3-2021

Examining How Standby Assets Impact Optimal Dispatching Decisions within a Military Medical Evacuation System Via a Markov Decision Process Model

Kylie Wooten

Examining How Standby Assets Impact Optimal Dispatching Decisions within a Military Medical Evacuation System via a Markov Decision Process Model

THESIS

Kylie Wooten, Capt, USAF

AFIT-ENS-MS-21-M-196

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENS-MS-21-M-196

EXAMINING HOW STANDBY ASSETS IMPACT OPTIMAL DISPATCHING DECISIONS WITHIN A MILITARY MEDICAL EVACUATION SYSTEM VIA A MARKOV DECISION PROCESS MODEL

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

Kylie Wooten, B.S.

Capt, USAF

25 March 2021

AFIT-ENS-MS-21-M-196

EXAMINING HOW STANDBY ASSETS IMPACT OPTIMAL DISPATCHING DECISIONS WITHIN A MILITARY MEDICAL EVACUATION SYSTEM VIA A MARKOV DECISION PROCESS MODEL

THESIS

Kylie Wooten, B.S.
Capt, USAF

Committee Membership:

Capt Phillip R. Jenkins, PhD
Chair

Dr. Matthew J. Robbins
Member

AFIT-ENS-MS-21-M-196

Abstract

The Army medical evacuation (MEDEVAC) system ensures proper medical treatment is readily available to wounded soldiers on the battlefield. The objective of this research is to determine which MEDEVAC unit to task to an incoming 9-line MEDEVAC request and where to station a single standby unit to maximize patient survivability. A discounted, infinite-horizon continuous-time Markov decision process model is formulated to examine this problem. We design, develop, and test an approximate dynamic programming (ADP) technique that leverages a least squares policy evaluation value function approximation scheme within an approximate policy iteration algorithmic framework to solve practical-sized problem instances. A computational example is applied to a synthetically generated scenario in Iraq. The optimal policy and ADP-generated policies are compared to a commonly practiced (i.e., myopic) policy. Examining multiple courses of action determines the best location for the standby MEDEVAC unit, and sensitivity analysis reveals that the optimal and ADP policies are robust to standby unit mission preparation times. The best performing ADP-generated policy is within 2.62% of the optimal policy regarding a patient survivability metric. Moreover, the ADP policy outperforms the myopic policy in all cases, indicating the currently practiced dispatching policy can be improved.

This research is dedicated to the men and women of the United States military that gave the ultimate sacrifice in service to our country. I hope this research is continued in order to provide the most efficient medical evacuation system possible for our deployed forces who risk their lives to defend this great nation.

Acknowledgements

Throughout the writing of this thesis I have received a great deal of support and assistance. I would first like to thank my advisor, Dr. Jenkins, for his mentorship, guidance, and patience. Without his supervision, this thesis would not have been possible, and I am thankful for having the opportunity to work with him. In addition, I would like to thank my committee member, Dr. Robbins. His expertise and professional dedication greatly assisted in the development of this research. Finally, I would like to thank all of my classmates who assisted me in and out of the classroom during my time here at AFIT. You all have made me a better student, analyst, and Air Force officer.

Kylie Wooten

Table of Contents

List of Figures

List of Tables

EXAMINING HOW STANDBY ASSETS IMPACT OPTIMAL DISPATCHING DECISIONS WITHIN A MILITARY MEDICAL EVACUATION SYSTEM VIA A MARKOV DECISION PROCESS MODEL

I. Introduction

The Army medical evacuation (MEDEVAC) system provides the necessary means to ensure proper medical treatment is readily available to wounded soldiers on the battlefield. MEDEVAC units rapidly respond to battlefield casualties and evacuate them to an appropriate, nearby medical treatment facility (MTF). Moreover, MEDEVAC units have dedicated on board medical personnel that provide en route medical care to casualties with an objective to maintain or improve the conditions of the casualties during evacuation. Senior military leaders and medical personnel are responsible for the management of scarce medical resources within the MEDEVAC system and determine how these resources are distributed and utilized during battlefield operations. Effective and efficient use of medical resources corresponds to higher soldier morale by demonstrating that specialized medical care is quickly available to the wounded (Department of the Army, 2019).

Although both air and ground evacuation platforms are incorporated in Army MEDEVAC units, the HH-60M Black Hawk helicopter is most often used to evacuate casualties. Whereas ground vehicles are hindered by roads, terrain, and possible traffic, rotary-wing aircraft (e.g., HH-60M) are able to fly directly to a casualty collection point (CCP) (i.e., where casualties are assembled for evacuation) and then fly directly to an MTF. Hence, air assets generally provide faster response times than ground assets, making them the preferred platform of choice. The HH-60M is also equipped

with a medical interior integrated with a litter system that enables the transport of up to six patients (Buckenmaier & Mahoney, 2015). These capabilities, combined with the protection of the Geneva Conventions from intentional enemy attack, provide HH-60M helicopters with the ability to evacuate patients to an MTF efficiently without enemy intervention.

Military medical planners are responsible for designing MEDEVAC systems and operations. For instance, air assets must be strategically stationed to maximize coverage while minimizing response time. CCPs also need to be predesignated in optimal locations, and they may or may not be staffed based on risk management and personnel availability. Determining a dispatching policy is another vital aspect of MEDEVAC planning. A dispatching order needs to be identified that maximizes patient survivability (or minimizes response time). However, the complexity and uncertainty of MEDEVAC missions makes dispatching decisions difficult. For example, aircraft reliability, enemy threat levels, personnel requirements, technical issues, and weather are possible sources of uncertainty that may impact dispatching decisions, which makes it difficult for medical planners to optimize MEDEVAC procedures.

An important difference between this thesis and other MEDEVAC research is the incorporation of a standby unit. The inclusion of a standby unit has not yet been researched for civilian or military emergency medical services (EMS) systems. A standby MEDEVAC unit is available to respond to 9-line MEDEVAC requests (i.e., requests for evacuation containing nine standardized lines of communication), but might do so at a slower rate than a primary unit. Standby units are co-located with primary units and may be tasked at anytime. For example a standby unit may be tasked to respond to a non-life-threatening request so that the primary unit may be reserved for a life-threatening request expected to occur in the near future. Likewise, if a primary unit is busy when a new request arrives, the standby unit can respond

to minimize response time. A standby unit can also be relocated at any time. For example, it may be beneficial to relocate a standby unit to a different staging facility if requests are more likely to arrive in a different area.

This thesis focuses on the decisions of which MEDEVAC unit to task to an incoming 9-line MEDEVAC request and where to station a single standby MEDEVAC unit. Similar to Jenkins (2017), admission control is incorporated, which allows any request to be rejected by the dispatching authority and handled by an outside organization. This enables the dispatching authority to reserve MEDEVAC units for higher priority requests. The decision of if and when to reject a request is incorporated into the dispatching policy. The reported dispatch policy is based on the location and status of MEDEVAC units as well as the location and priority level of an incoming 9-line MEDEVAC request (e.g., Priority I - Urgent and Priority II - Priority). The military often defaults to a myopic policy that is easy to implement, such as always tasking the MEDEVAC unit that is closest to the CCP regardless of important system characteristics (e.g., request priority level), but this is often not the optimal policy (Jenkins, 2017). Therefore, differences in the optimal policy and a myopic policy are explored. The reported policy also dictates where to place a standby MEDEVAC unit and when to task the standby unit.

An infinite horizon, continuous-time Markov decision process (MDP) is formulated to determine an optimal dispatching policy and standby operations that will maximize the expected total discounted reward (ETDR). Uniformization is applied to transform the continuous-time MDP to an equivalent, more easily analyzed discrete-time MDP. The location of primary MEDEVAC units and the locations wherein casualties occur are known. It is assumed that the standby MEDEVAC unit can be co-located with any primary unit. In addition to solving the MDP model to optimality, the MDP model is also solved via an approximate dynamic programming (ADP)

solution approach that utilizes a least squares policy evaluation (LSPE) value function approximation scheme within an approximate policy iteration (API) algorithmic framework. A computational example is applied to a synthetically generated scenario in Iraq, and the optimal policy is compared to a myopic policy and an ADP-generated policy calculated via API-LSPE.

This thesis is organized as such: Chapter II provides a review of research relating to EMS systems as well as MDP and ADP techniques. Chapter III outlines the MDP formulation developed to determine an optimal MEDEVAC dispatch policy and standby unit operations and the ADP formulation developed to determine a high-quality policy. Chapter IV covers an application of the formulated MDP based on a representative scenario in Iraq along with sensitivity analysis and excursions. Finally, Chapter V concludes the thesis and proposes directions for future research.

II. Literature Review

Over the last 50 years, ample research has been conducted on the optimization of both civilian and military EMS systems. This research is focused on decision-making regarding EMS system components such as the location of servers (e.g., ambulances) (Daskin & Stern, 1981; Rettke *et al.*, 2016; Jenkins, 2019), the number of servers per location (Zeto *et al.*, 2006; Fulton *et al.*, 2010), the server dispatching policy (Carter *et al.*, 1972; Bandara *et al.*, 2012; Jenkins *et al.*, 2021a,b), and a zone tessellation strategy for the service area (Mayorga *et al.*, 2013). Researchers must also identify which performance measure to focus on as the objective: response time thresholds (RTTs) or patient survivability rates (McLay & Mayorga, 2010). With the addition of a standby MEDEVAC unit, this thesis also includes the decisions of where to locate and when to task the standby unit. Although these decisions apply to both civilian and military research, the research in each field varies due to differing mission requirements.

Initial EMS research by Carter *et al.* (1972) explores the idea of an optimal dispatching policy. This research reveals that dispatching the nearest server to the service call does not always produce the lowest average response time. That is, dispatching the closest server, a policy that is easy to implement (often called a myopic policy), is not always the optimal policy. This insight confirms the need for EMS system optimization research and is revisited and confirmed by many sources. For instance, Bandara *et al.* (2012) shows that the myopic policy is suboptimal when service priority levels (e.g., high priority and low priority) are included in the MDP model formulation. This is especially evident in sensitivity analysis of priority level ratios; as the percentage of high priority service requests from a particular service zone increases, the optimal policy suggests to reserve the server that responds the fastest to the high priority requests that are most likely to occur. This is a key con-

cept for military MEDEVAC research because service requests include priority levels (e.g., Priority I - Urgent, Priority II - Priority, Priority III - Routine).

Although research relating to civilian services informs military medical planners, this thesis expands upon military MEDEVAC-specific research. The decision-making process is similar between civilian and military EMS systems, but the nature of military MEDEVAC missions present unique challenges (Jenkins *et al.*, 2020a,b). Military MEDEVAC systems are designed and set in place in a combat environment only when needed, but civilian medical systems are permanent establishments. For instance, in civilian EMS research, locations of hospitals are typically assumed to be known, but this is not always the case with military EMS systems. MTF placement is at the discretion of military medical planners (Rettke *et al.*, 2016). Similarly, civilian systems primarily use ground vehicles (ambulances) as servers, whereas the military MEDEVAC vehicle of choice is the HH-60M helicopter. Additional differences between civilian and military EMS systems are discussed by Keneally *et al.* (2016), Jenkins *et al.* (2021a), and Jenkins *et al.* (2021c). These differences warrant the research of military-specific EMS systems.

It is vital to select a suitable performance metric to produce an accurate model with meaningful results. Most EMS systems measure overall performance according to an RTT (McLay & Mayorga, 2010). An RTT is the maximum time for a server to respond to cover the request. When modeling EMS systems, RTTs are usually preferred over other metrics related to patient outcomes because they are easier to evaluate. Due to the nature of the differing mission requirements and structure, civilian and military EMS systems have different RTT requirements. In 2009, the Department of Defense mandated that the United States MEDEVAC system respond to critically injured combat casualties in 60 minutes or less, also known as the golden-hour rule (Kotwal *et al.*, 2016). According to Kotwal *et al.* (2016), after the mandate

there was an overall decrease in percentage killed in action (from 16.0% to 9.9%) and 52% median reduction in transport time.

Although RTTs seem to have many benefits, one common criticism relates to how well patient survivability is captured when utilizing RTTs. Since we ultimately seek to maximize patient survivability, a better performance measure is patient survivability rate. Research shows that performance measures based directly on patient survivability provide more accurate results than using RTTs (Bandara *et al.*, 2014). However, estimating patient survivability tends to be a difficult task due to the lack of available patient survival information due to privacy regulations (McLay & Mayorga, 2010). Also, a casualty may not be discharged and considered "survived" for several months and can transfer to different medical facilities while being treated, making the task of tracking casualty survivability tedious and difficult (Rettke *et al.*, 2016). Nevertheless, many researchers (McLay & Mayorga, 2010; Bandara *et al.*, 2012; Mayorga *et al.*, 2013; Bandara *et al.*, 2014) utilize patient survivability as the performance measure in studying the optimization of EMS systems. As such, one of the objectives of this thesis is to identify an optimal dispatching policy for MEDEVAC systems that maximize the probability of battlefield casualty survivability.

Although the topic of EMS systems has been the general focus of many articles, authors often distinguish their work by introducing unique system enhancements. EMS systems are complex and multifaceted, and most system distinctions warrant their own dedicated research. For instance, McLay & Mayorga (2013a) discuss how equity of servers and equity of customers is a key yet controversial factor when deciding how to allocate resources. This idea is explored specifically when determining how to dispatch ambulances to prioritize patients. McLay & Mayorga (2013a) formulate an MDP model with four types of equity constraints, and the respective optimal policies are compared in a computational example. Similarly, McLay & Mayorga (2013b)

introduce the idea of triage classification errors and the effects of over-responding versus under-responding to misclassified patients. The objective is to determine which ambulance to dispatch to arriving patients to maximize the expected coverage (e.g., the probability of achieving a predefined RTT) of high-risk patients.

A MEDEVAC-specific system enhancement is present in Keneally *et al.* (2016). Keneally *et al.* (2016) discusses the idea of armed escorts in a high threat environment as part of a MEDEVAC scenario, which introduces an additional complication to the EMS problem. The authors develop an MDP model to determine MEDEVAC dispatching policies in a combat environment with the added complication of threat conditions. In this situation, a MEDEVAC helicopter may require an armed escort. The MDP model indicates a dispatching order that maximizes steady-state utility. Computational examples are also included to investigate optimal policies in different threat environments.

Similar to Jenkins *et al.* (2018), this research includes admission control. Consideration of admission control and queuing can greatly improve the performance of this type of system (Stidham Jr., 2002). Admission control allows the decision-maker to observe the current state of the system when a 9-line MEDEVAC request arrives and decide whether to admit the request based on the request priority level and location of origin. Requests that are admitted enter the system and must be serviced immediately. Those that are rejected never enter the system and are handled by an outside organization. Without the inclusion of a system controller (i.e., the MEDEVAC dispatching authority), system behavior can be inconsistent with periods of long queues followed by periods of inactivity (Puterman, 1994).

When developing a decision rule or policy to optimize EMS systems, operations research methods such as linear programming, MDP techniques, and ADP techniques have been widely used. For example, Jenkins *et al.* (2020c) use a mixed integer linear

program to solve the MEDEVAC location-allocation problem, which determines the optimal placement of MEDEVAC assets. Bandara *et al.* (2012) use an MDP model to maximize patient survivability of a civilian EMS system by determining a dispatching order for ambulances. Similarly, Jenkins (2017) formulates and solves an MDP model to determine an optimal MEDEVAC dispatching policy. The model allows the dispatching authority to either accept, reject, or queue the incoming request for service. Patient survivability with respect to priority level is embedded in the model to maximize the overall patient survivability rate. Results indicate that the myopic policy is not always optimal; dispatching MEDEVAC units based on the priority level and zone of incoming requests increases the performance of the MEDEVAC system.

When researchers utilize MDPs to model EMS scenarios, uniformization is commonly applied to transform a continuous-time problem into a discrete-time problem. Uniformization is applied to a continuous-time MDP model to obtain a model with constant transition rates so that algorithms for discrete-time discounted models may be applied. The resulting optimal policy of the uniformized system is equivalent to the optimal policy of the original continuous-time system (Puterman, 1994). More details on how uniformization is applied to this thesis are discussed in Chapter III.

As the size of a problem increases, MDP techniques become less viable. In these cases, ADP is utilized to determine high-quality policies (Powell, 2011). For example, Rettke *et al.* (2016) utilize ADP to solve a realistic MEDEVAC scenario. More specifically, the authors formulate an MDP model of the MEDEVAC dispatching problem and apply an ADP solution approach that utilizes least squares temporal differences (LSTD) in an API framework to generate high-quality dispatching policies. The ADP-generated policies outperform a myopic approach by over 30% in regards to a life-saving metric. Similarly, Jenkins *et al.* (2021b) formulate an MDP model to solve the MEDEVAC dispatching, preemption-rerouting, and redeployment (DPR)

problem. The authors use a support vector regression (SVR) value function approximation (VFA) scheme within an API algorithmic framework to generate high-quality policies for a realistic scenario set in Azerbaijan. Their results from computational experiments indicate the ADP-generated policies significantly outperform the two benchmark policies considered.

ADP techniques are also seen in civilian EMS research such as Maxwell *et al.* (2010) and Schmid (2012). Maxwell *et al.* (2010) explores API to determine where we should redeploy (i.e., reposition) idle ambulances to maximize the number of calls reached within a delay threshold, also known as the ambulance redeployment problem. Initial results show that the ADP policy outperforms both a myopic policy and a static policy by 4.7% and 4.0%, respectively. Of note, this result is achieved without adding any extra resources to the EMS system. The algorithm parameters are tuned to minimize computational run time while maintaining a high-quality solution.

Schmid (2012) applies an ADP solution approach to minimize the average response time observed by properly dispatching and positioning ambulances within the Austrian EMS system. The author uses a pure aggregation approach with a generic approximate value iteration (AVI) algorithm. Since she is concerned with both ambulance relocation and dispatching, she first examined them individually. She tested three relocation strategies and found that the ADP relocation strategy decreases response time by 12.08%. She also found that the ADP dispatching strategy decreases response time by 12.89%. Combining these two approaches yields an improvement of 7% over Austria's current dispatching policy. Algorithmic parameters (e.g., step size, decay parameter, and aggregation) were chosen using a computational experiment.

This research features LSPE in an API framework to generate high-quality MEDEVAC dispatching policies. Summers *et al.* (2020) utilize this algorithm to generate a high-quality firing policy for interceptor allocation to incoming missiles. The ob-

jective is to minimize the expected total damage to defended assets over a sequence of engagements. This dynamic weapon target assignment problem is formulated as an MDP and solved via API-LSTD and API-LSPE. A computational experiment investigates problem features such as conflict duration, attacker and defender weapon sophistication, and defended asset values. A comparison of the ADP policies and two baseline policies show that the ADP policies outperform both baseline polices when conflict duration is short and attacker weapons are sophisticated. Although Summers *et al.* (2020) utilize API-LSPE to solve a different military application problem, the ADP methodology presented in their work informs this research due to similar algorithm implementation.

III. Methodology

This chapter outlines the MDP model for the military MEDEVAC dispatching problem along with a problem description and an ADP solution approach. The MDP model provides a framework with which a dynamic programming algorithm is used to compute an exact optimal policy for a small problem instance. The problem is then expanded, and an ADP technique is utilized to generate high-quality dispatching policies. The following MDP components are described in detail in this chapter: decision epochs, state space, action space, transition probabilities, rewards, objective, and optimality equations. This MDP model formulation provides the basis for the ADP solution approach discussed later in the chapter.

3.1 MDP Formulation

The general support aviation battalion (GSAB) manages all aerial operations, including Army HH-60M helicopters employed in MEDEVAC missions. Therefore, an Army aeromedical evacuation officer (AEO) that works within the GSAB serves as the decision-maker for the military MEDEVAC system (Department of the Army, 2019). When a 9-line MEDEVAC request is received, the AEO must quickly make a dispatching decision. Since any delay in decision-making could cause the casualty survivability rate to decrease, it is crucial to implement a dispatching policy that optimizes the evacuation of battlefield casualties to an appropriate, nearby MTF. Figure 1 depicts the MEDEVAC mission timeline. This timeline leverages the procedures outlined in the Army's Medical Evacuation Field Manual (Department of the Army, 2019) and other MEDEVAC research sources (Keneally *et al.*, 2016; Rettke *et al.*, 2016; Jenkins *et al.*, 2018).

A 9-line MEDEVAC request is transmitted in a standardized message format.

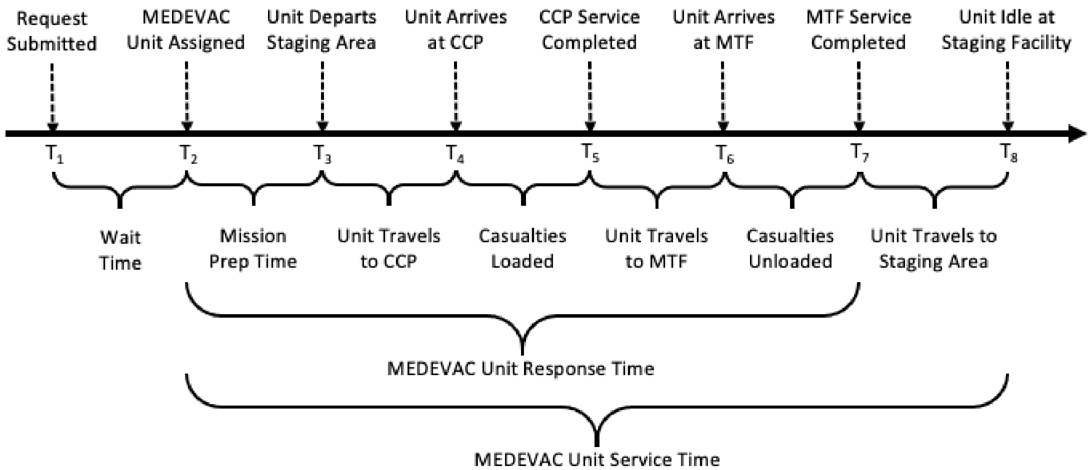

Figure 1. MEDEVAC Mission Timeline

The required information included in a 9-line MEDEVAC request is reported in the following order: the location of the pickup site, radio frequency and call sign, number of casualties by priority, special equipment required, number of casualties by type, security of pickup site, method of marking pickup site, casualty nationality and status, and chemical, biological, radiological, and nuclear contamination (Department of the Army, 2019). A Priority III – Routine evacuation is assigned to casualties that are deemed as minimally injured and unlikely to deteriorate and can tolerate an evacuation delay of up to 24 hours (De Lorenzo, 2003). These patients are typically evacuated by ground or waterborne assets. Since the focus of this research is aerial MEDEVAC missions, this thesis only considers 9-line MEDEVAC requests that are Priority I – Urgent or Priority II - Priority. The US Army defines these priority levels as follows (Department of the Army, 2019):

1. Priority I - Urgent: Assigned to emergency cases that should be evacuated as soon as possible and within a maximum of 60 minutes in order to save life, limb, or eyesight, to prevent complications of serious illness, or to avoid permanent disability.

2. Priority II - Priority: Assigned to sick and wounded personnel requiring prompt

medical care. This precedence is used when the individual should be evacuated within 240 minutes or his medical condition could deteriorate to such a degree that he will become Priority I - Urgent, or whose requirements for special treatment are not available locally, or who will suffer unnecessary pain or disability.

For the purpose of this research, if there are multiple casualties in a single request, the overall 9-line MEDEVAC request priority is based on the most time-sensitive casualty within the request. Casualties are evacuated to a CCP once the 9-line request has been communicated. The time at which the AEO receives the request is denoted by T_1.

The decision epochs of the MEDEVAC system are the points in time that require a decision and are denoted by $\mathcal{T} = \{1, 2, ...\}$. Two event types in the MEDEVAC system constitute all decision epochs. The first is the receipt of a 9-line MEDEVAC request. Upon receipt of the 9-line MEDEVAC request, the AEO must decide whether to admit the request based on the current status and location of the MEDEVAC units and the priority level and location of the request. If a Priority I request is expected to occur in the near future, the AEO may reject an incoming Priority II request from entering the system. If the AEO admits the request into the MEDEVAC system, another decision must be made as to which MEDEVAC unit should respond. If only one MEDEVAC unit is available, that unit will respond by default. However, if more than one unit is available, the AEO will make the decision based on the current location of the available MEDEVAC units and the priority and location (i.e., zone) of the request. The AEO assigns a MEDEVAC unit to the request at time T_2. A new request is automatically rejected if all MEDEVAC units are busy servicing other requests. It is assumed that rejected requests are still serviced by non-aerial MEDEVAC platforms. The second event that requires a decision is the change in status of a MEDEVAC unit from busy to idle upon completion of a mission. This

unit is either immediately retasked to respond to another request, or the AEO chooses for this unit to remain idle.

The state $S_t \in \mathcal{S}$ describes the status of all components of the MEDEVAC system at epoch $t \in \mathcal{T}$, which comprises MEDEVAC unit status and the current 9-line MEDEVAC request status. Let $S_t = (M_t, \hat{R}_t)$, wherein M_t represents the MEDEVAC status tuple at epoch t, and $\hat{R}_t$ represents the request arrival status tuple at epoch t. The tuple M_t denotes the status of all MEDEVAC units, which is given by $M_t = (M_{tm})_{m \in \mathcal{M}}$, wherein $\mathcal{M} = \{1, 2, ..., |\mathcal{M}|\}$ represents the set of MEDEVAC units. Let $\mathcal{Z} = \{1, 2, ..., |\mathcal{Z}|\}$ represent the zones (i.e., locations) from which requests originate. The variable $M_{tm} \in \{0\} \cup \mathcal{Z}$ contains the status of MEDEVAC unit $m \in \mathcal{M}$ at epoch t. A status of $M_{tm} = 0$ indicates unit m is idle at epoch t, and a status of $M_{tm} = z$, where $z \in \mathcal{Z}$, indicates unit m is servicing a zone z request. Similarly, $\hat{R}_t$ denotes the zone and priority level of a new request awaiting an admission decision at epoch t and is given by $\hat{R}_t = (\hat{Z}_t, \hat{K}_t)_{\hat{Z}_t \in \mathcal{Z}, \hat{K}_t \in \mathcal{K}}$, where $\mathcal{K} = \{1, 2, ..., |\mathcal{K}|\}$ represents the set of priority levels. The variable $\hat{Z}_t$ represents the zone of the new request, and the variable $\hat{K}_t$ represents the priority level of the new request. If there are no new 9-line MEDEVAC requests at epoch t, then $\hat{R}_t = (0, 0)$.

Unfortunately, as the number of state variables (e.g., number of MEDEVAC units, zones, and priority levels) increase, the size of the state space grows exponentially. This is commonly referred to as the *curse of dimensionality*, which refers to a large or uncountable state space, action space, or outcome space (Powell, 2011). The size of the state space is calculated as follows:

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

As the state space increases, exact dynamic programming techniques become intractable for analyzing large-scale realistic scenarios. Even so, small problem in-

stances are solved to optimality to glean insights about the structure of the optimal policy. Larger problem instances are also examined via the use of ADP techniques.

In order to model the arrival rates of urgent and priority 9-line MEDEVAC requests, the arrivals must be split into their respective categories. Splitting is used to generate two or more counting processes from a single Poisson process (Kulkarni, 2017). The single Poisson process of request arrivals is separated into multiple processes based on priority level and location using a splitting technique. Let $\{N(t') : t' \geq 0\}$ be the original Poisson process $PP(\lambda)$. This $PP(\lambda)$ counts the number of 9-line MEDEVAC request received by the AEO during the time interval $(0, t']$. The original counting process is split into multiple counting processes based on the zone and priority level of the request. Let $R = \{(z, k) : (z, k) \in \mathcal{Z} \times \mathcal{K}\}$ denote the set of request categories, with a total of $|\mathcal{R}| = |\mathcal{Z}||\mathcal{K}|$ different categories. The original counting process is split into $|\mathcal{R}|$ processes $\{N_{zk}(t') : t' \geq 0\}$, $\forall (z, k) \in \mathcal{R}$, where each request belongs to one and only one category. Each request is categorized using a Bernoulli splitting mechanism given the parameters $p_{zk} > 0$, $\forall$ $(z, k) \in \mathcal{R}$ such that $\sum_{(z,k)\in\mathcal{R}} p_{zk} = 1$, where p_{zk} is the probability that a request originates in zone z and is of priority k. The Berniolli splitting mechanism allows the characterization of each process $\{N_{zk}(t') : t' \geq 0\}$, $\forall$ $(z, k) \in \mathcal{R}$ as a Poisson process with parameter λp_{zk}, denoted as $PP(\lambda p_{zk})$.

When a 9-line MEDEVAC request is received, the AEO must observe the current state of the system and make an admission decision followed by a dispatching decision, if needed. The AEO's possible actions include accepting a request and assigning an available MEDEVAC unit to service the request, or rejecting the request from ever entering the system. Let $x_t^{reject} \in \{\Delta, 0, 1\}$ represent the admission control decision at decision epoch t. If the AEO chooses to admit the request, then $x_t^{reject} = 0$. If the AEO chooses to reject the request, then $x_t^{reject} = 1$. If there is not an arrival request

at epoch t (i.e., $\hat{R}_t = (0,0)$), an admission decision is not necessary and $x_t^{reject} = \Delta$.

If a request is admitted to the MEDEVAC system, the AEO must choose which idle MEDEVAC unit to task to service the request. Let $\mathcal{I}(S_t) = \{m : m \in \mathcal{M}, M_{tm} = 0\}$ denote the set of idle MEDEVAC units at state S_t. The arrival request dispatch decision variable is represented by the tuple $x_t^d = (x_{tm}^d)_{m \in \mathcal{I}(S_t)}$, which describes the AEO's dispatching decision for the arrival request at epoch t. If $x_{tm}^d = 1$, then MEDEVAC unit $m \in \mathcal{I}(S_t)$ is tasked to service the request $\hat{R}_t$ at epoch t, and 0 otherwise.

Let $x_t = (x_t^{reject}, x_t^d)$ denote the tuple of decision variables at epoch t. The AEO's decision is bound by the following constraint:

$$\sum_{x \in \mathcal{I}(S_t)} x_{tm}^d \leq \mathbb{I}_{\{\hat{R}_t \neq (0,0)\}} \mathbb{I}_{\{x_t^{reject}=0\}}. \tag{2}$$

Here, $\mathbb{I}_{\{\hat{R}_t \neq (0,0)\}}$ is an indicator function that takes the value of 1 if an incoming request is present. Similarly, $\mathbb{I}_{\{x_t^{reject}=0\}}$ is an indicator function that takes the value of 1 if the incoming request is admitted to the system. Equation 2 ensures at most one MEDEVAC unit is dispatched at time t. Moreover, a unit can only be tasked if a request is admitted to the system.

The set of available actions at epoch t is denoted as follows:

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

where Equation 2 must be satisfied. The first case shows the only feasible action is to transition with no changes if there is not a request arrival at decision epoch t. The second case shows the only feasible action is to reject a new request if one arrives but

all MEDEVAC units are busy servicing other requests. The final case shows the set of feasible actions if a new request arrives and at least one MEDEVAC unit is idle.

Once a MEDEVAC unit has been tasked to respond to a request, it departs the staging area in pursuit of the CCP, denoted as time T_3 in Figure 1. The time between the MEDEVAC unit tasking, T_2, and the MEDEVAC unit departure, T_3, is the total mission preparation time, which includes refueling and re-equipping. The mission preparation time of a primary MEDEVAC unit and a standby MEDEVAC unit differ due to the nature of a standby unit. Whereas primary MEDEVAC units are ready for a tasking at any moment, standby MEDEVAC units are on-call and operate as a backup to the primary unit. A standby unit might include aircrew or maintenance crew that are off-shift or an aircraft that is in need of pre-flight checks. These differences result in a longer mission preparation time for the standby unit to be fully equipped for departure to the CCP.

T_4 denotes the time at which the MEDEVAC unit arrives at the CCP. Casualties are loaded on board the aircraft, and the helicopter departs the CCP to proceed to an MTF, denoted as T_5. The destination MTF is chosen in a deterministic manner based on the location of the CCP and therefore is not included as a separate element of this model. T_6 denotes the time at which the MEDEVAC unit arrives at the MTF.

Upon arrival to the MTF, the MEDEVAC unit unloads casualties, and the responsibility of medical care is transferred to the MTF medical staff. The MEDEVAC unit then departs the MTF at time T_7. Although retasking the MEDEVAC unit at this point is possible, MEDEVAC units often travel back to their respective staging area for aircraft, equipment, or crew requirements (Rettke *et al.*, 2016). The mission is complete when the MEDEVAC unit arrives back at its staging area. The MEDEVAC unit's status changes from busy to idle upon arrival, and the unit is available for another tasking at time T_8. The total response time is defined as $T_7 - T_2$, and

the total service time is defined as $T_8 - T_2$. The MEDEVAC unit service time is an important metric because it comprises the time between the initial unit tasking (i.e., the moment the unit's status is changed to busy) and returning to the staging area (i.e., the moment the unit's status is changed back to idle). For the purpose of this model, MEDEVAC unit response and service times are assumed to be exponentially distributed.

State transitions are Markovian and occur with two events. Either a MEDEVAC unit completes a service or a 9-line MEDEVAC request arrives. Let μ_{mz} denote the service rate of MEDEVAC unit $m \in \mathcal{M}$ when servicing a request in zone $z \in \mathcal{Z}$, and let λ_{zk} denote the arrival rate of requests originating in zone $z \in \mathcal{Z}$ of priority $k \in \mathcal{K}$. Also, let $\mathcal{B}(S_t) = \{m : m \in \mathcal{M}, M_{tm} \neq 0\}$ denote the set of busy MEDEVAC units when the system is in state S_t at epoch t. If the MEDEVAC system is in state S_t and action x_t is taken, the system will immediately transition to a post-decision state, denoted as S_t^x (Powell, 2011). The time the system remains in the post-decision state before transitioning to the next pre-decision state (i.e., sojourn time) follows an exponential distribution with parameter $\beta(S_t, x_t)$. Simple calculations reveal that

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

If $\mathcal{B}(S_t) = \varnothing$ and $x_t^d = \{0\}^{|\mathcal{I}(S_t)|}$ (i.e., all MEDEVAC units are idle and no MEDEVAC units are tasked at time t), 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 9-line MEDEVAC request. Otherwise, $\mathcal{B}(S_t) \neq \varnothing$ and/or $x_t^d = \{0,1\}^{|\mathcal{I}(S_t)|}$ (i.e., at least one MEDEVAC unit is busy at time t and/or a MEDEVAC unit is tasked at epoch t). In this case, $\beta(S_t, x_t)$ represents the sojourn time for the state-action pairs wherein the next decision epoch occurs after either the arrival of a new request or a busy

MEDEVAC unit completes a service.

The transition probabilities of this system are summarized in terms of a infinitesimal generator (i.e., rate matrix) as follows:

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

wherein

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

denotes the probability of transitioning to state S_{t+1} from state S_t after taking action x_t. The post-decision state variable $M_{tm}^x \in \{0\} \cup \mathcal{Z}$ denotes the status of MEDEVAC unit $m \in \mathcal{M}$ when decision x_t is made at epoch t. Note that $p(S_{t+1}|S_t, x_t) = 0$; the system will always transition to a new state at the end of a sojourn time.

To perform subsequent analysis on a continuous-time MDP, uniformization is applied to transform the model to an equivalent discrete-time MDP. To uniformize the system, the maximum rate of transition is calculated as follows:

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

wherein

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

One restriction of a continuous-time model is the absence of self-transitions (i.e., transitioning from a state to itself). Applying uniformization (Puterman, 1994) elim-

inates this restriction and 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 system earns rewards when a MEDEVAC unit is dispatched to a 9-line MEDEVAC request. Several factors impact the amount of reward gained at each decision epoch: the zone and priority level of the 9-line MEDEVAC request and the staging area of the servicing MEDEVAC unit. Let $c(S_t, x_t) = \psi_{mzk}$ denote the immediate expected reward if MEDEVAC unit $m \in \mathcal{M}$ is tasked to service a request in zone $z \in \mathcal{Z}$ of priority $k \in \mathcal{K}$. Given that the MEDEVAC system seeks to service urgent and priority 9-line MEDEVAC requests within 60 and 240 minutes from notification (Department of the Army, 2019), respectively, the expected immediate reward is expressed as

$$\psi_{mzk} = \begin{cases} \delta e^{\frac{-\zeta_{mz}}{60}}, & \text{if } k = 1 \\ e^{\frac{-\zeta_{mz}}{240}}, & \text{if } k = 2 \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

The first case (i.e., if $k = 1$) represents the reward gained from responding to a Priority I - Urgent 9-line MEDEVAC request. The second case (i.e., if $k = 2$) represents the reward gained from responding to a Priority II - Priority 9-line MEDEVAC request. The third case occurs when a MEDEVAC unit is not dispatched to service a request at decision epoch t, which results in a reward of 0. The expected response time when MEDEVAC unit $m \in \mathcal{M}$ is tasked to service a request in zone $z \in \mathcal{Z}$ is denoted as ζ_{mz}. The tradeoff parameter $\delta \geq 1$ is utilized to vary the urgent to priority immediate

expected reward ratio. To convert the continuous-time reward function to a discrete-time one, uniformization is applied via

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

wherein $\alpha > 0$ denotes the continuous-time discounting rate.

The objective of the MEDEVAC system is to determine an optimal dispatching order that maximizes the ETDR attained by the system. Let $X^\pi(S_t) : \mathcal{S} \to \mathcal{X}_{S_t}$, $S_t \in \mathcal{S}$ represent the decision function, based on policy π, that maps the state space to the action space. More specifically, $X^\pi(S_t)$ indicates the action (i.e., $X^\pi(S_t) = x_t$) to take when the system is in state S_t at epoch t given policy π. The optimal policy, denoted as π^*, is sought from the class of policies $\big(X^\pi(S_t)\big)_{\pi\in\Pi}$ to maximize the ETDR. This objective is written as

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

wherein $\gamma = \frac{\nu}{\nu+\alpha}$ is the uniformized discount factor. The Bellman equation below is solved to obtain the optimal policy:

$$V(S_t) = \max_{x_t\in\mathcal{X}(S_t)}\Big(\tilde{r}(S_t, x_t) + \gamma\mathbb{E}\big(V(S_{t+1})\big|S_t, x_t\big)\Big). \tag{4}$$

The policy iteration algorithm is implemented in MATLAB 2020a to solve Equation 4, thereby solving the MDP model to optimality; see Algorithm 1. This algorithm yields an optimal policy π^* by first selecting an arbitrary policy (e.g., the myopic policy) in Step 1. This policy is evaluated in Step 2 by computing the ETDR, which yields a value function $V(S_t)$. Next, an improved policy is selected in Step 3. The algorithm terminates in Step 4 when a policy improvement is not available (i.e., the policy converges) (Puterman, 1994).

Algorithm 1 Policy Iteration Algorithm

1: Initialize $n = 0$ and select an arbitrary decision rule $X^{\pi_0}(S_t)$, where $\pi_0 \in \Pi$.
2: (Policy Evaluation) For each $S_t \in \mathcal{S}$, solve $V^n(S_t)$ via Equation 4.
3: (Policy Improvement) For each $S_t \in \mathcal{S}$, choose $X^{\pi_{n+1}}(S_t)$ to satisfy

$$X^{\pi_{n+1}}(S_t) = \underset{x_t \in \mathcal{X}_{S_t}}{\operatorname{argmax}} \Big\{ \tilde{r}(S_t, x_t) + \gamma \mathbb{E}\big[V^n(S_{t+1}|S_t, x_t)\big] \Big\},$$

setting $X^{\pi_{n+1}}(S_t) = X^{\pi_n}(S_t)$ if possible.
4: If $\pi_{n+1} = \pi_n$, stop and set $\pi^* = \pi_n$. Otherwise, increment n by 1 and return to Step 2.

3.2 ADP Formulation

Classic dynamic programming algorithms (e.g., policy iteration) employ enumeration of the state space. Therefore, only relatively smaller problem instances are computationally tractable due to the curse of dimensionality, as shown in Equation 1. More specifically, solving Equation 4 via exact dynamic programming methods is often impractical. Therefore, ADP is utilized to overcome the large state space that realistic MEDEVAC scenarios possess. We use LSPE combined with API to approximate a solution and generate high-quality policies for the MEDEVAC dispatching problem.

Utilizing the post-decision state variable yields lower computational effort because it eliminates the expectation within the Bellman equation (i.e., Equation 4) and reduces the size of the state space (Ruszczynski, 2010). This allows us to take advantage of approximation techniques. The state transition function is broken into two steps. First, an action is selected and the system transitions to the post-decision state variable, denoted as

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

Then, the system transitions to the next pre-decision state upon a sample realization

of exogenous information, denoted as

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

wherein W_{t+1} is a sample realization of exogenous information. For the MEDEVAC dispatching problem, all post-decision states take the form $S_t^x = (M_t, (0,0))$, because after a decision has been made, any previous 9-line MEDEVAC request information is cleared from the state space (i.e., $\hat{R}_t = (0,0)$). Let $\mathcal{S}^x$ denote the post-decision state space. The size of the post-decision state space is calculated as follows: $|\mathcal{S}^x| = (|\mathcal{Z}|+1)^{|\mathcal{M}|}$. Clearly $|\mathcal{S}^x| < |\mathcal{S}|$, which reduces the computational burden. The next pre-decision state depends on a sample realization of either a request arrival or a MEDEVAC unit service completion.

We proceed by modifying Equation 4 to incorporate the post-decision state variable. The value of being in post-decision state S_t^x is denoted as $V^x(S_t^x)$. The relationship between $V(S_t)$ and $V^x(S_t^x)$ is given by

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

Therefore, the Bellman equation around the post-decision state is given by

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

Note that the use of the post-decision state allows the expectation and maximization operators to be swapped. This exchange provides computational advantages because it allows explicit approximation of the expectation, which is statistically easier than estimating the entire expected value function as a function of decision x_t (Ruszczynski, 2010).

Next, we introduce the basis function that aids in training the value function. The basis function is denoted as $\phi(S_t^x)$ and contains features regarding the status of all MEDEVAC units in post-decision state S_t^x. The features that are captured in the basis function are the status of each MEDEVAC unit and all two-factor interactions. The basis function is expressed as

$$\phi(S_t^x) = \big[1, M_t, (M_{ti}M_{tj})\big]^\top \quad \forall\, i, j \in \mathcal{M}$$

wherein 1 is a bias term, M_t is the status of all MEDEVAC units, and $(M_{ti}M_{tj})$ contains two-factor interactions. Similar to linear regression, we seek to find a parameter vector, denoted as θ, using a set of observations that are created from the basis function $\phi(S_t^x)$. The θ column vector is used to predict the value of new observations of the state space. We specify this value function approximation as

$$\bar{V}^x(S_t^x|\theta) = \theta^\top \phi(S_t^x).$$

For a given θ (i.e., a fixed policy), decisions are made utilizing the decision function

$$X^\pi(S_t|\theta) = \underset{x_t \in \mathcal{X}_{S_t}}{\operatorname{argmax}} \Big\{ \tilde{r}(S_t, x_t) + \gamma \mathbb{E}\big[\bar{V}^x(S_t^x|\theta)\big] \Big\}. \tag{6}$$

Given Equation 6, the approximate post-decision state value function is given by

$$\bar{V}^x(S_{t-1}^x|\theta) = \mathbb{E}\Big[\tilde{r}\big(S_t, X^\pi(S_t|\theta)\big) + \gamma \bar{V}^x(S_t^x|\theta)|S_{t-1}^x\Big]. \tag{7}$$

We utilize an API algorithmic strategy to attain high-quality MEDEVAC dispatching policies. API is an algorithmic approach that is similar in structure to policy iteration (i.e., Algorithm 1). Like policy iteration, API produces a sequence of policies and approximate value functions via iterations with two phases. Within

the policy evaluation loop (i.e., the inner loop), the algorithm evaluates a fixed policy and updates the approximate value function parameters based upon observed results. The updated approximate value function is utilized in the subsequent policy improvement loop (i.e., next outer loop iteration) when the next fixed policy is evaluated. Algorithm 2 depicts the API algorithm.

Algorithm 2 API-LSPE Algorithm

1: Initialize θ (linear model coefficients).
2: **for** $n = 1$ to N **do** (Policy Improvement Loop)
3:     **for** $j = 1$ to J **do** (Policy Evaluation Loop)
4:         Generate a random post-decision state, $S^x_{t-1,j}$.
5:         Record basis function evaluation, $\phi(S^x_{t-1,j})$.
6:         Determine the set of next possible pre-decision states $\bar{\mathcal{S}} \in \mathcal{S}$ utilizing Equation 5.
7:         For each pre-decision state $S_{t,i} \in \bar{\mathcal{S}}$, determine decision x_t utilizing Equation 6 and compute $\hat{v}_{j,i}$ utilizing Equation 8.
8:         Determine the estimated value of being in post-decision state $S^x_{t-1,j}$ via Equation 9.
9:     **end for**
10:     Update θ utilizing Equations 10 and 11.
11: **end for**
12: Return the approximate value function $\bar{V}^x(\cdot\ |\theta)$.

API-LSPE uses a series of J inner loops to evaluate a set policy, and a series of N outer loops to seek further policy improvement. Upon initialization (Step 1) and beginning the algorithm loops (Steps 2 and 3), we choose a post-decision state from a set of J randomly generated post-decision states (Step 4). These random post-decision states are generated via Latin hypercube sampling (LHS). This sampling method generates well-spaced, uniform random samples for Monte Carlo procedures. The post-decision state is used to record the basis function evaluation in Step 5.

In Step 6, the set of next possible pre-decision states $\bar{\mathcal{S}} \subseteq \mathcal{S}$ is determined by leveraging the state transition function $S^{M,W}(S^x_{t-1}, W_t)$. Note that the distribution governing W_t is described by the transition probability function in Section 3.1. In Step 7, for each pre-decision state $S_{t,i} \in \bar{\mathcal{S}}$, $i = 1, 2, ..., |\bar{\mathcal{S}}|$, we solve the approximate

optimality equation via

$$\hat{v}_{j,i} = \tilde{r}\big(S_{t,i}, X^{\pi}(S_{t,i}|\theta)\big) + \gamma \bar{V}^{x}(S_{t,i}^{x}|\theta). \tag{8}$$

This equation represents the estimated value of being in post-decision state $S_{t-1,j}^{x}$ given the system evolves to pre-decision state $S_{t,i}$.

In Step 8, the value of being in post-decision state $S_{t-1,j}^{x}$ is calculated. This value estimate is more accurate because the algorithm computes and records the estimated values of being in all possible pre-decision states that the system could evolve to from post-decision state $S_{t-1,j}^{x}$ (Robbins *et al.*, 2020). The transition probability function indicates the likelihood of the system evolving to each of the pre-decision states in $\bar{\mathcal{S}}$. Utilizing this information, $\hat{v}_j$ is computed as

$$\hat{v}_j = \sum_{i=1}^{|\bar{\mathcal{S}}|} \tilde{p}\big(S_{t,i}|S_{t-1,j}^{x}\big)\hat{v}_{j,i}. \tag{9}$$

This process is repeated J times.

The policy improvement takes place in Step 10, wherein the θ vector is refined. The matrix Φ_{t-1} contains rows of basis function evaluations of the sampled post-decision states. Also, $\hat{V}_t$ is a vector that contains the estimated values calculated in Step 8. These are further defined as follows:

$$\Phi_{t-1} \triangleq \begin{bmatrix} \phi(S_{t-1,1}^{x})^{\top} \\ \vdots \\ \phi(S_{t-1,J}^{x})^{\top} \end{bmatrix}, \hat{V}_t \triangleq \begin{bmatrix} \hat{v}_1 \\ \vdots \\ \hat{v}_J \end{bmatrix}.$$

Since we are sampling the state space to approximate the Bellman equation, smoothing is required and is incorporated in Step 10. Here, we introduce a step size rule $\kappa = 1/n^{\rho}$ wherein $n = 1, 2, ..., N$ indicates the outer loop iteration and ρ is a tunable

parameter that controls the step size. The normalizing equation (i.e., Equation 10) and smoothing rule (i.e., Equation 11) utilized to update θ are defined as

$$\hat{\theta} = \left[(\Phi_{t-1})^\top (\Phi_{t-1}) + \mathbb{I}\eta\right]^{-1} (\Phi_{t-1})^\top \hat{V}_t, \tag{10}$$

$$\theta \leftarrow (1 - \kappa_n)\theta + \kappa_n \hat{\theta}. \tag{11}$$

In Equation 10, we utilize regularization with parameter η to ensure we do not overfit the data collected in any single policy-evaluation iteration. This process is repeated N times. In Step 12, the algorithm returns the approximate value function $\bar{V}^x(\cdot\,|\theta)$.

The tuning of algorithmic parameters is essential to achieve high-quality results. The tuning parameters that are examined within a computational experiment in Chapter IV are N number of outer loops, J number of inner loops, the ρ step size rule, and η regularization.

IV. Testing, Results, & Analysis

This chapter presents a representative MEDEVAC scenario utilized to demonstrate the applicability of the model described in Chapter III, to examine the behavior of the optimal policy, and to examine the quality of the policies generated from our ADP solution approach. Sensitivity analysis is conducted to identify significant model parameters that affect the policies generated by the ADP solution approach. We also analyze how standby unit mission preparation time affects the optimal policy and best-performing ADP-generated policy. Moreover, the representative scenario is expanded to a large scale scenario that cannot be solved to optimality and is solved via an ADP solution approach. The thesis utilizes a dual Intel Xeon E5-2687v3 workstation with 64 GB of RAM and MATLAB's Parallel Computing Toolbox to conduct the computational experiments and perform subsequent analysis.

4.1 Representative Scenario

The US invaded Iraq in 2003 in response to Saddam Hussein's continuous hindrance of United Nations (UN) inspections and disobedience of UN sanctions. This marked the beginning of the Iraq War. Violence began to decline in 2007, and in 2011, the US formally withdrew from Iraq (Augustyn *et al.*, 2020).

This research considers a notional planning scenario similar to operations conducted during the Iraq War. The computational examples described in Keneally *et al.* (2016) and Jenkins (2017) are leveraged as inspiration for the representative scenario described in this section. This scenario assumes a MEDEVAC system with six demand zones (i.e., zones from which 9-line MEDEVAC requests originate), three primary MEDEVAC units, and one standby MEDEVAC unit, and is hereafter referred to as the 6×4 case, wherein MEDEVAC 4 represents the standby unit. The place-

ment of MEDEVAC units and MTFs are meant to represent the operations during the Iraq War.

According to casualty data retrieved from White (2020), over 60% of all war-related casualties in Iraq from 2003 to 2011 occurred in the provinces of Baghdad and Anbar. These numbers account for all war-related deaths as the result of a hostile attack. Further investigation reveals that a large majority of these casualties occurred in major cities along the Euphrates River. Although real casualty data will not be used for this scenario (due to data restrictions), these numbers provide a representative sample of the threat level within each province of Iraq.

With influence from White (2020), future 9-line MEDEVAC requests are modeled with a Monte Carlo simulation via a Poisson cluster process. Casualty cluster centers (CCCs) are selected based on cities or areas in which a large number of casualties occurred during the Iraq War. The distribution of 9-line MEDEVAC request locations from a CCC is generated from a uniform distribution with respect to the distance of the request to the CCC. It is important to note that this input data will influence the policy generated by an MDP or ADP solution approach, so relevant scenario data is essential to generate a meaningful result. Figure 2 depicts the zone tessellation scheme and the placement of bases, MTFs, MEDEVAC units, and CCCs that will be used to generate the simulation data described throughout this section. MEDEVAC 1 is stationed at a coalition base, which includes a helicopter landing zone (HLZ) but no medical facilities. Both MEDEVAC 2 and MEDEVAC 3 are stationed at an MTF equipped with an HLZ. Note that MEDEVAC 4 (i.e., the standby unit) may be co-located with any primary unit; the location of the standby unit will be determined by the analysis described in Section 4.2.

The outputs of the Monte Carlo simulation include the probability of a 9-line MEDEVAC request originating in zone $z \in \mathcal{Z}$, the expected response time of MEDE-

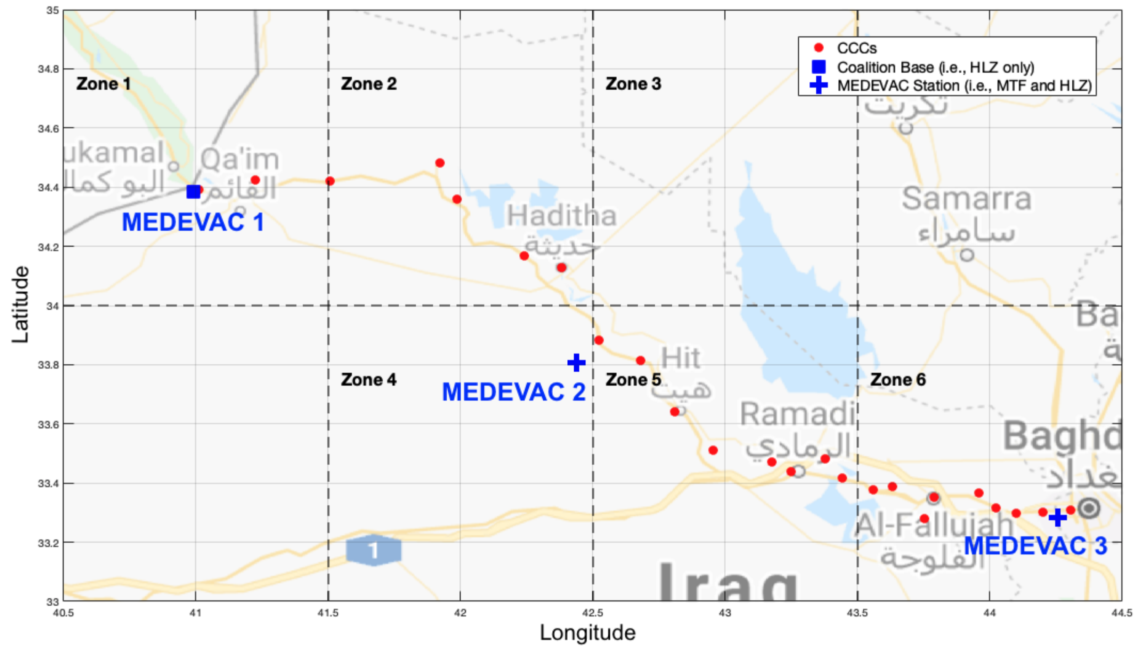

Figure 2. MEDEVAC locations, Zones, and CCCs in Iraq

VAC $m \in \mathcal{M}$ when servicing a request in zone $z \in \mathcal{Z}$, and the expected service time of MEDEVAC $m \in \mathcal{M}$ when servicing a request in zone $z \in \mathcal{Z}$. Although historical data suggests otherwise, an equal probability of high and low priority requests (i.e., urgent and priority requests) is assumed for the 6×4 case. That is, $p_{k_1} = p_{k_2} = 0.5$. Therefore, the probability of a request originating in zone $z \in \mathcal{Z}$ and is of priority $k \in \mathcal{K}$, p_{zk}, is calculated by multiplying the zone proportion and the priority level proportion. For instance, $p_{11} = p_{z_1} p_{k_1}$. Table 1 shows the request categorization proportions for the 6×4 case.

Table 1. 9-Line MEDEVAC Request Proportions by Zone-Priority Level

Zone z	Urgent	Priority	Total
Zone 1	0.05169	0.05169	0.10338
Zone 2	0.09449	0.09449	0.18899
Zone 3	0.00003	0.00003	0.00006
Zone 4	0.00599	0.00599	0.01199
Zone 5	0.15952	0.15952	0.31905
Zone 6	0.18826	0.18826	0.37653

The total response time comprises mission preparation time, travel time to the CCP, time to load the casualty onto the aircraft, travel time to the MTF, and time to unload the casualty at the MTF. Utilizing information from Bastian (2010), mission preparation time is set to 15 minutes, load time is set to 10 minutes, and unload time is set to five minutes. It is assumed that the standby unit has an additional 15 minute mission preparation period. All travel times are calculated using a flight speed of 156 knots and the distance between location coordinates. Casualties are simulated in random locations based on the predetermined CCCs, which induces variation in travel times and, therefore, variation in total response times. The resulting response times for MEDEVAC $m \in \mathcal{M}$ when servicing zone $z \in \mathcal{Z}$ are determined by averaging the simulated response times of the respective MEDEVAC-zone combination. The mean response times used as inputs to the MDP model are outlined in Table 2. Experimentation of three courses of action (COAs) will determine the best location for the standby MEDEVAC unit. Hence, response times are listed for three COAs, which correspond to standby unit information when co-located with MEDEVAC 1, 2, and 3, respectively.

Table 2. Expected Response Times (minutes)

	MEDEVAC, m			Standby Unit		
Zone z	1	2	3	COA 1	COA 2	COA 3
Zone 1	61.868	85.639	122.190	76.868	100.639	137.190
Zone 2	64.499	57.595	91.984	79.499	72.595	106.984
Zone 3	66.626	42.996	74.460	81.626	57.996	89.460
Zone 4	62.644	34.046	68.943	77.644	49.046	83.943
Zone 5	85.846	55.393	66.513	100.846	70.393	81.513
Zone 6	98.090	67.467	43.782	113.090	82.467	58.782

The total service time comprises response time and travel time back to the MEDEVAC's staging area. These travel times are calculated as described above. Note that some expected response times and expected service times are equivalent for certain MEDEVAC-zone combinations. This is because the staging area for these units is

co-located with an MTF, so the travel time back to the staging area is zero as long as that is the closest MTF to the casualty. The resulting service times for MEDEVAC $m \in \mathcal{M}$ when servicing zone $z \in \mathcal{Z}$ are determined by averaging the simulated service times of the respective MEDEVAC-zone combination. The mean service times that are used as inputs to the MDP model are outlined in Table 3. Standby unit service times are listed for three COAs as described above.

Table 3. Expected Service Times (minutes)

	MEDEVAC, m			Standby Unit		
Zone z	1	2	3	COA 1	COA 2	COA 3
Zone 1	92.583	85.639	159.140	107.583	100.639	174.140
Zone 2	95.214	57.595	128.940	110.214	72.595	143.940
Zone 3	97.341	42.996	111.410	112.341	57.996	126.410
Zone 4	93.360	34.046	105.900	108.360	49.046	120.900
Zone 5	126.140	65.013	93.845	141.140	80.013	108.845
Zone 6	165.610	104.420	43.782	180.610	119.420	58.782

Given the zone and priority level, the immediate expected reward for servicing a 9-line MEDEVAC request is calculated according to Equation 3. The 6×4 case utilizes $\delta = 10$, which rewards the servicing of urgent 9-line MEDEVAC requests more than priority 9-line MEDEVAC requests. Table 4 summarizes the computed immediate expected rewards, ψ_{mzk}.

The 6×4 case assumes a high operations tempo, indicated by a request arrival rate of $\lambda = \frac{1}{30}$. This indicates an average arrival rate of one request every 30 minutes. It is important to note that this arrival rate will influence the policy generated by an MDP or ADP model. Hence, operational planners should determine a reasonable request arrival rate prior to a planned combat operation for the proposed model to generate a meaningful result.

Table 4. **Immediate Expected Rewards**

		MEDEVAC, m			Standby Unit		
Zone z	Priority (k)	1	2	3	COA 1	COA 2	COA 3
Zone 1	Urgent (1)	3.5660	2.3995	1.3048	2.7772	1.8687	1.0162
	Priority (2)	0.7728	0.6999	0.6010	0.7259	0.6575	0.5646
Zone 2	Urgent (1)	3.4130	3.8292	2.1587	2.6581	2.9822	1.6812
	Priority (2)	0.7643	0.7866	0.6816	0.7180	0.7390	0.6403
Zone 3	Urgent (1)	3.2942	4.8841	2.8909	2.5655	3.8037	2.2515
	Priority (2)	0.7576	0.8360	0.7333	0.7117	0.7853	0.6888
Zone 4	Urgent (1)	3.5202	5.6698	3.1694	2.7415	4.4156	2.4683
	Priority (2)	0.7703	0.8677	0.7503	0.7236	0.8152	0.7049
Zone 5	Urgent (1)	2.3913	3.9724	3.3004	1.8623	3.0937	2.5703
	Priority (2)	0.6993	0.7939	0.7580	0.6569	0.7458	0.7120
Zone 6	Urgent (1)	1.9498	3.2483	4.8205	1.5185	2.5298	3.7542
	Priority (2)	0.6645	0.7549	0.8332	0.6242	0.7092	0.7828

4.2 Representative Scenario Results

In this section, we explore the results of utilizing the previously described MDP and ADP solution approaches to generate MEDEVAC dispatching policies for the 6×4 case. A list of parameters associated with the 6×4 case are outlined in Table 5. Utilizing these parameter settings and the zone-priority level proportions, expected response times, expected service times, and immediate expected rewards computed in the previous section, the optimal policy for the 6×4 case is computed via policy iteration (i.e., Algorithm 1). Moreover, ADP policies are generated via API-LSPE (i.e., Algorithm 2). Applying Equation 1 reveals that the size of the state space for the 6×4 case is 31,213. This result indicates that even for this relatively small scenario, the size of the state space is quite large and will increase drastically if elements are added (i.e., additional zones, MEDEVAC units, or priority levels).

For comparison purposes, the myopic policy is considered the baseline policy. The myopic policy suggests that the closest idle MEDEVAC unit to the casualty be tasked to respond, regardless of the request's zone or priority level. If the co-located primary and standby units are both idle, the myopic policy will always task the

Table 5. 6×4 **Case Parameter Settings**

Parameter	Description	Setting
λ	9-line MEDEVAC request arrival rate	$\frac{1}{30}$
$\|\mathcal{M}\|$	Number of MEDEVACs	4
$\|\mathcal{Z}\|$	Number of zones	6
$\|\mathcal{K}\|$	Number of priority levels	2
γ	Uniformized discount factor	0.99
δ	Weight for urgent requests	10

primary unit. Therefore, the standby unit will only be tasked if the primary unit is busy and if it is the closest unit to the casualty. The myopic policy also does not include admission control. This means that if at least one MEDEVAC unit is idle when a request is submitted to the system, the request must be serviced. The optimal policy's dispatching order is compared against the best performing ADP-generated policy and myopic policy to obtain insights as to where similarities and differences exist. Moreover, the optimality gap is computed to demonstrate whether a myopic policy is appropriate for the given 6×4 case.

4.2.1 MDP Results

Table 6 is used to compare the optimal and myopic policies to determine which primary MEDEVAC unit the standby unit should be co-located with to maximize ETDR. Note that the response times, service times, and immediate expected rewards for COA 1, COA 2, and COA 3 are used for this analysis, as described in Section 4.1. The ETDR for the optimal policy and myopic policy when the system is in an empty state $S_0 = ((0,0,0,0),(0,0))$ (i.e., all MEDEVAC units are idle and there are no 9-line MEDEVAC requests in the system) are displayed in Table 6, along with the optimality gap associated with the myopic policy. The results indicate that the best location for the standby unit is with MEDEVAC 1; the optimality gap is the largest, and the ETDR for both policies is the largest. The myopic policy has an optimality

gap of 10.00%, and the optimal policy yields an ETDR of 38.6669. Whereas this optimality gap may not seem large, the results indicate the optimal policy saves more lives.

Table 6. Comparison of ETDR & Optimality Gap

COA	Policy, π	$V^{\pi}(S_0)$	Optimality Gap
1	Optimal	38.6669	N/A
	Myopic	35.1488	10.00%
2	Optimal	34.0953	N/A
	Myopic	32.0347	6.43%
3	Optimal	36.2634	N/A
	Myopic	33.9255	6.89%

We also compare the three COAs by calculating rejection rates and standby unit tasking rates of their respective optimal policy. Rejection rates are calculated by determining the percentage of states in which the policy is to reject an incoming request. A lower rejection rate indicates a more efficient MEDEVAC system. Similarly, standby unit tasking rates are calculated by determining the percentage of states in which the policy is to task the standby unit. A higher standby unit tasking rate is desired to reserve primary MEDEVAC units for urgent requests expected to arrive in the near future. Table 7 compares these rates for the optimal policies associated with COA 1, COA 2, and COA 3. We see that COA 3 has the highest rejection rate and the lowest standby unit tasking rate, which makes this the least desirable COA. COAs 1 and 2 are similar; however, COA 1 utilizes the standby unit moreso than COA 2. Furthermore, Figure 3 shows the rejection rates for 9-line MEDEVAC requests originating in zone $z \in \mathcal{Z}$ for each COA. We see that COA 3 yields the highest rejection rates for all zones except Zone 6. We also see the trade off in rejection rates by zone between COA 1 and COA 2. The results shown in Table 6, Table 7, and Figure 3 confirm that the optimal location for the standby unit is with MEDEVAC 1. Therefore, all subsequent analysis assumes that the standby unit is co-located with

MEDEVAC 1.

Table 7. COA Comparison - Rejection Rates and Standby Tasking Rates

COA	Rejection Rate	Standby Tasking Rate
1	69.61%	9.36%
2	69.53%	8.98%
3	73.40%	5.15%

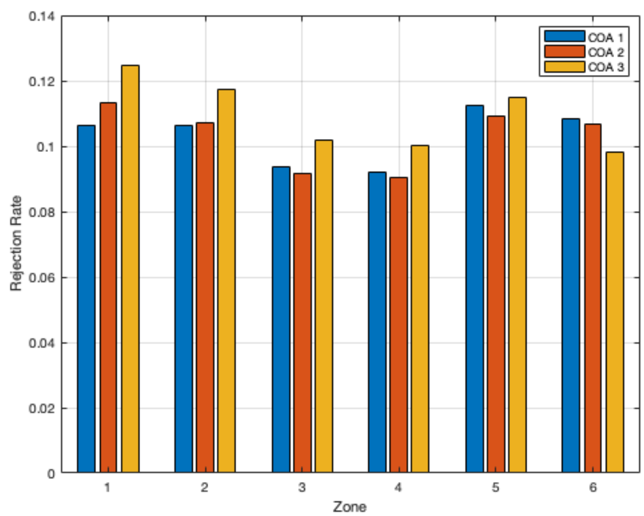

Figure 3. COA Comparison - Rejection Rates by Zone

4.2.2 ADP Results

An experimental design is constructed to explore different parameter settings for the implementation of Algorithm 2 to solve the 6×4 case. The tuning parameters within the computational experiment for API-LSPE are the number of outer loops (N), the number of inner loops (J), the step size rule (ρ), and the regularization (η). Table 8 shows the factor levels for each parameter in the computational experiment. Note that the number of inner loops J is a function of the size of the post-decision state space. In other words, $J \in \left\{\lceil 0.2 \times |\mathcal{S}^x| \rceil, \lceil 0.4 \times |\mathcal{S}^x| \rceil, \lceil 0.6 \times |\mathcal{S}^x| \rceil, \lceil 0.8 \times |\mathcal{S}^x| \rceil\right\}$, wherein $|\mathcal{S}^x| = 2401$. A full factorial design of these parameter settings yields 256

Table 8. Computational Experiment Parameter Levels

Parameter	Levels
N	20, 30, 40, 50
J	480, 960, 1441, 1921
ρ	0.3, 0.4, 0.6, 0.9
η	0, 0.001, 0.01, 0.1

different factor combinations, and five replications of each run is performed. The root mean squared error (RMSE) of each factor combination run is calculated over all $s \in \mathcal{S}$ to compare the value of each ADP policy to the value of the optimal policy. A smaller RMSE indicates that the ADP-generated value function is closer to the optimal value function. The RMSE of each replication is recorded to calculate the mean RMSE of the factor combination. The minimum RMSE observed for each factor combination and the variance of the RMSE of the five replications are also reported. The top 20 factor combinations in terms of mean RMSE are listed in Table 9. We see that the best ρ value is 0.9 because all 20 runs listed in the table have this ρ value. The algorithm also performs best when $N = 30$ and when J is smaller. However, the range of the mean RMSE over the top 20 parameter combinations is only 0.0144, which indicates this algorithm is fairly robust to parameter settings. The best parameter combination for API-LSPE when solving the 6×4 case is $N = 30, J = 480, \rho = 0.9, \eta = 0.01$. These parameter settings are utilized for all subsequent analysis. Note that the value function of the ADP policy is solved for exactly using policy evaluation.

The ETDR for the optimal policy, best performing ADP-generated policy, and myopic policy, denoted as π^*, π^{LSPE}, and π^{myopic}, respectively, when the system is in an empty state $S_0 = ((0,0,0,0),(0,0))$ is displayed in Table 10, along with the optimality gap associated with the ADP policy and myopic policy. These results, along with the RMSE displayed in Table 9, show that the ADP policy is an improvement over the myopic policy and close to the optimal policy. In fact, the ADP policy is a 7.18% improvement over the myopic policy.

Table 9. API-LSPE Computational Experiment Results

Run	ρ	η	J	N	Min RMSE	Mean RMSE	Var RMSE
1	0.9	0.01	480	30	1.0779	1.0959	0.0003
2	0.9	0.001	480	30	1.0779	1.0959	0.0003
3	0.9	0.1	480	30	1.0781	1.0965	0.0003
4	0.9	0.001	960	30	1.0880	1.1034	0.0003
5	0.9	0.01	960	30	1.0880	1.1034	0.0003
6	0.9	0.1	960	30	1.0888	1.1036	0.0003
7	0.9	0.001	1441	30	1.1011	1.1063	0.0001
8	0.9	0.01	1441	30	1.1011	1.1063	0.0001
9	0.9	0.1	1441	30	1.1015	1.1066	0.0001
10	0.9	0.1	1921	40	1.0756	1.1066	0.0005
11	0.9	0.001	1921	40	1.0755	1.1068	0.0005
12	0.9	0.01	1921	40	1.0755	1.1068	0.0005
13	0.9	0.1	480	50	1.0700	1.1086	0.0006
14	0.9	0.01	480	50	1.0704	1.1093	0.0005
15	0.9	0.001	480	40	1.0717	1.1094	0.0007
16	0.9	0.01	480	40	1.0717	1.1094	0.0007
17	0.9	0.1	480	40	1.0734	1.1101	0.0007
18	0.9	0.1	1441	40	1.0813	1.1102	0.0003
19	0.9	0.01	1441	40	1.0813	1.1103	0.0003
20	0.9	0.001	1441	40	1.0813	1.1103	0.0003

Table 10. Comparison of ETDR & Optimality Gap

Policy, π	$V^{\pi}(S_0)$	Optimality Gap
π^*	38.6669	N/A
π^{LSPE}	37.6790	2.62%
π^{myopic}	35.1488	10.00%

4.2.3 Policy Comparison

The workload of each MEDEVAC unit is an interesting performance measure, which reveals differences between policies in terms of effectiveness. This performance measure is constructed by calculating the steady state probabilities and adding them together for the states of interest (i.e., states wherein MEDEVAC unit $m \in \mathcal{M}$ is not idle). Figure 4 shows the long-run busy probabilities for all four MEDEVAC units (i.e., the fraction of time each MEDEVAC unit is busy servicing a request) for the optimal policy, best performing ADP-generated policy, and myopic policy. Figure 4

reveals that with the exception of the standby unit (i.e., MEDEVAC 4), the optimal policy has lower busy probabilities than the myopic policy. Also, with the exception of MEDEVAC 1, the ADP policy has lower busy probabilities than the myopic policy but higher busy probabilities than the optimal policy. However, both the optimal policy and ADP policy utilize the standby unit more than the myopic policy to lessen the workload of the primary units. This means that the optimal policy utilizes each MEDEVAC in the most efficient manner, followed by the ADP policy. This result is intuitive because the optimal and ADP policies have control over admission and dispatching rules whereas the myopic policy does not.

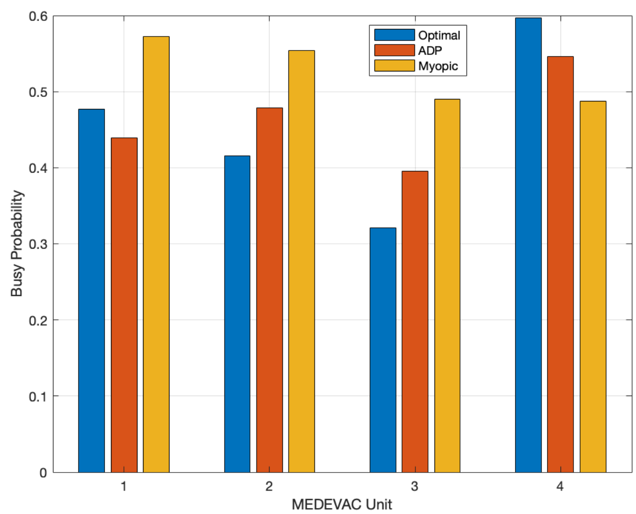

Figure 4. MEDEVAC Unit Busy Rate Comparison

The rate at which the policy rejects incoming 9-line MEDEVAC requests is another interesting performance measure used to compare policies. This performance measure is calculated as described in Section 4.2.1. Figure 5 shows the rejection rates for 9-line MEDEVAC requests originating in zone $z \in \mathcal{Z}$ for the optimal policy, best performing ADP policy, and myopic policy. Recall that the myopic policy only rejects a request when all MEDEVAC units are busy. However, the optimal policy and ADP policy may reject requests anytime to reserve MEDEVAC units for higher priority requests.

These results indicate that the optimal policy has the highest rejection rate overall, followed by the ADP policy, and the myopic policy has the lowest rejection rate. This result is expected because the optimal policy and ADP policy utilize admission control whereas the myopic policy does not. The optimal policy and ADP policy reject more requests from Zone 5 than any other zone. This is due to the high probability that a request originates in Zone 5, but no MEDEVAC unit is stationed in Zone 5. The optimal policy rejects more requests than the ADP policy when the request originates in Zones 1, 2, 5, and 6. However, the ADP policy rejects more requests from Zones 3 and 4 than the optimal policy.

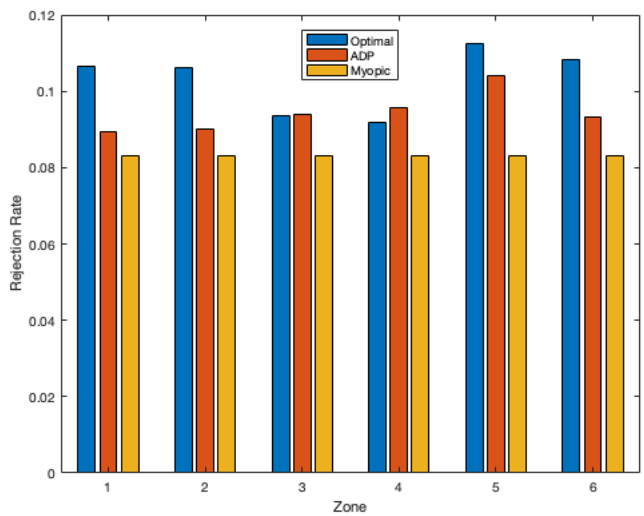

Figure 5. Rejection Rates by Zone

Table 11 outlines six scenarios wherein differences exist between the optimal decision rule, ADP decision rule, and myopic decision rule, denoted as $X^{\pi^*}(S_t)$, $X^{\pi^{\text{LSPE}}}(S_t|\theta)$, and $X^{\pi^{\text{myopic}}}(S_t)$, respectively. Recall that an action includes an admission decision, wherein a 1 indicates admission and 0 indicates rejection, and a dispatching decision (i.e., which unit to task if the request is admitted into the system). Differences in the policies stem from the utilization of the standby unit, admission control, and tasking a unit farther from the casualty. For example, in Scenario 1 the

co-located primary and standby MEDEVAC units (i.e., MEDEVAC 1 and MEDEVAC 4, respectively) are both idle, and a 9-line MEDEVAC request arrives that originates in Zone 6 and is Priority II (i.e., priority). The myopic policy suggests to task MEDEVAC 1 because it is the primary unit, but the optimal policy and ADP policy task the standby unit. In Scenario 2, MEDEVAC 1, MEDEVAC 2, and the standby unit are idle, and a 9-line MEDEVAC request arrives that originates in Zone 5 and is Priority II. The myopic policy suggests to task MEDEVAC 2 because it is the closest unit to the casualty, but the optimal policy and ADP policy task the standby unit. In cases like this, it is beneficial to reserve the primary units for Priority I (i.e., urgent) requests that are expected to occur in the near future. Examination of all three policies for the entire state space reveals that the differences in policies regarding standby unit operations can be generalized: The standby unit is only tasked to Priority II requests but is often tasked in place of any primary unit.

Table 11. Policy Differences

Scenario	$S_t = (M_t, \hat{R}_t)$	$X^{\pi^*}(S_t)$	$X^{\pi^{\text{LSPE}}}(S_t\|\theta)$	$X^{\pi^{\text{myopic}}}(S_t)$
1	((0,2,3,0),(6,2))	Dispatch MEDEVAC 4	Dispatch MEDEVAC 4	Dispatch MEDEVAC 1
2	((0,0,3,0),(5,2))	Dispatch MEDEVAC 4	Dispatch MEDEVAC 4	Dispatch MEDEVAC 2
3	((0,1,2,4),(5,2))	Reject	Reject	Dispatch MEDEVAC 1
4	((3,2,0,4),(1,2))	Reject	Reject	Dispatch MEDEVAC 3
5	((0,4,0,1),(5,2))	Dispatch MEDEVAC 1	Dispatch MEDEVAC 1	Dispatch MEDEVAC 3
6	((0,0,2,5),(2,2))	Dispatch MEDEVAC 1	Dispatch MEDEVAC 2	Dispatch MEDEVAC 2

The differences in policies are examined in terms of admission control. In Scenario 3, MEDEVAC 1 is the only idle unit, and a 9-line MEDEVAC request arrives that originates in Zone 5 and is Priority II. The myopic policy suggests to task MEDEVAC 1 because it is the only idle unit, but the optimal policy and ADP policy reject the request. Similarly, in Scenario 4, MEDEVAC 3 is the only idle unit, and a 9-line MEDEVAC request arrives that originates in Zone 1 and is Priority II. The myopic policy suggests to task MEDEVAC 3 because it is the only idle unit, but the optimal policy and ADP policy reject the request. In cases like this, it is beneficial to reject

Priority II requests to reserve the idle units for Priority I requests that are expected to occur in the near future rather than task a unit to respond to a request that requires a high service time. In general, admission control is used to reject Priority II requests when the service time of the idle units is too high or if only one unit is idle. Admission control is also used by the optimal policy to reject Priority I requests when the request originates in Zone 1 or Zone 2 and the only idle MEDEVAC unit is MEDEVAC 3, whereas the ADP policy does not reject any Priority I requests.

Lastly, the optimal policy will sometimes task a unit that is farther from the casualty to reserve another MEDEVAC unit for an anticipated request. In Scenario 5, MEDEVAC 1 and MEDEVAC 3 are idle, and a 9-line MEDEVAC request arrives that originates in Zone 5 and is Priority II. The myopic policy suggests to task MEDEVAC 3 because it is the closest unit to the casualty, but the optimal policy and ADP policy task MEDEVAC 1. In Scenario 6, MEDEVAC 1 and MEDEVAC 2 are idle, and a 9-line MEDEVAC request arrives that originates in Zone 2 and is Priority II. The myopic policy suggests to task MEDEVAC 2 because it is the closest unit to the casualty, but the optimal policy tasks MEDEVAC 1. Note that the ADP policy is different from the optimal policy in this scenario; it suggests to task MEDEVAC 2. In cases like this, the optimal policy tasks units in a way that bears in mind what is expected to occur in the future. In these two scenarios, MEDEVAC 2 and MEDEVAC 3 are reserved for the high rate of requests that originate in Zones 5 and 6.

4.3 Excursion - Standby Mission Preparation

The section considers the impact the standby unit mission preparation time has on the optimal policy and the best performing ADP policy when the MEDEVAC system is in an empty state. The same parameter settings from the 6×4 case are utilized for the standby mission preparation excursion; see Table 5. We also assume

that the standby unit is co-located with MEDEVAC 1 because this is the optimal location as indicated by the results in Table 6. However, we modify the standby unit mission preparation time from an additional 15 minutes by altering its response and service times. Standby unit additional mission preparation times from 5 minutes to 45 minutes in 5 minute increments were explored (i.e., total mission preparation times of 20 to 60 minutes), and the results are displayed in Figures 6 and 7.

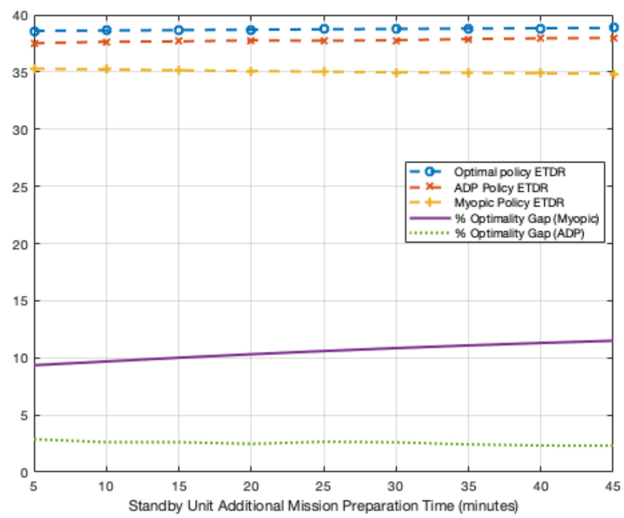

Figure 6. Comparison of ETDR & Optimality Gap

Figure 6 shows the ETDR for the optimal policy, ADP policy, and myopic policy when the system is in an empty state $S_0 = ((0,0,0,0),(0,0))$ for each standby unit mission preparation time. This graph shows that the myopic policy optimality gap increases as the standby unit mission preparation time increases. This is because the myopic policy ETDR decreases while the optimal policy ETDR increases as standby unit mission preparation time increases. The myopic policy ETDR trend is intuitive because while the myopic decision rule does not change, the reward associated with tasking the standby unit decreases, which yields a lower ETDR. However, it is interesting that the optimal policy ETDR actually increases by 0.26 when the standby unit additional mission preparation time increases from 5 to 45 minutes (i.e., a total

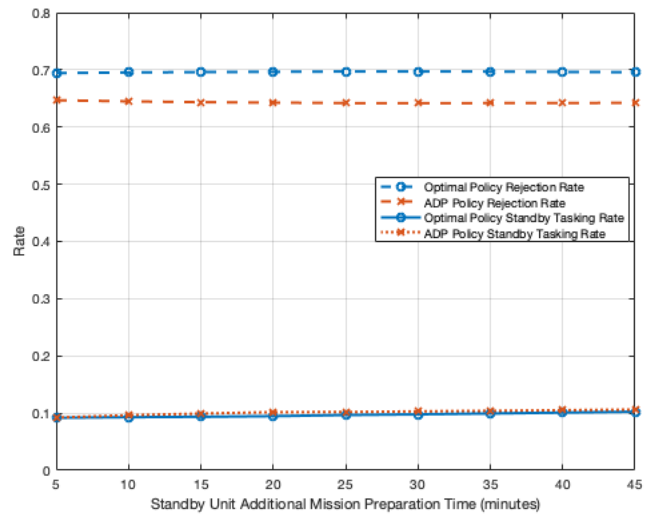

Figure 7. Rejection Rate and Standby Tasking Rate Comparison

mission preparation time from 20 to 60 minutes). This result shows the superiority of the optimal policy over the myopic policy. Furthermore, the ADP policy optimality gap decreases as the standby unit mission preparation time increases. This result pairs well with the results depicted in Figure 7.

Figure 7 shows the optimal policy and ADP policy rejection rates and standby unit tasking rates for standby unit additional mission preparation times of 5 minutes to 45 minutes. These rates are calculated as described in Section 4.2.1. This result indicates that as the standby unit mission preparation time increases, the rejection rate and standby unit tasking rate of the optimal policy and ADP policy slightly increase but remain consistent. Although the ADP policy has a lower rejection rate than the optimal policy, the difference between the two rates remains consistent as standby unit mission preparation time increases. Moreover, the two policies have nearly the same standby unit tasking rate for each standby unit mission preparation time. This shows that the optimal policy and ADP policy are fairly robust to standby unit mission preparation times.

4.4 Excursion - 38×4 case

This section expands the 6×4 case by altering the zone tessellation scheme to include 38 zones. This extended case is hereafter referred to as the 38×4 case. The location of all four MEDEVAC units remains the same as the 6×4 case, wherein the standby unit is co-located with MEDEVAC 1. Figure 8 depicts the zone tessellation scheme and the placement of bases, MTFs, MEDEVAC units, and CCCs for the 38×4 case. The Monte Carlo simulation described in Section 4.1 is used to calculate the probability of a 9-line MEDEVAC request originating in zone $z \in \mathcal{Z}$, the expected response time of MEDEVAC $m \in \mathcal{M}$ when servicing a request in $z \in \mathcal{Z}$, and the expected service time of MEDEVAC $m \in \mathcal{M}$ when servicing a request in $z \in \mathcal{Z}$.

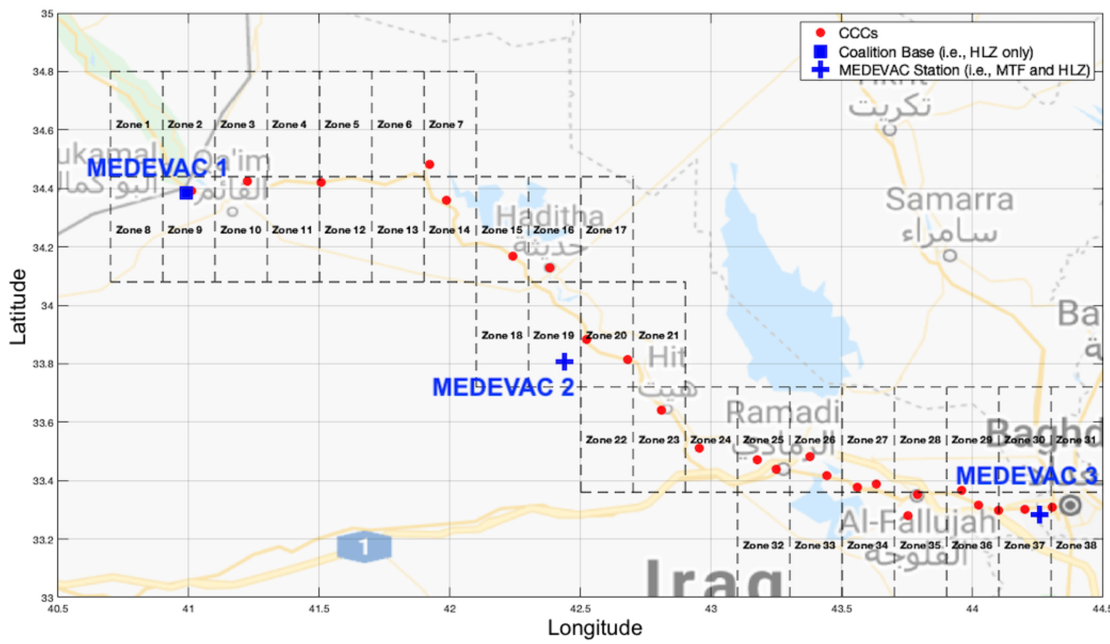

Figure 8. MEDEVAC locations, Zones, and CCCs in Iraq

A list of parameters associated with the 38×4 case are outlined in Table 12. Applying Equation 1 reveals that the size of the state space for the 38×4 case is 178,134,957. This result indicates that this scenario is too large to be solved to optimality. Hence, the 38×4 case is solved approximately via API-LSPE (i.e.,

Algorithm 2) utilizing the information in Table 12 and the outputs of the Monte Carlo simulation.

Table 12. 38×4 **Case Parameter Settings**

Parameter	Description	Setting
λ	9-line MEDEVAC request arrival rate	$\frac{1}{30}$
$\lvert\mathcal{M}\rvert$	Number of MEDEVACs	4
$\lvert\mathcal{Z}\rvert$	Number of zones	38
$\lvert\mathcal{K}\rvert$	Number of priority levels	2
γ	Uniformized discount factor	0.99
δ	Weight for urgent requests	10
p_{k_1}	Proportion of urgent requests	0.5
p_{k_2}	Proportion of priority requests	0.5

We utilize the solution approach described in Algorithm 2 along with the best parameter settings found in Section 4.2.2 to generate a policy for the 38×4 case. Unfortunately, this approach does not scale well to a larger problem size. A small computational experiment is conducted in an attempt to tune the algorithmic parameters and generate a high-quality policy. We explore a larger number of policy evaluation phases (i.e., J), a larger number of policy improvement phases (i.e., N), and a different step size rule (i.e., ρ) by running single algorithmic instances and changing one parameter setting each time. Since repeating the computational experiment described in Section 4.2.2 for this large-scale problem would take approximately 370 days to run on a single computer, we opted to explore a slightly different solution approach.

Recall that Algorithm 2 generates J post-decision states via LHS. Although this sampling method is sure to explore the entire post-decision state space in a uniform manner, it is possible to randomly sample states that the system is unlikely to visit. This means the algorithm may generate a policy based on states that are rarely encountered. We replace the LHS scheme with a trajectory-following scheme. Before entering the policy evaluation loop (i.e., the inner loop), we set the post-decision state

to an empty state $S_0 = ((0,0,0,0),(0,0))$. Upon completing Step 8, we determine the next pre-decision state that is most likely to occur based on arrival rates and service times. Then, we determine the set of feasible actions corresponding to the pre-decision state and choose one based on a uniform random sample. The randomly chosen action determines what post-decision state is used in Steps 5 – 8 of the next inner loop. These modifications ensure that the algorithm utilizes states the system is likely to visit while exploring the entire action space. The algorithmic parameters used to solve the 38×4 case via this modified solution approach are $N = 50, J = 50000, \rho = 0.9, \eta = 0.01$.

The policy generated via the aforementioned solution approach, denoted as π^{ADP}, and the myopic policy are evaluated utilizing a simulation-based approach wherein we apply common random numbers to reduce variance. Table 13 provides a summary of these results. The reported values are the average value over 200 simulations wherein the system starts in an empty state $S_0 = ((0,0,0,0),(0,0))$. The ADP policy attains a 3.61 ± 1.97 percent improvement over the myopic policy at the 95% confidence level. Moreover, a 95% confidence interval around the difference in the means (i.e., $34.5746 - 33.3713$) yields 1.20 ± 0.90. Since zero is not within the bounds of this interval, we conclude that the average value of the ADP policy is statistically greater than the average value of the myopic policy. These results suggest that our ADP solution approach obtains a high-quality policy. With the proper resources to conduct a full computational experiment to tune algorithmic parameters, we are confident that an even better policy can be generated.

Table 13. Comparison of ETDR & Percent Improvement

Policy, π	Avg $V^{\pi}(S_0)$	% Imp. over myopic
π^{ADP}	34.5746	3.61 ± 1.97
π^{myopic}	33.3713	N/A

Although we cannot calculate the long-run busy probabilities due to the size of the state space, we can compare the ADP policy and myopic policy in terms of rejection

rates and standby unit tasking rates. These rates are calculated as described in Section 4.2.1 and are outlined in Table 14. These results show that the ADP policy has slightly higher rejection and standby tasking rates. More specifically, the ADP policy rejects more Priority II requests than the myopic policy and tasks the standby unit to respond to more Priority II requests than the myopic policy. Although not as drastic, this trend is consistent with that of the 6×4 case. This shows that the ADP policy utilizes the MEDEVAC units and admission control more efficiently than the myopic policy.

Table 14. Rejection Rates and Standby Tasking Rates

Policy π	Priority (k)	Rejection Rate	Standby Tasking Rate
π^{ADP}	Urgent (1)	44.48%	1.19%
	Priority (2)	44.53%	1.25%
π^{myopic}	Urgent (1)	44.48%	1.20%
	Priority (2)	44.48%	1.20%

Table 15 outlines six scenarios wherein differences exist between the ADP decision rule and myopic decision rule. Differences in the policies stem from the utilization of the standby unit, admission control, and tasking a unit farther from the casualty. For example, in Scenario 1 the co-located primary and standby MEDEVAC units (i.e., MEDEVAC 1 and MEDEVAC 4, respectively) are both idle, and a 9-line MEDEVAC request arrives that originates in Zone 11 and is Priority II. The myopic policy suggests to task MEDEVAC 1 because it is the primary unit, but the ADP policy tasks the standby unit. In Scenario 2, MEDEVAC 3 and the standby unit are idle, and a 9-line MEDEVAC request arrives that originates in Zone 26 and is Priority II. The myopic policy suggests to task MEDEVAC 3 because it is the closest unit to the casualty, but the ADP policy tasks the standby unit. In cases like this, it is beneficial to reserve the primary units for Priority I requests that are expected to occur in the near future. Examination of the policies for the entire state space reveals that the differences in policies regarding standby unit operations can be generalized: The standby unit is

only tasked to Priority II requests but is often tasked in place of any primary unit.

Table 15. Policy Differences

Scenario	$S_t = (M_t, \hat{R}_t)$	$X^{\pi^{\text{ADP}}}(S_t\|\theta)$	$X^{\pi^{\text{myopic}}}(S_t)$
1	((0,1,7,0),(11,2))	Dispatch MEDEVAC 4	Dispatch MEDEVAC 1
2	((7,9,0,0),(26,2))	Dispatch MEDEVAC 4	Dispatch MEDEVAC 3
3	((19,38,0,17),(18,2))	Reject	Dispatch MEDEVAC 3
4	((0,36,20,37),(38,2))	Reject	Dispatch MEDEVAC 1
5	((0,0,0,4),(26,2))	Dispatch MEDEVAC 1	Dispatch MEDEVAC 3
6	((0,5,0,14),(21,2))	Dispatch MEDEVAC 1	Dispatch MEDEVAC 3

The differences in policies are also be examined in terms of admission control. In Scenario 3, MEDEVAC 3 is idle, and a 9-line MEDEVAC request arrives that originates in Zone 18 and is Priority II. The myopic policy suggests to task MEDEVAC 3 because it is the only idle unit, but the ADP policy rejects the request. Similarly, in Scenario 4, MEDEVAC 1 is idle, and a 9-line MEDEVAC request arrives that originates in Zone 38 and is Priority II. The myopic policy suggests to task MEDEVAC 1 because it is the only idle unit, but the ADP policy rejects the request. In cases like this, it is beneficial to reject certain requests to reserve the idle units for Priority I requests that are expected to occur in the near future rather than task a unit to respond to a request that requires a high service time. In general, admission control is used to reject Priority II requests if only one unit is idle.

Lastly, the ADP policy will sometimes task a unit that is farther from the casualty to reserve another MEDEVAC unit for an anticipated request. In Scenario 5, MEDEVAC 1, MEDEVAC 2, and MEDEVAC 3 are idle, and a 9-line MEDEVAC request arrives that originates in Zone 26 and is Priority II. The myopic policy suggests to task MEDEVAC 3 because it is the closest unit to the casualty, but the ADP policy tasks MEDEVAC 1. In Scenario 6, MEDEVAC 1 and MEDEVAC 3 are idle, and a 9-line MEDEVAC request arrives that originates in Zone 21 and is Priority II. The myopic policy suggests to task MEDEVAC 3 because it is the closest unit to the

casualty, but the ADP policy tasks MEDEVAC 1. In cases like this, the ADP policy tasks units in a way that bears in mind what is expected to occur in the future. In these two scenarios, MEDEVAC 3 is reserved for the high rate of requests that are expected to arrive in the southwestern area of interest.

V. Conclusions & Recommendations

This thesis examines the medical evacuation (MEDEVAC) dispatching problem. Specifically, we focus on decisions concerning which MEDEVAC unit to task to an incoming 9-line MEDEVAC request and where to station a single standby MEDEVAC unit. A standby MEDEVAC unit is available to respond to evacuation requests but might do so at a slower rate than a primary unit. The objective of this research is to maximize MEDEVAC system performance by determining a dispatching rule that maximizes battlefield casualty survivability rates. An infinite horizon, continuous-time Markov decision process (MDP) model is formulated to examine this problem. The MDP model incorporates admission control, which allows any request to be rejected by the dispatching authority and handled by an outside organization. The reported dispatching policy is based on the location and status of MEDEVAC units as well as the location and priority level of an incoming 9-line MEDEVAC request. To solve the MDP model, we apply policy iteration as well as an approximate dynamic programming (ADP) technique: a least squares policy evaluation (LSPE) value function approximation scheme within an approximate policy iteration (API) algorithmic framework. A computational example is applied to a synthetically generated scenario in Iraq, which represents combat operations during the Iraq War. The optimal policy and the ADP-generated policies are compared to a myopic (i.e., baseline) policy of dispatching the closest available MEDEVAC unit to the casualty. A computational excursion reveals how the standby unit mission preparation time affects the optimal policy, best-performing ADP-generated policy, and myopic policy. Moreover, the representative scenario is expanded to inspect the efficacy of the ADP solution approach when the problem is too large to solve to optimality.

A representative scenario set in Iraq is generated to explore the differences between the optimal, ADP, and myopic policies. This scenario assumes a MEDEVAC system

with six demand zones (i.e., zones from which 9-line MEDEVAC requests originate), three primary MEDEVAC units, and one standby MEDEVAC unit. The proposed MDP model is solved three times to determine the best location for the standby unit. The three solutions represent the MEDEVAC system performance when the standby unit is co-located with MEDEVAC Unit 1, 2, or 3. The best course of action (COA) is determined by examining the expected total discounted reward (ETDR) of each solution. The COA with the best ETDR is COA 1 (i.e., the standby unit is co-located with MEDEVAC 1). Further investigation of the optimal and myopic policy reveals that dispatching MEDEVAC units considering the entire MEDEVAC system state (i.e., the status of each MEDEVAC unit and the location and priority level of the incoming request) increases the casualty survivability rate. This scenario is also solved approximately utilizing API-LSPE. A computational experiment is conducted to explore different algorithmic parameter settings, and the best parameter combination is determined by analyzing root mean squared error (RMSE). More specifically, the parameter settings that yield a value function closest to the optimal value function are used for all subsequent analysis. The optimality gap associated with the best performing ADP-generated policy is 2.62%. Moreover, this policy is a 7.18% improvement over the myopic policy with respect to a life saving performance metric.

Differences in the optimal, ADP, and myopic policies stem from the utilization of the standby unit, admission control, and tasking a unit farther from the casualty. The optimal and ADP policies utilize the standby unit more than the myopic policy to lessen the workload of the primary units. Similarly, the optimal and ADP policies also reject more requests than the myopic policy to reserve idle units for high priority (i.e., urgent) requests expected to occur in the near future. Lastly, the optimal and ADP policies sometimes task a unit that is farther from the casualty to reserve another MEDEVAC unit for an anticipated request.

The representative scenario assumes the standby MEDEVAC unit mission preparation time is 15 minutes longer than that of a primary unit. We conduct an excursion that considers the impact the standby unit mission preparation time has on the optimal policy and best-performing ADP policy. Standby unit additional mission preparation times from 5 minutes to 45 minutes in 5 minute increments are explored (i.e., total mission preparation times of 20 to 60 minutes). Results show that the myopic policy optimality gap increases as the standby unit mission preparation time increases, whereas the ADP policy optimality gap decreases. This trend is due to a decreasing myopic policy value and a steady optimal policy value. Moreover, as the standby unit mission preparation time increases, the rejection rate and standby unit tasking rate of the optimal policy and ADP policy slightly increase but remain consistent. Although the ADP policy has a lower rejection rate than the optimal policy, the difference between the two rates remains consistent as standby unit mission preparation time increases. Also, the two policies have nearly the same standby unit tasking rate for each standby unit mission preparation time. This shows that the optimal policy and ADP policy are fairly robust to standby unit mission preparation times.

Finally, we expand the representative scenario by including 38 zones rather than 6. All other parameters remain the same as well as the location of all four MEDEVAC units. This scenario is too large to solve to optimality, which warrants the use of an ADP solution approach. Unfortunately, the ADP solution approach we implement for the baseline scenario does not scale well to a larger problem size. However, we alter the algorithm to include a trajectory-following state sampling scheme. This algorithm modification yields a policy that attains a 3.61 ± 1.97 percent improvement over the myopic policy at the 95% confidence level. Future research includes a full computational experiment to determine the best parameter settings for this modified

ADP solution approach.

The research presented in this thesis is of interest to military medical planners and dispatching authorities. If medical planners determine realistic model input parameters (e.g., request arrival rate, expected response times, and request proportions by zone-priority level), the proposed model can be applied to compare different dispatching policies for a variety of planning scenarios with fixed medical treatment facility (MTF) and MEDEVAC staging locations. Moreover, medical planners can evaluate different COAs for standby unit locations to maximize the overall performance of the medical system.

Although there are endless possibilities for model enhancements, the following are natural extensions of the proposed model. First, this MDP model assumes that each MEDEVAC unit must return to its staging area before another tasking. However, it is more realistic that a MEDEVAC unit is tasked immediately upon casualty delivery at an MTF. Also, this model does not include the option to relocate the standby unit. In some cases, it might be beneficial for the standby unit to relocate depending on the rate of requests that are expected to arrive in the near future. Finally, although we generate high-quality policies with the proposed ADP solution approach, other ADP algorithms should be explored that scale well to larger problems.

Bibliography

Augustyn, Adam, Zeidan, Adam, Zelazko, Alicja, Eldridge, Alison, McKenna, Amy, Tikkanen, Amy, Gadzikowski, Ann, Schreiber, Barbara A., Duignan, Brian, Mahajan, Deepti, Promeet, Dutta, Goldstein, Emily, Rodriguez, Emily, Gregersen, Erik, Shukla, Gaurav, Liesangthem, Gita, Lotha, Gloria, Young, Grace, Bolzon, Henry, Luebering, J.E., Wallenfeldt, Jeff, Hibler, Joan, Rafferty, John P., Gupta, Kanchan, Rogers, Kara, Chmielewski, Kenny, Manchanda, Kokila, Heintz, Kurt, Chaveriat, Laura, Dixon, Letricia, Matthias, Meg, Petruzzello, Melissa, Ray, Michael, Metych, Michele, Bauer, Patricia, Riley, Patrick, Bhathya, Piyush, Vasich, Sheila, Hollar, Sherman, Singh, Shiveta, Gupta, Shweta, & Seddon, Stephen. 2020. *Britannica Encyclopaedia: Iraq War*. Encyclopaedia Britannica, Incorporated.

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.

Buckenmaier, Chester, III, & Mahoney, Peter F. 2015. *Combat Anesthesia: The First 24 Hours*. Textbooks of Military Medicine. Office of the Surgeon General United States Army.

Carter, Grace M, Chaiken, Jan M, & Ignall, Edward. 1972. Response areas for two emergency units. *Operations Research*, **20**(3), 571–594.

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

De Lorenzo, Robert A. 2003. Military casualty evacuation: MEDEVAC. *Pages 45–59 of: Aeromedical Evacuation*. Springer.

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

Fulton, Lawrence V, Lasdon, Leon S, McDaniel Jr, Reuben R, & Coppola, M Nicholas. 2010. Two-stage stochastic optimization for the allocation of medical assets in steady-state combat operations. *The Journal of Defense Modeling and Simulation*, **7**(2), 89–102.

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), 1–12.

Jenkins, Phillip R, Robbins, Matthew J, & Lunday, Brian J. 2021a. Approximate Dynamic Programming for Military Medical Evacuation Dispatching Policies. *INFORMS Journal on Computing*, **33**(1), 2–26.

Jenkins, Phillip R., Robbins, Matthew J., & Lunday, Brian J. 2021b. Approximate dynamic programming for the military aeromedical evacuation dispatching, preemption-rerouting, and redeployment problem. *European Journal of Operational Research*, **290**(1), 132–143.

Jenkins, Phillip R., Robbins, Matthew J., & Lunday, Brian J. 2021c. Optimising Aerial Military Medical Evacuation Dispatching Decisions via Operations Research Techniques. *BMJ Military Health*, 1–3.

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

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.

Kulkarni, Vidyadhar G. 2017. *Modeling and Analysis of Stochastic Systems*. Chapman & Hall/CRC Texts in Statistical Science Series. CRC Press, Boca Raton, FL.

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

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

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

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

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

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

Puterman, Martin L. 1994. *Markov decision processes: discrete stochastic dynamic programming.* Hoboken, New Jersey: John Wiley & Sons.

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

Robbins, Matthew J, Jenkins, Phillip R, Bastian, Nathaniel D, & Lunday, Brian J. 2020. Approximate Dynamic Programming for the Aeromedical Evacuation Dispatching Problem: Value Function Approximation Utilizing Multiple Level Aggregation. *Omega*, **91**(102020), 1–17.

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.

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

Stidham Jr., Shaler. 2002. Analysis, Design, and Control of Queueing Systems. *Operations Research*, **50**(1), 197.

Summers, Daniel S., Robbins, Matthew J., & Lunday, Brian J. 2020. An approximate dynamic programming approach for comparing firing policies in a networked air defense environment. *Computers & Operations Research*, **117**, 104890.

White, Michael. 2020. *Iraq Coalition Casualty Count*. http://icasualties.org. Accessed: 2020-12-10.

Zeto, John, Yamada, Wade, & Collins, Gale. 2006. *Optimizing the emplacement of scarce resources to maximize the expected coverage of a geographically variant demand function*. Tech. rept. US Center for Army Analysis, Ft. Belvoir, VA.

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704–0188

1. REPORT DATE *(DD–MM–YYYY)*
25–03–2021

2. REPORT TYPE
Master's Thesis

3. DATES COVERED *(From — To)*
Sept 2019 — Mar 2021

4. TITLE AND SUBTITLE
Examining How Standby Assets Impact Optimal Dispatching Decisions within a Military Medical Evacuation System via a Markov Decision Process Model

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Wooten, Kylie E., 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
Wright-Patterson AFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT-ENS-MS-21-M-196

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
The Army medical evacuation (MEDEVAC) system ensures proper medical treatment is readily available to wounded soldiers on the battlefield. The objective of this research is to determine which MEDEVAC unit to task to an incoming 9-line MEDEVAC request and where to station a single standby unit to maximize patient survivability. A discounted, infinite-horizon continuous-time Markov decision process model is formulated to examine this problem. We design, develop, and test an approximate dynamic programming (ADP) technique that leverages a least squares policy evaluation value function approximation scheme within an approximate policy iteration algorithmic framework to solve practical-sized problem instances. A computational example is applied to a synthetically generated scenario in Iraq. The optimal policy and ADP-generated policies are compared to a commonly practiced (i.e., myopic) policy. Examining multiple courses of action determines the best location for the standby MEDEVAC unit, and sensitivity analysis reveals that the optimal and ADP policies are robust to standby unit mission preparation times. The best performing ADP-generated policy is within 2.62% of the optimal policy regarding a patient survivability metric. Moreover, the ADP policy outperforms the myopic policy in all cases, indicating the currently practiced dispatching policy can be improved.

15. SUBJECT TERMS
Markov decision processes, medical evacuation, admission control, approximate dynamic programming, least squares policy evaluation

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	Capt Phillip R. Jenkins, PhD, AFIT/ENS

19b. TELEPHONE NUMBER *(include area code)*
(937)255-3636, x4727; phillip.jenkins@afit.edu