The Impact of Threat Levels at the Casualty Collection Point on Military Medical Evacuation System Performance

THESIS

Nathaniel C. Dennie, 1st Lt, USAF

AFIT-ENS-MS-21-M-154

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENS-MS-21-M-154

THE IMPACT OF THREAT LEVELS AT THE CASUALTY COLLECTION POINT ON MILITARY MEDICAL EVACUATION SYSTEM PERFORMANCE

THESIS

Presented to the Faculty

Department of Operational Sciences

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

in Partial Fulfillment of the Requirements for the

Degree of Master of Science in Operations Research

Nathaniel C. Dennie, BS

1st Lt, USAF

March 25, 2021

AFIT-ENS-MS-21-M-154

THE IMPACT OF THREAT LEVELS AT THE CASUALTY COLLECTION POINT ON MILITARY MEDICAL EVACUATION SYSTEM PERFORMANCE

THESIS

Nathaniel C. Dennie, BS
1st Lt, USAF

Committee Membership:

Capt Phillip R. Jenkins, PhD
Chair

Dr. Matthew J. Robbins
Member

AFIT-ENS-MS-21-M-154

Abstract

One of the primary duties of the Military Health System is to provide effective and efficient medical evacuation (MEDEVAC) to injured battlefield personnel. To accomplish this, military medical planners seek to develop high-quality dispatching policies that dictate how deployed MEDEVAC assets are utilized throughout combat operations. This thesis seeks to determine dispatching policies that improve the performance of the MEDEVAC system. A discounted, infinite-horizon continuous-time Markov decision process (MDP) model is developed to examine the MEDEVAC dispatching problem. The model incorporates problem features that are not considered under the current dispatching policy (e.g., myopic policy), which tasks the closest-available MEDEVAC unit to service an incoming request. More specifically, the MDP model explicitly accounts for admission control, precedence level of calls, different asset types (e.g., Army versus Air Force helicopters), and threat level at casualty collection points. An approximate dynamic programming (ADP) algorithm is developed within an approximate policy iteration algorithmic framework that leverages kernel regression to approximate the state value function. The ADP algorithm is used to develop high-quality solutions for large scale problems that cannot be solved to optimality due to the *curse of dimensionality*. We develop a notional scenario based on combat operations in southern Afghanistan to investigate model performance, which is measured in terms of casualty survivability. The results indicate that significant improvement in MEDEVAC system performance can be obtained by utilizing either the MDP or ADP generated policies. These results inform the development and implementation of tactics, techniques and procedures for the military medical planning community.

This research is dedicated to the men and women, past, present and future, who have or will dedicate their lives in support of our country. May this research help lay the ground work needed to provide the best medical evacuation system possible for those who risk their lives in the defense of our country.

Acknowledgements

I would like to express my deep gratitude to Dr. Jenkins for his leadership and support. His mentorship and guidance over this past year has helped to shape me as a better analyst andofficer. I am extremely grateful that I was afforded the opportunity to work with him on this topic.

I also want to acknowledge my committee member, Dr. Matthew Robbins. His guidance and encouragement provided me the confidence to pursue this area of research. Without him none of this would be possible.

Finally I would like to thank family. Their on-going encouragement and support has given me the drive and motivation to conquer all tasks that are before me. My mother has always encouraged me to be persistent, and this lesson has served me well throughout this entire process. I would especially like to thank my cousin, who would take time away from his own Ph.D. studies to read and edit my drafts. My family is the reason I am here, and I would not be where I am without them.

Nathaniel C. Dennie

Table of Contents

Page

Abstract . . . iv

Dedication . . . v

Acknowledgements . . . vi

List of Figures . . . ix

List of Tables . . . x

I. Introduction . . . 1

1.1 Motivation . . . 1
1.2 Background . . . 2
1.3 Thesis Overview . . . 3

II. Literature Review . . . 5

2.1 EMS Systems . . . 5
2.1.1 Civilian . . . 7
2.1.2 Military . . . 9
2.2 Thesis Contribution . . . 12

III. Problem Description . . . 14

3.1 The MEDEVAC Dispatching Process . . . 14

IV. Methodology . . . 21

4.1 MDP Formulation . . . 21
4.1.1 Decision Epochs . . . 23
4.1.2 State Space . . . 24
4.1.3 Action Space . . . 25
4.1.4 Transitions . . . 27
4.1.5 Transition Probabilities . . . 28
4.1.6 Rewards . . . 28
4.1.7 Optimality Equation . . . 29
4.2 ADP Formulation . . . 30
4.2.1 Kernel Regression . . . 32
4.2.2 API . . . 33

Page

V. Testing, Results, & Analysis 37

5.1 Representative Scenario 37
5.2 Representative Scenario Results 41
5.2.1 Algorithmic Experimental Design 43
5.2.2 Policy Comparison 45
5.3 Excursions 49
5.3.1 Excursion 1 - Request Arrival Rate 49
5.3.2 Excursion 2 - MEDEVAC Asset Types 51
5.3.3 Excursion 3 - Threat Level Proportion 52
5.4 12-zone Scenario Results 53
5.5 34-zone Scenario Results 59
5.5.1 Algorithm Modification 60
5.5.2 34-zone Problem Results 62

VI. Conclusions & Recommendations 64

6.1 Conclusions 64
6.2 Recommendations for Future Research 65
Bibliography 67

List of Figures

Figure		Page
1	MEDEVAC Mission Timeline	17
2	Representative Scenario	38
3	6-zone Scenario	39
4	ADP Algorithm Performance	44
5	MEDEVAC Busy Rates by Policy	48
6	Arrival Rate Impact on MEDEVAC System	50
7	High Threat Proportion Impact on MEDEVAC System Performance	53
8	12-zone Scenario	54
9	MEDEVAC Busy Rates by Policy	58
10	34-zone Scenario	59

List of Tables

Table		Page
1	Evacuation Precedence Categories (Department of the Army, 2019)	16
2	Threat Levels	16
3	Threat level 1 (low threat) proportion of arrivals	40
4	Threat level 2 (high threat) proportion of arrival	40
5	MEDEVAC Response Times with low Threats	40
6	MEDEVAC Service Times with low Threats	41
7	Baseline parameter settings	42
8	Policy Comparison	43
9	Experimental Design Factor Levels	43
10	Experimental Design Results	45
11	Value Comparison	45
12	Scenario Description	46
13	Optimal Decisions	47
14	MEDEVAC Busy Rates	48
15	Arrival Rate Impact on $V^{\pi}(S_0)$ for the MEDEVAC System	50
16	Percent Increase over Myopic policy with Asset Replacement	52
17	Percent Increase over Myopic policy with respect to High Threat Proportion	52
18	12-zone parameter settings	55
19	12-zone Results	55
20	12-zone Optimal Decisions	57

Table		Page
21	MEDEVAC Busy Rates	58
22	34-zone parameter settings	60
23	34-zone Algorithm Settings	62
24	34-zone Results	62

THE IMPACT OF THREAT LEVELS AT THE CASUALTY COLLECTION POINT ON MILITARY MEDICAL EVACUATION SYSTEM PERFORMANCE

I. Introduction

1.1 Motivation

One of the primary objectives of the military's emergency medical service (EMS) response system is to evacuate injured personnel from the battlefield as quickly as possible. There are two options available to accomplish this task: (1) medical evacuation (MEDEVAC) and (2) casualty evacuation (CASEVAC). The first and most preferred option utilizes dedicated platforms (e.g., HH-60M Black Hawk helicopters) staffed with trained medical personnel that can effectively tend to patients while in transit to a medical treatment facility (MTF). The latter option is primarily utilized as a contingency and typically does not have medical personnel on board to provide the critical care needed while en route to an MTF (Department of the Army, 2019).

Sequential resource allocation decision-making within the military MEDEVAC system consists of determining which MEDEVAC unit (if any) to dispatch in response to a casualty event; this is commonly referred to as the MEDEVAC dispatching problem (Robbins *et al.*, 2018). These decisions are complicated due to the inherent uncertainties found within a MEDEVAC system (e.g., response time, service times, and request arrival rate). This thesis seeks to solve the MEDEVAC dispatching problem (i.e., determine an optimal MEDEVAC dispatching policy) via a Markov decision process (MDP) model.

The effectiveness of the MEDEVAC system enables the execution of combat operations in a deployed environment. For this reason, improvements are constantly being made. This thesis further develops the MEDEVAC dispatching literature by explicitly modeling casualty collection point (CCP) threat levels. Moreover, this research is the first to model a joint environment wherein both Air Force and Army MEDEVAC units are available to respond to requests for service. The results provide insights to the military medical planning community and inform the development of tactics, techniques and procedures for MEDEVAC operations.

1.2 Background

The concept of removing casualties from the battlefield was introduced during the American Civil War, and the practice of evacuating battlefield casualties has continually improved throughout each major United States (U.S.) conflict. Such evolution consists of using horse-drawn wagons for CASEVAC in the Spanish American War in 1898, motorized ground-vehicles for CASEVAC in World War I, helicopters for CASEVAC in World War II and the Korean War, and helicopters for MEDEVAC during the Vietnam War. The use of dedicated aeromedical helicopters for MEDEVAC continues to be the primary evacuation method for the U.S. military (Jenkins *et al.*, 2021a). Due to these ongoing improvements, the casualty survival rate has increased over the decades from 84% in Vietnam to 90% during the decade of conflict from 2001 to 2011 (Eastridge *et al.*, 2012).

MEDEVAC and CASEVAC operations utilize a plethora of vehicle types (e.g., trucks, ships, and helicopters). This thesis focuses on the aerial aspects of MEDEVAC operations (i.e., aeromedical helicopter operations), which are accomplished through the use of helicopter ambulances. Helicopters have the ability to fly directly to a point-of-injury (POI) or CCP that other platforms (e.g., ground vehicles or fixed-winged aircraft) may not be able to access or get to quickly. This aspect greatly increases a casualty's chances of survival, making helicopters the MEDEVAC vehicle of choice. The HH-60M Black Hawk helicopter in particular is specifically designed to support the MEDEVAC mission. These helicopters come equipped with the necessary resources (e.g., oxygen generator, integrated electrocardiogram (EKG) machine, electronically controlled litters, built-in external hoist, and an infrared system that can locate patients by their body heat) to give medical personnel the ability to simultaneously treat and transport casualties from a POI (or CCP) to an appropriate MTF (Jenkins, 2017). The HH-60G Pave Hawk, on the other hand, is designed to conduct personnel recovery missions under hostile conditions but can be used to support the MEDEVAC mission if needed.

There are three main aspects to consider when developing a MEDEVAC system: location, dispatching, and redeployment. The location of MEDEVAC units are usually determined while considering two objectives: maximizing coverage and minimizing response time subject to logistical, resource, and force protection constraints (Jenkins *et al.*, 2020c). Military dispatching authorities typically task MEDEVAC units to respond to incoming requests for service according to a closest-available dispatching policy, which, as the name suggests, automates the decision making process by simply dispatching the closest-available unit when requests are submitted to the system regardless of the precedence level of the casualties at the CCP or other system characteristics. Redeployment, while possible, poses challenges due to communication, resource, and availability issues. For this reason, redeployments are not typically performed and are not considered in this research.

1.3 Thesis Overview

This thesis focuses on a MEDEVAC dispatching problem in which a decision maker (i.e., dispatching authority) must decide which MEDEVAC unit (if any) to dispatch in response to a particular request for service. Redeployment is not considered, and the location of MTFs and MEDEVAC units are known.

An infinite-horizon, continuous-time Markov decision process (CTMDP) model is developed and transformed into an equivalent discrete-time MDP model via uniformization to determine the optimal dispatching policy that maximizes the expected total discounted reward earned by the system. A computational example is developed and applied to a MEDEVAC system forward deployed in Afghanistan in support of combat operations. Comparisons are made between the myopic policy (i.e., the closest-available dispatching policy) and the optimal policy generated by the MDP model.

This thesis contributes to the existing military EMS literature (e.g., Keneally *et al.* (2016); Rettke *et al.* (2016); Jenkins *et al.* (2018); Robbins *et al.* (2018); Jenkins *et al.* (2021b,a)), by explicitly modeling and accounting for threat levels at CCPs and different MEDEVAC asset types. These additions create a more realistic scenario in which the dispatching authority must not only take into account the precedence levels of incoming requests, availability of MEDEVAC units, future demand locations, and arrival rates, but also the potential of enemy threats into the dispatching decision as well as which asset type should be utilized to maximize the system's performance.

The remainder of this thesis is structured as follows: Chapter II provides a detailed review of relevant research pertaining to civilian and military emergency medical service (EMS) systems. Chapter III describes the MEDEVAC dispatching problem. Chapter IV presents the MDP formulation. Chapter V covers an application of the developed MDP model based on a representative scenario in southern Afghanistan. Chapter VI provides conclusions and areas for future research.

II. Literature Review

This chapter discusses research pertaining to EMS systems in both the civilian and military communities. In particular, this literature review focuses on research concerning the dispatching of civilian and military EMS vehicles upon receipt of service requests.

2.1 EMS Systems

The research pertaining to EMS operations can be traced back to the late 1960's. Within the EMS literature, the primary areas of focus include characteristics such as optimal location (e.g., Jarvis (1975); Daskin & Stern (1981); Bianchi & Church (1988)), allocation (e.g., Hall (1972); Berlin & Liebman (1974); Baker *et al.* (1989)), dispatch (e.g., Ignall *et al.* (1982); Swersey (1982); Green & Kolesar (1984)) and relocation of emergency vehicles (e.g., Kolesar & Walker (1974); Chaiken & Larson (1972); Berman (1981); Jenkins (2019)) to enhance the performance of the EMS system (Jenkins, 2017).

To analyze EMS systems, applied operations research techniques (e.g., stochastic modeling, queueing, discrete optimization, and simulation modeling) tend to be the tools of choice due to their ability to provide rigorous, defensible, and quantitative insights (Green & Kolesar, 2004). The goal of utilizing operations research techniques is to provide decision makers the ability to make data-driven decisions via the application of published models, but unfortunately this is not always the case due to limiting and, at times, unrealistic assumptions. However, Green & Kolesar (2004) reveal how research has positively influenced changes to EMS response systems. The research in this thesis intends to do the same for the military MEDEVAC community.

An optimality criterion (i.e., performance measure) must be established to optimize the performance of an EMS system. This is vital because it governs employment of EMS system resources and ultimately dictates patient survivability (McLay & Mayorga, 2010). EMS system performance is normally measured in terms of a response time threshold (RTT), which indicates the proportion of calls serviced within a given timeframe. RTTs are generally easy to evaluate, but there has been concerns about their ability to fully capture patient survivability. Due to this criticism, Erkut *et al.* (2008) recommend explicitly incorporating patient survivability. Estimating patient survivability has proven to be quite difficult; however, Erkut *et al.* (2008) propose the use of a monotonically decreasing function over time better describes the probability of patient survivability. This approach has been adopted and utilized in subsequent work (e.g., Bandara *et al.* (2012); Mayorga *et al.* (2013); Bandara *et al.* (2014); Grannan *et al.* (2015); Rettke *et al.* (2016); Jenkins *et al.* (2018); Robbins *et al.* (2018); Jenkins *et al.* (2021b,a)). Following suite, this thesis incorporates a survivability function to best represent patient survivability.

As it stands, most EMS response systems utilize a closest-available dispatching policy (i.e, a myopic policy). Although this policy simplifies the decision-making process, it is not always optimal and can be improved upon by taking into account important system characteristics such as patient precedence levels (Bandara *et al.*, 2012). The EMS dispatching literature focuses on developing feasible dispatching rules that render the highest utility based on the established optimality criterion.

Decisions concerning which EMS unit to dispatch upon a request for service must be made sequentially over time and under uncertainty. For this reason, many researchers utilize a dynamic programming approach to model this problem (Jenkins *et al.*, 2020b, 2021c). The following sections provide an extensive review of both civilian and military EMS response system research that leverage this approach.

2.1.1 Civilian

McLay & Mayorga (2013b) develop an MDP model to determine an optimal ambulance dispatching policy that incorporates the precedence levels of requests and the possibility of classification errors. The authors reveal alternative optimal solutions are possible based on the likelihood of classification errors. Moreover, the results indicate that relaxing the assumption of exponentially distributed service times has little impact on the MDP-generated optimal policy. McLay & Mayorga (2013a) expand upon the aforementioned MDP model to consider the problem feature of balancing equality and equity. More specifically, the authors examine tradeoffs between adopting a dispatching policy that decreases performance in rural, low-populated areas in favor of increased performance in higher populated areas. The authors formulate a constrained MDP model as a linear programming model to identify optimal dispatching policies and use it to analyze four different measures of equity.

Both McLay & Mayorga (2013b) and McLay & Mayorga (2013a) provide meaningful insights with regard to the ambulance dispatching problem. However, their research only allows for smaller scale problem instances to be evaluated due to the *curse of dimensionality*. The results garnered from these solutions are still of value, but larger problem instances should be evaluated to provide more realistic and meaningful insights. To accomplish this, researchers are utilizing approximate dynamic programming (ADP) approaches to generate high-quality results for practical-sized problems within a reasonable amount of computational time.

Maxwell *et al.* (2010) adopt an ADP solution approach, utilizing an approximate policy iteration (API) algorithm that seeks to determine a high-quality ambulance redeployment policy (i.e., the movement of ambulances that have just completed service at a hospital to service another request). The value function is approximated through an affine combination of deliberately designed, problem-specific basis functions. A least-squares policy evaluation (LSPE) technique is used within their API algorithm to update the basis function coefficients. Relocation and dispatching decisions are not considered, but the queueing of requests is allowed. The authors apply their solution methodology to two metropolitan EMS response scenarios and are able to achieve improved performance when compared to the benchmark policies.

Schmid (2012) models and solves an ambulance dispatching problem that accounts for both redeployment decisions and the queueing of requests. The author adopts an ADP solution approach with an approximate value iteration (AVI) algorithmic structure. The value function is approximated via a spatial and temporal aggregation scheme. Schmid (2012) demonstrates the efficacy of the solution approach by applying it to an EMS response scenario based in Vienna, Austria. The obtained results indicate improved performance when compared to a benchmark policy.

Nasrollahzadeh *et al.* (2018) investigate a modification of the ambulance dispatching problem, considering redeployment decisions and relocation decisions, and allowing for the queueing of requests. The authors develop an MDP model and apply it to an EMS response scenario set in Mecklenburg County, North Carolina. To obtain results with a computationally efficient manner, the authors adopt an ADP solution approach utilizing an API framework. The value function is approximated through an affine combination of deliberately designed, problem-specific basis functions and are approximated via a LSPE technique. The authors are able to obtain improved performance results when compared to multiple benchmark policies.

Park & Lee (2019) evaluate another variant of the ambulance dispatching problem. Their research builds upon previous work in the ambulance dispatching and redeployment literature by additionally considering a two-tiered ambulance system and patient classification errors. In a two-tiered ambulance system, different ambulances are specially equipped to handle different types of patient requests (e.g., basic life support (BLS) vehicles for non-emergency patient transport and advanced life support (ALS) for emergency-patient transport). The authors argue that incorporation of this system, consisting of both advanced and basic life support for emergency and non-emergency patient care, respectively, can provide a cost efficient medical service. However, a system like this requires accurate medical classifications to avoid serious complications. Within this system, the authors seek to determine how the optimal policy changes according to classification errors, and what type of classification decisions need to be made for ambiguous patients to minimize patient risk. An MDP model is developed and a mini-batch monotone-ADP approach is proposed to solve the problem. Computational experiments using realistic system dynamics based on historical data from Seoul, South Korea are developed and reveal that the ADP-generated optimal policy can reduce a patient's risk level index by an average of 11.2% when compared to a myopic policy.

Military and civilian EMS systems are similar in nature, as both are designed to address the transportation needs of time-sensitive patients to higher level MTFs. For this reason, the advancement in the civilian EMS dispatching literature has helped pave the foundation upon which the military MEDEVAC dispatching literature is able to evolve. Despite their similarities, several substantive differences remain that must be considered when examining the performance of a military EMS system (e.g., longer travel times, longer loading and unloading times, more complicated evacuation process) (Jenkins *et al.*, 2021a). To account for these differences, independent research has been dedicated to focus on the military MEDEVAC dispatching problem.

2.1.2 Military

Whereas many research papers have been written in relation to the improvement of military MEDEVAC system efficiency, very few investigate the decision of which

MEDEVAC unit to launch to a given prioritized request for service (Robbins *et al.*, 2018). The first of this kind appears to be Keneally *et al.* (2016).

Keneally *et al.* (2016) develop an MDP model that examines the military MEDEVAC dispatching policy in a combat environment. The proposed model indicates how to optimally dispatch MEDEVAC helicopters to casualty events to maximize system utility during steady-state combat operations. The utility gained from servicing a specific request depends on the number of casualties being evacuated, the precedence of the casualties, and the location of both the servicing vehicle and CCP. The authors apply their model to a notional scenario set in Afghanistan, wherein the MEDEVAC system is supporting counter-insurgency operations. The results indicate that the myopic policy is not always the best method for dispatching MEDEVAC helicopters.

Jenkins *et al.* (2018) expand upon the work conducted by Keneally *et al.* (2016) by incorporating admission control. This provides the dispatching authority with the ability to reject incoming requests for service, thereby reserving MEDEVAC units for higher precedence requests. A queueing system is also incorporated, allowing the dispatching authority to accept incoming requests regardless of the status of the MEDEVAC units and place them in a queue to be serviced later. Moreover, the authors utilize a survivability function based on response time instead of a RTT to model the MDP reward function. Similar to Keneally *et al.* (2016), Jenkins *et al.* (2018) conduct a computational experiment with their model based on counter-insurgency operations in Afghanistan. The authors conclude that the dispatching of MEDEVAC units with the consideration of precedence levels and locations of MEDEVAC units increases system performance.

Both Keneally *et al.* (2016) and Jenkins *et al.* (2018) develop an MDP model to determine an optimal dispatching policy for MEDEVAC units. Each effort performs small-scale computational experiments, wherein their respective MDP models are able to determine optimal solutions in a tractable amount of time. The results garnered provide valuable insights concerning MEDEVAC dispatching policies. However, despite these accomplishments, practical (i.e., large-scale) problem instances should be analyzed to obtain more realistic insights.

The MEDEVAC dispatching problem in practice has a high-dimensional state space. This renders exact solution methodologies (e.g., dynamic programming and linear programming) ineffective (Robbins *et al.*, 2018) due to the *curse of dimensionality*. Given this challenge, several researchers (e.g., Rettke *et al.* (2016), Robbins *et al.* (2018), Jenkins *et al.* (2021b), Jenkins *et al.* (2021a)) employ ADP approaches to obtain high quality dispatch policies relative to current practices (i.e., dispatching the closest-available unit).

Rettke *et al.* (2016) develop an approximate policy iteration (API) algorithm that utilizes least-squares temporal difference (LSTD) learning for policy evaluation to solve their MDP model of the MEDEVAC dispatching problem. The authors demonstrate the applicability of their model via a notional scenario representative of contemporary military operations in northern Syria. The authors obtain a solution that outperforms the myopic policy by over 30% with regards to a life saving performance metric.

Robbins *et al.* (2018) add to the MEDEVAC dispatching literature by examining a realistic large-scale combat scenario set in Afghanistan. The authors utilize a zone tessellation scheme within their model, similar to that used by U.S. military MEDEVAC practitioners. Utilizing a hierarchical aggregation value function approximation scheme within an API algorithmic framework, the authors obtain high quality solutions that substantially outperform the myopic policy and are within 1% of optimality.

In Jenkins *et al.* (2021a), the authors contribute to the MEDEVAC dispatching literature by formulating and utilizing an MDP model that incorporates previously examined problem features (i.e., admission control and queueing) as well as redeployment. The authors develop, test, and compare two distinct ADP solution approaches, both of which utilize an API algorithmic framework. The first approach uses LSTD learning for policy evaluation, whereas the second uses neural network (NN) based learning. The authors generate 30 different problem instances and are able to significantly outperform the closest-available benchmark policies, 90% and 80% of the time with their NN and LSTD techniques, respectively.

The research conducted in Jenkins *et al.* (2021b) defines and examines the MEDEVAC dispatching, preemption-rerouting and redeployment (DPR) problem. This effort formulates an MDP model and solves it via an ADP approach within an API framework that utilizes a support vector regression (SVR) value function approximation scheme. The DPR problem is a variation of the MEDEVAC dispatching problem that not only seeks to determine which MEDEVAC unit to dispatch upon receipt of a request, but also incorporates how a unit should redeploy upon finishing a service request. In this research, the authors apply their model to a notional scenario based on high-intensity combat operations to defend Azerbaijan against a notional aggressor. The ADP-generated policies obtained from the computational experiments significantly outperform two benchmark policies.

2.2 Thesis Contribution

The research in this thesis adds to the MEDEVAC dispatching literature by utilizing an MDP model that explicitly incorporates the threat level at CCPs and different MEDEVAC asset types, problem aspects that have yet to be considered, as well as previously examined problem features (e.g., admission control). Whereas previous research has accounted for the precedence level of MEDEVAC requests, none take into account the environment in which requests need to be serviced. Moreover, previous research has primarily focused on utilizing one MEDEVAC asset type, the HH-60M.

The HH-60Ms (Black Hawk) are a variant of the U.S. Army's HH-60 helicopter series and are specifically utilized for MEDEVAC operations. The HH-60Gs (Pave Hawk) are a gunship variant of the Black Hawk operated by the U.S. Air Force. The Pave Hawk primarily serves as the Air Force's premier combat search and rescue vehicle; however, there are times when it can be used to support MEDEVAC operations. Unlike the Black Hawk, the Pave Hawk comes equipped with two .50 caliber mini guns that enable it to enter hostile areas unaccompanied, regardless of threat level. Despite being owned by a different service, this thesis will directly model Pave Hawks to determine how its incorporation will affect dispatching policies. This addition enables more complexity to enter the model. Similar to Keneally *et al.* (2016), Jenkins *et al.* (2018), and Robbins *et al.* (2018), a computational experiment set in Afghanistan is examined to demonstrate the efficacy of the rendered solutions. Following from Rettke *et al.* (2016), Robbins *et al.* (2018), Jenkins *et al.* (2020c), Jenkins *et al.* (2021a), and Jenkins *et al.* (2021b), this research also develops and employs an ADP solution technique.

III. Problem Description

This chapter provides a detailed description of the military MEDEVAC dispatching process.

3.1 The MEDEVAC Dispatching Process

The Army Health System (AHS) is the Army component of the Department of Defense Military Health System (Department of the Army, 2019). One of the primary missions of the AHS is to provide MEDEVAC services across a range of military operations. To accomplish this, the Army's MEDEVAC system is comprised of dedicated air and ground evacuation platforms that have been designed, manned, and equipped to provide en route medical care to patients being evacuated (Department of the Army, 2019). Dedicated rotary-wing air ambulances are utilized for MEDEVAC missions and are commanded by the general support aviation battalion (GSAB) (Jenkins, 2017). The GSAB manages all activities related to the execution of aerial MEDEVAC operations (Department of the Army, 2019).

Within the GSAB, an Army aeromedical evacuation officer (AEO) acts as the MEDEVAC dispatching authority in a deployed military EMS system (Fish, 2014). AEOs direct the use of medical aircraft, personnel, and equipment in support of operational and strategic MEDEVAC procedures within a theater of operations. Upon the receipt of a MEDEVAC request, the AEO must decide which unit, if any, to dispatch. Delays in the decision-making process substantially decrease the patient's probability of survival. As such, it is imperative that the GSAB implements a dispatching policy resulting in high-quality and rapid evacuation of combat casualties from CCPs to appropriate MTFs.

When conducting an operation, if a unit sustains casualties that need to be evacuated, they will initiate medical evacuation operations by calling in a 9-line MEDEVAC request (Jenkins *et al.*, 2020a). A 9-line MEDEVAC request is transmitted in a standardized message format with a prescribed amount of information to aide in the process of transporting casualties. During wartime conditions, the information required in a 9-line MEDEVAC request is reported in the following order: the location of the pickup site (i.e., POI or CCP), radio frequency and call sign, number of casualties by precedence, special equipment required, number of casualties by type, security (i.e., threat level) at pickup site, method of marking pickup site, casualty nationality and status, and chemical, biological, radiological, and nuclear (CBRN) contamination (Jenkins, 2017). It is the responsibility of either the senior military member or medical person, if available, at the scene to identify the evacuation precedence category of each casualty and determine whether a 9-line MEDEVAC request is necessary (Jenkins, 2017). The overall precedence of a 9-line MEDEVAC request is based on the most time sensitive precedence of the casualties. Table 1 describes the different precedence levels that the U.S. military utilize in order to categorize MEDEVAC requests. Due to the low quantity and high demand of aerial MEDEVAC units, accurate precedence level assignment is essential due to the burden it can place on the system. Table 2 describes the different threat levels that can be reported in a 9-line MEDEVAC request. Once a request is submitted, it is then transmitted to a dispatching authority that is responsible for the execution of all MEDEVAC operations (i.e., GSAB).

In a combat scenario, requests for MEDEVAC units are typically made at the POI, once enemy fire has been suppressed (Jenkins, 2017). The requests are then transmitted through several layers of command before reaching an AEO working within the GSAB. The exact flow of information depends on the infrastructure within the command, the communication equipment available, and the command and control

Table 1. Evacuation Precedence Categories (Department of the Army, 2019)

Priority I - Urgent	Assigned to emergency cases that should be evacuated as soon as possible and within a maximum of one hour in order to save life, limb or eyesight and to prevent complications of serious illness and to avoid permanent disability.
Priority II - Priority	Assigned to sick and wounded personnel requiring prompt medical care. This precedence is used when the individual should be evacuated within four hours or if his medical condition could deteriorate to such a degree that he will become an Urgent precedence, or whose requirements for special treatment are not available locally, or who will suffer unnecessary pain or disability.
Priority III - Routine	Assigned to sick and wounded personnel requiring evacuation, but whose condition is not expected to deteriorate significantly. The sick and wounded in this category should be evacuated within 24 hours.

Table 2. Threat Levels

N	No enemy troops in the area
P	Possible enemy troops in the area (approach w/ caution)
E	Enemy troops in the area (approach w/ caution)
X	Enemy troops in the area (armed escort required)

organization of the MEDEVAC system (Rettke *et al.*, 2016). The procedures outlined in Department of the Army (2019) and the graphical representations developed by previous researchers (e.g., Keneally *et al.* (2016); Rettke *et al.* (2016); Jenkins *et al.* (2018)) were used as a basis for the MEDEVAC mission timeline depicted in Figure 1.

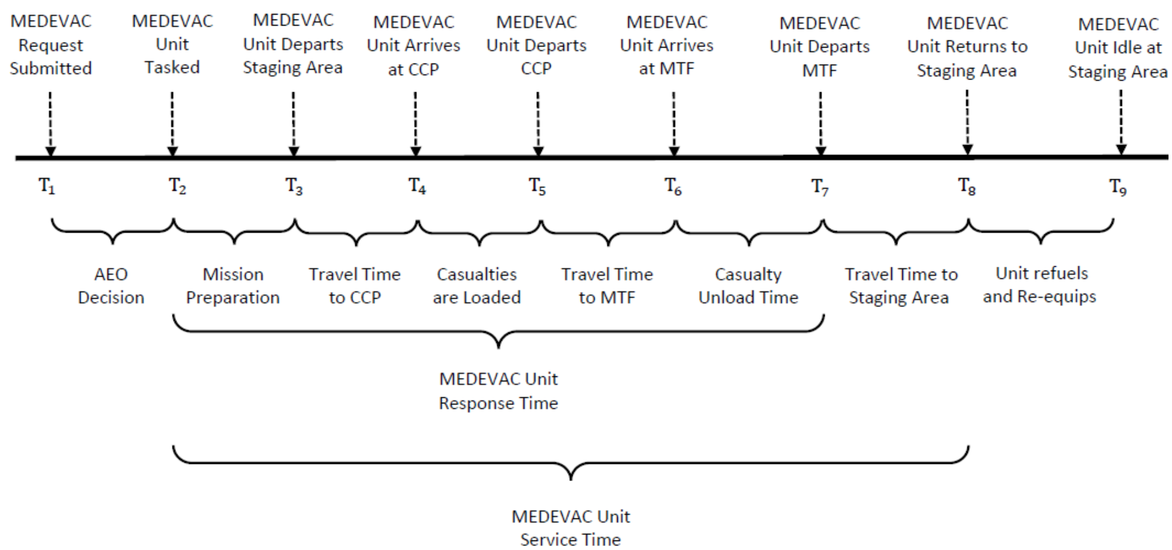


Figure 1. MEDEVAC Mission Timeline

Once a request has been submitted, as indicated by T_1, the casualty is taken to a pre-designated pick-up location established prior to conducting operations (i.e., CCP). Previous papers that evaluate the military MEDEVAC dispatching problem typically assume that CCPs are located in areas that are more secure and viable for helicopter landing, thus reducing the need for extra requirements prior to dispatching (e.g., armed-escorts, rescue hoists). This assumption, while reasonable, is not always the case. Indeed, line six of the 9-line MEDEVAC request message indicates the security (i.e., threat level) at the designated pick-up site, informing the dispatching authority of the needed security requirements prior to making a decision. This research relaxes the assumption that CCPs are always secure.

The time between T_1 and T_2 represents the AEO's decision time until they are able to task a helicopter. This period accounts for the time required to determine which MEDEVAC unit to dispatch; whether an armed escort is required; which armed escort team to assign, if required; and the time required to transmit the request information to the assigned MEDEVAC assets (Jenkins, 2017).

When a MEDEVAC unit is assigned, indicated by T_2, it will begin mission preparation (e.g., preparing medical equipment and personnel). Once the unit is ready, it dispatches and begins traveling towards the designated CCP, indicated by T_3. If an armed escort is required, it is typically dispatched with the appropriate MEDEVAC unit from the staging area. However, there are situations where the MEDEVAC unit may need to meet an armed escort en route to the CCP. Waiting for an armed escort may substantially increase travel time; however, an unarmed MEDEVAC unit (e.g., HH-60M Black Hawk) cannot enter a high threat level area without an armed escort.

T_4 denotes the time at which the MEDEVAC unit arrives at the designated CCP. Upon arrival, the MEDEVAC unit immediately begins initial treatment and loads casualties (Jenkins, 2017). Once all casualties are loaded, the MEDEVAC unit evacuates the patients to the closest MTF, indicated by T_5. It is important to note that in high conflict areas, the successful extraction of patients is not always guaranteed. Imposing threats can severely impact the helicopters ability to land and properly treat their intended patients. In certain circumstances a MEDEVAC unit may need to abort the mission due to intense hostility in the area and the potential risk of being shot down.

Upon successfully evacuating the patient, the MEDEVAC unit arrives at the MTF at time T_6. The unit then begins to unload casualties and transfer the responsibility of subsequent care over to the appropriate medical staff. Once all casualties are unloaded, the MEDEVAC unit departs the MTF and returns to its original staging area, indicated by T_7 and T_8 respectively. Once the MEDEVAC unit arrives back to its staging area the mission is complete. Upon completion of the mission, the MEDEVAC unit begins refueling and re-equipping and is ready to be dispatched again at time T_9. Although difficult, it is possible for the GSAB to task the MEDEVAC unit upon its arrival back to its staging facility at time T_7; however, this aspect of redeployment will not be considered in this thesis.

When considering dispatching policies, military medical planners must consider the measurement of the MEDEVAC system performance. McLay & Mayorga (2010) report that the most common method for evaluating civilian EMS systems focuses on response times (i.e., how long it takes an ambulance to reach a patient after receiving a call). Due to the fact that civilian EMS systems are evaluated on response time, they primarily focus on their ability to rapidly respond to cardiac arrest situations. These cases are emphasized due to the time-sensitive nature upon which they need to be serviced. It is also believed that if an EMS system is able to quickly respond to cardiac arrest patients, then it is more likely to be able to service similar life-threatening situations; thus, cementing response time as the main evaluation criteria for civilian EMS systems. However, because of the nature of military operations, the MEDEVAC system should not be measured using the same criteria.

Several different factors complicate the evacuation of casualties from a battlefield (e.g., enemy threat as well as longer travel, load, and unload times). Furthermore, the primary cause of death for battlefield casualties is blood loss, not cardiac arrest (Shackelford *et al.*, 2017). In an attempt to alleviate this issue, some MEDEVAC units have been outfitted with in-flight blood transfusion capabilities, but the majority have not. Although it is intuitive that early transfusion should help to diminish this issue, published data on pre-hospital transfusions do not demonstrate a survival advantage (Shackelford *et al.*, 2017). Smith *et al.* (2016) provide a systematic review that details limitations in pre-hospital transfusion trauma care research. Nonetheless, due to the high costs of implementation and assessment as well as the lack of supporting data, there has not been a change to the MEDEVAC system's evaluation measure. Therefore, it is vital for the MEDEVAC unit to not only arrive to the casualty, but also to stabilize and evacuate the casualty to an appropriate MTF.

Whereas civilian EMS systems measure response time as the time it takes to reach a patient after obtaining a call for service, response times for the military EMS system need to account for the time it takes the MEDEVAC unit to arrive and evacuate patients to an MTF. Therefore, it is appropriate to define the response time for a MEDEVAC unit as the time between T_2 and T_7 in Figure 1. Since a MEDEVAC unit must return to its home station after being dispatched, service time is defined as $T_8 - T_2$.

The objective of this thesis is to determine a policy that dispatches MEDEVAC units such that the expected total discounted reward earned by the system is maximized. Dispatching decisions are difficult due to the fact that subsequent casualty events and ensuing requests are not known beforehand. The stochasticity in this sequential decision-making problem stems from casualty demand and casualty event locations, as well as dispatch, travel, and services times (Robbins *et al.*, 2018). For this reason, the decision making authority (i.e., AEO) must have an established dispatching policy prior to the commencement of an operation. This research seeks to provide this policy via an MDP model.

To provide an accurate representation of the system, this thesis leverages information related to MEDEVAC dispatch, travel, and service times to parameterize the model. Moreover, the research conducted in this thesis uses stochastic simulation methods (i.e., Monte-Carlo) to model 9-line MEDEVAC request submissions. These features are incorporated into the MDP model, which is subsequently examined via a notional scenario based on combat operations in southern Afghanistan to demonstrate the efficacy of the rendered solutions.

IV. Methodology

This chapter provides the formulation for the Markov decision process (MDP) model of the military MEDEVAC dispatching problem under evaluation in this thesis.

4.1 MDP Formulation

A discounted, infinite-horizon continuous time MDP (CTMDP) model is formulated to determine the optimal dispatching policy for a deployed MEDEVAC system. The objective of the model is to develop a dispatching policy for the MEDEVAC system that maximizes the expected total discounted reward earned by the system.

The model in this thesis assumes that 9-line MEDEVAC requests arrive sequentially over time according to a Poisson process with parameter λ, $PP(\lambda)$. It is important to note that a $PP(\lambda)$ has independent and stationary increments. These assumptions are reasonable due to the nature of the environment in which the MEDEVAC system operates. During combat operations, there can potentially be a large number of violent interactions that take place resulting in different casualty events, thereby generating unrelated requests for service; thus, the number of arrivals that occur in disjoint time intervals are independent. The assumption of stationary increments is also valid due to that the implicit sizes, locations, and dispositions of forces are generally fixed with respect to time.

A MEDEVAC request that enters the system is characterized by the location of the casualty event (i.e., zone), its precedence level (i.e., *urgent*, *priority*, *routine*), and the threat level in which the casualty request arrived. The arrival of these requests are modeled using a *splitting* technique. Splitting refers to the generation of two or more counting processes from a single Poisson process (Kulkarni, 2017). Let the original counting process $\{N(t') : t' \geq 0\}$ denote the $PP(\lambda)$ that counts the number of 9-line MEDEVAC requests that enter the system within the time interval $(0, t']$. The original counting process can be segmented into different counting processes that are categorized by the zone $z \in \mathcal{Z} = \{1, 2, \dots, |\mathcal{Z}|\}$, the precedence level $k \in \mathcal{K} = \{1, 2, \dots, |\mathcal{K}|\}$, and the threat level $l \in \zeta = \{1, 2, \dots, |\zeta|\}$. We let $\mathcal{R} = \{(z, k, l) : (z, k, l) \in \mathcal{Z} \times \mathcal{K} \times \zeta\}$ denote our set of request categories. There are a total of $|\mathcal{R}| = |\mathcal{Z}||\mathcal{K}||\zeta|$ possible request categories. The original process is now split into $|\mathcal{R}|$ independent processes $\{N_{zkl}(t') : t' \geq 0\}, \forall (z, k, l) \in \mathcal{R}$. Since each request belongs to one and only one category, we obtain the following result:

$$N(t') = \sum_{(z,k,l) \in \mathcal{R}} N_{zkl}(t'). \tag{1}$$

The nature of the split processes $\{N_{zkl}(t') : t' \geq 0\}, \forall (z, k, l) \in \mathcal{R}$ depends on how the requests are categorized, which is conducted using a Bernoulli splitting mechanism. The Bernoulli splitting mechanism generates the splitting processes $\{N_{zkl}(t') : t' \geq 0\}, \forall (z, k, l) \in \mathcal{R}$ given parameters $p_{zkl} > 0, \forall (z, k, l) \in \mathcal{R}$ such that $\sum\limits_{(z,k,l) \in \mathcal{R}} p_{zkl} = 1$. Each request is independently categorized by its zone, precedence and threat level combination with probability p_{zkl}. Due to the splitting mechanism, each split process is characterized as a Poisson process with parameter λp_{zkl}, denoted as $PP(\lambda p_{zkl})$.

The MEDEVAC unit's service time comprises the time from initial assignment notification to the unit's return back to its original staging area. This thesis assumes that the service times for MEDEVAC units are exponentially distributed. Although some may argue that this assumption is unrealistic, McLay & Mayorga (2013b) conducted simulation-based EMS system analyses utilizing different types of service time distributions, determining that this assumption does not significantly impact the generation of optimal policies. This result suggests that despite the assumption of exponentially distributed service times, the solutions generated from our MDP model will still provide useful insights.

Now that the characteristics of the arrival process and service times have been introduced, we can proceed with the formulation of the MDP model. The MDP model components (i.e., decision epoch, state space, action space, transition probabilities, rewards, objective and optimality equation) are described in detail below. The following formulation is leveraged and adapted from Jenkins *et al.* (2018).

4.1.1 Decision Epochs

In an MDP model, the decision epochs are the points in time in which the decision maker (e.g., AEO) needs to make a decision. We define the set of decision epochs as $\mathcal{T} = \{1, 2, \ldots\}$. Events take place when the status of the MEDEVAC system changes. This occurs either due to the completion of a service request or the arrival of a 9-line MEDEVAC request.

The MDP model of the MEDEVAC system follows the properties of semi-Markov decision processes (SMDPs). SMDPs generalize MDPs by allowing, or requiring, the decision maker to choose actions whenever the status of the system changes; modeling the system evolution in continuous time; and by allowing the time spent in a particular state to follow an arbitrary probability distribution (Puterman, 2005). The MEDEVAC system MDP model is viewed as a CTMDP, a special case of an SMDP wherein the inter-transition times are exponentially distributed and decisions are made at each transition. In order to analyze a CTMDP, complex techniques are typically required. However, through the process of *uniformization*, we are able to obtain an equivalent discrete-time discounted model with constant transition rates (Puterman, 2005). This transformation allows the use and interpretation of discrete-time algorithms to be applied directly.

The model in this thesis is formulated as a CTMDP and transformed through the process of uniformization. The policy iteration algorithm is then applied to determine an optimal dispatching policy.

4.1.2 State Space

The state $S_t \in \mathcal{S}$ describes the status of the MEDEVAC system at epoch $t \in \mathcal{T}$. The MEDEVAC system state is represented by the tuple $S_t = \left(M_t, \hat{R}_t\right)$ wherein M_t represents the status of each MEDEVAC unit, and $\hat{R}_t$ represents the status of the arrival request at each epoch t.

The tuple M_t can be written as:

$$M_t = (M_{tm})_{m \in \mathcal{M}}, \tag{2}$$

where $\mathcal{M} = \{1, 2, \ldots, |\mathcal{M}|\}$ is the set of MEDEVAC units in the system. The state variable $M_{tm} \in \{0\} \cup \mathcal{Z}$ contains information in relation to a MEDEVAC unit $m \in \mathcal{M}$ at each epoch t. Each MEDEVAC unit can either be idle (i.e., $M_t = 0$) or servicing a particular zone, $\mathcal{Z}$ (i.e., $M_t = z$).

The request arrival status tuple $\hat{R}_t$ indicates whether or not there is a request waiting to be admitted into the system or denied. This tuple also provides the zone from which the request has originated, the precedence level of the request, and its threat level. Let $\hat{R}_t = (0, 0, 0)$ indicate that no requests are in the system at epoch t, otherwise:

$$\hat{R}_t = \left(\hat{Z}_t, \hat{K}_t, \hat{L}_t\right)_{\hat{Z}_t \in \mathcal{Z}, \hat{K}_t \in \mathcal{K}, \hat{L}_t \in \zeta}. \tag{3}$$

The random variables, $\hat{Z}_t, \hat{K}_t,$ and $\hat{L}_t$, correspond to the zone, precedence level, and threat level of the request arrival at epoch t, respectively. The information contained in $\hat{Z}_t, \hat{K}_t,$ and $\hat{L}_t$, only become known when a request occurs at epoch t. Until then, the information contained in these variables is uncertain.

The size of the state space $\mathcal{S}$ depends on $|\mathcal{M}|, |\mathcal{Z}|, |\mathcal{K}|$, and $|\zeta|$. The following expression provides the cardinality of the state space for the MEDEVAC system:

$$|\mathcal{S}| = \big(1 + |\mathcal{Z}|\big)^{|\mathcal{M}|}\big(1 + |\mathcal{Z}||\mathcal{K}||\zeta|\big). \tag{4}$$

Unfortunately, as more state variables are added, the size of the state space grows exponentially. When the state space is too large, exact dynamic programming algorithms become intractable. This is commonly referred to as the *curse of dimensionality*. As depicted in Chapter II of this thesis, approximation techniques exist and have been utilized to work around these issues.

4.1.3 Action Space

When a 9-line MEDEVAC request is transmitted, the dispatching authority needs to determine whether or not to reject the request (i.e., admission control decision), or which MEDEVAC unit to dispatch upon the request's acceptance (i.e., dispatching decision). In this thesis, there are two possible outcomes: the request is rejected from the system, or the request is accepted and a MEDEVAC unit is tasked to service it.

Let $x_t^A \in \{\Delta, 0, 1\}$ denote the admission control decision at epoch t. When an arrival request is not present at epoch t the system will continue to transition without any impact from x_t^A, indicated by $x_t^A = \Delta$. We let $x_t^A = 0$ denote an arrival request being admitted into the system and $x_t^A = 1$ denote an arrival request being rejected.

If a 9-line MEDEVAC request is admitted into the system, the AEO must determine which idle MEDEVAC unit to dispatch to service it. Let $\mathcal{I}(S_t) = \{m : m \in \mathcal{M}, M_{tm} = 0\}$ denote the set of all idle MEDEVAC units when the system is in state S_t at epoch t. If $\mathcal{I}(S_t) = \varnothing$ there are no idle MEDEVACs at epoch t, and the 9-line MEDEVAC request must be rejected (i.e., $x_t^A = 1$). The dispatching decision tuple to describe the AEO's decision can be written as

$$x_t^D = (x_{tm}^D)_{m \in \mathcal{I}(S_t)}. \tag{5}$$

The decision variable $x_{tm}^D = 1$ if MEDEVAC unit $m \in I(S_t)$ is dispatched to service the accepted arrival request at epoch t, $x_{tm}^D = 0$ otherwise.

Let $x_t = \left(x_t^A, x_t^D\right)$ denote a compact representation of the decision variables at epoch t. The following constraint is used to bound the decisions being made at each epoch t,

$$1 - x_t^A = \sum_{m \in \mathcal{I}(S_t)} x_{tm}^D. \tag{6}$$

This constraint indicates that if an arrival request is accepted into the system at epoch t, then a MEDEVAC unit must be dispatched to service it. The set of all possible actions when a decision is required is defined as follows

$$\mathcal{X}(S_t) = \begin{cases} (\Delta, \{0\}^{|\mathcal{I}(S_t)|}) & \text{if } \hat{R}_t = (0,0,0), \mathcal{I}(S_t) \neq \varnothing \\ (1, \{0\}^{|\mathcal{I}(S_t)|}) & \text{if } \hat{R}_t \neq (0,0,0), \mathcal{I}(S_t) = \varnothing \\ (\{0,1\}, \{0,1\}^{|\mathcal{I}(S_t)|}) & \text{if } \hat{R}_t \neq (0,0,0), \mathcal{I}(S_t) \neq \varnothing \end{cases} \cdot \tag{7}$$

The first case in Equation (7) represents the set of actions available to the AEO, when a MEDEVAC unit returns from servicing a request. The last two correspond to the set of all feasible actions the AEO can take, when a decision epoch occurs due to 9-line MEDEVAC request submission.

4.1.4 Transitions

Before introducing the transition probabilities, the notion of post decision states need to be introduced. Let μ_{mzl} denote the service rate of MEDEVAC unit $m \in \mathcal{M}$ when servicing a 9-line MEDEVAC request in zone $z \in \mathcal{Z}$ under threat level $l \in \zeta$, and $\mathcal{B}(S_t) = \{m : m \in \mathcal{M}, M_{tm} \neq 0\}$ denote the set of busy MEDEVAC units when the system is in state S_t at epoch t.

If the MEDEVAC system is in a pre-decision state S_t (i.e, the state of the system directly before a decision has been made) and action x_t is taken, the system will immediately transition to a post-decision state S_t^x (i.e., the state of the system directly after a decision has been made). The amount of time the system remains in post-decision state S_t^x before transitioning to pre-decision state S_{t+1} follows an exponential distribution with parameter $\beta(S_t, x_t)$; this is known as the sojourn time and can be written as

$$\beta(S_t, x_t) = \lambda + \sum_{m \in \mathcal{B}(S_t)} \mu_m, M_{tm} + \sum_{m \in \mathcal{I}(S_t)} \mu_{mzl} x_{tm}^D. \tag{8}$$

The probabilistic nature of the process is summarized in terms of an infinitesimal $|\mathcal{S}| \times |\mathcal{S}|$ generator matrix with components

$$G(S_{t+1}|S_t, x_t) = \begin{cases} -[1 - p(S_t^x|S_t, x_t)]\beta(S_t, x_t), & \text{if } S_{t+1} = S_t^x \\ p(S_{t+1}|S_t, x_t)\beta(S_t, x_t), & \text{if } S_{t+1} \neq S_t^x \end{cases}, \tag{9}$$

wherein $p(S_{t+1}|S_t, x_t)$ denotes the probability of the system transitioning to state S_{t+1} given that it was in state S_t and action x_t was taken (Jenkins *et al.*, 2018).

4.1.5 Transition Probabilities

To properly develop the transition probabilities and rewards associated with the model, uniformization is applied to obtain constant transition rates. This allows us to apply algorithms for discrete-time models directly. To uniformize the system, we calculate the maximum rate of transition as given by

$$\nu = \lambda + \sum_{m \in \mathcal{M}} \tau_m, \tag{10}$$

wherein

$$\tau_m = \max_{z \in \mathcal{Z}, l \in \zeta} \mu_{mzl}, \ \forall \ m \in \mathcal{M}, \forall \ l \in \zeta. \tag{11}$$

Applying the above uniformization principles render, the following transition probabilities:

$$\hat{p}(S_{t+1}|S_t, x_t) = \begin{cases} 1 - \frac{[1-p(S_t^x|S_t,x_t)]\beta(S_t,x_t)}{\nu} & \text{if } S_{t+1} = S_t^x \\ \frac{p(S_{t+1}|S_t,x_t)\beta(S_t,x_t)}{\nu} & \text{if } S_{t+1} \neq S_t^x \\ 0 & \text{Otherwise} \end{cases} \cdot \tag{12}$$

4.1.6 Rewards

The contribution function $c(S_t, x_t)$ captures the expected immediate reward (i.e., contribution) attained by the dispatching authority for making decision x_t when the system is in state S_t. Utilizing the framework provided in Jenkins *et al.* (2018), let $c(S_t, x_t) = \Phi_{mzkl}$ denote the immediate reward earned when the AEO dispatches MEDEVAC unit $m \in \mathcal{M}$ to service a request originating in zone $z \in \mathcal{Z}$ with precedence level $k \in \mathcal{K}$ and threat level $l \in \zeta$. The system receives rewards upon a unit being tasked. The recommended time requirements for each precedence level are outlined in Table 1. The continuous-time contribution function can be written as

$$\Phi_{mzkl} = \begin{cases} \eta e^{\frac{-\Psi_{mzl}}{60}}, & \text{if } k = 1 \ (i.e., urgent) \\ e^{\frac{-\Psi_{mzl}}{240}}, & \text{if } k = 2 \ (i.e., priority) \\ 0, & \text{Otherwise,} \end{cases} \tag{13}$$

wherein, Ψ_{mzl} is the expected response time for a particular MEDEVAC unit $m \in \mathcal{M}$ servicing a request in zone $z \in \mathcal{Z}$ with a threat level of $l \in \zeta$, and $\eta \geq 1$ is the incentive parameter used to vary the urgent-to-priority immediate expected reward ratio. Note that the system receives no reward for servicing routine requests. Due to the high operational tempo that is simulated in this thesis, we assume that routine requests will be of lesser concern to commanders who will prioritize servicing life threatening requests. For this reason the AEO will choose to let routine requests be serviced by other agencies (i.e., CASEVAC) instead of utilizing a MEDEVAC unit.

After applying uniformization, we obtain the following reward function,

$$\tilde{c}(S_t, x_t) = c(S_t, x_t)\frac{\alpha + \beta(S_t, x_t)}{\alpha + \nu}, \tag{14}$$

wherein $\alpha > 0$ denotes the continuous-time discounting rate, which indicates that the present value of one unit received t time units in the future equals $e^{-\alpha t}$. The discrete-time discount rate γ is obtained by setting $\gamma = \frac{\nu}{\nu+\alpha}$.

4.1.7 Optimality Equation

Let $X^{\pi}(S_t)$ be a decision function based on policy $\pi \in \Pi$ that prescribes dispatching decisions for each state $S_t \in \mathcal{S}$. The objective of our MDP model is to determine the optimal policy, π^*, from the class of policies $\pi \in \Pi$ that maximizes the expected total discounted reward earned by the MEDEVAC system. Our objective is expressed as follows

$$\max_{\pi\in\prod} \mathbb{E}^{\pi}\left[\sum_{t=1}^{\infty}\gamma^{t-1}\tilde{c}.(S_t, X^{\pi}(S_t))\right], \tag{15}$$

The optimal policy, π^*, is found by solving the following optimality equation

$$V(S_t) = \max_{x_t\in\mathcal{X}(s_t)}\left(\tilde{c}(S_t, x_t) + \gamma\mathbb{E}\left[V(S_{t+1})|S_t, x_t\right]\right). \tag{16}$$

The policy iteration algorithm is implemented in MATLAB 2020A to solve Equation (16).

4.2 ADP Formulation

The MDP model provides an appropriate mathematical framework for solving the MEDEVAC dispatching problem, but determining an optimal policy utilizing Equation 16 becomes computationally intractable when the cardinality of the state space (i.e., $|\mathcal{S}|$) is too large due to what is commonly referred to as the *curse of dimensionality*. However, as previously mentioned, large-scale problem instances allow for the incorporation of more realistic problem features into the model. To overcome the curse of dimensionality, this thesis leverages an ADP solution strategy to approximate the value function around the post-decision state variable.

Similar to Rettke *et al.* (2016), Robbins *et al.* (2018), and Jenkins *et al.* (2021b,a), a post-decision state convention is adopted to help alleviate the problems associated with the curse of dimensionality. The post-decision state refers to the status of the MEDEVAC system directly after the system is in pre-decision state S_t and action x_t is taken. This information enables the modification of the optimality equation to incorporate the post-decision state convention. Let

$$V^x(S_t^x) = \mathbb{E}\Big[V(S_{t+1})|S_t^x\Big] \tag{17}$$

denote the value of being in post-decision state $S_t^x \in \mathcal{S}^x$, where $\mathcal{S}^x$ is the post-decision state space. By substituting Equation 17 into Equation 16 we obtain the following approximate Bellman equation

$$V(S_t) = \max_{x_t \in \mathcal{X}(S_t)} \Big(\tilde{c}(S_t, x_t) + \gamma V^x(S_t^x)\Big). \tag{18}$$

Noting that the value of being in post-decision state S_{t-1}^x is given by

$$V^x(S_{t-1}^x) = \mathbb{E}\Big[V(S_t)|S_{t-1}^x\Big], \tag{19}$$

and substituting Equation 18 into Equation 19 renders the following optimality equation around the post-decision state

$$V^x(S_{t-1}^x) = \mathbb{E}\Big[\max_{x_t \in \mathcal{X}(S_t)} \tilde{c}(S_t, x_t) + \gamma V^x(S_t^x)|S_{t-1}^x\Big]. \tag{20}$$

Despite the computational advantages attained from utilizing the post-decision state convention, due to the size and dimensionality of the state space, solving Equation 19 remains computationally intractable. As such, an ADP technique is selected utilizing a value function approximation scheme that involves replacing the true value function with a statistical approximation. More specifically, we approximate the value function by leveraging kernel regression within an API algorithmic framework.

4.2.1 Kernel Regression

Kernel regression is an extension of the k-nearest neighbor approximation technique that forms an estimate of the state value function estimate by using a weighted sum of prior observations, which is generally written as

$$\bar{V}^x(\tilde{S}_t^x|\theta) = \frac{\sum_{j=1}^{|S^x|} K_h(\tilde{S}_t^x, S_{t,j}^x)\theta_j}{\sum_{j=1}^{|S^x|} K_h(\tilde{S}_t^x, S_{t,j}^x)}, \tag{21}$$

where $K_h(\tilde{S}_t^x, S_{t,j}^x)$ is a kernel function with bandwidth parameter h that determines the similarity between states $\tilde{S}_t^x$ and $S_{t,j}^x$ where $\tilde{S}_t^x$ corresponds to a specific state of interest and $S_{t,j}^x$ corresponds to a particular state within the post-decision state space. The parameter θ is a $|S^x|$ by 1 column vector of weights (Powell, 2011). There are many possible choices for the kernel function $K_h(\tilde{S}_t^x, S_{t,j}^x)$. The most popular being the Gaussian kernel (i.e, radial basis function), which is given by

$$K_h(\tilde{S}_t^x, S_{t,j}^x) = e^{-\left(\frac{||\tilde{S}_t^x - S_{t,j}^x||}{h}\right)^2},$$

where $||\cdot||$ is the Euclidean norm. The Gaussian kernel provides a smooth, continuous estimate of $\bar{V}^x(\tilde{S}_t^x|\theta)$.

The Gaussian kernel is primarily used on continuous data to provide a measurement of similarity. Due to the nominal nature of the post-decision state variable explored herein, the Aitchison and Aitken (AA) kernel is more appropriate. Aitchison & Aitken (1976) extend the kernel method of density estimation from continuous to multivariate binary space. More specifically the AA kernel is given by

$$K_h(\tilde{S}_t^x, S_{t,j}^x) = \begin{cases} 1-h & (S_{t,j}^x = \tilde{S}_t^x) \\ \frac{(h)}{|Z|} & (S_{t,j}^x \neq \tilde{S}_t^x) \end{cases}. \tag{22}$$

Incorporating the kernel regression approximation into Equation 20 renders the following post-decision state value function approximation

$$\bar{V}(S^x_{t-1}|\theta) = \mathbb{E}\left[\tilde{c}(S_t, X^\pi(S_t|\theta)) + \gamma \bar{V}^x(S^x_t|\theta)|S^x_{t-1}\right], \tag{23}$$

wherein action x_t is determined via the decision function

$$X^\pi(S_t|\theta) = \underset{x_t \in \mathcal{X}(s_t)}{\operatorname{argmax}} \left\{ \tilde{c}(S_t, x_t) + \gamma \bar{V}^x(S^x_t|\theta) \right\}. \tag{24}$$

4.2.2 API

After defining the decision function and the value function approximation setup, the process on how the value function approximation is updated is now presented. As previously stated, this thesis utilizes an API algorithmic strategy, the general structure of which is derived from exact policy iteration, wherein a sequence of value function approximations and policies are generated from two repeated and alternating phases: policy evaluation and policy improvement (Jenkins *et al.*, 2021b). Within the policy evaluation phase, the value of a fixed policy is approximated via simulation. Within the policy improvement phase, a new policy is generated by leveraging the information collected by the previous policy evaluation phase. The API algorithm herein is adapted from Rettke *et al.* (2016) and Jenkins *et al.* (2021a,b) and is displayed in Algorithm 1.

The API-KR algorithm starts by initializing θ, which is a vector corresponding to the weights associated with being in each post-decision state. The policy evaluation phase is then initiated, for each iteration $n = 1, 2, \ldots, N$, the following steps commence. A post-decision state, $S^x_{t-1,j}$, is randomly selected via a Latin hypercube sampling scheme. We then simulate the system evolving from post-decision state $S^x_{t-1,j}$ to a pre-decision state $S_{t,i}$.

Algorithm 1 Approximate Policy Iteration Kernel Regression (API-KR) Algorithm

Initialize θ
for $n = 1$ to N **do**
  for $j = 1$ to J **do**
    Generate a random post-decision state, $S^x_{t-1,j}$
    Determine set of possible pre-decision states $\bar{\mathcal{S}} \subseteq \mathcal{S}$ by utilizing the state transition function $S^{M,W}(S^x_{t-1}, W_j)$
    For each pre-decision state $S_{t,i} \in \bar{\mathcal{S}}$, $i = 1, 2, \ldots, |\bar{\mathcal{S}}|$, solve the approximate optimality equation using Equation (25) and record the estimated value $\hat{v}_{j,i}$ of being in post-decision state $S^x_{t-1,j}$ given the system evolves to pre-decision state $S_{t,i}$
    Determine and record the estimated value $\hat{v}_j$ of being in post decision state $S^x_{t-1,j}$ by computing the probability weighted sum of all $\hat{v}_{j,i}$ values using Equation (26)
  end for
  Update θ utilizing Equations (27)-(29)
end for
Return the approximate value function $\bar{V}^x(\cdot|\theta)$

The set of possible pre-decision states $\bar{\mathcal{S}} \subseteq \mathcal{S}$ is determined by leveraging the random variable W_j, which indicates the timing and type of the next system event. In conjunction with the system model $S^{M,W}(S^x_{t-1}, W_j)$ (i.e., the post-decision state to pre-decision state transition function), this information is used to identify the set of the next possible pre-decision states. The distribution governing W_j is conditioned on the post-decision state and is described by the transition probability function $\hat{p}(\cdot)$. Moreover, it is important to note that the conditional distribution governing W_j given the system is in pre-decision state $S_{t,j}$ and decision x_t is taken, is the same as the conditional distribution governing W_j given the post-decision state $S^x_{t,j}$. This equivalence enables us to utilize the post-decision state convention.

For each pre-decision state $S_{t,i} \subseteq \bar{\mathcal{S}}$ we solve the approximate optimality equation

$$\hat{v}_{j,i} = \tilde{c}(S_{t,i}, X^\pi(S_{t,i}|\theta)) + \gamma \bar{V}^x(S^x_{t,i}|\theta), \tag{25}$$

and record its value, the estimated value of transitioning to post decision state $S^x_{t,j}$

given that the system was previously in pre-decision state $S_{t,i}$. Using this value, we are able to calculate the value $\hat{v}_j$ of being in post-decision state $S^x_{t-1,j}$. To obtain a more accurate value of $\hat{v}_j$, the algorithm computes and records the estimated values of being in all possible pre-decision states that the system could evolve to from post-decision state $S^x_{t-1,j}$. Utilizing this information the value for $\hat{v}_j$ is computed as follows:

$$\hat{v}_j = \sum_{i=1}^{|\bar{S}|} \hat{p}(S_{t,i}|S^x_{t-1,j})\hat{v}_{j,i}. \tag{26}$$

When a policy evaluation phase is complete, we then move into a policy improvement phase wherein, for each iteration $n = 1, 2, \ldots, N$, the following steps occur. The sample estimate of θ is calculated via kernel regression using

$$\hat{\theta}_i = \frac{\sum_{j=1}^{J} K_h(S^x_{t,i}, S^x_{t,j})\hat{v}_j}{\sum_{j=1}^{J} K_h(S^x_{t,i}, S^x_{t,j})} \quad \text{for } i = 1, 2, \ldots, |\mathcal{S}^x| \tag{27}$$

A polynomial stepsize rule is then used to smooth in $\hat{\theta} = [\hat{\theta}]_{j\in\mathcal{S}^x}$ with the previous estimate. The stepsize rule is given by

$$\delta_n = \frac{1}{n^\kappa}, \tag{28}$$

wherein $\kappa \in (0, 1]$. The polynomial stepsize rule δ_n is an extension of the harmonic sequence and greatly impacts the algorithm's rate of convergence and attendant solutions. The rate at which δ_n declines as the policy improvement loop iterates depends on the value of κ. The smaller the value of κ, the slower the rate at which δ_n declines; however, the best value of κ depends on the given problem, making it a parameter that needs to be tuned (Powell, 2011).

Next, θ is updated using the following equation:

$$\theta \leftarrow \delta_n \hat{\theta} + (1 - \delta_n)\theta, \tag{29}$$

wherein the θ on the right hand side is the previous estimate based on the previous policy improvement iterations, and $\hat{\theta}$ is the new estimate from the current iteration. As the number of iterations n increases, we place less emphasis on the sample estimates (i.e., $\hat{\theta}$) and rely more on the estimate based on the first $n-1$ iterations (i.e., θ).

Once θ is updated via Equation 29, we have completed one policy improvement iteration of the API-KR algorithm. If $n < N$ then the algorithm continues by starting another policy evaluation phase. The parameters, $N, J, \kappa,$ and h are tunable, where N is the number of iterations of the policy improvement phase, J is the number of iterations of the policy evaluation phase, κ is the polynomial stepsize parameter, and h is the bandwidth parameter. Concluding after N policy improvement phases, the algorithm provides the recommended policy and approximate value function $\bar{V}^x(\cdot|\theta)$.

V. Testing, Results, & Analysis

This chapter presents a representative MEDEVAC planning scenario set in Southern Afghanistan to demonstrate the applicability of the policies generated by the formulated MDP model and ADP algorithm. This thesis utilizes an Intel Xeon Silver 4114 CPU workstation that has 64 GB of RAM and MATLAB's Parallel Computing Toolbox to conduct the computational experiments and analyses outlined in this chapter.

5.1 Representative Scenario

This thesis considers a notional planning scenario in which the U.S. is performing combat operations in support of the government of Afghanistan. The computational examples in Robbins *et al.* (2018) are closely followed and leveraged as a basis for the representative scenario examined herein. The location of the main coalition bases (i.e., larger bases that are able to host both a MEDEVAC helicopter landing zone (HLZ) and MTF) and forward operating bases (i.e., smaller bases that are only able to host a MEDEVAC HLZ) are established at likely military tactical sites. Figure 2 shows the location of the two main coalition bases that contain an MTF and MEDEVAC HLZ; the two forward operating bases that only have a MEDEVAC HLZ; and the 56 casualty cluster centers that depict the likely locations of violent confrontations between friendly and enemy forces.

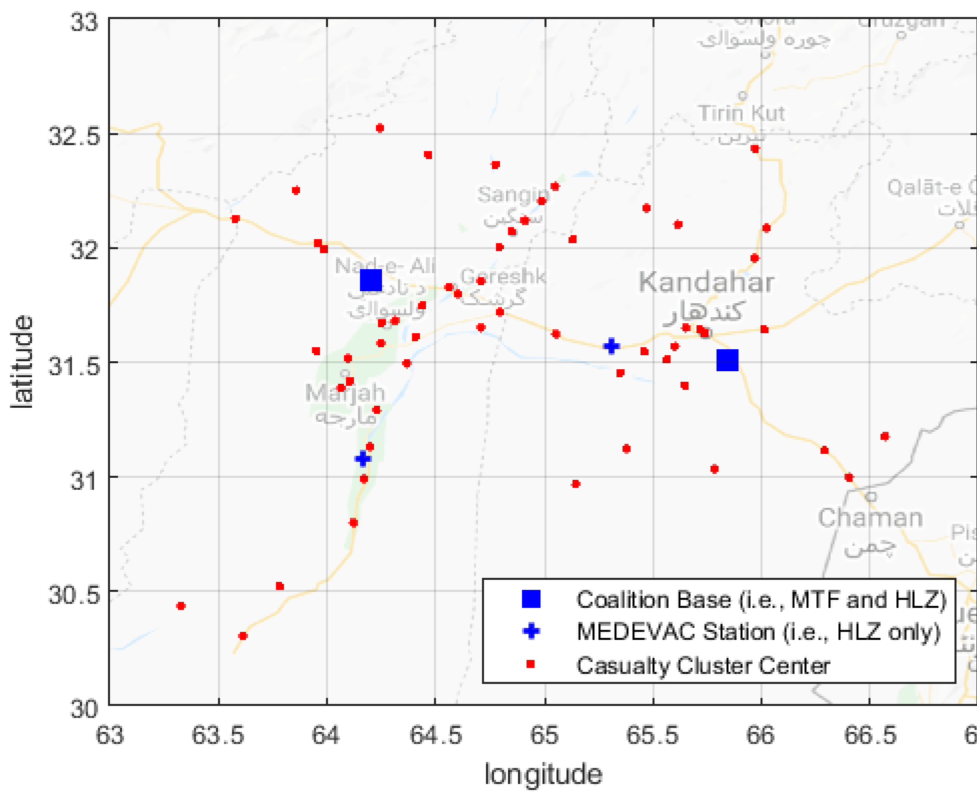


Figure 2. Representative Scenario

The notional scenario utilized in this thesis assumes a MEDEVAC system with four MEDEVAC HLZs (e.g., location of MEDEVAC units) and six demand zones (i.e., zones in which a 9-line MEDEVAC request can originate from) with two MTFs collocated at the main operating bases. The 6-zone problem instance utilizes the information in Figure 2 and tessellates the map into six different zones as depicted in Figure 3. Incorporating this tesselation scheme, along with the four MEDEVAC units and precedence and threat levels, result in the size of the state space being 60,025, which is calculated using Equation 4.

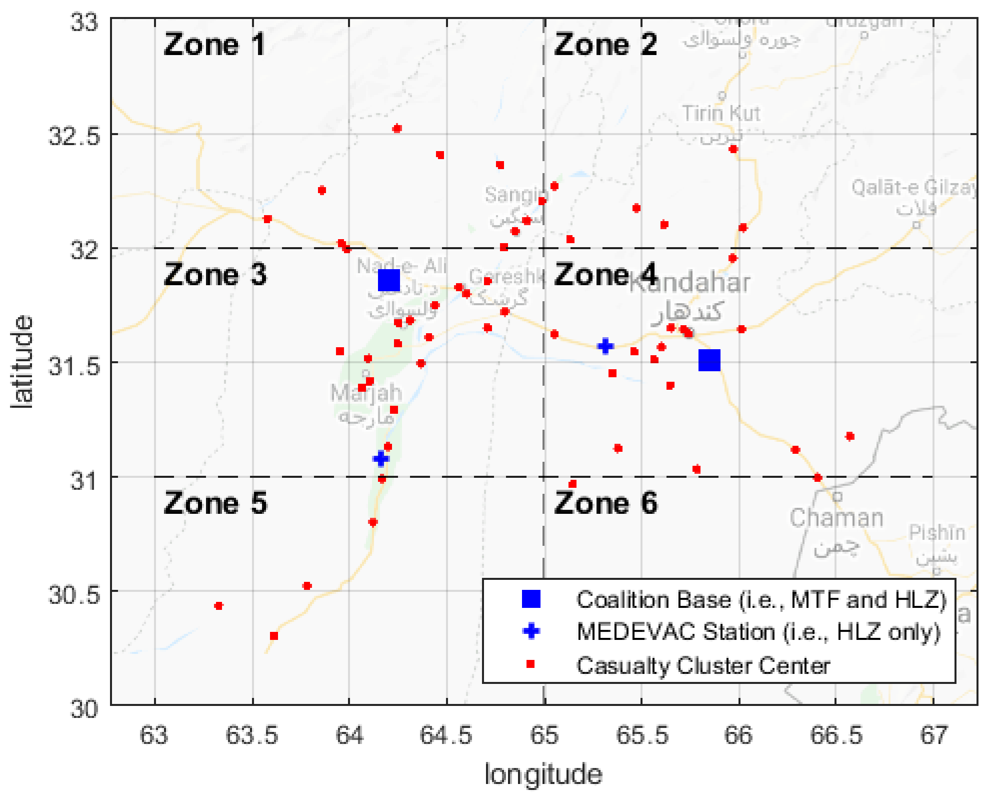


Figure 3. 6-zone Scenario

Each 9-line MEDEVAC request is independently categorized by its zone z, precedence level k (e.g., urgent and priority) and threat level (e.g., high and low). It is assumed that an urgent or priority casualty event is equally likely to take place. For simplicity, this thesis only models two threat levels (i.e., high and low). High threat levels result in scenarios wherein the MEDEVAC unit must be armed or have an armed escort before it arrives at the CCP. Low threat levels result in scenarios wherein armed escorts are not required. Tables 3 and 4 display the proportion of arrivals for both low threat and high threat scenarios, respectively, based on the priority level of the call.

Due to operational security, this thesis avoids using actual data from Afghanistan operations, but instead leverages a Monte Carlo simulation to obtain realistic response

Table 3. Threat level 1 (low threat) proportion of arrivals

	Priority, k	
Zone, z	1 (Urgent)	2 (Priority)
1	0.056	0.056
2	0.038	0.038
3	0.116	0.116
4	0.095	0.095
5	0.028	0.028
6	0.009	0.009

Table 4. Threat level 2 (high threat) proportion of arrival

	Priority, k	
Zone, z	1 (Urgent)	2 (Priority)
1	0.025	0.025
2	0.017	0.017
3	0.052	0.052
4	0.043	0.042
5	0.013	0.012
6	0.004	0.004

and service time data. The means of the response and service times for each zone are displayed in Tables 5 and 6, respectively.

Table 5. MEDEVAC Response Times with low Threats

	MEDEVAC, m			
Zone, z	1	2	3	4
1	62.212	71.726	67.733	86.353
2	83.995	84.797	63.804	72.799
3	52.217	52.513	58.804	77.184
4	79.055	71.555	49.855	51.608
5	104.701	75.499	99.167	123.201
6	94.267	77.125	65.879	71.172

The system receives a reward when a MEDEVAC unit is dispatched to service a 9-line request. These rewards are monotonically decreasing over time and are based

Table 6. MEDEVAC Service Times with low Threats

	MEDEVAC, m			
Zone, z	1	2	3	4
1	62.212	89.529	90.484	126.480
2	113.140	114.765	77.784	83.7808
3	52.245	70.328	81.546	117.282
4	118.815	105.954	60.640	51.9744
5	104.720	93.310	121.91	163.309
6	134.394	111.677	76.554	71.1707

upon the response times listed in Table 5 when the enemy threat is considered low. MEDEVACs 1 and 4 are both Pave Hawk units while MEDEVACs 2 and 3 are Black Hawk units. This distinction becomes especially important when high threat levels are considered. As previously mentioned, Pave Hawks are equipped with machine guns that allow them to enter hostile environments without an armed escort. This weapons load out, however, weighs them down, forcing them to move slower than their Black Hawk counterpart (Grannan *et al.*, 2015). This trade-off is seen when higher threat level calls occur. To account for higher threat levels, it is assumed that when an armed escort is required the Black Hawk units are delayed by approximately 30 minutes. This number is based off of real world accounts and is taken to be a realistic estimate. For simplicity this is directly added to the MEDEVAC units that correspond to Black Hawks (i.e., Units 2 and 3). Higher threat levels do not impact the Pave Hawk units since they are able to enter areas regardless of threat levels.

5.2 Representative Scenario Results

The notional scenario examined in this thesis assumes a high operational tempo (OPSTEMPO), where casualties and subsequent 9-line MEDEVAC requests arrive with a base line request arrival rate of $\lambda = \frac{1}{30}$ to represent a 9-line MEDEVAC arrival request of 1 every 30 minutes. Due to the high OPSTEMPO of the scenario, it is also assumed that high threat level cases are more likely to occur than low threat level cases, rendering a proportion of 0.6 for high threat level cases and 0.4 for low threat level cases. Table 7 depicts the list of parameters associated with this scenario along with their description and settings.

Table 7. Baseline parameter settings

Parameter	Description	Setting
λ	9-line MEDEVAC request arrival rate	$\frac{1}{30}$
$\lvert\mathcal{M}\rvert$	# of MEDEVAC units	4
$\lvert\mathcal{Z}\rvert$	# of zones	6
$\lvert\mathcal{K}\rvert$	# of precedence categories	2
$\lvert\zeta\rvert$	# of threat level categories	2
γ	Uniformized time discount rate	0.99
η	Weight for urgent requests	10

For comparison purposes a baseline myopic policy is developed that focuses on minimizing the response time that the MEDEVAC unit has to reach a 9-line casualty request. This policy may appear optimal upon first glance; however, without considering the implications of future requests we show that significant improvements can be made.

Table 8 compares the expected total discounted reward (ETDR) obtained by both the optimal and myopic policies when the system is in an idle state (i.e., $S_0 = \big((0,0,0,0),(0,0,0)\big)$). In general, the optimal policy prioritizes urgent level requests (i.e., $\hat{K}_t = 1$) by reserving Pave Hawk assets to only dispatch when the request is urgent and the threat level is high (i.e., $\hat{L}_t = 2$). In scenarios where all but one MEDEVAC asset is busy, the optimal policy will choose to reject priority level calls rather than service them like the myopic policy. There are some exceptions to this, however this is the general structure.

As shown in Table 8, the optimal policy significantly outperforms the myopic for the 6-zone problem instance with respect to the ETDR at the idle state (i.e.,

Table 8. Policy Comparison

Policy,π	$V^{\pi}(S_0)$	Optimality Gap
Optimal	40.954	N/A
Myopic	38.915	4.97%

$V^{\pi}(S_0)$). However, as the state space grows, the MEDEVAC dispatching problem becomes computationally intractable to solve to optimality. To alleviate this issue, we leverage the ADP algorithm developed in Chapter IV and compare it to the optimal and myopic policies to demonstrate its efficacy.

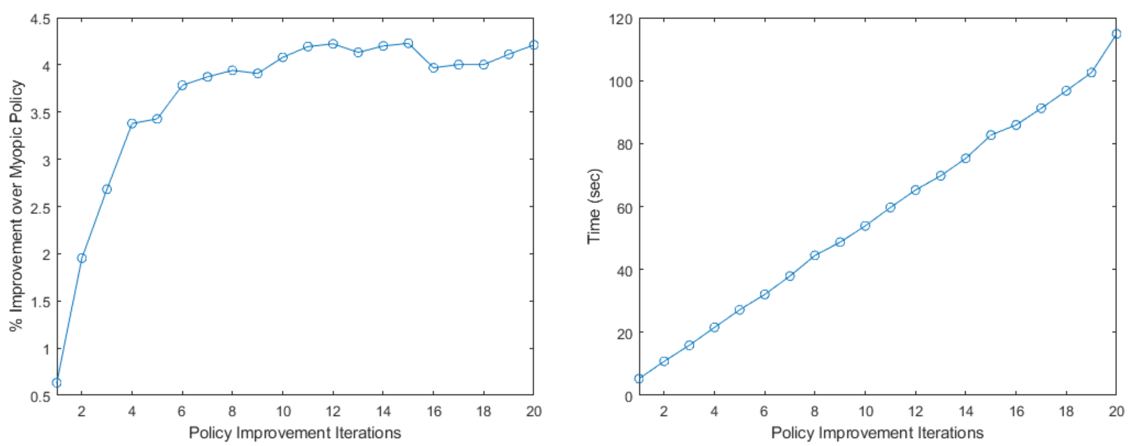
5.2.1 Algorithmic Experimental Design

Before this comparison can commence, as stated in Chapter IV, the ADP algorithm has several parameters that need to be properly tuned to render high-quality performance. To accomplish this, we conduct a full factorial 4^4 experiment. Table 9 shows the set of parameters along with the range in which they were examined. The chosen parameter test ranges are based upon previous research and the knowledge of how the kernel regression algorithm performs. In particular, the parameter J, number of policy evaluation iterations, was chosen based upon the size of the post-decision state space, $|\mathcal{S}^x|$. The selected values corresponds to 20-80% of $|\mathcal{S}^x|$. This was chosen in order to provide enough policy evaluation iterations to render the best policy, while simultaneously balancing computational run time.

Table 9. Experimental Design Factor Levels

Algorithm Parameters	Description	Levels
N	Policy Improvements	{10, 20, 30, 40}
J	Policy Evaluations	{481, 962, 1441, 1921}
κ	Step-size	{0.3, 0.5, 0.7, 0.9}
h	Kernel Bandwidth	{0.01, 0.03, 0.1, 0.3}

The experiment is ranked based on the best value rendered when the system is in the idle state. In total there were 256 different parameter pairings evaluated, the top 20 are displayed in Table 10. Looking at Table 10 we see that the top 20 results are all high quality policies that render results that are at least 98% optimal. Based on the values rendered by this experiment, it appears that the algorithm performs best with the following settings: $\kappa = 0.3$, $h = 0.01$, $J = 1921$, and $N = 20$. It is important to note that with multiple replications, the rankings have potential to change. However, due to time constraints only one replication is conducted. Despite being less than ideal, this result provides a good outline of which parameter settings allow the algorithm to perform the best. In particular it can be assumed that higher N values, policy improvement iterations, will yield better results. However, looking at Figure 4a we see that as N increases the percent improvement over the myopic policy begins to level out. Figure 4b shows the trade off between having higher N iterations substantially increases the computational run time.



(a) % Improvement **(b) Computation Time**

Figure 4. ADP Algorithm Performance

Once the experiment is complete, the best value parameter settings are selected along with their corresponding value. Table 11 compares the obtained ADP value to the optimal and myopic policy values. Despite being sub optimal the ADP algorithm is able to generate a policy that is over 99% optimal and outperforms the myopic by

Table 10. Experimental Design Results

κ	h	J	N	$V^{\pi}(S_0)$	% Optimal
0.3	0.01	1921	20	40.637	99.225
0.3	0.01	1921	30	40.593	99.118
0.5	0.01	1921	20	40.583	99.094
0.5	0.01	1921	30	40.571	99.064
0.5	0.01	1921	40	40.568	99.056
0.3	0.01	1921	40	40.555	99.023
0.7	0.01	1921	40	40.536	98.978
0.7	0.01	1921	30	40.528	98.959
0.3	0.03	1921	20	40.523	98.946
0.7	0.01	1921	20	40.521	98.942
0.3	0.03	1921	30	40.499	98.887
0.5	0.03	1921	20	40.488	98.861
0.5	0.03	1921	30	40.486	98.857
0.9	0.01	1921	40	40.480	98.840
0.3	0.01	1441	30	40.464	98.802
0.3	0.03	1921	40	40.458	98.788
0.9	0.01	1921	30	40.456	98.782
0.3	0.01	1441	20	40.455	98.779
0.5	0.03	1921	40	40.453	98.776
0.5	0.01	1441	30	40.447	98.760

4.47%.

Table 11. Value Comparison

Policy,π	$V^{\pi}(S_0)$	Optimality Gap
Optimal	40.954	N/A
ADP	40.637	0.77%
Myopic	38.915	4.97%

5.2.2 Policy Comparison

Four scenarios are examined in detail to provide a better understanding of the differences between the decisions generated by the optimal, ADP, and myopic policies.

The chosen scenarios were picked to represent each possible MEDEVAC status (i.e., all MEDEVAC units are idle, 3 MEDEVAC units are idle, 2 MEDEVAC units are idle, or only 1 MEDEVAC unit is idle). Table 12 provides a description of each of

Table 12. Scenario Description

Scenario	Description	States
A	All MEDEVACs are idle	$S_t \in \left\{(0,0,0,0),(\hat{Z}_t,\hat{K}_t,\hat{L}_t)\right\}$
B	MEDEVAC 1 is serving zone 1	$S_t \in \left\{(1,0,0,0),(\hat{Z}_t,\hat{K}_t,\hat{L}_t)\right\}$
C	MEDEVACs 1 and 2 are servicing zone 1	$S_t \in \left\{(1,1,0,0),(\hat{Z}_t,\hat{K}_t,\hat{L}_t)\right\}$
D	MEDEVACs 1, 2, and 3 are servicing zone 1	$S_t \in \left\{(1,1,1,0),(\hat{Z}_t,\hat{K}_t,\hat{L}_t)\right\}$

the chosen scenarios. Table 13 displays the optimal decisions for each scenario based upon the corresponding request status tuple. One asterisk (i.e., *) indicates that the particular decision differs from the myopic policy, two asterisks (i.e., **) indicate that the decision differs from the ADP policy, and three asterisks indicates a difference in all three policies (i.e., ***).

The results from Table 13 show that while the optimal and myopic policies have significant overlap, there are particular requests for which where the optimal policy deviates from acting myopically, rendering higher performance. As previously stated, the optimal policy chooses to prioritize urgent level requests in order to maximize the reward obtained by the system. For example, if the request is urgent, the optimal policy will choose to send the MEDEVAC with the lowest response time. If the request for service is priority, pending on how many MEDEVAC units are available and which zone the call arrived from, it most likely will choose to send a Black Hawk or reject the call. Another interesting result occurs when the only idle MEDEVAC is MEDEVAC 4 and a request comes from Zone 5 (i.e., $S_t \in \{(1,1,1,0),(5,\hat{K}_t,\hat{R}_t)\}$). Due to the very high response time MEDEVAC 4 has when servicing Zone 5, as shown in Table 5, the optimal policy chooses to ignore (i.e., reject) these calls regardless of the precedence level of the request. Despite there being some distinct differences, the

Table 13. Optimal Decisions

	Scenario			
Request Status, $\hat{R}_t$	A	B	C	D
(1,1,1)	1	3**	3	4
(1,1,2)	1	3***	3*	4
(1,2,1)	2***	3**	3	Δ*
(1,2,2)	2***	3**	3*	Δ*
(2,1,1)	3	3	3	4
(2,1,2)	4	4	4	4
(2,2,1)	3	3	3	Δ*
(2,2,2)	3*	3*	3*	Δ*
(3,1,1)	1	2	3	4
(3,1,2)	1	2	3***	4
(3,2,1)	2*	2	3	Δ*
(3,2,2)	2*	2	3	Δ*
(4,1,1)	3**	3*	4	4
(4,1,2)	4	4	4	4
(4,2,1)	3*	3*	4	4
(4,2,2)	4	4	4	4
(5,1,1)	2	2	3	Δ***
(5,1,2)	2	2	3***	Δ***
(5,2,1)	2	2	Δ*	Δ*
(5,2,2)	2	2	Δ*	Δ*
(6,1,1)	3	3	3	4
(6,1,2)	4	4	4	4
(6,2,1)	3	3	3	Δ*
(6,2,2)	3*	3*	4**	Δ*

ADP policy follows a structure that is very similar to that of the optimal policy.

The workload of each MEDEVAC unit is another interesting measurement of performance that varies among each of the policies. Figure 5 and Table 14 display the long run busy probabilities for each of the MEDEVAC units under the different policies. Examination of Figure 5, indicates that the optimal policy prefers to dispatch MEDEVACs 2 and 3 (i.e., Black Hawks) more than MEDEVACs 1 and 4 (i.e., Pave

Hawks). As previously stated, the optimal policy prefers to hold onto the Pave Hawk assets and only dispatch them when the threat level is high and the precedence level is urgent. If the precedence level is priority the optimal policy will generally choose to either dispatch a Black Hawk or reject the call. However, the myopic policy chooses to dispatch regardless of the precedence of level. Due to the high threat combat environment, the myopic policy will choose to dispatch the Pave Hawks more often since they can move faster under high threat levels.

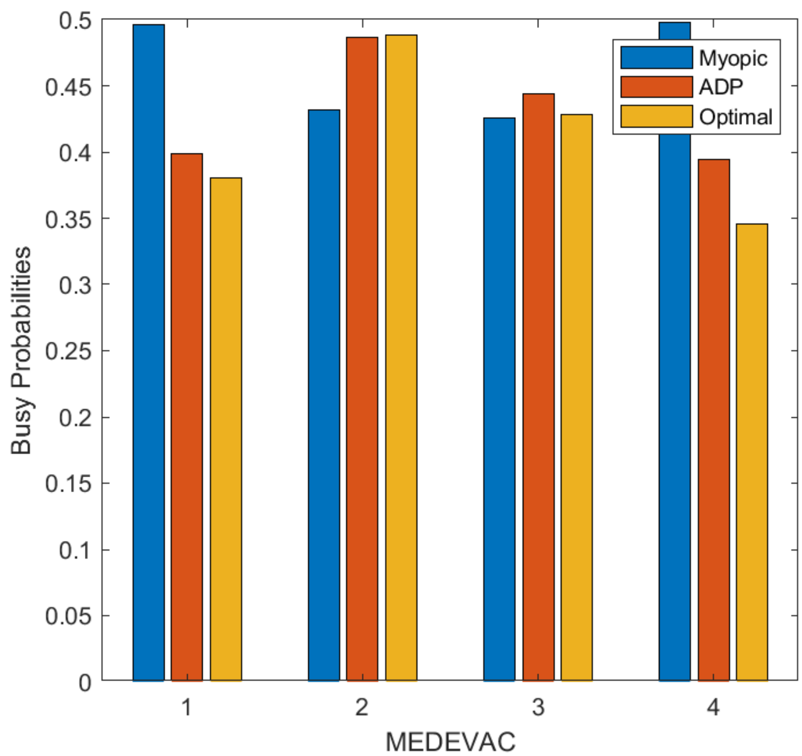


Figure 5. MEDEVAC Busy Rates by Policy

Table 14. MEDEVAC Busy Rates

	Policies		
MEDEVAC	Myopic	ADP	Optimal
1	0.50	0.40	0.38
2	0.43	0.49	0.49
3	0.43	0.44	0.43
4	0.50	0.39	0.35

Examination of Figure 5 and Table 14 indicates that the busy rates generated by the ADP policy are similar in structure to that of the optimal policy, which corresponds to our previous results. Similar to the optimal policy, the ADP policy reduces the busy rates of MEDEVACs 1 and 4 by choosing to save them for high threat and urgent precedence calls. As such, the optimal and ADP policies reduce the amount that each of the MEDEVAC units are used, decreasing the overall burden to the system.

5.3 Excursions

This section explores how the optimal, ADP, and myopic policies change as key problem features of the representative scenario are varied. Unless otherwise stated, the following excursions utilize the same baseline problem features that are listed in Table 7.

5.3.1 Excursion 1 - Request Arrival Rate

The casualty arrival rate, λ, is an important problem feature that severely impacts the MEDEVAC system. To examine the impact of this problem feature, this section explores the values obtained by all three policies when the MEDEVAC is in the idle state. The same parameter settings displayed in Table 7 are used in this excursion, with exception of the arrival rate, λ. Table 15 displays the different λ values that are examined along with the corresponding idle state values for the optimal, myopic, and ADP policies.

The results demonstrate that when the casualty arrival rate begins to decrease, the optimal and ADP policies render values that are very similar to the myopic policy. However, when the casualty arrival rate is high, the optimal and ADP policies vastly outperform the myopic. Figure 6 provides a visual representation of this effect.

Table 15. Arrival Rate Impact on $V^{\pi}(S_0)$ for the MEDEVAC System

	Policies		
$\frac{1}{\lambda}$	Optimal	ADP	Myopic
10	62.38	58.37	43.47
20	51.20	50.20	45.09
30	40.95	40.63	38.91
40	32.84	32.67	32.05
50	26.73	26.61	26.33
60	22.17	22.07	21.92

Overall, it appears that the optimal and ADP policies provide little value over the myopic policy when the operations tempo is low (i.e., $\lambda = \frac{1}{60}$). However, when the operations tempo is high (i.e., $\lambda = \frac{1}{10}$) a significant improvement can be demonstrated.

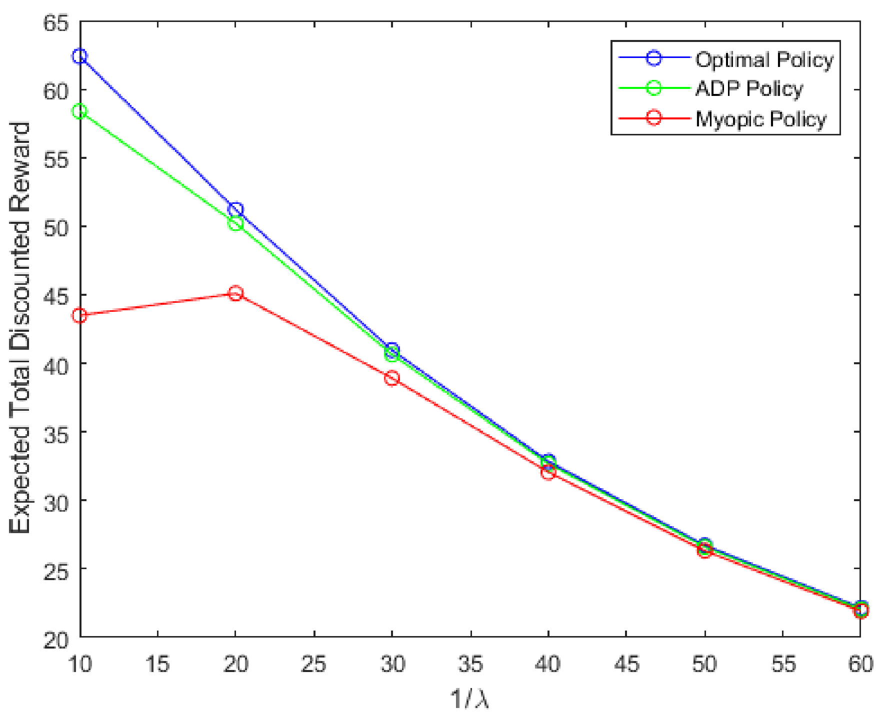


Figure 6. Arrival Rate Impact on MEDEVAC System

5.3.2 Excursion 2 - MEDEVAC Asset Types

The U.S. Army and Air Force currently employ Black Hawk and Pave Hawks, respectively, to conduct MEDEVAC and personnel recovery missions. However, the replacements for these aircraft are currently in development and are beginning to be implemented. The assets currently competing to replace the Army's Black Hawk units are the Sikorsky-Boeing 1 (SB-1) Defiant and the Bell V-280 Valor. Jenkins *et al.* (2021a) demonstrate that the V-280 Valor renders higher performance than the SB1-Defiant. Following this, the V-280 Valor will be the main focus of analysis for this excursion. For the Air Force, the HH-60W's (i.e., Whiskeys) are designed to replace their current fleet of Pave Hawks rendering increased flight time and speed. Using the information collected from Jenkins *et al.* (2021a) and Grannan *et al.* (2015), the speeds for these new assets are approximated and simulated to determine their impact on the system. We conduct two experiments to determine how the implementation of these new assets impact the MEDEVAC system.

The first experiment examines how the system performs when all four assets are replaced. Although ideal, the replacement of all four assets may not be feasible due to either time or monetary constraints. The second experiment takes this into consideration and only looks at the replacement of one asset for each service branch. For this experiment, Assets 1 and 2 are replaced for the Air Force and Army respectively. These assets were selected due to their high utilization rates depicted in Figure 5 and Table 14. The percent increase over the myopic policy for both experiments, with regards to ETDR are displayed in Table 16.

With the improved flight speeds, the MEDEVAC assets are able to service requests and become idle again at a faster rate. This improves the myopic policy's performance and decreases the percent improvement gained from the ADP and optimal policies. With that said, significant improvement can still be gained from implementing either

Table 16. **Percent Increase over Myopic policy with Asset Replacement**

	Policies	
Category	Optimal	ADP
Full Replacement	4.15%	3.30%
Half Replacement	4.97%	4.13%
No Replacement	5.24%	4.47%

the optimal or ADP policies.

5.3.3 Excursion 3 - Threat Level Proportion

The proportion of high threat level calls (i.e., $l = 2$) that enter the system can vastly change how each of the policies perform. If this proportion is increased, then the Pave Hawk units have a more significant role due to their decreased response times. However, if this proportion is decreased, the Black Hawks will most likely be burdened with the majority of the work load. To explore this, the parameters settings displayed in Table 7 are held constant with the exception of the proportion threat level (i.e., p_{zk1} and p_{zk2}). Keeping everything else constant, as the proportion of high threat level calls increased, the more the optimal and ADP policies began to outperform the myopic. Table 17 and Figure 7 depict the upward trend corresponding to an increase in high threat level calls.

Table 17. Percent Increase over Myopic policy with respect to High Threat Proportion

	Policies	
p_{zk2}	Optimal	ADP
0.3	3.88%	2.91%
0.4	4.28%	3.36%
0.5	4.63%	3.88%
0.6	5.24%	4.42%
0.7	6.01%	4.78%
0.8	6.90%	5.82%
0.9	7.88%	6.54%

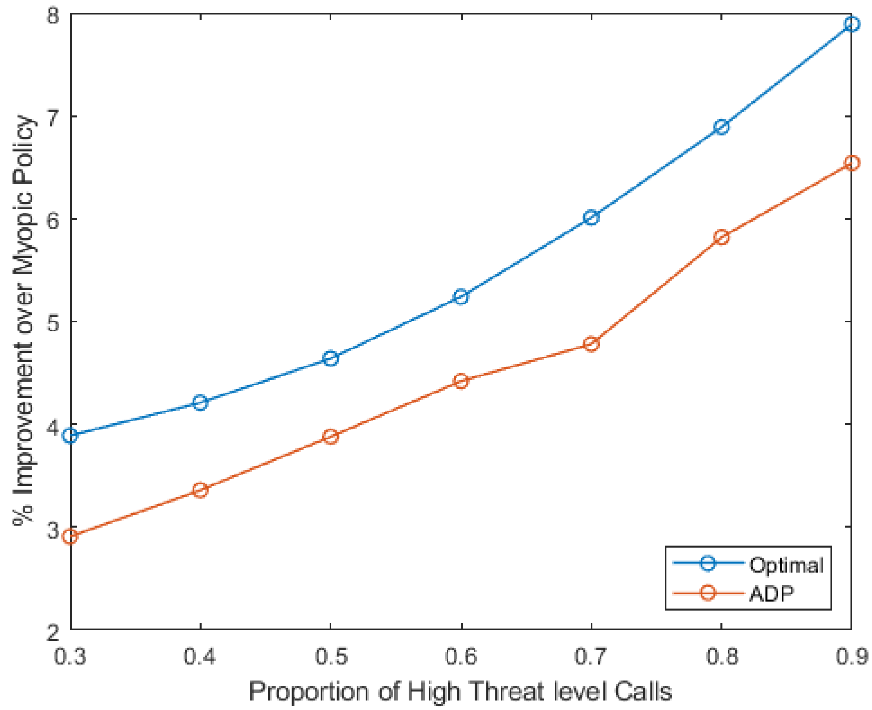


Figure 7. High Threat Proportion Impact on MEDEVAC System Performance

The proportion of high threat arrivals significantly impacts the MEDEVAC system. As the proportion of high threat level requests increases, the adoption of the optimal or ADP policies can render significant improvement over the myopic policy.

5.4 12-zone Scenario Results

This scenario expands the original 6-zone case by expanding the original zone tessellation scheme to incorporate 6 additional zones. The MTFs and MEDEVAC stations are located in the same positions as in the original problem instance. Figure 8 provides a visual representation of the new 12-zone tessellation scheme.

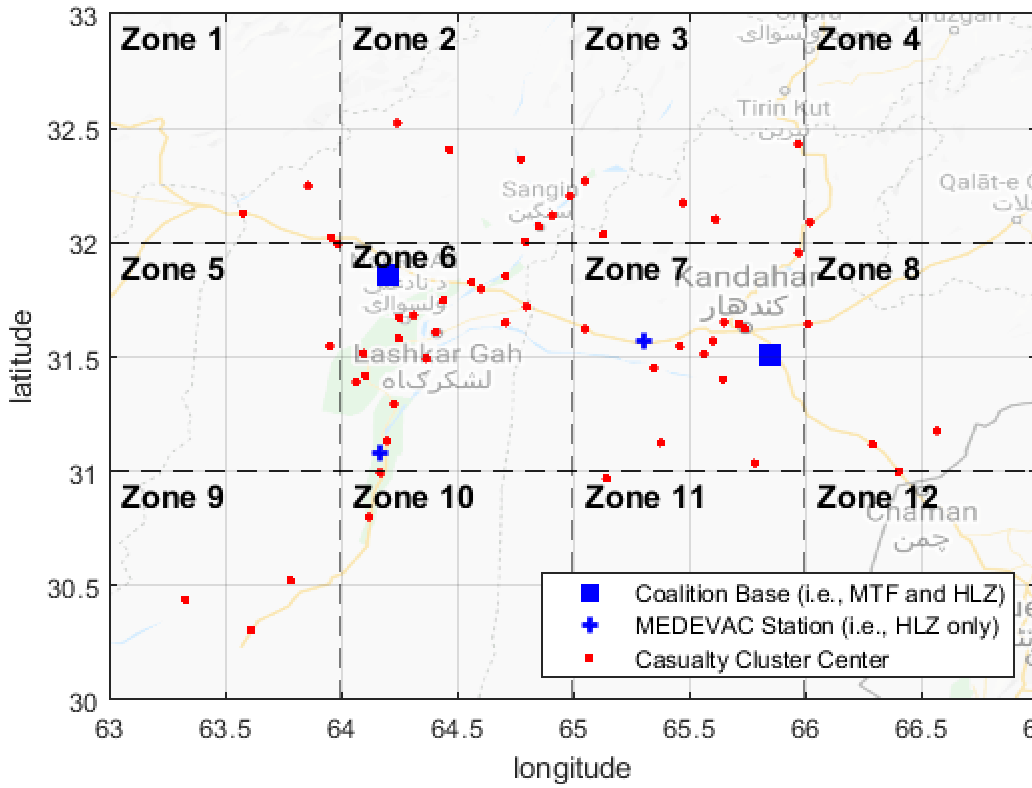


Figure 8. 12-zone Scenario

Similar to the 6-zone case, a Monte Carlo simulation is utilized to determine the service rates, response times, and casualty arrival probabilities by zone. Utilizing Equation 4, the cardinality of the state space $|\mathcal{S}|$ is determined to be 1,399,489. While significantly larger than the original 6-zone case, this problem instance is still small enough to be solved to optimality, but large enough to provide convincing results. The list of parameters associated with this new problem instance is displayed in Table 18.

Due to time constraints, the best performing ADP parameter settings used for the 6-zone case are used again for this problem instance. Moreover, if granted more time, a proper experiment could be conducted to properly tune this algorithm for the new problem instance. However, these parameter settings provide a baseline for the algorithm's performance. Table 19 compares the values obtained by the optimal,

Table 18. 12-zone parameter settings

Parameter	Description	Setting
λ	9-line MEDEVAC request arrival rate	$\frac{1}{30}$
$\lvert\mathcal{M}\rvert$	# of MEDEVAC units	4
$\lvert\mathcal{Z}\rvert$	# of zones	12
$\lvert\mathcal{K}\rvert$	# of precedence categories	2
$\lvert\zeta\rvert$	# of threat level categories	2
γ	Uniformized time discount rate	0.99
η	Weight for urgent requests	10

ADP, and myopic policies for this problem instance when the system is in the idle state. The optimal policy is able to obtain a 5.38% improvement over the myopic

Table 19. 12-zone Results

Policy,π	$V^{\pi}(S_0)$	Optimality Gap
Optimal	38.687	N/A
ADP	38.323	0.059%
Myopic	36.709	5.113%

policy and the ADP policy is able to obtain a 4.39% improvement. These results indicate that the optimal and ADP algorithms scale well to the increased problem size.

Similar to the 6-zone case, we compare the decisions generated by all three policies to gain insight into how they each differ from one another. The scenarios outlined in Table 12 are used for comparison for this problem instance. Table 20 displays the optimal decisions for each scenario based upon the corresponding request status tuple. The asterisk notation defined prior is utilized again to differentiate between the different policies. Despite being a larger problem instance the optimal policy generated in the 12-zone case follows a similar pattern as the one in the 6-zone case. The optimal policy chooses to reserve Pave Hawk units to save them for high-threat level and urgent requests. If a priority request arrives from the same zone in which a

Pave Hawk asset is located, the system will choose to dispatch the Pave Hawk unit due to how fast it can service the request and return to being idle. There are some exceptions to this rule, but this is the general structure of the policy.

Table 20. 12-zone Optimal Decisions

	Scenario			
Request Status, $\hat{R}_t$	A	B	C	D
(1,1,1)	1	2	3	4
(1,1,2)	1	2	3***	4
(1,2,1)	1	2**	3	Δ*
(1,2,2)	1	2**	3	Δ*
(2,1,1)	1**	3	3	4
(2,1,2)	1	3***	3*	4
(2,2,1)	2***	3**	3	Δ*
(2,2,2)	2***	3***	3*	Δ*
(3,1,1)	3	3	3	4
(3,1,2)	4	4	4	4
(3,2,1)	3	3	3	Δ*
(3,2,2)	3*	3*	3*	Δ*
(4,1,1)	3	3**	3	4
(4,1,2)	4	4	4	4
(4,2,1)	3*	3***	3***	Δ***
(4,2,2)	3***	3***	4	Δ***
(5,1,1)	1	2	3	4
(5,1,2)	1	2	3*	4
(5,2,1)	2***	2	3	Δ*
(5,2,2)	1	2	3	Δ*
(6,1,1)	1	2	3	4
(6,1,2)	1	2*	3***	4
(6,2,1)	2***	2	3	Δ*
(6,2,2)	2***	2	3	Δ*
(7,1,1)	3	3	3	4
(7,1,2)	4	4	4	4
(7,2,1)	3*	3***	4**	4
(7,2,2)	3***	4	4**	4
(8,1,1)	4	4	4	4
(8,1,2)	4	4	4	4
(8,2,1)	4	4**	4	4
(8,2,2)	4	4**	4	4
(9,1,1)	2	2	3	Δ*
(9,1,2)	2	2	Δ***	Δ*
(9,2,1)	2	2	Δ***	Δ*
(9,2,2)	2	2	Δ*	Δ*
(10,1,1)	2	2	3	Δ***
(10,1,2)	2*	2	3*	Δ***
(10,2,1)	2	2	3	Δ*
(10,2,2)	2	2	Δ***	Δ*
(11,1,1)	3	3	3	4
(11,1,2)	4	4	4	4
(11,2,1)	3	3	3	Δ*
(11,2,2)	3***	3*	3***	Δ*
(12,1,1)	4	4	4	4
(12,1,2)	4	4	4	4
(12,2,1)	3*	3*	4	Δ***
(12,2,2)	4**	4**	4	Δ***

We next examine the workload for each MEDEVAC unit under the different policies to determine how the workload for the MEDEVAC units scale with a larger problem instance. Figure 9 and Table 21 display the long run busy probabilities for each of the MEDEVAC units under the three different policies. The utilization rates for each of the MEDEVAC units in the 12-zone case are similar to those displayed in Figure 5 and Table 14 for the 6-zone case.

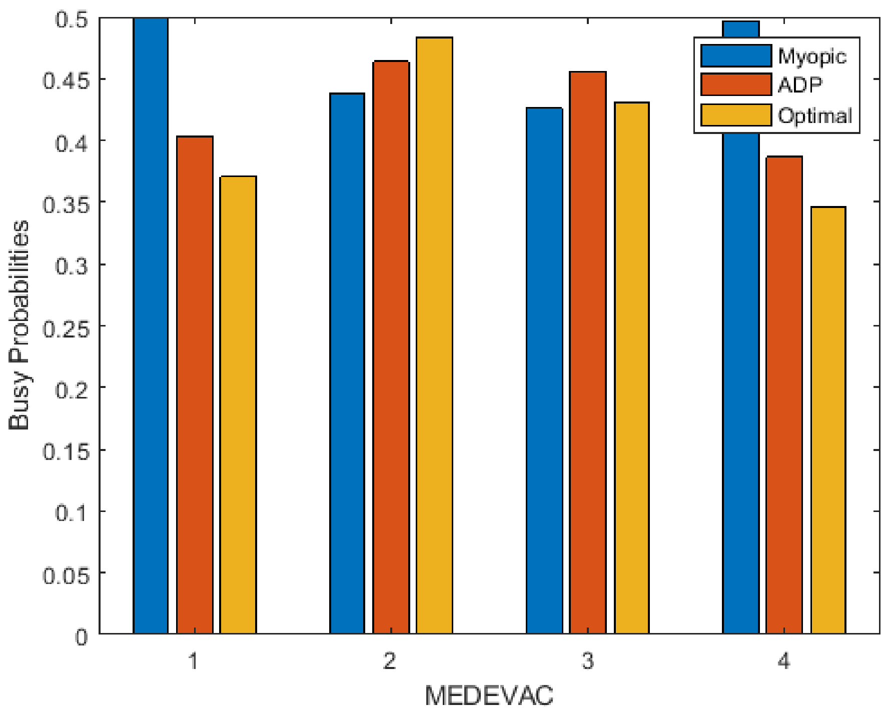


Figure 9. MEDEVAC Busy Rates by Policy

Table 21. MEDEVAC Busy Rates

	Policies		
MEDEVAC	Myopic	ADP	Optimal
1	0.50	0.40	0.37
2	0.44	0.46	0.48
3	0.42	0.45	0.43
4	0.50	0.39	0.35

For both scenarios, the optimal and ADP policies are able to reduce the overall workload of the MEDEVAC system. Overall the structure of the policies generated by the 12-zone case appears very consistent with the policies generated by the 6-zone case.

5.5 34-zone Scenario Results

To demonstrate the scalability of our solution approach, this section develops and examines the 34-zone problem instance by altering the zone tessellation scheme to expand the original problem instance from 6 zones to 34. Figure 10 provides a visual representation of the 34-zone tessalation scheme. The location of the MTFs and MEDEVAC stations are the same.

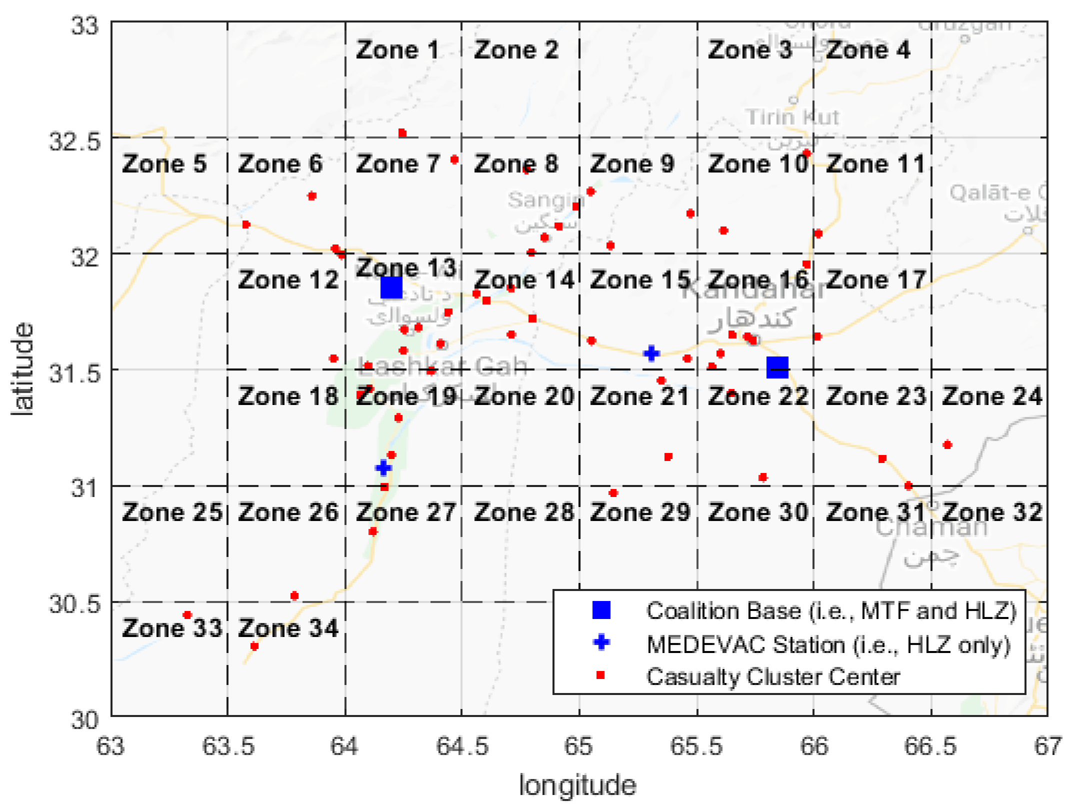


Figure 10. 34-zone Scenario

Similar to the 6-zone and 12-zone cases, a Monte Carlo simulation is used to determine the service rates, response times, and zone probabilities associated with a 9-line request. Table 22 displays a list of parameter settings used to develop the 34-zone case.

Table 22. 34-zone parameter settings

Parameter	Description	Setting
λ	9-line MEDEVAC request arrival rate	$\frac{1}{30}$
$\lvert\mathcal{M}\rvert$	# of MEDEVAC units	4
$\lvert\mathcal{Z}\rvert$	# of zones	34
$\lvert\mathcal{K}\rvert$	# of precedence categories	2
$\lvert\zeta\rvert$	# of threat level categories	2
γ	Uniformized time discount rate	0.99
η	Weight for urgent requests	10

Applying Equation 4 reveals that the cardinality of the state space, $|\mathcal{S}|$, for this problem instance is 205,585,625. Whereas this problem expands the baseline problem instance into a more realistic scenario, it is now too large to be solved to optimality. To overcome this issue, we utilize the ADP algorithm displayed in Algorithm 1 along with the best parameter settings depicted in Table 9 to generate a policy. Under normal circumstances, this solution approach would be enough to generate results that outperform the myopic policy. However, due to the size of the problem, modifications are made to the algorithm to render results within a reasonable computation time.

5.5.1 Algorithm Modification

Due to the large problem size, the original parameter settings used are modified to render results in a computationally reasonable time. This new problem instance is too large to be solved using the original formulation displayed in Algorithm 1. In particular, the cardinality of the new post-decision state space, $|\mathcal{S}^x|$, is 1,500,625. Due to this large size, it is computationally expensive for the kernel regression algorithm to approximate θ using the entire range of post-decision states. To alleviate this issue, a dictionary that contains a specific range of reference points, D, is used to approximate the desired theta values. We also utilize a trajectory sampling scheme that focuses on the idle state. For this reason, the dictionary primarily contains a set of reference points corresponding to the states that the system is most likely to transition to from the idle state, substantially reducing the size of the vector θ. Moreover, we also utilize a substantially smaller amount of policy evaluation iterations, J. It is important to note that better results may be garnered with proper tuning and experimentation. However, due to time constraints, our efforts are focused on developing new results that outperform the myopic policy to demonstrate the efficacy of our solution approach. The modified algorithm and new parameter settings are displayed in Algorithm 2 and Table 23 respectively.

Algorithm 2 Trajectory Following KR-API Algorithm

Generate set of dictionary points D
Initialize θ
for $n = 1$ to N **do**
    Set $S^x_{t-1,j} = S_0$
    for $j = 1$ to J **do**
        Determine set of possible pre-decision states $\bar{\mathcal{S}} \subseteq \mathcal{S}$ by utilizing the state transition function $S^{M,W}(S^x_{t-1}, W_j)$
        For each pre-decision state $S_{t,i} \in \bar{\mathcal{S}}$, $i = 1, 2, \ldots, |\bar{\mathcal{S}}|$, solve the approximate optimality equation using Equation (25) and record the estimated value $\hat{v}_{j,i}$ of being in post-decision state $S^x_{t-1,j}$ given the system evolves to pre-decision state $S_{t,i}$
        Determine and record the estimated value $\hat{v}_j$ of being in post decision state $S^x_{t-1,j}$ by computing the probability weighted sum of all $\hat{v}_{j,i}$ values using Equation (26)
        Simulate the transition to next pre-decision state $S_{t,j}$
        Randomly select a feasible action and transition $S_{t,j}$ to $S^x_{t,j}$
        Set $S^x_{t-1,j} = S^x_{t,j}$
    end for
    Update θ utilizing Equation 29
end for
Return the approximate value function $\bar{V}(\cdot|\theta)$

Table 23. 34-zone Algorithm Settings

Algorithm Parameters	Description	Levels
N	Policy Improvements	5
J	Policy Evaluations	5000
$\|D\|$	Dictionary size	7501
κ	Step-size	0.3
h	Kernel Bandwidth	0.01

5.5.2 34-zone Problem Results

A simulation is developed to evaluate and compare the ETDR generated by both the myopic and ADP policies when in the idle state. We simulate a trajectory of $1,500$ events to produce a reasonable approximation. Moreover, because we are discounting, the γ term in our objective function becomes so small that longer simulation trajectories do not impact the measure of performance. The simulation is conducted $1,000$ times to reduce variation.

The value rendered by each simulation corresponds to the value of the system when in the idle state (i.e., $\bar{V}^{\pi}(S_0|\theta)$). The average value across each of the $1,000$ runs for both the myopic and ADP policies are generated and displayed in Table 24.

Table 24. 34-zone Results

Policy	$\bar{V}^{\pi}(S_0\|\theta)$	% Improvement
Myopic	29.856	
ADP	30.198	1.144 $\pm$0.792

As shown in Table 24 the ADP algorithm is able to generate a 1.144% improvement across the 1,000 runs. A confidence interval for the comparison of two means is conducted and renders the bounds of $(0.0108, 0.6725)$, indicating that there is a statistical difference between the average values generated by both policies. While there is statistical significance, after considering all of the modifications and training needed to be done, a 1% improvement may not be of practical significance to a decision maker. With that said, this improvement demonstrates that our modified algorithm can still render results that outperform the myopic policy. Given more time, a proper experiment can be conducted to properly tune the parameter settings of the algorithm to render improved performance.

VI. Conclusions & Recommendations

This thesis examines the MEDEVAC dispatching problem. The objective of this research is to develop policies that improve the performance of the MEDEVAC system, ultimately increasing the survivability of battlefield casualties. The research conducted herein extends the work performed by prior researchers by incorporating two different problem features that have yet to be explored.

6.1 Conclusions

The research in this thesis adds to the MEDEVAC dispatching literature by developing an MDP model that explicitly incorporates the threat level at the CCP and different MEDEVAC asset types, problem aspects that have yet to be considered, as well as previously examined problem features (e.g., admission control).

As demonstrated by previous research, solving large scale problem instances are important to demonstrate the efficacy of the developed solution techniques. To accomplish this an API algorithm that utilizes a kernel regression approximation scheme to approximate the post-decision state value function is developed and utilized herein. To demonstrate the applicability of our MDP model and ADP algorithm, a notional scenario is constructed based upon combat operations in southern Afghanistan. Three different problem instance sizes of this scenario were examined, the 6-, 12- and 34-zone cases. Our modeling and solution approach identified high-quality MEDEVAC dispatching policies in a computationally efficient manner that scaled well with increased problem size.

The results demonstrate that the optimal and ADP policies are able to obtain a 5.24% and 4.42% improvement over the myopic policy respectively in regards to the ETDR of the idle state in the 6-zone problem instance. For the 12-zone problem instance, a 5.388% and 4.211% improvement with respect to the ETDR of the idle state are obtained for the optimal and ADP policies respectively. The evaluation of these problem instances demonstrate the scalability of our solution approach, which in turn encouraged the development and evaluation of the 34-zone case, a problem instance that that is too large to be solved to optimality. To accomplish this task, algorithmic modifications are conducted to solve the problem in a computationally tractable amount of time. With these modifications, we were able to obtain a 1.144% improvement over the myopic policy. Whereas this may not be of practical significance, it encourages the development of a experimental design to properly tune the algorithmic parameters and render better results.

A series of sensitivity analyses and computational excursions were also conducted to identify how the generated policies change with respect to different problem parameter settings. The results from these experiments indicate that, in a high OPSTEMPO environment where the rate of casualties and proportion of high threat level calls are high, the ADP and optimal policies are able to vastly outperform the myopic policy. With improved MEDEVAC assets that have faster flight speeds, the gap between the optimal and myopic policy decreases; however, significant improvements can still be made. The research presented herein is of particular interest to the military and civilian medical planners and dispatch authorities and should in turn influence the development of future tactics, techniques, and procedures.

6.2 Recommendations for Future Research

An immediate extension to this work includes the proper evaluation of the 34-zone problem instance and proper algorithmic parameter tuning for the modified algorithm. Further extensions include relaxing the zone tessellation scheme and allowing for casualties to occur at any location within a region of interest. This aspect incorporates another layer of realism into the problem, which in turn can render insightful results. Due to the size of the problem, a queue was not examined; however, moving forward this would be another interesting incorporation that should improve the analysis of the MEDEVAC dispatching problem. Also, the examination of different ADP strategies and parameter tuning approaches may render results that outperform the strategies presented in this research. Lastly, as it stands the MEDEVAC dispatching problem has primarily been modeled using an off-line model based approach. However, these approaches are only as good as the model in which they are based off of. Future research should consider the development of a model-free on-line algorithm, which is able to adapt as different problem features change.

Bibliography

Aitchison, John, & Aitken, Colin GG. 1976. Multivariate binary discrimination by the kernel method. *Biometrika*, **63**(3), 413–420.

Baker, Joanna R, Clayton, Edward R, & Taylor III, Bernard W. 1989. A non-linear multi-criteria programming approach for determining county emergency medical service ambulance allocations. *Journal of the Operational Research Society*, **40**(5), 423–432.

Bandara, Damitha, Mayorga, Maria E, & McLay, Laura A. 2012. Optimal dispatching strategies for emergency vehicles to increase patient survivability. *International Journal of Operational Research*, **15**(2), 195–214.

Bandara, Damitha, Mayorga, Maria E, & McLay, Laura A. 2014. Priority dispatching strategies for EMS systems. *Journal of the Operational Research Society*, **65**(4), 572–587.

Berlin, Geoffrey N, & Liebman, Jon C. 1974. Mathematical analysis of emergency ambulance location. *Socio-Economic Planning Sciences*, **8**(6), 323–328.

Berman, Oded. 1981. Repositioning of distinguishable urban service units on networks. *Computers & Operations Research*, **8**(2), 105–118.

Bianchi, Geoffrey, & Church, Richard L. 1988. A hybrid FLEET model for emergency medical service system design. *Social Science & Medicine*, **26**(1), 163–171.

Chaiken, Jan M, & Larson, Richard C. 1972. Methods for allocating urban emergency units: a survey. *Management Science*, **19**(4), 110–130.

Daskin, Mark S, & Stern, Edmund H. 1981. A hierarchical objective set covering model for emergency medical service vehicle deployment. *Transportation Science*, **15**(2), 137–152.

Department of the Army, United States. 2019. *FM 4-02.2 Medical Evacuation*.

Eastridge, Brian J, Mabry, Robert L, Seguin, Peter, Cantrell, Joyce, Tops, Terrill, Uribe, Paul, Mallett, Olga, Zubko, Tamara, Oetjen-Gerdes, Lynne, Rasmussen, Todd E, *et al.* 2012. Death on the battlefield (2001–2011): implications for the future of combat casualty care. *Journal of Trauma and Acute Care Surgery*, **73**(6), S431–S437.

Erkut, Erhan, Ingolfsson, Armann, & Erdoğan, Güneş. 2008. Ambulance location for maximum survival. *Naval Research Logistics*, **55**(1), 42–58.

Fish, Peter N. 2014. *Army medical officer's guide*. Stackpole Books.

Grannan, Benjamin C, Bastian, Nathaniel D, & McLay, Laura A. 2015. A maximum expected covering problem for locating and dispatching two classes of military medical evacuation air assets. *Optimization Letters*, **9**(8), 1511–1531.

Green, Linda, & Kolesar, Peter. 1984. A comparison of the multiple dispatch and M/M/c priority queueing models of police patrol. *Management Science*, **30**(6), 665–670.

Green, Linda V, & Kolesar, Peter J. 2004. Anniversary article: Improving emergency responsiveness with management science. *Management Science*, **50**(8), 1001–1014.

Hall, William K. 1972. The application of multifunction stochastic service systems in allocating ambulances to an urban area. *Operations Research*, **20**(3), 558–570.

Ignall, E, Carter, G, & Rider, K. 1982. An algorithm for the initial dispatch of fire companies. *Management Science*, **28**(4), 366–378.

Jarvis, James Patrick. 1975. *Optimization in stochastic service systems with distinguishable servers.* Ph.D. thesis, Massachusetts Institute of Technology.

Jenkins, Phillip R. 2017. *Using Markov decision processes with heterogeneous queueing systems to examine military MEDEVAC dispatching policies.* M.Phil. thesis, Air Force Institute of Technology.

Jenkins, Phillip R. 2019. *Strategic location and dispatch management of assets in a military medical evacuation enterprise.* Ph.D. thesis, Air Force Institute of Technology.

Jenkins, Phillip R, Robbins, Matthew J, & Lunday, Brian J. 2018. Examining military medical evacuation dispatching policies utilizing a Markov decision process model of a controlled queueing system. *Annals of Operations Research*, **271**(2), 641–678.

Jenkins, Phillip R, Lunday, Brian J, & Robbins, Matthew J. 2020a. Aerial MEDEVAC Operations. *Phalanx*, **53**(1), 63–66.

Jenkins, Phillip R., Lunday, Brian J., & Robbins, Matthew J. 2020b. Artificial Intelligence for Medical Evacuation in Great-Power Conflict. *War on the Rocks*.

Jenkins, Phillip R, Lunday, Brian J, & Robbins, Matthew J. 2020c. Robust, Multi-Objective Optimization for the Military Medical Evacuation Location-Allocation Problem. *Omega*, **97**(2020), 1–12.

Jenkins, Phillip R, Robbins, Matthew J, & Lunday, Brian J. 2021a. Approximate dynamic programming for military medical evacuation dispatching policies. *INFORMS Journal on Computing*, **33**(1), 2–26.

Jenkins, Phillip R, Robbins, Matthew J, & Lunday, Brian J. 2021b. Approximate Dynamic Programming for the Military Aeromedical Evacuation Dispatching, Preemption-Rerouting, and Redeployment Problem. *European Journal of Operations Research*, **290**(1), 132–143.

Jenkins, Phillip R, Robbins, Matthew J., & Lunday, Brian J. 2021c. Optimising aerial military medical evacuation dispatching decisions via operations research techniques. *BMJ Mil Health*, 1–3.

Keneally, Sean K, Robbins, Matthew J, & Lunday, Brian J. 2016. A Markov decision process model for the optimal dispatch of military medical evacuation assets. *Health Care Management Science*, **19**(2), 111–129.

Kolesar, Peter, & Walker, Warren E. 1974. An algorithm for the dynamic relocation of fire companies. *Operations Research*, **22**(2), 249–274.

Kulkarni, Vidyadhar G. 2017. *Modeling and Analysis of Stochastic Systems*. 3rd edn. Crc Press.

Maxwell, Matthew S, Restrepo, Mateo, Henderson, Shane G, & Topaloglu, Huseyin. 2010. Approximate dynamic programming for ambulance redeployment. *INFORMS Journal on Computing*, **22**(2), 266–281.

Mayorga, Maria E, Bandara, Damitha, & McLay, Laura A. 2013. Districting and dispatching policies for emergency medical service systems to improve patient survival. *IIE Transactions on Healthcare Systems Engineering*, **3**(1), 39–56.

McLay, Laura A, & Mayorga, Maria E. 2010. Evaluating emergency medical service performance measures. *Health Care Management Science*, **13**(2), 124–136.

McLay, Laura A, & Mayorga, Maria E. 2013a. A dispatching model for server-to-customer systems that balances efficiency and equity. *Manufacturing & Service Operations Management*, **15**(2), 205–220.

McLay, Laura A, & Mayorga, Maria E. 2013b. A model for optimally dispatching ambulances to emergency calls with classification errors in patient priorities. *IIE Transactions*, **45**(1), 1–24.

Nasrollahzadeh, Amir Ali, Khademi, Amin, & Mayorga, Maria E. 2018. Real-Time Ambulance Dispatching and Relocation. *Manufacturing & Service Operations Management*, **20**(3), 467–480.

Park, Seong Hyeon, & Lee, Young Hoon. 2019. Two-Tiered Ambulance Dispatch and Redeployment considering Patient Severity Classification Errors. *Journal of Healthcare Engineering*, **2019**.

Powell, Warren B. 2011. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. 2nd edn. Princeton, New Jersey: John Wiley & Sons.

Puterman, Martin L. 2005. *Markov decision processes: discrete stochastic dynamic programming*. Hoboken, New Jersey: John Wiley & Sons.

Rettke, Aaron J, Robbins, Matthew J, & Lunday, Brian J. 2016. Approximate dynamic programming for the dispatch of military medical evacuation assets. *European Journal of Operational Research*, **254**(3), 824–839.

Robbins, Matthew J, Jenkins, Phillip R, Bastian, Nathaniel D, & Lunday, Brian J. 2018. Approximate Dynamic Programming for the Aeromedical Evacuation Dispatching Problem: Value Function Approximation Utilizing Multiple Level Aggregation. *Omega*, **91**(2020), 1–17.

Schmid, Verena. 2012. Solving the dynamic ambulance relocation and dispatching problem using approximate dynamic programming. *European Journal of Operational Research*, **219**(3), 611–621.

Shackelford, Stacy A, Del Junco, Deborah J, Powell-Dunford, Nicole, Mazuchowski, Edward L, Howard, Jeffrey T, Kotwal, Russ S, Gurney, Jennifer, Butler, Frank K, Gross, Kirby, & Stockinger, Zsolt T. 2017. Association of prehospital blood product transfusion during medical evacuation of combat casualties in Afghanistan with acute and 30-day survival. *JAMA*, **318**(16), 1581–1591.

Smith, Iain M, James, Robert H, Dretzke, Janine, & Midwinter, Mark J. 2016. Prehospital blood product resuscitation for trauma: a systematic review. *Shock (Augusta, Ga.)*, **46**(1), 3.

Swersey, Arthur J. 1982. A Markovian decision model for deciding how many fire companies to dispatch. *Management Science*, **28**(4), 352–365.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704–0188

1. REPORT DATE *(DD–MM–YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From — To)*
25–03–2021	Master's Thesis	September 2019 — March 2021

4. TITLE AND SUBTITLE

The Impact of Threat Levels at the Casualty Collection Point on Military Medical Evacuation System Performance

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Dennie, Nathaniel C., Lt, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

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

Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT-ENS-MS-21-M-154

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

Joint Artificial Intelligence Center
Rebecca E. Lee, Product Manager, JAIC
122 S. Clark Street
Crystal City, VA 22202
rebecca.e.lee.20.civ@mail.mil

10. SPONSOR/MONITOR'S ACRONYM(S)

JAIC

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

12. DISTRIBUTION / AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A:
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

This work is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

14. ABSTRACT

One of the primary duties of the Military Health System is to provide effective and efficient medical evacuation (MEDEVAC) to injured battlefield personnel. To accomplish this, military medical planners seek to develop high-quality dispatching policies that dictate how deployed MEDEVAC assets are utilized throughout combat operations. This thesis seeks to determine dispatching policies that improve the performance of the MEDEVAC system. A discounted, infinite-horizon continuous-time Markov decision process (MDP) model is developed to examine the MEDEVAC dispatching problem. The model incorporates problem features that are not considered under the current dispatching policy (e.g., myopic policy), which tasks the closest-available MEDEVAC unit to service an incoming request. More specifically, the MDP model explicitly accounts for admission control, precedence level of calls, different asset types (e.g., Army versus Air Force helicopters), and threat level at casualty collection points. An approximate dynamic programming (ADP) algorithm is developed within an approximate policy iteration algorithmic framework that leverages kernel regression to approximate the state value function. The ADP algorithm is used to develop high-quality solutions for large scale problems that cannot be solved to optimality due to the *curse of dimensionality*. We develop a notional scenario based on combat operations in southern Afghanistan to investigate model performance, which is measured in terms of casualty survivability. The results indicate that significant improvement in MEDEVAC system performance can be obtained by utilizing either the MDP or ADP generated policies. These results inform the development and implementation of tactics, techniques and procedures for the military medical planning community.

15. SUBJECT TERMS

Markov decision processes, medical evacuation, admission control, priority dispatching, policy iteration, approximate dynamic programming

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON / 19b. TELEPHONE NUMBER *(include area code)*
U	U	U	UU	83	Dr. Phillip R. Jenkins, AFIT/ENS; (937) 255-3636, x4727; phillip.jenkins@afit.edu