Analyzing the Impact of Blood Transfusion Kits and Triage Misclassification Errors for Military Medical Evacuation Dispatching Policies via Approximate Dynamic Programming

THESIS

Channel A. Rodriguez, Capt, USAF

AFIT-ENS-MS-22-M-166

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENS-MS-22-M-166

ANALYZING THE IMPACT OF BLOOD TRANSFUSION KITS AND TRIAGE MISCLASSIFICATION ERRORS FOR MILITARY MEDICAL EVACUATION DISPATCHING POLICIES VIA APPROXIMATE DYNAMIC PROGRAMMING

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

Channel A. Rodriguez, BS

Capt, USAF

March 24, 2022

AFIT-ENS-MS-22-M-166

ANALYZING THE IMPACT OF BLOOD TRANSFUSION KITS AND TRIAGE MISCLASSIFICATION ERRORS FOR MILITARY MEDICAL EVACUATION DISPATCHING POLICIES VIA APPROXIMATE DYNAMIC PROGRAMMING

THESIS

Channel A. Rodriguez, BS
Capt, USAF

Committee Membership:

Maj Phillip R. Jenkins, PhD
Chair

Dr. Matthew J. Robbins
Member

AFIT-ENS-MS-22-M-166

Abstract

Members of the armed forces greatly rely on having an effective and efficient medical evacuation (MEDEVAC) process for evacuating casualties from the battlefield to medical treatment facilities (MTF) during combat operations. This thesis examines the MEDEVAC dispatching problem and seeks to determine an optimal policy for dispatching a MEDEVAC unit, if any, when a 9-line MEDEVAC request arrives, taking into account triage classification errors and the possibility of having blood transfusion kits on board select MEDEVAC units. A discounted, infinite-horizon continuous-time Markov decision process (MDP) model is formulated to examine such a problem and compare generated dispatching policies to the myopic policy of sending the closest available unit. We utilize an approximate dynamic programming (ADP) technique that leverages a random forest value function approximation within an approximate policy iteration algorithmic framework to develop high-quality policies for both a small-scale problem instance and a large-scale problem instance that cannot be solved to optimality. A representative planning scenario involving joint combat operations in South Korea is developed and utilized to investigate the differences between the various policies. Results from the analysis indicate that applying ADP techniques can improve current practices by as much as 29% with regard to a life-saving performance metric. This research is of particular interest to the military medical community and can inform the procedures of future military MEDEVAC operations.

This work is dedicated to all who have contributed to this great nation's defense. I hope this research helps, even if just a little, in our ability as a military to rescue the men and women willing to make the ultimate sacrifice for their country.

Acknowledgements

Foremost, I would like to extend my sincerest gratitude to Dr. Jenkins for his guidance throughout this research process and in helping me advance my academic capabilities.

I also want to acknowledge one of my professors and committee member, Dr. Matthew Robbins. His instruction provided me the knowledge to continue in this area of research.

Finally, I want to express my appreciation and gratitude for my wife and children for their understanding and patience over the last 18 months. Without your support, this journey would not have been possible.

Channel A. Rodriguez

Table of Contents

List of Figures

List of Tables

ANALYZING THE IMPACT OF BLOOD TRANSFUSION KITS AND TRIAGE MISCLASSIFICATION ERRORS FOR MILITARY MEDICAL EVACUATION DISPATCHING POLICIES VIA APPROXIMATE DYNAMIC PROGRAMMING

I. Introduction

1.1 Motivation and Background

Members of the armed forces greatly rely on having an effective and efficient emergency medical service (EMS) system for evacuating combat casualties from the battlefield to medical treatment facilities (MTFs). Unfortunately, the number of combat casualties that occur in today's wars is still high. For example, the United States military has seen more than 20,000 service members wounded in combat throughout operation Enduring Freedom and nearly 32,000 service members wounded during operation Iraqi Freedom (DCAS, 2021). To further exacerbate this problem, the number of dedicated military resources available to evacuate combat casualties is limited. As such, it is vital for senior military leaders and medical planners to carefully manage the use of dedicated evacuation assets to minimize the negative impacts resulting from casualty events (e.g., loss of limb or life).

The United States (U.S.) Army has two options for evacuating combat casualties: (1) casualty evacuation (CASEVAC) and (2) medical evacuation (MEDEVAC). Although both options utilize a variety of platforms for evacuating casualties (e.g., ground and aerial vehicles), MEDEVAC platforms come equipped with medical professionals on board to administer life-saving treatments to the casualty while en route to the MTF whereas CASEVAC platforms do not come equipped with medical professionals. Current Army Health System policy mandates the use of MEDEVAC assets over CASEVAC except for when MEDEVAC assets are overwhelmed or when the injury sustained is less severe (Department of the Army, 2019).

Both ground and aerial platforms are utilized for MEDEVAC, but the majority of MEDEVAC missions utilize helicopters. Helicopters are able to fly in direct paths to casualty collection points (CCPs), ultimately reducing the time between injury and surgical intervention, which is vital to increasing the probability of survival (Eastridge *et al.*, 2012). The U.S. Army utilizes the HH-60M Black Hawk, which is capable of providing medical support such as expeditious delivery of whole blood, biological, and medical supplies to meet critical requirements; rapid movement of medical personnel and accompanying equipment to meet the requirement for mass casualty situations, reinforcement, or emergency situations; and movement of patients between hospitals, aeromedical staging facilities, hospitals ships, casualty receiving and treatment ships, seaports, and railheads (Department of the Army, 2019). Black Hawks are able to transport groups of 11 fully-equipped soldiers at a time while cruising at a speed of 282 kilometers per hour (USAASC, 2021). Such capabilities, along with a launch time of less than 7 minutes, enable HH-60M Black Hawk helicopters to evacuate casualties to MTFs efficiently and provide adequate treatment simultaneously (Jenkins, 2017).

The U.S. Navy utilizes the MH-60R Seahawk helicopter to execute their sea control missions, to include MEDEVAC operations (Hernandez *et al.*, 2010). Although the Seahawk is primarily used for anti-submarine warfare, it is still equipped with a rescue hoist for lifting casualties to the helicopter and can carry up to 5 additional passengers. With the ability to carry external fuel tanks for extended range, the Seahawk is able to transport combat casualties to MTFs while maintaining a cruising speed of 234 kilometers per hour (U.S. Naval Academy, 2021).

While traditional medical facilities are stationary, the U.S. Navy has the capability to bring medical services wherever they deploy. The U.S. Naval Ship (USNS) MERCY (T-AH 19) and the USNS COMFORT (T-AH 20) are both U.S. Navy hospital ships that provide mobile, flexible, and responsive medical and surgical care (Department of the Navy, 2019). The T-AHs are designed to be an afloat MTF, capable of housing up to 1000 patients. Each of these ships have 12 fully-equipped operating rooms, digital radiological services, a medical laboratory, a pharmacy, an optometry lab, a CAT-scan, and two oxygen producing plants. A fully manned hospital ship consists of approximately 1300 personnel and, when operating at that level, it is comparable to a continental U.S. general hospital. Furthermore, a helicopter landing deck is available on each ship with sufficient space for military helicopters to land (Military Sealift Command, 2020).

The MEDEVAC system is complex and should be carefully designed and developed to maximize effectiveness and efficiency. Several important decisions include the location of helicopters and CCPs as well as the evacuation dispatching policy, which dictates which (if any) unit to task to service evacuation requests. Helicopters should be strategically positioned in a manner that facilitates maximum coverage of the CCPs but also minimizes the time to evacuate the casualties to an appropriate MTF. The dispatching policy is integral for handling the varying levels of MEDEVAC requests that enter a time-sensitive system and is the primary focus of this thesis. A complex policy may take longer to implement and/or cause confusion, therefore simple policies may be preferred, such as the currently practiced myopic policy, which entails dispatching the closest-available MEDEVAC unit to service an incoming request, regardless of the triage level (e.g., urgent, priority, and routine). Unfortunately, simple policies are typically suboptimal, especially in high-intensity scenarios (i.e., when the number of request entering the system in a short amount of time is high). Determining optimal and/or high quality dispatching policies for military MEDEVAC is commonly referred to as the MEDEVAC dispatching problem.

The MEDEVAC dispatching problem has been thoroughly researched over the past decade (e.g., Rettke *et al.* (2016); Robbins *et al.* (2020); Jenkins *et al.* (2020b); Jenkins *et al.* (2021a); Dennie (2021)), but most authors assume the reported triage level is accurate, which is not always true. Increased stress levels in a deployed environment can affect the decision-making process, including the ability to properly triage a casualty (Porcelli & Delgado, 2017). Graves *et al.* (2021) researched the MEDEVAC dispatching problem while accounting for errors in the true classification of a casualty and incorporating blood transfusion kits on board select MEDEVAC units. As such, this thesis builds on their work, taking into account the probability the true triage classification of a casualty may not be what was reported, as well as including blood transfusion kits. The inclusion of blood transfusion kits on MEDEVAC units allows for casualties to receive vital medical treatment prior to arriving at an MTF, thereby giving them a higher probability of survival. This research also accounts for admission control as in past works (e.g., Jenkins *et al.* (2018) and Robbins *et al.* (2020)). With admission control, the dispatching authority has the freedom to choose if an incoming request will be serviced or not based on the data provided from the MEDEVAC request. If the decision is to service the request, the dispatching authority determines which available asset to send, diverging from the common myopic policy of sending the closest available unit.

A discounted, infinite-horizon Markov decision process (MDP) model is formulated to determine how to dispatch MEDEVAC assets in response to service requests with an objective of maximizing the expected total discounted reward generated by the system. The system earns a reward when a MEDEVAC unit is dispatched to service a MEDEVAC request based on the response time and triage level. A notional scenario, based in South Korea, is developed to provide an appropriate context for comparing dispatching policies. Two problem instances (e.g., small and large) are created from the scenario with the smaller problem being solved to optimality utilizing exact dynamic programming methods. Due to the cardinality of the state-space, the policy for the larger scaled problem is determined by using approximate dynamic programming (ADP) techniques and compared against the myopic policy.

1.2 Organization of the Thesis

The remainder of this thesis is structured as follows. Chapter II reviews the literature pertaining to the MEDEVAC dispatching problem. Chapter III details the MEDEVAC dispatching problem as well as the MDP and ADP formulations developed to solve it. Chapter IV examines the efficacy of the MDP and ADP solution approach on a synthetically generated notional scenario based on high-intensity combat operations in South Korea. Chapter V summarizes key points from this thesis and provides areas for future research.

II. Literature Review

Research related to this thesis involves emergency medical services (EMS) in both the civilian and military communities. As such, this literature review has two components: (1) civilian EMS, and (2) military MEDEVAC.

2.1 Civilian EMS

EMS research consists of, but is not limited to, facility location (Baker *et al.*, 1989), dispatching policies for ambulances (Bandara *et al.*, 2014), and fleet size for EMS vehicles (Lim *et al.*, 2011). Solving these problems are complicated by stochastic elements that must be addressed (e.g., the location of a service request). Modeling techniques used in this field of research include Markov decision process (MDP), simulation, and mathematical programming (Jenkins *et al.*, 2020c).

The myopic dispatching policy in most EMS and MEDEVAC systems simply dispatches the closest available unit. This policy is easy to implement but does not always render optimal results (Kuisma *et al.*, 2004). When considering the various factors that determine the amount of time it takes to travel from one location to another (e.g., road congestion, traffic accidents, road maintenance), dispatching the closest-available unit may lead to sub-optimal results (Mukhopadhyay *et al.*, 2019). Mukhopadhyay *et al.* (2019) approach this problem by utilizing real-time streaming data to update their online incident model. The authors employ a Monte-Carlo Tree Search in a Semi-Markov decision process (SMDP) framework to make the problem computationally tractable and find high-quality solutions, which lead to significant improvements in responder dispatching policies when compared to a myopic approach.

Another vital detail that can be overlooked is the capacity of an emergency room. While minimizing the time it takes a patient to get from the point-of-injury to the hospital is important, getting them to surgical intervention is more critical. This process may be delayed when an ambulance arrives to a hospital with an injured patient only to wait because all emergency rooms are occupied, a problem known as ambulance offload delay (AOD) (Almehdawe *et al.*, 2013). This delay not only affects the patient but also the availability of that ambulance to service other EMS requests. Li *et al.* (2021) formulate a discrete time, infinite-horizon, discounted MDP model to determine when to send certain patients to out-of-region emergency departments, enduring a longer travel time not only to gain a shorter offload time but to avoid over burdening the emergency department. Their results suggest that both the EMS system and patients benefit from the policy determined by their model.

Triage classification is also an important detail to consider when making dispatching decisions. EMS dispatching decisions are made with an assumption that the risk or injury severity (i.e., triage level) of the patient is known by the dispatcher. The reported classification may be falsely reported, which may lead to a different dispatching decision. McLay & Mayorga (2013) looked at developing an optimal dispatching policy for ambulances knowing that the information relayed to the dispatcher is not always correct. They modeled this problem with an infinite-horizon, average reward MDP to maximize the long-run average customer utility over the true customer risk levels. Their results revealed that it is not always optimal to dispatch the closest-available ambulance even for patients classified as high risk when considering classification errors.

Performance measures evaluate the efficacy of a given system and as such need to be established in a manner that directly assesses how well a system achieves its goals. The objective of any EMS system is to ultimately save the lives of the patients they are treating. 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 (Dennie, 2021). Although RTTs are easy to obtain and understand, Bandara *et al.* (2012) propose using a performance measure that more closely relates to patient outcomes. The authors formulate a discounted, infinite-horizon MDP to generate dispatching policies based on the severity of the call and evaluate them in terms of the patient's survivability instead of response times. The results show an increase in the average survival probability of the patients for dispatching policies that take into account the severity of incoming calls. Furthermore, results indicate that more lives can be saved, while maintaining the same inventory of paramedic units, by implementing the optimal dispatching policy.

One of the decisions that dispatching authorities face is deciding which ambulance to dispatch to respond to an emergency call. Schmid (2012) examines this decision as well as ambulance relocation, which consists of determining where to relocate an ambulance after it has serviced a request. Although an ambulance will typically return to its home base, Schmid (2012) relaxes this constraint to improve the performance of the system and its capability to serve emergency requests. The author utilizes approximate dynamic programming (ADP) techniques on real-world data to compute high-quality solutions. By deviating from the norm of dispatching the closest-available ambulance and relocating back to home-base, the ADP model renders an improvement of 12.89% in the average response time over the myopic policy.

2.2 Military MEDEVAC

Bradley *et al.* (2017) note that military medical research has been conducted for over a century. MEDEVAC research in particular has been heavily looked into over the past decade (e.g., Malsby III *et al.* (2013), Keneally *et al.* (2016), Rettke *et al.* (2016), Jenkins *et al.* (2018), Jenkins (2019), Jenkins *et al.* (2020a)). Although the military community benefits from research conducted around civilian EMS response systems, there are differences unique to the armed forces that necessitate further analysis (Wooten, 2021). For instance, injuries sustained from improvised explosive devices (IEDs) have produced injury patterns never seen before (Bradley *et al.*, 2017). Other differences include determining MEDEVAC unit locations and MTF locations as well as accounting for when a casualty is in a high threat area, thereby necessitating an armed escort to accompany the MEDEVAC unit (Keneally *et al.*, 2016).

Early works of MEDEVAC research leveraging an MDP framework include those by Keneally *et al.* (2016) and Rettke *et al.* (2016). Keneally *et al.* (2016) solve their instance of the MEDEVAC dispatching problem by using a relative value iteration dynamic programming algorithm on combat scenario examples based in Afghanistan. The authors use the steady-state system utility as a performance metric to compare different dispatching policies. Their results indicate that an optimal policy yields a higher steady-state utility by 0.01 when compared to a myopic policy and 0.09 when compared to an intra-zone policy (i.e., MEDEVAC units are restricted to operate in specific zones).

Jenkins *et al.* (2018) contribute to this area of research by incorporating admission control and queuing, allowing the dispatching authority the flexibility to accept or reject an incoming MEDEVAC request based on the current state of the system (i.e., MEDEVAC unit availability and request status). This capability permits the dispatching authority to reserve a MEDEVAC unit for when a higher precedence request arrives. Utilizing an MDP framework, a discounted, infinite-horizon model is formulated to solve this problem. The results from this research illustrate that a myopic policy of sending the closest-available unit is not always optimal. Instead, examining the entire state of the system, with admission control, proves to be optimal when the flight speed of a helicopter is not at full potential due to various factors (e.g., atmospheric, environmental, or mechanical issues) or when intra-zone policies are enforced.

Jenkins *et al.* (2021b) further advance the work surrounding MEDEVAC problems by examining the MEDEVAC dispatching, preemption-rerouting, and redeployment (DPR) problem. The authors include fuel constraints within their model, which combined with the inclusion of DPR, yields a higher fidelity model and allows for improved decision making. Another distinguishing component of Jenkins *et al.* (2021b) from other MEDEVAC dispatching problems is the utilization of a support vector regression value function approximation scheme within an approximate policy iteration algorithmic framework. Results reveal that as the rate at which MEDEVAC requests enter the system increases, the performance gap between the ADP policy and the myopic policy (i.e., the currently practiced closest-available dispatching policy) increases substantially.

Sequential resource-allocation decision-making for MEDEVAC systems involves balancing when to dispatch a helicopter in response to a MEDEVAC request knowing that future requests are highly probable. This uncertainty complicates the decision of which action to take to maximize the reward of the system. As such, MDP-models are common in this area of research (e.g., Jenkins *et al.* (2021a), Robbins *et al.* (2020)). Jenkins *et al.* (2021a) state that although MDP models are well-suited for such problems, high dimensionality and uncountable state space render classical dynamic programming solution methods intractable. Instead, Jenkins *et al.* (2021a) resort to ADP solution methods to generate high-quality dispatching policies relative to the myopic dispatching policy. Utilizing an approximate policy iteration algorithmic framework, the authors compare two distinct ADP solution methods. The first algorithm uses least-squares temporal differences (LSTD) learning for policy evaluation, whereas the second algorithm uses neural network (NN) learning.

In like manner, Robbins *et al.* (2020) implement an approximate policy iteration algorithm to solve their MEDEVAC problem instance; however, their algorithm utilizes a multiple level aggregation scheme to approximate the post-decision state value function. Their modeling technique adopts a discrete zone tessellation scheme (e.g., 6-zone, 12-zone, and 34-zone) that conforms with the current practices in the U.S. military rather than a continuous-time model such as Jenkins *et al.* (2018) and Rettke *et al.* (2016). The results from the ADP model yields improvements as high as 12% over the myopic policy in the 34-zone case.

Wooten (2021) explores the scenario of adding a standby unit to the MEDEVAC dispatching problem. This provides the flexibility to relocate the standby unit as needed to staging areas with a higher influx of MEDEVAC requests, allowing for faster response times. Wooten (2021) formulates a discounted, infinite-horizon continuous-time MDP model and generates an optimal solution via policy iteration. She also applies an ADP technique that leverages a least squares policy evaluation value function approximation scheme within an approximate policy iteration algorithmic framework to solve a large problem instance representing an Iraq situation. Her findings indicate that the ADP-generated dispatching policies outperform the myopic policies in every case.

One of the objectives when considering optimal dispatching policies, in both civilian and military EMS, has been minimizing the time between critical injury and definitive care. As such, the Secretary of Defense in 2009 mandated that the United States MEDEVAC system respond to critically injured combat casualties in 60 minutes or less (Kotwal *et al.*, 2016). The study conducted by Kotwal *et al.* (2016) reported that after the mandate, the median transport time was reduced by over 50%.

In continued efforts to reduce the time from when a battlefield casualty receives medical treatment, military medical personnel have established standard operating procedures to administer blood transfusion by flight medics on-board MEDEVAC helicopters while en route to an MTF, a practice which has been in place since 2010 (Bradley *et al.*, 2017).

This research contributes to past military EMS research by further exploring the MEDEVAC dispatching problem. Building on the work done by Graves *et al.* (2021), this research examines similar problem features (i.e., admission control, triage misclassification errors, and on board blood transfusion kits). The previous author looked at a small scale problem instance and consequently was able to solve her MDP model to optimality. In contrast, this thesis continues those efforts by amplifying the state space to the point where ADP solution techniques are essential due to the cardinality of the state space. Specifically, the ADP framework used herein also utilizes a supervised machine learning technique (i.e., random forest) to generate a policy when the problem is computationally intractable. Furthermore, we explore a new scenario that combines the assets and capabilities of the U.S. Navy and U.S. Army, allowing for a more realistic representation of the joint environment in which the U.S. military operates.

III. Methodology

This chapter details the MEDEVAC process and the modeling techniques used to address the MEDEVAC dispatching problem examined herein.

3.1 Problem Description

Dedicated Army rotary-wing medical aircraft (i.e., air ambulances) fall under the mission command of the general support aviation battalion (GSAB). The GSAB is charged with positioning air ambulances (e.g., HH-60M Black Hawk helicopters) where they can best support timely and responsive evacuation (Department of the Army, 2019). The aviation commander, within the GSAB, considers the collective risk assessment of the mission and determines final execution or launch authority whenever a MEDEVAC request is submitted. For aerial MEDEVAC missions, the medical approval authority is accomplished by verifying the details of the MEDEVAC request with the policy contained in the medical rules of eligibility. If a MEDEVAC unit is available to dispatch, the dispatching authority must then decide which unit (if any) to dispatch. This sequential resource allocation decision-making of dispatching MEDEVAC units in response to a casualty event within the military aerial MEDEVAC system is known as the MEDEVAC dispatching problem (Robbins *et al.*, 2020).

When a service member sustains an injury in combat requiring the need for medical evacuation, a 9-line MEDEVAC request is submitted. This request contains, but is not limited to, the following information: location of the pickup site, triage classification of each injured individual, special equipment required, and security of pickup site (Wooten, 2021). It is the responsibility of the medical person at the scene to identify the evacuation triage level of each casualty and conclude if a MEDEVAC request is appropriate. If there is not a medical person at the scene, then this responsibility falls to the senior military member onsite (Jenkins, 2017). Table 1 provides a description of the three different triage levels (i.e., urgent, priority, and routine). Since the preferred method of evacuation for all casualty types is by air ambulance (Department of the Army, 2019), this thesis considers all triage levels.

Table 1. Triage Levels for Evacuation (Department of the Army, 2019)

Triage Category	**Description**
Priority I - Urgent	Is 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	Is 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	Is 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

Figure 1 outlines the MEDEVAC mission timeline. The time at which the dispatching authority receives a MEDEVAC request is denoted as T_1. If the MEDEVAC request is accepted, a MEDEVAC unit is assigned, and it begins necessary preparations (e.g., preparing medical equipment and personnel) to depart the staging area, which is indicated as T_2. Once the unit has completed its preparations, it begins traveling toward the designated pick-up site, known as the casualty collection point (CCP), which is denoted as T_3. This thesis assumes that CCPs are located in secure regions; therefore, MEDEVAC units do not require extra accommodations for servicing the request (e.g., armed-escorts, rescue hoists).

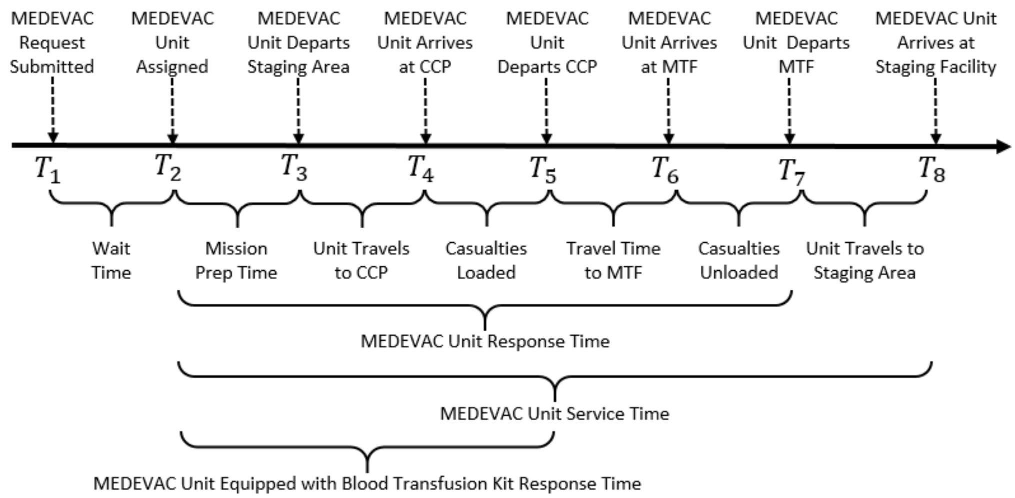

Figure 1. MEDEVAC Mission Timeline

T_4 indicates the time the MEDEVAC unit arrives at the CCP. The casualties are loaded, and the MEDEVAC unit proceeds to depart the CCP, denoted as T_5, and begins traveling toward a medical treatment facility (MTF), which is selected in a deterministic manner based on the location of the CCP. This thesis utilizes a monotonically decreasing reward function based in part on response time; therefore, the MEDEVAC unit travels to the closest MTF (in terms of response time). T_6 marks the MEDEVAC unit arriving at the MTF where the casualties are unloaded, and the responsibility of medical care is transferred to the MTF medical staff. Following the transfer of casualties, the unit returns to its original staging area for refueling and re-equipping and is ready to be re-tasked for another request, denoted by T_7 and T_8, respectively (Dennie, 2021).

The total response time for a MEDEVAC unit that is not equipped with a blood transfusion kit (BTK) is defined as $T_7 - T_2$, whereas the total response time for a unit with a BTK on board is defined as $T_5 - T_2$. Garrett (2013) reports that blood loss is the primary cause of death for soldiers killed in action, accounting for approximately 85% of them. A MEDEVAC unit equipped with a BTK is able to administer life-saving medical treatment the moment the casualty is loaded onto the helicopter, thereby having a faster response time. The total service time, regardless of the equipment on board, is defined as $T_8 - T_2$, which is also the total time the MEDEVAC unit's status is considered busy and not idle.

3.2 MDP Formulation

A discounted, infinite-horizon continuous-time MDP model is formulated to determine which MEDEVAC unit, if any, to dispatch in response to a given 9-line MEDEVAC request. The objective of the MDP model is to generate an optimal dispatching policy that maximizes the expected total discounted reward over an infinite horizon.

The U.S. Army utilizes a three-category casualty triage rubric (i.e., urgent, priority, and routine) when submitting a 9-line MEDEVAC request (Department of the Army, 2019). Prior works have excluded the routine triage level in their models (e.g., Rettke *et al.* (2016); Graves *et al.* (2021); Wooten (2021)). Although a routine triage level evacuation is assigned to minimally injured casualties and typically results in other forms of evacuation (i.e., CASEVAC), this thesis focuses on the possibility of triage misclassifications; therefore, all three triage levels are considered.

The 9-line MEDEVAC requests are assumed to arrive according to a Poisson process with parameter λ, denoted as $PP(\lambda)$. A splitting technique is utilized to model these arrivals, which are characterized by the location of the casualty event (i.e., zone), its reported triage level (i.e., urgent, priority, or routine), and the true triage level. Splitting is a technique used to generate two or more counting processes from a single Poisson process (Kulkarni, 2017). Let $\{N(t') : t' \geq 0\}$ denote the original counting process $PP(\lambda)$, which counts the number of 9-line requests that enter the MEDEVAC system during a given time interval $(0, t']$. The MEDEVAC requests are split into multiple processes categorized by the zone $z \in \mathcal{Z} = \{1, 2, \ldots, |\mathcal{Z}|\}$, the reported triage level $k \in \mathcal{K} = \{1, 2, \ldots, |\mathcal{K}|\}$, and the true triage level $c \in \mathcal{K}$. Let $\mathcal{R} = \{(z, k, c) : (z, k, c) \in \mathcal{Z} \times \mathcal{K} \times \mathcal{K}\}$ denote the set of request categories, for a total of $|\mathcal{R}| = |\mathcal{Z}||\mathcal{K}|^2$ possible request categories. The original counting process is split into $|\mathcal{R}|$ processes $\{N_{zkc}(t') : t' \geq 0\}, \forall (z, k, c) \in \mathcal{R}$, where each request belongs to one and only one of the categories. This generates the following result

$$N(t') = \sum_{(z,k,c) \in \mathcal{R}} N_{zkc}(t').$$

All requests are split using a Bernoulli splitting mechanism given parameters $p_{z,k,c} > 0, \forall (z, k, c) \in \mathcal{R}$ such that $\sum\limits_{(z,k,c) \in \mathcal{R}} p_{z,k,c} = 1$. The parameter $p_{z,k,c}$ indicates the probability of seeing a request in zone z with a reported triage level k and a true triage level c. The Bernoulli splitting mechanism yields counting processes $\{N_{zkc}(t') : t' \geq 0\}, \forall (z, k, c) \in \mathcal{R}$ where each is a Poisson process with parameter $\lambda p_{z,k,c}$, denoted as $PP(\lambda p_{z,k,c})$.

The decision epochs of this MDP model are the points in time that the aviation commander needs to make a decision. Decision epochs occur whenever a MEDEVAC request enters the system or a MEDEVAC unit completes servicing a request. The set of decision epochs is denoted as $\mathcal{T} = \{1, 2, \ldots\}$.

The state of the MEDEVAC system at decision epoch $t \in \mathcal{T}$ is described by the tuple $S_t = (M_t, R_t)$. The first component, M_t, represents the MEDEVAC status tuple and is defined as

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

where $\mathcal{M} = \{1, 2, \ldots, |\mathcal{M}|\}$ represents the set of MEDEVAC units in the system. The state variable $M_{tm} \in \{0\} \cup \mathcal{Z}$ provides the status of MEDEVAC unit $m \in \mathcal{M}$ at epoch t. When $M_{tm} = 0$, unit m is idle, meaning it is available to service a request and when $M_{tm} \in \mathcal{Z}$, unit m is busy servicing a zone $z \in \mathcal{Z}$ request.

The second component in the tuple S_t is R_t, which provides the details of a MEDEVAC request awaiting a decision from the aviation commander. More specifically, R_t provides the zone from which the request originates from, the reported triage level, and the true triage level and is denoted as

$$R_t = (Z_t, K_t, C_t)_{Z_t \in \mathcal{Z}, K_t \in \mathcal{K}, C_t \in \mathcal{K}}.$$

The components Z_t, K_t, and C_t correspond to the request zone, reported triage level, and true triage level respectively. $R_t = (0, 0, 0)$ indicates there is no request pending a decision in the system. When $R_t \neq (0, 0, 0)$, a pending request is in the system. Moreover, when $R_t \neq (0, 0, 0)$ and $C_t \neq K_t$, a classification error has occurred, and the dispatching authority has incorrect data about the triage level of the request.

The size of the state space can be calculated using the following equation

$$|\mathcal{S}| = (1 + |\mathcal{Z}|)^{|\mathcal{M}|}(1 + |\mathcal{Z}||\mathcal{K}|^2). \tag{1}$$

As more state variables are added (e.g., number of MEDEVAC units, zones, and triage levels), the size of the state space grows exponentially. If the size of the state space grows too large, then exact dynamic programming techniques become intractable, a phenomenon known as the *curse of dimensionality*. In such problem instances, approximate dynamic programming (ADP) techniques are required to gain insights, which are explored in this thesis.

When a MEDEVAC request is submitted, the dispatching authority has to decide whether or not they will admit the request (i.e., admission control). If all MEDEVAC units are busy servicing requests, then the only option is to reject the request (i.e., there is no queue). If there is at least one MEDEVAC unit available, then the dispatching authority needs to decide if they will reject the request due to a belief that a higher triage level request (e.g., urgent) will enter the system soon, or accept the request and choose which unit to dispatch. Whereas a myopic policy dispatches the closest available unit, an optimal policy may dispatch a different unit to reserve closer units for future requests.

Let $x_t^{reject} \in \{\Delta, 0, 1\}$ denote the admission control decision at epoch $t \in \mathcal{T}$, where $x_t^{reject} = 0$ denotes the decision to accept the request in the system. When $x_t^{reject} = 1$, the MEDEVAC request is rejected, either because all units are busy or to reserve the unit(s) for future requests. When there is no request in the system at epoch t (i.e., $R_t = (0, 0, 0)$), the system transitions without any impact from the admission control decision, denoted by $x_t^{reject} = \Delta$.

If the decision is made to accept a request (i.e., $x_t^{reject} = 0$), the next decision is to decide which unit to send. Let $\mathcal{I}(S_t) = \{m : m \in \mathcal{M}, M_{tm} = 0\}$ represent the set of idle MEDEVAC units at state S_t and let $x_t^d = (x_{tm}^d)_{m\in\mathcal{I}(S_t)}$ represent the dispatch decision variable tuple. If $x_{tm}^d = 1$, then MEDEVAC unit $m \in \mathcal{I}(S_t)$ is tasked and dispatched to service the request R_t at epoch t, otherwise $x_{tm}^d = 0$.

The decision variables at epoch t are denoted by the tuple $x_t = (x_t^{reject}, x_t^d)$. The dispatch decision variables are bounded by the following constraint

$$\sum_{m\in\mathcal{I}(S_t)} x_{tm}^d = I_{\{x_t^{reject}=0\}}. \tag{2}$$

The indicator function $I_{\{x_t^{reject}=0\}}$ takes the value of 1 when an incoming request is admitted into the system. This constraint ensures that a single MEDEVAC unit is tasked to service accepted MEDEVAC requests.

The action space for a given state, S_t, at epoch t subject to Constraint 2 is

$$\mathcal{X}_{S_t} = \begin{cases} (\Delta, \{0\}^{|\mathcal{I}(S_t)|}) & \text{if } R_t = (0,0,0) \\ (1, \{0\}^{|\mathcal{I}(S_t)|}) & \text{if } R_t \neq (0,0,0), \mathcal{I}(S_t) = \varnothing \\ (\{0,1\}, \{0,1\}^{|\mathcal{I}(S_t)|}) & \text{if } R_t \neq (0,0,0), \mathcal{I}(S_t) \neq \varnothing \end{cases} .$$

The first case accounts for when there is no service request in the system; therefore, the only action available is to transition with no changes. The second case indicates that there is a request in the system, but the set of idle MEDEVAC units is empty, meaning every unit is busy servicing other requests. The only action available in this case is to reject the request. The final case represents when the system has a MEDEVAC request and at least one MEDEVAC unit available. The available actions are to reject the request and not dispatch a unit or accept the request and dispatch one of the available units, subject to Constraint (2).

There are two events that cause the MEDEVAC system to transition: (1) a MEDEVAC request is submitted and (2) a MEDEVAC unit completes service.

Let $\mathcal{B}(S_t) = \{m : m \in \mathcal{M}, M_{tm} \neq 0\}$ denote the set of busy MEDEVAC units (i.e., MEDEVAC units servicing a request) when the system is in state S_t at epoch t, and let μ_{mz} denote the service rate of MEDEVAC unit $m \in \mathcal{M}$ servicing a request in zone $z \in \mathcal{Z}$. When the MEDEVAC system is in state S_t and takes action x_t at epoch t, the system immediately transitions to a post-decision state, denoted as S_t^x. The time the system remains in the post-decision state prior to transitioning to the next pre-decision state S_{t+1} (i.e., sojourn time) is exponentially distributed with parameter $\beta(S_t, x_t)$. The state-action sojourn time can be calculated as follows

$$\beta(S_t, x_t) = \lambda + \sum_{m \in \mathcal{B}(S_t)} \mu_{m,M_{tm}} + \sum_{m \in \mathcal{I}(S_t)} x_{tm}^d \mu_{m,Z_t}.$$

If $\mathcal{B}(S_t) = \varnothing$ and $x^d_{tm} = 0 \; \forall m \in \mathcal{M}$, indicating that all MEDEVAC units are idle and no units have been dispatched, then $\beta(S_t, x_t)$ represents the sojourn time for the state-action pairs wherein the next decision epoch occurs upon the arrival of a new MEDEVAC service request. Otherwise, $\mathcal{B}(S_t) \neq \varnothing$ and/or a unit is tasked to service an incoming request. In this case, $\beta(S_t, x_t)$ represents the sojourn time for the state-action pairs wherein the next decision epoch occurs after any event (i.e., a MEDEVAC request arrival or a MEDEVAC unit completing service). The probabilistic nature of the process can be summarized in terms of an infinitesimal $|\mathcal{S}| \times |\mathcal{S}|$ generator matrix as follows (Jenkins *et al.*, 2018)

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

wherein

$$p(S_{t+1}|S_t, x_t) = \begin{cases} \frac{\lambda_{zkc}}{\beta(S_t, x_t)}, & \text{if } R_{t+1} = (z, k, c), z \in \mathcal{Z}, k \in \mathcal{K}, c \in \mathcal{K} \\ \frac{\mu_{mz}}{\beta(S_t, x_t)}, & \text{if } R_{t+1} = (0, 0, 0), M_{m,t+1} = 0, M^x_{tm} = z, m \in \mathcal{M}, z \in \mathcal{Z} \\ 0, & \text{otherwise} \end{cases}$$

denotes the probability that the system transitions to state S_{t+1} by taking action x_t given that it is currently in state S_t. The post-decision state variable $M^x_{tm} \in \{0\} \cup \mathcal{Z}$ contains the information regarding MEDEVAC unit $m \in \mathcal{M}$ when decision x_t is made at epoch t. Note that $p(S^x_t|S_t, x_t) = 0$, indicating that the system will not occupy the same state, but rather will transition to a different state at the end of a sojourn in state S^x_t.

Leveraging the process of uniformization, the continuous-time MDP is transformed into a discrete-time MDP. Puterman (2005) notes that as long as the infinitesimal generator is not altered when applying uniformization, the discrete-time MDP will have the same probabilistic structure and allow for easier subsequent analysis. First, a maximum rate of transition ν must be determined, satisfying

$$\nu \geq \lambda + \sum_{m \in \mathcal{M}} \tau_m,$$

wherein

$$\tau_m = \max_{z \in \mathcal{Z}} \mu_{mz}, \ \forall m \in \mathcal{M}.$$

Through uniformization, the system state is observed more frequently than in the original system, allowing it to now have self-transitions. This transformation may be viewed as inducing extra (i.e., "fictitious") transitions from a state to itself. The discrete-time MDP yields the following transition probabilities:

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

The MEDEVAC system earns a reward when the dispatching authority decides to accept a 9-line MEDEVAC request and dispatches a MEDEVAC unit. The amount of reward that the system earns is captured in the contribution function $C(S_t, x_t)$ and varies based on the triage level of the request (i.e., k), the location (i.e., z), and the specific MEDEVAC unit tasked (i.e., m). The contribution function is defined as

$$C(S_t, x_t) = w_k u_k(r_{mz}), \tag{3}$$

where k is the triage level reported in the MEDEVAC request (i.e., $k = k_t$) and r_{mz} is the expected response time for the tasked MEDEVAC unit $m \in \mathcal{M}$ (i.e., $x^d_{tm} = 1$) to service a call originating in zone $z \in \mathcal{Z}$ (i.e., $z_t = z$). The k-utility function $u_k(r_{mz})$ is a monotonically decreasing function that renders a reward as a function of the response time r_{mz} and the reported triage level k. Therefore, for a fixed triage level k, servicing a MEDEVAC request with a faster expected response time yields a higher immediate reward. The weight parameter w_k scales the immediate reward between the different triage levels. Keeping all else equal, servicing an urgent request earns a higher reward than servicing a priority request, which earns a higher reward than servicing a routine request.

We employ the uniformization process once more to transform the continuous-time contribution function into an equivalent, discrete-time contribution function, denoted as

$$\tilde{C}(S_t, x_t) = C(S_t, x_t)\frac{\tilde{\gamma} + \beta(S_t, x_t)}{\tilde{\gamma} + \nu},$$

where $\tilde{\gamma} > 0$ represents the continuous-time discount rate. The discrete-time discount rate is determined by setting $\gamma = \frac{\nu}{\nu + \tilde{\gamma}}$.

The objective of the MDP formulation described is to maximize the expected total discount reward that the MEDEVAC system earns over an infinite horizon. Let $X^\pi(S_t)$ be a decision function that determines the action x_t the system takes in state $S_t \in \mathcal{S}$ according to policy π. Therefore, the objective is to determine the optimal policy, π^*, from the class of policies, $\pi \in \Pi$, to maximize the expected total discounted reward over an infinite horizon. The objective can be expressed as

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

The following optimality equation is used to solve for the optimal policy, π^*,

$$V(S_t) = \max_{x_t \in \mathcal{X}_{S_t}} (\tilde{C}(S_t, x_t) + \gamma \mathbb{E}[V(S_{t+1})|S_t, x_t]). \tag{4}$$

3.3 ADP Formulation

Although the previously described MDP model is appropriate for this problem, solving Equation (4) to generate an optimal policy becomes computationally intractable due to the large cardinality of the state space (i.e., $|\mathcal{S}|$). Instead of solving Equation 4 with exact dynamic programming methods, we employ approximation techniques to overcome the *curse of dimensionality* that MEDEVAC scenarios tend to possess. This thesis applies an ADP strategy that utilizes a random forest regression value function approximation scheme around the post-decision state variable within an approximate policy iteration (API) algorithmic framework to generate high-quality solutions to the MEDEVAC dispatching problem.

We adopt a post-decision state convention because of its two-fold computational improvement: (1) the reduction of the state space dimensionality and (2) the modification of the optimality equation (i.e., Equation (4)), which removes the expectation from within the maximum operator (Ruszczynski, 2010). The post-decision state can be denoted as

$$S_t^x = S^{M,x}(S_t, x_t),$$

which expresses the state of the MEDEVAC system immediately after the system was in pre-decision state S_t and action x_t was taken. From the post-decision state, the system transitions to the next pre-decision state upon a sample realization of exogenous information, expressed as

$$S_{t+1} = S^{M,W}(S_t^x, W_{t+1}). \tag{5}$$

W_{t+1} represents the exogenous information, which is either the details of a MEDEVAC request that has entered the system or a MEDEVAC unit that has finished servicing a request.

The new state convention requires a modification to the optimality equation. Let

$$\begin{aligned} V^x(S_t^x) &= \mathbb{E}[V(S_{t+1})|S_t, x_t] \\ &= \mathbb{E}[V(S_{t+1})|S_t^x] \end{aligned}$$

denote the value of being in post-decision state S_t^x. This new equation can now be incorporated into Equation (4) as follows:

$$V(S_t) = \max_{x_t \in \mathcal{X}_{S_t}} (\tilde{C}(S_t, x_t) + \gamma V^x(S_t^x)). \quad (6)$$

Note that the value of being in post-decision S_{t-1}^x can be written as

$$V^x(S_{t-1}^x) = \mathbb{E}[V(S_t)|S_{t-1}^x]. \quad (7)$$

Substituting Equation (6) into Equation (7) results in the optimality equation around the post-decision state variable

$$V^x(S_{t-1}^x) = \mathbb{E}\left[\max_{x_t \in \mathcal{X}_{S_t}} (\tilde{C}(S_t, x_t) + \gamma V^x(S_t^x))|S_{t-1}^x\right]. \quad (8)$$

The key difference between the post-decision state optimality equation (i.e., Equation (8)) and the pre-decision state optimality equation (i.e., Equation (4)) is the exchange of the maximization and expectation operators. Ruszczynski (2010) explains that this convention is statistically much easier to average the optimal value, which amounts to simple mean estimation, than to estimate the entire expected value function as a function of x_t. Therefore, this exchange of operators yields computational advantages over the pre-decision state optimality equation.

Although implementing a post-decision state convention delivers computational savings, the size and dimensionality of the state space still render Equation (8) intractable for larger-scale problems. We proceed with our value function approximation approach to generate approximate solutions to Equation (8). Let $\bar{V}^x(S_t^x)$ denote the approximate value of being in post-decision S_t^x, where

$$\bar{V}^x(S_{t-1}^x) = \mathbb{E}\left[\max_{x_t \in \mathcal{X}_{S_t}} (\tilde{C}(S_t, x_t) + \gamma \bar{V}^{x,n-1}(S_t^x))|S_{t-1}^x\right], \tag{9}$$

and $\bar{V}^{x,n-1}(S_t^x)$ is the approximate value of being in state S_t^x from the (n-1)th iteration from Algorithm (1), which is explained in detail later. Decisions are made using policy

$$\bar{X}^\pi(S_t) = \underset{x_t \in \mathcal{X}_{S_t}}{\operatorname{argmax}}(\tilde{C}(S_t, x_t) + \gamma \bar{V}^{x,n-1}(S_t^x)).$$

We obtain an estimate for $\bar{V}^x(S_t^x)$ by utilizing a random forest algorithm. Random forest is a supervised learning algorithm that builds an ensemble of decision trees and merges them to get a more accurate and stable prediction. We can predict the value of being in a post decision state using the following equation

$$\dot{\bar{V}}^x(S_t^x) = \frac{1}{F}\sum_{f=1}^{F} G_f(S_t^x), \tag{10}$$

where F is the number of decision trees in the forest. The function G_f denotes a single decision tree and can be further expressed as

$$G_f(S_t^x) = \sum_{\eta=1}^{H} c_\eta I(S_t^x \in Q_\eta),$$

where I is an indicator function taking a value of 1 if S_t^x is part of region Q_η (0 otherwise), c_η denotes the mean response of region Q_η, and H represents the total number of regions. A decision tree splits samples into regions by selecting a feature that maximizes the reduction in mean square error over all splitting candidates. The mean response for region Q_η is calculated as follows,

$$c_\eta = \frac{1}{|Q_\eta|} \sum_{j \in Q_\eta} \hat{v}_j^n I(S_{t,i}^x \in Q_\eta),$$

where $\hat{v}^n$ is a vector containing the estimated values of being in post-decision states that have been sampled in iteration n (see Algorithm 1).

Now we continue with the implementation of an API algorithmic strategy to attain high-quality MEDEVAC dispatching policies. API is based on the structure of exact policy iteration, wherein two sequences are alternated in repeating fashion, as shown in Algorithm 1. Policy evaluation (i.e., the inner loop) is the first sequence, which consists of evaluating a fixed policy via simulation and incrementally updating the approximate value function parameters based upon observed results. The second sequence in the API algorithm is policy improvement wherein the next iteration of the inner loop uses the updated approximate value function from the previous iteration.

The random forest API (RF-API) algorithm begins by initializing $\bar{V}^{x,0}$ to zero. After initializing, the algorithm enters the policy improvement loop (i.e., Step 3) wherein a pre-determined number of post-decision states (i.e., J) are sampled and evaluated via iteration (i.e., Step 4). The post-decision states are selected by means of Latin hypercube sampling (LHS), a sampling method that generates well-spaced, uniform random samples for Monte Carlo procedures (Wooten, 2021). A post-decision state is selected out of the sample in Step 5.

For the selected post-decision state, the set of feasible next pre-decision states is calculated using the state transition function, $S_t = S^{M,W}(S_{t-1}^x, W_t)$. For each pre-

Algorithm 1 Random Forest Approximate Policy Iteration (RF-API) Algorithm

1: Initialize $n = 0$
2: Initialize $\bar{V}^{x,0}(\cdot)$ to zero.
3: **for** $n = 1$ to N **do** (Policy Improvement Loop)
4:     **for** $j = 1$ to J **do** (Policy Evaluation Loop)
5:         Generate a random post-decision state, $S^x_{t-1,j}$.
6:         Determine the set of next pre-decision states $\bar{\mathcal{S}} \subseteq \mathcal{S}$ using Equation (5).
7:         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 (11) and record the estimated value $\hat{v}_{j,i}$ of being in post-decision state $S^x_{t-1,j}$, given the system transitions to pre-decision state $S_{t,i}$.
8:         Compute and record the estimated value $\hat{v}^n_j$ utilizing Equation (12).
9:     **end for**
10:     Compute $\dot{\bar{V}}^x$ using Equation (10).
11:     Update $\bar{V}^{x,n}$ using Equation (13).
12: **end for**
13: Return the approximate value function $\bar{V}^{x,N}(\cdot)$.

decision state $S_{t,i} \in \bar{\mathcal{S}}, i = 1, 2, \ldots, |\bar{\mathcal{S}}|$, we solve the following equation

$$\hat{v}_{j,i} = \max_{x_t \in \mathcal{X}_{S_t}} \left(\tilde{C}(S_{t,i}, x_{t,i}) + \gamma \bar{V}^{x,n-1}(S^x_{t,i}) \right), \tag{11}$$

which represents the estimated value of transitioning from pre-decision $S_{t,i}$ to post-decision $S^x_{t,i}$. Note that the approximate value function $\bar{V}^{x,n-1}$ is from the previous policy improvement iteration (i.e., $n-1$). This indicates that during the first iteration (i.e., $n = 1$), $\bar{V}^{x,1-1} = \bar{V}^{x,0} = 0$. Therefore, decisions are made based on the action that maximizes the contribution function during the first policy improvement loop, which is equivalent to acting myopicly.

Advancing with Step 8, the value of being in a post-decision state $S^x_{t-1,j}$ is computed using all $|\bar{\mathcal{S}}|$ pre-decision state values generated in Step 7 and leveraging the transition probability function $\tilde{p}$, which indicates the probability of transitioning from a post-decision state, $S_{t-1,j}$, to a pre-decision state, $S_{t,i}$, $\forall S_{t,i} \in \bar{\mathcal{S}}$. Combining this information, we compute $\hat{v}_j^n$ as follows,

$$\hat{v}_j^n = \sum_{i=1}^{|\bar{\mathcal{S}}|} \tilde{p}(S_{t,i}|S_{t-1,j}^x)\hat{v}_{j,i}. \tag{12}$$

After the conclusion of the policy evaluation loop, the policy improvement loop continues by constructing a random forest model (i.e., $\dot{\bar{V}}^x$) utilizing the sampled post-decision states $S_{t-1,j}^x, j = 1, 2, \ldots, J$, as the feature space and the estimated values $v_j^n, j = 1, 2, \ldots, J$, as the associated responses via Equation (10). We use a polynomial step-size rule to smooth in the new estimate of the value function approximation (i.e., $\dot{\bar{V}}^x$) with the previous estimate (i.e., $\bar{V}^{x,n-1}$) in Step 11. The step-size rule is expressed as

$$\alpha_n = \frac{1}{n^\kappa},$$

where $\kappa \in [0, 1]$ and n is the policy improvement loop counter. Using the step-size rule, the updated approximate value function is denoted as

$$\bar{V}^{x,n} = (1 - \alpha_n)\bar{V}^{x,n-1} + \alpha_n \dot{\bar{V}}^x. \tag{13}$$

This iterative process is repeated N times at which point the algorithm returns the final approximate value function in Step 14. The RF-API algorithm requires tuning parameters to achieve quality results, which is further explored in Chapter IV.

IV. Testing, Results, and Analysis

This chapter illustrates the applicability of the Markov decision process (MDP) model by examining a theoretical scenario of interest to the military medical community and evaluating the policies generated by approximate dynamic programming (ADP). The varying policies (i.e., myopic, optimal, and ADP) are compared when the problem instance is tractable to gain insight on the approximation of the ADP solution to the optimal policy. Furthermore, the representative scenario is scaled and expanded such that it can no longer be solved to optimality, but is approximated using the techniques described in Chapter III. All computational efforts were solved using an Intel Xeon Silver 4114 CPU with 64 GB of RAM, while leveraging MATLAB's Parallel Computing Toolbox.

4.1 Representative Scenario

The origin of the hostile tension between the Democratic People's Republic of Korea (DPRK) (i.e., North Korea) and the Republic of Korea (ROK) (i.e., South Korea) dates back to the Korean War, which began mid-way through the twentieth century. The strongest of the attacks from the DPRK troops during the three year conflict was aimed at capturing the city of Seoul in the South. This invasion from the North led to other countries getting involved, including the United States (U.S.). Although the conflict ended in an armistice and was never officially terminated, the death toll on both sides was great. The U.S. lost over 35,000 military members in combat with an additional 100,000 wounded (Millet, 2021).

Following the armistice, the demilitarized zone (DMZ) was established along the 38th parallel separating the DPRK from the ROK. The areas north and south of the DMZ are heavily fortified with both sides maintaining large contingents of troops

(Britannica, 2020). Over the years, U.S. and ROK armed forces have trained together in exercises, such as the Combined Command Post Training (Maxwell, 2020), to maintain necessary readiness for defending South Korea from an attack by the North Korean military.

The theoretical planning scenario being explored in this thesis considers joint combat operations in South Korea, with a heavy presence near the DMZ. The joint aspect leverages assets from both the U.S. Army and Navy (i.e, HH-60M Black Hawk and MH-60R Seahawk helicopters). This scenario assumes a MEDEVAC system with four demand zones (i.e., zones from which 9-line MEDECAC requests originate), four MEDEVAC unit staging areas (i.e., the locations in which the MEDEVAC units are stationed), two medical treatment facilities (MTFs), and two hospital ships. The location of coalition bases (i.e., larger military bases with space for both a helicopter landing zone (HLZ) and MTF) and forward operating bases (i.e., smaller bases that are only able to host a HLZ) are established at likely military tactical sites as shown in Figure 2. The MEDEVAC units utilizing Black Hawks (i.e., MEDEVAC unit -BH) are located in Zones 1 and 3 and the MEDEVAC unit utilizing a Seahawk (i.e., MEDEVAC unit -SH) is located in Zone 2. Two MEDEVAC units are co-located with an MTF (i.e., MEDEVAC units 1 and 2), thereby allowing them to be equipped with blood transfusion kits. The hospital ships are stationed on the eastern and western borders of the country, giving MEDEVAC units more options when determining where to evacuate a casualty. Both the MTFs and hospital ships are equipped with the necessary resources to treat any and all casualties, eliminating the need to transfer casualties due to capacity limits. Therefore, only the proximity of an MTF to a casualty collection point (CCP) is utilized to determine where a MEDEVAC unit will evacuate casualties.

Using past MEDEVAC research as a frame of reference (i.e., Jenkins *et al.* (2018)

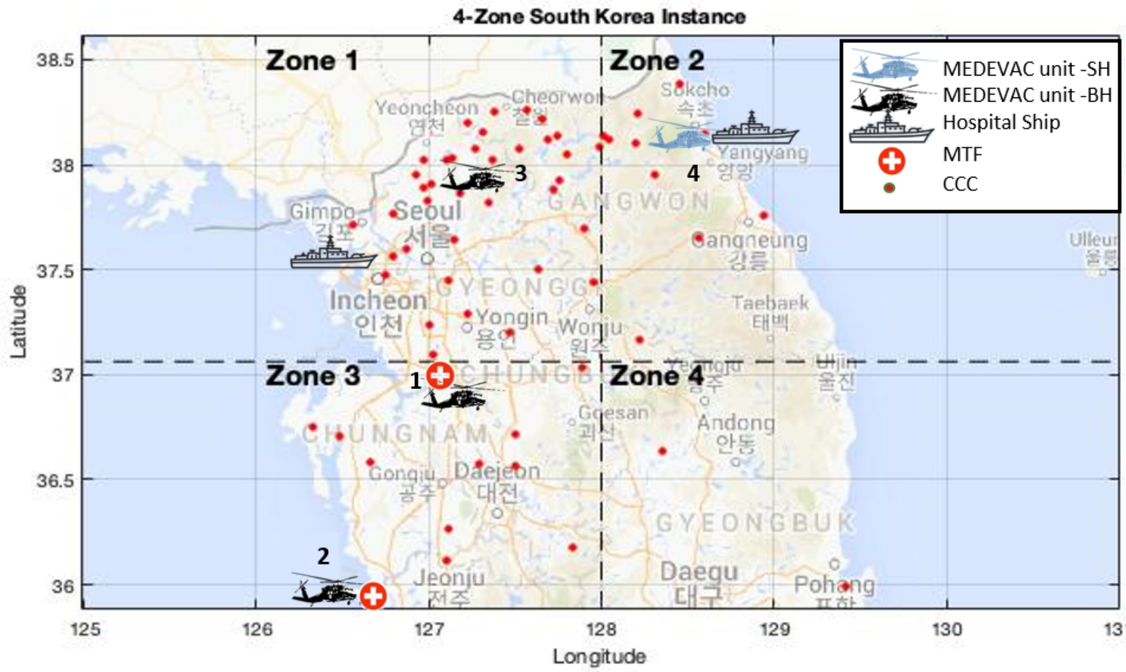

Figure 2. MEDEVAC locations, zones, and CCCs

and Wooten (2021)), future 9-line MEDEVAC requests are modeled with a Monte Carlo simulation via a Poisson cluster process. Casualty cluster centers (CCCs) are strategically selected as areas that could face a large number of casualties given an attack. This thesis utilizes the geographic coordinates of major and minor ROK military bases as CCCs. The distribution of 9-line MEDEVAC request locations from a given CCC is generated on a uniform distribution with respect to the distance of the request to the CCC. The quantity and location of the chosen CCCs directly affect the generated dispatching policy. Therefore, future use of this model must be modified for its intended scenario to develop meaningful results.

Table 2 displays the proportion of requests that originate from each zone. The capital, and most populated city, in South Korea (i.e., Seoul) accounts for the majority of the nation's population. Therefore, the zone Seoul belongs to in Figure 2 (i.e., Zone 1) has the highest proportion of MEDEVAC requests as we would expect a greater volume of casualties in that area.

Table 2. 9-Line MEDEVAC Requests Proportion by Zone

Zone	Proportion
1	0.6157
2	0.1672
3	0.1840
4	0.0331

A key component of this thesis is accounting for MEDEVAC requests being reported with an incorrect triage level. Whereas Graves *et al.* (2021) assumes service request triage levels can only be overestimated (i.e., casualty is less severe than reported), this thesis also assumes a service request can be underestimated (i.e., casualty is more severe than reported). Empirical studies done on hospital emergency department triage misclassifications have shown that patients are misclassified 9-15% of the time, depending on nurse experience (Saghafian *et al.*, 2014). Leveraging that data and combining it with this high operations tempo scenario justifies the use of higher misclassification rates. Table 3 outlines the triage misclassification rates used, where each triage level (i.e., urgent, priority, and routine) is misclassified at a rate of 40%, 30%, and 0.2% respectively.

Table 3. Misclassification Rates ϕ_{kc}

	Truth, c		
Reported, k	Urgent	Priority	Routine
Urgent	0.60	0.35	0.05
Priority	0.05	0.70	0.25
Routine	0.001	0.001	0.998

The arrival rate for MEDEVAC requests, λ, is estimated by military medical planners based on how often they expect a request to enter the system. The baseline MEDEVAC request arrival rate $\lambda = \frac{1}{30}$ indicates one request every 30 minutes on average according to a Poisson process. Table 4 outlines the proportion of 9-line MEDEVAC request arrivals that originate in zone $z \in \mathcal{Z}$ having a reported triage level $k \in \mathcal{K}$, and a true triage level $c \in \mathcal{K}$.

Table 4. MEDEVAC Request Proportions by Zone and Triage Level

	Zone, z			
Request, $R_t = (z, k, c)$	1	2	3	4
(z,1,1)	0.2501	0.0652	0.0717	0.0130
(z,1,2)	0.1401	0.0380	0.0419	0.0075
(z,1,3)	0.0200	0.0054	0.0060	0.0011
(z,2,1)	0.0077	0.0021	0.0023	0.0004
(z,2,2)	0.1078	0.0293	0.0322	0.0058
(z,2,3)	0.0385	0.0104	0.0115	0.0021
(z,3,1)	0.0000	0.0000	0.0000	0.0000
(z,3,2)	0.0000	0.0000	0.0000	0.0000
(z,3,3)	0.0612	0.0167	0.0184	0.0033
Total	0.6157	0.1672	0.1840	0.0331

As depicted in Figure 1 in Chapter III, the accounted time for a MEDEVAC unit equipped with a blood transfusion kit (BTK) to respond to a MEDEVAC request includes the mission preparation time, travel time to the CCP, and time to load the casualty onto the helicopter. Incorporating parameter settings from Bastian (2010) and Wooten (2021), the mission preparation time is set to 15 minutes, load time is set to 10 minutes, and unload time is set to five minutes. When the MEDEVAC unit is not equipped with the BTK, the accounted response time also includes time to travel to the MTF and time for unloading the casualty at the MTF in addition to the previously mentioned events. The Monte Carlo simulation described earlier generates casualties based on the CCCs and calculates the response time for each MEDEVAC unit. These data points are averaged to get expected response times for each MEDEVAC unit $m \in \mathcal{M}$ to respond to a MEDEVAC request originating in zone $z \in \mathcal{Z}$ as shown in Table 5.

Table 5. Expected Response Times (minutes)

	Zone, z			
MEDEVAC unit, m	1	2	3	4
1	46.5000	58.9820	37.9980	62.2010
2	71.1390	82.0260	45.3790	69.1250
3	55.1840	65.1730	73.7650	116.2200
4	75.6830	53.2230	89.3210	116.9500

From the same simulation, the expected service times are generated in like manner, as presented in Table 6. For MEDEVAC units equipped with a BTK, the total service time is comprised of the response time, travel time to the MTF, time to unload the casualty at the MTF, and time to travel back to the staging area. The total service time for MEDEVAC units not equipped with a BTK consist of the response time and the time to travel back to the staging area.

Table 6. Expected Service Times (minutes)

	Zone, z			
MEDEVAC unit, m	1	2	3	4
1	87.2100	112.1500	60.4730	116.8100
2	133.4500	158.1900	81.9150	148.2800
3	76.7550	94.6460	101.2100	140.7200
4	112.2000	57.2380	144.1100	150.4000

Whenever a MEDEVAC unit is dispatched to service a request, a reward is earned based on the distance traveled (i.e, response time of dispatching MEDEVAC unit m) and the request priority (i.e., triage level k). The weights for the different triage levels from Equation (3), w_1, w_2, and w_3, are set to 0.9009, 0.0901, and 0.009, respectively, so that urgent requests are prioritized first, followed by priority and routine requests. The expected immediate rewards are displayed in Table 7.

Table 7. Expected Immediate Rewards

		MEDEVAC, m			
Zone, z	Priority, k	1	2	3	4
1	1(Urgent)	0.8710	0.7328	0.8393	0.6900
	2(Priority)	0.0796	0.0741	0.0777	0.0730
	3(Routine)	0.0088	0.0087	0.0088	0.0087
2	1(Urgent)	0.8200	0.6213	0.7806	0.8479
	2(Priority)	0.0768	0.0716	0.0754	0.0781
	3(Routine)	0.0088	0.0087	0.0088	0.0088
3	1(Urgent)	0.8886	0.8740	0.7088	0.5315
	2(Priority)	0.0815	0.0799	0.0735	0.0700
	3(Routine)	0.0089	0.0088	0.0087	0.0087
4	1(Urgent)	0.8008	0.7501	0.1747	0.1664
	2(Priority)	0.0761	0.0745	0.0639	0.0638
	3(Routine)	0.0088	0.0087	0.0086	0.0086

4.2 Representative Scenario Results

The current practice (i.e., myopic policy) of the MEDEVAC system resorts to sending the nearest available MEDEVAC unit in response to a 9-line MEDEVAC request. This policy does not include admission control, therefore, it does not consider the impact of sending a unit now has on future decisions nor does it consider key information before making the decision (e.g., triage level of the casualty or the number of units available). Although the myopic policy does not consider the triage level of the casualty, that information is what determines the reward for the action. When policies are generated, they are evaluated based on the true triage level of the casualty, not what was reported.

The initial state of the system $S_0 = (M_0, R_0) = ((0,0,0,0),(0,0,0))$ indicates that all MEDEVAC units are idle and available and there is no MEDEVAC request in the system. The value associated with being in the initial state is used when making comparisons between policies. The parameter settings used in this scenario are displayed in Table 8. The 4-zone problem instance is solved twice to optimality via policy iteration. The first solution formulates a policy based on the reported casualty triage level (i.e., $\text{Optimal}_{reported}$), whereas the second solution generates the policy based on the true triage level (i.e., Optimal_{truth}) and then both are evaluated on the true triage level. Both of these policies are compared to the myopic policy and the results are shown in Table 9.

Table 8. 4-Zone problem instance 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	4
$\lvert\mathcal{K}\rvert$	# of triage levels	3
γ	Continuous time discount rate	0.001
w_1	Weight for urgent requests	0.9009
w_2	Weight for priority requests	0.0901
w_3	Weight for routine requests	0.009

Table 9. 4 Zone Policy Comparison

Policy, π	$V^{\pi}(S_0)$	% Improvement over Myopic	% Optimal_{truth}
Optimal_{truth}	2.91	13.30%	-
$\text{Optimal}_{reported}$	2.76	7.51%	94.92%
Myopic	2.57	-	88.38%

The RF-API algorithm has several parameters that require tuning in order for it to generate a high-quality approximate solution. The parameters for this algorithm include the number of policy improvement loops (i.e., N), the number of policy evaluation loops (i.e., J), the power parameter within the polynomial step-size function (i.e., κ), and the number of trees in the random forest (i.e., F). Four factor levels were chosen for each of these parameters based on the results from preliminary experiments. More specifically, a 4^4 full factorial experimental design is constructed and evaluated. Table 10 shows the different factor levels for the parameters and Table 11 displays the top 20 results from the experiment.

Table 10. Experimental Design Factor Levels

Algorithm Parameters	Description	Levels
N	Policy Improvements	{10, 15, 20, 25}
J	Policy Evaluations	{100, 200, 300, 400}
κ	Step-size	{0.01, 0.1, 0.2, 0.5 }
F	Trees in Random Forest	{75, 100, 125, 150 }

Using the best parameter settings found in the experiment (i.e., Run 1 in Table 11), we generate two polices for the MEDEVAC dispatching problem. The value of being in the starting state and following policy ADP_{truth} is shown in Table 12 along with the value of following policy $\text{ADP}_{reported}$. Again, both policies are evaluated off the truth data, but policy ADP_{truth} is constructed assuming truth knowledge whereas policy $\text{ADP}_{reported}$ is constructed using the reported information.

The RF-API algorithm generates a policy using the parameters listed above in 14.67 seconds with the computing capabilities stated at the start of the chapter. The $\text{ADP}_{reported}$ policy is 93.87% optimal, when compared to the Optimal_{truth} policy. Although this policy can be improved, it still outperforms the myopic policy by 6.37%. The difference between the *truth* and *reported* policies for both optimal and ADP is the cost of misclassification. For the given misclassification rates defined in Ta-

Table 11. RF-API Computational Experiment Results

Run	N	J	κ	F	V(1)	%-Optimal$_{truth}$
1	20	300	0.2	125	2.7295	93.87%
2	10	400	0.01	75	2.7290	93.85%
3	20	400	0.1	100	2.7277	93.81%
4	20	400	0.2	100	2.7277	93.81%
5	25	400	0.1	100	2.7275	93.80%
6	15	400	0.1	125	2.7272	93.79%
7	20	400	0.1	75	2.7271	93.79%
8	25	400	0.1	75	2.7267	93.77%
9	10	300	0.2	100	2.7267	93.77%
10	10	400	0.01	125	2.7266	93.77%
11	20	400	0.5	75	2.7262	93.75%
12	15	400	0.01	100	2.7261	93.75%
13	15	300	0.1	100	2.7260	93.75%
14	15	400	0.01	125	2.7260	93.75%
15	15	300	0.2	125	2.7257	93.74%
16	15	400	0.5	75	2.7256	93.73%
17	25	400	0.5	75	2.7255	93.73%
18	20	300	0.2	100	2.7253	93.72%
19	15	400	0.01	75	2.7253	93.72%
20	10	400	0.01	100	2.7250	93.71%

Table 12. 4 Zone Policy Comparison Pt. 2

Policy, π	$V^{\pi}(S_0)$	% Improvement over Myopic	% Optimal$_{truth}$
Optimal$_{truth}$	2.91	13.30%	-
Optimal$_{reported}$	2.76	7.51%	94.92%
ADP$_{truth}$	2.86	11.35%	98.36%
ADP$_{reported}$	2.73	6.37%	93.89%
Myopic	2.57	-	88.38%

ble 3, the results in Table 12 indicate that the cost of misclassification between the Optimal$_{truth}$ and Optimal$_{reported}$ is approximately 5% and the cost of misclassification between the Optimal$_{truth}$ and ADP$_{reported}$ is approximately 6%. These results highlight the importance for proper triage categorization at the point-of-injury prior to submitting a 9-line MEDEVAC request.

4.2.1 Policy Comparison

Four scenarios are explored in further detail to highlight the differences and similarities between the three policies. Table 13 indicates the state of the system for each scenario along with the corresponding action from each policy. The first scenario has

Table 13. Policy Differences

Scenario	$S_t = (M_t, R_t)$	$X^{\pi^{\text{Myopic}}}(S_t)$	$X^{\pi^{\text{Optimal}_{reported}}}(S_t)$	$X^{\pi^{\text{ADP}_{reported}}}(S_t)$
1	((0,0,0,0), (1,1,1))	Dispatch MEDEVAC 1	Dispatch MEDEVAC 3	Dispatch MEDEVAC 3
2	((1,0,4,0), (3,2,3))	Dispatch MEDEVAC 2	Reject request	Dispatch MEDEVAC 4
3	((0,4,1,0), (2,1,1))	Dispatch MEDEVAC 4	Dispatch MEDEVAC 4	Dispatch MEDEVAC 4
4	((z_{t1},z_{t2},0,z_{t3}), (4,k_t, c_t))	Dispatch MEDEVAC 3	Reject request	Reject request

all MEDEVAC units available for service and a request originating from Zone 1 with a reported and true triage level of 1 (i.e., urgent). The myopic policy seeks to maximize the immediate reward, which in this case would be to dispatch MEDEVAC unit 1 according to Table 7. Both the Optimal$_{reported}$ and ADP$_{reported}$ policies differ from the myopic by choosing to dispatch MEDEVAC unit 3. Although the immediate reward for sending MEDEVAC unit 3 is 4% less than sending unit 1, this action allows the system to reserve MEDEVAC unit 1 for later use.

The second scenario has MEDEVAC units 2 and 4 idle with MEDEVAC unit 1 servicing a request in Zone 1 and MEDEVAC unit 3 servicing a request in Zone 4. The current request in the system includes a casualty with a reported triage level of 2 (i.e., priority), but a true triage level of 3 (i.e., routine) in Zone 3. Because there are units available to dispatch, the myopic policy dispatches the nearest available unit which is MEDEVAC unit 2. The reward for a servicing a priority triage level is significantly less than an urgent request, as such, the action for the Optimal$_{reported}$ policy is to reject the request despite having two available units. Interestingly, the ADP$_{reported}$ policy differs from both the myopic and Optimal$_{reported}$ policies by tasking MEDEVAC unit 4 to service the request. This indicates that although the ADP$_{reported}$

does not perform optimally, it still has enough information on the system to make the decision to reserve MEDEVAC unit 2 for a more severe request that is closer to its staging facility.

The state of the system in scenario 3 consists of MEDEVAC units 1 and 4 being idle, MEDEVAC unit 2 servicing a request in Zone 4, and MEDEVAC unit 3 servicing a request in Zone 1. The request in the system stems from Zone 2 and has a reported and true triage level of 1 (i.e., urgent). Given that MEDEVAC unit 4 is located in Zone 2, all three policies task unit 4 to service the request. This results reveals the importance each policy attributes to urgent-level requests.

The final scenario has a single unit available to dispatch (i.e., MEDEVAC unit 3) with a MEDEVAC request coming from Zone 4. Both the $\text{Optimal}_{reported}$ and $\text{ADP}_{reported}$ policies take the same action in this scenario, regardless of the MEDEVAC request triage level or the location of the other MEDEVAC units, as long as they are not idle. Both the $\text{Optimal}_{reported}$ and $\text{ADP}_{reported}$ policies choose to reject the request and save the last MEDEVAC unit for a request that comes from one of the other zones, whereas the myopic policy tasks the last available unit to service the request. This can be explained by the fact that Zone 4 is the only zone without a MEDEVAC unit within its boundaries, therefore units have to travel further distances to service requests. Based on the baseline parameter settings, the $\text{Optimal}_{reported}$ and $\text{ADP}_{reported}$ policies recognize that it is more beneficial to let requests be serviced by outside agencies (i.e., CASEVAC) when the system is in this particular state rather than tasking its last remaining MEDEVAC unit.

One of the advantages of leveraging a random forest technique is the ability to cross-validate throughout the training process. The out of bag (OOB) mean square error (MSE) and the training MSE are shown in Figure 3. As illustrated in Figure 3, the final random forest model that is utilized to generate the $\text{ADP}_{reported}$ policy is adequate and is neither over-fitted nor under-fitted given the small values in the y-axis (i.e., less than 0.2 MSE) and the proximity of the training MSE to the OOB MSE.

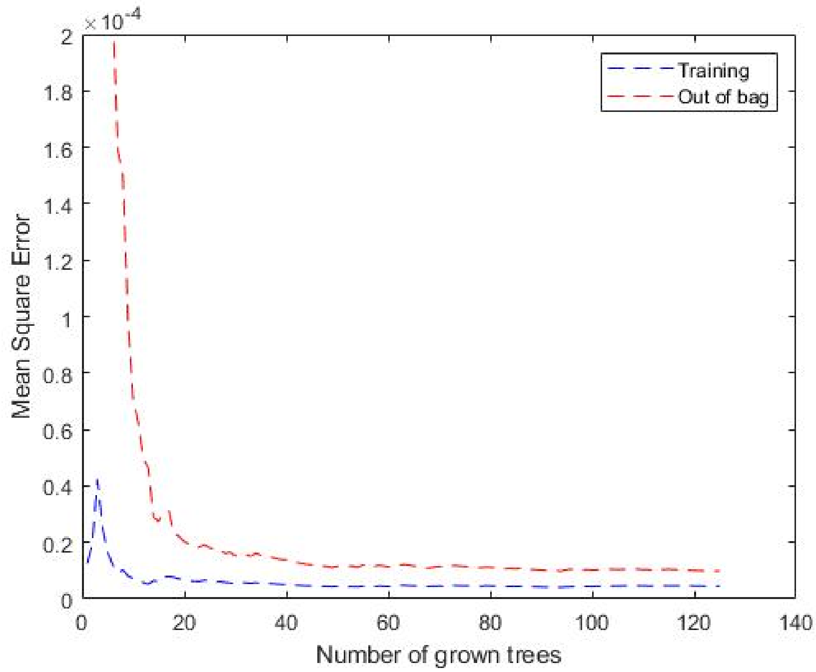

Figure 3. Out of Bag Error

As mentioned in Chapter III, the random forest algorithm takes in as inputs the post-decision status of the MEDEVAC system (i.e., S_t^x). Recall that the status of the MEDEVAC system is comprised of the MEDEVAC status tuple (i.e., M_t^x) and the request status tuple (i.e., R_t^x). The request status tuple in a post-decision state is always empty (i.e., $R_t^x = (0, 0, 0)$). Therefore, the model generated is ultimately based on the status of the MEDEVAC units (i.e., M_t^x). Figure 4 displays the importance of the features (i.e., the MEDEVAC units) within the model. MEDEVAC unit 1 has the highest feature importance value. The map displayed in Figure 2 shows that MEDEVAC unit 1 is located in Zone 3 just south of the border between Zones 1 and 3. Furthermore, that unit is also equipped with a BTK, reducing the total amount of time it takes for that unit to respond to a request as it is able to begin life saving

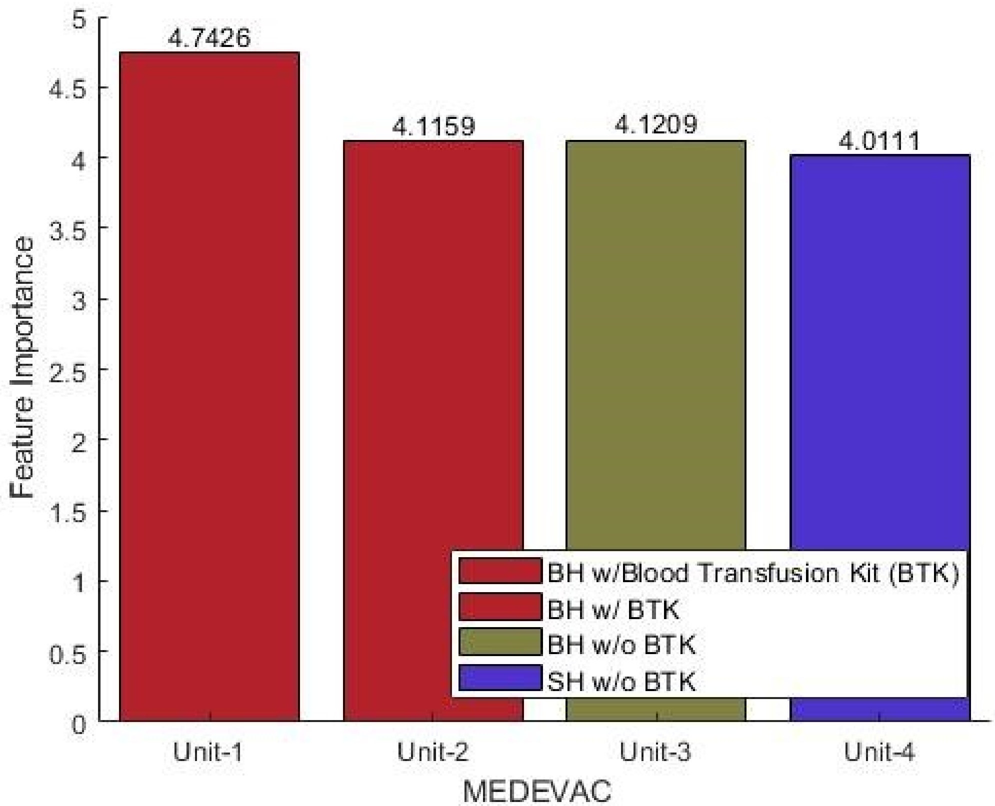

Figure 4. Feature Importance Plot

treatment upon arrival. Recall that approximately 62% of all requests originate from Zone 1 as noted in Table 2. Although MEDEVAC unit 3 is located in Zone 1, it is not equipped with a BTK. Its longer response time is the cause for the lower feature importance score. We see that having a BTK on-board does have an impact on the model generated. Despite MEDEVAC 2 being located in the southwest region of Zone 3, it was considered almost as important as MEDEVAC unit 1 due to the BTK.

4.3 Excursion - Arrival Rate

We further examine the effect that the MEDEVAC request arrival rate (i.e., λ) has on all three policies. Various values of λ are tested and the resulting policies are compared. All problem instance parameter settings defined in Table 8 remain the same with the exception of λ. Moreover, the best-tuned ADP algorithmic settings from the baseline scenario (i.e., Run 1 in Table 10) are utilized. The value of the starting state (i.e., $S_t = ((0,0,0,0),(0,0,0))$) for each request arrival rate is displayed in Table 14. As the arrival rate speeds up, the gap between the ADP$_{reported}$ policy and

Table 14. Arrival Rate Impact on ADP

	$V^{\pi}(S_0)$			ADP$_{reported}$ Performance	
$\frac{1}{\lambda}$	Optimal$_{truth}$	ADP$_{reported}$	Myopic	% Improvement over Myopic	% Optimal$_{truth}$
10	13.11	9.72	7.50	29.68%	74.13%
20	5.18	4.46	3.96	12.56%	86.18%
30	2.91	2.73	2.57	6.37%	93.89%
40	1.92	1.85	1.81	2.48%	96.19%
50	1.39	1.35	1.34	0.88%	97.42%
60	1.06	1.04	1.03	0.60%	98.42%

the myopic policy increases. With more requests entering the system, sub-optimal actions (i.e., dispatching the nearest available MEDEVAC unit) get compounded and produce inferior results. Moreover, the gap between the Optimal$_{truth}$ and ADP$_{reported}$

policies increases as arrival rate increases. This indicates that changing the arrival rate requires another experimental design to re-tune the parameters to achieve better results with regard to the optimality gap. As the arrival rate slows down, the difference between the Optimal_{truth}, $\text{ADP}_{reported}$, and myopic policies becomes negligible. For example, when the arrival rate is $\lambda = \frac{1}{60}$, implementing an $\text{ADP}_{reported}$ policy would only improve upon the myopic policy by 0.60% and it would be a more complex policy. More notable improvements with the $\text{ADP}_{reported}$ and the Optimal_{truth} policies exist as the MEDEVAC request arrival rate increases.

The change in the MEDEVAC request arrival rate also impacts which features are important to the algorithm. Figure 5 illustrates the importance of each feature (i.e., MEDEVAC unit status) based on the MEDEVAC request arrival rate. In every

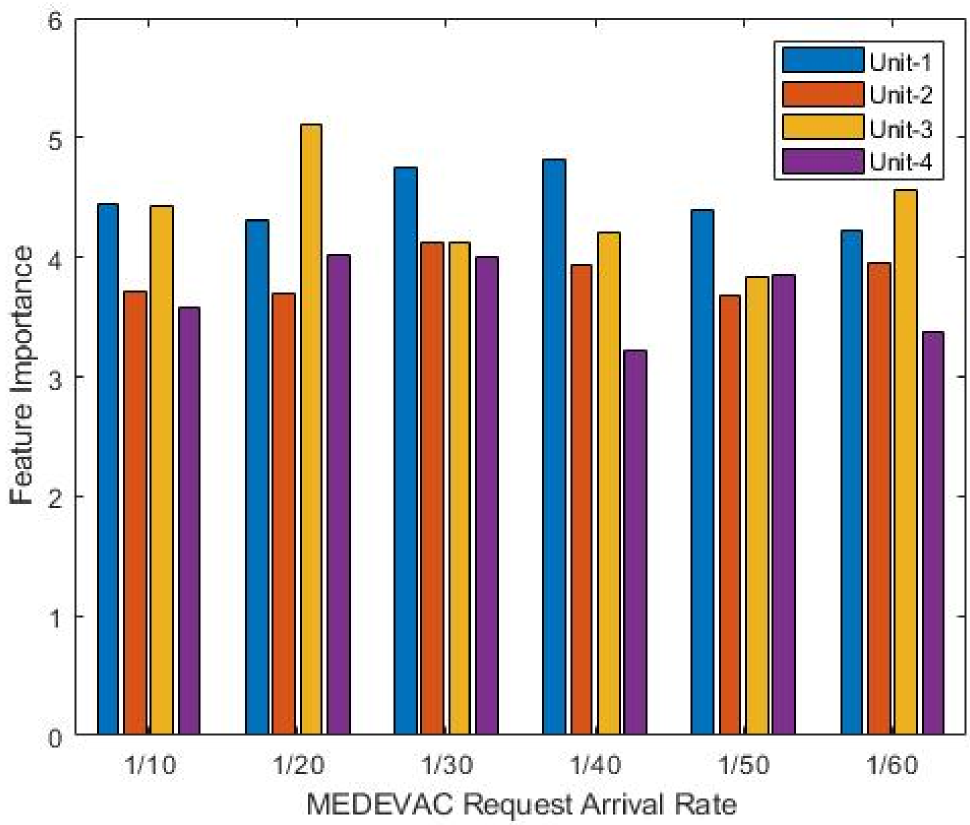

Figure 5. Feature Importance Plot for all λ

case, the top features were either the status of MEDEVAC unit-1 or unit-3. These are the units closest to Zone 1, which is the zone that has the most requests. From this insight, we see that our algorithm uses the location of the MEDEVAC units to develop a high-quality policy.

4.4 Excursion - V22 Osprey

As mentioned in Chapter I, the Navy currently uses the MH-60R Seahawk helicopter to conduct their MEDEVAC operations. In 2021, the Navy successfully conducted the inaugural landing of a V-22 Osprey aircraft on the deck of the U.S. Naval Ship MERCY (Correll, 2021). This accomplishment allows for the possibility of using the V-22 Osprey to evacuate patients to the hospital ships. The V-22 Osprey combines the advantages of airplanes (i.e., faster speeds and increased payload capacity) with the hovering ability of helicopters. With a cruising speed of 493 kilometers per hour, the Osprey outperforms not only the Seahawk, but the Black Hawk as well (Freudenrich, 2001). This aircraft can carry up to 24 passengers and is also equipped with rescue hoists making it an ideal platform for medical rescue operations.

We now explore the scenario where the V-22 Osprey is used in our baseline problem, replacing the Naval aircraft. Figure 6 displays the map of the problem scenario with a V-22 Osprey in Zone 2. The expected response times and service times for MEDEVAC unit 4 are reduced due to the increased cruising speed. Tables 15 and 16 reflect the updated changes. All other parameters listed in Table 8 remain the same.

Any positive changes to the MEDEVAC system result in a smaller gap between the myopic policy and the $\text{ADP}_{reported}$ policy. In this case, the V-22 Osprey improves the system by increasing the speed of the fourth MEDEVAC unit which reduces the time it takes a casualty to be transported to an MTF. Table 17 displays the value of the starting state for each of the policies. Again, both the Optimal_{truth} and $\text{ADP}_{reported}$

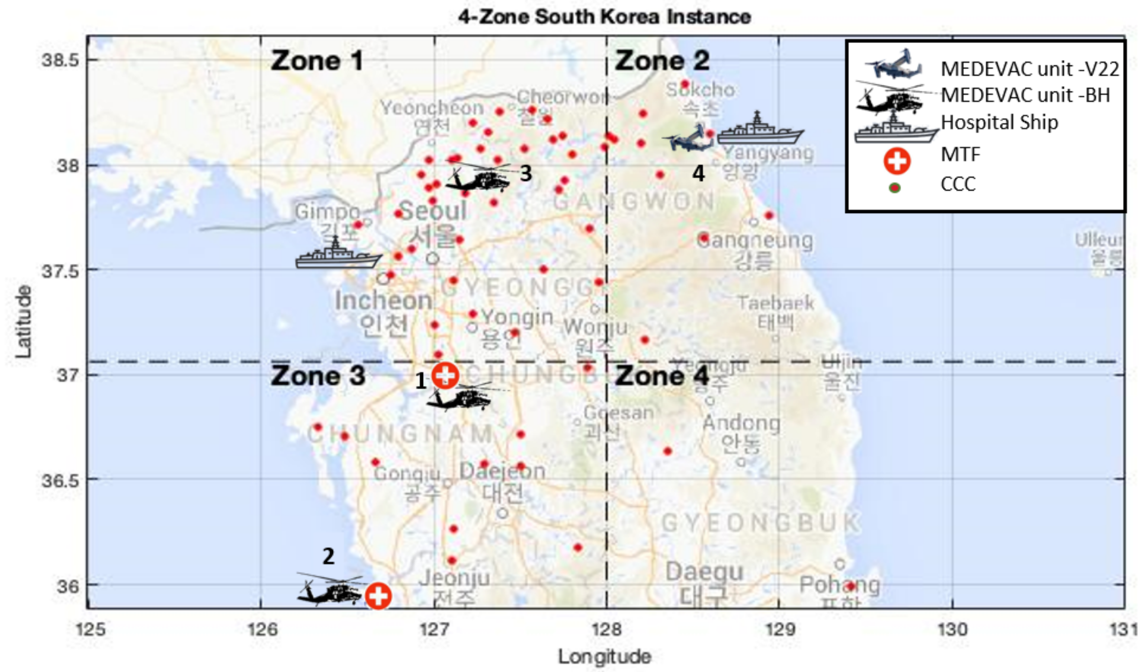

Figure 6. Excursion - MEDEVAC locations, zones, and CCCs

policies are generated on the reported triage classification but they are evaluated on the truth. The Optimal$_{reported}$ policy outperforms the myopic policy by 5.8% while the ADP$_{reported}$ policy outperforms the myopic by 4.73%. The ADP$_{reported}$ policy is 94.74% optimal, which promotes the efficacy of using random forest within the API algorithm to approximate the optimal policy.

Table 15. Expected Response Times (minutes) with V-22 Osprey

	Zone, z			
MEDEVAC unit, m	1	2	3	4
1	46.5000	58.9820	37.9980	62.2010
2	71.1390	82.0260	45.3790	69.1250
3	55.1840	65.1730	73.7650	116.2200
4	65.5050	46.2940	81.6380	92.7920

Table 16. Expected Service Times (minutes) with V-22 Osprey

	Zone, z			
MEDEVAC unit, m	1	2	3	4
1	87.2100	112.1500	60.4730	116.8100
2	133.4500	158.1900	81.9150	148.2800
3	76.7550	94.6460	101.2100	140.7200
4	82.8050	48.1950	107.5900	108.6400

Table 17. 4 Zone Policy Comparison with V-22 Osprey

Policy, π	$V^{\pi}(S_0)$	% Improvement over Myopic	% Optimal
Optimal_{truth}	3.04	13.30%	-
$\text{Optimal}_{reported}$	2.91	5.80%	95.72%
ADP_{truth}	3.01	9.48%	99.01%
$\text{ADP}_{reported}$	2.88	4.73%	94.74%
Myopic	2.75	-	90.46%

4.5 Excursion - 20 Zone

This section expands the original 4-zone problem into 20 zones. In doing so, we achieve greater fidelity in the response times and service times for all MEDEVAC units. Recall that the response and service times are calculated by averaging the results from a Monte Carlo simulation. By having a smaller geographic zone, the averages become more representative of the true response and service time, as noted in Tables 18 and 19.

Table 18. Expected Response Times (minutes) 20 Zone Scenario

	MEDEVAC unit, m			
Zone, z	1	2	3	4
1	49.755	74.174	50.824	78.262
2	53.217	78.406	59.499	73.453
3	59.498	84.046	61.434	51.155
4	64.991	88.354	61.368	34.832
5	43.629	67.392	47.313	74.990
6	46.067	71.094	53.345	72.062
7	53.170	76.609	63.053	54.860
8	63.161	84.049	71.491	47.351
9	33.783	57.624	50.213	72.748
10	33.546	58.311	54.378	70.316
11	44.896	65.190	73.384	71.923
12	66.168	82.022	95.485	75.604
13	36.656	43.398	71.745	91.933
14	34.230	47.830	68.284	78.972
15	48.496	60.747	88.764	85.798
16	70.084	80.729	115.630	98.309
17	40.205	36.449	77.812	93.597
18	43.181	38.103	82.296	93.036
19	48.647	48.904	95.734	98.936
20	74.623	77.365	140.920	128.640

The location of the MEDEVAC units, MTFs, Hospital Ships, and CCCs are unchanged from the baseline problem as shown in Figure 7. The number of zones directly impacts the size of the state space. Utilizing Equation (1) from Chapter III we calculate the size of the state space to be 35,201,061. This problem suffers from the *curse of dimensionality* and is computationally intractable. Therefore, we cannot

Table 19. Expected Service Times (minutes) 20 Zone Scenario

	MEDEVAC unit, m			
Zone, z	1	2	3	4
1	86.742	130.79	67.299	121.37
2	107.3	153.82	83.079	94.358
3	115.38	162.79	91.361	52.225
4	112.46	158.69	91.296	35.901
5	74.342	117.73	63.787	118.09
6	87.87	133.76	73.479	104.78
7	110.8	157.11	92.971	55.974
8	116.47	160.23	101.42	48.42
9	52.101	98.298	69.315	113.88
10	47.357	97.348	76.246	109.37
11	83.219	127.98	98.127	97.687
12	133.61	172.33	125.41	76.673
13	53.714	85.567	93.613	131.05
14	48.459	87.389	90.252	117.95
15	76.993	114.57	110.73	124.78
16	148.87	182.38	145.56	99.378
17	73.927	60.649	116.91	148.41
18	82.936	58.927	126.02	152.13
19	87.918	91.524	128.51	147.9
20	154.8	181.26	168.09	142.81

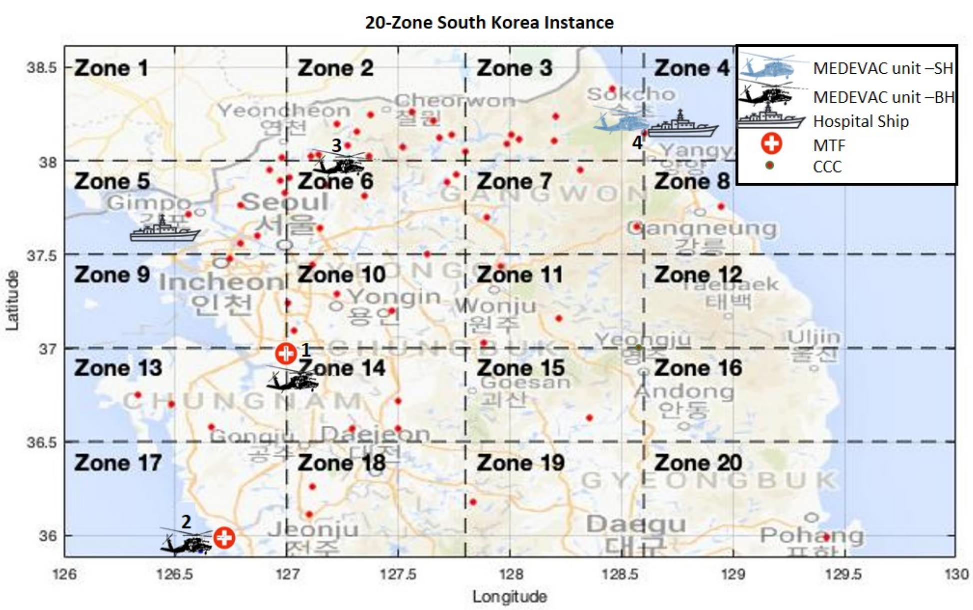

Figure 7. Excursion - MEDEVAC locations, zones, and CCCs

determine the optimal policy, but instead approximate a solution using the RF-API algorithm.

Prior to re-tuning the ADP algorithmic parameter settings, we run the RF-API algorithm for the 20-zone scenario using the best parameters from the 4-zone design of experiment (i.e., displayed in Run 1 in Table 11). The $\text{ADP}_{reported}$ policy generated using these settings achieves an 8.64% ± 0.09% improvement over the myopic policy with 95% confidence. Although these results demonstrate a high-quality solution, re-tuning the ADP algorithmic parameters may lead to an even better result. Due to the nature of this problem (i.e., the choice of a dispatching policy can result in more or less lives saved), it is necessary to re-tune the algorithmic parameter settings to explore other ADP-generated policies that may yield better improvement over the myopic policy.

To explore other policies, we design a 4^3 full factorial design computational experiment for the 20-zone scenario problem, which is displayed in Table 20. These

Table 20. 20-Zone Experimental Design Factor Levels

Algorithm Parameters	Description	Levels
N	Policy Improvements	{3, 6, 9}
J	Policy Evaluations	{15,000, 20,000, 25,000}
κ	Step-size	{0.05, 0.10, 0.15}
F	Trees in Random Forest	{75, 100, 150}

specific factor levels were chosen after performing exploratory evaluations. Note that the time to run one design point ranges from 8-12 minutes given our computational resources.

We develop a simulation model and perform 50 replications of each parameter setting combination identified in Table 20 to evaluate the ETDR for the ADP_{truth}, $\text{ADP}_{reported}$, and myopic policies. The best parameter settings associated with the $\text{ADP}_{reported}$ policy from the experiment are $N = 6$, $J = 20,000$, $\kappa = 0.05$, and

$F = 150$. The results from this combination of parameter settings for the ADP_{truth} and $\text{ADP}_{reported}$ policies with 95% confidence are displayed in Table 21.

Table 21. 20 Zone Policy Comparison

Policy, π	% Improvement over Myopic
ADP_{truth}	$18.09\% \pm 0.14\%$
$\text{ADP}_{reported}$	$12.95\% \pm 0.10\%$

The results from Table 21 show that the tuned $\text{ADP}_{reported}$ policy is able to achieve nearly 13% improvement over the myopic policy. Furthermore, the results for the ADP_{truth} policy show that if zero misclassifications exist then over 18% improvement can be attained over the myopic policy. This highlights not only the importance of tuning parameters to achieve meaningful results, but also the effectiveness of the RF-API algorithm for the MEDEVAC dispatching problem. As noted in Section 4.3, we would expect the percent improvement over the myopic policy to increase as the intensity of combat operations escalates and MEDEVAC requests are submitted with high frequency and less interarrival times.

V. Conclusions and Recommendations

The purpose of this thesis is to examine a variation of the MEDEVAC dispatching problem and develop a dispatching rule that can more efficiently and effectively utilize MEDEVAC units to maximize battlefield casualty survivability rates. This research directly builds upon the work previously done by Graves *et al.* (2021) by incorporating similar problem features (i.e., triage misclassifaction errors and blood transfusion kits). An infinite horizon, continuous time Markov decision process (MDP) model is formulated to examine this problem. The dispatching policy is generated based on the reported information pertaining to the MEDEVAC request (i.e., originating zone and triage classification) while also considering the status of the MEDEVAC system (i.e., the status of the individual units).

In small problem instances, optimal policies can be obtained using exact dynamic programming techniques. As more realistic features are added to a problem, the state space grows and the problem becomes computationally intractable, creating a need to approximate a solution. This thesis adds to the MEDEVAC literature by utilizing a random forest value function approximation within an approximate policy iteration algorithm to produce high-quality solutions to the MEDEVAC dispatching problem.

We develop a notional scenario to demonstrate the applicability of our model and solution methodology. Our scenario deviates from the norm of middle east operations and instead looks at a scenario based out of South Korea, while also incorporating a joint military environment. In this environment, we utilize Naval hospital ships as MTFs, a feature that has not been explored in the MEDEVAC dispatching problem.

In the small-scale problem problem instance we utilize the ADP algorithm to generate a dispatching policy and compare it to the optimal policy and the currently practiced policy. The ADP policy is approximately 94% optimal and outperforms the myopic policy by over 6%. This problem is then enlarged by expanding the number of zones from 4 to 20, thereby increasing our state space and subjecting ourselves to the *curse of dimensionality.* Although, we cannot solve this scenario to optimality, the small-scale demonstrates that the ADP solution can achieve high-quality solutions that outperform a myopic approach when properly tuned. After conducting a computational experiment to determine parameter settings, we generate an ADP policy that outperforms the myopic policy by nearly 13%.

As part of a sensitivity analysis, the arrival rate of MEDEVAC requests is altered to evaluate the effect it has on the differences between the optimal, ADP, and myopic policies. When MEDEVAC requests enter the system more quickly, ceteris paribus, the improvement that both the optimal and ADP polices offer over the myopic policy increases. At the same time, the ADP policies also begin to diminish in terms of its proximity to the optimal policy. This highlights the importance of properly tuning the RF-API algorithm to achieve high-quality results. As the MEDEVAC request arrival rate decreases, the difference between the optimal and ADP policies and the myopic policy decrease as well.

5.1 Recommendations

This research is valuable to the military community and the leaders therein who establish the dispatching policy for utilizing MEDEVAC assets. This work can assist and influence the manner in which MEDEVAC units are dispatched to maximize the survivability of battlefield casualties.

The analysis conducted in this thesis can be enhanced in future research. The following two recommendations are extensions to this line of research. First, the dispatching policies in this model were generated based on the reported triage classification. If the dispatching authority has a prior belief about the true classification of the request, then an expectation can be taken, which would likely improve the policy. Second, the unique distinction between a hospital ship and a general hospital is the fact that it is mobile. This problem could be mixed with a resource relocation problem to determine when and where to move the hospital ship(s) to reduce the time it takes MEDEVAC units to deliver a casualty to a MTF.

Bibliography

Almehdawe, Eman, Jewkes, Beth, & He, Qi-Ming. 2013. A Markovian queueing model for ambulance offload delays. *European Journal of Operational Research*, **226**(3), 602–614.

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.

Bastian, Nathaniel D. 2010. A robust, multi-criteria modeling approach for optimizing aeromedical evacuation asset emplacement. *The Journal of Defense Modeling and Simulation*, **7**(1), 5–23.

Bradley, Matthew, Nealeigh, Matthew, Oh, John S., Rothberg, Philip, Elster, Eric A., & Rich, Norman M. 2017. Combat casualty care and lessons learned from the past 100 years of war. *Current problems in surgery*, **54**(6), 315–351.

Britannica, The Editors of Encyclopaedia. 2020 (Feb). *Demilitarized zone.*

Correll, Diana Stancy. 2021 (Apr). *V-22 Osprey conducts inaugural landing on Hospital Ship Mercy's new Flight Deck.*

DCAS. 2021 (Jul). *U.S. Military Casualties - Operation Enduring Freedom (OEF) Wounded in Action.*

Dennie, Nathaniel C. 2021. *The Impact of Threat Levels at the Casualty Collection Point on Military Medical Evacuation System Performance.* M.Phil. thesis, Air Force Institute of Technology.

Department of the Army, United States. 2019 (July). *Field Manual 4-02.2, Medical Evacuation.*

Department of the Navy, United States. 2019 (Jan). *PATIENT MOVEMENT.*

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.

Freudenrich, Craig. 2001 (Apr). *How the V-22 Osprey Works*.

Garrett, Michael X. 2013. *USCENTCOM review of MEDEVAC procedures in Afghanistan*. Tech. rept. United States Central Command.

Graves, Emily S., Jenkins, Phillip R., & Robbins, Matthew J. 2021. Analyzing the Impact of Triage Classification Errors on Military Medical Evacuation Dispatching Policies. *Pages 1–12 of: 2021 Winter Simulation Conference (WSC)*. IEEE.

Hernandez, Marc, Jackson, Henry, Meza, Oscar, McKenney, Craig, Morgan, Rebecca, Palermo, Billy, Roach, Sommer, Spaterna, Al, & Wathen, Diane. 2010. *MH-60 Seahawk/MQ-8 Fire Scout Interoperability*. Tech. rept. NAVAL POSTGRADUATE SCHOOL MONTEREY CA DEPT OF SYSTEMS ENGINEERING.

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 Decision-making under Uncertainty to Save Lives. *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**, 102088.

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 Operational Research*, **290**(1), 132–143.

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.

Kotwal, Russ S, Howard, Jeffrey T, Orman, Jean A, Tarpey, Bruce W, Bailey, Jeffrey A, Champion, Howard R, Mabry, Robert L, Holcomb, John B, & Gross, Kirby R. 2016. The effect of a golden hour policy on the morbidity and mortality of combat casualties. *JAMA Surgery*, **151**(1), 15–24.

Kuisma, Markku, Holmström, Peter, Repo, Jukka, Määttä, Teuvo, Nousila-Wiik, Maria, & Boyd, James. 2004. Prehospital mortality in an EMS system using medical priority dispatching: a community based cohort study. *Resuscitation*, **61**(3), 297–302.

Kulkarni, Vidyadhar G. 2017. *Modeling and analysis of stochastic systems: Third edition*. CRC Press.

Li, Mengyu, Carter, Alix, Goldstein, Judah, Hawco, Terence, Jensen, Jan, & Vanberkel, Peter. 2021. Determining ambulance destinations when facing offload delays using a Markov decision process. *Omega*, **101**, 102251.

Lim, Cheng Siong, Mamat, Rosbi, & Braunl, Thomas. 2011. Impact of Ambulance Dispatch Policies on Performance of Emergency Medical Services. *IEEE Transactions on Intelligent Transportation Systems*, **12**(2), 624–632.

Malsby III, Robert F, Quesada, Jose, Powell-Dunford, Nicole, Kinoshita, Ren, Kurtz, John, Gehlen, William, Adams, Colleen, Martin, Dustin, & Shackelford, Stacy. 2013. Prehospital blood product transfusion by US Army MEDEVAC during combat operations in Afghanistan: a process improvement initiative. *Military Medicine*, **178**(7), 785–791.

Maxwell, David. 2020 (Sep). *ROK/US combined training protects integrity of the OPCON transition process*.

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

Military Sealift Command. 2020 (April). *Hospital Ships T-AH*.

Millet, Allan R. 2021. Korean War. *Encyclopedia Britannica*, Jun.

Mukhopadhyay, Ayan, Pettet, Geoffrey, Samal, Chinmaya, Dubey, Abhishek, & Vorobeychik, Yevgeniy. 2019. An online decision-theoretic pipeline for responder dispatch. *Proceedings of the 10th ACM/IEEE International Conference on Cyber-Physical Systems*.

Porcelli, Anthony J, & Delgado, Mauricio R. 2017 (Apr). *Stress and Decision Making: Effects on Valuation, Learning, and Risk-taking.*

Puterman, Martin L. 2005. *Markov Decision Processes: Discrete Stochastic Dynamic Programming.* 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. 2020. Approximate dynamic programming for the aeromedical evacuation dispatching problem: Value function approximation utilizing multiple level aggregation. *Omega*, **91**, 102020.

Ruszczynski, Andrzej. 2010. Commentary–Post-decision states and separable approximations are powerful tools of approximate dynamic programming. *INFORMS Journal on Computing*, **22**(1), 20–22.

Saghafian, Soroush, Hopp, Wallace J, Van Oyen, Mark P, Desmond, Jeffrey S, & Kronick, Steven L. 2014. Complexity-augmented triage: A tool for improving patient safety and operational efficiency. *Manufacturing & Service Operations Management*, **16**(3), 329–345.

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

U.S. Naval Academy. 2021. *Aviation Warfare.*

USAASC. 2021. Black Hawk Utility Helicopter - UH/HH-60. *United States Army Acquisition Support Center.*

Wooten, Kylie. 2021. *Examining How Standby Assets Impact Optimal Dispatching Decisions within a Military Medical Evacuation System via a Markov Decision Process Model.* M.Phil. thesis, Air Force Institute of Technology.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704–0188

1. REPORT DATE *(DD–MM–YYYY)*
24–03–2022

2. REPORT TYPE
Master's Thesis

3. DATES COVERED *(From — To)*
September 2020 — March 2022

4. TITLE AND SUBTITLE

Analyzing the Impact of Blood Transfusion Kits and Triage Misclassification Errors for Military Medical Evacuation Dispatching Policies via Approximate Dynamic Programming

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Rodriguez, Channel A., Capt, 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-22-M-166

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

14. ABSTRACT

Members of the armed forces greatly rely on having an effective and efficient medical evacuation (MEDEVAC) process for evacuating casualties from the battlefield to medical treatment facilities (MTF) during combat operations. This thesis examines the MEDEVAC dispatching problem and seeks to determine an optimal policy for dispatching a MEDEVAC unit, if any, when a 9-line MEDEVAC request arrives, taking into account triage classification errors and the possibility of having blood transfusion kits on board select MEDEVAC units. A discounted, infinite-horizon continuous-time Markov decision process (MDP) model is formulated to examine such problem and compare generated dispatching policies to the myopic policy of sending the closest available unit. We utilize an approximate dynamic programming (ADP) technique that leverages a random forest value function approximation within an approximate policy iteration algorithmic framework to develop high-quality policies for both a small-scale problem instance and a large-scale problem instance that cannot be solved to optimality. A representative planning scenario involving joint combat operations in South Korea is developed and utilized to investigate the differences between the various policies. Results from the analysis indicate that applying ADP techniques can improve current practices by as much as 29% with regard to a life-saving performance metric. This research is of particular interest to the military medical community and can inform the procedures of future military MEDEVAC operations.

15. SUBJECT TERMS

Markov decision processes, medical evacuation, approximate dynamic programming, policy iteration, random forest

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
U	U	U	UU	70	Dr. Phillip R. Jenkins, AFIT/ENS
					19b. TELEPHONE NUMBER *(include area code)* (937) 255-3636, x4727; phillip.jenkins@afit.edu

Standard Form 298 (Rev. 8–98)
Prescribed by ANSI Std. Z39.18