

Shortening Emergency Medical Response Time with Joint Operations of UAVs with Ambulances

Xiaoquan Gao

School of Industrial Engineering, Purdue University, gao568@purdue.edu,

Nan Kong

Weldon School of Biomedical Engineering, Purdue University, nkong@purdue.edu,

Paul Griffin

Department of Industrial and Manufacturing Engineering, Pennsylvania State University, pmg14@psu.edu,

Problem definition: Uncrewed aerial vehicles (UAVs) are transforming emergency service logistics applications across sectors, offering easy deployment and rapid response. In the context of emergency medical services (EMS), UAVs have the potential to augment ambulances by leveraging bystander assistance, thereby reducing response times for delivering urgent medical interventions and improving EMS outcomes. Notably, the use of UAVs for opioid overdose cases is particularly promising as it addresses the challenges faced by ambulances in delivering timely medication. This study aims to optimize the integration of UAVs and bystanders into EMS in order to minimize average response times for overdose interventions. **Methodology:** We formulate the joint operation of UAVs with ambulances through a Markov decision process (MDP) that captures random emergency vehicle travel times and bystander availability. We apply an approximate dynamic programming (ADP) approach to mitigate the solution challenges from high-dimensional state variables and complex decisions through a neural-network-based approximation of the value functions (NN-API). To design the approximation, we construct a set of basis functions based on queueing and geographic properties of the UAV-augmented EMS system. **Managerial implications:** The simulation results suggest that our NN-API policy tends to outperform several noteworthy rule- and optimization-based benchmark policies in terms of accumulated rewards, particularly for situations that are primarily characterized by high request arrival rates and a limited number of available ambulances and UAVs. The results also demonstrate the benefits of incorporating UAVs into the EMS system and the effectiveness of an intelligent real-time operations strategy in addressing capacity shortages, which are often a problem in rural areas of the US. Additionally, the results provide insights into specific contributions of each dispatching or redeployment strategy to overall performance improvement.

Key words: Approximate dynamic programming, large-scale MDP, uncrewed aerial vehicles, dispatching, simulation

History:

1. Introduction

1.1. Motivation

The US is facing a severe opioid crisis, with over 210 opioid overdose deaths reported each day (CDC 2022). Opioid overdoses can lead to respiratory depression and cardiac arrest, and without timely

intervention, the chances of survival decrease by up to 10% per minute (Cao 2005). Brain damage can occur after four minutes, and death can occur within six to eight minutes later (Doe-Simkins et al. 2009). In many cases, trained first responders are unable to reach the patient in time to administer naloxone, typically administered as a nasal spray, and provide ventilation, to prevent death. The bystander-enabled uncrewed aerial vehicle (UAV) delivery system is one potential approach to mitigate this problem. In such a UAV-augmented EMS system, 9-1-1 dispatchers can dispatch a UAV and direct a bystander to render an emergency response to the patient while emergency medical service (EMS) personnel are en route.

UAV use, as pilotless aircraft, has seen a rapid expansion of applications in recent years. In the US, this is largely facilitated by the increasingly specified guidelines and relaxed regulations of the US Federal Aviation Administration (FAA) on UAV airspace and operations. For example, Amazon Prime Air, Amazon's special service that delivers packages within 30 minutes, was granted operation by the FAA in mid-2020 to test order delivery via UAVs (Reuters 2020). In addition to parcel delivery, UAV use has a wide range of applications, including land surveillance, wildlife tracking, search and rescue operations, disaster response, and border patrol (Everaerts et al. 2008). UAVs are particularly well-suited for these tasks due to their ability to cover large areas quickly and efficiently, as well as their ability to access areas that may be difficult or dangerous for humans to reach. Since 2016, a number of east and central African countries have collaborated with Zipline, the world's largest automated delivery system designer, manufacturer, and operator, to deliver blood supplies, reducing the delivery time from four hours to 15 minutes in some cases (World Health Organization 2019). With state-of-the-art technologies, UAVs designed for on-demand commodity delivery can fly up to an hour and reach distances of up to 45 miles while carrying necessary payloads for emergency responses, such as automated external defibrillators for out-of-hospital cardiac arrests (Boutilier et al. 2017), blood transfusion toolkits for trauma injuries (Ling and Draghic 2019), and naloxone nasal spray for opioid overdoses (Ornato et al. 2020). All these recent developments, including the improvement of technology, decreases in cost, and changing regulations, will enhance the potential use of UAVs in EMS delivery.

This technology is especially promising given the inherent need for a rapid response to enhance patient outcomes, particularly in remote or hard-to-reach areas. For example, UAVs can respond faster in urban environments that present barriers to emergency services, such as heavy traffic congestion. Additionally, the scarcity of ambulances in many US rural counties can lead to extremely long response times, with one in ten patients waiting nearly 30 minutes for EMS arrival (Mell et al. 2017). Compared to increasing the number of ambulances and corresponding medical personnel, incorporating UAVs for medical delivery is a more realistic and cost-effective approach to optimize medical resources. For emergencies such as opioid overdoses and out-of-hospital cardiac arrests, it

is expected that UAV-delivered medical interventions can significantly save critical response times and avert life-threatening conditions. In practice, medical UAVs will be equipped with audio or video assistance devices, such as cameras, to help bystanders quickly assess the situation and follow instructions. For example, Zipline and Intermountain Healthcare have implemented drone deliveries in the Salt Lake Valley to reach patients and deliver medication faster, without requiring patients to travel to a clinic or hospital (Utah Public Radio 2022). In the realm of EMS delivery, there have been ongoing efforts to encourage bystander intervention, such as educational campaigns (Lockey et al. 2021), training programs (Clark et al. 2014), and liability protection laws (Latimore and Bergstein 2017). More recently, mobile phone applications, e.g., UnityPhilly (Schwartz et al. 2020), have focused on connecting layperson first responders with people experiencing overdoses. Furthermore, research has demonstrated the benefit of layperson-initiated overdose reversal through the administration of naloxone before the arrival of an ambulance or first responders (Schwartz et al. 2020).

The main objectives of our work are twofold: first, to demonstrate the benefits of using UAVs to deliver life-saving medication, and second, to provide guidance for EMS agencies on how to incorporate UAVs into their operations. By highlighting the potential advantages of using UAVs, we hope to encourage their wider adoption in the EMS field. At the same time, our work aims to provide practical guidance for how to effectively incorporate UAVs into EMS operations, including strategies for dispatching and redeployment.

In the US, 911 calls are typically received by a public safety answering point (PSAP), which determines the nature of the emergency, and next decides either to dispatch responders immediately or transfer the call to a specialized secondary PSAP. We assume that both UAVs and ambulances are managed in a centralized scheme and aim to improve EMS outcomes through the joint operation of both types of vehicles, including dispatching and redeployment. We develop a Markov Decision Processes (MDP) framework to capture the interplay between these decisions, spatially distributed stochastic arrivals of requests, and the state of the EMS system involving the concurrent use of UAVs and ambulances. The process flow is shown in Figure 1. When a witness calls 9-1-1 to report a case, a request enters the EMS system. During the call, the dispatcher will ask about the presence of bystanders nearby in addition to following the current protocol. Based on the status of the system, a dispatching decision is made, which may include the dispatch of an ambulance or a UAV followed by an ambulance, and the specific ambulance (and UAV) to be dispatched. When the UAVs or ambulances have completed their service, they are redeployed to one of their bases to be better prepared for upcoming requests. Note that UAV dispatching is helpful if and only if at least one bystander is willing to help when the UAV arrives at the emergency scene. The probability that at least one bystander is willing to help is estimated based on the number of bystanders acquired

during the call and the likelihood that one of them is willing to retrieve and administer the EMS toolkit.

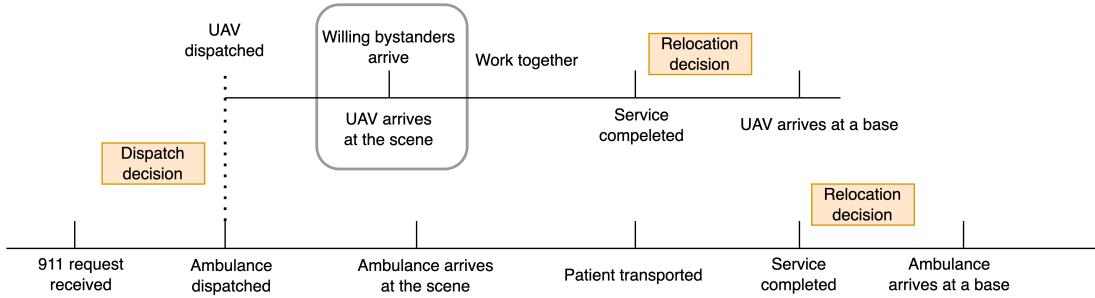


Figure 1 An illustration of UAV-based EMS process timeline.

The most straightforward dispatching strategy is to dispatch the UAV closest to the request in order to minimize the response time and maximize the chance of survival (Kim et al. 2009). However, this myopic strategy is typically suboptimal, and a more sophisticated strategy would improve outcomes (Jagtenberg et al. 2015). Because ambulance and UAV resources are limited, minimization of response time for the current requests may lead to a much longer response time for future requests, i.e., a nuanced balance should be considered between current and future requests. Additionally, we consider redeployment decisions for UAVs and ambulances. Redeploying UAVs and ambulances to one of their bases allows for charging, replenishment, and personnel rest. Proper redeployment decisions aim to balance the distribution of available UAVs and ambulances over the bases for time-varying demand. On top of this delicate trade-off, uncertainties in travel time, bystander willingness to assist, and emergency request volume and their locations should be modeled explicitly in the MDP framework.

In summary, the integration of UAVs and ambulances brings increased flexibility to EMS delivery, but also introduces unique challenges in making intelligent operational decisions in real time. The operational strategy/policy suggested by our model and algorithm addresses the fundamental tradeoff between response times for current and future requests, while taking into account various uncertainties and the use of heterogeneous servers (ambulances and UAVs).

1.2. Main Contributions and Results

1. We extend the MDP-based analytics framework to consider a range of real-time operational decisions in a dynamic, coordinated logistics system with both conventional (ambulances) and augmenting (UAVs) delivery vehicles. Given the two types of delivery vehicles, we consider two types of requests that can be answered in the logistic system, differing in the use of UAVs. Our

approach also accounts for additional sources of uncertainty. In addition to uncertain request arrival times and locations, we use a delayed reward function to reflect the uncertainty of UAV and ambulance travel times. Yet, we preserve the model's high fidelity, including the presence of bystanders.

2. We apply an approximate dynamic programming (ADP) approach, and design a tractable approximate policy iteration algorithm for the complex stochastic dynamic optimization problem, which employs value function approximation via neural networks. Based on the spatial and temporal characteristics of the EMS system, we design a set of effective basis functions to enhance the algorithm's performance. Our basis functions are novel and distinct from the literature in three aspects: (a) we consider the heterogeneous nature of the ambulance-UAV system and approximate coverage accordingly, especially on the future missed call rate, (b) we approximate the average response time to adapt to the objective of minimizing response times, beyond maximizing the number of requests served, and (c) we consider not only spatial features but also temporal features of the system, i.e., availability reduction. Additionally, the neural network representation of the value function offers a much richer class of nonlinear functions and can be trained iteratively.

3. We acquire important managerial implications about real-time operations for coordinated EMS logistics. Our case studies are based on historical data from the state of Indiana for emergency naloxone administration for opioid overdoses and the National EMS Information System (NEMSIS) database. Simulation results suggest that our policy consistently outperforms several noteworthy rule- and optimization-based benchmark policies in terms of accumulated rewards. This superiority is particularly pronounced when the request arrival rate is high and the availability of ambulances and UAVs is limited. Our results also highlight the benefits of using UAVs in the EMS system and the effectiveness of an intelligent real-time operations strategy in addressing capacity shortages, which are a common challenge in rural areas of the United States. Additionally, our results provide valuable insights into the contributions made by each dispatching or redeployment strategy. We also identify scenarios where the use of the NN-API is strongly recommended, as it significantly outperforms benchmark policies, and situations where a simple static policy can perform comparably to the NN-API.

2. Literature Review

To the best of our knowledge, our work is the first that combines the real-time operations of UAVs and ambulances for time-sensitive logistics. Previous research has primarily focused on individual aspects of the problem, such as ambulance dispatching and redeployment using static and dynamic policies, optimization models for vehicle-mix in EMS, and the use of UAVs for emergency response. Our research builds upon these efforts by also considering the coordination optimization between UAVs and ambulances.

There is a wealth of literature on ambulance operations management, with many studies focusing on minimizing average response time, as in our research. Early work in this area primarily utilized static policies, including threshold-based policies and policies derived through integer programming (IP) or mixed-integer programming (MIP). For example, Daskin (1983), Marianov and ReVelle (1996) used an IP-based approach and generalized the maximal expected covering location problem (MEXCLP) for public service facility location analysis. Alternatively, the threshold-based policy is also widely used in ambulance dispatching and redeployment. One representation of the threshold policy is the “preparedness” measure proposed by Andersson and Värbrand (2007), which is used to evaluate the ability of an EMS system to serve potential patients.

In the EMS system, patient numbers are highly uncertain and so pre-planned scheduling or operation solutions may not optimally respond to fluctuating situations. Therefore, real-time decision-making is required, which considers systems dynamics such as time-varying demand (emergency calls), time-varying traffic, and different intervention times required by patients. As a result, several researchers have explored the benefit of dynamic dispatching and redeployment optimization, using assumptions such as exponential service time and no-buffer request queue, McLay and Mayorga (2013a), McLay and Mayorga (2013b) and Jagtenberg et al. (2017) built several MDP models and solved them to optimality for small-scale instances. These exactly solved MDPs highlighted the value and suboptimality of the closest idle dispatching policy and how various equity formulations affect the underlying dispatching policies. Recent advances in approximate dynamic programming (ADP) have improved our ability to solve large-scale problems efficiently. For example, Schmid (2012) and Jenkins et al. (2020) approximated value functions with tabular ADP. Among previous work using ADP, our work is most closely related to Maxwell et al. (2010) and Nasrollahzadeh et al. (2018), which proposed novel basis functions based on the underlying problem structure to approximate value functions. Our work differs in that we consider the joint operation of two delivery modes for EMS logistics. The joint operation requires the consideration of additional novel basis functions for heterogeneous service providers. We refer to two review papers – Aringhieri et al. (2017), and Bélanger et al. (2019) – for comprehensive discussions on optimizing location, redeployment, and dispatching decisions for emergency medical vehicles.

Another stream of literature relates to vehicle mix and response to multiple requests for EMS. Similar to our work, most of these papers differentiate vehicles by their service capability for different types of patients. Previous papers consider multiple responses in the context of deterministic and probabilistic maximal covering ambulance location problems (Schilling et al. 1979, ReVelle and Marianov 1991). McLay (2009) proposed the maximum expected coverage location problem with two types of servers (MEXCLP2) to efficiently deploy two types of medical units, i.e., Advanced

Life Support (ALS) and Basic Life Support (BLS), to serve multiple types of customers. For UAV-ambulance coordination, Shin et al. (2022) developed a modeling framework to optimize a network of drones, bystanders, and ambulances for cardiac arrest response, taking into account the availability of bystanders. In addition to location problems, researchers have also worked on dispatching and redeployment problems involving multiple types of ambulances with stochastic programming (SP) and MDP. Boujema et al. (2020) addressed the ambulance redeployment planning problem in a two-tiered EMS using a two-stage SP model, with the first stage addressing redeployment decisions and the second stage addressing dispatching decisions. Similarly, Yoon et al. (2021) formulated a two-stage SP problem for location and dispatching decisions considering prioritized emergency patients, and also extended the model to incorporate non-transport vehicles, similar to UAVs in our model. For real-time operations, Chong et al. (2016) and Yoon and Albert (2020, 2021) constructed MDP models to optimize the dispatching of multiple types of vehicles to (prioritized) patients, demonstrating structural properties. For example, the optimal policy is a control-limit policy, which is more likely to send an ALS unit to calls when more ALS units are available. However, to make the MDP tractable, the authors only decided whether to dispatch an ALS or a BLS, but not which specific unit to dispatch. With ADP-based solution techniques, we are able to model and improve upon dispatching and redeployment decisions with greater specificity, including the specific UAV and ambulance to dispatch and the specific base to redeploy the UAV or ambulance to.

Unlike previous work on vehicle mix, we model the information of bystanders for UAV dispatch, which is unique. And to our best knowledge, among studies of multiple types of ambulances, only Park and Lee (2019) considered the real-time dispatching that makes specific dispatching and redeployment decisions. The authors leveraged ADP with state aggregation and monotonicity-preserving projection operators to solve the complex MDP model. In our work, we maintain a high level of fidelity without using state aggregation and address the curse of dimensionality through neural network-based value function approximation (VFA). We capture differences between UAVs and ambulances by using different transition dynamics and integrating them into the VFA using a queueing model of heterogeneous servers.

Recent studies have explored the feasibility of delivering medical equipment via UAVs, including flotation devices (Claesson et al. 2017a), automated external defibrillators (AED) (Boutilier et al. 2017, Claesson et al. 2017b) and blood products (Amukele et al. 2017). In addition to strategic decisions, Chu et al. (2021) developed UAV dispatching rules based on the difference between predicted ambulance response time and calculated UAV response time for each out-of-hospital cardiac arrest. There are two categories of existing literature that investigate the joint optimization of UAV

planning and operations management. The first category concerns situations where UAVs are carried by truck and dispatched from the truck near the service location. Typical applications include precision agriculture, package delivery, oceanographic sampling, forest fire, or oil spill monitoring (Tokekar et al. 2016, Fawaz et al. 2017, Jia and Zhang 2017). The second category of studies considers cases where UAVs and ground vehicles perform independent tasks, similar to the situation in our work. However, most previous studies in this category only consider facility locations at the strategic level and UAV allocation at the tactical level (Dorling et al. 2016, Agatz et al. 2018). The studies most similar to ours are Ulmer and Thomas (2018) and Chen et al. (2019). Ulmer and Thomas (2018) explored the addition of UAVs to conventional vehicles for the same-day delivery problem. Using an MDP model, the authors presented a dynamic vehicle routing problem with heterogeneous fleets, where the decisions were to reject an order and assign a UAV or a ground vehicle. To address the curse of dimensionality, the authors adopted policy function approximation (PFA) based on the insight that distant customers should generally be served by UAVs and closer customers served by conventional vehicles. Chen et al. (2019) extended the above work to include additional information on resource availability and demand as well and implemented a Deep Q -Learning (DQN). However, only the acceptance and general assignment decisions were made in both studies above, i.e., the order/request should be either served by UAVs or conventional vehicles. An important difference in our work is that we seek to optimize the dispatching and redeployment decisions of UAVs and ambulances with potentially multiple responses to each EMS request. This difference in the decisions significantly expands our MDP model, making it very difficult to parameterize the policy function directly and use the value function of state-action pairs. Additionally, our application emphasizes the “time-criticality” of the service, with emergency response time being the key objective. Therefore, the objective function is different, and different basis functions are required to approximate the value function.

3. MDP Model

This section presents an infinite-horizon average-cost Markov decision process (MDP) formulation. We adopt event-driven modeling to incorporate on-demand UAV and ambulance dispatching and redeployment decisions in the model and capture the EMS system evolution. Events are triggered by changes in the status of UAVs, ambulances, and requests. Let $\mathcal{N} := \{1, 2, \dots, N\}$ be the set of demand nodes, $\mathcal{M}^u := \{1, 2, \dots, M^u\}$ be the set of UAV charging stations, and $\mathcal{M}^a := \{1, 2, \dots, M^a\}$ be the set of ambulance bases, e.g., hospitals or bases of private EMS agencies. We consider a total of L^u UAVs and L^a ambulances. The home base of UAV l is denoted by $h_l^u \in \{1, 2, \dots, M^u\}$ and the home base of ambulance l is denoted by $h_l^a \in \{1, 2, \dots, M^a\}$. Let $\mathcal{A}^u(s)$ be the set of available UAVs, i.e., $\mathcal{A}^u(s) := \{l : r_l^u = 0\}$, where r_l^u is the remaining time in its current status of UAV l ; similarly,

let $\mathcal{A}^a(s)$ be the set of available ambulances, i.e., $\mathcal{A}^a(s) := \{l : r_l^a = 0\}$, where r_l^a is the remaining service time in its current status of ambulance l .

We divide 9-1-1 requests into two types based on whether UAVs can be of help: (1) Type 1: requests that UAVs can serve as the first response such as opioid overdose and out-of-hospital cardiac arrest; and (2) Type 2: requests that only ambulances can help such as massive hemorrhage. We assume that arrivals of EMS requests of Type 1 and Type 2 follow Poisson processes with rates λ^u and λ^a , respectively. We make dispatching decisions for both types of requests. For Type-1 requests, we need to choose between a single-ambulance response and a UAV-ambulance sequence response. For Type-2 requests, we only consider which ambulance to dispatch. We also make redeployment decisions each time a UAV or ambulance finishes the current emergency response task and returns to a base for replenishment, recharging, and personnel rest. In the following model description, we assume that all the information obtained from the 9-1-1 call, such as request location, request type, and bystander information, is accurate.

3.1. State Space

The state space is composed of five parts: vectors $B^u = (b_1^u, b_2^u, \dots, b_{L^u}^u)$, $B^a = (b_1^a, b_2^a, \dots, b_{L^a}^a)$, $C = (c_1, c_2, \dots, c_K)$, e and τ , where b_l^u , $l = 1, \dots, L^u$, contains information about the state of the l^{th} UAV; b_l^a , $l = 1, \dots, L^a$, contains information about the state of the l^{th} ambulance; c_j , $j = 1, \dots, J$, contains information about the j^{th} request; e denotes the event type and τ corresponds to the current time. Therefore, the state space of the system is represented by $\mathcal{S} := \{\mathbf{s} = (\tau, e, B^u, B^a, C)\}$.

The status of UAV l is given by $b_l^u = (d_l^u, r_l^u, f_l^u)$, $l = 1, \dots, L^u$, where $d_l^u \in \{1, 2, \dots, N\}$ is the destination of each UAV. For this work, it is sufficient to consider four possibilities for the status of UAVs, i.e., $f_l^u \in \{0, 1, 2, 3\}$, where 0 indicates that the UAV is available at the base; 1 indicates that the UAV is going to a request location; 2 indicates that the UAV is serving a request on the scene; 3 indicates that the UAV is returning to a base.

The state of ambulance l is given by $b_l^a = (d_l^a, r_l^a, f_l^a)$, $l = 1, \dots, L^a$, where $d_l^a \in \{1, 2, \dots, N\}$ is the destination of each ambulance. Assume that all the opioid-overdosed patients require transport to hospitals, and all other patients require transporting to hospitals with probability p^t . For this work, it is sufficient to consider five possibilities for the status of an ambulance, i.e., $f_l^a \in \{0, 1, 2, 3, 4\}$, where 0 indicates that the ambulance is available at the base; 1 indicates that the ambulance is going to a request location; 2 indicates that the ambulance is serving a request on the scene; 3 indicates that the ambulance is going to the hospital; 4 indicates that the ambulance is returning to a base.

A request j is represented by $c_j = (g_j, q_j, o_j, \omega_j)$, $j = 1, \dots, J$, where $g_j \in \{1, 2, \dots, N\}$, is the request location, q_j is the arrival time of the request. For this work, $o_j \in \{1A, 1B, 2\}$ denotes the type and

status of the request, where $o_j = 1A$ ($o_j = 1B$) implies UAVs are qualified for the first response and the request is waiting for the first (follow-up) response; $o_j = 2$ implies only ambulances are qualified for serving the request. If request j is finished, it would be marked as “served” and immediately removed from the request set. In addition, ω_j represents bystander helping probability at request j . Specifically, $\omega_j = 1 - (1 - p^b)^{N^b}$ is the probability that at least one bystander is willing to help when the UAV arrives, where N^b is the estimated total number of bystanders, and p^b is the probability that each bystander is still willing to help when the UAV arrives. The bystander willingness p^b can be estimated through interviews (Lankenau et al. 2013).

An event is represented by $e, e \in E$, where E is the set of all possible event types. Without loss of generality, we assume that decisions are made at transition times. In our model, transition times are associated with the following events: (1) request j arrives; (2) ambulance l is in transit to request j ; (3) ambulance l arrives at the location of request j and starts service; (4) ambulance l finishes serving request j at scene; (5) ambulance l finishes serving request j at a hospital; (6) ambulance l arrives at a base; (7) UAV l is in transit to request j ; (8) UAV l arrives at the location of request j and starts service; (9) UAV l finishes serving request j at the scene and in transit to a base; (10) UAV l arrives at a base.

3.2. Action Space

We consider a loss system without request queues, i.e., a request will be outsourced to a nearby EMS agency if there are no available UAVs and ambulances in our system. Another option is to place requests in a queue when all servers are busy. Bandara et al. (2014) showed that the strategy performance relationship remains the same for systems allowing and not allowing request queueing, but the overall system performance with queueing is lower due to increased vehicle utilization. Therefore, we consider outsourcing requests that arrive when no ambulance is available. The action space is described with three event-based cases based on the type of actions required.

Case 1 If request j arrives ($e = 1$), the decision-maker has three types of decisions: (1) whether to outsource the request; (2) which ambulance to immediately dispatch to serve the request; (3) which UAV to immediately dispatch and which ambulance to dispatch as a follow-up.

Define $X_{l,j}^u = 1$ if UAV l is dispatched to request j , and $X_{l,j}^u = 0$ otherwise. Also, define $X_{l,j}^a = 1$ if ambulance l is dispatched to request j , and $X_{l,j}^a = 0$ otherwise. Therefore, if event $e = 1$, the action space is given by

$$A_1(s) := \left\{ (X_{l,j}^u, X_{l,j}^a) : \sum_{l \in \mathcal{A}^u(s)} X_{l,j}^a \leq 1, \right. \\ \sum_{l \in \mathcal{A}^u(s)} X_{l,j}^u \leq \sum_{l \in \mathcal{A}^a(s)} X_{l,j}^a, \\ \left. X_{l,j}^u, X_{l,j}^a \in \{0, 1\} \right\},$$

where the first constraint states that for each request, at most one ambulance is dispatched and the second constraint states that the number of UAVs dispatched should not exceed the number of ambulances dispatched. The two constraints together limit the dispatching decisions into three types: (1) dispatching one UAV and one ambulance; (2) dispatching only one ambulance, and (3) dispatching no UAV or ambulance, i.e., outsourcing the request.

Case 2 If event $e \in \{5, 9\}$, the decision is to determine to which base to redeploy the ambulance/UAV.

If event $e = 5$, i.e., an ambulance finishes service, let $Z_{l,b}^a = 1$ if ambulance l is redeployed to base b , and $Z_{l,b}^a = 0$ otherwise. Then the action space is given by

$$A_2(s) := \left\{ (Z_{l,b}^a) : \sum_{b \in \mathcal{M}^a} Z_{l,b}^a = 1 \right\},$$

which ensures that ambulance l is redeployed to only one base.

If event $e = 9$, i.e., a UAV finishes service, let $Z_{l,b}^u = 1$ if UAV l is redeployed to base b , and $Z_{l,b}^u = 0$ otherwise. Then the action space is given by

$$A_3(s) := \left\{ (Z_{l,b}^u) : \sum_{n \in \mathcal{M}^u} Z_{l,b}^u = 1 \right\},$$

which ensures that UAV l is redeployed to only one base.

Case 3 If the event $e \in \{2, 3, 4, 6, 7, 8\}$, we set $A(s) = \emptyset$, i.e., no action will be taken.

3.3. Transitions

Let s_k be the state of the system when the k th event happens. The evolution of state s_k can be characterized by action a_k , and random element $\omega(s_k, a_k)$, and a function F , i.e., $s_{k+1} = F(s_k, a_k, \omega(s_k, a_k))$. We assume the on-scene time follows a lognormal distribution (Ingolfsson et al. 2008).

3.4. One-step Reward Function

We consider maximizing the average health outcome as the primary objective function in our optimization framework. The health outcome is modeled as a decreasing function of the response time because the likelihood of survival decreases with the time it takes to receive medical treatment (Blackwell and Kaufman 2002, Wilde 2013), especially in cases of opioid overdose where every minute is critical. Let $h(s_k, a_k, s_{k+1})$ denote the cost or reward of a transition from s_k to s_{k+1} , when action a_k is taken. The system only incurs a cost or gains a reward at event $e \in \{1, 3, 8\}$.

When $e = 1$, i.e., request j arrives, a penalty will be incurred if the request is outsourced, specifically,

$$h(s_k, a_k, s_{k+1}) = \begin{cases} C_o & \text{if } \sum_{l \in \mathcal{A}^u(s)} X_{l,j}^u = 0, \sum_{l \in \mathcal{A}^a(s)} X_{l,j}^a = 0; \\ 0 & \text{otherwise,} \end{cases} \quad (1)$$

where C_o is the penalty for health outcome decrease from delay in response time caused by outsourcing request j .

When $e = 8$, i.e., a UAV arrives at the scene,

$$h(s_k, a_k, s_{k+1}) = \begin{cases} g_1(\tau - q_j) \cdot \mathbf{1}_{\{N_j \geq 1\}} & \text{if } o_j = 1A; \\ 0 & \text{otherwise,} \end{cases} \quad (2)$$

where $\tau - q_j$ represents the first response time of request j ; $g_1(\cdot)$ denotes the reward function with respect to the response time for opioid overdoses, e.g., $g_1(t) = \frac{[T_1-t]^+}{T_1}$, where T_1 is the response time threshold, e.g., $T_1 = 8$ minutes; random variable $N_j \sim \text{Binom}(N^b, p^b)$ is the number of willing bystanders when a UAV arrives at the scene. If $o_j \neq 1A$, then the UAV arrives later than the ambulance, or the UAV is dispatched for a request of Type 2. In this case, the UAV dispatching becomes redundant.

When $e = 3$, i.e., an ambulance arrives at the scene,

$$h(s_k, a_k, s_{k+1}) = \begin{cases} g_1(\tau - q_j) & \text{if } o_j = 1A; \\ g_2(\tau - q_j) & \text{if } o_j = 2; \\ 0 & \text{if } o_j = 1B, \end{cases} \quad (3)$$

where $g_2(\cdot)$ denotes the reward function with respect to the response time for general, e.g., $g_2(t) = \frac{[T_2-t]^+}{T_2}$, where T_2 is the response time threshold, e.g., $T_2 = 12$ minutes; $o_j = 1A$ and $o_j = 2$ denotes the case where the ambulance serves as the first response; $o_j = 1B$ denotes the case where the ambulance serves as the follow-up response.

3.5. Optimality Criterion

The *expected* average reward value of a policy π is defined for all $s_0 \in \mathcal{S}$ as

$$v_g(\pi, s_0) := \lim_{T \rightarrow \infty} \frac{1}{T} \mathbf{E}_{s_t, a_t, s_{t+1}} \left[\sum_{t=0}^{T-1} h(s_t, a_t, s_{t+1}) | s_0, \pi \right], \quad (4)$$

where $h(s_t, a_t, s_{t+1})$ is the one-step cost/reward, which is defined in Section 3.4. The limit in Equation (4) exists for a stationary policy when the MDP is unichained (Puterman (2014) Section 8.3.3). Assuming the Markov chain under the policy π is unichain, we have $v_g(\pi, s_0) = v_g(\pi)$, $\forall s_0 \in \mathcal{S}$. Then the optimal policy π^* is the policy that satisfies the average-reward Bellman *optimality* equation as follows,

$$v_b(\pi^*, s) + v_g(\pi^*) = \max_{a \in \mathcal{A}} \sum_{s' \in \mathcal{S}} p(s'|s, a) [h(s, a, s') + v_b(\pi^*, s')], \quad \forall s \in \mathcal{S}, \quad (5)$$

where $v_b(\cdot)$ is the optimal relative value function, $h(s, a, s')$ is the one-step cost/reward, and $p(s'|s, a)$ is the transition probability. For notational simplicity, we omit π^* and denote the optimal average reward as v_g in the following sections. We refer to Cavazos-Cadena (1991), Cavazos-Cadena and Sennott (1992) for summaries of results on existence conditions for discrete-time average cost MDPs with countable state space and finite action sets.

4. Approximate Solutions

A conventional method to solve Equation (5) is through policy iteration (PI). The PI algorithm starts with a random policy, then computes the value function of that policy (step 1: policy evaluation), and then determines a new, improved policy based on the previous value function (step 1: policy improvement). These two steps are repeated iteratively until the policy converges. However, to perform steps 1 and 2, it is first necessary to parameterize the associated transition matrix $\mathcal{S} \times \mathcal{A} \times \mathcal{S} \rightarrow \mathcal{R}$ and reward matrix $\mathcal{S} \times \mathcal{A} \rightarrow \mathcal{R}$. In our MDP model, the state space $|\mathcal{S}|$ is unbounded as the time variable τ is continuous. Even without τ , the dimension of the state space grows exponentially with the number of UAVs and ambulances. This makes it infeasible to store all $v_b(s), s \in \mathcal{S}$, not to mention enumerating the state space to solve the Bellman equations (5) to optimality. To tackle the curse of dimensionality resulting from the need to enumerate the state space, we conduct a simulation-based approximate policy iteration (API). The main framework of the API algorithm is described in Section 4.1.

To represent value functions of the high-dimensional state space, we approximate the relative value function, i.e., $v_b(s), s \in \mathcal{S}$, with a neural network model of a finite set of basis functions (or features). That is, for each $s \in \mathcal{S}$,

$$v_b(s) \approx Z(\Phi(s)) = Z(\phi_1(s), \phi_2(s), \dots, \phi_f(s)). \quad (6)$$

Here, $\Phi(s) = (\phi_1(s), \phi_2(s), \dots, \phi_f(s))$ is the set of f basis functions; and $Z(\cdot)$ is a model mapping basis functions to the relative value function. In our algorithm, $Z(\cdot)$ represents a neural network model. Using Equation (6), the approximate relative value function is determined by the neural network model $Z(\cdot)$ and a set of pre-specified basis functions, which are described in detail in Sections 4.1 and 4.2, respectively.

4.1. Approximate Policy Iteration

The neural network-based approximate policy iteration (NN-API) algorithm is extended from the basic policy iteration (PI) algorithm and adapted in the following four aspects.

Simulation-based policy evaluation. The core concept of approximate dynamic programming is to follow a sample path rather than enumerating the state space to update value functions. In NN-API, sample paths are generated based on the current policy and predefined distributions of randomness. Only rewards from the states visited on the sample paths would be used to update $Z(\cdot)$.

Post-decision state variable. A post-decision state variable is the state of the system after we have made a decision but before any new information has arrived (Powell 2007). Rather than estimating the expectation of the value around the next pre-decision state s_{k+1} , we directly

estimate $\bar{V}(s_k^a)$ for the post-decision state s_k^a . That is, we make decisions by optimizing $\hat{v}_k^n = \min_{a_k} \left(C(s_k^n, a_k) + \gamma \bar{V}_k^{n-1}(S^{M,a}(s_k, a_k)) \right)$, instead of estimating the expectation and optimizing $\tilde{v}_k^n = \min_{a_k} \left(C(s_k^n, a_k) + \gamma \mathbb{E}\{\bar{V}_{k+1}^{n-1}(S^{M,W}(s_k^n, a_k, W_{k+1}))\} \right)$. Here M and W represent our model and the uncertainty in the model, respectively; $S^{M,a}$ denotes the system state immediately after decision a and $S^{M,W}$ denotes the system state after the uncertainty W is realized.

Average reward computation. For the average reward v_g computation, instead of solving the Poisson equation (5), NN-API estimates v_g iteratively with $v_g \leftarrow v_g + \beta_g \Delta_g$, where $\beta_g = \frac{1}{n_u+1}$ (n_u is the number of updates of v_g so far), and $\Delta_g = r(s, a) - v_g$.

Value function representation. We use a feed-forward neural network to approximate the value function, rather than using a tabular form. The neural network consists of three layers: an input layer, a hidden layer, and an output layer. The information provided to the input layer is a set of $|\phi|$ basis functions associated with a post-decision state s^a . The hidden layer consists of a set of nonlinear activation units, and the size of this layer is a tunable parameter. The output layer produces a single scalar output by applying the activation function, which is the final approximation for the value function with respect to the input. For more information on the design and training of the neural network, see Section 1 of the Online Supplement.

The above four solution ideas help alleviate the curse of dimensionality and are incorporated into the framework of the NN-API (Algorithm 1) by using temporal difference (TD) learning.

4.2. Basis Functions

A key difficulty in the design of value function approximation is to select a set of basis functions that enables us to approximate the downstream costs. Based on the problem property, we conjecture the following basis functions (Section 4.2.1 - 4.2.6). We adopt an iterative process of testing and refining to identify effective basis functions for our ADP implementation. While simulation-based comparison would need to verify the effectiveness of these basis functions, we ensure that the constructed basis functions have the same monotonicity as the optimal value function at the design phase. Specifically, the optimal value function is monotone with bystander helping probability and availability of ambulances/UAVs, as stated in Proposition 1.

Proposition 1. *The optimal value function for the average-reward MDP described in Section 3 has the following properties:*

1. *It increases with an increase in the bystander helping probability w_j , given the assumption that the on-scene time of the UAV is insignificant.;*
2. *It increases when the system has one additional available UAV or ambulance.*

The proof of proposition 1 is based on the coupling arguments between the systems under optimal and suboptimal policies; see Section 2 of the Online Supplement for details. To further

Algorithm 1: NN-API: Neural network-based approximate policy iteration

Result: A trained neural network whose input is basis functions and output is an approximate value function.

Construct the basis functions $\Phi = (\phi_1, \phi_2, \dots, \phi_f)'$;

Initialize the neural network $Z_0(\Phi)$ using a myopic policy. Initialize the average reward

$$\hat{v}_g^{1,1} = 0.$$

for $n = 1, 2, \dots, N$ **do**

Policy evaluation starts.

Sample an initial state $S^{n,1}$ and choose a sample path ω^n ;

for $m = 1, 2, \dots, M$ **do**

Compute $a^{n,m} = \arg \min_{a \in \mathcal{A}^{n,m}} \left(C(s^{n,m}, a) + Z_{n-1} \left(\Phi^{n,m} (S^{M,a}(S^{n,m}, a)) \right) \right)$;

Compute $S^{n,m+1} = S^M(S^{n,m}, a^{n,m}, W_m(\omega^n))$;

Compute $\hat{v}_g^{n,m+1} = \hat{v}_g^{n,m} + \beta_g (C(s^{n,m}, a) - C(s^{n,m}, a^{n,m}))$, $\beta_g = 1/((n-1)M + m + 1)$;

end

Let Ψ^n be a $m \times F$ matrix where the $(i,j)^{th}$ entry is given by $\phi_k^{n,i}$;

Let \hat{V}^n be a vector of m dimensions with element,

$\hat{v}^{n,m} = C(s^{n,m}, a^{n,m}) - \hat{v}_g^{n,M} + Z_{n-1} \left(\Phi^{n,m} (S^{M,a^{n,m}}(S^{n,m}, a^{n,m})) \right)$;

Policy evaluation ends.

Retrain the neural network model with feature Ψ^n and label \hat{V}^n . Denote the updated neural network model with Z_n ; Policy improvement.

end

refine the design of basis functions, we follow Nasrollahzadeh et al. (2018) by approximating the system dynamics by constructing a queueing model. We use an $M/G/c/c$ queue to approximate our system since we do not consider putting 9-1-1 requests in the queue as reasonable. Another adaptation from the fact that we cannot treat UAVs and ambulances as homogeneous servers with the same service time distribution. Thus, based on the approximations provided by Fakinos (1980), we construct basis functions by dealing with the queueing system with heterogeneous servers. In Section 4.2.1 - 4.2.6, when mentioning “server”, we refer to both UAVs and ambulances for Type-1 requests and ambulances for Type-2 requests.

4.2.1. Expected delayed rewards When dispatching decisions are made (i.e., $e = 1$), the immediate reward only includes the penalty for outsourcing requests. The reward for serving the request is not realized until a UAV or ambulance arrives at the scene (i.e., $e = 8$ or $e = 3$). To reflect the direct impact of the dispatching decision on the reward, we include the expected reward as one of the basis functions. We denote the expected travel time for UAVs to travel between locations d_1 and d_2 as $\hat{t}^u(d_1, d_2)$, and the expected travel time for ambulances as $\hat{t}^a(d_1, d_2)$.

If both a UAV and an ambulance are dispatched, i.e., $X_{l_1,j}^u = 1, X_{l_2,j}^a = 1$,

$$\phi_1(s) = \begin{cases} \omega_j \cdot g_1(\hat{t}^u(d_{l_1}^u, g_j)) & \text{if } o_j = 1A, \hat{t}^u(d_{l_1}^u, g_j) < \hat{t}^a(d_{l_2}^a, g_j), \\ g_1(\hat{t}^a(d_{l_1}^a, g_j)) & \text{if } o_j = 1A, \hat{t}^u(d_{l_1}^u, g_j) \geq \hat{t}^a(d_{l_2}^a, g_j), \\ g_2(\hat{t}^a(d_{l_2}^a, g_j)) & \text{otherwise,} \end{cases} \quad (7)$$

where $o_j = 1A, \hat{t}^u(d_{l_1}^u, g_j) < \hat{t}^a(d_{l_2}^a, g_j)$ refers to the scenario where UAVs are eligible for the first response and the dispatched UAV is expected to arrive earlier than the dispatched ambulance; $g(\hat{t}^u(d_{l_1}^u, g_j))$ represents the reward from UAV's first response time if there are more than one willing bystander (with probability ω_j); $o_j = 1A, \hat{t}^u(d_{l_1}^u, g_j) \geq \hat{t}^a(d_{l_2}^a, g_j)$ refers to the scenario where UAVs are eligible for the first response, but the dispatched ambulance is expected to be the first response. Otherwise, $o_j = 2$, i.e., the dispatched ambulance, would serve as the first response.

If only an ambulance is dispatched, i.e., $\sum_l X_{l,j}^u = 0, X_{l_2,j}^a = 1$, a delayed reward from ambulance response is expected, i.e.,

$$\phi_1(s) = \begin{cases} g_1(\hat{t}^a(d_{l_2}^a, g_j)) & \text{if } o_j = 1A, \\ g_2(\hat{t}^a(d_{l_2}^a, g_j)) & \text{if } o_j = 2. \end{cases} \quad (8)$$

If the request is outsourced, i.e., $\sum_l X_{l,j}^u = 0, \sum_l X_{l,j}^a = 0$, the penalty for outsourcing would be incurred immediately when the dispatching decision is made so that there would not be any delayed rewards, i.e.,

$$\phi_1(s) = 0. \quad (9)$$

4.2.2. Uncovered request rate This basis function captures the rate of request arrivals that cannot be reached within the response time threshold by any of the available servers. Let $\mathcal{A}^u(s)$ and $\mathcal{A}^a(s)$ be the set of available UAVs and ambulances, respectively, when the system state is s . Specifically, $\mathcal{A}^u(s) = \{l | f_l^u = 0\}$ and $\mathcal{A}^a(s) = \{l | f_l^a = 0\}$. Then the coverage of demand node i can be written as

$$N_i(s) = \sum_{l \in \mathcal{A}^u(s)} \mathbf{1}_{\{d(d_l^u(s), i) \leq \Delta\}} + \sum_{l \in \mathcal{A}^a(s)} \mathbf{1}_{\{d(d_l^a(s), i) \leq \Delta\}}. \quad (10)$$

We can then compute the rate of request arrivals that are not covered by any available servers with

$$\phi_2(s) = \sum_{i \in \mathcal{N}} \lambda_i \mathbf{1}_{\{N_i(s) = 0\}}. \quad (11)$$

Note that a Type-1 request is considered as covered either when (i) an ambulance is expected to reach it within time T_A or (ii) a UAV is expected to reach it within T_A , and an ambulance is expected to reach it within T_B .

4.2.3. Future uncovered request rate While making redeployment decisions, the state where the redeployed server reaches its new base is more important than the current state. This basis function is parallel to the second basis function, but it replaces the current location of redeployed servers by its destination if we are making redeployment decisions, i.e., $e \in \{5, 9\}$. Denote the future state with the redeployed server arriving at the base by s' . Then the future uncovered rate can be written as

$$\phi_3(s') = \sum_{i \in \mathcal{N}} \lambda_i \mathbf{1}_{\{N_i(s')=0\}}, \quad (12)$$

where the coverage $N_i(s')$ is defined in the same way as in (10) except that s is replaced by s' . When making dispatching decisions, we do not perform the replacement because the server will still be unavailable when it arrives at its destination. In other words, with the “future uncovered request rate”, we would like to maximize the future coverage when the redeployed server becomes available.

4.2.4. Future missed call rate This basis function captures the rate that a request is outsourced or delayed because all the servers are busy with other requests, which is represented as:

$$\phi_4(s) = \sum_{i=1}^N \lambda_i P_i(s), \quad (13)$$

where $P_i(s)$ is the probability that all servers that can reach a request at demand node i are busy with other requests. We estimate $\{P_i(s), i = 1, \dots, N\}$ by treating the request service processes in different demand areas as *Erlang loss systems*. In an Erlang loss system with arrival rate λ , service rate μ , and n servers, the steady-state probability of losing a request is given by $\psi(\lambda, \mu, n) = \frac{(\lambda/\mu)^n / n!}{\sum_{k=0}^n (\lambda/\mu)^k / k!}$. In an EMS system, arrivals follow a Poisson distribution, which satisfies the assumptions of Erlang loss systems. Service times include response time (base to the scene), on-scene time, transport time (scene to hospital), and transition time (hospital to base). However, in our model, service time distributions for ambulances and UAVs are not identical. Fakinos (1980) generalized the Erlang B-formula for the case of heterogeneous servers. Specifically, for a $M/G/k/k$ blocking system with heterogeneous servers, denote the arrival rate at demand node i with λ_i , average service time $\beta_{i,j}, j = 1, \dots, k$ and their product $\rho_{i,j} = \lambda_i \beta_j, j = 1, \dots, k$. Then the probability of blocking is

$$P_i(s) = B_{i,k}(\rho_{i,1}, \dots, \rho_{i,k}) = \frac{\frac{1}{k!} \rho_{i,1} \rho_{i,2} \cdots \rho_{i,k}}{\sum_{v=0}^k \frac{(k-v)!}{k!} \sum_{j_1 < \dots < j_v} \rho_{i,j_1} \rho_{i,j_2} \cdots \rho_{i,j_k}},$$

where $j_1 < j_2 < \dots < j_v$ is a permutation of v servers in $\{1, 2, \dots, k\}$. To estimate parameters in the generalized Erlang B-formula, we let \mathcal{L}_i be the set of available servers that can serve a request in each demand node i within the threshold response time so that $\mathcal{L}_i(s) = \{l \in \mathcal{A}(s) : d(d_l^{a/u}, i) \leq \Delta\}$. Then, we use $k = |\mathcal{L}_i(s)|$ as the number of servers in the Erlang loss system for demand area i .

4.2.5. Average response time This basis function captures the average response time of the two closest servers to each request. We only count the number of servers satisfying certain conditions in the previous three basis functions. With “average response time”, we emphasize the exact distance, directly affecting the response time. Denote by ρ_{li} the probability that server l is dispatched to the request at node i (Chelst and Jarvis 1979). Then the average travel time to each demand node is:

$$T_i = \frac{\sum_l \rho_{li} t_{li}}{\sum_l \rho_{li}},$$

where t_{li} is the average travel time from l to i . The dispatching probability is estimated by the hypercube queueing model developed by Larson (1974), which characterizes the operations of an EMS system with a multiserver-queueing system comprising distinguishable servers. And we estimate the dispatching probability following the approximation procedures described in Larson (1975). Thus the demand-weighted average response time can be written as

$$\phi_5(s) = \sum_{i \in \mathcal{N}} \lambda_i T_i = \sum_{i \in \mathcal{N}} \lambda_i \frac{\sum_l \rho_{li} t_{li}}{\sum_l \rho_{li}}. \quad (14)$$

4.2.6. Availability reduction The previous five basis functions capture the spatial features of the system, whereas this one captures the temporal aspect of availability reduction associated with the traveling of UAVs and ambulances. The travel time matters since UAVs and ambulances are unavailable to serve requests during the travel. For example, consider a scenario where an ambulance can be redeployed to two bases, A and B. Redeployment to A results in a slightly higher coverage but requires a much longer travel time. Without this basis function, we will choose to redeploy the ambulance to base A, but it may not always be the optimal choice. Therefore, we represent availability reduction caused by the travel time of UAVs and ambulance redeployment with the following basis function,

$$\phi_6(s) = \begin{cases} t(d_l(s), d_l(s')) & e \in \{5, 9\} \text{ and } \sum_{b \in \mathcal{M}^a} Z_{l,b}^a = 1 \text{ or } \sum_{b \in \mathcal{M}^u} Z_{l,b}^u = 1, \\ 0 & \text{otherwise.} \end{cases} \quad (15)$$

5. Numerical Experiments

In this section, we present a simulation-based comparative study with realistic UAV design parameters and real EMS data from the National EMS Information System (NEMSIS) database and naloxone administration heatmap in the state of Indiana. We start by introducing the experimental setup and benchmark policies. Then we compare the policies’ performance in scenarios with varying numbers of UAVs, service areas, and base locations. When we investigate the influence of one factor, everything else is fixed to ensure a fair comparison. Next, in Sections 5.6 and 5.7, we analyze the underlying mechanism behind the superior performance of NN-API and provide insights into situations where a simpler approach, such as the static ad-hoc policy, exhibits comparable performance to the NN-API.

5.1. Experimental Setup

In this section, we present an overview of our experimental setup and the calibration criteria employed for the determination of base locations, bystander willingness, EMS demands, performance metrics, and the validation of results. For a comprehensive understanding of the estimation and evidential support for other parameters such as fleet sizes and time parameters, please refer to Section 3 of the Online Supplement. To ensure statistical confidence in our comparisons, we set the simulation time horizon to one day and conducted 400 replications for each case.

Base Locations and Initial Layouts To evaluate the efficiency of dynamic operations of the joint EMS system, we consider the home bases and initial layouts of UAVs and ambulances as predetermined. We leverage a status quo layout: ambulance bases are located at hospitals, police stations and fire departments; bases of UAVs are located at low-ambulance-coverage areas. Initially, numbers of ambulances and UAVs located at each base are proportional to the demand density of the base location. We investigate the impact of optimizing base locations in Section 5.5.

Bystander Willingness The willingness of bystanders to provide assistance in emergency situations exhibits significant variation, as reported in the literature with estimates ranging from 27% to 76% (Kerr et al. 2009, Strang et al. 2000, Barbic et al. 2020). This variability can be attributed to several factors, including sociodemographic characteristics, prior witnessing experience, prior overdose experience, perceived risk of arrest, and the specific location of the overdose incident (Tobin et al. 2005, Burn 2017). Considering the wide range of estimates, we model bystander willingness using a uniform distribution within the interval of [0.2, 0.8].

EMS Demands For our case studies, we extract EMS request distribution from the naloxone administration heatmap in the state of Indiana, with the assumption that opioid overdose requests share the same distribution with other EMS requests. The geographic distribution of these request incidences (measured by the centrality of the distribution) is consistent with the rural-urban area classification (Figure 2). Accordingly, we consider three catchment areas based on their geographic delineation: Marion County (urban), Tippecanoe County (semi-urban), and Marshall County (rural).

Performance Metrics We calculate six metrics: accumulated total rewards, average response time for all requests and for opioid overdose requests, fraction of calls served within the response time threshold for all requests and for opioid overdose requests, and fraction of outsourced requests.

Results Validation The response time of the practical policy is in line with the performance of the current EMS system. Specifically, median EMS arrival times for all call types are between seven and eight minutes (Mell et al. 2017). In rural, remote, geographically challenging, or high-traffic urban areas, this response time can average more than 14 minutes (Hanna 2018).

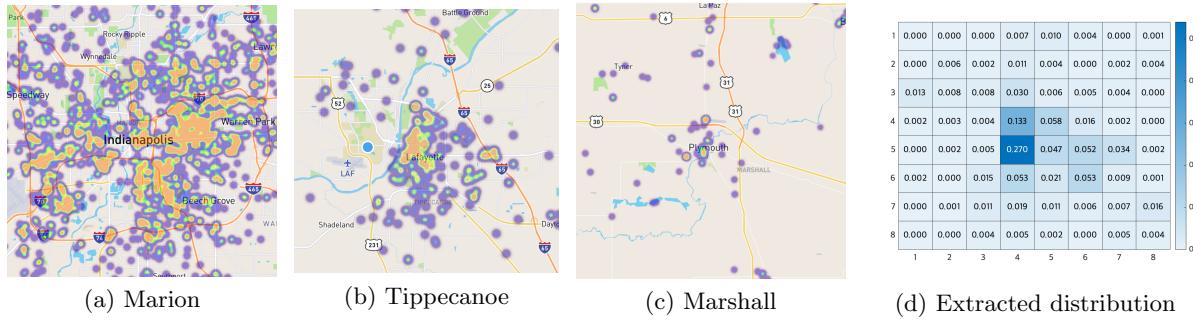


Figure 2 Spatial distribution of naloxone administration for opioid overdose.

In Figures (a)-(c), the yellowish colors (lighter parts in black-and-white mode) represent the highest incident arrival rates, while purplish colors (darker parts in black-and-white mode) represent medium incident arrival rates. Figure (d) shows the probability of an incident occurring in each grid, based on the incident arrival probability distribution heatmap for Tippecanoe County. For example, the value “0.270” in the center grid indicates that there is a 0.27 probability of the next incident occurring in that grid in Tippecanoe County.”

5.2. Benchmark Policies

In this section, we introduce several policies that are used as benchmarks in our study.

Static ad-hoc policy. In current practice, experienced dispatchers make ambulance dispatching decisions in the following ad-hoc manner (Schmid 2012). They can view a dashboard with a regional map that shows the position and status of each server. In case of an emergency, the closest available server is usually dispatched. Servers will return to their home bases after serving a request.

Dynamic heuristic policy. When an EMS request is received, the closest available server is dispatched. The redeployment policy uses the heuristic developed by Jagtenberg et al. (2015), which is based on solving the maximum expected coverage location problem (MEXCLP). This heuristic is easy to implement and has shown good performance for instances considered by the authors. When a server finishes its current task, it will be redeployed to a base that results in the largest marginal contribution to coverage according to the MEXCLP model.

Maxwell-ADP. Maxwell et al. (2010) and Nasrollahzadeh et al. (2018) proposed two ADP-based policies for optimal ambulance redeployment decisions and/or dispatching decisions. We regard them as a predecessor of our solution methodology and thus use them as benchmark policies. We employ the basis functions and coefficient training algorithms proposed by Maxwell et al. (2010) and Nasrollahzadeh et al. (2018), while keeping our modeling framework which includes the system dynamics and the objective function.

L-ADP. In this ADP-based benchmark, we employ the basis functions developed in Section 4.2. However, for feature combination and ADP training, we adopt a linear representation for the value

function approximation and choose the coefficient training algorithms in Maxwell et al. (2010) and Nasrollahzadeh et al. (2018).

5.3. Benefit of Introducing UAVs

In this section, we explore the benefit of the joint operation of UAVs with ambulances. The comparison is based on the arrival profile in Tippecanoe County. We consider three settings of the fleet size, i.e., no UAVs, 8 UAVs, and 16 UAVs. Table 1 suggests that the NN-API policy achieves a shorter response time than the benchmark in all the tested scenarios (the complete comparison results can be found in Section 4.1 of the Online Supplement).

Table 1 Average response time (minutes) comparison (95% confidence interval) among different numbers of UAVs in Tippecanoe County (semi-urban) and Marshall County (rural).

Tippecanoe County		General	Opioid	Marshall County		General	Opioid
0 UAV	NN-API	10.8±0.3	8.3±0.4	0 UAV	NN-API	12.6±0.5	10.8±0.5
	Static	13.7±0.3	12.8±0.5		Static	13.5±0.6	11.5±0.7
8 UAVs	NN-API	9.1±0.4	5.7±0.3	8 UAVs	NN-API	7.8±0.3	5.4±0.3
	Static	10.3±0.2	7.7±0.4		Static	9.3±0.4	7.9±0.5

Table 1 demonstrates that the incorporation of UAVs into an EMS system can significantly reduce the response time, especially for opioid overdose requests. The response time reduction is more significant in rural areas, where ambulance resources are typically limited. For both counties, the response times can be reduced by over 30% with the use of UAVs, even when using a simple operational strategy. With our proposed strategy, response times could be reduced by as much as 50%. This magnitude of reduction in response time could potentially save more lives, as UAVs would be able to arrive at the scene faster and provide life-saving treatment in a timely manner. This could be especially beneficial for opioid overdose victims, who might not have survived without prompt medical intervention. Moreover, Table 1 also demonstrates comparable performances between the NN-API policy with no UAV and static policy with 8 UAVs in Tippecanoe county (semi-urban). This suggests that intelligent real-time operations can compensate for the lack of EMS resources. In general, the addition of UAVs to the EMS system can improve the overall performance and efficiency of emergency response.

5.4. Sensitivity Analysis on Service Areas

In this section, we investigate the sensitivity of the policy performances with respect to the urban-rural delineation of the service areas. As shown in Figure 3, a service area that is more urban (e.g., Marion County) tends to have a higher volume of requests and more pockets of spatially concentrated requests, whereas a service area that is more rural (e.g., Marshall County) tends to be the reverse. Table 2 summarizes our results, suggesting that the NN-API policy performs consistently

better than the benchmarks in terms of accumulated rewards. The complete comparison results can be found in Section 4.2 of the Online Supplement.

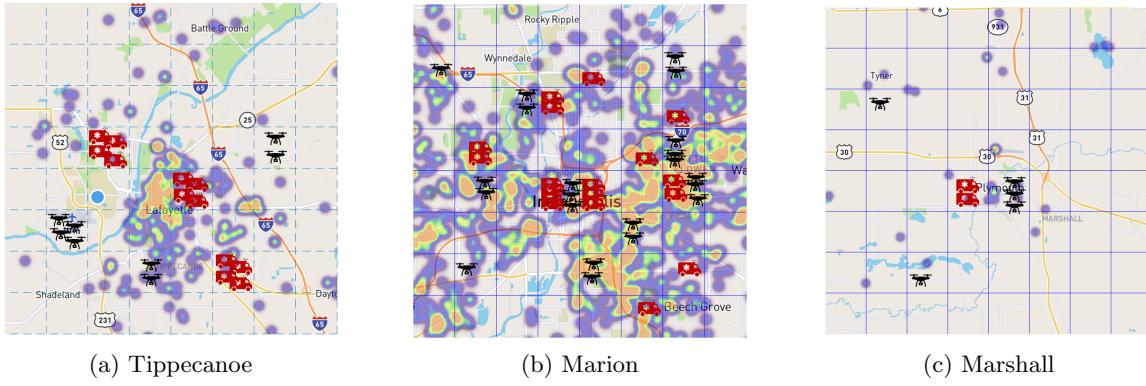


Figure 3 Spatial distribution of naloxone administration and settings of ambulances and UAVs in Tippecanoe, Marion and Marshall Counties.

Table 2 Performance comparison (95% confidence interval) among different service areas.

		Reward	Response Time (Minutes)	Respond within threshold (%)	Outsourced (%)
		General	Opioid	General	Opioid
Tippecanoe	NN-API	2132±40	9.1±0.4	76.6±0.4	78.8 ±0.4
	Static	1882±39	10.3±0.2	74.9±0.5	74.6±0.7
Marion	NN-API	5420±133	6.8±0.3	70.2±0.5	72.6±0.6
	Static	3672±104	7.9±0.6	64.5±0.7	61.2±0.7
Marshall	NN-API	135±10	12.6±0.6	72.6±1.2	74.1±1.7
	Static	112±9	13.5±1.1	67.5±1.1	62.2±1.8

Table 2 suggests that in areas with high demand, such as Marion County, more requests arrived continuously. Hence, ADP-based policies take advantage of considering future requests and lead to significantly better performance, as indicated by a p -value of less than 0.01 in a paired t -test. In low-demand cases like Marshall County, the intervals between requests were longer, making the decision process more similar to a one-shot decision. In these cases, there is barely any benefit from sacrificing current rewards to prepare for future demand, resulting in similar performances between the NN-API policy and the static policy, as indicated by a p -value of 0.12 in a paired t -test.

5.5. Performance Improvement with Optimized Base Locations

In this section, we investigate the impact of the optimization of the home base locations on the policy performance. We use a maximal coverage location problem (MCLP) to jointly optimize the locations of ambulance and UAV bases for maximum coverage. More details about this location model can be found in Section 4.3.1 in the Online Supplement. We compare the performance of our

NN-API policy with the benchmark policies under four different sets of base locations in Tippecanoe County (Figure 4): (a) both random, where both ambulance and UAV bases are randomly located across the service area; (b) status quo-supplementary, which is the baseline setting used in the previous two sections, where ambulance bases are located at hospitals, fire departments, and police stations, while UAV bases are located in low-ambulance-coverage areas; (c) status quo-optimized, where ambulance bases are fixed at status quo locations and only UAV bases are optimized using an MCLP; (d) both optimized, where both ambulance and UAV bases are jointly optimized using a larger-scale MCLP.

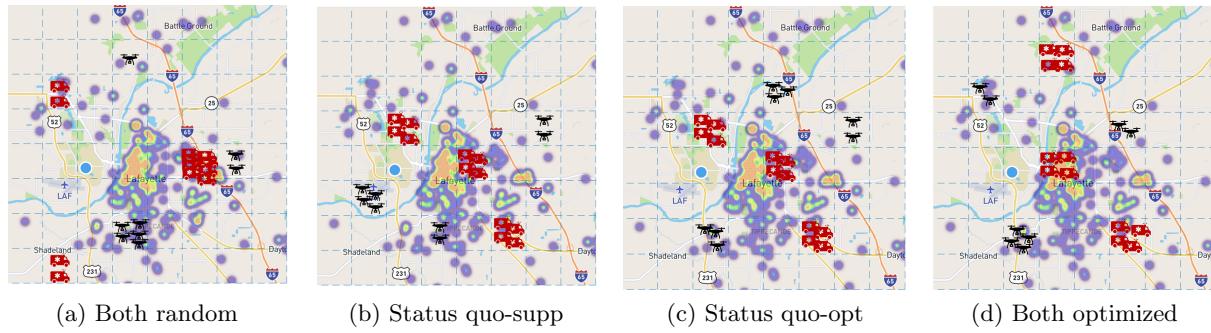


Figure 4 Various settings of base locations of ambulances and UAVs in Tippecanoe County.

To simplify the result presentation and clarify the comparison across different settings, we only present the performance of the NN-API and static ad-hoc policies in Table 3. Full comparative results can be found in Section 4.3.2 of the Online Supplement.

Table 3 Performance comparison (95% confidence interval) under different base locations and initial layouts at Tippecanoe County.

		Reward	Response Time (Minutes)		Respond w/in threshold (%)		Outsourced (%)
			General	Opioid	General	Opioid	
Both random	NN-API	1975±44	11.5±0.4	7.4±0.3	76.3±0.4	78.6±0.5	15.2±0.3
	Static	1735±38	14.1±0.6	8.7±0.4	74.0±0.4	69.5±0.6	17.5±0.4
Status quo-supplementary	NN-API	2132±40	9.1±0.4	5.7±0.3	76.6±0.4	78.8 ±0.4	13.9±0.4
	Static	1882±39	10.3±0.2	7.7±0.4	74.9±0.5	74.6±0.7	14.8±0.4
Status quo-optimized	NN-API	2169±41	9.1±0.4	5.6±0.3	77.1±0.3	79.1 ±0.4	13.9±0.4
	Static	1916±40	10.0±0.2	7.5±0.4	75.3±0.4	75.2±0.6	14.3±0.5
Both optimized	NN-API	2192±41	9.0±0.4	5.5±0.3	77.8±0.3	79.4±0.4	13.6±0.5
	Static	1932±39	9.9±0.2	7.4±0.4	75.4±0.4	75.9±0.6	14.1±0.4

Table 3 shows that our NN-API policy consistently outperforms the benchmark policies in terms of accumulated rewards by reducing response times, particularly for opioid overdoses. This suggests that even with pre-determined base locations, there is always room for improvement in real-time operations, which our NN-API policy can achieve. Furthermore, the table demonstrates a significant

improvement from the *Both random* setting to the *status quo-supplementary* setting under both the NN-API and static policies, highlighting the importance of base locations. Base locations determine the feasible choices for dispatching and relocation decisions, and optimizing them could further improve performance and shorten response times, especially for the static policy. While optimizing base locations can improve coverage, the impact of such optimization for the NN-API policy may be limited to Tippecanoe County due to the centralized pattern of requests. The *status quo-supplementary* setting yields base locations that adequately cover high-demand areas and optimizing base locations would primarily benefit low-demand areas through enhancing coverage. Therefore, the resulting improvement in average response time may not be significant. Nevertheless, optimizing base locations in regions with more dispersed demands could lead to more significant reductions in response time.

5.6. Mechanism behind Performance Superiority

Our experiments in Sections 5.3 - 5.5 consistently demonstrate that the NN-API policy outperforms the benchmark policies, as shown in Tables 1 - 3. The NN-API policy is particularly effective at reducing response times for opioid overdose emergency requests. Table 4 provides an example of the performance comparison between the NN-API policy and all the tested benchmarks. The superior performance of the NN-API policy is driven by the consideration of the following components: dynamic redeployment, sequential events and long-term cost, heterogeneity between UAVs and ambulances, and the complex relationship between value functions and basis functions. The benefits of considering these components are verified by the comparison between different benchmarks.

Table 4 Performance comparison (95% confidence interval) among all benchmark policies at Tippecanoe county with realistic base locations and a moderate number of UAVs.

	Reward	Response Time (Minutes)		Respond within threshold (%)		Outsourced (%)
		General	Opioid	General	Opioid	
Static ad-hoc	1882±39	10.3±0.2	7.7±0.4	74.9±0.5	74.6±0.7	14.8±0.4
Dynamic heuristic	1943±41	9.8±0.2	7.6±0.3	74.9±0.4	74.6±0.7	14.8±0.4
Maxwell-ADP	2002±39	9.3±0.3	7.1±0.4	75.4±0.4	75.1±0.3	14.8±0.2
L-ADP	2090±38	9.7±0.3	6.9±0.3	75.9±0.3	76.2±0.4	14.7±0.3
NN-API	2132±40	9.1±0.4	5.7±0.3	76.6±0.4	78.8±0.4	13.9±0.4

The ad-hoc static policy generally performs poorly compared to the other policies. With the same dispatching policy, the heuristic policy can achieve a shorter response time by maximizing coverage in dynamic redeployment. In general, due to the same dispatching policy, the performances of the static policy and the dynamic heuristic policy do not differ a lot from each other. In most of the test scenarios, the three ADP-based policies significantly outperform the other two policies, with a much shorter response time, particularly for opioid overdose emergency requests. This is because

these policies consider long-term rewards and are able to learn towards an objective function that gives higher weights to opioid overdose cases. Among the three ADP-based policies, the NN-API and L-ADP policies can lead to higher rewards in most of the test scenarios. While Maxwell-ADP occasionally outperforms L-ADP, the NN-API policy, which uses a well-trained neural network, can consistently improve on the performance of L-ADP and outperform Maxwell-ADP. The superior performance of the NN-API policy can be attributed to two factors: the design of the basis functions and the modeling flexibility of the neural network. The basis functions we use take into account the heterogeneity of ambulances and UAVs in terms of service time and incorporate temporal features of the system. This allows our solution method to approximate the value function more accurately, even with a linear approximation model.

5.7. Managerial Insights on Policy Selection

To facilitate the integration of UAVs in EMS and provide operational guidance for policy selection, we collect and analyze 60 sets (environment scenarios) of results. These sets cover a range of demand patterns (centralized, scattered), demand rates (low, medium, high), ambulance and UAV quantities (low, medium, high), and base locations (optimized or unoptimized). Our collected results demonstrate that the NN-API policy outperforms other policies in 59 out of 60 scenarios, with statistical significance. The L-ADP policy and static ad-hoc policy show comparable performance to the NN-API policy in 28 and 11 scenarios, respectively. These findings reaffirm the superiority of the NN-API policy, while also suggesting that the adoption of a “simpler” policy may yield sufficiently good performance under specific circumstances. Thus, we investigate two questions: (1) (NN-best) Under which circumstances is the NN-API policy recommended? (2) (Static-NN-comp) When can the static ad-hoc policy achieve comparable performance to the NN-API policy? To answer these questions, we utilize exact logistic regression with regularization to examine the relationship between the scenario settings and the policy performance. For categorical variables, default values are used: scattered demand patterns, medium demand rates, medium ambulance and UAV quantities, and unoptimized base locations.

The regression coefficients for “NN-best” are shown in Table 5a. These results indicate that when demand is not low and ambulance and UAV availability is limited, adopting the NN-API policy is strongly recommended as it significantly outperforms all other policies. These findings align with the observations made in Sections 5.3-5.5, which suggest the use of the NN-API policy in scenarios where dispatching and redeployment decisions are particularly challenging. This typically occurs when consecutive high-demand situations occur and resources are scarce, where each dispatching decision significantly affects the response time of subsequent events.

Table 5b presents the regression coefficients for “Static-NN-comps,” where the coefficient for “demand_high” is regularized to zero. The regression results suggest that the static ad-hoc policy

Table 5 Regularized Exact Logistic Regression Coefficients for “NN-best” and “Static-NN-comps”. Bold*p-values indicating significance at the 95% confidence level.*

	(a) NN-best			(b) Static-NN-comps		
	Coef.	Std.Err.	P-value	Coef.	Std.Err.	P-value
demand_central	0.037	0.359	0.918	-2.530	0.121	0.000
demand_high	-0.119	0.439	0.786	0.000	-	-
demand_low	-1.229	0.531	0.021	1.055	0.153	0.000
fleet_high	0.375	0.430	0.383	-0.073	0.105	0.485
fleet_low	1.737	0.468	0.016	-0.635	0.110	0.000
base_optimized	0.131	0.387	0.735	0.853	0.099	0.000

has the potential to achieve performance comparable to the NN-API policy under specific conditions. These conditions include low and scattered demand, sufficient availability of ambulances and UAVs, and optimized base locations. On the contrast to the favorable scenarios for NN-best, the Static-NN-comps requires an adequate number of ambulances and UAVs, along with low demand that allows for less consideration of subsequent requests. Additionally, in the case of scattered demand, redeploying to the closest base becomes less disadvantageous as predicting the location of the next request becomes more challenging compared to centralized demand scenarios. Further, optimizing the base locations enhances the potential of improving performance with the static policy. Under optimized base locations, each base can effectively cover its designated area, thereby rendering the redeployment to home bases a satisfactory strategy.

6. Conclusions

In this paper, we develop an ADP approach that makes intelligent real-time decisions in UAV-augmented logistic operations, with a focus on the emergency response to opioid overdose incidents. We formulate this problem with an MDP model that captures the complex system evolution and interplay between the high-dimensional state and a variety of actions. In the face of the solution challenges, we adopt an ADP-based solution approach, which is shown to be efficient in solving real-sized instances. We extend the literature on emergency response operations management by incorporating autonomous vehicles into the modeling. Additionally, we extensively investigate the use of basis functions and neural networks in the ADP framework.

With a detailed event-based simulation model, we evaluate the system performance metrics of the NN-API policy, together with various benchmark policies proposed in the previous literature. We use historical data on naloxone administration for opioid overdose in Indiana and the National EMS information system (NEMSIS) database for our case studies. Our comparative results suggest that our NN-API policy significantly improves system performance over the benchmark policies, especially when the base locations are scattered and the incident rate is high. This improvement is mainly due to the following reasons. First, by adopting an MDP framework, we incorporate future

rewards into consideration when making both dispatching and redeployment decisions. Second, by explicitly considering the heterogeneity of UAVs and ambulances, we can better approximate the value functions. Additionally, our use of basis functions designed to minimize response times and consider the heterogeneity of the system leads to better performance compared to the benchmarks in the literature that only consider a single type of server. Finally, the use of neural networks helps improve value function approximation in a wide range of scenarios. Notably, when demand is significant and ambulance and UAV availability is limited, we strongly recommend the adoption of the NN-API policy due to its outstanding performance surpassing all other tested policies.

In brief, to implement the NN-API algorithm, we first obtain catchment area details from the partnering EMS agency, including the locations of their anticipated UAV and ambulance bases. Using this information, we calibrate the EMS logistics system simulator and develop prediction models for travel times, 9-1-1 request times and locations, and bystander response rates. We then train our recommended policies using the calibrated simulator, our algorithm, and the prediction models, taking into account the level of uncertainty and capacity of the ambulances and UAVs. These policies provide dispatching and redeployment recommendations to the dispatcher based on the status of the ambulances and UAVs (i.e., location and idle/busy status) and the location of any incoming requests.

In this work, we demonstrate the benefits of using UAVs to deliver life-saving medication and provided guidance on how EMS agencies can incorporate UAVs into their operations. Our findings highlight the potential advantages of using UAVs in the EMS field and offer practical guidance on how to effectively utilize UAVs, including strategies for dispatching and redeployment. We hope that this work will contribute to the wider adoption of UAVs in EMS.

Our future research will be in the following directions. First, we will incorporate additional realistic features of UAVs and the EMS system, such as time-varying weather conditions and non-stationary, request distributions in the future. For example, individuals are likely to be at different places over our considered time horizon, e.g., the workplace during daytime and home during nighttime. Second, we will investigate the issue of EMS access equity across diverse communities within a service area, for which we will formulate constrained MDPs with constraints on the allowable system outcome inequity. We will develop an ADP approach to the resultant constrained MDPs. Finally, it is also worth investigating the optimal locations of ambulances and UAV bases given a RL-based operational strategy instead of assuming a closest dispatch rule in the facility location phase. This will allow us to further improve system efficiency and reduce response times.

References

- Niels Agatz, Paul Bouman, and Marie Schmidt. Optimization approaches for the traveling salesman problem with drone. *Transportation Science*, 52(4):965–981, 2018.

- Timothy Amukele, Paul M Ness, Aaron AR Tobian, Joan Boyd, and Jeff Street. Drone transportation of blood products. *Transfusion*, 57(3):582–588, 2017.
- Tobias Andersson and Peter Värbrand. Decision support tools for ambulance dispatch and relocation. *Journal of the Operational Research Society*, 58(2):195–201, 2007.
- Roberto Aringhieri, Maria Elena Bruni, Sara Khodaparasti, and J Theresia Van Essen. Emergency medical services and beyond: Addressing new challenges through a wide literature review. *Computers & Operations Research*, 78:349–368, 2017.
- Damitha Bandara, Maria E Mayorga, and Laura A McLay. Priority dispatching strategies for ems systems. *Journal of the Operational Research Society*, 65(4):572–587, 2014.
- David Barbic, Kevin Duncan, Ryan Trainor, Emily A Ertel, Megan K Enos, Hannah Philips, Floyd Besserer, Brian Grunau, Andrew Kestler, Jim Christenson, et al. A survey of the public’s ability to recognize and willingness to intervene in out-of-hospital cardiac arrest and opioid overdose. *Academic Emergency Medicine*, 27(4):305–308, 2020.
- Valérie Bélanger, A Ruiz, and Patrick Soriano. Recent optimization models and trends in location, relocation, and dispatching of emergency medical vehicles. *European Journal of Operational Research*, 272(1):1–23, 2019.
- Thomas H Blackwell and Jay S Kaufman. Response time effectiveness: comparison of response time and survival in an urban emergency medical services system. *Academic Emergency Medicine*, 9(4):288–295, 2002.
- Rania Boujema, Aida Jebali, Sondes Hammami, and Angel Ruiz. Multi-period stochastic programming models for two-tiered emergency medical service system. *Computers & Operations Research*, 123:104974, 2020.
- Justin J Boutilier, Steven C Brooks, Alyf Janmohamed, Adam Byers, Jason E Buick, Cathy Zhan, Angela P Schoellig, Sheldon Cheskes, Laurie J Morrison, and Timothy CY Chan. Optimizing a drone network to deliver automated external defibrillators. *Circulation*, 135(25):2454–2465, 2017.
- Shawn Meghan Burn. Appeal to bystander interventions: A normative approach to health and risk messaging. In *Oxford research encyclopedia of communication*. 2017.
- Xi-Ren Cao. Basic ideas for event-based optimization of markov systems. *Discrete Event Dynamic Systems*, 15(2):169–197, 2005.
- Rolando Cavazos-Cadena. Recent results on conditions for the existence of average optimal stationary policies. *Annals of Operations Research*, 28(1):3–27, 1991.
- Rolando Cavazos-Cadena and Linn I Sennott. Comparing recent assumptions for the existence of average optimal stationary policies. *Operations research letters*, 11(1):33–37, 1992.
- National Center for Health Statistics CDC. Drug overdose deaths in the u.s. top 100,000 annually, 2022. URL https://www.cdc.gov/nchs/pressroom/nchs_press_releases/2021/20211117.htm.

- Kenneth Chelst and James P Jarvis. Estimating the probability distribution of travel times for urban emergency service systems. *Operations Research*, 27(1):199–204, 1979.
- Xinwei Chen, Marlin W. Ulmer, and Barrett W. Thomas. Deep Q -learning for same-day delivery with a heterogeneous fleet of vehicles and drones. Preprint, submitted October 25, 2019, <https://arxiv.org/abs/1910.11901>, 2019.
- Kenneth C Chong, Shane G Henderson, and Mark E Lewis. The vehicle mix decision in emergency medical service systems. *Manufacturing & Service Operations Management*, 18(3):347–360, 2016.
- Jamal Chu, KH Benjamin Leung, Paul Snobelen, Gordon Nevils, Ian R Drennan, Sheldon Cheskes, and Timothy CY Chan. Machine learning-based dispatch of drone-delivered defibrillators for out-of-hospital cardiac arrest. *Resuscitation*, 162:120–127, 2021.
- A Claesson, L Svensson, P Nordberg, M Ringh, M Rosenqvist, T Djärv, J Samuelsson, O Hernborg, P Dahlbom, A Jansson, et al. Drones may be used to save lives in out of hospital cardiac arrest due to drowning. *Resuscitation*, 114:152–156, 2017a.
- Andreas Claesson, Anders Bäckman, Mattias Ringh, Leif Svensson, Per Nordberg, Therese Djärv, and Jacob Hollenberg. Time to delivery of an automated external defibrillator using a drone for simulated out-of-hospital cardiac arrests vs emergency medical services. *Jama*, 317(22):2332–2334, 2017b.
- Angela K Clark, Christine M Wilder, and Erin L Winstanley. A systematic review of community opioid overdose prevention and naloxone distribution programs. *Journal of addiction medicine*, 8(3):153–163, 2014.
- Mark S Daskin. A maximum expected covering location model: Formulation, properties and heuristic solution. *Transportation Science*, 17(1):48–70, 1983.
- Maya Doe-Simkins, Alexander Y Walley, Andy Epstein, and Peter Moyer. Saved by the nose: bystander-administered intranasal naloxone hydrochloride for opioid overdose. *American journal of public health*, 99(5):788–791, 2009.
- Kevin Dorling, Jordan Heinrichs, Geoffrey G Messier, and Sebastian Magierowski. Vehicle routing problems for drone delivery. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 47(1):70–85, 2016.
- Jurgen Everaerts et al. The use of unmanned aerial vehicles (uavs) for remote sensing and mapping. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37 (2008):1187–1192, 2008.
- D Fakinos. The m/g/k blocking system with heterogeneous servers. *Journal of the Operational Research Society*, 31(10):919–927, 1980.
- Wissam Fawaz, Ribal Atallah, Chadi Assi, and Maurice Khabbaz. Unmanned aerial vehicles as store-carry-forward nodes for vehicular networks. *IEEE Access*, 5:23710–23718, 2017.
- M Hanna. Utility of unmanned aircraft systems (drones) in inner city emergent response during peak rush hour. *Annals of Emergency Medicine*, 72(4):S77–S78, 2018.

- Armann Ingolfsson, Susan Budge, and Erhan Erkut. Optimal ambulance location with random delays and travel times. *Health Care management science*, 11(3):262–274, 2008.
- Caroline J Jagtenberg, Sandjai Bhulai, and Robert D Van der Mei. An efficient heuristic for real-time ambulance redeployment. *Operations Research for Health Care*, 4:27–35, 2015.
- Caroline J Jagtenberg, Sandjai Bhulai, and Robert D Van der Mei. Dynamic ambulance dispatching: Is the closest-idle policy always optimal? *Health Care Management Science*, 20(4):517–531, 2017.
- Phillip R Jenkins, Matthew J Robbins, and Brian J Lunday. Approximate dynamic programming for military medical evacuation dispatching policies. *INFORMS Journal on Computing*, 33(1):2–26, 2020.
- Shucong Jia and Lin Zhang. Modelling unmanned aerial vehicles base station in ground-to-air cooperative networks. *IET Communications*, 11(8):1187–1194, 2017.
- Debra Kerr, Paul Dietze, ANNE-MAREE KELLY, and Damien Jolley. Improved response by peers after witnessed heroin overdose in melbourne. *Drug and alcohol review*, 28(3):327–330, 2009.
- Daniel Kim, Kevin S Irwin, and Kaveh Khoshnood. Expanded access to naloxone: Options for critical response to the epidemic of opioid overdose mortality. *American Journal of Public Health*, 99(3):402–407, 2009.
- Stephen E Lankenau, Karla D Wagner, Karol Silva, Aleksandar Kecojevic, Ellen Iverson, Miles McNeely, and Alex H Kral. Injection drug users trained by overdose prevention programs: responses to witnessed overdoses. *Journal of Community Health*, 38(1):133–141, 2013.
- Richard C Larson. A hypercube queuing model for facility location and redistricting in urban emergency services. *Computers & Operations Research*, 1(1):67–95, 1974.
- Richard C Larson. Approximating the performance of urban emergency service systems. *Operations research*, 23(5):845–868, 1975.
- Amanda D Latimore and Rachel S Bergstein. “caught with a body” yet protected by law? calling 911 for opioid overdose in the context of the good samaritan law. *International Journal of Drug Policy*, 50:82–89, 2017.
- Geoffrey Ling and Nicole Draghic. Aerial drones for blood delivery. *Transfusion*, 59(S2):1608–1611, 2019.
- Andrew S Lockey, Terry P Brown, Jason D Carlyon, and Claire A Hawkes. Impact of community initiatives on non-ems bystander cpr rates in west yorkshire between 2014 and 2018. *Resuscitation Plus*, 6:100115, 2021.
- Vladimir Marianov and Charles ReVelle. The queueing maximal availability location problem: A model for the siting of emergency vehicles. *European Journal of Operational Research*, 93(1):110–120, 1996.
- Matthew S Maxwell, Mateo Restrepo, Shane G Henderson, and Huseyin Topaloglu. Approximate dynamic programming for ambulance redeployment. *INFORMS Journal on Computing*, 22(2):266–281, 2010.
- Laura A McLay. A maximum expected covering location model with two types of servers. *IIE Transactions*, 41(8):730–741, 2009.

- Laura A McLay and Maria E Mayorga. A dispatching model for server-to-customer systems that balances efficiency and equity. *Manufacturing & Service Operations Management*, 15(2):205–220, 2013a.
- Laura A McLay and Maria E Mayorga. A model for optimally dispatching ambulances to emergency calls with classification errors in patient priorities. *IIE Transactions*, 45(1):1–24, 2013b.
- Howard K Mell, Shannon N Mumma, Brian Hiestand, Brendan G Carr, Tara Holland, and Jason Stopryra. Emergency medical services response times in rural, suburban, and urban areas. *JAMA surgery*, 152(10):983–984, 2017.
- Amir Ali Nasrollahzadeh, Amin Khademi, and Maria E Mayorga. Real-time ambulance dispatching and relocation. *Manufacturing & Service Operations Management*, 20(3):467–480, 2018.
- Joseph P Ornato, Alan X You, Gray McDiarmid, Lori Keyser-Marcus, Aaron Surrey, James R Humble, Sirisha Dukkipati, Lacy Harkrader, Shane R Davis, Jacob Moyer, et al. Feasibility of bystander-administered naloxone delivered by drone to opioid overdose victims. *The American Journal of Emergency Medicine*, 38(9):1787–1791, 2020.
- Seong Hyeon Park and Young Hoon Lee. Two-tiered ambulance dispatch and redeployment considering patient severity classification errors. *Journal of Healthcare Engineering*, 2019, 2019.
- Warren B Powell. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*, volume 703. John Wiley & Sons, Hoboken, NJ, 2007.
- Martin L Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, New York, 2014.
- Reuters. Amazon receives U.S. regulatory approval to start drone delivery trials. Accessed February 6, 2021, <https://www.reuters.com/article/us-amazon-prime-air/amazon-receives-u-s-regulatory-approval-to-start-drone-delivery-trials-idUSKBN25R2NG>, Aug 2020.
- Charles ReVelle and Vladimir Marianov. A probabilistic fleet model with individual vehicle reliability requirements. *European Journal of Operational Research*, 53(1):93–105, 1991.
- David Schilling, D Jack Elzinga, Jared Cohon, Richard Church, and Charles ReVelle. The team/fleet models for simultaneous facility and equipment siting. *Transportation science*, 13(2):163–175, 1979.
- Verena Schmid. Solving the dynamic ambulance relocation and dispatching problem using approximate dynamic programming. *European journal of operational research*, 219(3):611–621, 2012.
- David G Schwartz, Janna Ataiants, Alexis Roth, Gabriela Marcu, Inbal Yahav, Benjamin Cocchiaro, Michael Khalemsky, and Stephen Lankenau. Layperson reversal of opioid overdose supported by smartphone alert: A prospective observational cohort study. *EClinicalMedicine*, 25:100474, 2020.
- Jungeun Shin, Lavanya Marla, and Justin Boutilier. Heterogeneous facility location under coordination and collaboration: The case of ambulance-bystander-drone coordination. 6 2022. Available at SSRN: <https://ssrn.com/abstract=3519243>.

- John Strang, David Best, Lan-Ho Man, Alison Noble, and Michael Gossop. Peer-initiated overdose resuscitation: fellow drug users could be mobilised to implement resuscitation. *International Journal of Drug Policy*, 11(6):437–445, 2000.
- Karin E Tobin, Melissa A Davey, and Carl A Latkin. Calling emergency medical services during drug overdose: an examination of individual, social and setting correlates. *Addiction*, 100(3):397–404, 2005.
- Pratap Tokek, Joshua Vander Hook, David Mulla, and Volkan Isler. Sensor planning for a symbiotic UAV and UGV system for precision agriculture. *IEEE Transactions on Robotics*, 32(6):1498–1511, 2016.
- Marlin W Ulmer and Barrett W Thomas. Same-day delivery with heterogeneous fleets of drones and vehicles. *Networks*, 72(4):475–505, 2018.
- Jared Gereau Utah Public Radio. Intermountain healthcare partners with zipline for drone delivery service, Oct 2022. URL <https://www.upr.org/utah-news/2022-10-10/intermountain-healthcare-partners-with-zipline-for-drone-delivery-service>.
- Elizabeth Ty Wilde. Do emergency medical system response times matter for health outcomes? *Health economics*, 22(7):790–806, 2013.
- World Health Organization. Drones take Rwanda's national blood service to new heights. Accessed February 6, 2021, <https://www.who.int/news-room/feature-stories/detail/drones-take-rwandas-national-blood-service-to-new-heights>, June 2019.
- Soovin Yoon and Laura A Albert. A dynamic ambulance routing model with multiple response. *Transportation Research Part E: Logistics and Transportation Review*, 133:101807, 2020.
- Soovin Yoon and Laura A Albert. Dynamic dispatch policies for emergency response with multiple types of vehicles. *Transportation Research Part E: Logistics and Transportation Review*, 152:102405, 2021.
- Soovin Yoon, Laura A Albert, and Veronica M White. A stochastic programming approach for locating and dispatching two types of ambulances. *Transportation Science*, 55(2):275–296, 2021.

Online Supplement - Shortening EMS Response Time with joint operations of UAVs with ambulances

1 Neural Network

In this section, we describe the design and training of the neural network (NN) used in our NN-API algorithm. The NN is used to approximate the value of being in post-decision state S_t^a . It consists of three layers: an input layer, a hidden layer, and an output layer.

According to Heaton (2008), an NN with one or two hidden layers is sufficient to approximate any continuous functions, while a deeper NN can automatically extract features for improved performance. In our case, we have manually constructed the features (i.e., basis functions) based on our understanding of the dispatching and relocation problem for heterogeneous servers. Therefore, a shallow NN is sufficient to approximate the value function as a nonlinear function of the basis functions.

After preliminary tests, we find that the performance of NN with one hidden layer is comparable to those with two hidden layers. Considering interpretability of the model and its training cost, we focus on choosing the number of hidden neurons and training the coefficients for the one-hidden-layer NN. The hidden layer comprises a set of activation units $\mathcal{H} = \{1, 2, \dots, |\mathcal{H}|\}$, i.e., nonlinear perceptron nodes. The input of the hidden layer is a linear combination of the input layer, i.e., basis function, which is given by

$$Y_h^{(2)}(S_t^a) = \sum_{f \in \mathcal{F}} \Theta^{(1)} \phi_f^s(S_t^a), \forall h \in \mathcal{H},$$

where $\Theta^{(1)} \equiv [\Theta_{f,h}^{(1)}]_{f \in \mathcal{F}, h \in \mathcal{H}}$ is an $|\mathcal{F}| \times |\mathcal{H}|$ matrix of weights controlling the function mapping from the input layer to the hidden layer. A nonlinear logistic sigmoid activation function

$$\sigma(y) = \frac{1}{1 - e^{-y}}$$

is applied to each $Y_h^{(2)}(S_t^a)$ to produce the inputs for the hidden layer, which are given by

$$Z_h^{(2)}(S_t^a) = \sigma(Y_h^{(2)}(S_t^a)), \forall h \in \mathcal{H}.$$

The hidden layer produces a single scalar output, which is given by

$$Y^{(3)}(S_t^a) = \sum_{h \in \mathcal{H}} \Theta^{(2)} Z_h^{(2)}(S_t^a),$$

where $\Theta^{(2)} \equiv [\Theta_h^{(2)}]_{h \in \mathcal{H}}$ is an $|\mathcal{H}| \times 1$ matrix of weights controlling the function mapping from the hidden layer to the output layer. The output layer produces a single scalar output by applying the sigmoid activation function to $Y^{(3)}(S_t^a)$, which gives the approximated value function:

$$\bar{V}^a(S_t^a | \Theta) = \sigma(Y^{(3)}(S_t^a))$$

wherein $\Theta = (\Theta^{(1)}, \Theta^{(2)})$ is the parameter representing the NN weights.

It is important to note that before using the basis functions in the NN model, we scale them using a normalization procedure. This involves transforming each value of the $|\mathcal{F}|$ features by subtracting its mean and dividing by its range. Scaling the inputs has several benefits. According to Hastie et al. (2009), it ensures that all input dimensions are treated equally in the regularization process, allows for the selection of meaningful initial weights, and enables more effective optimization when updating the NN weights. In addition, it helps prevent one input dimension from dominating the others during training.

Given Θ , our *NN – API* algorithm makes decisions utilizing the decision function,

$$\pi(S_t|\Theta) = \arg \max_{a \in \mathcal{A}_t} \{C(S_t, x_t) + \bar{V}^a(S_t^a|\Theta)\}.$$

The NN-API algorithm begins by initializing the weight matrices $\Theta^{(1)}$ and $\Theta^{(2)}$, which together form the mapping function Θ , and an initial fixed policy. To initialize the weights, we follow the recommendation of Hastie et al. (2009) and select small, random values near zero. This weight initialization policy has been shown to improve the NN model performance when the weights are later updated in the policy-improvement phase using a quasi-Newton optimization solution procedure. By starting with small, random values, we can avoid the possibility of getting stuck in a local minimum and allow the optimization algorithm to find the global minimum more efficiently.

In the policy evaluation phase, the following steps occur at each iteration. We randomly select a post-decision state $S_{t-1,m}^a$ and compute the associated scaled basis function evaluation $\phi^s(S_{t-1,m}^a)$. Utilizing the NN, we compute the value function approximation $\bar{V}^a(S_{t-1,m}^a|\Theta)$ and record the value. We then simulate the system evolving from post-decision state S_{t-1}^a to a pre-decision state S_t , and we compute (and record) a sample realization of the value attained from the current policy by solving

$$\hat{v} = \max_{a \in \mathcal{A}_t} \{C(S_t, x_t) + \bar{V}^a(S_t^a|\Theta)\} - \hat{v}_g,$$

where \hat{v}_g is the estimated average reward. At each iteration, we collect a total of M sample realizations of the value attained by following the current policy.

In the policy improvement phase, the following steps occur for each iteration. We normalize the value function sample realizations collected in the just-completed iteration of the policy-evaluation phase, so the NN can be properly fit to the collected approximate value function data. We seek Θ -values that optimize the model fitting. We utilize a regularized, mean-squared error measure of the loss function:

$$L(\Theta) = \frac{1}{2M} \sum_{m=1}^M (\hat{v}_m - \bar{V}^x(S_{t-1,m}^a|\Theta))^2 + \frac{\eta}{2M} \left(\sum_{f \in \mathcal{F}} \sum_{h \in \mathcal{H}} (\theta_{f,h}^{(1)})^2 + \sum_{h \in \mathcal{H}} (\Theta_h^{(2)})^2 \right).$$

The penalty term in the loss function prevents overfitting the data and reduces the generalization error. The regularization (i.e., weight decay) parameter $\eta \geq 0$ is a tunable parameter; larger η -values will tend to shrink Θ toward zero. Then the new sample estimate of Θ is determined by Newton's methods as a minimizer of:

$$\hat{\Theta} = \arg \min_{\Theta} L(\Theta).$$

Once Θ is updated via the above equation, we have completed one iteration of the policy-improvement phase of the NN-API algorithm.

As shown in Figure 1, the training loss in our NN model rapidly converges to zero during each policy improvement phase, owing to the straightforward structure of NNs.

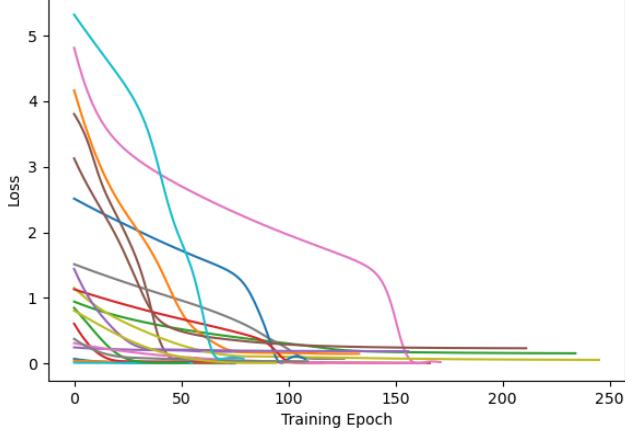


Figure 1: The convergence of loss functions in policy improvement phases.

2 Proof of Proposition 1

2.1 Value Function Monotonicity on Bystander Helping Probability

Consider two states, s_0 and $s_0 + \delta w_j$, with $\delta w_j > 0$. We prove $V(s_0) \leq V(s_0 + \delta w_j)$ by discussing the optimal action at these two states based on the Poisson equation:

$$V^*(s) + V_g = \max_{a \in \mathcal{A}} \sum_{s'} P(s'|s, a)[h(s, a, s') + V^*(s')], \quad \forall s \in \mathcal{S}.$$

The difference in bystander helping probability only affects the value function when the dispatching decision is made, i.e., the event type is “request j arrives”. Thus, we discuss the possible nine combinations of dispatching decisions at s_0 and $s_0 + \delta w_j$.

Case 1. $\pi^*(s_0) = \pi^*(s_0 + \delta w_j)$ = “outsourcing” or “dispatching an ambulance” Since varying the bystander helping probability does not affect the operation of ambulances, the same ambulance should be dispatched in s_0 and $s_0 + \delta w_j$. Thus, $\forall s' \in \mathcal{S}$, we have

$$\begin{aligned} P(s'|s_0, a) &= P(s'|s_0 + \delta w_j, a) \\ h(s_0, a, s') &= h(s_0 + \delta w_j, a, s') \end{aligned}$$

From the Poisson equation, we have $V^*(s_0) = V^*(s_0 + \delta w_j)$.

Case 2. $\pi^*(s_0) = \pi^*(s_0 + \delta w_j)$ = “dispatching a UAV and an ambulance” Since the effect of the bystander helping probability on UAV operation is consistent for all UAVs, the same UAV should be dispatched in s_0 and $s_0 + \delta w_j$. So is the ambulance that should be dispatched. We couple the subsequent systems following s_0 and $s_0 + \delta w_j$ as they have identical incident arrivals, corresponding traveling times, corresponding on-scene times, etc. Then the only difference between the two systems is the bystander availability when the UAV arrives at the scene since the on-scene time of UAVs is negligible. Then the expected rewards when the UAV arrives at the scene are:

$$\begin{aligned} c_U(s_0) &= w_j \cdot g(t_0) \\ c_U(s_0 + \delta w_j) &= (w_j + \delta w_j) \cdot g(t_0). \end{aligned}$$

Thus, $c_U(s_0 + \delta w_j) > c_U(s_0)$. Since all the other subsequent costs/rewards are the same, we have $V^*(s_0) < V^*(s_0 + \delta w_j)$.

Case 3. $\pi^*(s_0)$ = “dispatching an ambulance” and $\pi^*(s_0 + \delta w_j)$ = “dispatching a UAV and an ambulance” Define a suboptimal policy π_1 as

$$\pi_1(s) = \begin{cases} \text{dispatching an ambulance,} & s = s_0 + \delta w_j; \\ \pi^*(s_0), & \text{otherwise.} \end{cases}$$

From the optimality of π^* , we have $V_1(s_0 + \delta w_j) \leq V^*(s_0 + \delta w_j)$.

Since varying the bystander helping probability does not affect the operation of ambulances, the ambulance dispatched under $\pi_1(s_0 + \delta w_j)$ is the same as that under $\pi^*(s_0)$. So are the subsequent systems. Thus, we have $V_1(s_0 + \delta w_j) = V^*(s_0)$. This implies that $V^*(s_0) = V_1(s_0 + \delta w_j) \leq V^*(s_0 + \delta w_j)$.

Case 4. $\pi^*(s_0)$ = “outsourcing” and $\pi^*(s_0 + \delta w_j)$ = “dispatching a UAV and an ambulance” Define a suboptimal policy π_1 as

$$\pi_1(s) = \begin{cases} \text{outsourcing,} & s = s_0 + \delta w_j; \\ \pi^*(s_0), & \text{otherwise.} \end{cases}$$

From the optimality of π^* , we have $V_1(s_0 + \delta w_j) \leq V^*(s_0 + \delta w_j)$.

The subsequent systems following s_0 under π^* and following $s_0 + \delta w_j$ under π_1 are identical. Thus, we have $V_1(s_0 + \delta w_j) = V^*(s_0)$, which implies that $V^*(s_0) = V_1(s_0 + \delta w_j) \leq V^*(s_0 + \delta w_j)$.

Case 5. $\pi^*(s_0)$ = “dispatching a UAV and an ambulance” and $\pi^*(s_0 + \delta w_j)$ = “outsourcing” or “dispatching an ambulance” We prove it is impossible that $\pi^*(s_0)$ = “dispatching a UAV and an ambulance” and $\pi^*(s_0 + \delta w_j)$ = “dispatching an ambulance”.

Define two suboptimal policies π_1 as

$$\pi_1(s) = \begin{cases} \text{dispatching an ambulance,} & s = s_0; \\ \pi^*(s_0), & \text{otherwise.} \end{cases}$$

$$\pi_1(s) = \begin{cases} \text{dispatching a UAV and an ambulance,} & s = s_0 + \delta w_j; \\ \pi^*(s_0), & \text{otherwise.} \end{cases}$$

Consider the two systems (1) following s_0 under π_1 and (2) following $s_0 + \delta w_j$ under π^* , where both policies choose to dispatch an ambulance. Since varying the bystander helping probability does not affect the operation of ambulances, the ambulance dispatched under $\pi_1(s_0)$ is the same as that under $\pi^*(s_0 + \delta w_j)$. So are the subsequent systems. Hence, we have $V_1(s_0) = V^*(s_0 + \delta w_j)$.

Consider the two systems (1) following s_0 under π^* and (2) following $s_0 + \delta w_j$ under π_2 , where both policies choose to dispatch a UAV and an ambulance. Since the effect of bystander helping probability on UAV operation is consistent for all UAVs, the same UAV should be dispatched in s_0 and $s_0 + \delta w_j$. So as the dispatched ambulance. We couple the subsequent systems following s_0 and $s_0 + \delta w_j$ as they have identical incident arrivals, corresponding traveling times, corresponding on-scene times, etc. Then the only difference between the two systems is the bystander availability when the UAV arrives at the scene since the on-scene time of UAVs is negligible. Then the expected rewards when the UAV arrives at the scene are:

$$\begin{aligned} c_U(s_0) &= w_j \cdot g(t_0), \\ c_U(s_0 + \delta w_j) &= (w_j + \delta w_j) \cdot g(t_0). \end{aligned}$$

Thus, $c_U(s_0 + \delta w_j) > c_U(s_0)$. Since all the other subsequent costs/rewards are the same, we have $V^*(s_0) < V_2(s_0 + \delta w_j)$.

From the optimality of π^* at s_0 , we have $V_1(s_0) \leq V^*(s_0)$. This further implies that

$$V^*(s_0 + \delta w_j) = V_1(s_0) \leq V^*(s_0) < V_2(s_0 + \delta w_j),$$

which contradicts with the optimality of π^* at $s_0 + \delta w_j$. Therefore, it is impossible that $\pi^*(s_0) = \text{“dispatching a UAV and an ambulance”}$ and $\pi^*(s_0 + \delta w_j) = \text{“dispatching an ambulance”}$.

It can be proved with a similar argument that it is impossible that $\pi^*(s_0) = \text{“dispatching a UAV and an ambulance”}$ and $\pi^*(s_0 + \delta w_j) = \text{“outsourcing”}$.

Case 6. $\pi^*(s_0) = \text{“dispatching an ambulance”}$ and $\pi^*(s_0 + \delta w_j) = \text{“outsourcing”}$ or $\pi^*(s_0) = \text{“outsourcing”}$ and $\pi^*(s_0 + \delta w_j) = \text{“dispatching an ambulance”}$ Since varying the bystander helping probability does not affect the operation of ambulances, these two combinations of optimal actions are impossible to occur. \square

2.2 Value Function Monotonicity on Ambulances/UAVs Availability

Without loss of generality, we consider two states s_0 and s_1 , where ambulance 1 is busy at s_0 and idle at s_1 , and everything else is the same for s_0 and s_1 .

Define a suboptimal policy π_1 as

$$\pi_1(s) = \begin{cases} \pi^*(s_0), & s = s_1; \\ \pi^*(s), & s \neq s_1. \end{cases}$$

That is, we ignore the additional available ambulance and only consider dispatching ambulances that are available at s_0 if a request arrives.

Then the subsequent system following s_1 under π_1 is identical to that following s_0 under π^* . Thus, we have

$$V^*(s_0) = V_1(s_1) \leq V^*(s_1).$$

\square

3 Experimental Setup Details

Fleet sizes Finding out how many ambulances there are serving one particular EMS agency in the US is difficult due to the various ways EMS is run. In some places, ambulances are part of the fire department’s responsibility. In some other places, ambulance services are run by private companies. There is also a third service mode, like the city of Boston, where Police, Fire, and Ambulance are all working for the municipality. According to National Association of State EMS Officials and others (2020), there are 2384 EMS vehicles in the state of Indiana, with ambulances making up 84% of them. We assume that the number of ambulances is proportional to the population size. We set up the number of ambulances in Marion, Tippecanoe, Johnson, and Marshall counties of Indiana as 288, 55, 47, and 13, respectively. Using similar methods, we estimate the numbers of EMS agencies <https://www.countyoffice.org/in-ambulance/> and numbers of 9-1-1 responses <https://www.in.gov/dhs/ems/ems-data/>. A summary of the estimates is shown in Table 1. For UAVs, we assume that charging and supplement stations are located in the areas with the highest opioid overdose rates.

UAV Parameters We consider top-of-range consumer UAVs that are capable of flying for up to one hour with an average speed of 45 miles per hour. Our preliminary work (see Gao et al. (2020)) indicates that the type and capacity of UAVs typically only influence the magnitude of the performance gaps between policies, rather than the relative rankings of the performances.

County	EMS agencies	Ambulances	Daily responses
Marion	13	288	302
Tippecanoe	3	55	57
Johnson	2	47	49
Marshall	5	13	14

Table 1: Estimated numbers of EMS agencies, numbers of ambulances, numbers of 9-1-1 responses in Marion, Tippecanoe, Johnson, and Marshall counties in the state of Indiana, 2020.

EMS Demands The pattern of the demand differs not only by its geographic distribution but also by its overall volume of requests (Table 2). We look into the 2020 NEMSIS database established by the National Highway Traffic Safety Administration (NHTSA) to determine the proportion of type-1 requests to all EMS requests. In total, 33,042,532 events with complete entries are included in the 2020 NEMSIS data, with 214,295 of them (0.65%) connected to ICD-10 codes implying opioid-associated disorders (F11 codes) and poisoning by (and adverse effects of) opioid-related medicines (T40 codes). To make the performance comparison more evident, we increase the percentage of Type-1 requests in all EMS requests to 10%, which is reasonable in terms of including additional causes of Type-1 requests, such as out-of-hospital cardiac arrest (UAV-delivered AED) and traumatic injury (UAV-delivered blood transfusion kit).

For our case studies, we extract real-world data from the naloxone administration heatmap in the state of Indiana¹. This heatmap provides information on suspected opioid overdose incidence locations from 2014 to 2021. These encounters (demands) were initially requested by patients themselves and bystanders. We partition each county into an 8×8 grid and extract the geographic distribution of the request arrivals by analyzing the RGB features of the corresponding heatmap. We thus simulate geographic-specific demand arrivals based on the heatmap of the four Indiana counties. Each of these demand nodes is represented by the centroid of a 1×1 square on the grid.

Heatmap used	Country type	Overall arrival rate	Requests Distribution
Marion County	urban	high	dense everywhere
Tippecanoe County	semi-urb	medium	dense only in the center
Johnson County	semi-urb	medium	dense in the north and center
Marshall County	rural	low	scattered

Table 2: Features of different incidences distribution.

Time parameters In the simulator, we use $t = t_0 + \frac{s}{v} + \epsilon$ to model the travel times of UAVs and ambulances, where t_0 refers to chute time, s is the travel distance (Manhattan distance for ambulances and Euclidean distance for UAVs), v is the travel speed, and $\epsilon \sim N(0, \sigma^2)$ is a random disturbance reflecting the influence of traffic congestion, building obstruction, and weather condition. The average service time derived using NEMSIS data is shown in Table 3. The on-scene time varies mildly depending on the specific location of the request. As a result, we use an empirical distribution generated from NEMSIS data to sample the on-scene times (Figure 2). Due to a lack of data, we assume the same on-scene time distribution for UAVs and ambulances. According to the most up-to-date EMS monthly report of the State of Indiana Emergency Medical Services (August 2021), patient transport is required in 75,776 (81.4%) EMS missions out of a total of 93,094

¹<https://www.in.gov/recovery/naloxone/heatmap.html>

valid ones. We thus assume that patients need transportation with a probability of 0.8. Since incorporating the hospital choice for ambulances is impractical, we sample ambulance transport time straightly from the NEMSIS data (Figure 2), even though the transport time is likely also dependent on the distance between the requests and hospitals.

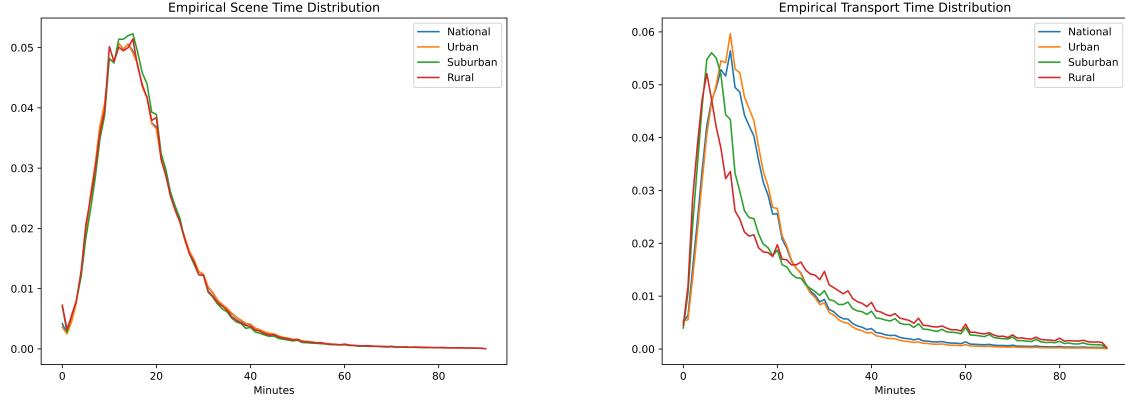


Figure 2: Empirical distribution of on-scene time and transportation time extracted from NEMSIS data in different types of areas.

Area	Decision time	Response time	On-scene time	Transport time
Urban	2.8	11.6	18.4	15.9
Semi-urban	2.4	12.6	18.1	21.6
Rural	2.1	13.6	18.3	26.0
Wilderness	2.5	16.8	19.3	32.4

Decision time: time between the 9-1-1 call and an emergency vehicle notified for dispatching. Response time: time between the 9-1-1 call and the vehicle arrival on the scene. On-scene time: time between the vehicle arriving and leaving the scene. Transport time: time between the vehicle leaving the scene and arriving at the hospital. Source: 2020 NEMSIS database.

Table 3: Ambulance service time distribution (in minutes) estimated from NEMSIS database.

Simulation parameters The time horizon for the simulation is set to be one day. This configuration balances the requirement of setting up the infinite-horizon MDP problem and that of making realistic UAV logistic operations. Finally, we simulate each setting (i.e., one service area type, one fleet size, and one overall arrival rate level) for 400 replications to ensure statistical confidence in the comparisons. Taking EMS delivery in Tippecanoe county (a semi-urban county) during regular hours and with a moderate number of ambulances as an example, we show in Figure 3 that the length of simulation duration and the number of simulation replications are sufficient to stabilize the average reward, which leads to meaningful comparisons.

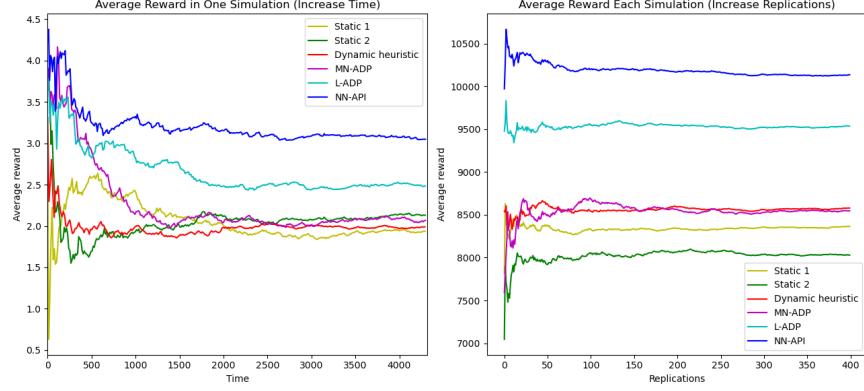


Figure 3: Left: The trend of average reward in one simulation replication, i.e., $\frac{\sum_{\tau=1}^t h(t)}{t}$ vs. t , where t is time. Right: The trend of average total rewards, i.e., $\frac{\sum_{k=1}^n \sum_{\tau=1}^T h(t)}{n}$ vs. n , where n is the index of simulation, and T is simulation length.

4 Sensitivity Analysis

4.1 Varying Fleet Sizes

In this section, we present the complete comparison results on the outcomes of our NN-API policy and benchmark policies under various fleet sizes (Table 4).

		Reward	Response Time (Min)		Respond w/in threshold (%)		Outsourced (%)
			General	Opioid	General	Opioid	
Tippecanoe 0 UAV	NN-API	1849±36	10.8±0.5	8.3±0.4	73.5±0.4	70.8±0.5	15.0±0.5
	L-ADP	1752±36	12.1±0.6	10.2±0.5	72.8±0.5	69.9±0.6	15.8±0.6
	Maxwell-ADP	1726±36	12.1±0.4	11.6±0.6	71.9±0.5	69.3±0.5	16.7±0.6
	Heuristic	1683±35	12.6±0.5	12.4±0.6	71.3±0.5	67.9±0.7	16.7±0.5
	Static	1651±35	13.7±0.3	12.8±0.5	71.1±0.6	67.5±0.8	16.9±0.6
Tippecanoe 8 UAVs	NN-API	2132±40	9.1±0.4	5.7±0.3	76.6±0.4	78.8±0.4	13.9±0.4
	L-ADP	2090±38	9.7±0.3	6.9±0.3	75.9±0.3	76.2±0.4	14.7±0.3
	Maxwell-ADP	2002±39	9.3±0.3	7.1±0.4	75.4±0.4	75.1±0.3	14.8±0.2
	Heuristic	1943±41	9.8±0.2	7.6±0.3	74.9±0.4	74.6±0.7	14.8±0.4
	Static	1882±39	10.3±0.2	7.7±0.4	74.9±0.5	74.6±0.7	14.8±0.4
Marshall 0 UAV	NN-API	105±14	12.6±0.5	10.8±0.5	69.3±0.5	67.2±0.8	17.0±0.8
	L-ADP	103±14	11.6±0.6	10.9±0.7	68.8±0.6	64.5±0.5	17.3±0.7
	Maxwell-ADP	99±13	11.7±0.6	11.3±0.6	67.3±0.5	63.5±0.5	17.2±0.8
	Heuristic	99±14	12.8±0.7	11.4±0.7	67.6±0.6	63.2±0.5	17.2±0.8
	Static	98±14	13.5±0.6	11.5±0.7	67.7±0.7	63.1±0.6	17.3±0.8
Marshall 8 UAVs	NN-API	173±16	7.8±0.3	5.4±0.3	78.5±0.6	75.6±0.8	15.3±0.7
	L-ADP	154±16	8.1±0.4	6.1±0.5	76.1±0.6	74.8±0.6	15.9±0.6
	Maxwell-ADP	149±14	9.3±0.5	7.8±0.6	71.6±0.5	71.9±0.6	16.5±0.7
	Heuristic	147±15	9.3±0.5	7.8±0.4	71.1±0.6	70.8±0.6	16.7±0.7
	Static	144±14	9.3±0.4	7.9±0.5	70.2±0.6	69.3±0.5	16.8±0.7

Table 4: Performance comparison (95% confidence interval) under different fleet sizes.

4.2 Varying Service Areas

In this section, we present the complete comparison results on the outcomes of our NN-API policy and benchmark policies under various service areas (Table 5).

		Reward	Response Time (Min)	Respond w/in threshold (%)	Outsourced (%)		
		General	Opioid	General	Opioid		
Tippecanoe	NN-API	2132±40	9.1±0.4	5.7±0.3	76.6±0.4	78.8±0.4	13.9±0.4
	L-ADP	2090±38	9.7±0.3	6.9±0.3	75.9±0.3	76.2±0.4	14.7±0.3
	Maxwell-ADP	2002±39	9.3±0.3	7.1±0.4	75.4±0.4	75.1±0.3	14.8±0.2
	Heuristic	1943±41	9.8±0.2	7.6±0.3	74.9±0.4	74.6±0.7	14.8±0.4
	Static	1882±39	10.3±0.2	7.7±0.4	74.9±0.5	74.6±0.7	14.8±0.4
Marion	NN-API	5420±133	6.8±0.3	4.9±0.2	70.2±0.5	72.6±0.6	21.9±0.4
	L-ADP	5317±129	6.9±0.4	5.2±0.5	71.9±0.5	71.8±0.6	22.9±0.5
	Maxwell-ADP	3999±118	7.4±0.6	6.6±0.4	66.6±0.6	64.7±0.5	22.9±0.5
	Heuristic	3683±108	7.7±0.5	6.9±0.6	65.4±0.7	62.9±0.6	25.2±0.7
	Static	3672±104	7.9±0.6	7.4±0.5	64.5±0.7	61.2±0.7	25.9±0.8
Marshall	NN-API	135±10	12.6±0.6	8.1±0.6	72.6±1.2	74.1±1.7	17.5±1.0
	L-ADP	125±10	10.9±0.9	8.5±0.7	71.7±0.9	73.1±1.6	17.3±0.8
	Maxwell-ADP	121±10	10.9±1.0	8.7±0.9	72.1±0.9	69.9±1.5	17.6±1.0
	Heuristic	118±9	11.3±1.0	8.9±0.9	71.2±1.0	67.5±1.6	17.8±0.9
	Static	112±9	13.5±1.1	9.2±0.8	67.5±1.1	62.2±1.8	18.1±0.9

Table 5: Performance comparison (95% confidence interval) under different service areas.

4.3 Performance Improvement with Optimized Base Locations

In this section, we first introduce the maximal coverage location problem (MCLP) for the ambulance bases and UAV bases optimization. Then we present the complete comparison results on the outcomes of our NN-API policy and benchmark policies under various base locations and initial layouts (Table 7).

4.3.1 Base Location Optimization

In this section, we first present an adapted version of the classic MCLP for incorporation of the joint coverage of ambulances and UAVs.

Following is a presentation of the variables, parameters and the optimization model:

Objective function

$$\max \sum_k h_k W_k \quad (1)$$

Symbol	Meaning
Decision Variables	
X_i	$X_i = 1$ if an ambulance base is located at node i
Y_j	$Y_j = 1$ if a UAV base is located at node j
Auxiliary Variables	
W_k	$W_k = 1$ if demand node k is covered
Z_{ij}	$Z_{ij} = X_i Y_j$ denotes the joint location by the ambulance base X_i and UAV base Y_j
Parameters	
h_k	Demand (request density) at node k
a_{ki}	$a_{ki} = 1$ if an ambulance base at i covers demands at node k
b_{kij}	$b_{kij} = 1$ if an ambulance base at i and a UAV base at j cover demands at node k
M_a, M_u	Numbers of ambulance bases and UAV bases to be located
N	Number of nodes in the network. In our case, $N = 64$.

Table 6: Notation in the base location optimization

Subject to

$$W_k - \sum_i a_{ki} X_i - \sum_i \sum_j b_{kij} Z_{ij} \leq 0, \quad k = 1, \dots, N \quad (2)$$

$$\sum_i X_i \leq M_a \quad (3)$$

$$\sum_j Y_j \leq M_u \quad (4)$$

$$Z_{ij} \leq X_i \quad i, j = 1, \dots, N \quad (5)$$

$$Z_{ij} \leq Y_j \quad i, j = 1, \dots, N \quad (6)$$

$$Z_{ij} \geq X_i + Y_j - 1 \quad i, j = 1, \dots, N \quad (7)$$

$$X_i, Y_j, Z_{ij}, W_k \in \{0, 1\}, \quad i, j, k = 1, \dots, N \quad (8)$$

The objective (1) is to maximize the weighted coverage $\sum_k h_k W_k$, where h_k represents the demand density at node k and W_k is an auxiliary variable denoting the binary coverage. Constraint (2) ensures that the coverage variable $W_k = 1$ only if the demand node k is covered by an ambulance base i or under joint coverage of ambulance base i and UAV base j . Constraints (3) and (4) ensure the total numbers of ambulance bases and UAV bases do not exceed M_a and M_u , respectively. Constraints (5) - (7) ensure the binary variable $Z_{ij} = X_i Y_j$. Constraints (8) ensure variables X_i, Y_j, Z_{ij}, W_k are binary.

For the coverage definition, a node is covered if an ambulance base is located within 8 minute or a UAV base is located within 8 minutes and an ambulance base is within 20 minutes (Doe-Simkins et al. 2009, Byrne et al. 2019).

For the layout with status quo ambulance bases and optimized UAV bases, we modify the decision variables X_i to predetermined parameters based on the status quo setting.

For the joint location problem, the number of decision variables is $N + N$, in our case, $2N = 128$. The number of constraints is $N + 2 + 3N^2$, in our case, 12354. We need to pre-compute and store values of a_{ki} and b_{kij} , of sizes 4096 and 262144, for which we can use a sparse matrix, since most of the elements are zero. Both binary linear programs can be solved efficiently within one minute by Gurobi 10.0.1.

4.3.2 Detailed Results with Various Sets of Base Locations

In this section, we present the detailed results in Table 7 under various sets of base locations under four settings: (a) (Status quo-supplementary) bases of ambulances are located at hospitals, fire departments, and police stations, while bases of UAVs are located at low-ambulance-coverage areas. (b) (Both random) bases of ambulances and UAVs are randomly located across the service area. (c) (Both optimized) bases of ambulances and UAVs are both optimized using an MCLP with joint coverage. (d) (Status quo-optimized) bases of ambulances are located at hospitals, fire departments, and police stations, while bases of UAVs are optimized based on the smaller MCLP.

	Reward	Response Time (Min)		Re in threshold (%)		Outsourced (%)	
		General	Opioid	General	Opioid		
Both random	NN-API	1975±44	11.5±0.4	7.4±0.3	76.3±0.4	78.6±0.5	15.2±0.3
	L-ADP	1889±42	12.1±0.5	7.9±0.4	75.9±0.5	76.2±0.6	15.7±0.6
	Maxwell-ADP	1819±41	12.9±0.5	8.3±0.4	74.7±0.6	72.3±0.5	17.1±0.5
	Heuristic	1755±39	13.9±0.7	8.5±0.5	74.4±0.5	70.2±0.7	17.5±0.5
	Static	1735±38	14.1±0.6	8.7±0.4	74.0±0.4	69.5±0.6	17.5±0.4
Status quo -supp	NN-API	2132±40	9.1±0.4	5.7±0.3	76.6±0.4	78.8±0.4	13.9±0.4
	L-ADP	2090±38	9.7±0.3	6.9±0.3	75.9±0.3	76.2±0.4	14.7±0.3
	Maxwell-ADP	2002±39	9.3±0.3	7.1±0.4	75.4±0.4	75.1±0.3	14.8±0.2
	Heuristic	1943±41	9.8±0.2	7.6±0.3	74.9±0.4	74.6±0.7	14.8±0.4
	Static	1882±39	10.3±0.2	7.7±0.4	74.9±0.5	74.6±0.7	14.8±0.4
Status quo -opt	NN-API	2169±41	9.1±0.4	5.6±0.3	77.1±0.3	79.1±0.4	13.9±0.3
	L-ADP	2119±39	9.5±0.3	5.5±0.3	76.4±0.3	78.5±0.4	13.7±0.3
	Maxwell-ADP	2014±41	9.7±0.3	6.0±0.3	76.3±0.4	78.5±0.3	13.6±0.2
	Heuristic	1973±42	9.7±0.2	7.5±0.3	75.9±0.4	75.6±0.6	14.8±0.4
	Static	1916±40	10.0±0.2	7.5±0.4	75.3±0.4	75.2±0.6	14.3±0.5
Both optimized	NN-API	2192±41	9.0±0.4	5.5±0.3	77.8±0.3	79.4±0.4	13.6±0.5
	L-ADP	2156±38	8.9±0.3	5.9±0.3	77.5±0.4	78.3±0.4	13.3±0.3
	Maxwell-ADP	2043±40	9.0±0.3	6.9±0.5	76.4±0.4	77.6±0.3	13.9±0.3
	Heuristic	1995±41	9.7±0.2	7.4±0.3	75.9±0.4	76.1±0.6	13.8±0.3
	Static	1932±39	9.9±0.2	7.4±0.4	75.4±0.4	75.9±0.6	14.1±0.4

Table 7: Performance comparison (95% confidence interval) under different base locations and initial layouts at Tippecanoe County.

References

- James P Byrne, N Clay Mann, Mengtao Dai, Stephanie A Mason, Paul Karanicolas, Sandro Rizoli, and Avery B Nathens. Association between emergency medical service response time and motor vehicle crash mortality in the united states. *JAMA surgery*, 154(4):286–293, 2019.
- Maya Doe-Simkins, Alexander Y Walley, Andy Epstein, and Peter Moyer. Saved by the nose: bystander-administered intranasal naloxone hydrochloride for opioid overdose. *American journal of public health*, 99(5):788–791, 2009.
- Xiaoquan Gao, Nan Kong, and Paul M Griffin. Dynamic optimization of drone dispatch for substance overdose rescue. In *2020 Winter Simulation Conference (WSC)*, pages 830–841. IEEE, 2020.
- Trevor Hastie, Robert Tibshirani, Jerome H Friedman, and Jerome H Friedman. *The elements of statistical learning: data mining, inference, and prediction*, volume 2. Springer, 2009.
- Jeff Heaton. *Introduction to neural networks with Java*. Heaton Research, Inc., 2008.

National Association of State EMS Officials and others. National emergency medical services assessment, 2020.