Approximate Dynamic Programming for the Dispatch of Military Medical Evacuation Assets

Aaron J. Rettke, Matthew J. Robbins*, Brian J. Lunday

*Air Force Institute of Technology, Department of Operational Sciences, 2950 Hobson Way,
Wright-Patterson AFB, OH 45433, United States*

Abstract

Military medical planners must consider the dispatching of aerial military medical evacuation (MEDEVAC) assets when preparing for and executing major combat operations. The launch authority seeks to dispatch MEDEVAC assets such that prioritized battlefield casualties are transported quickly and efficiently to nearby medical treatment facilities. We formulate a Markov decision process (MDP) model to examine the MEDEVAC dispatching problem. The large size of the problem instance motivating this research suggests that conventional exact dynamic programming algorithms are inappropriate. As such, we employ approximate dynamic programming (ADP) techniques to obtain high quality dispatch policies relative to current practices. An approximate policy iteration algorithmic strategy is applied that utilizes least squares temporal differencing for policy evaluation. We construct a representative planning scenario based on contingency operations in northern Syria both to demonstrate the applicability of our MDP model and to examine the efficacy of our proposed ADP solution methodology. A designed computational experiment is conducted to determine how selected problem features and algorithmic features affect the quality of solutions attained by our ADP policies. Results indicate that the ADP policy outperforms the myopic policy (i.e., the default policy in practice) by up to nearly 31 percent with regard to a lifesaving performance metric, as compared for a baseline scenario. Moreover, the ADP policy provides decreased MEDEVAC response times and utilization rates. These results benefit military medical planners interested in the development and implementation of cogent MEDEVAC tactics, techniques, and procedures for application in combat situations with a high operations tempo.

Keywords: Approximate dynamic programming, Emergency medical dispatch, Least squares temporal differences, Markov decision processes, Military medical evacuation (MEDEVAC)

---

*Corresponding author.

Email addresses: matthew.robbs@afit.edu (Matthew J. Robbins),
brian.lunday@afit.edu (Brian J. Lunday)

---

1. Introduction

United States military medical planners must consider the dispatching of aerial military medical evacuation (MEDEVAC) assets when preparing for and executing major combat operations. MEDEVAC involves the rapid transport of prioritized battlefield casualties from the point of injury (POI) to the nearest appropriate medical treatment facility (MTF) by dedicated, standardized MEDEVAC platforms with onboard medical personnel providing care en route.

Aerial MEDEVAC operations involve the use of dedicated helicopters, specifically the UH-60 A/L Blackhawk. Helicopters are well suited for MEDEVAC operations because of their ability to travel faster, farther, and access terrain not accessible to ground vehicles. The ability to simultaneously treat and quickly transport casualties from the POI to an MTF greatly increases the probability of survival for casualties. Recent experience indicates that survivability of injuries on the battlefield is at a historic high; 90% of all casualties survive in post 9/11 combat operations, compared to 84% in Vietnam and 80% in World War II (Eastridge et al., 2012). This improvement is attributed primarily to the speed with which casualties are able to receive proper medical attention.

Military medical planners must consider many important aspects when designing a deployed military emergency medical service (EMS) system (Department of the Army, 2007). Such aspects include determining MEDEVAC unit locations, identifying a MEDEVAC dispatching policy, and considering the redeployment or relocation of MEDEVAC units after a mission starts or terminates. The determination of locations for MEDEVAC assets to support combat operations is influenced by the tactical situation. Location decisions require a balance between maximizing coverage and minimizing response time subject to resource, force protection, and logistical constraints. The decision of which MEDEVAC unit to launch (i.e., dispatch) to a given service call is an important aspect of the military EMS system and is the focus of this paper. The MEDEVAC launch authority (i.e., the dispatching authority) typically defaults to a myopic policy in which the nearest MEDEVAC unit is launched to a POI regardless of its categorized priority (among three possible tiers). Redeployment of MEDEVAC assets from an MTF directly to a new POI without returning to the MEDEVAC unit’s originating base is possible; however, refueling and medical resupply requirements, armed escort availability, communication challenges, and crew basing limitations often make MEDEVAC redeployment difficult, and it is not often performed within a theater of operations. For similar reasons, the temporary relocation of idle MEDEVAC assets to a different base is also not routinely performed.

In this paper, we consider the MEDEVAC dispatching problem in which a dispatching authority must decide how to dispatch MEDEVAC units to prioritized requests for service. We develop a Markov decision process (MDP) model of the MEDEVAC dispatching problem. The large size of the motivating problem instance yields a high-dimensional and uncountable state space, rendering

classical dynamic programming methods inappropriate. As such, we apply approximate dynamic programming (ADP) techniques to obtain high quality dispatch policies. We develop an approximate policy iteration (API) algorithm that utilizes least-squares temporal difference (LSTD) learning for policy evaluation. We define a set of basis functions within a linear architecture to approximate the value function around the post-decision state. To demonstrate the applicability of our MDP model and to examine the efficacy of our proposed solution methodology, we construct a notional, representative planning scenario based on contingency operations in northern Syria. A designed computational experiment is conducted to determine how selected problem features and algorithmic features affect the quality of solutions attained by our ADP policies.

The unique military aspect of the MEDEVAC dispatching problem warrants discussion. The lack of unclassified historical, extensible data presents difficulties for those interested in examining military EMS systems. We cannot simply analyze computer-aided dispatch (CAD) data to obtain the necessary parameters for our MDP model. When seeking to inform decisions regarding the dispatching of MEDEVAC assets (e.g., in anticipation of a major combat operation), it is necessary to have notions of the anticipated mission, enemy disposition, terrain, weather, friendly forces disposition, logistical support availability, time constraints, and civil authority considerations. Within a military context, *disposition* denotes whether a unit is passively defending an area, actively patrolling it, or conducting some other deliberate operation. Calls for service result from the interaction of the opposing forces in hostile actions. Such actions are very fluid with respect to both location and intensity. Even when the sizes, locations, and dispositions of friendly forces are relatively stable, the locations and dispositions of adversary forces are dynamic and challenging to predict. We develop a simulation of the MEDEVAC process that implicitly considers such information. For example, deliberate selection of the casualty cluster centers within the Hawkes process (Hawkes, 1971) governing casualty event arrivals embeds the information concerning the interaction of friendly and adversary forces. The Hawkes process assumes that subsequent events are more likely to occur in close proximity to prior events, and so it works well when simulating the spatiotemporal aspects of both criminal and terrorist activity (Egesdal et al., 2010; White et al., 2013); moreover, Lewis et al. (2012) show that this process well represents casualty incidents within a country fighting an insurgency. We utilize the MEDEVAC process simulation within our ADP algorithm to obtain simulated trajectories of the system and to measure the performance of the MEDEVAC dispatch policies obtained from our ADP algorithm.

The remainder of this paper is organized as follows. Section 2 presents a review of pertinent literature concerning civilian and military EMS systems. Several ADP papers that inform the development of our model and solution methodology are also reviewed. Section 3 presents a description of the MEDEVAC dispatching problem. Section 4 describes the MDP model formulation of the MEDEVAC dispatching problem and presents our ADP approach. In Section 5 we demonstrate the applicability of our model and examine the efficacy of our proposed solution methodology. Section 6 provides conclusions and

directions for future research.

2. Literature Review

Two streams of literature inform our work. The first stream of literature relates to emergency medical service (EMS) systems, with an emphasis on the dispatching of emergency response vehicles to calls for service. The second stream of literature involves selected works concerning approximate dynamic programming (ADP).

2.1. EMS systems

Past research endeavors examine both civilian and military EMS systems. Important aspects considered include determining server locations; defining the number and sizes of response districts (if a service area partitioning scheme is adopted), the number of servers per server location, and the server dispatch policy; possibly repositioning servers as a consequence of call initiation or service completion; and considering whether to focus on response time or patient survivability as the objective. Military medical planners also examine optimal placement of medical treatment facility (MTF) locations when designing a military medical evacuation (MEDEVAC) system, whereas hospital locations are typically given as fixed in civilian EMS studies. Various operations research methods are employed to examine EMS systems. These methods include, but are not limited to, discrete optimization, stochastic modeling, queueing, and simulation modeling.

Research from the 1970's and 1980's focuses primarily on civilian EMS systems and examines such aspects as the optimal placement (Fitzsimmons, 1973; Kolesar and Blum, 1973), relocation (Kolesar and Walker, 1974), and dispatch of emergency vehicles (Green and Kolesar, 1984; Ignall et al., 1982; Larson, 1972; Swersey, 1982) to improve the performance of the EMS system. See Green and Kolesar (2004) for an account of early emergency service management papers. Few papers seek to improve the performance of the military EMS (i.e., MEDEVAC) system. In this paper, we focus on the dispatch of emergency response via aerial MEDEVAC to prioritized requests for service. Examining the dispatch policy of emergency response vehicles requires a dynamic and stochastic approach. Moreover, consideration of casualty priority classification is important. Many EMS system dispatch policies do not account for the priority level of the patient (Bandara et al., 2014). This omission results in the nearest emergency response unit fulfilling the service requirement without regard to the void created in the system by that unit's temporary absence. Known as a myopic policy, such procedures are shown to be inadequate by many researchers (Carter et al., 1972; Kuisma et al., 2004; Nicholl et al., 1999); it follows that better informed priority dispatching can save more lives on the battlefield.

Another important aspect of the dispatching problem is the selection of the system optimality criterion (or performance measure). Selection of an appropriate EMS system performance measure is important since it governs employment of EMS system resources and ultimately determines patient survivability

(McLay and Mayorga, 2010). EMS system performance is most often measured in terms of response time threshold (RTT), indicating the proportion of calls that are serviced within a fixed timeframe. A service call is considered *covered* if served by the stated time threshold. An often employed MEDEVAC system performance measure is the proportion of Priority I (Urgent) calls served in less than 60 minutes. While RTTs are easy to evaluate and easy to explain, one important criticism relates to how well the performance measure captures patient survival, the fundamental objective of the EMS system. For example, a system would receive credit for a MEDEVAC arriving at the MTF 59 minutes and 59 seconds after the call, but not for arriving at exactly 60 minutes. Yet, the medical literature (Fitch, 2005) suggests that there would be little difference in patient survival between two such cases. Erkut et al. (2008) compares the utilization of RTTs to patient survival rates in an ambulance location problem; the authors recommend incorporating patient survivability explicitly and avoiding the use of RTTs, if possible. Unfortunately, as McLay and Mayorga (2010) note, estimating patient survivability is quite difficult. For our purposes, survival is defined in terms of discharge from the military medical system, which might occur several months later and after many transfers from field hospitals in the combat zone to larger, permanent facilities in the United States. It would be very difficult to retroactively determine MEDEVAC responses that resulted in survival.

Despite such challenges, Erkut et al. (2008) propose the use of a monotonically decreasing function over time for the probability of patient survival as a measure of performance. Subsequent work by Bandara et al. (2012), Mayorga et al. (2013), Bandara et al. (2014), and Grannan et al. (2015) similarly use survivability functions instead of the traditional RTT as a performance measure for their model. We also follow Erkut et al. (2008) by assuming a function that is monotonically decreasing in response time to model casualty outcome. A challenge with the use of a survivability function is the difficulty finding empirical evidence to support a particular functional form. Eastridge et al. (2012) provide extensive statistics about combat deaths, but the response times are not known, preventing the creation of a high confidence survivability function from their results. Although Feero et al. (1995) examine EMS response times relative to trauma survivability, their work is limited to response times under eight minutes. MEDEVAC units typically need to travel significantly further than civilian EMS units, and so response times are notably longer. The MEDEVAC system seeks to respond within 60 minutes from notification to drop-off of the Priority I patient at an MTF (Garrett, 2013). The stated RTTs for each of the casualty priority levels (i.e., 60 minutes, 240 minutes, and 24 hours (Department of the Army, 2007)) serves to inform the development of our casualty outcome functions.

Bandara et al. (2012) mention several studies in which the EMS system greatly improves patient survival probability if priority level is considered when deciding which emergency response vehicle to dispatch. The authors construct an MDP model that focuses on the priority level of an emergency call. Their study indicates an optimal policy in which the closest available (i.e., idle) unit is

dispatched to the most urgent call and the next closest available unit to the less urgent call, regardless of the call order. While this result may seem intuitive, this dispatch policy quickly becomes complex. For example, it may be optimal to dispatch a more distant vehicle if the closer vehicle is more likely to receive a higher priority call. This policy essentially rations the closer vehicle in anticipation of a more urgent request. For problems with several more service zones and ambulances, the policy might not be as intuitive, although EMS systems may stand to benefit greatly from the employment of an optimal policy versus a myopic policy. [Bandara et al. \(2014\)](#) develop a simulation model to evaluate dispatch policies for EMS systems. The authors measure performance in terms of patient survival probability. Their results also indicate that dispatching the closest emergency response vehicle is not always optimal and that dispatching vehicles considering priority levels leads to an increase in the average survival probability of patients. The authors utilize their results to develop a heuristic that could be applied to the analysis of large-scale systems.

[Mayorga et al. \(2013\)](#) examine the dispatch policy within the EMS system wherein performance is measured in terms of the probability of patient survival. Prior to examining the dispatch policies, however, the authors obtain solutions for the number of districts and district location decisions by developing a construction heuristic. Their research provides more depth than previous studies by analyzing the dispatch policies for inter-district and intra-district situations. An intra-district situation occurs when a response vehicle services the call within its own district, whereas an inter-district situation occurs when all response vehicles within a district area are busy and the call must be serviced by a response vehicle from a different district. For the inter-district policy, either a myopic policy (i.e., the closest vehicle responds) or a heuristic policy (e.g., as developed in [Bandara et al. \(2012\)](#)) is implemented. Whereas the myopic policy is the most widely applied policy in EMS systems, the heuristic policy considers the priority of the call as well as the workload of each crew. For such an implementation, a utilization factor is included to account for the workload of each crew. For the intra-district policy, two policies are considered. The first policy assumes that a sister emergency service (e.g., fire department or police department) will respond. The second policy uses the heuristic policy that [Bandara et al. \(2012\)](#) developed and allows a response vehicle from another district to cross boundary lines.

[McLay and Mayorga \(2013b\)](#) examine the dispatch policy within the EMS system while considering patient classification errors. [McLay and Mayorga \(2013b\)](#) differs from [Mayorga et al. \(2013\)](#) and [Bandara et al. \(2012\)](#) in that they consider patient classification errors and employ an optimality criterion based on RTT rather than patient survivability. They focus on the patients' urgency level with an overall objective of maximizing the average long-run utility of the EMS system, rewarding the expected coverage of high-risk patients. [McLay and Mayorga \(2013a\)](#) examine the dispatch policy within a general server-to-customer service system (e.g., an EMS system) while considering the issue of balancing equity and efficiency. They develop a constrained variant of the Markov decision problem introduced by [McLay and Mayorga \(2013b\)](#) and formulate a

linear programming model to solve the constrained problem. Their objective is to dispatch ambulances to patients to maximize average total reward subject to minimum standards of equity. They report results while considering four different notions of equity.

EMS system research exists that focuses on military MEDEVAC systems. [Fulton et al. \(2010\)](#) and [Bastian \(2010\)](#) examine the allocation of MEDEVAC units in steady-state combat operations. [Fulton et al. \(2010\)](#) present a stochastic optimization model that locates deployable hospitals, reallocates hospital beds, and determines the location of emergency response vehicles (both aerial and ground MEDEVAC assets) prior to MEDEVAC system initialization. Their objective function minimizes total penalty-weighted expected service times. The penalty weight captures patient severity, derived from the historical data collection of patient injury severity scores from combat operations in the Iraq theater. [Bastian \(2010\)](#) describes a multi-criteria modeling approach that optimizes the emplacement of MEDEVAC assets. The author considers casualty demand coverage, MEDEVAC spare capacity, and site attack vulnerability. [Keneally et al. \(2014\)](#) develop an MDP model to examine MEDEVAC dispatch policies in the Afghanistan theater. The authors consider priority classifications and a reward function based on RTT, and they simulate a Hawkes process for casualty generation. They observe that myopic policies do not always lead to optimal EMS system performance. [Grannan et al. \(2015\)](#) examine location decisions in the Afghanistan theater. The authors propose a binary integer programming model to optimally locate two types of aerial MEDEVAC assets. The underlying queueing dynamics of their MEDEVAC system are captured using an approximate spatial queueing model, with the results serving to parameterize their integer program. The results obtained in the aforementioned research endeavors serve to inform decision-making by military medical planners.

2.2. Approximate Dynamic Programming

The decisions concerning which MEDEVAC to dispatch to a request for service must be made sequentially over time and under uncertainty. Since dispatch decisions impact the capability of the MEDEVAC system to serve future requests, we must account for how current decisions affect the future state of the system. As such, we formulate a Markov decision process (MDP) model of the MEDEVAC dispatching problem. Unfortunately, due to the well-known curses of dimensionality, an optimal policy cannot be identified using classical exact dynamic programming algorithms. Instead, we employ an ADP solution methodology to solve the MEDEVAC dispatching problem. For an introduction to ADP from an operations research perspective we refer the reader to [Powell \(2009, 2011, 2012\)](#). Earlier contributions from an engineering controls perspective ([Bertsekas and Tsitsiklis, 1996](#)) and from a computer science (artificial intelligence) perspective ([Sutton and Barto, 1998](#)) also provide excellent overviews.

Two general algorithmic strategies exist for obtaining approximate solutions to our computational stochastic optimization problem: approximate value iteration (AVI) and approximate policy iteration (API). The interested reader

is referred to Bertsekas (2011) for a detailed discussion concerning API. We utilize an API algorithmic strategy to obtain a policy that maps the system state (e.g., MEDEVAC locations and service status, casualty event locations and priorities) to an action (e.g., dispatching a specific MEDEVAC unit to service a new or queued call). Powell (2012) discusses four classes of policies: myopic cost function approximation, lookahead policies, policy function approximations, and policies based on value function approximations. We construct high-quality dispatch policies based on value function approximations. Our approximation strategy involves the design of an appropriate set of basis functions for use within a linear architecture. Moreover, we approximate the value function around the post-decision state. First introduced by Van Roy et al. (1997), the post-decision state allows for modification of Bellman’s equation to obtain an equivalent, deterministic expression. Use of the post-decision state addresses the curse of dimensionality with respect to the outcome space (Powell, 2011). Within the policy evaluation step of our API algorithm, we update the value function approximation for a fixed policy using least squares temporal differencing (LSTD). Introduced by Bradtke and Barto (1996), LSTD is a computationally efficient method for estimating the adjustable parameters when using a linear architecture with fixed basis functions to approximate the value function for a fixed policy. Lagoudakis and Parr (2003) extend the LSTD algorithm to include the consideration of state-action pairs. We implement a variant of the LSTD algorithm that utilizes value function approximations around the post-decision state, as recommended by Powell (2011) and demonstrated in Scott et al. (2014).

Two recent papers utilize ADP to examine ambulance repositioning (i.e., relocation or redeployment) and dispatching in civilian EMS systems. These two papers greatly informed the development of our paper and warrant discussion.

Maxwell et al. (2010) utilize an API algorithmic strategy to determine high quality ambulance redeployment decisions. At each system event occurrence, such as a new call arrival or an ambulance arrival at a hospital after transporting a patient, a decision is made to redeploy any idle ambulances to better cover the service area. The authors seek policies that minimize expected total discounted cost and define the single-period cost function in terms of an urgent call not being covered. That is, their system sustains a cost if the RTT requirement is not met. They demonstrate significant improvement in EMS response at two metropolitan cities, comparing their ADP redeployment policy to the currently practiced static, myopic policy. Our paper differs from Maxwell et al. (2010) in a number of ways. We approximate the value function around the post-decision state, whereas they use the pre-decision state. We employ temporal differencing in the policy evaluation step of our API algorithm, whereas they do not. We define a reward function based on casualty survival, whereas they define cost in terms of RTT. We focus on dispatching decisions, whereas they focus on redeployment decisions.

Schmid (2012) adopts an AVI algorithmic strategy to determine high quality relocation and dispatching decisions. At each system event occurrence, a decision is made to (1) relocate any idle ambulances to better cover the service

area and (2) dispatch available ambulances to requests for service. Polices are sought that minimize expected total discounted cost, wherein the single-period cost function is defined in terms of response times. The author demonstrates improvement in EMS response for the city of Vienna, Austria, reporting that the ADP policy reduces average response time from 4.6 minutes to 4.01 minutes, a 12.9% reduction. Our paper differs from Schmid (2012) in the application and corresponding model, as well as in the methodology we apply to solve it. We consider a military application of the dispatching problem that requires implicit consideration of escort vehicle (i.e., attack helicopters) availability. This military aspect of the problem is captured within our simulation in terms of the random dispatch time. In a civilian context, such a phenomenon would be akin to a requirement for police escort for some service calls. The most significant difference involves the algorithmic strategy; we employ API, whereas the author uses AVI. Although the author does approximate the value function around the post-decision state as we do in this research, we utilize a linear model with a designed set of basis functions to approximate the value function, whereas the author employs an aggregation scheme.

3. Problem Description

The aeromedical evacuation coordination center (AECC) is the primary decision-making authority within the military medical evacuation (MEDEVAC) system, responsible for monitoring and managing all activities related to the execution of aeromedical evacuation operations (Department of the Army, 2007). In a deployed military emergency medical service (EMS) system, the AECC acts as the MEDEVAC dispatching authority. When a casualty occurs and a ‘9-line’ MEDEVAC request is received, a decision must be made quickly regarding which MEDEVAC asset (if any) to dispatch. Delays in decision making affect casualty survivability. Thus, it is critical that the AECC employs a dispatching policy that results in the quick and efficient transport of high priority battlefield casualties from the point of injury (POI) to the nearest medical treatment facility (MTF). The MEDEVAC dispatching process is depicted in Figure 1.

Figure 1: MEDEVAC Dispatch Timeline

A 9-line MEDEVAC request is transmitted over secure radio on a dedicated MEDEVAC frequency and contains a standard, prescribed amount of information. Such information includes the location of the POI, radio frequency and call sign of the requesting unit, number of patients by precedence, special equipment needed, casualty nationalities, and threat conditions at the POI. When making the MEDEVAC request, the on site medic and unit leader apply a three category casualty triage rubric to govern evacuation precedence ([Department of the Army, 2007](#)):

- **Priority I, Urgent:** Assigned to emergency cases that should be evacuated as soon as possible and *within a maximum of 1 hour* in order to save life, limb, or eyesight, to prevent complications of serious illness, or to avoid permanent disability.
- **Priority II, Priority:** Assigned to sick and wounded personnel requiring prompt medical care. This precedence is used when the individual should be evacuated *within 4 hours* or when an individual's medical condition could deteriorate to such a degree that he or she will become an *Urgent* precedence, or whose requirements for special treatment are not available locally, or who will suffer unnecessary pain or disability.
- **Priority III, Routine:** Assigned to sick and wounded personnel requiring evacuation but whose condition is not expected to deteriorate significantly. The sick and wounded in this category should be evacuated *within 24 hours*.

MEDEVAC requests are typically transmitted from the POI, through intermediary headquarters, to the AECC at higher headquarters. The particular information flow depends on the communications equipment available to the requesting unit, the communication infrastructure within the command, and the command and control organization of the MEDEVAC system. The time at which the MEDEVAC request is received by the AECC is denoted by T_1 . The AECC may or may not immediately assign a MEDEVAC unit¹ to the request, depending on the location of the request, the precedence categories of the casualties, the status of the MEDEVAC units, and any pre-existing requests in the system. If a suitable idle MEDEVAC unit is available, it is directed at time T_2 to service the request. The AECC transmits the 9-line MEDEVAC request information to the assigned MEDEVAC unit via the command's communication system. The time at which the MEDEVAC unit departs its station for the POI is denoted by T_3 . The amount of time required between receipt of the 9-line MEDEVAC request at the AECC, T_1 , and the MEDEVAC unit departure, T_3 , is the dispatch time. The dispatch time comprises the process duration of determining which MEDEVAC unit to dispatch; whether an armed escort is required; which armed escort unit to dispatch, if required; notifying the MEDEVAC unit; and preparing the medical personnel, medical equipment, and the MEDEVAC helicopters for the mission.

The MEDEVAC unit arrives on the ground at the POI site at time T_4 and begins treating and loading casualties immediately. The time at which initial treatment and loading ends and the MEDEVAC unit departs the POI site is denoted by T_5 . The destination MTF is selected in a deterministic manner based on the location of the POI. The MEDEVAC unit travels to the nearest MTF, arriving at time T_6 . The casualties are then unloaded and receive subsequent treatment from the medical staff at the MTF. Once the casualties are unloaded, the MEDEVAC unit departs the MTF and travels back to its own station (or staging area). The time at which the MEDEVAC unit completes the unloading process and departs the MTF is denoted by T_7 . The time at which the MEDEVAC unit arrives at its station is denoted by T_8 . Upon arrival, the MEDEVAC unit's mission is complete, and it becomes available for dispatch once again. Note that travel times from the MEDEVAC station to the POI site, from the POI site to an MTF, and from the MTF to the MEDEVAC station are expected to vary based on battlefield conditions (e.g., weather conditions, enemy disposition, air density due to flight altitude, and the amount of equipment being transported).

MEDEVAC units are not dispatched to POI sites directly from nearby MTFs immediately after unloading casualties from the prior mission, nor are they repositioned to other MEDEVAC stations; MEDEVAC units must return to their own stations. This restriction results from concerns about low fuel levels, on-board medical equipment configurations, crew bed down limitations, and other

---

¹We note that, in U.S. military doctrine, a MEDEVAC unit is typically comprised of two MEDEVAC helicopters when executing missions due to a mutual support safety requirement.

logistical issues. Predominantly, MEDEVAC units must return to their stations to refuel and, if necessary, reconfigure onboard medical equipment.

EMS systems typically refer to response time as the amount of time required to reach the patient after receiving an emergency call. According to [McLay and Mayorga \(2010\)](#), the rapid response to cardiac arrest situations is a primary focus of the (civilian) EMS system. This rationale exists because the EMS system is often evaluated on how quickly it responds to emergency cardiac arrest calls since there is effective treatment for cardiac arrests, and they are time-sensitive. Also, if the EMS system can respond quickly enough to a cardiac arrest call, it is more likely to be successful with similar life-or-death situations. Therefore, it is quite intuitive that the response time for a civilian EMS system is typically defined as the time between the receipt of the emergency call and the time the first emergency response vehicle arrives at the injury site.

However, the performance of the MEDEVAC system cannot be evaluated using the same measures as the EMS system since several additional factors are involved when medically evacuating a casualty from a battlefield. Not only can the load times, travel times, and unload times be much greater and vary much more, the primary cause of death on the battlefield is blood loss, not cardiac arrest. [Garrett \(2013\)](#) reports that 85% of soldiers killed in action (KIA) were a direct result of blood loss. Very recent improvements have been made in this area by equipping MEDEVAC units with in-flight blood transfusion capabilities, but not enough data has been generated to alter the MEDEVAC system’s evaluation measure at the time of this research ([Malsby III et al., 2013](#)). Thus, we consider it far more critical to stabilize and transport casualties to the nearest MTF (with an available blood transfusion capability) and into surgery rather than to simply reach the casualties quickly. Accordingly, we define the response time for a MEDEVAC unit as $T_6 - T_1$. Moreover, we define service time for a MEDEVAC unit as $T_8 - T_2$, the time expended servicing the call.

The overall objective of the MEDEVAC system is to dispatch MEDEVAC units so as to minimize the response times for all MEDEVAC requests, with a particular focus on the response times of the *urgent* and *priority* casualties. The decisions concerning which MEDEVAC to dispatch to a 9-line request for service must be made sequentially over time, under uncertainty. Casualty events and the subsequent MEDEVAC requests are not known *a priori* and become known to the MEDEVAC system only upon receipt. Once a MEDEVAC request is initiated, the assigned MEDEVAC unit must respond promptly, moving to the POI and transporting the casualties to the nearest MTF while providing medical care en route. The stochastic nature of the problem results from the uncertain manner in which casualty events manifest. Moreover, the dispatch time, travel times, and service times are stochastic. Information about dispatch, travel, and service times is available and can be leveraged to parameterize our models. However, as noted previously, the challenge for military medical planners is identifying a dispatch policy *before* the commencement of a major combat operation. In such a situation, no casualty event data exists. We employ a spatiotemporal point process, a Hawkes process ([Hawkes, 1971](#)), to model casualty event arrivals. Deliberate selection of the casualty event cluster centers

within the Hawkes process embeds the military planners’ expert judgment concerning the future interaction of friendly and adversary forces that results in the occurrence of casualty events.

4. Methodology

This section describes the Markov decision process (MDP) model formulation of the medical evacuation (MEDEVAC) dispatching problem. The approximate dynamic programming (ADP) methodology utilized to obtain high quality solutions to the problem is also presented.

4.1. MDP Formulation

The objective of the MDP model is to determine which MEDEVAC unit to dispatch in response to a given 9-line MEDEVAC request in order to maximize expected total discounted reward over an infinite horizon. The reward function is monotonically decreasing in response time to model casualty outcome. We assume casualty events arrive sequentially over time and consider two of the three casualty event precedence categories: *urgent* and *priority* (Department of the Army, 2007). Because the focus of rapid aerial MEDEVAC support is on life-saving evacuations, casualty events having a *routine* category are assumed to be handled by ground MEDEVAC or casualty evacuation (i.e., evacuation provided by non-medical personnel and equipment). Hence, we exclude *routine* calls for service from our analysis.

A self-exciting spatiotemporal point process called a Hawkes process is utilized to generate casualty events over time. The Hawkes process is commonly applied in seismology, where it captures the notion that earthquakes can have aftershocks, and those aftershocks can have aftershocks, and so on (Hawkes and Adamopoulos, 1973). In our context, the occurrence of casualty events in a self-exciting point process causes other nearby casualty events to be more likely to occur due to the presumption of continued violent interactions between friendly and adversary forces. Using military planning guidelines for combat operations, we identify casualty event cluster centers that correspond both to locations at which friendly operations to secure the populace and transport supplies are certain and at which enemy attacks are likely to be mounted while reducing the risk of exposure by enemy forces. We proceed by generating a sequence of casualty event centers from a stationary Poisson process on the set of identified casualty event cluster centers with arrival rate λ . We then generate clusters of offspring casualty events around each center in the sequence. The collection of all generated events forms the Hawkes process. See Kroese and Botev (2014) for a full description of the generation procedure.

Should a casualty event (i.e., MEDEVAC request or call) arrive when all MEDEVAC units are busy, the request is placed in a queue to be serviced later. Incorporating a queue in the model allows the possibility of waiting before launching a MEDEVAC. The decision to wait can be advantageous if a low priority casualty event has occurred while many other MEDEVAC units

are busy. It may be better to remain on standby until other MEDEVAC units become available in case a high priority casualty event occurs. All casualties are evacuated to the nearest medical treatment facility (MTF). After casualties are delivered to an MTF, the MEDEVAC unit returns to its originating station (or base).

We leverage [Maxwell et al. \(2010\)](#) when developing the state space of our MDP model. The system state is represented by the tuple $s = (\tau, \epsilon, M, Q)$, wherein τ corresponds to the current system time and ϵ corresponds to the current event. The two main system components are (1) the MEDEVAC status tuple $M = (m_1, m_2, \dots, m_a)$, wherein m_i contains information about the i^{th} MEDEVAC unit and a represents the number of MEDEVAC units in the system, and (2) the call queue status tuple $Q = (q_1, q_2, \dots, q_b)$, wherein q_j contains information about the j^{th} call in the queue and b represents the maximum number of calls allowed in the queue. The state of MEDEVAC unit i is given as a tuple $m_i = (\sigma_i, d_i, t_i)$, wherein σ_i is the status of the MEDEVAC unit, d_i is the expected time at which the current mission is completed (i.e., casualties delivered to the nearest MTF), and t_i is the starting time of the MEDEVAC unit's mission. After a MEDEVAC unit delivers its onboard casualties to an MTF, it returns to base where it then becomes available to launch. The status of the MEDEVAC unit σ_i can be "idle", "en route to call j ", "at call j POI", "en route to MTF with call j casualties", "at MTF", or "returning to base". If the MEDEVAC unit is not idle, t_i corresponds to the starting time of the movement; otherwise, t_i represents the time of the current event cycle. The state of call j in the queue is $q_j = (\delta_j, l_j, \zeta_j, \eta_j)$, where δ_j is the status of the call in the j^{th} position, l_j is the location of the call, ζ_j is the time the call arrived in the system, and η_j is the priority of the call. The status of a call δ_j can be "queued" or "assigned to MEDEVAC unit i ." Once a MEDEVAC unit is launched to service a call, the queue status is updated. The call is removed from the queue once the servicing MEDEVAC unit delivers the call's casualties to the MTF.

Events are triggered by changes in the status of a MEDEVAC or an arrival of a call. We assume MEDEVAC dispatch decisions only occur when an event occurs. Although it is possible to reroute a MEDEVAC mid-flight, delay and confusion in communication can cause large problems and this practice is not typical in such operations.

To capture dispatching decisions we let $\mathcal{M}(s) = \{i : \sigma_i = \text{"idle"}\}$ denote the set of MEDEVACs available for dispatching and $\mathcal{Q}(s) = \{j : \delta_j = \text{"queued"}\}$ denote the set of casualty events awaiting service when the system is in state s . Let $x_{ij}(s) = 1$ if MEDEVAC unit $i \in \mathcal{M}(s)$ is dispatched to call $j \in \mathcal{Q}(s)$ when the system is in state s , and 0 otherwise. Further, let $x(s) = \{x_{ij}(s) : i \in \mathcal{M}(s), j \in \mathcal{Q}(s)\}$. The set of feasible decisions can then be written as:

$$\mathcal{X}(s) = \left\{ x(s) \in \{0, 1\}^{|\mathcal{M}(s)| \times |\mathcal{Q}(s)|} : \sum_{i \in \mathcal{M}(s)} \sum_{j \in \mathcal{Q}(s)} x_{ij}(s) \leq 1 \right\}. \quad (1)$$

The trajectory of the system is given by $\{(s_k, x_k) : k = 1, 2, \dots\}$ where s_k is the state of the system and x_k is the decision after the k th event has occurred. The following recursion captures the dynamics of the system: $s_{k+1} = \psi(s_k, x_k, \omega(s_k, x_k))$, where $\omega(s_k, x_k)$ is a random element of an appropriate space representation of the stochastic process related to casualty event arrivals and MEDEVAC travel, load delays, and unload delays, and where ψ is the transfer function.

The reward function $r(s_k, x_k, s_{k+1})$ is defined as follows:

$$r(s_k, x_k, s_{k+1}) = \begin{cases} \beta_1 (0.99)^{\frac{t_{k+1} - \zeta_j}{60}} & \text{if } \eta_j = \text{"urgent"} \text{ and } \epsilon(s_{k+1}) = \text{"MEDEVAC } i \\ & \text{delivers call } j \text{ at MTF"} \\ \beta_2 (0.99)^{\frac{t_{k+1} - \zeta_j}{240}} & \text{if } \eta_j = \text{"priority"} \text{ and } \epsilon(s_{k+1}) = \text{"MEDEVAC } i \\ & \text{delivers call } j \text{ at MTF"} \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

The system receives rewards upon delivery of casualties to an MTF based on the time the call has been in the system. Time requirements are outlined by the [Department of the Army \(2007\)](#); the MEDEVAC system seeks to respond within 60 minutes from notification to MTF drop-off for *urgent* casualties and within 240 minutes from notification to MTF drop-off for *priority* casualties. The reward obtained by servicing higher priority casualties decays quickly; however, servicing higher priority casualties obtains higher reward and therefore are often served first. For lower priority casualties, the decay allows enough time to wait before being forced to launch immediately, which increases the flexibility of the system. The β parameters model the value trade-off between servicing *urgent* and *priority* casualty events. Note that no rewards are earned for *routine* calls. In the high operational tempo combat scenarios of interest in this paper, *urgent* and *priority* calls are assumed to occur with high frequency, and commanders will have less interest in servicing non-life threatening injuries using a limited fleet of aerial MEDEVAC helicopters.

A policy $\pi(s) \in \mathcal{X}(s)$ maps the state space to the action space, indicating the action to take when the system is in state s . When following policy π , the state trajectory of the system $\{s_k^\pi : k = 1, 2, \dots\}$ evolves according to $s_{k+1}^\pi = \psi(s_k^\pi, \pi(s_k^\pi), \omega(s_k^\pi, \pi(s_k^\pi)))$. The expected total discounted reward earned by the system when following policy π and starting from initial state s is

$$J^\pi(s) = \mathbb{E} \left[\sum_{k=1}^{\infty} \gamma^{k-1} r(s_k^\pi, \pi(s_k^\pi), s_{k+1}^\pi) | s_1^\pi = s \right], \quad (3)$$

where $\gamma \in [0, 1)$ is the discount factor. We adopt an event-based discounting scheme to model the notion that servicing recent calls is more important than servicing future calls. However, we also utilize a relatively high discount factor (e.g., 0.98) in our computational experiments to incentivize the system to position itself to efficiently respond to future calls for service and to not simply act to obtain the best immediate reward (e.g., by sending the nearest MEDEVAC no matter the state of the system) thereby sacrificing future rewards.

The optimal policy π^* maximizes the expected total discounted reward and satisfies the following optimality equation

$$J(s) = \max_{x \in \mathcal{X}(s)} \{ \mathbb{E} [r(s, x, \psi(s, x, \omega(s, x))) + \gamma J(\psi(s, x, \omega(s, x)))] \}. \quad (4)$$

Unfortunately, computing an optimal policy using Equation (4) is difficult. Even if we discretize the state space to obtain a finite representation, the high dimensionality of the state space renders the problem of determining an optimal policy computationally intractable. We instead propose an ADP approach. Utilizing a value function approximation approach, high quality suboptimal policies are constructed and compared to a simple myopic policy, which is often employed in practice.

4.2. ADP Formulation

We employ an approximate policy iteration (API) algorithmic strategy to construct high-quality dispatch policies based on value function approximations. To obtain such policies we must determine approximate solutions to Equation (4). We proceed by employing a modified version of the optimality equation that uses a *post-decision state* variable convention. The post-decision state s_k^x refers to the state of the system after being in state s_k and taking action $x_{ij}(s_k)$. The post-decision state variable provides tremendous computational advantages as its use eliminates the embedded expectation within the optimality equation. The value of being in pre-decision state s_k is denoted by $J_k(s_k)$, and the value of being in post-decision state s_k^x is denoted by $J_k^x(s_k^x)$. The relationship between $J_k(s_k)$ and $J_k^x(s_k^x)$ is then given by

$$J_k^x(s_k^x) \triangleq \mathbb{E} [J_{k+1}(s_{k+1}) | s_k^x]. \quad (5)$$

The optimality equation in terms of the post-decision state variable is

$$J_{k-1}^x(s_{k-1}^x) = \mathbb{E} \left[\max_{x \in \mathcal{X}(s)} (r(s_k, x, s_{k+1}) + \gamma J_k^x(s_k^x)) \middle| s_{k-1}^x \right]. \quad (6)$$

Despite the dimensionality reduction attained by using the post-decision state variable, the optimality equation remains computationally intractable. We proceed by defining a fixed set of basis functions to approximate the post-decision state value function, $J_k^x(s_k^x)$. Let $\phi_f(s_k^x)$ be a *basis function*, where $f \in \mathcal{F}$ is a *feature* and $\mathcal{F}$ is the set of features. Identification of basis functions and features that are important to a particular problem can be difficult but is key to obtaining a quality approximation. Equation (7) shows the linear approximation architecture we adopt. Let

$$\bar{J}_k^x(s_k^x; \theta) \triangleq \sum_{f \in \mathcal{F}} \theta_f \phi_f(s_k^x) = \theta^\top \phi(s_k^x), \quad (7)$$

where $\phi(s_k^x)$ is a column vector with elements $\{\phi_f(s_k^x)\}_{f \in \mathcal{F}}$ and θ is a column vector of basis function weights. By substituting the value function approxi-

mation (7) into the modified optimality equation (6), we obtain the following expression for our value function approximation

$$\theta^\top \phi(s_{k-1}^x) = \mathbb{E} [r(s_k, \pi(s_k; \theta), s_{k+1}) + \gamma \theta^\top \phi(s_k^x) \mid s_{k-1}^x], \quad (8)$$

where, for a given weight vector θ , the decision is given by the policy function

$$\pi(s_k; \theta) = \arg \max_{x \in \mathcal{X}(s_k)} \{R(s_k, x) + \gamma \theta^\top \phi(s_k^x)\}. \quad (9)$$

Note that we let $R(s_k, x)$ denote the expected reward $\mathbb{E} [r(s_k, x, s_{k+1}) \mid s_k, x]$.

Selection of the set of basis functions is an important aspect of our approximation strategy. We proceed by discussing the development of the basis functions $\{\phi_f(\cdot) : f \in \mathcal{F}\}$ utilized in our approximation of the post-decision state value function, Equation (7). Basis functions can be very difficult to develop. Much trial and error went into the identification of effective basis functions for our application, and so we first discuss the logic underlying each one we utilize. The first basis function utilizes an indicator variable, denoting whether MEDEVAC unit i is available. Let

$$\phi_{1,i} = \begin{cases} 1 & \text{if } \sigma_i = \text{"idle"} \\ 0 & \text{otherwise.} \end{cases} \quad i = 1, 2, \dots, a \quad (10)$$

The second basis function captures the expected time until MEDEVAC unit i becomes available. Let

$$\phi_{2,i} = \begin{cases} d_i - \tau + d_{i,j}^r & \text{if } \sigma_i \neq \text{"idle"} \\ 0 & \text{otherwise,} \end{cases} \quad i = 1, 2, \dots, a \quad (11)$$

where $d_{i,j}^r$ represents the expected time for MEDEVAC unit i to return to base after delivering the casualties of its current call (call j) to the nearest MTF. This computation relies on l_j , the POI location of the call MEDEVAC unit i is currently serving, to identify the destination MTF and compute the expected travel time from the MTF to the MEDEVAC unit's base. The expected time $d_{i,j}^r$ is added to the $d_i - \tau$ term, the expected time taken for the MEDEVAC unit to complete its current mission.

The third basis function captures the expected time from the current event time τ until MEDEVAC unit i arrives at the nearest MTF with its assigned casualties. Let

$$\phi_{3,i} = \begin{cases} d_i - \tau & \text{if } \sigma_i \neq \text{"idle"} \\ 0 & \text{otherwise.} \end{cases} \quad i = 1, 2, \dots, a \quad (12)$$

The next three basis functions capture information about the calls in the queue and the calls currently being served. Basis function $\phi_{4,i,j}$ captures the expected total time, including all travel, load, and unload times, that call j will be in the system if it is served by MEDEVAC unit i . Let

$$\phi_{4,i,j} = \begin{cases} \zeta_j + d_i - t_i & \text{if } \sigma_i = \text{"enroute"} \\ & \text{to call } j \\ 0 & \text{otherwise.} \end{cases} \quad i = 1, 2, \dots, a; \quad j = 1, 2, \dots, b \quad (13)$$

Basis function $\phi_{5,i,j}$ captures the priority of the call (call j) being served by MEDEVAC i . Let

$$\phi_{5,i,j} = \begin{cases} \eta_j & \text{if } \sigma_i = \text{"enroute"} \\ & \text{to call } j \\ 0 & \text{otherwise.} \end{cases} \quad i = 1, 2, \dots, a; \quad j = 1, 2, \dots, b \quad (14)$$

Basis function $\phi_{6,i,j}$ represents the expected time in system for call j if it is assigned to MEDEVAC i . Let

$$\phi_{6,i,j} = d_{i,j}^s + \phi_{2,i} \quad i = 1, 2, \dots, a; \quad j = 1, 2, \dots, b, \quad (15)$$

where $d_{i,j}^s$ represents the expected MEDEVAC service time for MEDEVAC unit i to serve call j . We utilize polynomial functions comprised of these basis functions to approximate the post-decision state value function.

Having stipulated the policy function and the value function approximation upon which it is based, we proceed by discussing the manner in which the value function approximation is updated. We employ an API algorithmic strategy, the general structure of which is similar to exact policy iteration. Within the policy evaluation step of our API algorithm, we update the value function approximation for a fixed policy using least squares temporal differencing (LSTD). Introduced by [Bradtke and Barto \(1996\)](#), LSTD is a computationally efficient method for estimating the adjustable parameters (e.g., θ), when using a linear architecture with fixed basis functions to approximate the value function for a fixed policy. The API algorithm we employ is adapted in part from [Scott et al. \(2014\)](#) and is shown in Algorithm 1.

The algorithm begins with an initial θ vector, representing an initial base policy. The performance evaluation of the current policy proceeds as follows. A post-decision state is randomly sampled, and the value $\phi(s_{k-1,h}^x)$ is recorded. Next, we simulate one event forward and determine the best decision as per Equation (9), recording the associated expected reward $R(s_{k,h}, x)$ and basis function evaluations of the post-decision state, $\phi(s_{k,h}^x)$. A total of $\mathcal{H}$ temporal difference sample realizations are collected, where $R(s_{k,h}, \pi(s_{k,h}; \theta)) + \gamma \theta^\top \phi(s_{k,h}^x) - \theta^\top \phi(s_{k-1,h}^x)$ is the h th temporal difference, given the parameter vector θ .

After obtaining the $\mathcal{H}$ temporal difference sample realizations, we conduct the policy improvement steps of the API algorithm. We compute $\hat{\theta}$, a sample estimate of θ , by regressing the $h = 1, \dots, \mathcal{H}$ basis function evaluations of the post-decision states $\phi(s_{k-1,h}^x)$ and $\phi(s_{k,h}^x)$ against the reward $R(s_{k,h}, x)$. We

---

Algorithm 1 Approximate Policy Iteration Algorithm

```

1: Initialize  $\theta$ 
2: for  $g=1$  to  $\mathcal{G}$  (Policy Improvement Loop)
3:   for  $h=1$  to  $\mathcal{H}$  (Policy Evaluation Loop)
4:     Simulate a random post-decision state,  $s_{k-1,h}^x$ 
5:     Record  $\phi(s_{k-1,h}^x)$ 
6:     Simulate transition to next event, obtain pre-decision state  $s_{k,h}$ 
7:     Determine decision  $x=\pi(s_{k,h};\theta)$  using Eq. (9)
8:     Record  $R(s_{k,h},x)$ 
9:     Record  $\phi(s_{k,h}^x)$ 
10:   end for
11:   Update  $\theta$  using Eq. (16) and Eq. (18)
12: end for
13: End

```

---

conduct a least squares regression so that the sum of the temporal differences over the $\mathcal{H}$ inner loop simulations (which approximates the expectation) is equal to zero. To compactly state our update of θ , we denote basis function matrices and reward vectors as follows. Let

$$\Phi_{k-1} \triangleq \begin{bmatrix} \phi(s_{k-1,1}^x)^\top \\ \vdots \\ \phi(s_{k-1,\mathcal{H}}^x)^\top \end{bmatrix}, \quad \Phi_k \triangleq \begin{bmatrix} \phi(s_{k,1}^x)^\top \\ \vdots \\ \phi(s_{k,\mathcal{H}}^x)^\top \end{bmatrix}, \quad R_k \triangleq \begin{bmatrix} R(s_{k,1}) \\ \vdots \\ R(s_{k,\mathcal{H}}) \end{bmatrix},$$

where matrices Φ_{k-1} and Φ_k consist of rows of basis function evaluations of the sampled post-decision states, and R_k is the reward vector for the sampled events. The parameter vector $\hat{\theta}$ is then computed as follows.

$$\hat{\theta} = \left[(\Phi_{k-1} - \gamma\Phi_k)^\top (\Phi_{k-1} - \gamma\Phi_k) \right]^{-1} (\Phi_{k-1} - \gamma\Phi_k)^\top R_k \quad (16)$$

We employ a generalized harmonic stepsize rule to smooth in the new observation $\hat{\theta}$ with the previous estimate θ . The stepsize rule is given by

$$\alpha_g = \frac{\kappa}{\kappa + g - 1}. \quad (17)$$

The stepsize rule α_g greatly influences the rate of convergence of the algorithm and therefore impacts the attendant solutions. Increasing the stepsize parameter κ slows the rate at which the smoothing stepsize drops to zero. Selecting an appropriate value for κ requires understanding the rate of convergence of the application. Some problems allow convergence to good solutions fairly quickly (tens to hundreds of iterations) whereas others require much more effort (thousands of iterations). We observe that relatively lower κ -values work best in our application.

The update equation for θ is given by

$$\theta \leftarrow \theta(1 - \alpha_g) + \hat{\theta}(\alpha_g). \quad (18)$$

On the right hand side of Equation (18), θ is the previous estimate and is based on information from all previous outer loop iterations; $\hat{\theta}$ is our estimate from the current outer loop iteration. As the number of iterations g increases, we place less emphasis on any single estimate and more emphasis on the estimate based on information from the first $g - 1$ iterations.

Upon obtaining an updated (and smoothed) parameter vector θ , we have completed one policy improvement iteration of the algorithm. The parameters $\mathcal{G}$ and $\mathcal{H}$ are tunable, where $\mathcal{G}$ is the number of policy improvement iterations completed and $\mathcal{H}$ is the number of policy evaluation iterations completed.

4.3. Simulation

This section presents the MEDEVAC process simulation model that supports the design and development of high quality MEDEVAC dispatch policies. The simulation model enables evaluation of policies identified by our ADP algorithm as well as the evaluation of the myopic policy typically utilized in practice by the aeromedical evacuation coordination center (AECC). For the myopic policy, MEDEVAC requests are served in decreasing order of priority with a first-in-first-out policy for like priorities. Note that we also employ the simulation model within the ADP algorithm itself when we randomly sample a post-decision state and simulate forward one system event.

The simulation model we develop encompasses the operations of a MEDEVAC system, incorporating information obtained from military doctrine (Department of the Army, 2007), interviews with US Army Blackhawk helicopter pilots (Fett, 14 August 2013; Flores, 16 August 2013), and our own personal experiences. In the representative scenario considered in the subsequent section, placement of the MTFs follows doctrine. Moreover, the deliberate selection of the casualty cluster centers within the Hawkes process governing casualty event arrivals embeds the information concerning the expected interaction of friendly and adversary forces.

We utilize the MEDEVAC process simulation within our ADP algorithm to obtain simulated trajectories of the system and to measure the performance of the MEDEVAC dispatch policies obtained from our ADP algorithm. The flow of control for the MEDEVAC process simulation model is shown in Figure 2. We employ an event-scheduling approach to simulation modeling, using a next-event time-advance mechanism.

Figure 2: Simulation Flow Chart

In the initialization routine, the simulation is initiated by setting the simulation clock to zero, setting the system state to all MEDEVAC units sitting “idle” at their own base and an empty call queue, and setting statistical counters to zero. Statistical counters refer to the variables used to store system performance information such as accumulated rewards, casualty event response times, and MEDEVAC utilization (busy) rates. The event list is initiated with a casualty event. A Hawkes process with arrival rate λ governs casualty event generation. The Hawkes process models situations where subsequent events are likely to occur in close proximity to the prior event. In the timing routine, the next event (i.e., event k) of type ϵ , is selected from the event list and the simulation clock is advanced to the appropriate simulation time. In the event type ϵ routine, a dispatch decision is made according to the stipulated MEDEVAC dispatch policy. The system state and statistical counters are updated accordingly. We generate future events and add them to the event list. To obtain these future events, random variates capturing the MEDEVAC mission travel, load, and unload times are generated from probability distributions depicting the real-world MEDEVAC processes. We then check to see if the simulation is over, terminating the simulation at the 24-hour mark. If the simulation continues, we return to the timing routine to determine which event on the event list occurs next and the simulation proceeds.

5. Computational Results

In this section, we demonstrate the applicability of our model to a problem of interest to the military medical planning community and examine the efficacy of our proposed solution methodology. We present a representative military medical evacuation (MEDEVAC) planning scenario as the basis for our analysis. We examine different features of the MEDEVAC dispatching problem and of our approximate dynamic programming (ADP) algorithm to gain insight regarding the performance of our overall proposed approach. In particular, a computational experiment is designed and conducted to obtain insights regarding solution quality and computational effort. Moreover, we perform a series of sensitivity analyses to focus our study on specific problem features and algorithmic features of interest. For the computational experiments, we utilize a dual Intel Xeon E5-2650v2 workstation having 192 GB of RAM and MATLAB’s parallel computing toolbox.

5.1. Representative Scenario

We consider a notional planning scenario in which a coalition of allied countries performs peacekeeping operations in response to aggression by Islamic State militants in northern Syria. The locations for MEDEVAC bases are likely key military tactical sites. Casualty event cluster centers for use within the Hawkes process are selected and weighted by projected enemy locations. Figure 3 shows the 18 casualty cluster centers, five MEDEVAC stations, and two medical treatment facilities (MTFs). Each MTF has a MEDEVAC station co-located with it.

Figure 3: Representative Scenario Disposition

We assume a high operations tempo with a baseline casualty event arrival rate of $\lambda = \frac{1}{60}$, representing an average casualty event arrival rate of one event per hour. Military medical planners in cooperation with military intelligence and operational planners should be able to stipulate a reasonable arrival rate in advance of a planned combat operation based on the size, equipment, and disposition of friendly and adversary forces. Any MEDEVAC is allowed to service any casualty event, and casualty events need not be served as soon as they arrive. In this scenario, we assume *urgent* and *priority* class casualty events arrive in equal proportion. *Routine* events are not considered due to the high operations tempo. In such scenarios, *routine* events would be serviced by non-MEDEVAC or ground MEDEVAC assets. The reward function utilizes $\beta_1 = 10$ and $\beta_2 = 2$ which rewards the servicing of *urgent* casualty events much more than *priority* casualty events. We further assume the UH-60 A/L Blackhawk acts as the dedicated aeromedical MEDEVAC asset.

5.2. Experimental Design

In our computational experiments, we measure performance of an ADP policy in terms of the percent increase in total discounted reward over that obtained by use of the default myopic policy. When using the myopic policy, the system dispatches the closest idle MEDEVAC unit. An important problem feature of interest is the casualty event arrival rate λ . The algorithmic features of interest

Table 1: Experimental Design Factor Levels

λ	$\mathcal{G}$	$\mathcal{H}$
$\frac{1}{30}$	5	5,000
$\frac{1}{60}$	10	10,000
$\frac{1}{120}$	20	20,000

include the number of policy improvement iterations $\mathcal{G}$ and the number of policy evaluation iterations $\mathcal{H}$. In our experimental design we utilize a 3rd order polynomial of the basis functions to approximate the post-decision state value function. We select algorithmic factor levels based on our initial experiences with implementing the algorithm. We observe that computational effort beyond $\mathcal{G} = 20$ policy iterations obtains little improvement. Similarly, after $\mathcal{H} = 20,000$ temporal difference samples, we observe little improvement in solution quality.

During our initial investigations, we also explored the issue of whether to employ smoothing for the θ parameter. In our experiments, smoothing outperforms non-smoothing. Figure 4 illustrates a typical observed result. Indeed, we observed such a result across all experimental levels and for all runs. As such, we utilize an aggressive smoothing function (i.e., $\kappa = 1$) when implementing our algorithm and do not report further results concerning smoothing.

Figure 4: ADP Performance: Smoothing vs. Non-Smoothing

Informed by our initial investigations, we design a 3^3 full factorial computational experiment to examine solution quality and computation time at different levels for factors λ , $\mathcal{G}$, and $\mathcal{H}$. Table 1 shows the factor levels. The purpose of the experiment is to inform the selection of appropriate algorithm parameter values for $\mathcal{G}$ and $\mathcal{H}$ for subsequent sensitivity analysis. Moreover, we seek general insight regarding the performance of the ADP algorithm for MEDEVAC dispatching problem instances with decreased and increased intensity in terms of the casualty event arrival rate, λ .

5.3. Experimental Results

Table 2 reports the results from the experimental design. Starting from the left, the first three columns indicate the factor levels, the fourth column indicates the 95% confidence interval for ADP algorithm performance, expressed in terms of percentage improvement over the myopic policy (with respect to total discounted reward), the fifth and sixth columns report the mean response times (i.e., length of time from a 9-line call submission until casualties are dropped off at an MTF) experienced by *urgent* and *priority* calls, respectively, the seventh column reports MEDEVAC unit busy rates, and the eighth column reports computation times.

Table 2: Experimental Design Results

$\frac{1}{\lambda}$	Outer $\mathcal{G}$	Inner $\mathcal{H}$	Improvement Over Myopic	Urgent Response Time (mins)	Priority Response Time (mins)	MEDEVAC Busy	Computation Time (secs)
30	5	5000	48.8% $\pm$ 0.9%	182.5	398.0	90.7%	18.9
30	5	10000	57.8% $\pm$ 0.9%	187.9	375.0	90.4%	37.5
30	5	20000	51.4% $\pm$ 0.9%	177.8	397.2	90.6%	74.8
30	10	5000	51.7% $\pm$ 1.0%	186.8	388.2	90.6%	37.3
30	10	10000	61.9% $\pm$ 0.9%	191.7	369.3	88.7%	74.8
30	10	20000	59.6% $\pm$ 0.9%	190.5	369.7	89.6%	150.3
30	20	5000	52.8% $\pm$ 0.9%	190.8	377.4	90.5%	74.8
30	20	10000	50.4% $\pm$ 0.9%	186.7	386.5	90.6%	149.9
30	20	20000	53.7% $\pm$ 0.9%	192.0	373.5	90.5%	300.2
60	5	5000	24.2% $\pm$ 1.0%	170.0	234.3	84.1%	19.2
60	5	10000	26.8% $\pm$ 1.1%	166.7	232.8	83.2%	37.8
60	5	20000	25.3% $\pm$ 1.0%	167.0	240.4	84.2%	75.0
60	10	5000	24.9% $\pm$ 1.1%	170.4	230.1	83.1%	37.2
60	10	10000	27.0% $\pm$ 1.1%	168.2	232.0	82.4%	74.9
60	10	20000	30.8% $\pm$ 1.1%	154.0	235.4	81.0%	149.9
60	20	5000	26.8% $\pm$ 1.0%	177.1	265.4	80.1%	74.7
60	20	10000	27.1% $\pm$ 1.0%	167.9	230.8	82.0%	150.6
60	20	20000	28.2% $\pm$ 1.0%	165.6	237.8	81.4%	300.1
120	5	5000	1.1% $\pm$ 1.4%	115.2	120.3	34.8%	19.2
120	5	10000	-0.6% $\pm$ 1.5%	116.6	121.7	35.0%	37.8
120	5	20000	0.7% $\pm$ 1.6%	113.3	117.5	32.8%	75.3
120	10	5000	-5.3% $\pm$ 1.6%	121.5	132.3	36.4%	37.5
120	10	10000	0.0% $\pm$ 1.5%	113.9	121.0	32.2%	74.6
120	10	20000	-0.7% $\pm$ 1.5%	117.2	123.7	35.4%	149.6
120	20	5000	-4.9% $\pm$ 1.6%	117.9	127.0	33.8%	75.4
120	20	10000	-3.1% $\pm$ 1.5%	119.3	145.4	32.4%	150.1
120	20	20000	-2.1% $\pm$ 1.5%	116.2	129.2	32.7%	300.6

We highlight the best performing algorithmic features with respect to improvement for event arrival rates $\lambda = \frac{1}{30}$ and $\lambda = \frac{1}{60}$. For a casualty event arrival rate of $\lambda = \frac{1}{120}$, the ADP policy generally performs the same as the myopic policy. In a sense, this is an intuitive result. As the likelihood of an event arrival decreases, it is less risky to simply send the nearest MEDEVAC unit to service a call, regardless of its priority category. As such, the myopic policy is best, and the ADP policy will not outperform it. As the casualty event arrival rate λ increases, the ADP policy increasingly outperforms the myopic policy across all investigated algorithmic experimental factor levels. The ADP policy

successfully captures important nuances of the problem, whereby MEDEVAC rationing occurs in anticipation of the arrival of *urgency* calls. A general trend observed for the first two casualty event arrival rates is that $\mathcal{G} = 10$ policy improvement loops and at least $\mathcal{H} = 10,000$ policy evaluation loops achieve the best performance. We observe that computation time scales closely with the product $\mathcal{G} \cdot \mathcal{H}$, as expected.

The experimental results indicate that setting $\mathcal{G} = 10$ and $\mathcal{H} = 20,000$ is appropriate for subsequent investigations. We proceed by investigating our baseline case more thoroughly.

5.4. Baseline Scenario Results and Analysis

Table 3: Baseline Scenario Parameters

Parameter	Problem Features	
	Description	Setting
λ	Casualty Event Arrival Rate	$\frac{1}{60}$
β_1	Weight for <i>urgent</i> event	10
β_2	Weight for <i>priority</i> event	2
Parameter	Algorithmic Features	
	Description	Setting
$\mathcal{G}$	Policy Improvements	10
$\mathcal{H}$	Policy Evaluations	20,000

We utilize the problem and algorithmic parameters shown in Table 3 when investigating the baseline scenario. The medial casualty event arrival rate of $\lambda = \frac{1}{60}$ is selected as the base case arrival rate, as it models a high intensity conflict scenario of interest to military medical planners. The weights $\beta_1 = 10$ and $\beta_2 = 2$ capture the medical planners' strong preference for servicing *urgent* casualty events over *priority* casualty events. We select $\mathcal{G} = 10$ and $\mathcal{H} = 20000$ as they are the best performing algorithmic parameters for $\lambda = \frac{1}{60}$.

Table 4 shows the performance of the ADP policy against the myopic policy for the baseline scenario. We consider basis function polynomials up to 3rd order. We stopped at the 3rd order because 4th order polynomials caused singularity issues when performing matrix operations to compute θ , indicating a linear dependency issue among the basis function vectors. Utilizing a 3rd order polynomial of basis functions to approximate the value of a post-decision state obtains the best ADP policy performance, outperforming the myopic policy by nearly 31%. The mean casualty event response times for the two priority classification levels and the mean MEDEVAC *busy* times (which include time spent traveling back to base even if they are not actively serving a casualty event) also outperform the myopic policy. The ADP policy focuses on servicing *urgent* casualty events first, as noted by shorter wait times; moreover, MEDEVACs are being utilized more efficiently as shown by the lower average busy percentage.

Table 4: Baseline Scenario Performance

Policy	Number of Basis Functions	Improvement Over Myopic	<i>Urgent</i> Response Time (mins)	<i>Priority</i> Response Time (mins)	MEDEVAC Busy %
Myopic	-	-	285.4	286.7	89.3%
ADP, 1st Order	160	17.7% $\pm$ 1.4%	167.0	299.3	87.8%
ADP, 2nd Order	320	26.0% $\pm$ 1.5%	165.5	248.9	82.3%
ADP, 3rd Order	480	30.8% $\pm$ 1.1%	154.0	235.4	81.0%

A discussion concerning the *urgent* and *priority* response times is warranted. Even when implementing the best ADP policy, our simulation results indicate a mean response time of 154.0 minutes for *urgent* evacuations. While the standard stipulated by the United States Secretary of Defense is 60 minutes, we note that the North Atlantic Treaty Organization (NATO) standard is 90 minutes (Cordell et al., 2008). Moreover, the Department of the Navy adheres to a standard of 120 minutes, as indicated in the field medical service technician student manual definitions (Department of the Navy, 2013). Regardless of the desired goal, if the high dispatching response times are unacceptable to military medical planners and their commanding officers, then the placement of additional MTFs and/or more MEDEVAC units should be considered. Conversely, medical asset resource allocation and placement decisions, determined prior to the dispatching problem analysis, may have to be revisited. Analysis of the θ vector can inform such follow-up decisions.

Examination of the basis function θ coefficients from the best performing 3rd order ADP policy provides the following insights. The basis function ϕ_5 , which captures the interaction between MEDEVAC unit i and the priority of the casualty event it is serving, has the largest impact on the policy. This result is indicated by the relatively larger positive values of the θ coefficients associated with ϕ_5 , as compared to the θ coefficients of the other basis functions. The number and proximity of casualty event cluster centers to a MEDEVAC base show an increase in magnitude for their respective basis function θ coefficients. Moreover, MEDEVAC units co-located with the MTF show a similar increase in magnitude, despite being further away from large groupings of casualty event cluster centers. MEDEVAC units with low θ coefficient magnitudes (for their associated basis functions) may be candidates for permanent relocation, should medical planners consider such an issue. Such MEDEVAC units contribute very little to the value of being in a particular system state and may be more productive if stationed at a different location. MEDEVAC units located at more productive sites have higher θ -values (for their associated basis functions) since operating from their location (and attendant ability to service nearby 9-line calls) provides value to the system (as indicated by positive, higher θ -values). Interestingly, ϕ_1 shows low statistical significance and also has small magnitudes for its coefficients for all MEDEVACs. This occurs because the ADP policy does not seek to penalize MEDEVACs for being idle, which could force them to otherwise launch in response to low priority and/or distant casualty events. Examination of the superlative ADP policy indicates that it is best to launch MEDEVACs which are close to casualty events and co-located with hospitals only on *urgent* casualty events and to hold these units in reserve when lower

priority casualty events occur.

The benefits of the ADP policy become more evident upon considering two vignettes (i.e., possible system states) in which the myopic and the ADP policies differ and recognizing the advantage afforded by the ADP in those vignettes. Consider the following *reserve vignette* in which a *priority*-level casualty event occurs in a high-demand area, after which an *urgent*-level event occurs in the same area. Initially, all MEDEVACs are idle. Under the myopic policy, the closest MEDEVAC is dispatched in response to the first (*priority*) event. Subsequently, the next-closest (yet, somewhat distant) MEDEVAC is dispatched in response to the second (*urgent*) event. The system receives a lower reward when responding to the *urgent* event due to the longer response time of the next-closest MEDEVAC. Under the ADP policy, the next-closest MEDEVAC is dispatched in response to the first (*priority*) event as dispatching the closest (in-area) MEDEVAC would place the system in a disadvantageous position, as indicated by the value function approximation. The closest MEDEVAC is dispatched in response to the second (*urgent*) event. The system receives a higher reward when responding to the *urgent* event due to the closest MEDEVAC's shorter response time. In dispatching the next-closest MEDEVAC to the first (*priority*) event, the closest MEDEVAC is essentially held in reserve in anticipation of servicing the likely imminent arrival of an *urgent* casualty event. When the *urgent* casualty event then occurs, the system is well positioned to dispatch the closest MEDEVAC and earn a higher reward.

Consider the following *wait vignette* in which a casualty event occurs in a high-demand area, after which a second casualty event occurs in a distant high-demand area. Initially, all MEDEVACs are busy with the exception of the MEDEVAC in the distant high-demand area. Moreover, the closest MEDEVAC (in the first event's area) is about to complete service, prior to the arrival of the second event. Under the myopic policy, the distant MEDEVAC is dispatched in response to first event. After the closest MEDEVAC completes its service, it is dispatched in response to the second event. Both calls must wait for a distant MEDAVAC, resulting in longer than necessary response times and attendant low rewards. Indeed, the system realizes a low reward with a high utilization rate. Under the ADP policy, the dispatching authority waits and queues the first call because the decision to dispatch the distant MEDEVAC in response to the first event would place the system in a disadvantageous position, as indicated by the value function approximation. The ADP policy accounts for the fact that the closest MEDEVAC is about to complete its current call and will then become available. After waiting, the closest MEDEVAC completes service and is then immediately dispatched to service the queued first call. The distant MEDEVAC initially remains idle and is dispatched in response to the second event. The system realizes a higher reward with a lower utilization rate, as compared to the myopic policy.

We next examine the performance of the ADP policy for different casualty event arrival rates by adjusting λ . Figure 5 shows the performance of the ADP policy against the myopic policy. As λ decreases (to the right in Figure 5), the frequency with which casualties arrive decreases and the ADP policy no longer

Figure 5: Sensitivity of Performance to Casualty Event Arrival Rate

outperforms the myopic policy. As casualty events arrive at a slower rate, the reward obtained by holding MEDEVAC aircraft in reserve is diminished. At a casualty event arrival rate of one every two hours (i.e., $\lambda = \frac{1}{120}$), we observe a reduction in performance of the ADP compared to the myopic policy, perhaps showing the limitation of the set of basis functions we have defined and their resulting value function approximations. Alternatively, as mentioned previously, as λ approaches zero, the difference in performance between the optimal and myopic policies becomes negligible; for such low intensity operations we would not expect to see a difference in performance between the ADP policy and the myopic policy.

Figure 6: Sensitivity of Performance to Proportion of *Urgent* Casualty Events

The impact of changing the proportion of *urgent* casualty events is shown in Figure 6. The ADP policy outperforms the myopic policy for all proportions, increasingly so when there are more frequent *urgent* casualty events occurring.

The ADP policy efficiently manages resources by not immediately sending the nearest available MEDEVAC, but rather waiting and/or dispatching a MEDEVAC unit that is further away from the casualty event and that is also not close to a high rate casualty cluster area.

Our final investigation considers the impact of replacing the UH-60 A/L Blackhawk with an improved aeromedical aircraft. There are experimental rotary wing aircraft which could potentially be put into service and which can travel significantly faster than the UH-60 A/L Blackhawk. To examine the impact of these aircraft on MEDEVAC system performance, we adjust the average speed at which the MEDEVAC aircraft can travel. Other random variables modeling the MEDEVAC process remain the same. Table 5 shows the results obtained by increasing average airspeed, where airspeed is indicated as a percentage increase over the airspeed obtained by the current dedicated aeromedical asset, the UH-60 A/L Blackhawk.

Table 5: Increased Airspeed Scenario Performance

Myopic Policy					
Airspeed Increase	Improvement Over Baseline	<i>Urgent</i> Response Time (mins)	<i>Priority</i> Response Time (mins)	MEDEVAC Busy	
Baseline	-	285.4	286.7	89.3%	
25%	50.1%	168.5	166.8	78.6%	
50%	89.5%	104.8	104.1	61.6%	
75%	126.1%	77.2	76.8	47.8%	

ADP Policy					
Airspeed Increase	Improvement Over Baseline	<i>Urgent</i> Response Time (mins)	<i>Priority</i> Response Time (mins)	MEDEVAC Busy	
Baseline	30.8%	154.0	235.4	81.0%	
25%	68.3%	116.8	140.3	69.6%	
50%	102.8%	90.8	103.4	56.6%	
75%	127.1%	74.8	81.5	46.9%	

It is reasonable to assume that new rotary wing aircraft designs can achieve increased average airspeeds of 25% to 50% compared to the currently fielded UH-60 A/L Blackhawk, which uses a power plant designed prior to 1989 ([Federation of American Scientists, 2015](#)). This increase in speed has a significant impact on the overall performance for both ADP and myopic policies. When implementing the ADP policy, an increase in airspeed of 25% results in mean response times of 116.8 minutes and 140.3 minutes for *urgent* and *priority* casualty events, respectively. The reduction in mean response time from 235.4 minutes to 140.3 minutes is quite substantial for the *priority* class. Although we observe diminishing returns for the ADP compared to the myopic policy as airspeed increases (i.e., increased airspeed helps the myopic policy more than it helps the ADP policy), the MEDEVAC dispatching authority still benefits greatly from high quality dispatching solutions provided by the ADP algorithm.

When we examine computational effort more closely, we find the ADP so-

lution only requires about 150 seconds run time for the baseline scenario. The largest computational effort comes from running the simulation to obtain a reasonable measure of performance. For each event in the simulation, the best action must be determined, as indicated by the policy function (Equation (9)). In addition to these computations, we require 500 runs to achieve our desired level of confidence. The computation times for the simulation using the myopic and 3rd order ADP policies are shown in Table 6.

Table 6: Simulation run times

Policy	Computation Time (secs)	Runs
Myopic	34.6	500
ADP	884.9	500

The computation time required for computing the decision using Equation (9) is about 25 times that of the myopic policy. Despite this burden, it is still possible to compute the ADP policy and run the simulation in under 20 minutes. Moreover, once a satisfactory ADP policy is obtained, MEDEVAC dispatch decisions can be computed nearly instantaneously. These results are promising, as our proffered approach could readily be adapted and applied in current operational planning situations to yield timely results and inform military medical planning decisions.

6. Conclusions

This paper examines the medical evacuation (MEDEVAC) dispatching problem. The intent of the research is to determine policies that improve the performance of a deployed military emergency medical service system and ultimately to increase the survivability of battlefield casualties. Development of a Markov decision process (MDP) model of the MEDEVAC dispatching problem enables examination of many different military medical planning scenarios. Solving the MDP requires the design, development, and implementation of an approximate dynamic programming (ADP) algorithm. We develop an approximate policy iteration algorithm that utilizes least squares temporal difference learning for policy evaluation. We define a set of basis functions within a linear approximation architecture to approximate the value function around the post-decision state. To demonstrate the applicability of our MDP model and to examine the efficacy of the policies produced by our ADP algorithm, we construct a notional, representative military medical planning scenario based on contingency operations in northern Syria.

The ADP policy is able to increase overall expected total discounted reward by nearly 31% compared to the myopic policy in our baseline scenario. Moreover, MEDEVAC *busy* time was decreased by over 8%, indicating a more efficient utilization of MEDEVAC aircraft. Examination of basis function coefficients determined by our ADP algorithm revealed MEDEVAC aircraft in close proximity to higher likelihood casualty clusters are more valuable than aircraft

based further away. This is an intuitive result. These higher value MEDEVAC units should likely not be dispatched for low priority casualty events when the lower value MEDEVACs may be dispatched instead. The ADP approximation architecture is able to capture the overall time it would take for any MEDEVAC unit to service any casualty event. This is an important result, as a MEDEVAC unit may become available that could service a casualty event faster than an *idle* MEDEVAC that is further away. The average cruise speed of the MEDEVAC aircraft greatly impacts system performance. Results indicate that a 25% increase in speed increases expected total discounted reward by 50%. Even with such increased aircraft speed greatly improving myopic policy performance, the ADP policy still outperforms the myopic policy.

The research presented in this paper is of interest to military medical planners and relevant combatant command dispatch authorities. Military medical planners can apply this model and solution approach to compare policies for planning scenarios with fixed medical treatment facility (MTF) and MEDEVAC station locations. Moreover, planners may utilize the model to evaluate different potential MTF and MEDEVAC station allocation and location schemes to maximize overall MEDEVAC system performance. Military staffs can leverage current military intelligence and operational experiences to identify areas in which casualties are likely to occur. Once these areas are identified, planners can make informed decisions about the value of each MEDEVAC unit and ultimately achieve greater utilization of resources. Our results indicate the criticality of the MEDEVAC platform’s travel speed. Military medical planners and acquisition specialists (i.e., those personnel responsible for implementing new technology into the military) can leverage this model to examine the impact of speed and capacity. This comparison informs the future design and development of a replacement for the UH-60 A/L. Perhaps a mix of large capacity UH-60 A/L’s and a new lower-capacity, high-speed design would improve overall casualty survivability.

One limiting assumption made in this paper is that MEDEVAC units must return to their own station to refuel and re-equip after delivering patients to the MTF prior to servicing a queued call. Some bases may have both an MTF and a co-located MEDEVAC station. On such bases, it is reasonable for MEDEVACs based elsewhere to refuel after delivering patients to the MTF and then immediately proceed to service a queued call. Our restriction for MEDEVACs to return to their own bases to refuel before servicing a subsequent demand represents a simplifying assumption for this initial effort. Modifying our problem formulation and the corresponding MDP model to allow the refueling and dispatch of MEDEVACs immediately after patient delivery at an MTF with a co-located helicopter base is intended as a future extension to this work.

Other extensions to our approach include the explicit modeling of the number of casualties at the casualty event as well as the capacity of the MEDEVAC platform. One could then investigate the trade-offs associated with employing a smaller capacity, but faster MEDEVAC platform versus the currently fielded larger capacity, but slower MEDEVAC platforms. Two important aspects pertaining to the overall design of a deployed MEDEVAC system also remain of

interest: determining the number of MEDEVAC units and selecting the home stations for each unit. Examining the impact of these two issues on the best dispatching policy and demonstrating the advantage of the ADP policy over the myopic policy when making such decisions would be a valuable future contribution.

Implementing an ADP policy into active dispatch operations at the aeromedical evacuation coordination center (AECC) is a difficult proposition. The myopic policy is often used in practice because it is simple to implement and performs fairly well as long as casualty events arrive less frequently. In the representative scenario we analyzed, the ADP policy and the myopic policy performed equally well when casualty events arrived every two hours. However, when casualty events arrived more frequently, the myopic policy underperformed. The important point for the military medical planning community is recognizing that a myopic policy is not always best and understanding that there are operations research techniques available that may improve MEDEVAC system performance. Yet, important challenges remain. Although this research demonstrates the potential improvement of an ADP policy over a myopic policy, it is subject to challenges concerning changing battlefield conditions. Indeed, although a MEDEVAC dispatching decision can be computed nearly instantaneously using a particular ADP policy, if the assumptions embedded within the MDP model change (e.g., with respect to the casualty event arrival process or the locations and numbers of MEDEVAC units and MTFs), then a new ADP policy must be determined. Recognizing that a new ADP policy must be determined and then actually conducting an analysis to determine the new ADP policy would pose a great challenge for AECC personnel. Further investigation on this matter will be examined in a sequel to this research.

Acknowledgments

The views expressed in this paper are those of the authors and do not reflect the official policy or position of the United States Army, United States Air Force, Department of Defense, or the United States Government. The authors are grateful to the United States Army Medical Evacuation Proponency Directorate for its encouragement and feedback on this line of research. The authors would like to thank the coordinating editor and two anonymous referees for their insightful comments and suggestions that resulted in an improved manuscript.

References

- Bandara, D., Mayorga, M. E., McLay, L. A., 2012. Optimal dispatching strategies for emergency vehicles to increase patient survivability. *International Journal of Operational Research* 15 (2), 195–214.
- Bandara, D., Mayorga, M. E., McLay, L. A., 2014. Priority dispatching strategies for EMS systems. *Journal of the Operational Research Society* 65 (4), 572–587.

- Bastian, N. D., 2010. A robust, multi-criteria modeling approach for optimizing aeromedical evacuation asset emplacement. *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology* 7 (1), 5–23.
- Bertsekas, D. P., 2011. Approximate policy iteration: A survey and some new methods. *Journal of Control Theory and Applications* 9 (3), 310–335.
- Bertsekas, D. P., Tsitsiklis, J. N., 1996. *Neuro-Dynamic Programming*. Athena Scientific.
- Bradtke, S. J., Barto, A. G., 1996. Linear least-squares algorithms for temporal difference learning. *Machine Learning* 22 (1-3), 33–57.
- Carter, G. M., Chaiken, J. M., Ignall, E., 1972. Response areas for two emergency units. *Operations Research* 20 (3), 571–594.
- Cordell, B. R., Cooney, M., Beijer, D., 2008. Audit of the effectiveness of command and control arrangements for medical evacuation of seriously ill or injured casualties in Southern Afghanistan 2007. *Journal of the Royal Army Medical Corps* 154 (4), 227–230.
- Department of the Army, U. S., 2007. Field Manual 4-02.2, Medical Evacuation.
- Department of the Navy, U. S., 2013. Field Medical Training Battalion Student Manual.
- Eastridge, B. J., Mabry, R. L., Seguin, P., Cantrell, J., Tops, T., Uribe, P., Mallett, O., Zubko, T., Oetjen-Gerdes, L., Rasmussen, T. E., et al., 2012. Death on the battlefield (2001–2011): implications for the future of combat casualty care. *Journal of Trauma and Acute Care Surgery* 73 (6), S431–S437.
- Egesdal, M., Fathauer, C., Louie, K., Neuman, J., Mohler, G., Lewis, E., 2010. Statistical and stochastic modeling of gang rivalries in Los Angeles. *SIAM Undergraduate Research Online* 3, 72–94.
- Erkut, E., Ingolfsson, A., Erdoğan, G., 2008. Ambulance location for maximum survival. *Naval Research Logistics (NRL)* 55 (1), 42–58.
- Federation of American Scientists, 2015. UH-60 Black Hawk. “<http://fas.org/man/d0d-101/sys/ac/uh-60.htm>”, online, accessed 15-July-2015.
- Feero, S., Hedges, J. R., Simmons, E., Irwin, L., 1995. Does out-of-hospital ems time affect trauma survival? *The American Journal of Emergency Medicine* 13 (2), 133–135.
- Fett, G., 14 August 2013. Personal communication. Major Fett, US Army, has served in MEDEVAC and Air Assault units, amassing over 900 hours of flight time in the Army UH-60 Blackhawk helicopter, including over 300 combat hours of flight time conducting missions during two one-year deployments in support of Operation Iraqi Freedom and Operation New Dawn.

- Fitch, J., 2005. Response times: myths, measurement, and management. *Journal of Emergency Medical Services* 30 (9), 47–56.
- Fitzsimmons, J. A., 1973. A methodology for emergency ambulance deployment. *Management Science* 19 (6), 627–636.
- Flores, C., 16 August 2013. Personal communication. Major Flores, US Army, has served in Army Aviation units, amassing over 600 hours of flight time in the Army UH-60 Blackhawk helicopter, including over 350 combat hours of flight time conducting missions during two one-year deployments in support of Operation Iraqi Freedom.
- Fulton, L. V., Lasdon, L. S., McDaniel, R. R., Coppola, M. N., 2010. Two-stage stochastic optimization for the allocation of medical assets in steady-state combat operations. *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology* 7 (2), 89–102.
- Garrett, M. X., 2013. USCENTCOM Review of MEDEVAC Procedures in Afghanistan. Tech. rep., United States Central Command.
- Grannan, B. C., Bastian, N. D., McLay, L. A., 2015. A maximum expected covering problem for locating and dispatching two classes of military medical evacuation air assets. *Optimization Letters* 9 (8), 1511–1531.
- Green, L. V., Kolesar, P. J., 1984. A comparison of the multiple dispatch and m/m/c priority queueing models of police patrol. *Management Science* 30 (6), 665–670.
- Green, L. V., Kolesar, P. J., 2004. Anniversary article: improving emergency responsiveness with management science. *Management Science* 50 (8), 1001–1014.
- Hawkes, A. G., 1971. Spectra of some self-exciting and mutually exciting point processes. *Biometrika* 58 (1), 83–90.
- Hawkes, A. G., Adamopoulos, L., 1973. Cluster models for earthquakes - regional comparisons. *Bull. Int. Statist. Inst* 45 (3), 454–461.
- Ignall, E., Carter, G., Rider, K., 1982. An algorithm for the initial dispatch of fire companies. *Management Science* 28 (4), 366–378.
- Keneally, S. K., Robbins, M. J., Lunday, B. J., 2014. A Markov decision process model for the optimal dispatch of military medical evacuation assets. *Health Care Management Science*, 1–19.
- Kolesar, P., Blum, E. H., 1973. Square root laws for fire engine response distances. *Management Science* 19 (12), 1368–1378.
- Kolesar, P., Walker, W. E., 1974. An algorithm for the dynamic relocation of fire companies. *Operations Research* 22 (2), 249–274.

- Kroese, D. P., Botev, Z. I., 2014. Spatial process simulation. In: Schmidt, V. (Ed.), *Stochastic Geometry, Spatial Statistics and Random Fields: Models and Algorithms*. Springer-Verlag, Berlin, Ch. 12, pp. 369–404.
- Kuisma, M., Holmström, P., Repo, J., Määttä, T., Nousila-Wiik, M., Boyd, J., 2004. Prehospital mortality in an EMS system using medical priority dispatching: a community based cohort study. *Resuscitation* 61 (3), 297–302.
- Lagoudakis, M. G., Parr, R., 2003. Least-squares policy iteration. *The Journal of Machine Learning Research* 4, 1107–1149.
- Larson, R. C., 1972. *Urban Police Patrol Analysis*. MIT Press, Cambridge, MA.
- Lewis, E., Mohler, G., Brantingham, P. J., Bertozzi, A. L., 2012. Self-exciting point process models of civilian deaths in Iraq. *Security Journal* 25 (3), 244–264.
- Malsby III, R. F., Quesada, J., Powell-Dunford, N., Kinoshita, R., Kurtz, J., Gehlen, W., Adams, C., Martin, D., Shackelford, S., 2013. Prehospital blood product transfusion by US Army MEDEVAC during combat operations in Afghanistan: A process improvement initiative. *Military Medicine* 178 (7), 785–791.
- Maxwell, M. S., Restrepo, M., Henderson, S. G., Topaloglu, H., 2010. Approximate dynamic programming for ambulance redeployment. *INFORMS Journal on Computing* 22 (2), 266–281.
- Mayorga, M. E., Bandara, D., McLay, L. 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, L. A., Mayorga, M. E., 2010. Evaluating emergency medical service performance measures. *Health Care Management Science* 13 (2), 124–136.
- McLay, L. A., Mayorga, M. E., 2013a. A dispatching model for server-to-customer systems that balances efficiency and equity. *Manufacturing & Service Operations Management* 15 (2), 205–220.
- McLay, L. A., Mayorga, M. E., 2013b. A model for optimally dispatching ambulances to emergency calls with classification errors in patient priorities. *IIE Transactions* 45 (1), 1–24.
- Nicholl, J., Coleman, P., Parry, G., Turner, J., Dixon, S., 1999. Emergency priority dispatch systems: a new era in the provision of ambulance services in the UK. *Pre-hospital Immediate Care* 3, 71–5.
- Powell, W. B., 2009. What you should know about approximate dynamic programming. *Naval Research Logistics* 56 (3), 239–249.
- Powell, W. B., 2011. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*, 2nd Edition. John Wiley & Sons.

- Powell, W. B., 2012. Perspectives of approximate dynamic programming. *Annals of Operations Research* 13 (2), 1–38.
- Schmid, V., 2012. Solving the dynamic ambulance relocation and dispatching problem using approximate dynamic programming. *European Journal of Operational Research* 219 (3), 611–621.
- Scott, W., Powell, W., Moazeni, S., 2014. Least squares policy iteration with instrumental variables vs. direct policy search: Comparison against optimal benchmarks using energy storage. Tech. rep., Princeton University.
- Sutton, R. S., Barto, A. G., 1998. *Reinforcement Learning: An Introduction*. MIT Press.
- Swersey, A. J., 1982. A Markovian decision model for deciding how many fire companies to dispatch. *Management Science* 28 (4), 352–365.
- Van Roy, B., Bertsekas, D. P., Lee, Y., Tsitsiklis, J. N., 1997. A neuro-dynamic programming approach to retailer inventory management. In: *Proceedings of the IEEE Conference on Decision and Control*. Vol. 4. IEEE, pp. 4052–4057.
- White, G., Porter, M. D., Mazerolle, L., 2013. Terrorism risk, resilience and volatility: A comparison of terrorism patterns in three southeast Asian countries. *Journal of Quantitative Criminology* 29 (2), 295–320.