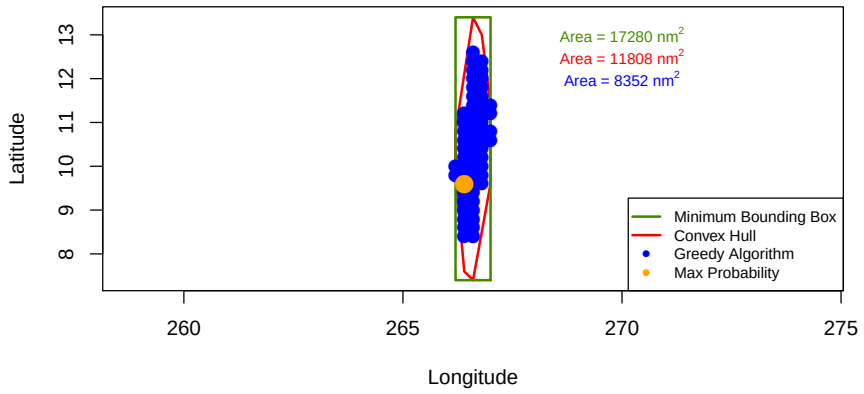
Figure 4.14 shows the results of the three methods over time. The AOU found using the minimum bounding box tends to result in a considerably higher area since it encompasses the entire



(a) Hour 20



(b) Hour 28



(c) Hour 70

Figure 4.13: Examples of the minimum bounding box, convex hull, and greedy algorithm applied to a target's probability mass over several time steps of a target's path. The minimum bounding box and convex hull encompass the entire probability mass at each time step, while the greedy algorithm results in the smallest AOU.

probability mass and additional area with zero probability. The convex hull performs slightly better but it is still greater than the greedy approach. By taking a probability threshold of 90%, we are able to significantly reduce the size of the AOU while still ensuring we capture a significant amount of the probability mass. Additionally, the greedy algorithm allows for flexibility and if desired, the user may select a threshold of 100%.

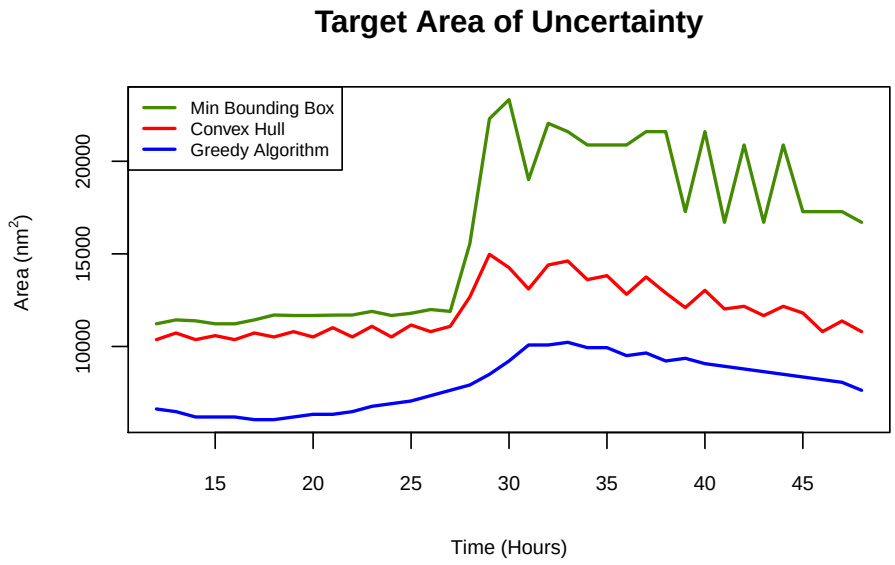


Figure 4.14: Comparison of the AOUs resulting from the three methods explored. The greedy algorithm consistently produces areas less than the other two approaches, while the minimum bounding box results in significantly higher areas.

With the AOU specified for each time step of the target's path, we now desire to create a piecewise constant area for the entire journey. The area analysis utilizes the functionality of MARS and automatically fits a piecewise linear model to the area data. The model smooths the area data but does not produce the piecewise constant areas required. Figure 4.15 shows the MARS model fit to the areas from the greedy algorithm. This implementation of MARS forces a minimum span between knots equal to the departure time uncertainty. The minimum span prevents MARS from introducing knots at points just before or after a new leg.

Additionally, for each constant speed and heading leg, we build a separate model since we desire to create the minimum number of new legs based on area, and we know the number of legs based solely on speed and heading changes. Figure 4.15 depicts the MARS models fit to

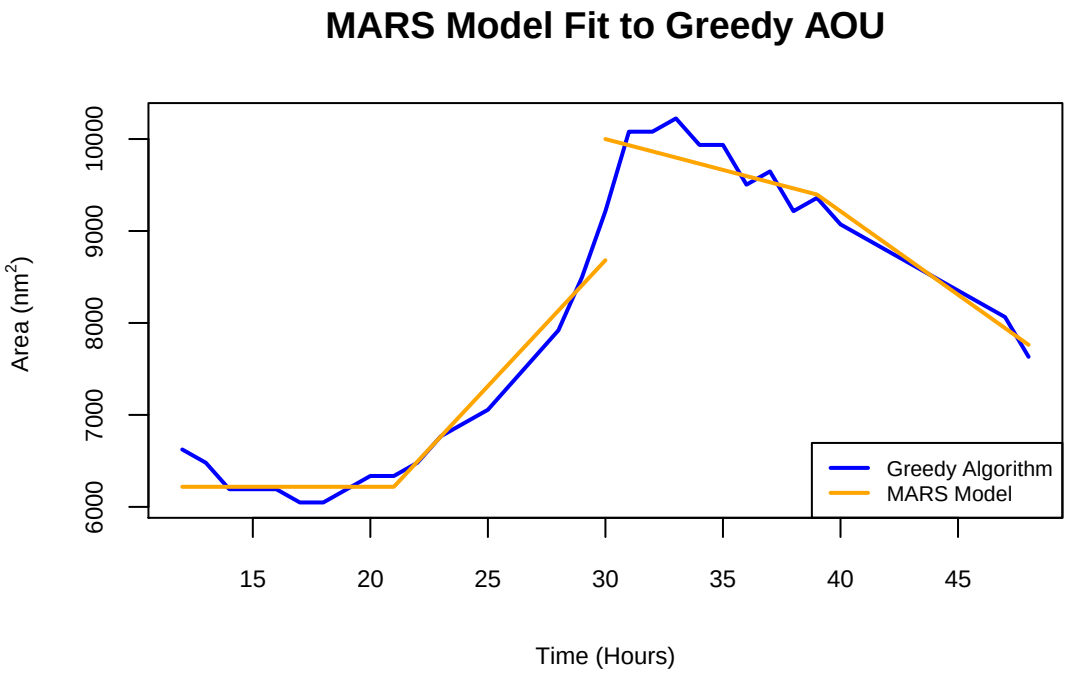


Figure 4.15: A MARS model smooths the AOU over time. A different model is built for each leg with a constant speed and heading. The fitted values from the model are further analyzed to produce a piecewise constant AOU.

each leg of the target's path. To determine the number of additional legs based on a change in the AOU, we calculate the maximum number of legs that we can fit in each constant speed-heading leg based on two different criteria. First, we use the minimum span to determine the number of additional legs that can be fit into each constant speed-heading leg. This is simply the length of the leg in time divide by the departure time uncertainty. The second method takes a user selected percent change in area, which is currently set at 25%, and determines the number of times a 25% change in area occurs. Then we take the minimum between the two methods. This gives us the number of new legs to insert based on area changes. We perform this procedure on each constant speed-heading leg. Considering the slope of the MARS model, we assign areas to the new legs at the maximum area of the new interval. Figure 4.16 shows the resulting piecewise constant area.

In this particular case, there are two constant speed-heading legs where the course changes at hour 30. Based on the AOU analysis, we divide the first constant speed-heading leg into two

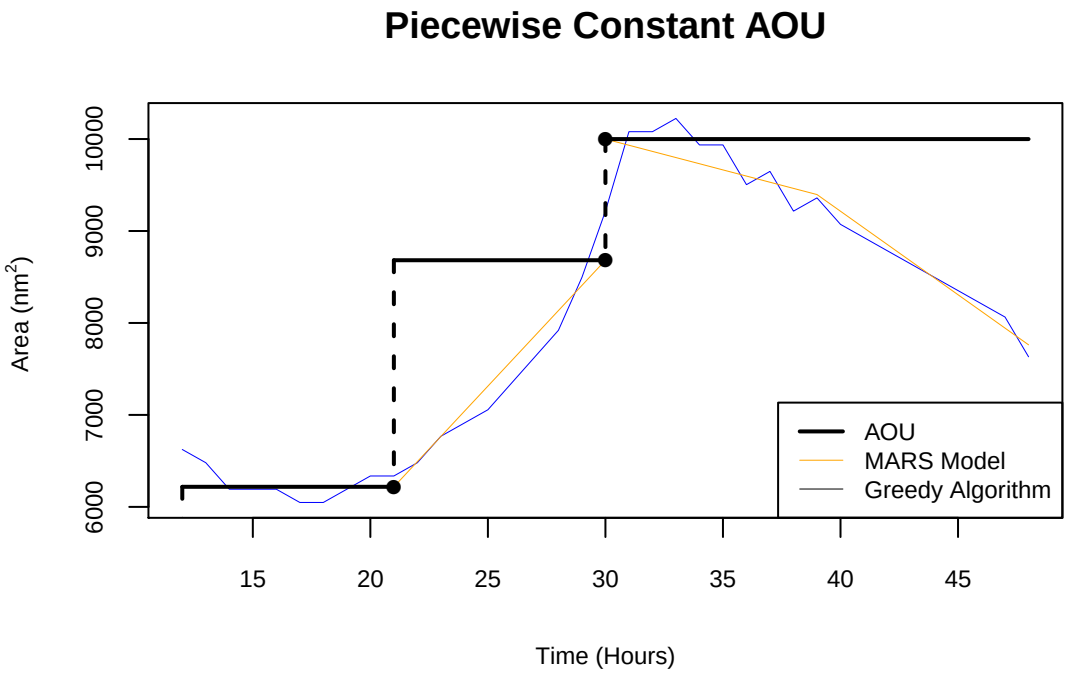


Figure 4.16: Illustration of a piecewise constant AOU.

segments. The first segment ends at hour 21, at which point, the AOU increases. The second constant speed-heading leg remains unchanged because the AOU does not change by more than 25%.

4.5 Straight Line Model and Illustration

Assume there is one target of interest that travels from the origin to the destination with a constant speed and heading. Figure 4.17 shows the target's track from departure location to arrival location. We illustrate the path-finding model by first entering the case parameters into the enhanced probability model. The case parameters as shown in Table 4.3 include uncertainty in departure and arrival locations and departure time.

Running the probability model produces the probability data over time necessary for the path-finding model. Figure 4.18 depicts several hours of the target's location as it moves through the AOR. In a worst case, this is all the JIATF South planners have and they desire to determine the information in Table 4.3.

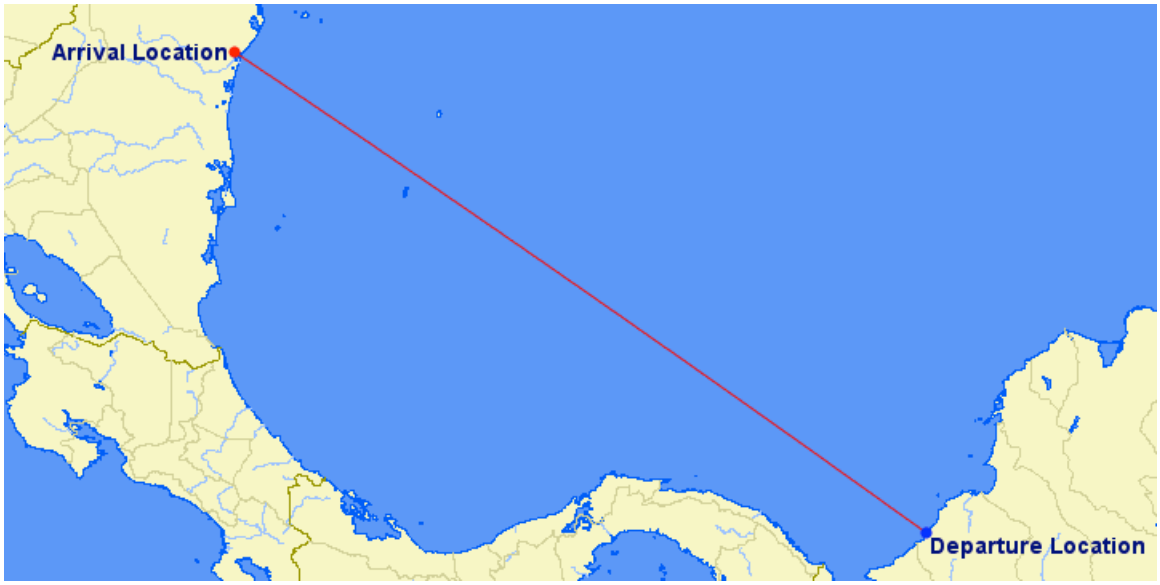
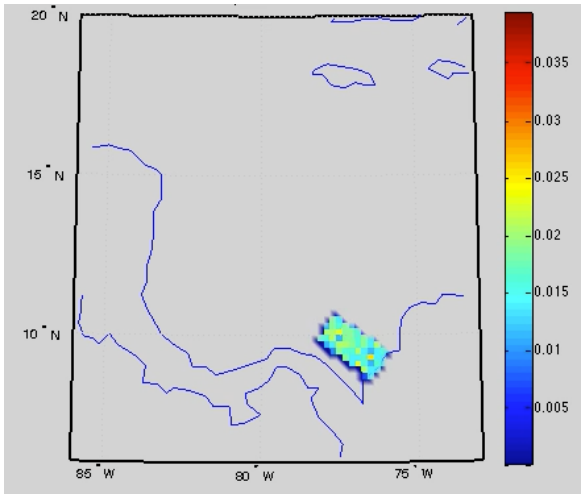


Figure 4.17: Illustration of a target's track from departure to arrival location. The target follows a direct path at a constant speed to illustrate the simplest path-finding case.

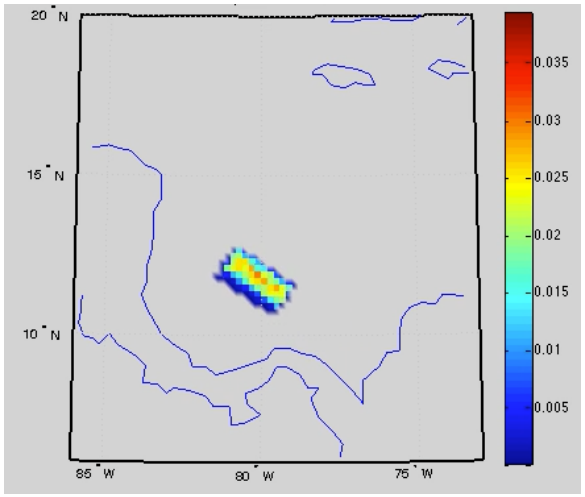
Approximate payload	5,000 kg
Speed	15 knots
Origin	76.3° W 9° N
Departure uncertainty	±30 nm, Uniform
Destination	83.5° W 14° N
Arrival uncertainty	±30 nm, Uniform
Departure time	33 hours
Departure time uncertainty	±4 hours, Uniform
Intel certainty	95%

Table 4.3: Parameters for a simple straight path from origin to destination with constant velocity. Uncertainty data is listed for departure and arrival locations and departure time.

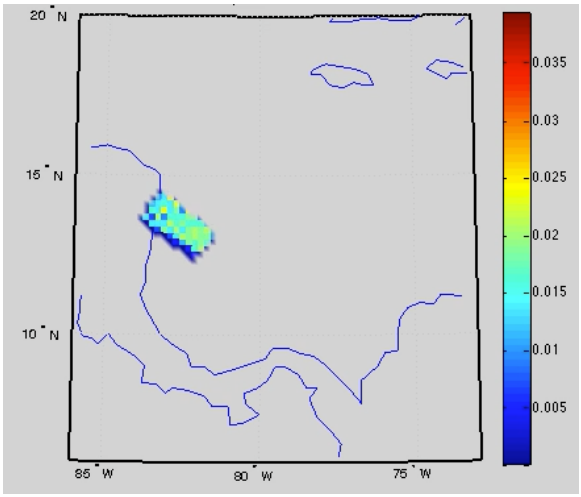
In this case, the probability model generates the heat map for 78 hours using the associated parameters from Table 4.3. In the context of this thesis, hour 1 represents the first simulated time step and hour 78 represents the last time step. This results in 78 probability data files, one for each time step, that we read into the path-finding model. Each file contains a probability for each grid cell in the AOR corresponding to the likelihood the target is in that grid cell at the associated time. The path-finding model analyzes the probabilities to establish the times listed in Table 4.2 using Equations 4.8 through 4.13. Table 4.4 lists the resulting times.



(a) Hour 37



(b) Hour 52



(c) Hour 64

Figure 4.18: A simple one target scenario produced in the enhanced probability model. The snapshots are at the latest departure time (hour 37), an intermediate point (hour 52), and at the earliest arrival time (hour 64).

Once the first departure time (τ_j^{min}) and last arrival time (τ_j^{max}) are known, the path-finding model proceeds by applying MARS to the probability data between these times. The resulting form of the MARS model is

$$\hat{x}_t = 280.42 - 0.205(t-49)_+ + 0.205(49-t)_+ \tag{4.14}$$

and

Time	Actual	Model Result
$\tau^{earliestdep}$	29 hours	29 hours
First τ^{min}	37 hours	37 hours
Last τ^{max}	64 hours	64 hours
$\tau^{latestarr}$	NA	72 hours
$\tilde{\tau}$	8 hours	8 hours
τ	33 hours	33 hours

Table 4.4: Case time information from path-finding model as compared to the actual inputs for a target traveling along a straight path.

$$\hat{y}_t = 12.13 + 0.142(t-55)_+ - 0.142(55-t)_+. \tag{4.15}$$

From Equations 4.14 and 4.15, we see MARS introduces the basis functions $(t-49)_+$ and $(49-t)_+$ in the x direction and $(t-55)_+$ and $(55-t)_+$ in the y direction. Further examination reveals the coefficients associated with the reflected pairs have equal magnitude but opposite sign. This is a result of the MARS model fitting procedure. Continuing the path-finding algorithm with this model results in a track composed of three legs but the same speed and heading on each leg. In this particular case, it is easy to determine that these points are not truly changes in speed or heading. In a more general case, it is necessary to determine if a point introduced by MARS is really a change in speed or heading.

We apply a filter to the knots or cuts introduced by MARS. The filter first examines the change in slope at each reflected pair. If the absolute difference in slope for both models changes by less than 0.01 the filter does not add them to the expected track and path-finding continues with the expected target locations as defined in Section 4.2. In cases where the reflected pairs or cuts are legitimate, we add an additional data field to the expected track data. This data field is identified by the associated time and populated with the result of the basis function for each time step in the path. For the moment let us consider the reflected pair at hour 49 as a legitimate time to split the journey in the x direction into two legs. The expected track data would then get an additional data field containing the result of $(t-49)_+$ and the final regression model is constructed with the new data.

In the current case, the filter determines that the points at time 49 and 55 are not valid based on the change in slope. These are not added to the expected data and the algorithm continues the

path-finding procedure using the expected locations in the x and y directions. At this point of the procedure, we know the target is traveling from the departure location to the arrival location on a straight path at a constant speed. Recall that the optimization model requires a constant AOU so we proceed to determine the AOU throughout the target's path.

The associated plot of the AOU and the resultant MARS model fit is shown in Fiugre 4.19. The overall change in area during the leg is less than the selected threshold of 25%, thus the AOU is chosen as the maximum area during the path. In this case, we capture 90% or more of the probability mass during the entire journey.

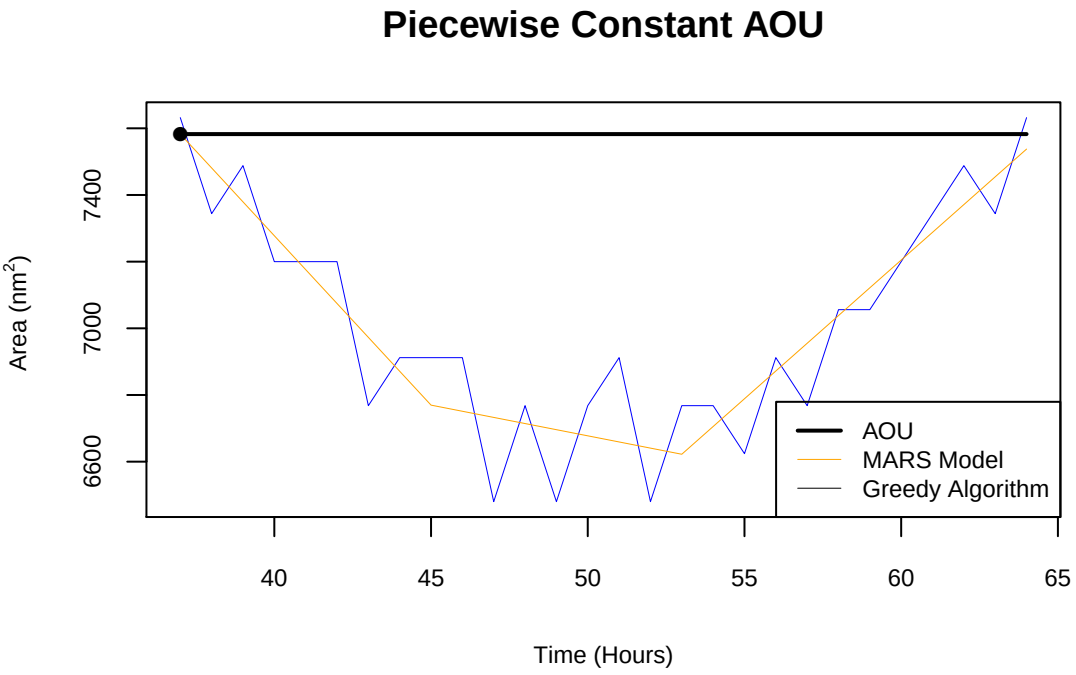


Figure 4.19: Resulting AOU throughout one target's straight path and associated MARS model fit to the greedy area data.

With the entire track examined for speed, heading, and area changes, we proceed to build the final ordinary least squares model using the expected location and time data. For this particular case with no waypoints and no changes in the AOU, the final model takes the following form

$$\hat{x} = 290.5 - 0.205t \tag{4.16}$$

and

$$\hat{y} = 4.3 + 0.142t. \tag{4.17}$$

Based on the times in Table 4.4 and Equations 4.16 and 4.17, we determine the remaining parameters utilizing the equations described in Section 4.4. Table 4.5 compares the results of the path-finding model with the associated inputs to the probability model. The path-finding model determines the target's departure location, speed, and associated departure and arrival times extremely well. When comparing the destinations, remember that the path-finding model uses the probability data for the time periods when the probabilities sum to one. Thus, the model determines an arrival location that is actually on the water vice on the coast as depicted by the actual location. An accurate estimate of the origin is required for the optimization model, but an accurate estimate of the destination is not required as long as its along the same track.

Parameter	Actual	Model Result
Speed	15 knots	14.981 knots
Origin	76.3° W 9° N	76.3° W 9° N
Destination	83.5° W 14° N	82.657° W 13.414° N
AOU	7200 nm^2	7582.65 nm^2
$\tau^{earliestdep}$	29 hours	29 hours
τ^{min}	37 hours	37 hours
τ^{max}	64 hours	64 hours
$\tilde{\tau}$	$\pm$4 hours	$\pm$4 hours
τ	33 hours	33 hours

Table 4.5: Comparison of actual parameters and path-finding model results for the simple target scenario.

Figure 4.20 shows the final model fit to plots of the expected track through the AOR and over time.

4.6 Multiple Leg Model and Illustration

Recall the target with two legs from Figure 4.1. Table 4.6 lists the associated parameters that we enter into the enhanced path-finding model.

Again we show in Figure 4.21 several hours of the probability mass as it moves through the AOR. The probability model generates a heat map for 120 hours of the target's journey and the

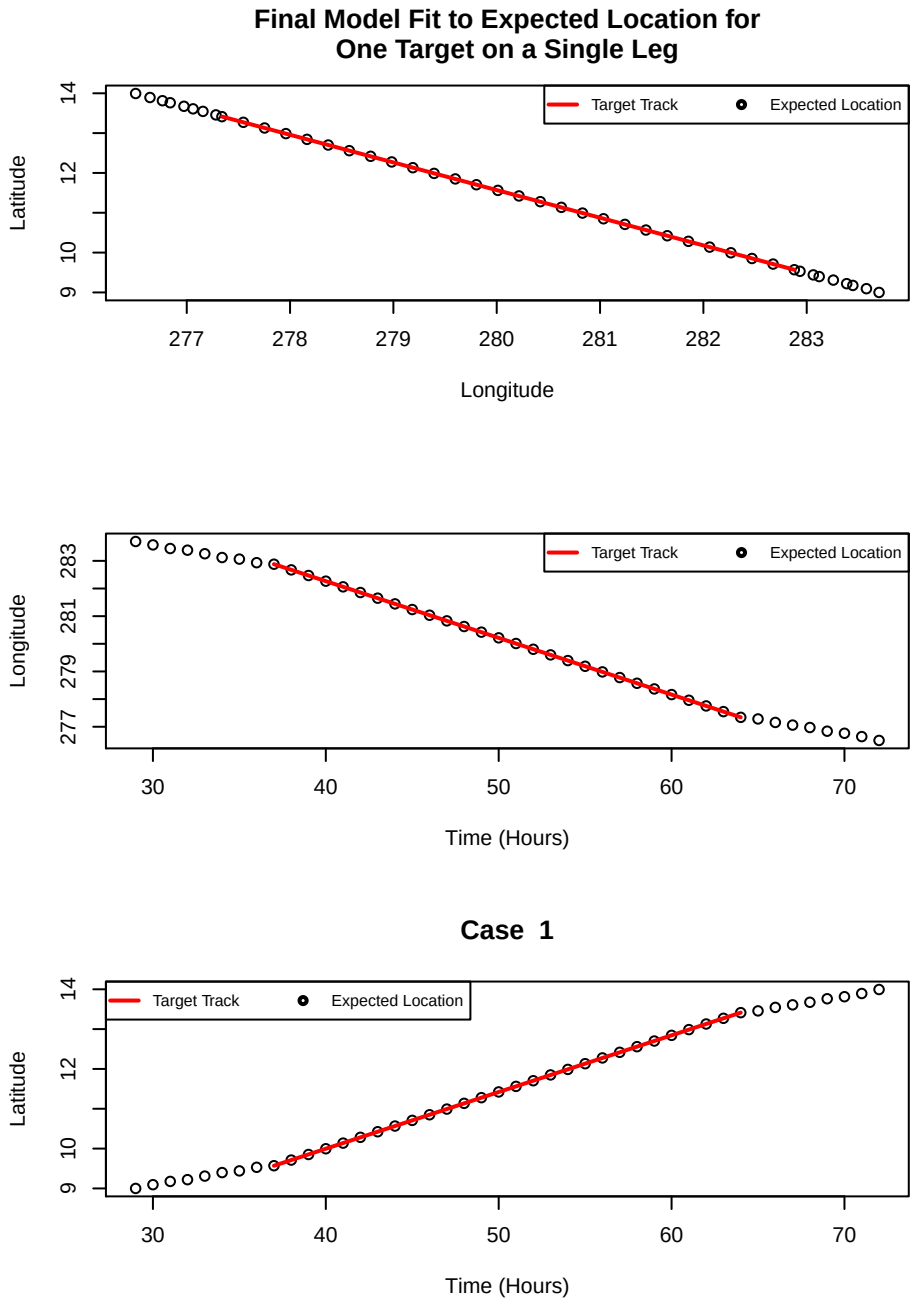


Figure 4.20: Output of the path-finding model, showing the target's track through the AOR and the target's track in both the x and y direction with respect to time.

resulting probability files feed into the path-finding model. Based on the probability data, we determine the initial and final departure and arrival times which are listed in Table 4.7.

Approximate payload	2,500 kg
Speed	15 knots
Origin	77° W 5° N
Departure uncertainty	±30 nm, Uniform
Waypoint	86° W 3° N
Waypoint uncertainty	±60 nm
Destination	92.2° W 14.5° N
Arrival uncertainty	±50 nm, Uniform
Departure time	10 hours
Departure time uncertainty	±3 hours, Uniform
Intel certainty	50%

Table 4.6: Parameters for a multiple leg path from origin to destination. Uncertainty data and associated distributions are listed for departure location, arrival location, and departure time.

Time	Actual	Model Result
$\tau^{earliestdep}$	7 hours	7 hours
First τ^{min}	13 hours	13 hours
Last τ^{max}	96 hours	91 hours
$\tau^{latestarr}$	NA	107 hours
$\tilde{\tau}$	6 hours	6 hours
τ	10 hours	10 hours

Table 4.7: Case time information from path-finding model as compared to the actual inputs for a target traveling along a path with multiple legs.

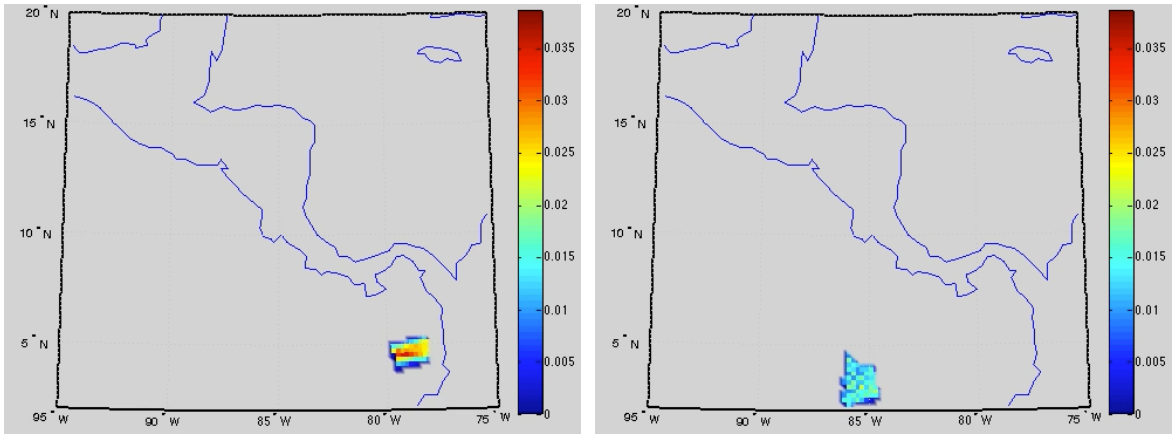
Next, we proceed with the application of MARS between time 13 and time 91. The resulting form of the MARS model is

$$\hat{x_t} = 274.36 - 0.123(t-45)_+ + 0.249(45-t)_+ \tag{4.18}$$

and

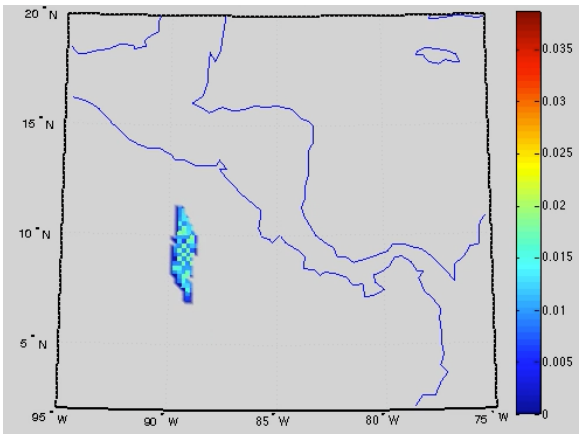
$$\hat{y_t} = 2.77 + 0.052(t-37)_+ + 0.170(t-49)_+ + 0.058(49-t)_+. \tag{4.19}$$

In this case, MARS introduces waypoints at hours 37, 45, and 49. We examine the validity of each waypoint by looking at the change in slope and the maximum time for the entire probability



(a) Time step 18

(b) Time step 45



(c) Time step 75

Figure 4.21: A one target, multiple leg scenario produced in the enhanced probability model. Snapshots are from various times during the target's route.

mass to travel through the respective waypoint. This time is then compared to the amount of time until the next waypoint. If the maximum time is greater than the time to the next waypoint, then we discard the next waypoint and assume it is a artifact of the probability mass moving through a turn.

Here, the change in slope is significant, thus solidifying that we have at least one valid waypoint. Next we determine which of the three are valid. We assume that the first waypoint, chronologically, is valid and check if the maximum time to move through the turn is less than the time to the next waypoint. Figure 4.22 illustrates the probability mass at hour 45. The current waypoint (at hour 45) and the next waypoint at hour 49 are highlighted. We desire to determine the length of time for this probability mass to move through the hour 45 waypoint.

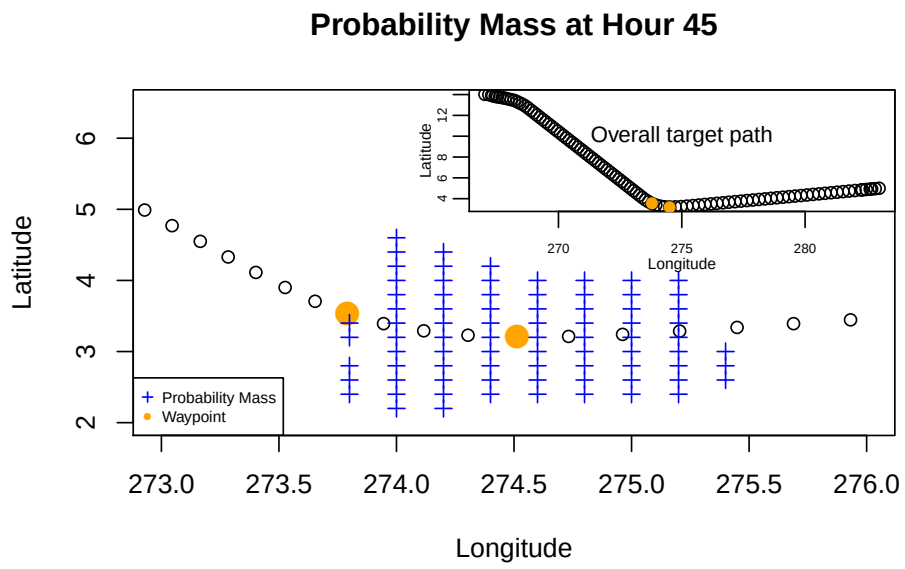


Figure 4.22: Depiction of the probability mass in the one target multiple leg case. The probability mass is shown at hour 45, which is the time of the current waypoint. We rotate this probability mass to align with the target's course en route to the next potential waypoint.

Specifically, we rotate the probability mass along the target's path to determine the maximum distance between x coordinates. This rotation aligns the probability mass with the target's track, allowing us to use the target's overall speed in the time calculation. This distance is then divided by target speed to produce the time to complete the turn. At hour 37, the time to complete the turn is 6.8 hours, which is less than the time to the next waypoint (8 hours), thus the waypoint at hour 45 is kept. The turn time at hour 45 is 7.4 hours and is greater than the time to the next waypoint (4 hours), therefore we discard hour 49 as a waypoint and conclude the target path consists of three legs based on speed and heading changes. Figure 4.23 illustrates the rotated probability mass and the associated horizontal distance. The final path based on constant speed and heading legs contains two valid waypoints with the first at hour 37 and the second at hour 45.

The AOU analysis follows which determines if the path requires additional legs based on significant changes in the AOU. The associated plot of the AOU and the resultant MARS model fit is shown in Figure 4.24. In this case, we analyze each leg separately using MARS.

The resultant MARS models fit three different linear piecewise segments to the area data. The

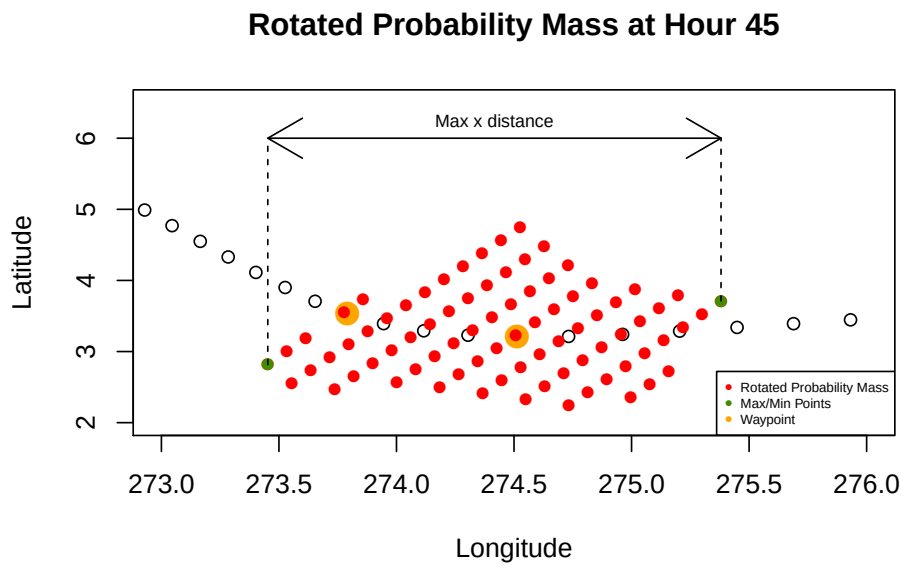


Figure 4.23: Depiction of the probability mass in the one target multiple leg case. The rotated probability mass is shown at hour 45, which is the time of the current waypoint. The rotated probability mass aligns with the target's course, which allows us to calculate the time for the target to move the maximum x distance using overall target speed.

first segment begins at hour 13 and ends at hour 37. During this time, the area changes by more than 25%, thus the model divides the constant speed heading leg into two legs based on the change in area. The area on the other two legs, from 37 to 45 and from 45 to 91, changes by less than 25% and are not further divided. Figure 4.25 depicts the associated times and areas of the new legs.

With the addition of new legs based on area changes, the path now consists of four legs. The new leg does not change the speed or heading of the target but only represents the location that the AOU increases by more than 25%. As a result, the final model accounts for the speed and heading changes identified by MARS but it is unnecessary to insert the points of area change. The final path-finding model is of the form

$$\hat{x_t} = 285.4 - 0.242t - 0.024(t-37)_+ + 0.145(t-45)_+ \tag{4.20}$$

and

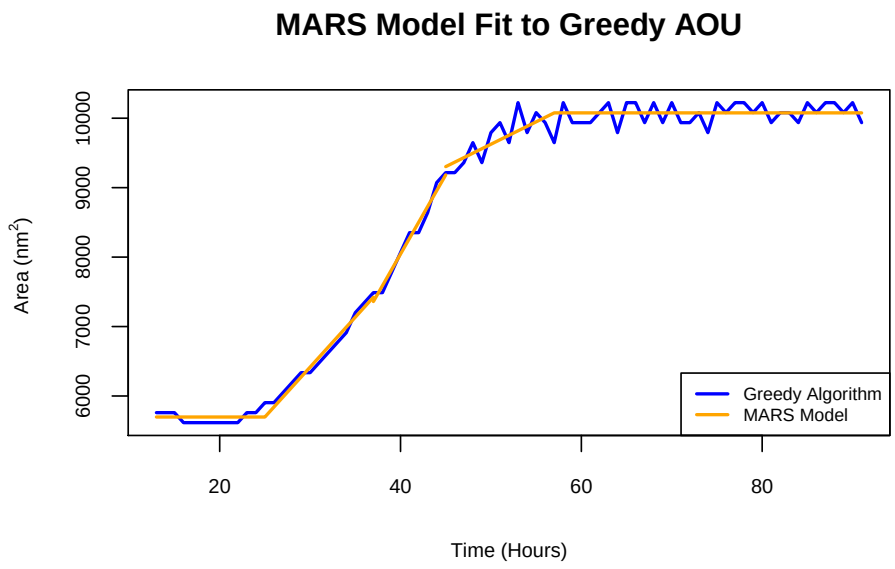


Figure 4.24: Resulting AOU throughout one target's multiple leg path and associated model fit to the greedy area data.

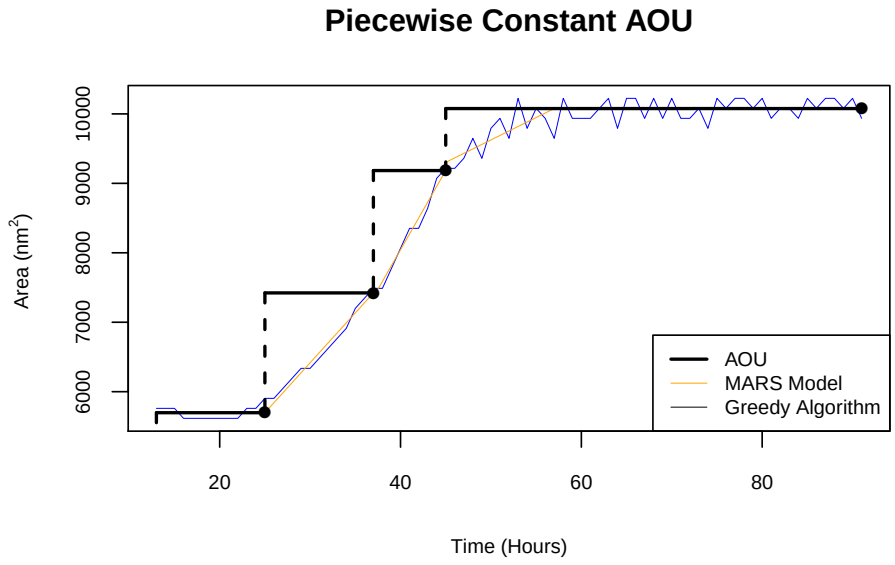


Figure 4.25: Resulting piecewise constant area for one target multiple leg case.

$$\hat{y}_t = 5.47 - 0.050t - 0.055(t-37)_+ + 0.320(t-45)_+. \tag{4.21}$$

The mechanics to determine the remaining case parameters are the same as those in Section 4.5 and Table 4.8 lists the model output and the actual parameters. Figure 4.26 shows the final regression model fit to the expected track over time.

	Leg 1		Leg 2	
Parameters	Actual	Model	Actual	Model
Speed	15 knots	14.843 knots	NA	14.843 knots
Origin	77° W 5° N	77.015° W 4.964° N	NA	80.649° W 4.211° N
Destination	86° W 3° N	80.649° W 4.211° N	NA	83.555° W 3.608° N
AOU	5400	5697	NA	7422
τ^{min}	13	13	NA	25
τ^{max}	46.88	25	NA	37
$\tilde{\tau}$	6	6	NA	6
τ	10	10	NA	25
	Leg 3		Leg 4	
Parameters	Actual	Model	Actual	Model
Speed	NA	17.175 knots	15 knots	14.754 knots
Origin	NA	83.555° W 3.608° N	86° W 3° N	85.684° W 2.763° N
Destination	NA	85.684° W 2.763° N	92.2° W 14.5° N	91.264° W 12.603° N
AOU	NA	9184	10800	10077
τ^{min}	NA	37	46.878	45
τ^{max}	NA	45	96.138	91
$\tilde{\tau}$	NA	6	6	6
τ	NA	37	46.878	45

Table 4.8: Comparison of actual parameters and path-finding model results for one target multiple leg scenario.

4.7 Results

The core effort of this thesis concerns the need to provide target parameters to the optimization model in Section 2.3. Thus far, we demonstrate the capability to take intelligence data concerning departure and arrival locations, speed, and departure time to produce a probabilistic heat map over time. With the heat map as input, we show the ability to determine the necessary target parameters to build a search vehicle allocation plan. This section briefly describes the results of the optimization model with our input parameters verse the results using the actual target parameters for several target scenarios.

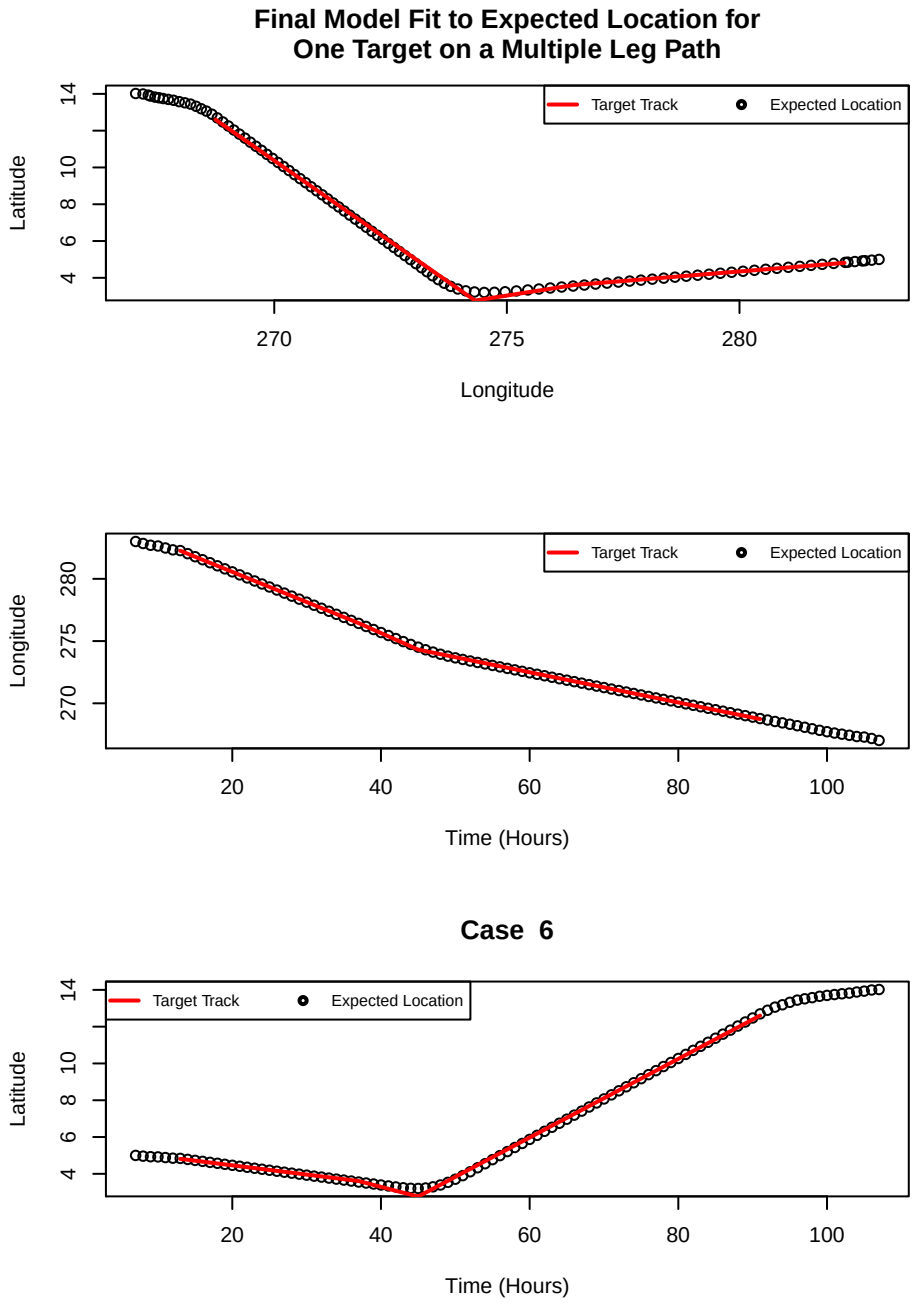


Figure 4.26: Output of the path-finding model for a multiple leg scenario, showing the target's track through the AOR and the target's track in both the x and y directions with respect to time.

4.7.1 Optimization Results

Consider a scenario with two targets, where target 1 is an SPSS in the eastern Pacific and target 2 is a go-fast in the Caribbean Sea. Target 1 follows a track consisting of two legs and target two

is on a track with a single leg. Figure 4.27 depicts the actual track and the track approximated by the path-finding model.

Figure 4.27: Case scenario with two targets where target 1 is a SPSS operating in the eastern Pacific and target 2 is a go-fast operating in the Caribbean Sea.

Figure 4.28 depicts the optimal search plans based on the actual track and the results of our path-finding model. In both cases the searcher leaves from home base and travels to search for target 2 and then to search for target 1. The searcher's track follows the points in sequential order, starting and ending at home. In each instance the search locations and times of the search are nearly identical, in addition to the expected amount of drugs found. The fact that the two search plans are so similar suggests that our path-finding model adequately determines the target's track since we are comparing optimal search plans with the parameters input into the enhanced probability model and the output from our path-finding model. This example illustrates a key link between our probability and path-finding models with the search vehicle optimization model.

Figure 4.28: The optimal search plan based on the results of the path-finding model and the actual track. The two search plans are nearly identical. The searcher leaves from home base and travels to search for target 2 and the to search for target 1. The searcher's track follows the points in sequential order starting and ending at home. The green track represents the search plan based on the true intelligence inputs, and the track in orange represents the search plan based on our path-finding model.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 5:
Conclusion

Illicit drug-trafficking is a major concern of the United States and continued expansion of drug-trafficking significantly affects the national security interests of the United States. This thesis addresses a particular aspect of illegal drug-trafficking in the eastern Pacific and Caribbean Sea. We focus on taking intelligence information concerning a drug trafficker's transit route to create a spatial temporal probability map and subsequently determine the trafficker's most likely route through the region. Last, we address the accuracy of our results by comparing the optimal search plan with our path-finding model as input and with the true case parameters as input to the optimization model.

We begin with a fundamental probability model to account for uncertainty in the departure and arrival locations of a single drug trafficker who travels along a linear path. Next we enhance this model to account for additional uncertainty in the trafficker's speed and departure time. Additionally, we build in functionality to handle multiple drug traffickers, multiple courses and speeds, and the ability to output results showing the expected amount of drugs.

With uncertainty in the target's transit route represented by several random variables, we create a probability map of the target's track over time. Using these probabilities, we apply regression techniques to determine the characteristics of the target's track needed to run a search vehicle optimization model. The optimization model requires inputs for each segment of the target's path where on each segment the target is traveling along a constant heading, with a constant speed, and a constant AOU. Departure time and arrival time on each segment identifies where one segment ends and the next begins. Segment changes occur based on changes in speed, heading, or AOU. Together these route characteristics represent the inputs to the search vehicle optimization model. Our procedure results in an optimal search plan sufficiently close to one developed using the probability model inputs.

We illustrate the models' functionality with two case studies of varying complexity. First, a simple straight path between the origin and destination shows the path-finding model's ability to identify the target's starting location, destination, speed, and associated travel times. Combining these route characteristics, we are able to reproduce the target's likely track from start to finish. Supplying our path-finding results into the search vehicle allocation model shows similar search

plan results. Next, we demonstrate a more complex scenario with a target traveling to the destination with a course change during the journey. Our model accurately identifies this course change and reproduces the target's track. Again, the resulting search plans are quite similar. These two case studies are representative of actual drug trafficker routes but additional analysis is needed to validate the model's ability.

Our probability model and path-finding model establish additional areas of research to aid JIATF South in their mission. We address the issue surrounding the AOU over time with three different methods where the chosen method involves a greedy algorithm. While this method results in areas much smaller than the other two approaches, our overall approach was relatively informal and further research could apply rigorous mathematical methods to the AOU problem.

We examine several cases with various target characteristics and compare our path-finding model results to the original probability model inputs. Our limited comparisons demonstrate promising results but a more formal comparison is necessary to uncover any modeling shortfalls that may be present. Additionally, we compare a few search vehicle plans using our path-finding model and the original case inputs. Again, a more thorough investigation should be undertaken, possibly involving design of experiments or other factors to determine where the model might not perform as well.

We do not address cases with a large departure uncertainty. These are the cases referred to as "channel" cases in this thesis. They consist of scenarios where a target has a large departure window, on the order of ± 24 hours or more. The probability model creates a long channel of probability that grows from the origin and reaches the destination, effectively connecting the two locations. In other words, the probability mass is never completely on the water and the target can be anywhere along the channel from the earliest departure time to the latest arrival time. The optimization model handles these cases in a different manner and requires different inputs, thus follow-on analysis to develop different algorithms considering these cases is necessary.

Our probability model does not account for the impact of weather on the drug trafficker's movements. A comparison of the path-finding model results utilizing input from our model and one with weather characteristics would highlight whether or not those items are important factors for consideration.

REFERENCES

An, W., Ayala, D. F. M., Sidoti, D., Mishra, M., Han, X., Pattipati, K. R., Regnier, E. D., Kleinman, D. L., & Hansen, J. A. (2012). Dynamic asset allocation approaches for counter-piracy operations. In *2012 15th International Conference on Information Fusion*, 1284–1291. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6289955

Bessman, D. L. (2010, September). *Optimal interdiction of an adaptive smuggler*. (Master's thesis, Naval Postgraduate School). Retrieved from http://hdl.handle.net/10945/5129

Fraser, D. M. (2010). Command strategy 2020 partnership for the Americas. Retrieved from http://www.comisede.com/paises/EEUU/PENTAGONO/Command

Friedman, J. H. (1991). Multivariate adaptive regression splines. *The Annals of Statistics*, 19(1), 1–67. Retrieved from http://projecteuclid.org

Gift, P. (2010, September). *Planning for an adaptive evader with application to drug interdiction operations*. (Master's thesis, Naval Postgraduate School). Retrieved from http://hdl.handle.net/10945/5161

Government Accountability Office (GAO). (2009). *Drug control: U.S. counternarcotics cooperation with Venezuela has declined* (GAO-09-806). Washington, DC: Author. Retrieved from http://www.gao.gov/products/GAO-09-806

Hansen, J., Jacobs, G., Hsu, L., Dykes, J., Dastugue, J., Allard, R., Barron, C., Lalejini, D., Abramson, M., Russel, S., & Mittu, R. (2011). Information domination: Dynamically coupling metoc and intel for improved guidance for piracy interdiction. *2011 NRL Review*, 109–115. Retrieved from http://www.nrl.navy.mil/media/publications/nrl-review/

Hastie, T. (1984, November). *Principle curves and surfaces*. (Doctoral dissertation, Stanford University). Retrieved from http://www.slac.stanford.edu/pubs/slacreports/slac-r-276.html

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). New York, NY: Springer.

Huebner, K., Ruthotto, S., & Kragic, D. (2008). Minimum volume bounding box decomposition for shape approximation in robot grasping. In *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*, 1628–1633. doi:10.1109/ROBOT.2008.4543434

Johnson, D. E. & Cohen, E. (1998). A framework for efficient minimum distance computations. In *Robotics and Automation, 1998. Proceedings. 1998 IEEE International Conference on*, (4), 3678–3684. doi:10.1109/ROBOT.1998.681403

Koopman, B. O. (1946). *Search and screening* (11000056.00). Retrieved from http://www.cna.org/sites/default/files/research/1100005600.pdf

Kraul, C. (2010, July 6). Ecuador police seize 100-foot narco-submarine being built secretly. *Los Angeles Times*. Retrieved from http://articles.latimes.com/2010/jul/06/world/la-fg-ecuador-narco-sub-20100706

Lempitsky, V., Kohli, P., Rother, C., & Sharp, T. (2009). Image segmentation with a bounding box prior. In *Computer Vision, 2009 IEEE 12th International Conference on*, 277–284. doi:10.1109/ICCV.2009.5459262

National Drug Intelligence Center (NDIC). (2008, December). *National drug threat assessment 2009* (2008-Q0317-005). Retrieved from http://www.justice.gov/archive/ndic/pubs31/31379/31379p.pdf

National Drug Intelligence Center (NDIC). (2011, August). *National drug threat assessment 2011* (2011-Q317-001). Retrieved from http://www.justice.gov/archive/ndic/pubs44/44849/44849p.pdf

National Security Staff (2011). Strategy to combat transnational organized crime: Addressing converging threats to national security. Retrieved from http://www.whitehouse.gov/administration/eop/nsc/transnational-crime

Office of National Drug Control Policy (2013). Transit zone operations. Retrieved from http://www.whitehouse.gov/ondcp/transit-zone-operations

Pfeiff, D. M. (2009, June). *Optimizing employment of search platforms to counter self-propelled semi-submersibles*. (Master's thesis, Naval Postgraduate School). Retrieved from http://hdl.handle.net/10945/4697

Pietz, J. (2013, September). *A generalized orienteering problem for optimal search and interdiction planning*. (Doctoral dissertation, Naval Postgraduate School). Retrieved from email.

Pietz, J. & Royset, J. O. (2013). Generalized orienteering problem with resource dependent rewards. *Naval Research Logistics*, 60(4), 294–312. doi:10.1002/nav.21534

Posture statement of General Douglas M. Fraser, United States Air Force, Commander, United States Southern Command before the House Armed Services Committee, 112th Cong., (2012). Retrieved from http://www.southcom.mil/newsroom/Documents/SOUTHCOM_2012_Posture_Statement.pdf

Posture statement of General John F. Kelly, United States Marine Corps, Commander, United States Southern Command before the 113th Congress House Armed Services Committee, 113th Cong., (2013). Retrieved from http://www.southcom.mil/newsroom/Pages/2013-Posture-Statement-to-Congress.aspx

Rozen, N. (2009, December). *Sensor-interceptor operational policy optimization for maritime interdiction missions*. (Master's thesis, Naval Postgraduate School). Retrieved from http://hdl.handle.net/10945/4370

Seelke, C. R., Wyler, L. S., Beittel, J. S., & Sullivan, M. P. (2011). *Latin America and the Caribbean: Illicit drug trafficking and U.S. counterdrug programs*. (R41215). Retrieved from http://www.fas.org/sgp/crs/row/R41215.pdf

United States Coast Guard (USCG). (2009). Dangerous self-propelled semi submersibles proliferating rapidly. Retrieved from http://www.uscg.mil/history/allen/messages/message25.pdf

United States Department of State (2012). *2012 international narcotics control strategy report*. Retrieved from http://www.state.gov/j/inl/rls/nrcrpt/2012/

United States Senate Caucus on International Narcotics Control (2012). *Preventing a security crisis in the Caribbean.* Retrieved from http://www.feinstein.senate.gov/public/index.cfm/files/serve/?File_id=90bb66bc-3371-4898-8415-fbfc31c0ed24

United States Southern Command (2013). Countering transnational organized crime. Retrieved from http://www.southcom.mil/ourmissions/Pages/Countering%20Transnational%20Organized%20Crime.aspx

Washburn, A. R. (2002). *Search and detection*. Linthicum, MD, Institute for Operations Research and the Management Sciences.

THIS PAGE INTENTIONALLY LEFT BLANK

Initial Distribution List

1. Defense Technical Information Center
   Ft. Belvoir, Virginia
2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California