Location-allocation models for casualty response planning during catastrophic health events

Caunhye, Aakil M

2014

Caunhye, A. M. (2014). Location-allocation models for casualty response planning during catastrophic health events. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/61672>

<https://doi.org/10.32657/10356/61672>

LOCATION-ALLOCATION MODELS FOR CASUALTY RESPONSE PLANNING DURING CATASTROPHIC HEALTH EVENTS

Aakil M. CAUNHYE

School of Mechanical and Aerospace Engineering

A thesis submitted to the Nanyang Technological University
in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

2014

“There is nothing noble in being superior to your fellow man; true nobility is being superior to your former self.”

Ernest Hemingway

“If we knew what it was we were doing, it would not be called research, would it?”

Albert Einstein

Abstract

Catastrophic health events are natural or man-made incidents that result in casualty numbers that overwhelm the immediate response capabilities of healthcare systems. These events overpower hospitals even when the latter trigger contingency capacity surges that are usually sufficient to deal with regular disasters. We seek to build effective location-allocation plans for the overwhelming number of casualties expected in these events by using alternative care facilities to address surge capacity issues and incorporating triage and the movement of self-evacuees for effective treatment allocation.

We first tailor and apply our framework in a deterministic location-allocation model for catastrophic radiological incidents, which are events whereby the release of radioactive material leads to significant consequences to people, the environment, and facilities. We formulate a mixed integer linear programming model to locate alternative care facilities, allocate casualties for triage, and allocate triaged casualties for treatment to minimize the total weighted casualty transportation time. The model is applied to the study of two separate radiological dispersal device incidents, both based on Department of Homeland Security's National Planning Scenario 11. We generate the optimal plan for each incident and use sensitivity analyses to draw insights on facility budgeting and triage capacity allocation at hospitals. These insights lead to some response planning rules of thumb.

With the above model as a basis, we incorporate data uncertainties, using probabilistically distributed scenarios, and the time sensitivity of information to create

a three-stage stochastic programming location-allocation model for general catastrophic health events. In the first stage, the model locates alternative care facilities prior to any scenario realization. In the second stage, it allocates casualties for triage, given initial damage scenarios involving uncertainties in casualty demands, transportation times, and numbers of self-evacuees. In the third stage, it allocates casualties for treatment, given triage scenarios and initial damage scenarios. Model solution time increases exponentially with the number of binary location variables. Given that large solution times are unacceptable in disaster response, we propose an algorithm based on Benders' decomposition to obtain good solutions fast. The algorithm uses valid inequalities to reduce feasibility cuts, flow cover inequalities to generate stronger cuts, and perturbed dual subproblems to reduce the effect of degeneracy. We implement the model and algorithm in the case study of an earthquake situation in the southernmost part of the San Andreas Fault. We compare the effectiveness of our algorithm to classical Benders' decomposition and perform sensitivity analyses on facility budgets and capacities to obtain interesting insights.

Acknowledgements

I want to extend my sincere thanks and appreciation to my advisor, Prof. Xiaofeng Nie. His continuous encouragement, guidance, and untainted trust made this work possible. He gave me the primary motivation and inspiration for this research and has guided me, with his broad knowledge and insight, from the start of this journey to the end. My thanks also goes to Prof. Mingzhe Li whose collaboration and wealth of experience have helped me improve my work on radiological incidents and my own general level of thought. Also, I owe a big thanks to my fellow friends at Nanyang Technological University and the various staff members who have made for a fun environment that is conducive to learning and doing research.

I would also like to express my deepest thanks to my mum and dad and my two brothers, Ajmal and Irshad. Though from afar, they have always been by my side, pushing, helping, and encouraging me. Their unconditional love and support have helped me grow as a person and, no less, as a researcher. Our Sunday Skype conversations made me start the school week in the right mood.

Contents

Abstract	ii
Acknowledgements	iv
List of Figures	vii
List of Tables	viii
1 Introduction	1
1.1 CHEs	1
1.2 Response Planning Issues	3
1.2.1 Surge Capacity	4
1.2.2 Effective Treatment Allocation	5
1.2.3 Self-Evacuees	6
1.3 Methodology	7
2 Literature Review	11
2.1 ACF Location	12
2.2 Treatment Planning with Triage	14
2.3 Other Relevant Models	17
2.3.1 Pandemic Outbreaks and Bioterrorism Attacks	17
2.3.2 Other Operations	21
2.4 Summary of the Literature	25
3 Deterministic Model for Radiological Incidents	26
3.1 Introduction	26
3.2 Problem Description	29
3.2.1 Demarkation of Radiation Zones	29
3.2.2 Facilities	32
3.2.3 Framework	34
3.3 Model Formulation	37
3.4 Case Studies	47
3.4.1 Los Angeles	50

3.4.2	New York City	52
3.4.3	Results and Analyses	53
3.4.3.1	Sensitivity Analysis on the Budget of Type I ACFs under Variable Total Resources	56
3.4.3.2	Sensitivity Analysis on the Budget of Type I ACFs under Fixed Total Resources	57
3.4.3.3	Sensitivity Analysis on the Allocated Triage Ca- pacity at Nonspecialized Medical Centers	58
3.4.4	Further Insights	60
4	Stochastic Programming Model	62
4.1	Introduction	63
4.2	Problem Description	65
4.2.1	Facilities and Casualties	66
4.2.2	Uncertainty Characterization	67
4.2.3	Framework	68
4.3	Model Formulation	70
4.4	Algorithm	81
4.4.1	Master Problem	84
4.4.2	Flow Cover Inequalities	93
4.4.3	Dual Subproblem	95
4.4.4	Algorithm Steps	99
4.5	Case Study	100
4.5.1	Optimal Plan	105
4.5.2	Algorithm Performance	105
4.5.3	ACF Budgeting and Capacity Sensitivity Analysis under Fixed Total Resource	109
4.5.4	ACF Budgeting and Capacity Sensitivity Analysis under Variable Total Resource	111
4.5.5	Conversion of High-Priority Treatment Capacity to Triage Capacity at Hospitals	112
4.5.6	Conversion of High-Priority Treatment Capacity to Low- Priority Treatment Capacity at Hospitals	114
4.5.7	Conversion of Triage Capacity to Low-Priority Treatment Capacity at Hospitals	115
4.5.8	Insights	118
5	Conclusions	119
5.1	Summary	119
5.2	Future Work	123

List of Figures

1.1	Number of reported natural disasters from 1900-2011 (from [1]) . . .	2
1.2	Number of people reported affected by natural disasters from 1900-2011 (from [1])	3
1.3	General modeling framework	9
3.1	Demarkation of radiation zones	30
3.2	Modified framework for radiological incidents	36
3.3	Radiation zones for Los Angeles (left) and New York City (right) .	48
3.4	Facilities in the Los Angeles case study	51
3.5	Facilities in the New York City case study	53
3.6	Optimal plans for the Los Angeles (top) and New York City (bottom) case studies	54
3.7	Sensitivity analysis on the budget of type I ACFs under variable total resources	57
3.8	Sensitivity analysis on the budget of type I ACFs under fixed total resources	58
3.9	Sensitivity analysis on the allocated triage capacity at nonspecialized medical centers	59
4.1	Scenario tree with notations	67
4.2	Example for triage satisfaction in intersecting polygons	93
4.3	Region under study	104
4.4	Illustration of the optimal plan	106
4.5	Algorithm performance	107
4.6	Sensitivity analysis on the budget and triage-to-treatment capacity ratio for fixed total resource	110
4.7	Sensitivity analysis on the budget of available ACFs for variable total resource	111
4.8	Sensitivity analysis on the conversion of high-priority treatment capacity to triage at hospitals	113
4.9	Sensitivity analysis on the conversion of high-priority treatment capacity to low-priority treatment capacity at hospitals	116
4.10	Sensitivity analysis on the conversion of triage capacity to low-priority treatment capacity at hospitals	117

List of Tables

3.1	Demand data for both case studies	49
3.2	Medical center capacities for the Los Angeles case study	51
3.3	Medical center capacities for the New York City case study	52
4.1	Variation of the optimal value and solution time with $ N $	106
4.2	Solutions over time for different configurations of our algorithm . .	108

To Ajmal, dad, Irshad, and mum

Chapter 1

Introduction

In this chapter, we define catastrophic health events (CHEs), discuss their impacts on society, and address the three major issues involved in generating proper response plans for them, namely surge capacity addition, effective treatment allocation, and the movement of self-evacuees. We discuss how this thesis uses the concept of alternative care facilities (ACFs) and triage to address these issues and then provide a general framework as a basis for tactical casualty response planning. The aim is to build effective location-allocation plans for the overwhelming number of casualties expected in CHEs.

1.1 CHEs

CHEs are natural or man-made incidents that result in numbers of ill or injured people (casualties) sufficient to overwhelm the immediate capabilities of emergency response and healthcare systems. These events overpower hospitals even when

the latter trigger contingency capacity surges that are usually sufficient to deal with regular disasters. We can think of a CHE as being an extreme form of mass casualty events. Examples of CHEs include nuclear detonations, severe widespread pandemics, and strong earthquakes. Over the years, the number of CHEs has increased and with the growing threats of terrorism, climate change, and new pandemics, there is a greater need to prepare for these events. The past decade itself has seen a drastic rise in the number of natural disasters (see Figure 1.1), many of which have had catastrophic effects, such as the Sichuan Earthquake, the Haiti Earthquake, and the Indian Ocean Tsunami, as shown by the dramatic rise in the number of affected people from disasters (see Figure 1.2).

FIGURE 1.1: Number of reported natural disasters from 1900-2011 (from [1])

In January 2010, [Toner et al. \[2\]](#) released a report, commissioned by the U.S. Department of Health and Human Services, calling CHEs the next challenge of healthcare preparedness. They detailed the need to prepare healthcare systems to respond to CHEs and provided several recommendations for the various issues involved. This thesis tackles the issue of tactical healthcare response to CHEs. More specifically, the main goal is to **develop a methodology which can be**

FIGURE 1.2: Number of people reported affected by natural disasters from 1900-2011 (from [1])

used to build effective location-allocation plans for healthcare response to CHEs. Proper location-allocation plans can help reduce mortalities and suffering, and aid in understanding the response supply chain better.

1.2 Response Planning Issues

Such plans need to address three major issues: 1) surge capacity, 2) effective treatment allocation, and 3) self-evacuee movement. Surge capacity issues are important to tackle the question: How can one prepare a healthcare system to provide for the overwhelming casualty counts? Effective treatment allocation involves determining effective allocation plans for casualty treatment. Self-evacuees constitute a significant portion of almost every disaster. Incorporating them in planning is paramount to generating effective response plans.

1.2.1 Surge Capacity

The conventional methods of adding surge capacity are facility-based, such as augmenting hospital capacities by adding beds and equipment, recalling off-duty personnel, maximizing resource utilization, and discharging patients [3]. In many cases, these methods are impractical. For example, a May 2008 survey by the U.S. House of Representative Committee on Oversight and Government Reform revealed that trauma centers in many major cities are operating at more than 100% capacity [see 4]. Trauma centers in Washington D.C., for instance, operate at 214% of their capacity, which consume any disaster surge capability. The very definition of CHEs implies that the healthcare system has to deal with cases where such facility-based surge solutions are inadequate. One way of addressing this issue is by using community-based surge solutions, meaning regional cooperation among healthcare facilities by sharing resources. This work focuses on such solutions via the use of ACFs, such as schools, parks, and stadiums, to stage casualty triage and treatment. The literature on mass casualty management and surge capacity widely advocates the use of ACFs to supplement hospital capacities during incidents which are as large in scale as CHEs [see 2, 5–8].

ACFs are usually manned by disaster medical assistance teams, military crews, and other people with basic medical training. With such personnel, they can take over triage duties, treat low-acuity and ambulatory casualties, stabilize the conditions of critically injured casualties, accommodate patients transferred from hospitals, as well as administer prophylactic treatments and vaccinations during pandemics. When emergency responders set up ACFs, they are effectively adding

new facilities and thus extra capacity to perform these activities. Moreover, they are relieving hospitals of tasks for which they are not indispensable, thereby allowing them to liberate capacity for the treatment of critically affected casualties. These casualties require the kind of specialist attention and equipment that ACFs and other facilities cannot provide.

1.2.2 Effective Treatment Allocation

Casualties are expected to have afflictions bearing different levels of seriousness. For proper response planning, it is essential to include a triage process to separate and prioritize casualties. Triage is the running of initial diagnostics on casualties so as to determine the types of treatment they need and the urgency with which they need these treatments. During triage, physicians, nurses, or other medically trained personnels assign priorities to casualties based on the seriousness of their conditions. With these priorities, it becomes easier to efficiently separate and queue casualties for treatment. Triage is essential during most disasters since resources are insufficient to treat all casualties immediately [9]. The literature on mass casualty management extensively advocates the planning of triage for effective treatment provision during mass casualty incidents [see 10–12]. The catastrophic scale of CHEs amplifies the importance of including triage in response planning. In incidents of such scale, existing medical facilities cannot handle all casualties and need additional capacity from ACFs, which is why it is essential to identify casualties whose afflictions are treatable at ACFs. In addition, without triage, we may end up with many low-acuity casualties occupying capacities

ahead of critically afflicted ones or the latter being transported long distances for treatment.

1.2.3 Self-Evacuees

In this work, self-evacuees refer to casualties who take matters in their own hands and seek treatment without the help of emergency responders. Self-evacuees can overwhelm resources and block the way for more severe casualties [13]. In most disasters, they form the most significant portion of casualties. The September 11, 2001 terrorist attack on the World Trade Center counted 85% of its victims as self-evacuees [14]. In the 2004 Madrid Bombing, the Gregorio Marañon University General Hospital, the closest hospital from the attack site, reported that around 78% of the casualties it received were self-evacuees [15]. The 2005 London Bombing produced the largest number of casualties in the UK since World War II, 86% of which were self-evacuees [16]. The problem of incorporating the movement of self-evacuees in response planning is one of the least analyzed in the literature. These casualties cause massive strain on hospitals. In some cases, they can be severely afflicted while in others, they may not be affected at all. In most of the literature, the “walking wounded” (those with minor injuries and who seek treatment on their own) are what constitute self-evacuees. However, here we also include those who are severely injured and are transported by volunteers without directions from emergency responders.

1.3 Methodology

Without a proper location-allocation plan, the benefits of ACFs, triage, and incorporating self-evacuees in planning will, in all probability, fail to materialize. For example, during the 1995 Oklahoma City bombing, emergency medical teams set up an ACF in the form of a triage center but failed to triage around half of the casualties because they initially placed the triage center too far from the attack site [17]. Consequently, casualties and volunteers overlooked the triage center and flocked to the nearest hospitals instead. This unexpected influx of casualties disrupted treatment provision at hospitals. Aside from avoiding such difficulties, an appropriate location-allocation plan, in tandem with a solid coalition among healthcare facilities, can improve coordination among the facilities, ensure that casualty loads are matched with available resources, and allocate scarce resources in an optimal way [18].

In this thesis, we use optimization modeling to build location-allocation plans. The concept of location-allocation planning, first described by Cooper [19], started with the problem of locating facilities and allocating service capabilities to them (for a review of early location-allocation models, see [20]). With time, the problem definition has evolved to cover a wide range of applications, some of which are directly related to disasters, such as disaster evacuation planning [see 21] and disaster supplies prepositioning [see 22]. In our location-allocation framework, “location” means choosing, among a set of possibilities, where to place ACFs and “allocation” refers to the decision on how many casualties to send to which facility for which process. Casualties are grouped according to geographical locations

called demand points. At each demand point, a fraction of the casualties self-evacuate, while the remaining wait for emergency responders. The total number of self-evacuees from and non self-evacuees at a demand point is called its demand. Every casualty has to go through triage, where the seriousness of his/her injury or affliction is determined, prior to being allocated for treatment based on the triage result. Self-evacuees are unaware of the triage capacities of their destination facilities when they self-evacuate. Because of that, the number of self-evacuees at a facility can exceed its triage capacity. When that happens, the excess self-evacuees must be allocated to other facilities where triage capacity is available.

Casualties who require urgent life-saving treatment are sent to hospitals (in this thesis, a hospital is any facility which can treat critically afflicted casualties) while the rest are treated at ACFs or hospitals. A system of distinct facility types providing different sets of services is called a hierarchical system in the literature. Many have used hierarchical location-allocation models for healthcare systems, especially for healthcare delivery in developing nations (for a review of these models, refer to [Rahman and Smith \[23\]](#)). Hierarchical healthcare systems are known to improve the equity in health provision and to reduce morbidity and mortality [\[23\]](#). Figure 1.3 is a flowchart showing the general framework based upon which optimization models are built.

Our models minimize the expected weighted total transportation time of casualties. Weights add systemic casualty prioritization in planning. Ideally, we would like to find an optimal value that minimizes the transportation time of low-acuity

FIGURE 1.3: General modeling framework

casualties and that of seriously afflicted casualties simultaneously. But this optimal value often does not exist. The weighted sum method is a popular easy-to-use method in the literature to find Pareto optimal solutions when the ideal optimal value cannot be reached [see 24–26]. A solution point is Pareto optimal if it is not possible to move from that point and improve at least one objective function without detriment to any other objective function [26]. The weighted sum method guarantees that the solution obtained is Pareto optimal.

Chapter 2 reviews the literature on optimization models for casualty response planning during disasters. Chapter 3 illustrates our response mechanism with a deterministic linear programming model for catastrophic radiological incidents.

In Chapter 4, we incorporate uncertainties involved in chaotic CHE environments and formulate a three-stage stochastic programming model for our general response planning framework. Chapter 5 summarizes our findings in this thesis and suggests future work directions.

Chapter 2

Literature Review

In this section, we compile a review of the literature on optimization models for disaster casualty response. The literature on disaster response planning is extensive. [Caunhye et al. \[27\]](#) provide an in-depth review of the optimization models for disaster response. However, only four of their models have specific links to casualty response. The rest involve the distribution of commodities, such as tents, food, and clothing, to people affected by disasters. In this review, we focus on casualty response. We classify the optimization models found into categories related to our theme of study. We find the literature on casualty response to be much less extensive than commodity distribution. Our literature search shows no model incorporating self-evacuees, or combining ACFs and triage for response planning. We break down the review into three categories: 1) models that consider the location of ACFs, 2) models that take into account triage results of casualties to plan for treatment, and 3) models that carry out casualty response planning without

ACFs or triage. For the third section, we find a sizeable body of literature related to pandemic and bioterrorism response.

2.1 ACF Location

The literature uses many terms interchangeably for ACFs or facilities that are conceptually close to ACFs. Most location-only models for ACFs are problems where facilities are located to cover casualty demand [see 28–30]. The only model we have found that incorporates allocation decisions with location decisions is in Yi and Özdamar [31].

Drezner [28] and Drezner et al. [29] use the term “casualty collection points” for ACFs. In their disaster concept, all other facilities have become non-operational and only casualty collection points are left to provide treatment for casualties. There is no mention of triage and transportation planning. Drezner [28] formulates models to locate casualty collection points via five means: p -median, p -center, p -maxcover, min-variance, and Lorenz curve. Drezner et al. [29] follow up the study by elaborating on a multi-objective problem. They consider five objectives for their model: p -median, p -center, two max-cover, and min-variance. The model seeks to minimize the maximum percentage deviation of these individual objectives from their best possible values. Though the concept of casualty collection points surfaces in their study, it makes no provision for a coalition among healthcare facilities. It assumes that all facilities other than casualty collection points are unusable or inaccessible and that casualties themselves choose where to go to.

In essence, the study assumes that casualties will choose the nearest casualty collection points and that the latter will be able to provide the required treatment, irrespective of the injury seriousness.

Another term that is close in meaning to ACF is emergency medical service facility, or EMS facility. Although the literature on EMS vehicle or facility location is extensive, most of the body of work focuses on day-to-day medical emergencies. For instance, EMS vehicles can be ambulances which are located to attend to emergency cases, such as cardiac arrests, or severe injuries, when needed. However, we have found two papers where the types of EMS discussed are facilities rather than vehicles and which purposely build location models for large-scale emergencies. [Jia et al. \[30\]](#) establish a general location framework for EMS facilities during large-scale emergencies. Their model locates facilities to cover demand points so as to achieve maximum efficiency in covering demand. The demand at different demand points is uncertain, with every demand point ascribed a likelihood of being struck by a large-scale emergency. The model also takes into account quality of coverage, which is defined as the number of facilities covering a demand point. [Huang et al. \[32\]](#) build a location model for EMS facilities in disaster situations. The model is close to a p -center problem with the additional assumption of possibility of facility failure to respond to demand. Furthermore, the model applies to situations where the emergency has rendered residents' closest facilities inept.

The only ACF location model accompanied with allocation planning that we have found is in [Yi and Özdamar \[31\]](#). The authors formulate a capacitated dynamic

model that routes different modes of transport through a network to deliver various types of commodities to people and move casualties of different priorities to temporary or permanent emergency sites after a disaster. We can equate these emergency sites to ACFs. The model minimizes the delay in the arrival of commodities at aid centers and in the provision of healthcare for the injured. The authors treat casualty priorities as known once the casualties are discovered. Fundamentally, this would amount to on-the-spot triage instead of in-facility triage. On-the-spot triage can suffice for small-scale disasters but will likely be rendered untenable in CHEs. The authors also consider the reallocation of service capacities among medical facilities. Initially, their model treats vehicles and casualties as integer-valued commodities. Thereafter, they use these integer solutions to construct vehicle routes and pick up/delivery instructions via a set of inequalities and an algorithm. Also, emergency sites are assumed to be able to treat any casualty no matter what their priorities are. The only distinction made between casualties of different priorities occur in the objective function where weights are added for prioritization.

2.2 Treatment Planning with Triage

Models for treatment planning mainly deal with dynamic treatment planning, with the possibility of casualty priorities changing over time. The two methods used are queueing theory [see 33–35], where triage decisions influence how casualties are queued in a system, and mathematical programming [see 36, 37], where planning for treatment is based upon triage decisions.

Gong and Batta [33], Argon et al. [34], and Kilic et al. [35] use queueing theory to prioritize casualties for treatment. Gong and Batta [33] develop a preemptive two-priority single-server queueing system to study the treatment of casualties in disaster response. Due to the preemptive service discipline, the server chooses to tackle demand of high-priority casualties until the number of low-priority casualties in the queue exceeds a threshold value. When that happens, the server shifts its total attention to low-priority casualties. In this system, the server never shares its attention between the two priorities. In addition, a major assumption is that service rates for both high- and low-priority casualties are the same.

Argon et al. [34] use queueing systems to tackle the problem of treatment planning of triaged casualties during mass casualty incidents. Their model allows for impatient casualties, implying those who leave the system if not taken care of within a time limit. The objective is to determine how to order casualties in a queue to minimize abandonments. Service time paradigms are varied so as to compare resulting policies. For example, if casualties can be ordered in such a way that the job with the shortest lifetime also has the shortest service time, the optimal policy will be to give the highest priority to the job with the shortest lifetime.

Kilic et al. [35] use a queueing system with a non-preemptive two-priority S -server configuration, and finite capacity to determine the treatment rate for each priority category following a mass casualty event. The model considers variable overall service rates that can be adjusted by increasing or decreasing the number of medical teams available. It seeks to minimize both the expected value of the square of the difference between the number of servers and the number of patients

in the system, and the cost of serving these patients over a determined time period by varying the treatment rate of every medical team. The model also incorporates the possibility of casualty priorities changing over time.

[Sacco et al. \[36\]](#) and [Wilson et al. \[37\]](#) use mathematical programming techniques to build triage systems. [Sacco et al. \[36\]](#) use a linear programming model to develop a triage system that maximizes the expected number of survivors given constraints on the timing and availability of resources. They call the system a resource-constrained triage system because of its application to disasters where resources are very limited or overwhelmed by demand. The model finds how many casualties to assign for treatment in different time periods, given initial triage scores. It takes into account time-dependent survival probabilities of casualties by considering the deterioration of their states over time. Victim deterioration estimates are based on expert opinions and cater only to casualties with penetrative injuries. Therefore, the model is applicable to severe casualties. The authors simulate and compare the performance of their model with a generic triage system called Simple Triage and Rapid Treatment (START) and find that their model mitigates the limitations of START with characteristics such as an improved sorting system for casualties and more accurate victim groupings, among others.

[Wilson et al. \[37\]](#) use a combinatorial multi-objective lexicographic model for decision support during mass casualty incidents. The model has several key problem features, including the uncertain health seriousness of casualties and the choice of appropriate hospitals for casualties. It uses estimates of task durations and travel times to schedule response operations. The model has three main objectives which

are the minimization of fatalities, suffering, and inverse efficiency. Being lexicographic, it assumes that the minimization of fatalities is infinitely more important to the emergency response decision maker than the minimization of suffering, which in turn is infinitely more important than the minimization of inverse efficiency. The way the model incorporates triage is through sorting to determine which casualties are close to being fatalities.

2.3 Other Relevant Models

There exist many models for casualty response that do not delve into either triage or ACF location. We divide these models into two groups, namely, those dealing with pandemic outbreaks or large-scale bioterrorism and other models that tackle a variety of other operations, such as ambulance positioning and search-and-rescue.

2.3.1 Pandemic Outbreaks and Bioterrorism Attacks

There is a sizeable body of work concerned with vaccine or drug distribution in cases of pandemic outbreaks or large-scale bioterrorism attacks. Most of these models incorporate a process of epidemic transmission and spread among populations. The issues discussed are drug allocation games [see 38–40], vaccine dispensing [see 41, 42], medical supplies stockpiling and distribution [see 43], medical worker assignment [see 44], and vehicle routing [see 45].

Sun et al. [38] analyze a drug allocation game among countries for prophylactic antiviral drugs to slow down the spread of an influenza pandemic. Because drug

stockpiles are presently scattered in different countries, the outbreak of an epidemic gives rise to a game in which each country must make decisions about how best to allocate its own stockpile in order to protect its population. The paper produces the first insights on how countries can use their stockpiles of drug to cover other populations. Its model captures three critical sources of uncertainty – the number of initial infections, the spread of the disease, and drug efficacy – while seeking to minimize the average number of infectives over two time periods. The authors use a two-period multivariate Reed-Frost model to represent the spread of the epidemic within and across countries at its onset. They find that when between-country infection rates are low, there exists a unique Pareto optimal solution. Furthermore, they find conditions under which centralized planning bears improvement over the decentralized one and vice-versa. [Wang et al. \[39\]](#) discuss a similar problem but for cases where countries make selfish (decentralized) decisions. The authors find that, if the starting infection level is low, countries will allocate their resources so as to bring the effective reproduction ratio below one and avoid a major outbreak. The authors also derive necessary and sufficient conditions under which decentralized allocation decisions would match centralized ones.

[Parnell et al. \[40\]](#) use a multi-objective defender-attacker-defender game model for bioterrorism risk mitigation decisions. The game is between an attacker (the terrorist) and a defender (the government), the defender being the leader. The defender makes the first decision about whether to add the Bio Watch program to a city and the amount of bio-agent vaccines to store in reserve. Based on those decisions, the attacker chooses the agent of attack and the target population size.

There is some unpredictability in the ability to acquire the agent and the effects of the agent. Finally, the defender gets to choose whether to deploy the vaccine reserve to mitigate casualties, limited by a cost budget. The model objectives concern the number of casualties and the economic impact, both of which are weighted and joined together to form a value function. The authors apply the model to a fictitious scenario and run sensitivity analyses on such parameters as the budget and the probabilities of acquiring different agents. They find combinations of probabilities of acquiring different agents that will cause the risk to change.

[Patel et al. \[41\]](#) integrate a mathematical programming model with stochastic epidemic simulations to find optimal vaccine distribution plans that minimize the number of illnesses or deaths in the population, given limited quantities of vaccines, for a possible influenza pandemic. Due to the probabilistic and non-linear nature of the epidemic process, the authors develop a genetic algorithm and a random mutation hill climbing algorithm and compare their abilities to find the optimal solution.

[Rachaniotis et al. \[42\]](#) produce a deterministic compartmental model to schedule limited resources for vaccination during a large-scale epidemic infection and minimize the total number of new infections. The model is appropriate for large populations, where random interactions can be averaged out. In addition, it incorporates the idea of increasing loss rates as more susceptibles become infected and more time and effort become necessary for epidemic containment. The model is applied to an H1N1 vaccination case in Greece and its efficiency compared to current practices and random solutions. The authors demonstrate the superiority

of the model compared to random solutions, both in terms of infection rates and completion times of vaccine distribution.

Zaric et al. [43] propose a dynamic model for cost-benefit analysis of various medical supplies stockpiling and distribution strategies following an anthrax attack. The model divides the population into compartments among which there are disease transmission rates. It uses this technique to tackle anthrax disease progression, prophylaxis and treatment, the distribution of supplies from storage facilities to local communities, inventory stockpiling, and antibiotic dispensing. The authors find from their model that, stockpiling large quantities of medical supplies is unlikely to be the most effective means of reducing mortality from an attack. They suggest that communities should instead focus on creating sufficient capacity for dispensing prophylaxis when an attack happens. Another interesting insight is that mortality rates are affected by the number of unexposed individuals seeking prophylaxis and treatment.

Lee et al. [44] combine simulation and optimization modeling to minimize the total cost of assigning workers to stations to attend to casualties from a large-scale bioterrorism or infectious disease incident. The total cost function includes costs associated with waiting time, average queue length, average utilization rate, average system time, and average casualty throughput. The decision support mechanism is applied to an anthrax drill in Dekalb county and compared to other schemes in other counties engaging in the same drill. DeKalb is found to achieve the highest throughput with labor usage at or below that of the other counties.

[Shen et al. \[45\]](#) propose a two-stage stochastic vehicle routing problem for large-scale bioterrorism emergencies. The first stage, called a planning stage, uses a chance-constrained programming technique with uncertainties in demand and travel times to route vehicles and minimize unmet demand and vehicle arrival time at demand nodes. The second stage, called an operational stage, provides delivery requirements (the load on each vehicle), as well as adjustments to first-stage route planning, maintaining the same objectives as the first stage. A tabu search heuristic is proposed to solve the first-stage vehicle routing problem. The authors use randomly generated problem instances to test the effectiveness of their heuristic and compare the effect of different recourse strategies.

2.3.2 Other Operations

This subsection compiles a variety of other operations used as part of casualty response planning for disasters. These include ambulance location [see [46–48](#)], search-and-rescue [see [49](#)], helicopter logistics [see [50](#)], and medical supply allocation to casualties [see [51, 52](#)].

[Gong and Batta \[46\]](#), [Zayas-Cabán et al. \[48\]](#), and [Jotshi et al. \[47\]](#) model ambulance location and allocation during disasters. [Gong and Batta \[46\]](#) formulate ambulance allocation and reallocation models for a disaster relief operation. The first allocation of ambulances consists of determining the correct number of ambulances to place in population clusters at the beginning of the casualty response process. Based on the growth of clusters after the disaster, ambulances are re-allocated so as to make full use of them and serve new clusters. Two iterative

models are built for location and reallocation to optimize the makespan and the weighted total flow time, respectively. These models are used in the case study of the Northridge Earthquake. When applied on several randomly generated cases, the models produce very consistent results, indicating robustness.

Jotshi et al. [47] investigate the problem of optimizing the available resources to minimize the response time for casualty pickup or delivery and to maximize the life expectancy of casualties. They input information about different entities (casualties, roads, emergency vehicles, clusters, and hospitals) and use data fusion to estimate various parameters that are then used to develop a routing methodology for emergency vehicles. Routing involves dispatching vehicles to casualty pickup locations followed by delivery to appropriate hospitals. Factors affecting routing results include waiting time at hospital emergency rooms, hospital capacity, and distance. The routing methodology is used to calculate shortest paths, taking into account possible congestion and road damage for three main dispatch decisions: to casualties, to clusters, and to hospitals.

Zayas-Cabán et al. [48] model emergency medical service vehicles as a multi-server queueing system which is used by a centralized planner to allocate those vehicles to areas struck by a catastrophic event. They also propose a clearing system model to dynamically control the added resources, which is solved via a Markov decision process. Their model minimizes the time taken (or the cost incurred) in returning the system to normalcy. Due to the large problem scale, the authors develop an algorithm called the “buddy system” which analyzes pairs of cities instead of individual ones to reduce solution time.

[Fiedrich et al. \[49\]](#) build a model to perform three operations: search-and-rescue, stabilization of infrastructures that have been damaged, and immediate rehabilitation of transportation lifelines. The objective is to minimize the total number of fatalities in the initial rescue period after strong earthquakes. The authors find the best resource allocation plan that reduces 1) delays in search-and-rescue operations, 2) lack of rescue attempts, 3) failures of stabilizing work, and 4) delayed transport. They focus on operations that, if met with failure or obstructions, can cause death counts to rise. The model is an integer programming model that tends to have large solution times even for moderately sized cases. Because of that, the authors develop heuristic procedures based on simulated annealing and tabu search and compare their performances.

[Barbarosoğlu et al. \[50\]](#) develop a two-level model to plan the use of helicopters for disaster relief. The top level makes assignments of: helicopters of various types to operation bases, pilots to helicopters, and the number of tours to be performed by each helicopter. The total cost of these assignments is minimized. The base level transports commodities to demand nodes and casualties to healthcare nodes, while considering re-fueling schedules and capacities, such that the makespan of the whole operation is minimized. To make the two models solvable as one, a coordination heuristic is devised with the aim of generating non-dominated solutions for the monolithic problem.

[Mete and Zabinsky \[51\]](#) formulate a model that brings resources to casualties instead of the other way round which is the subject of our research. There is a broad range of research done in resource allocation in the aftermath of events, but these

resources are not only medical resources but also others such as food and clothing, used to deal with affected people (see [27] for a comprehensive review). Since our focus is on healthcare response to casualties, we restrict our review to models that exclusively consider medical resources. [Mete and Zabinsky \[51\]](#) plan the allocation of medical supplies to casualties using a two-stage stochastic programming model. In the first stage, warehouses are selected and inventory levels of medical supplies allocated to them. The second stage devises allocation plans for medical supplies from warehouses to hospitals under disaster scenarios so as to minimize the total transportation duration and the penalty of unfulfilled demand. Disaster scenarios contain probabilistically distributed data and represent uncertainties in disaster situations. The authors input the optimal second-stage allocation result (from solving the deterministic equivalent of the stochastic programming model) into another optimization model to obtain a vehicle routing plan.

[Najafi et al. \[52\]](#) propose a robust optimization approach to schedule movement of disaster relief materials as well as injured people under demand and capacity uncertainties. The model can be used to route different capacitated transportation modes and has three objectives that mix humanitarian and cost issues. It minimizes the total lead time of unsatisfied demand, the waiting time of injured people, and the vehicle movement over time. The objective functions are arranged in a hierarchy, based on which the model is broken down into three smaller models and a solution methodology derived to reconcile the three models.

2.4 Summary of the Literature

Our literature review shows a general fragmentation of research on ACF location and triage. Models that locate ACFs tend to leave out triage and focus on using covering-location principles to serve casualties, without being specific about what the service would be. Similarly, models that plan treatment tend to omit the triage planning process. In CHEs, both ACFs and triage are important components and must be planned in tandem. In addition, these two components work hand in hand, since ACFs can provide triage services. Another observation from our literature review concerns the lack of models incorporating the movement of self-evacuees in planning. As we have mentioned, self-evacuees form the most significant portion of casualties in CHEs and these must be included in planning. Our aim in this thesis is to construct the first optimization models that combine ACFs, triage, and self-evacuee movement in a single tactical framework for casualty response.

Chapter 3

Deterministic Model for Radiological Incidents

In this chapter, we formulate a deterministic linear programming location-allocation model for casualty response during catastrophic radiological incidents. We start by motivating the work and its application, before describing the response framework and model formulation. We apply the formulation to response planning for two radiological dispersal device (RDD) case studies. Insights from the optimal plans and sensitivity analyses help devise some rules of thumb.

3.1 Introduction

A radiological incident is an event whereby the release of radioactive material leads to significant consequences to people, the environment, and facilities. Most definitions lump together the words “radiological” and “nuclear” under the common

term “nuclear and radiological accidents” (see [53] for example). In this thesis, we make a clear distinction between these two events and focus on catastrophic radiological incidents. Nuclear incidents involve nuclear fission and highly destructive explosions that instantly devastate people and buildings because of extreme heat and blast. Examples include nuclear detonations and explosions via improvised nuclear devices. Radiological incidents involve much smaller explosions, if any, and no nuclear fission. The main concern in such events is the dissemination of radioactive material. Sources of radioactive dissemination include attacks on nuclear power plants, nuclear reactor meltdowns caused by human error or natural disasters, and RDDs, or “dirty bombs” as they are known. Historical examples of radiological incidents are: the Three Mile Island accident in 1979, the Chernobyl disaster in 1986, and the Fukushima Daiichi nuclear disaster in 2011.

Radiological incidents have the potential to be catastrophic. Casualty estimates due to an explosive RDD, which depend on many factors, such as the type and size of the device and weather conditions, can reach from the hundreds to thousands [54]. A Cesium-137 RDD, for example, can have catastrophic effects because of the penetrating nature, high dispersibility, high solubility, and long half-life (30 years) of Cesium-137 radionuclides. In addition to casualty counts, tens to hundreds of thousands of people may require evacuation and decontamination in the aftermath of a radiological incident. Within 36 hours after the Chernobyl disaster in April 1986, 115,000 people had to be evacuated because of dangerous levels of radioactive contamination [55]. Increasing threats of terrorism and the higher

likelihood of natural disasters – such as the Tohoku earthquake and tsunami responsible for the Fukushima Daiichi disaster – due to climate change have recently raised concerns over radiological incidents. In 2007, two separate U.S. Homeland Security Presidential Directives were released, providing suggestions on measures to respond to catastrophic health events resulting from chemical, biological, radiological, and nuclear agents [see 56, 57]. Various reports have ensued, with focus on radiological incidents [see 58, 59]. Some in the research community have evoked the imperative of studying casualty response to nuclear/radiological incidents, ever since the September 11 attacks on the World Trade Center [see 60–62]. Until now, research has come out with frameworks, such as the “RTR” framework [63], the radiological dispersal device playbook [64], and the bioterrorism preparedness and response framework [65], to triage, transport, and treat casualties during radiological catastrophes. Albeit providing guidelines on casualty response, these frameworks leave many issues, such as how many casualties to transport from one facility to another, open to ad hoc decision making. This can lead to ineffective implementation. Quantitative research is necessary.

This chapter focuses on using optimization modeling, in the form of a deterministic location-allocation model, to systematically plan tactical casualty response during catastrophic radiological incidents. The model provides a location plan for ACFs and an allocation plan for casualties, while taking into account casualty triage. It addresses several important logistics challenges, some of them emphasized in a recent position paper by the Society for Medical Decision Making on health sector disaster response [66]. Note that in emergency logistics, “response” refers

to critical short-term actions [27]. This means that our model excludes long-term actions, such as medical follow-up of contaminated casualties for possible radiation-induced cancer or permanent relocation of casualties after evacuation.

3.2 Problem Description

Our model is an optimization model that makes two sets of decisions, namely, where to locate ACFs and how many casualties to allocate from one place to another. The model generates a location-allocation plan that minimizes the total weighted transportation time. We define relevant terms here and then explain the details of the model in subsequent subsections. The term “casualty” refers to anyone found within regions where radiation doses are high enough to cause harm if no treatment or decontamination is carried out, or anyone injured if the radiological incident involves a blast. Casualties in radiation zones require evacuation and at least some form of decontamination or treatment. Some casualties take matters in their own hands and look for treatment without the help of emergency responders. We refer to these casualties as self-evacuees.

3.2.1 Demarkation of Radiation Zones

The first step in producing effective response plans is to understand the impact of radiological incidents. To this end, we define three zones of concern around the radiological source, based on the radiation dose rate: the inner zone, the intermediate zone, and the outer zone. The demarkation of radiation zones as reported in

Figure 3.1 has been widely suggested in reports and papers on radiological incidents (see [63, 64, 67] for example). Defining boundaries in radiation zones helps understand the characteristics of casualties, determine the level of protection necessary for emergency responders, and identify safe zones.

FIGURE 3.1: Demarkation of radiation zones

The inner zone, with radiation rate $\geq 10 \text{ rem/hr}$, is usually where we expect to find most of the casualties with advanced stages of acute radiation syndrome (ARS). Should the radiological incident involve explosives, it is also where we would expect to find most casualties with blast injuries [63]. ARS can be diagnosed using radiation dosimetry and blood counts. According to the Centers for Disease Control and Prevention Fact Sheet on ARS [68], it takes as little as 30 rem of absorbed radiation for mild symptoms of ARS to appear. Absorbed doses between 30 rem and 600 rem usually result in hematopoietic syndrome, which includes symptoms such as nausea and vomiting related to falling blood counts. We call these symptoms “mild ARS.” Beyond 600 rem, the hematopoietic syndrome becomes more severe

and other afflictions, such as organ toxicity, cutaneous syndrome, gastrointestinal syndrome, and neurovascular syndrome, begin to appear [69]. Casualties with absorbed doses greater than 600 rem are struck with what we call “severe ARS.” These casualties require immediate treatment, usually with prolonged Cytokine therapy or stem cells/platelets transfusion [70], and follow-up, preferably at specialized medical facilities. In the intermediate and outer zones, we do not expect casualties suffering from severe ARS. With a radiation rate upper bound of 10 rem/hr, it would require more than 60 hours of exposure for the contraction of severe ARS. We can therefore safely assume that casualties in the intermediate and outer zones either suffer from mild ARS or light external contamination. Mild ARS does not require specialized treatment. Administration of antibiotics and prophylaxis, such as Prussian Blue for Cesium-137 contamination and Potassium Iodide for Iodine-125/131 radionuclides, together with a good rest, will do the trick [71]. Casualties with light external contamination, who only have light radiation contamination on their skin, need prevention rather than cure. Remedies include removal of contaminated clothes and items and a thorough soap/detergent wash. The radiation dose rate boundaries shown in Figure 3.1 act as a guide for catastrophic radiological incidents. We can use situation-dependent boundary values for other radiological incidents.

Casualties in the radiation zones described above are of two types: self-evacuees and non self-evacuees. Self-evacuees seek treatment on their own and may originate from all three zones. Those who originate from the outer or intermediate zone either have mild ARS or light external contamination, whereas those from the

inner zone may also have severe ARS or blast injuries. Non self-evacuees stay in radiation zones waiting for emergency responders to direct them through the response process. Non self-evacuees in the outer and intermediate zones have either mild ARS or light external contamination, whereas those in the inner zone may also be struck with severe ARS or blast injuries.

3.2.2 Facilities

Our response system consists of two main types of facilities: ACFs and medical centers. The main purpose of ACFs is to perform triage. We divide ACFs into two subcategories, namely, type I and type II. At type I ACFs, medically-trained military/paramilitary personnel perform such procedures as examination of vital signs (presence of fever or hypotension), dosimetry measurements, and complete blood counts to check for signs of ARS. In addition, type I ACFs are equipped with necessary provisions to decontaminate casualties who have light external contamination. Some type I ACFs already exist. These existing ACFs are any in-place facilities, such as clinics, dispensaries, and hospitals, that have the ability to perform the activities required of type I ACFs.

Type II ACFs deal with non self-evacuees in the inner zone. This is because the inner zone contains all casualties with severe ARS or blast injuries. Exclusive ACFs which operate under greater urgency will provide faster and smoother response to these casualties. Type II ACFs basically act as transition points where casualties undergo quick triage via dosimeter readings to check for signs of severe ARS before being sent for further tests and treatment. In addition, these ACFs

possess necessary supplies to administer simple pretreatment preparations, such as injury disinfection or bandaging of blast victims. We can safely assume that there are no existing type II ACFs.

After going through either type I or type II ACFs, casualties, if necessary, are transported to medical centers. Medical centers, chiefly concerned with providing treatment, are also subdivided into two classes: specialized medical centers and nonspecialized medical centers. Specialized medical centers treat severe ARS or blast injuries, whereas nonspecialized ones attend to mild ARS. This classification ensures that casualties only occupy capacity at medical centers that are indispensable for their treatment. For instance, having mild ARS casualties occupying capacity at specialized medical centers may result in them waiting long periods (being given a lower priority than those with severe ARS) for treatment or receiving treatment ahead of casualties with severe ARS. Our classification of ACFs and medical centers ensures that the response system is less stretched and is considerate of treatment urgency.

Nonspecialized medical centers need to be equipped with antibiotics, prophylaxis, and means to administer short-term Cytokine therapy to casualties having mild ARS. These medical centers may also apportion some resources to triage casualties. Specialized medical centers, on the other hand, must have a range of specialized equipment to perform activities such as examination of vital signs (fever, hypotension, and orthostasis), skin examination (erythema, blistering, onycholysis, and edema), neurologic examination (motor or sensory deficit, papilledema,

ataxia, and mental status and cognition), abdominal examination (pain or tenderness) [70], and treatment of the various syndromes and injuries (which may require surgical intervention) detected.

3.2.3 Framework

In our model, we group casualties in close geographical proximity together and represent them as points called demand points. The number of casualties that a demand point represents is called its demand. Demand includes both casualties who self-evacuate and those who do not. To begin the framework description, we elicit two simplifying assumptions: 1) All severe ARS or blast casualties are in the inner zone and 2) all severe ARS or blast injuries require specialized medical center treatment. In broad strokes, the response framework contains three major processes: locating ACFs, moving casualties to ACFs for triage, and using triage results to move casualties to medical centers for treatment.

After the radiological incident, emergency responders move to set up new ACFs. Eligible locations can be schools, parks, or stadiums found outside radiation zones. We choose up to a maximum budgeted number of ACFs (for each of type I and type II ACFs) to set up, limited by resource availability. Here, resources are military/paramilitary teams and medical supplies deployed to ACFs. While new ACFs are being set up, some casualties self-evacuate. These casualties cause massive strain on their destination facilities. They are unaware of the triage capacities of their destination facilities when they evacuate. Because of that, the number of self-evacuees at a facility may exceed its triage capacity. To keep casualty counts

under control and account for self-evacuees in our response plan, we allocate all excess self-evacuees to other type I ACFs where triage capacity is available. Excess self-evacuees are not sent to type II ACFs because these are transition points reserved for non self-evacuees inside the inner zone, who are more likely to have serious afflictions. It is important not to block capacity for these casualties. Non self-evacuees in the intermediate and outer zones are sent to type I ACFs for triage.

Triage involves categorizing casualties as either severe ARS/blast injuries, or mild ARS, or light external contamination. Based on triage results, casualties are allocated to medical centers for treatment, making sure to obey treatment capacity restrictions. Severe ARS or blast casualties are allocated to specialized medical centers and mild ARS casualties to nonspecialized medical centers. Those with light external contamination receive soap wash at their current respective ACFs. The latter casualties do not require resource-intensive treatment. In fact, ACFs can discharge most of them and allow them to self-decontaminate to save staff utilization [72]. Figure 3.2 illustrates the location-allocation framework used by our model.

FIGURE 3.2: Modified framework for radiological incidents

3.3 Model Formulation

Consider the situation where a catastrophic radiological incident results in three sets of demand points, D_ρ , where $\rho = A, B$, or C for the outer, intermediate, or inner zones, respectively. Our casualty response system consists of a set of potential locations for type I ACFs, N_I , for type II ACFs, N_{II} , and a set of existing ACFs, E , among which there is a set, Λ , of those to which self-evacuees move, which we refer to as self-evacuee destinations. In reality, self-evacuees can move to any existing medical facility, be it an existing ACF or a medical center. To make a medical center a self-evacuee destination, the planner can simply place an existing ACF at the same location as the medical center in the model and include it in the set of self-evacuee destinations. This modeling trick of using only existing ACFs as self-evacuee destinations is important to create a well-structured separation between the process of triage and that of treatment (only ACFs perform triage). Our model also has a set of specialized medical centers, H_S , and nonspecialized medical centers, H_{NS} . Facility sets are determined after the disaster. This means that they omit possible ACF locations, existing ACFs, and medical centers that have been rendered unusable by the incident.

We capture location decisions using binary variables. Y_j equals 1 if an ACF is deployed at location j and 0 otherwise. Resources being limited, the total number of ACFs, for each type, must be restricted to within a budget. The constraint $\sum_{j \in N_I} Y_j \leq V$ limits the total number of type I ACFs located in potential locations to a budget V . The constraint $\sum_{j \in N_{II}} Y_j \leq G$ sets a budget of G type II ACFs.

Let X_{ij} be the number of casualties transported from demand point i to ACF (new or existing) j . To make sure to triage every non self-evacuee in the outer and intermediate zones, we set $\sum_{j \in N_I \cup E} X_{ij} = \left(1 - \sum_{j \in \Lambda} \alpha_{ij}\right) d_i$, $\forall i \in D_\rho$, where $\rho = A$ or B , d_i is the corresponding demand at point i , and α_{ij} is the proportion of casualties at i who self-evacuate to destination j , meaning that $\left(1 - \sum_{j \in \Lambda} \alpha_{ij}\right) d_i$ is the number of non self-evacuees left at demand point i . Non self-evacuees from the inner zone are sent to type II ACFs for triage, leading to the constraints $\sum_{j \in N_{II}} X_{ij} = \left(1 - \sum_{j \in \Lambda} \alpha_{ij}\right) d_i$, $\forall i \in D_C$. Type II ACFs are temporary transition points set up from scratch and for which we assume there are no existing facilities.

Self-evacuees can overwhelm their destination facilities. A self-evacuee destination is considered overwhelmed when the number of self-evacuees it receives exceeds its allocated triage capacity. When that happens, the excess self-evacuees must be moved to other ACFs for triage. For moral reasons, we assume that self-evacuee destinations do not refuse triage (that is, transfer self-evacuees to other ACFs) when capacity is available. The total number of self-evacuees present at a self-evacuee destination j' is $\sum_{i \in D_A \cup D_B \cup D_C} \alpha_{ij'} d_i$. If that number is greater than the triage capacity $u_{j'}$, the excess self-evacuees are transferred to other type I ACFs, else the number of casualties transferred out, $\sum_{j \in N_I \cup E \setminus \{j'\}} W_{j'j}$, where $W_{j'j}$ is the number of casualties transported from j' to another ACF j , is zero, as expressed by the constraints $\sum_{j \in N_I \cup E \setminus \{j'\}} W_{j'j} = \max \left\{ 0, \sum_{i \in D_A \cup D_B \cup D_C} \alpha_{ij'} d_i - u_{j'} \right\}$, $\forall j' \in \Lambda$. Notice that only type I ACFs take care of self-evacuees.

The total number of casualties transferred to an ACF must not exceed its triage capacity. New ACFs are manned by military/paramilitary teams equipped with

their medical supplies. We assume that all these teams are homogeneous and have equal medical capabilities. This simplifies the model and makes all newly set up type I ACFs of equal triage capacities. a is the triage capacity of every new type I ACF and M is a big number representing the infinite triage capacity of every type II ACF. Remember that type II ACFs are transition points that perform non resource-intensive tasks. $\sum_{i \in D_A \cup D_B} X_{ij} + \sum_{j' \in \Lambda} W_{j'j}$ is the total number of casualties received for triage at new type I ACFs or existing type I ACFs that are not self-evacuee destinations. $\sum_{i \in D_A \cup D_B} X_{ij}$ is the number of non self-evacuees moved from the outer or intermediate zone and $\sum_{j' \in \Lambda} W_{j'j}$ is the number of excess self-evacuees received from overwhelmed self-evacuee destinations. The constraints $\sum_{i \in D_A \cup D_B} X_{ij} + \sum_{j' \in \Lambda} W_{j'j} \leq aY_j, \forall j \in N_I$ and $\sum_{i \in D_A \cup D_B} X_{ij} + \sum_{j' \in \Lambda} W_{j'j} \leq u_j, \forall j \in E \setminus \Lambda$ express the respective triage capacity limitations. If ACF j is a self-evacuee destination, $\sum_{i \in D_A \cup D_B \cup D_C} \alpha_{ij} d_i$ self-evacuees are already present. $\sum_{i \in D_A \cup D_B} X_{ij} + \sum_{j' \in \Lambda \setminus \{j\}} W_{j'j} - \sum_{j' \in N_I \cup E \setminus \{j\}} W_{jj'} \leq u_j - \sum_{i \in D_A \cup D_B \cup D_C} \alpha_{ij} d_i$ is its triage capacity constraint. The term $\sum_{j' \in N_I \cup E \setminus \{j\}} W_{jj'}$ represents the number of excess self-evacuees transferred out to other type I ACFs. For type II ACFs, the total number of casualties are all that come from demand points in the inner zone, giving $\sum_{i \in D_C} X_{ij} \leq MY_j, \forall j \in N_{II}$ as triage capacity constraints.

After triage, depending on affliction seriousness, casualties fall into the three classifications: severe ARS/blast injuries, mild ARS, and light external contamination. Casualties who are in the third category are treated and discharged at their current respective ACFs since they require minimal time and resources. Most of them can perform their own soap or detergent wash. Casualties with

mild ARS are sent to nonspecialized medical centers for treatment, whereas those with severe ARS or blast injuries are transferred to specialized medical centers. We make two simplifying assumptions in our model: 1) that the proportion of casualties in each respective triage category only depends on the radiation zone they are from and 2) that emergency responders randomly pick casualties for transportation. Random picking means that if the proportion of casualties with mild ARS at a demand point is f and x casualties are transported from this demand point to an ACF, the number of mild ARS casualties among those transported will be fx . Let f_ρ be the proportion of mild ARS among non self-evacuees in zone ρ and f_ρ^{SE} be the proportion of mild ARS among self-evacuees from ρ . The total number of mild ARS casualties at a new type I ACF j is $\sum_{i \in D_A} f_A X_{ij} + \sum_{i \in D_B} f_B X_{ij} + \sum_{j' \in \Lambda} \phi_{j'}^{SE} W_{j'j}$, where $\phi_{j'}^{SE} = \left(\sum_{i \in D_A} f_A^{SE} \alpha_{ij'} d_i + \sum_{i \in D_B} f_B^{SE} \alpha_{ij'} d_i + \sum_{i \in D_C} f_C^{SE} \alpha_{ij'} d_i \right) / \left(\sum_{i \in D_A \cup D_B \cup D_C} \alpha_{ij'} d_i \right)$, $\forall j' \in \Lambda$. $\phi_{j'}^{SE}$ essentially calculates the proportion of mild ARS among self-evacuees at destination j' . Here we expose an inherent assumption which is that the movement of self-evacuees is random with respect to affliction seriousness. This means that if, for example, f_A^{SE} of the self-evacuees from the outer zone have mild ARS and $\alpha_{ij'} d_i$ self-evacuees move from demand point $i \in D_A$ to destination j' , $f_A^{SE} \alpha_{ij'} d_i$ of these self-evacuees have mild ARS.

If Z_{pjk} is the number of casualties of seriousness p , where p is an index that takes a value of 0 for light external contamination casualties, 1 for mild ARS casualties, and 2 for severe ARS or blast casualties, transported from ACF j to medical center k , the constraints $\sum_{k \in H_{NS}} Z_{1jk} = \sum_{i \in D_A} f_A X_{ij} + \sum_{i \in D_B} f_B X_{ij} +$

$\sum_{j' \in \Lambda} \phi_{j'}^{SE} W_{j'j}, \forall j \in N_I \cup E \setminus \Lambda, \sum_{k \in H_{NS}} Z_{1jk} = \sum_{i \in D_A} f_A X_{ij} + \sum_{i \in D_B} f_B X_{ij} +$
 $\sum_{j' \in \Lambda \setminus \{j\}} \phi_{j'}^{SE} W_{j'j} - \sum_{j' \in N_I \cup E \setminus \{j\}} \phi_j^{SE} W_{jj'} + \sum_{i \in D_A} f_A^{SE} \alpha_{ij} d_i + \sum_{i \in D_B} f_B^{SE} \alpha_{ij} d_i +$
 $\sum_{i \in D_C} f_C^{SE} \alpha_{ij} d_i, \forall j \in \Lambda, \text{ and } \sum_{k \in H_{NS}} Z_{1jk} = \sum_{i \in D_C} f_C X_{ij}, \forall j \in N_{II}$ mean that
 all mild ARS casualties at every ACF should be transferred to nonspecialized medical centers. By the same token, the constraints $\sum_{k \in H_S} Z_{2jk} = \sum_{j' \in \Lambda} \theta_{j'}^{SE} W_{j'j},$
 $\forall j \in N_I \cup E \setminus \Lambda, \sum_{k \in H_S} Z_{2jk} = \sum_{j' \in \Lambda \setminus \{j\}} \theta_{j'}^{SE} W_{j'j} - \sum_{j' \in N_I \cup E \setminus \{j\}} \theta_j^{SE} W_{jj'} +$
 $\sum_{i \in D_C} g^{SE} \alpha_{ij} d_i, \forall j \in \Lambda, \text{ and } \sum_{k \in H_S} Z_{2jk} = \sum_{i \in D_C} g X_{ij}, \forall j \in N_{II}$ ensure the
 transfer of all casualties with severe ARS or blast injuries, who all (self-evacuees or non self-evacuees) originate from the inner zone, to specialized medical centers. In these constraints, g is the proportion of severe ARS or blast casualties among non self-evacuees in the inner zone, g^{SE} is the proportion of severe ARS or blast injuries among self-evacuees from the inner zone, and $\theta_{j'}^{SE},$
 where $\theta_{j'}^{SE} = \left(\sum_{i \in D_C} g^{SE} \alpha_{ij'} d_i \right) / \left(\sum_{i \in D_A \cup D_B \cup D_C} \alpha_{ij'} d_i \right), \forall j' \in \Lambda,$ is the proportion of severe ARS or blast casualties among the self-evacuees at destination $j'.$ Medical centers are also capacitated. $\sum_{j \in N_I \cup N_{II} \cup E} Z_{1jk} \leq m_k, \forall k \in H_{NS}$ and $\sum_{j \in N_I \cup N_{II} \cup E} Z_{2jk} \leq m_k, \forall k \in H_S$ guarantee that the numbers of casualties transferred to nonspecialized and specialized medical centers do not exceed treatment capacities.

The objective of the model is to minimize the weighted total transportation time of casualties. q_p is the weight associated with casualties of seriousness $p.$ Obviously, the more severe the affliction, the higher the weight, meaning that $q_2 > q_1 > q_0.$ A higher weight means that every unit of transportation time increases the objective value more, creating a model (which tries to minimize the

objective value) where higher-weight casualties have systemic prioritization over lower-weight ones. In the outer and intermediate zones, casualties either have mild ARS or light external contamination. $\sum_{i \in D_\rho} \sum_{j \in N_I \cup E} [q_0(1 - f_\rho) + q_1 f_\rho] t_{ij} X_{ij}$ is the total weighted transportation time of non self-evacuees for triage for $\rho = A$ or B , where t_{ij} is the transportation time from node i to node j . Of course, affliction severities are supposed to become known only after triage, but because of the random picking assumption, we can trace back casualty seriousness to the origin. $\sum_{i \in D_C} \sum_{j \in N_{II}} [q_0(1 - f_C - g) + q_1 f_C + q_2 g] t_{ij} X_{ij}$ is the total weighted transportation time of non self-evacuees for triage for the inner zone. $\sum_{j' \in \Lambda} \sum_{j \in N_I \cup E \setminus \{j'\}} [q_0(1 - \phi_{j'}^{SE} - \theta_{j'}^{SE}) + q_1 \phi_{j'}^{SE} + q_2 \theta_{j'}^{SE}] t_{j'j} W_{j'j}$ is the total weighted transportation time of excess self-evacuees for triage. After triage, affliction seriousness becomes known, with $\sum_{j \in N_I \cup N_{II} \cup E} \sum_{k \in H_{NS}} q_1 t_{jk} Z_{1jk}$ and $\sum_{j \in N_I \cup N_{II} \cup E} \sum_{k \in H_S} q_2 t_{jk} Z_{2jk}$ being the total weighted transportation times of casualties with mild ARS and severe ARS/blast injuries, respectively, for treatment. We do not have any component for the transportation of casualties with light external contamination for treatment because these casualties obtain treatment at the same respective ACFs which have triaged them. The objective function is the summation of the total weighted transportation times of non self-evacuees for triage, excess self-evacuees for triage, and casualties for treatment.

The list of notations is summarized as follows:

Sets

D_ρ : Demand points in zone ρ ($\rho = A$ for the outer zone, $= B$ for the intermediate zone, $= C$ for the inner zone)

E : Existing ACFs

N_I : Potential locations for type I ACFs

N_{II} : Potential locations for type II ACFs

H_{NS} : Nonspecialized medical centers

H_S : Specialized medical centers

Λ : Self-evacuee destinations

Parameters

M : Large number

V : Total number of type I ACFs to locate

G : Total number of type II ACFs to locate

q_p : Weight associated with casualties of seriousness p ($p = 0$ for light external contamination, $= 1$ for mild ARS, $= 2$ for severe ARS or blast injuries)

a : Triage capacity of new type I ACFs

u_j : Triage capacity of existing ACF j

d_i : Demand at point i

α_{ij} : Proportion of casualties at demand point i who self-evacuate to self-evacuee destination j

f_ρ : Proportion of casualties with mild ARS among non self-evacuees in zone ρ

f_ρ^{SE} : Proportion of casualties with mild ARS among self-evacuees from zone ρ

g : Proportion of casualties with severe ARS or blast injuries among non self-evacuees in the inner zone

g^{SE} : Proportion of casualties with severe ARS or blast injuries among self-evacuees from the inner zone

m_k : Treatment capacity of medical center k

t_{ij} : Time taken to transport a casualty from any node i to any node j

ϕ_j^{SE} : Proportion of mild ARS among self-evacuees at self-evacuee destination j

$$\phi_j^{SE} = \frac{\sum_{i \in D_A} f_A^{SE} \alpha_{ij} d_i + \sum_{i \in D_B} f_B^{SE} \alpha_{ij} d_i + \sum_{i \in D_C} f_C^{SE} \alpha_{ij} d_i}{\sum_{i \in D_A \cup D_B \cup D_C} \alpha_{ij} d_i} \quad \forall j \in \Lambda$$

θ_j^{SE} : Proportion of severe ARS or blast injuries among self-evacuees at self-evacuee destination j

$$\theta_j^{SE} = \frac{\sum_{i \in D_C} g^{SE} \alpha_{ij} d_i}{\sum_{i \in D_A \cup D_B \cup D_C} \alpha_{ij} d_i} \quad \forall j \in \Lambda$$

Decision Variables

$$Y_j = \begin{cases} 1 & \text{if a new ACF is set up at location } j \\ 0 & \text{otherwise} \end{cases}$$

X_{ij} : Number of casualties transported from demand point i to ACF j

$W_{j'j}$: Number of casualties transported from self-evacuee destination j' to ACF j

Z_{pjk} : Number of casualties of seriousness p transported from ACF j to medical center k .

Deterministic Model

$$\begin{aligned} \min \quad & \sum_{i \in D_A} \sum_{j \in N_I \cup E} \left[q_0 (1 - f_A) + q_1 f_A \right] t_{ij} X_{ij} + \sum_{i \in D_B} \sum_{j \in N_I \cup E} \left[q_0 (1 - f_B) + q_1 f_B \right] t_{ij} X_{ij} \\ & + \sum_{i \in D_C} \sum_{j \in N_{II}} \left[q_0 (1 - f_C - g) + q_1 f_C + q_2 g \right] t_{ij} X_{ij} + \sum_{j' \in \Lambda} \sum_{j \in N_I \cup E \setminus \{j'\}} \left[q_0 (1 - \phi_{j'}^{SE} - \theta_{j'}^{SE}) \right. \\ & \left. + q_1 \phi_{j'}^{SE} + q_2 \theta_{j'}^{SE} \right] t_{j'j} W_{j'j} + \sum_{j \in N_I \cup N_{II} \cup E} \sum_{k \in H_{NS}} q_1 t_{jk} Z_{1jk} + \sum_{j \in N_I \cup N_{II} \cup E} \sum_{k \in H_S} q_2 t_{jk} Z_{2jk} \end{aligned}$$

$$\text{s.t.} \quad \sum_{j \in N_I} Y_j \leq V \quad (3.1)$$

$$\sum_{j \in N_{II}} Y_j \leq G \quad (3.2)$$

$$\sum_{j \in N_I \cup E} X_{ij} = \left(1 - \sum_{j \in \Lambda} \alpha_{ij} \right) d_i \quad \forall i \in D_A \cup D_B \quad (3.3)$$

$$\sum_{j \in N_{II}} X_{ij} = \left(1 - \sum_{j \in \Lambda} \alpha_{ij} \right) d_i \quad \forall i \in D_C \quad (3.4)$$

$$\sum_{j \in N_I \cup E \setminus \{j'\}} W_{j'j} = \max \left\{ 0, \sum_{i \in D_A \cup D_B \cup D_C} \alpha_{ij'} d_i - u_{j'} \right\} \quad \forall j' \in \Lambda \quad (3.5)$$

$$\sum_{i \in D_A \cup D_B} X_{ij} + \sum_{j' \in \Lambda} W_{j'j} \leq aY_j \quad \forall j \in N_I \quad (3.6)$$

$$\sum_{i \in D_A \cup D_B} X_{ij} + \sum_{j' \in \Lambda} W_{j'j} \leq u_j \quad \forall j \in E \setminus \Lambda \quad (3.7)$$

$$\sum_{i \in D_A \cup D_B} X_{ij} + \sum_{j' \in \Lambda \setminus \{j\}} W_{j'j} - \sum_{j' \in N_I \cup E \setminus \{j\}} W_{jj'} \leq u_j - \sum_{i \in D_A \cup D_B \cup D_C} \alpha_{ij} d_i \quad \forall j \in \Lambda \quad (3.8)$$

$$\sum_{i \in D_C} X_{ij} \leq MY_j \quad \forall j \in N_{II} \quad (3.9)$$

$$\sum_{k \in H_{NS}} Z_{1jk} = \sum_{i \in D_A} f_A X_{ij} + \sum_{i \in D_B} f_B X_{ij} + \sum_{j' \in \Lambda} \phi_{j'}^{SE} W_{j'j} \quad \forall j \in N_I \cup E \setminus \Lambda \quad (3.10)$$

$$\begin{aligned} \sum_{k \in H_{NS}} Z_{1jk} &= \sum_{i \in D_A} f_A X_{ij} + \sum_{i \in D_B} f_B X_{ij} + \sum_{j' \in \Lambda \setminus \{j\}} \phi_{j'}^{SE} W_{j'j} - \sum_{j' \in N_I \cup E \setminus \{j\}} \phi_j^{SE} W_{jj'} \\ &+ \sum_{i \in D_A} f_A^{SE} \alpha_{ij} d_i + \sum_{i \in D_B} f_B^{SE} \alpha_{ij} d_i + \sum_{i \in D_C} f_C^{SE} \alpha_{ij} d_i \quad \forall j \in \Lambda \end{aligned} \quad (3.11)$$

$$\sum_{k \in H_{NS}} Z_{1jk} = \sum_{i \in D_C} f_C X_{ij} \quad \forall j \in N_{II} \quad (3.12)$$

$$\sum_{k \in H_S} Z_{2jk} = \sum_{j' \in \Lambda} \theta_{j'}^{SE} W_{j'j} \quad \forall j \in N_I \cup E \setminus \Lambda \quad (3.13)$$

$$\sum_{k \in H_S} Z_{2jk} = \sum_{j' \in \Lambda \setminus \{j\}} \theta_{j'}^{SE} W_{j'j} - \sum_{j' \in N_I \cup E \setminus \{j\}} \theta_j^{SE} W_{jj'} + \sum_{i \in D_C} g^{SE} \alpha_{ij} d_i \quad \forall j \in \Lambda \quad (3.14)$$

$$\sum_{k \in H_S} Z_{2jk} = \sum_{i \in D_C} g X_{ij} \quad \forall j \in N_{II} \quad (3.15)$$

$$\sum_{j \in N_I \cup N_{II} \cup E} Z_{1jk} \leq m_k \quad \forall k \in H_{NS} \quad (3.16)$$

$$\sum_{j \in N_I \cup N_{II} \cup E} Z_{2jk} \leq m_k \quad \forall k \in H_S \quad (3.17)$$

$$Y_j \in \{0, 1\} \quad \forall j \in N_I \cup N_{II} \quad (3.18)$$

$$X_{ij} \geq 0 \quad \forall i \in D_A \cup D_B, j \in N_I \cup E \text{ and } \forall i \in D_C, j \in N_{II} \quad (3.19)$$

$$W_{j'j} \geq 0 \quad \forall j' \in \Lambda, j \in N_I \cup E \setminus \{j'\} \quad (3.20)$$

$$Z_{1jk} \geq 0 \quad \forall j \in N_I \cup N_{II} \cup E, k \in H_{NS} \quad (3.21)$$

$$Z_{2jk} \geq 0 \quad \forall j \in N_I \cup N_{II} \cup E, k \in H_S. \quad (3.22)$$

Constraints (3.1) and (3.2) set the budgets for type I and type II ACFs, respectively. Constraints (3.3) ensure the triage of all non self-evacuees from the outer and intermediate zones at type I ACFs. Constraints (3.4) guarantee the triage of all non self-evacuees in the inner zone at type II ACFs. Constraints (3.5) ensure the transfer of all excess self-evacuees, if there are, from self-evacuee destinations to other type I ACFs. Constraints (3.6), (3.7), (3.8), and (3.9) are triage capacity constraints. Constraints (3.10), (3.11), and (3.12) make sure that, after triage at ACFs, all casualties with mild ARS are allocated to nonspecialized medical centers. Constraints (3.13), (3.14), and (3.15) make sure that all casualties with severe ARS or blast injuries are sent to specialized medical centers. Constraints (3.16) and (3.17) are treatment capacity constraints. Constraints (3.18) are binary constraints for location variables. Constraints (3.19), (3.20), (3.21), and (3.22) impose nonnegativity bounds on all allocation variables.

3.4 Case Studies

We implement our model in the study of two separate RDD incidents, one in Los Angeles and the other in New York City. Each incident is based on U.S. Department of Homeland Security’s National Planning Scenario 11 (NPS 11) “Radiological Attack – Radiological Dispersal Devices” [73]. In NPS 11, terrorists smuggle Cesium Chloride into the country to make explosive RDDs which they intend to trigger in various densely populated regions, thereby releasing highly dispersible and penetrative Cs-137. In our situation, we consider a single RDD source for each incident. In NPS 11, an RDD explosion is expected to cause 270 blast injuries and contaminate 20,000 people in an area of around 36 city blocks. A large RDD is expected to create an inner zone of 600 m boundary [54]. The intermediate zone serves as a buffer zone/transition area between the inner and outer zones [67]. In our case studies, we suppose that it is midway between the inner and outer zones. Using an average city block area of $150 \times 150 \text{ m}^2$ and a prevailing wind direction from the west-south-west for Los Angeles [74] and from the west for New York City [75], we come up with elliptical radiation zones shown in Figure 3.3 for both cities.

Assuming casualties are uniformly distributed throughout the radiation zones, the numbers of casualties in the inner, intermediate, and outer zones, based on the areas of the zones, are 575 (excluding blast injuries), 6,100, and 13,325, respectively. In our case studies, we suppose there are no casualties with light external contamination in the inner zone. We place a reasonable number of demand points throughout the radiation zones (see Figure 3.3), each sharing equally the number

FIGURE 3.3: Radiation zones for Los Angeles (left) and New York City (right)

of casualties in its respective radiation zone. It is to be noted that the choice of demand points is arbitrary here. The aggregation of casualties into demand points is itself a self-contained field of research, one that is not the focus of this thesis. Different aggregation schemes result in different aggregation errors. For discussions on these errors and on how to reduce them, in the location-allocation context, the reader can refer to [Goodchild \[76\]](#) and [Fotheringham et al. \[77\]](#).

In virtually all terrorist attacks, 80% or more of victims self-evacuate before the attack zone is contained [78]. Let the proportions of self-evacuees in the intermediate and outer zones be 80% and 95%, respectively. We assume that there are no self-evacuees from the inner zone and that, at each demand point in the intermediate or outer zone, self-evacuees move to the closest (by travel time) existing ACF. In NPS 11, the percentage of casualties with internal contamination in radiation zones (non self-evacuees) is 40% to 60%. All internally contaminated casualties either have mild or severe ARS. Suppose 60% and 40% of non self-evacuees have

mild ARS in the intermediate and outer zones, respectively and the percentages of self-evacuees with mild ARS from the intermediate and outer zones are 30% and 20%, respectively. Using plume modeling, the Agency for Healthcare Research and Quality estimates that, out of all non self-evacuees with internal contamination, 0.5% will have severe ARS following an RDD situation [79]. The total number of non self-evacuees with internal contamination being 1,574, we can expect 8 severe ARS cases, all in the inner zone, and 1,566 mild ARS cases. Including blast casualties, the inner zone thus has 278 casualties, 32.9% of its total casualty count, who require specialized treatment. To ensure systemic prioritization of casualties, we use weights of 1, 10, and 100 for casualties suffering from light external contamination, mild ARS, and severe ARS or blast injuries, respectively. Transportation times are obtained using the ArcGIS 9.3 Network Analyst tool. Table 3.1 summarizes the above described demand data. Note that the demand data come from the RDD incident provided by NPS 11 and are the same for both case studies.

TABLE 3.1: Demand data for both case studies

	Zones		
	Inner	Intermediate	Outer
Demand per demand point	845	1,017	1,333
% of self-evacuees	0	80	95
% of mild ARS among self-evacuees	–	30	20
% of mild ARS among non self-evacuees	67.1	60	40

Assume that the triage capacity of every newly set up type I ACF is 2,500 and that emergency responders have a budget of eight type I ACFs to locate in nearby parks. Treatment capacities for nonspecialized medical centers are proportional to their number of staff members and for specialized medical centers to their number of beds. We obtain data for personnel and bed counts from the American Hospital

Association data viewer [see 80]. In the following two subsections, we describe data that are specific to each case study.

3.4.1 Los Angeles

The vicinity of the radiological incident in Los Angeles has one clinic (Good Hope), which we treat as an existing ACF, and seven nonspecialized medical centers, which can treat mild ARS. Suppose each nonspecialized medical center allocates a capacity equal to one tenth of its treatment capacity for triage. For lack of data on Good Hope, we assume a triage capacity of 500.

Severe ARS needs specialized treatment for various syndromes. We use Ronald Reagan UCLA Medical Center and Cedars-Sinai Medical Center, which are nationally ranked in more than ten specialties and are, according to U.S. News and World Report [81], the two best hospitals in Los Angeles, for such treatment. These two medical centers are also designated trauma centers and can provide treatment for blast injuries. Because these medical centers are far from the affected region, military helicopters, with an average speed of 387 km/hr, are needed to transport casualties to them. For this reason, type II ACFs must be placed in large open areas for helicopter access. We identify four such areas, all of them school yards. A budget of one type II ACF is assumed. Table 3.2 summarizes the data set and Figure 3.4 illustrates the facilities involved in the Los Angeles case study (specialized medical facilities are omitted because they are too far). When information to calculate the capacity of a hospital is unavailable, we assume a value of 500.

3.4.2 New York City

The vicinity of the radiological incident in New York City has six nonspecialized medical centers. Similar to the Los Angeles case, we assume that each of these medical centers allocates a capacity equal to one tenth of its treatment capacity for triage. New York Presbyterian Medical Center at Columbia University and Skirball Institute of Biomolecular Medicine (also known as NYU Langone Medical Center) are, according to U.S. News and World Report [82], the two best hospitals in New York City and are used as specialized medical centers. New York Presbyterian Medical Center at Columbia University is far from the attack site and will be accessed, if needed, by helicopters. We identify three large areas, in Madison Square Garden, United Nations Headquarters, and Rockefeller University, with enough space for helicopter access and where type II ACFs can be set up. Similar to the Los Angeles case, a budget of one type II ACF is assumed. Table 3.3 summarizes the data set (HSS represents Hospital for Special Surgery) and Figure 3.5 illustrates the facilities involved in the New York City case study (New York Presbyterian Medical Center is missing because it is too far).

TABLE 3.3: Medical center capacities for the New York City case study

	Nonspecialized medical centers			
	Coler-Goldwater	HSS	Bellevue	Roosevelt
Treatment capacity	2,902	4,572	5,378	6,871
Triage capacity	290	457	537	687

	Nonspecialized medical centers	
	Mount Sinai	Sloan-Kettering
Treatment capacity	8,553	15,193
Triage capacity	855	1,519

	Specialized medical centers	
	Skirball	New York Presbyterian
Treatment capacity	161	452

FIGURE 3.5: Facilities in the New York City case study

3.4.3 Results and Analyses

In this subsection, we obtain results and insights by analyzing the optimal location-allocation plan for each case study and by running sensitivity analyses on the budget of type I ACFs and on nonspecialized medical center triage capacities. For budget sensitivity analyses, we study both the situation where the total capacity is variable, meaning that fewer or more ACFs can be set up, and the situation where the total capacity is fixed and we can choose whether to locate more but smaller ACFs or less but bigger ACFs. The optimal location-allocation plans for both case studies are shown in Figure 3.6.

FIGURE 3.6: Optimal plans for the Los Angeles (top) and New York City (bottom) case studies

We code our models in C++ and use Concert Technology to solve them with IBM-ILOG CPLEX 12.2 on a 3 GHz dual-core processor with 4 GB of RAM. Solution times are very small, with the processor solving our Los Angeles model in 0.09 seconds and our New York City model in 0.06 seconds. In each case study, there are two medical centers overwhelmed with self-evacuees (Good Samaritan and Metropolitan for Los Angeles and Roosevelt and Mount Sinai for New York City). From the optimal plans in Figure 3.6, we observe the transfer of excess self-evacuees from these overwhelmed medical centers to other ACFs or medical centers for triage. In fact, many type I ACFs are located near overwhelmed hospitals to handle, in many cases, exclusively, excess self-evacuees. In our optimal plans, the nearest type I ACF to an overwhelmed medical center is always used to triage excess self-evacuees and those found to have mild ARS are sent back to the medical center for treatment.

Non self-evacuees are allocated, in large majority, to ACFs for triage. Some are, it must be noted, moved to nonspecialized medical centers (Sunray, Shriners, and Saint Vincent for Los Angeles and Sloan-Kettering for New York City) and these casualties stay at their respective medical centers for treatment as well. In the inner zone, all casualties are assigned to type II ACFs for triage, before being sent to either nonspecialized or specialized medical centers for treatment depending on their triage results. For each of our optimal plans, we notice that the type II ACF is set up at the possible location that is nearest to the inner zone demand point. This is because the travel times of casualties with severe ARS or blast injuries are heavily weighted for prioritization and the helicopter travel times from

potential type II ACF locations to specialized medical centers are very small. These aspects result in the nearest location to the inner zone demand point offering the shortest demand point-to-ACF-to-specialized medical center travel time for the heavily prioritized severe ARS or blast casualties. Note that, in the New York City optimal plan, severe ARS or blast casualties are not sent to Skirball Institute of Biomolecular Medicine even if it seems closer to the chosen type II ACF than New York Presbyterian Medical Center at Columbia University because the latter is accessed by helicopter instead of land transportation. This gives a shorter travel time from the chosen type II ACF to New York Presbyterian Medical Center at Columbia University than to Skirball Institute of Biomolecular Medicine.

3.4.3.1 Sensitivity Analysis on the Budget of Type I ACFs under Variable Total Resources

Analyzing the available budget of type I ACFs is important for efficient planning. Figure 3.7 shows that increasing the budget improves the optimal value, but up to a limit. Beyond a budget of ten type I ACFs for Los Angeles and 8 type I ACFs for New York City, the optimal values remain constant. If the planner has more than ten military teams to deploy in the Los Angeles case, it would fare better by using only ten teams and redeploying the extra teams for other purposes, such as search-and-rescue from radiation zones and capacity improvement at medical centers. The budget limit depends on the number of additional good locations (potential type I ACF locations which, when chosen, reduce the total transportation times for triage or treatment). Los Angeles and New York City have different topographies,

leading to different numbers of good locations available. This is the reason for the difference in their budget limits even if the casualty counts are the same.

FIGURE 3.7: Sensitivity analysis on the budget of type I ACFs under variable total resources

3.4.3.2 Sensitivity Analysis on the Budget of Type I ACFs under Fixed Total Resources

The previous budget sensitivity analysis assumes that emergency planners have additional resources to add new ACFs to the arsenal already present. Here, we ask the question of whether it is better to locate more but smaller, or less but bigger ACFs, given that the total resources remains fixed. This means that if responders want to change the budgeted number by δ (can be a positive or negative number), every ACF will have a triage capacity equal to $\frac{8}{8+\delta} \times 2500$. Figure 3.8 shows the optimal value variations for both case studies when the budget is changed and the total resource availability is fixed. A bigger budget of smaller ACFs tends to improve the optimal solution up to a point. Beyond this point, the optimal value worsens. With increasing budget, more locations can be chosen, but at the expense of some casualties having to travel longer because of smaller capacities.

In both case studies, we would significantly improve response planning by setting up two (instead of eight) ACFs, each with a capacity of 10,000.

FIGURE 3.8: Sensitivity analysis on the budget of type I ACFs under fixed total resources

3.4.3.3 Sensitivity Analysis on the Allocated Triage Capacity at Non-specialized Medical Centers

One of the factors that the planner can control is the allocated triage capacity at nonspecialized medical centers. Here, we investigate how the optimal value changes when these allocated triage capacities are varied. We employ the simplifying assumption that a medical center adds/removes a unit of triage capacity by removing/adding a unit of treatment capacity. Figure 3.9 shows the optimal value variations. Because the range of possible triage capacities varies greatly from one medical center to another, we provide separate figures for clearer illustration.

Sunray and Shriners in Los Angeles fare better when they use all their resources for treatment because 1) they are far from the radiation zones, 2) they are not close to any facility that receives self-evacuees, and 3) they do not receive any self-evacuees themselves. Therefore, they do not expect to receive casualties for

FIGURE 3.9: Sensitivity analysis on the allocated triage capacity at nonspecialized medical centers

triage. These three conditions act as a rule of thumb to know when a hospital should allocate most, if not all, resources to treatment duties. For instance, should the planner consider using an additional medical center further to the left of Sunray in the response process, that medical center could safely use all of its resources for treatment. For the remaining medical centers, the general trend is an improving plan with increasing triage capacity, up to a point beyond which the plan starts to worsen. We notice that, for an overwhelmed medical center, this point is generally at a triage capacity where it can treat all its triaged casualties. If the allocated triage capacity is smaller than this, the number of excess self-evacuee transfers is

higher, whereas if it is larger, triaged self-evacuees must endure additional transfers to other medical centers for treatment.

3.4.4 Further Insights

Our optimal plans show the importance of ACFs in adding surge capacity and improving efficiency. ACFs can be placed just outside radiation zones for fast transportation of casualties out of danger. In addition, ACFs help overwhelmed facilities by offering alternative triage destinations for excess self-evacuees. We also observe the importance of triage in response planning. Because of triage, we can assign weights to prioritize casualties and allocate casualties to different destinations based on triage results, thereby ensuring better resource planning.

Based on our sensitivity analyses, we can develop four response planning rules of thumb which are: 1) If excess resources are available, limit the budget of type I ACFs to a certain value and reassign the remaining resources to other duties, such as search-and-rescue or capacity building at new or existing facilities, 2) if total resources are fixed, find the right balance between the choice of more locations and the increasing transportation times due to smaller capacities, 3) if a hospital is far from radiation zones, not close to any facility that receives self-evacuees, and is not expected to receive any self-evacuees, use most of its resources for treatment, and 4) for an overwhelmed medical center, allocate a triage capacity where it can treat all its triaged casualties.

Models such as ours can help central planners respond effectively to radiological incidents, avert deaths, reduce suffering, and understand the response supply chain better. They make decisions less ad hoc and allow planners to understand what is necessary for proper response to large-scale radiological incidents. The importance is emphasized even further in the current climate, where the rise of terrorism is raising concerns over the next radiological attack being more in the offing than ever.

Chapter 4

Stochastic Programming Model

Chapter 3 presents a deterministic model for responding to radiological incidents. This chapter addresses the additional challenge of taking into account uncertainties and the time sensitivity of information in response planning for a broader scope of CHEs than only radiological incidents. To this end, we build a three-stage stochastic programming location-allocation model, with uncertainties characterized as discrete probabilistically distributed scenarios. As model solution times are problematic, we also devise an algorithm based on Benders' decomposition to generate good solutions fast. We apply our model and algorithm in a case study of an earthquake event in the southernmost part of the San Andreas Fault. Numerous sensitivity analyses reveal interesting managerial insights.

4.1 Introduction

CHE environments are chaotic. More often than not, data is unavailable, and even when available, lacks accuracy. Damaged or congested communication lifelines complicate the flow of information to emergency responders and with little, or inaccurate information, planning and utilization of resources will be less effective. Following the 1995 bombing of the Murrah Federal Building in Oklahoma City, telephone lines were overloaded and hospital emergency radio systems broke down in 12 of 15 receiving hospitals [83]. In the aftermath of the World Trade Center attack, initial rescue efforts suffered greatly because of maimed telephone lines and overwhelming cellular traffic [84]. While proposals exist to improve post-disaster communication, such as using the internet and barcode-tagging casualties [85], they will not eliminate all data uncertainties and it remains paramount to develop planning methods that account for these uncertainties.

Time sensitivity of information is another major source of concern in disasters [86]. Usually, information becomes more accurate with time. Right after an incident, chaos reigns, people are in motion, casualties are undiscovered (in earthquakes, for instance, many lay injured under rubble), and triage information is unknown. On-the-ground information is unlikely to represent actual needs and if emergency responders were to generate plans based on such information, they may be blindsided with a totally different reality once they activate the response process. With the passage of time, the situation settles, some communication lifelines are restored, and data and information exchanges become more accurate.

One major limitation of the model we formulated in Chapter 3 is that it is deterministic and inherently assumes that data is known. Our aim in this chapter is to build a response planning model that takes into account data uncertainty and time sensitivity of information. The idea is to have plans that are robust to uncertainty and at the same time allow for flexibility when new information comes in. The use of stochastic programming is very appropriate to generate such plans [87]. In our case, the construction of such plans is done via a three-stage stochastic programming model with recourse, where uncertainty is characterized by probabilistically distributed disaster scenarios. In the first stage, the model locates ACFs, in the second stage it allocates casualties for triage, and in the third stage it allocates casualties for treatment. Allocation decisions are generated for different disaster scenarios, whilst location decisions are made irrespective of and robust to these scenarios. This means that the model generates a response plan that allows emergency responders to setup (and fix) ACFs right after the disaster and wait for more information on actual damages to choose the allocation plan for the scenario that most closely matches these damages.

As well as being stochastic, the model is mixed integer in nature. An increase in the number of location variables (possible ACF locations) results in an exponential rise in the solution time. In disaster situations where time is critical, long solution times are unacceptable. In addition, they can be detrimental to the practicality of the resulting response plan. Chances are that, while the model is being solved, events such as deteriorating casualty conditions and volunteer interference will render the response plan ineffective. One historical incident where long delays

affected response plan execution is the 1995 sarin attack of Tokyo subway. During this event, emergency responders set up a triage center for victims more than two hours after the attack. This led to deteriorating victim conditions and to some victims dying on their way to hospitals. In addition, by the time the triage center was operational, ambulances and volunteers had moved most nonambulatory victims to hospitals and ambulatory victims had already fled the scene in search of medical treatment [88]. To address the large solution time issue, we propose an algorithm, based on Benders' decomposition [89], to obtain good response plans fast. The algorithm is an iterative process where binary variables, the problematic variables, are isolated in a master model and solved separately from continuous allocation variables. Emergency responders can pick the best solution reached at the time limit they preset. The challenge of constructing the algorithm is in generating fast improving solutions.

4.2 Problem Description

Our response plan consists of the deployment of ACFs, the allocation of casualties to these ACFs for triage, and the assignment of casualties, depending on the triage results, to hospitals or ACFs for treatment. To liberate capacity at hospitals and improve the effectiveness of the plan, we also incorporate patient allocation from hospitals to ACFs. As mentioned in the previous section, a CHE leaves behind a chaotic environment, with much uncertainty associated with the information needed for response planning. The uncertainties we consider are those related to the number of casualties, the number of self-evacuees, transportation time,

and triage. The objective of the response plan is to minimize the expected total weighted transportation time of casualties.

4.2.1 Facilities and Casualties

We start explaining the framework by elaborating on the two physical features that concern us in a CHE environment, namely, casualties and facilities. A CHE results in casualties among which, some wait for emergency responders for rescue (non self-evacuees), while others take matters in their own hands and self-evacuate to seek treatment (self-evacuees). Self-evacuees and non self-evacuees can be subdivided into two classes depending on affliction seriousness: low-priority and high-priority. Those whose afflictions pose a serious threat to their lives are considered high-priority, while the rest are classed as low-priority.

Facilities fall into two main groups: ACFs and hospitals. ACFs are facilities which can perform triage, treat low-priority casualties, and stabilize the conditions of high-priority ones. Some ACFs are already present, such as still-functional family clinics and dispensaries. Others need to be set up from scratch at available places such as schools and stadiums by emergency response teams. Hospitals are facilities which have the necessary resources to treat high-priority casualties. While it is possible for them to allocate extra resources for other purposes, such as triage and low-priority treatment, their primary focus remains on high-priority treatment. A setting whereby ACFs share the workload for low-priority treatment while hospitals focus on high-priority treatment reduces strain on both casualties and hospitals by liberating capacity and reducing workloads and waiting times.

4.2.2 Uncertainty Characterization

We can best represent our data uncertainties in a scenario tree like the one in Figure 4.1. $\omega \in S$ is called an initial damage scenario and it has a probability of occurrence Pr_ω . Each initial damage scenario has a unique set of values of casualty count, proportion of self-evacuees, and transportation times. Having scenarios with probabilities of occurrence is how we characterize uncertainties in those three parameters. Every ω is associated with a set of triage scenarios I_ω . We denote the probability that a triage scenario β_ω occurs given that initial damage scenario ω has realized by $Pr_{\beta_\omega|\omega}$. $\beta_\omega \in I_\omega$ contains a unique set of values for the proportion of low-priority and high-priority afflictions among the casualties in ω , thus representing triage uncertainties. For the remainder of this chapter, we adopt the convention that ω is a random variable and ω is a realization of ω . Similarly, β_ω is a realization of the random variable β_ω .

FIGURE 4.1: Scenario tree with notations

When a CHE strikes, emergency responders set out to collect data using means such as satellite imaging, the internet, and on-the-ground surveys. With an approximate assessment of the situation, they draw up a scenario tree with initial damage scenarios and triage scenarios to cover possible data variations. The scenario tree data, together with other necessary parameters, are input in the stochastic programming model, which is then solved to obtain fixed location decisions and scenario-dependent allocation decisions. With time, as telephone lifelines are restored and the chaos and panic subside, more accurate data come in, revealing which of the initial damage scenarios and of the triage scenarios come closest to the actual situation and thus, which allocation decisions must be implemented.

4.2.3 Framework

The location-allocation framework contains, as the name suggests, location decisions and allocation decisions. Location decisions involve choosing, among a number of possible locations, where to set up ACFs. The number of ACFs that can be set up, which we call the ACF budget, is limited by resource availability. Here, resources are equipped military/paramilitary teams dispatched to man ACFs. Once ACFs in place, emergency responders decide where to allocate casualties for triage, making sure not to exceed triage capacity limits. Non self-evacuees wait for directions from emergency responders to move through the response process. To avoid the issues faced during the 1995 Oklahoma City bombing where the triage center was placed too far from the attack site, facilities used to triage non self-evacuees must be within a close enough travel time, which we call the covering

time, from them. Not all casualties wait for directions from emergency responders. Some self-evacuate and these casualties cause massive strain on their destination facilities. They are unaware of the triage capacities of these facilities and because of that, they may flock in numbers that exceed triage capacities. Facilities that have excess self-evacuees are said to be overwhelmed. For proper resource management, excess self-evacuees from overwhelmed facilities must be allocated to other facilities where triage capacity is available.

Allocation of non self-evacuees to covering-time-bound facilities and excess self-evacuees to non-overwhelmed facilities for triage depends on the number of non self-evacuees, the number of self-evacuees, and transportation times. These parameters are uncertain and their uncertainties are characterized as initial damage scenarios. As an illustration, two initial damage scenarios for an earthquake may have different transportation times due to the possibility of roads being broken, resulting in some facilities being covering-time-bound in one initial damage scenario and not the other. Triage allocation decisions must be generated for each initial damage scenario, so that when accurate data become available, emergency responders can implement the allocation plan for the scenario that is closest to reality.

After triage, casualties become classed as either low-priority or high-priority, depending on affliction seriousness. High-priority casualties are those who require hospital treatment. Low-priority casualties are treatable at both ACFs and hospitals. The proportion of low-priority and high-priority casualties is an uncertain

parameter whose possible variations are captured in triage scenarios. For each scenario, a plan is generated to allocate triaged casualties for treatment depending on triage results and respecting treatment capacity bounds. Clearly, treatment allocation depends on where casualties have been allocated for triage and is therefore conditional on the realized initial damage scenario. To improve response effectiveness, the plan also incorporates the transfer of hospital patients to ACFs to liberate hospital treatment capacity.

4.3 Model Formulation

In our model, we group casualties in close geographical proximity together and represent them as points called demand points. The number of casualties that a demand point represents is called its demand. Demand includes both non self-evacuees and self-evacuees. Consider the situation after a CHE has struck, leaving a set of demand points, D . Our casualty response system consists of a set of potential ACF locations, N , a set of existing ACFs, E , and a set of hospitals, H , all determined after the disaster. $\Lambda(\omega)$ is the set of existing ACFs to which self-evacuees move, which we refer to as self-evacuee destinations, in initial damage scenario ω . In reality, self-evacuees can move to any existing facility, be it an ACF or a hospital. To make a hospital, for example, a self-evacuee destination, the planner can simply place an existing ACF at the same location as the hospital in the model and include it in the set of self-evacuee destinations. This modeling trick of using existing ACFs as proxy facilities is useful to create a well-structured separation between the process of triage and that of treatment, with ACFs being

the sole provider of triage capabilities. The same trick is used to equip hospitals with triage capacity or low-priority treatment capacity.

In the first stage, the model makes location decisions for ACFs. The binary decision variable, Y_j , denotes whether or not an ACF is deployed at a potential location j . Resources being limited, the total number of ACFs set up, $\sum_{j \in N} Y_j$, must be within a budget, V . In this stage, the model minimizes $\mathbb{E}[Q_1(\mathbf{Y}, \boldsymbol{\omega})]$, the expected total weighted casualty transportation time resulting from recourse actions in subsequent stages, where $Q_1(\mathbf{Y}, \boldsymbol{\omega})$ is the second-stage recourse function.

The actions taken in the second stage are allocation decisions for triage for every initial damage scenario. Let $X_{ij}(\omega)$ be a decision variable for the number of non self-evacuees to allocate from demand point i to ACF j in initial damage scenario ω . If $d_i(\omega)$ is the demand at i in scenario ω and $\alpha_{ij}(\omega)$ is the proportion of casualties who self-evacuate from i to self-evacuee destination j in ω , $\left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega)\right] d_i(\omega)$ is the number of non self-evacuees left at i . All these non self-evacuees must be allocated to covering-time-bound ACFs for triage. If $\Pi_i(\omega)$ is the set of ACFs found within the (pre-specified) covering time of demand point i in initial damage scenario ω , $\sum_{j \in \Pi_i(\omega)} X_{ij}(\omega)$ is the total number of non self-evacuees allocated from i to its covering-time-bound ACFs. This number must equal $\left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega)\right] d_i(\omega)$.

Unlike non self-evacuees, self-evacuees do not stay at demand points but move to existing ACFs. These casualties can overwhelm their destination ACFs. A self-evacuee destination is considered overwhelmed when the number of self-evacuees it receives exceeds its allocated triage capacity. When that happens, the excess

self-evacuees must be allocated to other ACFs for triage. For moral reasons, we assume that self-evacuee destinations do not refuse triage (that is, transfer self-evacuees to other ACFs) when capacity is available, even though without that assumption, it may be possible to obtain better optimal solutions. The total number of self-evacuees present at a self-evacuee destination j' in initial damage scenario ω is $\sum_{i \in D} \alpha_{ij'}(\omega) d_i(\omega)$. If that number is greater than the triage capacity $u_{j'}$, the excess self-evacuees, $\sum_{i \in D} \alpha_{ij'}(\omega) d_i(\omega) - u_{j'}$, are allocated to other ACFs. Otherwise, no self-evacuees are transferred out of j' . If $W_{j'j}(\omega)$ is a decision variable representing the number of casualties to allocate from self-evacuee destination j' to another ACF j , the total number of outbound self-evacuee transfers from j' is $\sum_{j \in N \cup E} W_{j'j}(\omega)$. This number must equal $\max \left\{ \sum_{i \in D} \alpha_{ij'}(\omega) d_i(\omega) - u_{j'}, 0 \right\}$, meaning that it is equal to zero if j' is not overwhelmed and to $\sum_{i \in D} \alpha_{ij'}(\omega) d_i(\omega) - u_{j'}$ otherwise.

Allocation planning for non self-evacuees and excess self-evacuees must obey triage capacity constraints. New ACFs are manned by military/paramilitary teams equipped with their medical supplies. For simplicity, we assume that all these teams have equal medical capabilities and therefore, that all newly set up ACFs have the same triage capacity, denoted by a . A newly set up ACF j receives $\sum_{i \in D} X_{ij}(\omega)$ non self-evacuees from demand points and $\sum_{j' \in \Lambda(\omega)} W_{j'j}(\omega)$ self-evacuees from overwhelmed self-evacuee destinations. The constraint $\sum_{i \in D} X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} W_{j'j}(\omega) \leq aY_j$ ensures that the sum of casualties that ACF j receives does not exceed its triage capacity. If j is an existing ACF that is not a self-evacuee

destination, $\sum_{i \in D} X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} W_{j'j}(\omega) \leq u_j$ expresses its triage capacity constraint. However, if j is a self-evacuee destination, it contains $\sum_{i \in D} \alpha_{ij}(\omega) d_i(\omega)$ self-evacuees and if this number is overwhelming, excess self-evacuees must be transferred out. The triage capacity constraint for j is then $\sum_{i \in D} X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} W_{j'j}(\omega) - \sum_{j' \in N \cup E} W_{jj'}(\omega) \leq u_j - \sum_{i \in D} \alpha_{ij}(\omega) d_i(\omega)$. The term $\sum_{j' \in N \cup E} W_{jj'}(\omega)$ is the number of excess self-evacuees allocated to other ACFs. It takes on a positive value if j is overwhelmed and, as a result, $u_j - \sum_{i \in D} \alpha_{ij}(\omega) d_i(\omega)$ is negative.

As we have described, the second stage of our model makes allocation decisions for triage under resource limitations. The objective of the second stage is to minimize the sum of the total transportation times of non self-evacuees and excess self-evacuees, in addition to the expected transportation time from recourse actions to be taken in the third stage. Mathematically, the objective function is $\sum_{i \in D} \sum_{j \in N \cup E} t_{ij}(\omega) X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} \sum_{j \in N \cup E} t_{j'j}(\omega) W_{j'j}(\omega) + \mathbb{E} \left[Q_2(\mathbf{Y}, \mathbf{X}, \boldsymbol{\omega}, \boldsymbol{\beta}_{\boldsymbol{\omega}}) | \boldsymbol{\omega} = \omega \right]$, where $t_{ij}(\omega)$ is the time taken to transport unit casualty from point i to point j in initial damage scenario ω and $\mathbb{E} \left[Q_2(\mathbf{Y}, \mathbf{X}, \boldsymbol{\omega}, \boldsymbol{\beta}_{\boldsymbol{\omega}}) | \boldsymbol{\omega} = \omega \right]$ is the conditional expectation of the third-stage recourse function $Q_2(\mathbf{Y}, \mathbf{X}, \boldsymbol{\omega}, \boldsymbol{\beta}_{\boldsymbol{\omega}})$ over the set of triage scenarios given initial damage scenario ω . Note that $t_{ij}(\omega) \rightarrow \infty$ if $i = j$.

The third-stage actions are treatment allocation decisions. Treatment allocation decisions depend on triage scenarios. After triage, casualties become classed as either low-priority or high-priority. Triage scenarios characterize possible variations in the ratios of low-priority to high-priority casualties among the non self-evacuees

and among the self-evacuees of demand points. Because triage happens at ACFs, we need to find the ratio of high-priority to low-priority casualties at each ACF to devise an allocation plan for casualties to treatment destinations. For that, we make an assumption that casualties were picked randomly for triage allocation. This means that if the proportion of low-priority non self-evacuees at demand point i in triage scenario β_ω is $f_i(\beta_\omega)$ and $X_{ij}(\omega)$ casualties were transported from i to ACF j , the number of low-priority casualties among those transported is $f_i(\beta_\omega)X_{ij}(\omega)$. In the same vein, we assume that self-evacuees move to ACFs in a manner that is random with respect to affliction seriousness. Therefore, the proportion of low-priority self-evacuees at a destination j' , which we denote by $\phi_{j'}^{SE}(\beta_\omega)$, is $\sum_{i \in D} f_i^{SE}(\beta_\omega) \alpha_{ij'}(\omega) d_i(\omega) / \sum_{i \in D} \alpha_{ij'}(\omega) d_i(\omega)$, where $f_i^{SE}(\beta_\omega)$ is the proportion of low-priority casualties among self-evacuees from demand point i in triage scenario β_ω . With these two assumptions, we can obtain the ratios of low-priority to high-priority casualties at ACFs and make treatment allocation decisions using these information.

Let $Z_{plm}(\beta_\omega)$ be a decision variable for the number of casualties of priority $p \in P$ to allocate from point $l \in K$ to point $m \in K$ in triage scenario β_ω , where P is the set of priorities and K is the set of ACFs and hospitals. We denote $p = 0$ and $p = 1$ for low-priority and high-priority casualties, respectively. In our model, low-priority casualties are treated at ACFs. Remember that existing ACFs can be proxies for hospitals. ACFs have treatment capacity bounds. The treatment capacity for every new ACF is b and that for an existing ACF j is v_j . The number of low-priority casualties at an ACF j is $\sum_{i \in D} f_i(\beta_\omega) X_{ij}(\omega) +$

$\sum_{j' \in \Lambda(\omega)} \phi_{j'}^{SE}(\beta_\omega) W_{j'j}(\omega)$ if j is either a new ACF or an existing ACF that is not a self-evacuee destination. If that number is larger than the treatment capacity of j , excess low-priority casualties must be allocated to other ACFs for treatment. On the other hand, if the number of low-priority casualties at ACF j is less than its treatment capacity, there are resources available to treat more low-priority casualties from other ACFs. The net number of low-priority casualties at ACF j , $\sum_{i \in D} f_i(\beta_\omega) X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} \phi_{j'}^{SE}(\beta_\omega) W_{j'j}(\omega) + \sum_{l \in K} Z_{0lj}(\beta_\omega) - \sum_{j' \in N \cup E} Z_{0jj'}(\beta_\omega)$, must be less than or equal to its treatment capacity. The term $\sum_{l \in K} Z_{0lj}(\beta_\omega) - \sum_{j' \in N \cup E} Z_{0jj'}(\beta_\omega)$ is the net transfer of low-priority casualties in or out of j . Note that the term $\sum_{l \in K} Z_{0lj}(\beta_\omega)$ also allows for hospital patients to be transferred to ACF j . If j is a self-evacuee destination, the corresponding treatment capacity constraint is $\sum_{i \in D} f_i(\beta_\omega) X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} \phi_{j'}^{SE}(\beta_\omega) W_{j'j}(\omega) - \sum_{j' \in N \cup E} \phi_j^{SE}(\beta_\omega) W_{jj'}(\omega) + \sum_{l \in K} Z_{0lj}(\beta_\omega) - \sum_{j' \in N \cup E} Z_{0jj'}(\beta_\omega) \leq v_j - \sum_{i \in D} f_i^{SE}(\beta_\omega) \alpha_{ij}(\omega) d_i(\omega)$.

Whilst the treatment of low-priority casualties occurs at ACFs, that of high-priority casualties requires the specialist resources of hospitals. If an ACF j is newly set up or is an existing ACF that is not a self-evacuee destination, it contains $\sum_{i \in D} [1 - f_i(\beta_\omega)] X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} [1 - \phi_{j'}^{SE}(\beta_\omega)] W_{j'j}(\omega)$ high-priority casualties. These casualties must be allocated to hospitals for treatment, meaning that $\sum_{k \in H} Z_{1jk}(\beta_\omega)$ must equal the total number of high-priority casualties at j . If ACF j is a self-evacuee destination, it contains $\sum_{i \in D} [1 - f_i(\beta_\omega)] X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} [1 - \phi_{j'}^{SE}(\beta_\omega)] W_{j'j}(\omega) - \sum_{j' \in N \cup E} [1 - \phi_j^{SE}(\beta_\omega)] W_{jj'}(\omega) + \sum_{i \in D} [1 - f_i^{SE}(\beta_\omega)] \alpha_{ij}(\omega) d_i(\omega)$ high-priority casualties, who all require hospital treatment.

Hospital resources are limited. A hospital k has an initial treatment capacity of m_k . We call the treatment capacity “initial” because the hospital can liberate more capacity by transferring out patients. The treatment capacity constraint of hospital k is thus $\sum_{j \in N \cup E} Z_{1jk}(\beta_\omega) - \sum_{j \in N \cup E} Z_{0kj}(\beta_\omega) \leq m_k$, where $\sum_{j \in N \cup E} Z_{0kj}(\beta_\omega)$ is the capacity freed by transferring patients to ACFs. We assume a one-to-one correspondence, meaning that one less patient frees up capacity for one high-priority casualty. High-priority casualties need hospital beds for treatment. Patients who are eligible for transfer are those occupying beds and whose health statuses are safe enough that they can be moved to ACFs for treatment continuation. One less of these patients means an extra bed for a high-priority casualty. The number of patient transfers cannot exceed g_k , the number of transfer-eligible patients at hospital k .

The objective of the third stage is to minimize, $\sum_{p \in P} \sum_{l \in K} \sum_{m \in K} q_p t_{lm}(\omega) Z_{plm}(\beta_\omega)$, the total weighted time to transport: 1) casualties for treatment and 2) patients for treatment continuation. q_p is a weight associated with casualties of priority p . The more severe the affliction, the higher the weight, meaning that $q_1 > q_0$. Weights multiply transportation times and a higher weight means that every unit of transportation time increases the objective value more, leading to a model (which has a minimization objective) where higher-weight casualties have systemic prioritization over lower-weight ones. If $q_1 > q_0 > 1$, the model prioritizes treatment allocation over triage allocation.

The mathematical formulation of the three-stage stochastic programming model, with the full list of notations, is:

Sets

Δ : All nodes

S : Initial damage scenarios

I_ω : Triage scenarios associated with initial damage scenario ω

D : Demand points

N : Potential locations for new ACFs

E : Existing ACFs

H : Hospitals

K : All facilities ($N \cup E \cup H$)

$\Pi_i(\omega)$: ACFs that are within the covering time from demand point i in initial damage scenario ω

$\Lambda(\omega)$: Self-evacuee destinations in initial damage scenario ω

P : Priorities

Parameters

$d_i(\omega)$: Demand at point i in initial damage scenario ω

$\alpha_{ij}(\omega)$: Proportion of self-evacuees who move from demand point i to self-evacuee destination j in initial damage scenario ω

V : Total number of ACFs to locate

a : Triage capacity of new ACFs

u_j : Triage capacity of existing ACF j

b : Treatment capacity of new ACFs

v_j : Treatment capacity of existing ACF j

g_k : Number of transfer-eligible patients at hospital k

m_k : Capacity of hospital k

q_p : Weight associated with priority p

$f_i(\beta_\omega)$: Proportion of low-priority casualties among non self-evacuees at demand point i in triage scenario β_ω

$f_i^{SE}(\beta_\omega)$: Proportion of low-priority casualties among self-evacuees from demand point i in triage scenario β_ω

$\phi_j^{SE}(\beta_\omega)$: Proportion of low-priority casualties among self-evacuees at self-evacuee destination j in triage scenario β_ω

$$\phi_j^{SE}(\beta_\omega) = \frac{\sum_{i \in D} f_i^{SE}(\beta_\omega) \alpha_{ij}(\omega) d_i(\omega)}{\sum_{i \in D} \alpha_{ij}(\omega) d_i(\omega)} \quad \forall j \in \Lambda(\omega)$$

$t_{ij}(\omega)$: Time taken to transport unit casualty from point $i \in \Delta$ to point $j \in \Delta$ in initial damage scenario ω

Decision Variables

$$Y_j = \begin{cases} 1 & \text{if a new ACF is set up at location } j \\ 0 & \text{otherwise} \end{cases}$$

$X_{ij}(\omega)$: Number of non self-evacuees to allocate from demand point i to ACF j in initial damage scenario ω

$W_{j'j}(\omega)$: Number of self-evacuees to allocate from self-evacuee destination j' to ACF j in initial damage scenario ω

$Z_{plm}(\beta_\omega)$: Number of casualties of priority p to allocate from point $l \in K$ to point $m \in K$ in triage scenario β_ω .

Three-Stage Stochastic Programming Model

$$\begin{aligned} \min \quad & \mathbb{E}[Q_1(\mathbf{Y}, \boldsymbol{\omega})] \\ \text{s.t.} \quad & \sum_{j \in N} Y_j \leq V \end{aligned} \tag{4.1}$$

$$Y_j \in \{0, 1\} \quad \forall j \in N, \tag{4.2}$$

$$\text{where } Q_1(\mathbf{Y}, \omega) = \min \sum_{i \in D} \sum_{j \in N \cup E} t_{ij}(\omega) X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} \sum_{j \in N \cup E} t_{j'j}(\omega) W_{j'j}(\omega) \\ + \mathbb{E} \left[Q_2(\mathbf{Y}, \mathbf{X}, \omega, \beta_\omega) | \omega = \omega \right]$$

$$\text{s.t. } \sum_{j \in \Pi_i(\omega)} X_{ij}(\omega) = \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega) \right] d_i(\omega) \quad \forall i \in D \quad (4.3)$$

$$\sum_{j \in N \cup E} W_{j'j}(\omega) = \max \left\{ \sum_{i \in D} \alpha_{ij'}(\omega) d_i(\omega) - u_{j'}, 0 \right\} \quad \forall j' \in \Lambda(\omega) \quad (4.4)$$

$$\sum_{i \in D} X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} W_{j'j}(\omega) \leq aY_j \quad \forall j \in N \quad (4.5)$$

$$\sum_{i \in D} X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} W_{j'j}(\omega) \leq u_j \quad \forall j \in E \setminus \Lambda(\omega) \quad (4.6)$$

$$\sum_{i \in D} X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} W_{j'j}(\omega) - \sum_{j' \in N \cup E} W_{jj'}(\omega) \leq u_j \\ - \sum_{i \in D} \alpha_{ij}(\omega) d_i(\omega) \quad \forall j \in \Lambda(\omega) \quad (4.7)$$

$$X_{ij}(\omega) \geq 0 \quad \forall i \in D, j \in N \cup E \quad (4.8)$$

$$W_{j'j}(\omega) \geq 0 \quad \forall j' \in \Lambda(\omega), j \in N \cup E, \quad (4.9)$$

$$\text{where } Q_2(\mathbf{Y}, \mathbf{X}, \omega, \beta_\omega) = \min \sum_{p \in P} \sum_{l \in K} \sum_{m \in K} q_p t_{lm}(\omega) Z_{plm}(\beta_\omega)$$

$$\text{s.t. } \sum_{i \in D} f_i(\beta_\omega) X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} \phi_{j'}^{SE}(\beta_\omega) W_{j'j}(\omega) + \sum_{l \in K} Z_{0lj}(\beta_\omega) \\ - \sum_{j' \in N \cup E} Z_{0jj'}(\beta_\omega) \leq bY_j \quad \forall j \in N \quad (4.10)$$

$$\sum_{i \in D} f_i(\beta_\omega) X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} \phi_{j'}^{SE}(\beta_\omega) W_{j'j}(\omega) + \sum_{l \in K} Z_{0lj}(\beta_\omega) \\ - \sum_{j' \in N \cup E} Z_{0jj'}(\beta_\omega) \leq v_j \quad \forall j \in E \setminus \Lambda(\omega) \quad (4.11)$$

$$\begin{aligned}
& \sum_{i \in D} f_i(\beta_\omega) X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} \phi_{j'}^{SE}(\beta_\omega) W_{j'j}(\omega) - \sum_{j' \in N \cup E} \phi_j^{SE}(\beta_\omega) W_{jj'}(\omega) \\
& + \sum_{l \in K} Z_{0lj}(\beta_\omega) - \sum_{j' \in N \cup E} Z_{0jj'}(\beta_\omega) \leq v_j \\
& - \sum_{i \in D} f_i^{SE}(\beta_\omega) \alpha_{ij}(\omega) d_i(\omega) \quad \forall j \in \Lambda(\omega)
\end{aligned} \tag{4.12}$$

$$\begin{aligned}
& \sum_{i \in D} [1 - f_i(\beta_\omega)] X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} [1 - \phi_{j'}^{SE}(\beta_\omega)] W_{j'j}(\omega) \\
& = \sum_{k \in H} Z_{1jk}(\beta_\omega) \quad \forall j \in N \cup E \setminus \Lambda(\omega)
\end{aligned} \tag{4.13}$$

$$\begin{aligned}
& \sum_{i \in D} [1 - f_i(\beta_\omega)] X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} [1 - \phi_{j'}^{SE}(\beta_\omega)] W_{j'j}(\omega) \\
& - \sum_{j' \in N \cup E} [1 - \phi_j^{SE}(\beta_\omega)] W_{jj'}(\omega) + \sum_{i \in D} [1 - f_i^{SE}(\beta_\omega)] \alpha_{ij}(\omega) d_i(\omega) \\
& = \sum_{k \in H} Z_{1jk}(\beta_\omega) \quad \forall j \in \Lambda(\omega)
\end{aligned} \tag{4.14}$$

$$\sum_{j \in N \cup E} Z_{0kj}(\beta_\omega) \leq g_k \quad \forall k \in H \tag{4.15}$$

$$\sum_{j \in N \cup E} Z_{1jk}(\beta_\omega) - \sum_{j \in N \cup E} Z_{0kj}(\beta_\omega) \leq m_k \quad \forall k \in H \tag{4.16}$$

$$Z_{plm}(\beta_\omega) \geq 0 \quad \forall p \in P, l \in K, m \in K. \tag{4.17}$$

Constraint (4.1) limits the number of new ACFs that can be deployed. Constraints (4.2) set location decision variables to binary. Constraints (4.3) ensure the allocation of all non self-evacuees to covering-time-bound ACFs for triage. Constraints (4.4) guarantee the transfer of all excess self-evacuees from self-evacuee destinations to other ACFs. (4.5), (4.6), and (4.7) are triage capacity constraints. Constraints (4.8) and (4.9) set the non-negativity of second-stage decision variables, which are allowed continuous values instead of integer ones for ease of solution.

Constraints (4.10), (4.11), and (4.12) are low-priority treatment capacity constraints. Constraints (4.13) and (4.14) ensure that all high-priority casualties are treated at hospitals. Constraints (4.15) guarantee that, at every hospital, we cannot transfer out more than the available number of low-priority patients to ACFs. Constraints (4.16) are hospital treatment capacity constraints after patient out-transfers. Constraints (4.17) set the non-negativity of third-stage decision variables.

4.4 Algorithm

The literature on stochastic programming models widely uses methods such as Benders' decomposition [see 89] and the integer L-shaped method [see 90], which take advantage of the dual block-angularity of such models, to improve solution times [91]. In this work, we use a Benders' decomposition approach, which we improve with several structural properties of our model, to generate good solutions fast.

The problem that the algorithm has to solve is the deterministic equivalent of our stochastic programming model. Its formulation is as follows:

$$\begin{aligned}
\max \quad & - \sum_{\omega \in S} \sum_{i \in D} \sum_{j \in N \cup E} Pr_{\omega} t_{ij}(\omega) X_{ij}(\omega) - \sum_{\omega \in S} \sum_{j' \in \Lambda(\omega)} \sum_{j \in N \cup E} Pr_{\omega} t_{j'j}(\omega) W_{j'j}(\omega) \\
& - \sum_{\omega \in S} \sum_{\beta_{\omega} \in I_{\omega}} \sum_{p \in P} \sum_{l \in K} \sum_{m \in K} Pr_{\beta_{\omega}|\omega} Pr_{\omega} q_p t_{lm}(\omega) Z_{plm}(\beta_{\omega}) \\
\text{s.t.} \quad & \sum_{j \in \Pi_i(\omega)} X_{ij}(\omega) = \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega) \right] d_i(\omega) \quad \forall i \in D, \omega \in S
\end{aligned} \tag{4.18}$$

$$\sum_{j \in N \cup E} W_{j'j}(\omega) = \max \left\{ \sum_{i \in D} \alpha_{ij'}(\omega) d_i(\omega) - u_{j'}, 0 \right\} \quad \forall j' \in \Lambda(\omega), \omega \in S \quad (4.19)$$

$$\sum_{i \in D} X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} W_{j'j}(\omega) \leq aY_j \quad \forall j \in N, \omega \in S \quad (4.20)$$

$$\sum_{i \in D} X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} W_{j'j}(\omega) \leq u_j \quad \forall j \in E \setminus \Lambda(\omega), \omega \in S \quad (4.21)$$

$$\begin{aligned} \sum_{i \in D} X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} W_{j'j}(\omega) - \sum_{j' \in N \cup E} W_{jj'}(\omega) &\leq u_j \\ - \sum_{i \in D} \alpha_{ij}(\omega) d_i(\omega) &\quad \forall j \in \Lambda(\omega), \omega \in S \end{aligned} \quad (4.22)$$

$$\begin{aligned} \sum_{i \in D} f_i(\beta_\omega) X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} \phi_{j'}^{SE}(\beta_\omega) W_{j'j}(\omega) + \sum_{l \in K} Z_{0lj}(\beta_\omega) \\ - \sum_{j' \in N \cup E} Z_{0jj'}(\beta_\omega) &\leq bY_j \quad \forall j \in N, \omega \in S, \beta_\omega \in I_\omega \end{aligned} \quad (4.23)$$

$$\begin{aligned} \sum_{i \in D} f_i(\beta_\omega) X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} \phi_{j'}^{SE}(\beta_\omega) W_{j'j}(\omega) + \sum_{l \in K} Z_{0lj}(\beta_\omega) \\ - \sum_{j' \in N \cup E} Z_{0jj'}(\beta_\omega) &\leq v_j \quad \forall j \in E \setminus \Lambda(\omega), \omega \in S, \beta_\omega \in I_\omega \end{aligned} \quad (4.24)$$

$$\begin{aligned} \sum_{i \in D} f_i(\beta_\omega) X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} \phi_{j'}^{SE}(\beta_\omega) W_{j'j}(\omega) - \sum_{j' \in N \cup E} \phi_j^{SE}(\beta_\omega) W_{jj'}(\omega) \\ + \sum_{l \in K} Z_{0lj}(\beta_\omega) - \sum_{j' \in N \cup E} Z_{0jj'}(\beta_\omega) &\leq v_j \\ - \sum_{i \in D} f_i^{SE}(\beta_\omega) \alpha_{ij}(\omega) d_i(\omega) &\quad \forall j \in \Lambda(\omega), \omega \in S, \beta_\omega \in I_\omega \end{aligned} \quad (4.25)$$

$$\begin{aligned} \sum_{i \in D} [1 - f_i(\beta_\omega)] X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} [1 - \phi_{j'}^{SE}(\beta_\omega)] W_{j'j}(\omega) \\ - \sum_{k \in H} Z_{1jk}(\beta_\omega) &= 0 \quad \forall j \in N \cup E \setminus \Lambda(\omega), \omega \in S, \beta_\omega \in I_\omega \end{aligned} \quad (4.26)$$

$$\sum_{i \in D} [1 - f_i(\beta_\omega)] X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} [1 - \phi_{j'}^{SE}(\beta_\omega)] W_{j'j}(\omega)$$

$$\begin{aligned}
& - \sum_{j' \in N \cup E} \left[1 - \phi_j^{SE}(\beta_\omega) \right] W_{jj'}(\omega) - \sum_{k \in H} Z_{1jk}(\beta_\omega) \\
& + \sum_{i \in D} \left[1 - f_i^{SE}(\beta_\omega) \right] \alpha_{ij}(\omega) d_i(\omega) = 0 \quad \forall j \in \Lambda(\omega), \omega \in S, \beta_\omega \in I_\omega \quad (4.27)
\end{aligned}$$

$$\sum_{j \in N \cup E} Z_{0kj}(\beta_\omega) \leq g_k \quad \forall k \in H, \omega \in S, \beta_\omega \in I_\omega \quad (4.28)$$

$$\sum_{j \in N \cup E} Z_{1jk}(\beta_\omega) - \sum_{j \in N \cup E} Z_{0kj}(\beta_\omega) \leq m_k \quad \forall k \in H, \omega \in S, \beta_\omega \in I_\omega \quad (4.29)$$

$$X_{ij}(\omega) \geq 0 \quad \forall i \in D, j \in N \cup E, \omega \in S \quad (4.30)$$

$$W_{jj'}(\omega) \geq 0 \quad \forall j' \in \Lambda(\omega), j \in N \cup E, \omega \in S \quad (4.31)$$

$$Z_{plm}(\beta_\omega) \geq 0 \quad \forall p \in P, l \in K, m \in K, \omega \in S, \beta_\omega \in I_\omega. \quad (4.32)$$

Benders' decomposition works by separating the deterministic equivalent into a master problem and a dual subproblem. The master problem contains only binary location variables, which are the complicating variables. The dual subproblem is the dual of the deterministic equivalent wherein the location variables have been turned into parameters. The algorithm starts by solving the master problem to obtain a location solution. The solution is fed into the dual subproblem as a set of parameters, after which the dual subproblem is solved. If the dual subproblem is unbounded, which happens when the location solution is infeasible for the deterministic equivalent, the location solution is input into a modified dual subproblem to generate extreme rays, which are then used to construct feasibility cuts. These feasibility cuts are added to the master problem. If the dual subproblem is feasible, its solution is used to construct optimality cuts, which are added to the master problem. The whole process of solving and cutting the master problem continues until the master solution converges to that of the dual subproblem. [Benders \[89\]](#)

proves that convergence occurs at the optimal solution of the original problem.

Three major issues in Benders' decomposition impede its ability to generate good solutions fast. The first issue is that feasibility cuts are weak and require the solution of an additional problem. As a remedy, we modify the master problem so as to reduce the number of possible location solutions that are infeasible for the deterministic equivalent. The modification consists of adding valid inequalities that ensure that new ACFs located provide enough capacity to triage all casualties and treat all low-priority ones when all existing ACFs are maximally utilized. The second issue is that optimality cuts can also be weak. To improve their strength, we add flow cover inequalities to the deterministic equivalent to tighten the polyhedron of its linear relaxation. Flow cover inequalities are especially effective for CHE situations where most existing ACFs are overwhelmed with casualties. [Magnanti and Wong \[92\]](#) prove that tightening the polyhedron of the linear relaxation yields stronger cuts. The third issue is degeneracy, meaning that there may be multiple possible cuts for a location solution. To tackle this issue, we perturb the dual subproblem so that it generates maximal nondominated Benders' cuts [see [93](#)], which are stronger cuts, in degenerate cases.

4.4.1 Master Problem

The unmodified master problem is $\max \left\{ O : \sum_{j \in N} Y_j \leq V, Y_j \in \{0, 1\}, \forall j \in N \right\}$.

This problem, aside from the budget constraint, has no imposed restrictions on choosing new ACFs. By adding appropriate valid inequalities, it is possible to

tighten the feasible region of the master problem and reduce the number of possible location solutions that are infeasible for the deterministic equivalent. Since the dual subproblem is the dual of the deterministic equivalent with location variables set as parameters, adding these valid inequalities ensures fewer possible location solutions that make the dual subproblem unbounded and therefore, fewer feasibility cuts.

Proposition 4.1. *The following inequalities:*

$$\begin{aligned} \sum_{j \in \Pi_i(\omega) \cap N} aY_j &\geq \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega)\right] d_i(\omega) - \sum_{j \in \Pi_i(\omega) \cap E \setminus \Lambda(\omega)} u_j \\ &- \sum_{j \in \Pi_i(\omega) \cap \Lambda(\omega)} \max \left\{ u_j - \sum_{i' \in D} \alpha_{i'j}(\omega) d_{i'}(\omega), 0 \right\} \quad \forall i \in D, \omega \in S, \end{aligned} \quad (4.33)$$

$$\begin{aligned} \sum_{j \in \Theta_i(\omega) \cap N} aY_j &\geq \sum_{i' \in \Omega_i(\omega) \cup \{i\}} \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{i'j}(\omega)\right] d_{i'}(\omega) - \sum_{j \in \Theta_i(\omega) \cap E \setminus \Lambda(\omega)} u_j \\ &- \sum_{j \in \Theta_i(\omega) \cap \Lambda(\omega)} \max \left\{ u_j - \sum_{i' \in D} \alpha_{i'j}(\omega) d_{i'}(\omega), 0 \right\} \quad \forall i \in \{\bar{i} \in D : \Omega_{\bar{i}}(\omega) \neq \emptyset\}, \\ \omega &\in S, \end{aligned} \quad (4.34)$$

$$\sum_{j \in N} aY_j \geq \sum_{i \in D} d_i(\omega) - \sum_{j \in E} u_j \quad \forall \omega \in S, \quad (4.35)$$

and

$$\begin{aligned} \sum_{j \in N} bY_j &\geq \sum_{i \in D} f_i(\beta_\omega) \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega)\right] d_i(\omega) + \sum_{i \in D} \sum_{j \in \Lambda(\omega)} f_i^{SE}(\beta_\omega) \alpha_{ij}(\omega) d_i(\omega) \\ &+ \max \left\{ \sum_{i \in D} \left[1 - f_i(\beta_\omega)\right] \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega)\right] d_i(\omega) + \sum_{i \in D} \sum_{j \in \Lambda(\omega)} \left[1 - f_i^{SE}(\beta_\omega)\right] \alpha_{ij}(\omega) d_i(\omega) \right. \\ &\left. - \sum_{k \in H} m_k, 0 \right\} - \sum_{j \in E} v_j \quad \forall \omega \in S, \beta_\omega \in I_\omega, \end{aligned} \quad (4.36)$$

where $\Theta_i(\omega) = \bigcup_{i' \in \Omega_i(\omega) \cup \{i\}} \Pi_{i'}(\omega)$, $\Omega_i(\omega) = \{i' \in D \setminus \{i\} : \Pi_{i'}(\omega) \cap \Pi_i(\omega) \neq \emptyset \wedge \Pi_{i'}(\omega) \cap \Pi_i(\omega) \not\subseteq \Gamma(\omega)\}$, and $\Gamma(\omega) = \{j \in \Lambda(\omega) : \sum_{i \in D} \alpha_{ij}(\omega) d_i(\omega) \geq u_j\}$, are valid inequalities for our deterministic equivalent.

Proof. The demand satisfaction constraints in the deterministic equivalent are

$$\sum_{j \in \Pi_i(\omega)} X_{ij}(\omega) = \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega)\right] d_i(\omega) \quad \forall i \in D, \omega \in S.$$

These can be re-written as

$$\sum_{j \in \Pi_i(\omega) \cap N} X_{ij}(\omega) + \sum_{j \in \Pi_i(\omega) \cap E \setminus \Lambda(\omega)} X_{ij}(\omega) + \sum_{j \in \Pi_i(\omega) \cap \Lambda(\omega)} X_{ij}(\omega) = \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega)\right] d_i(\omega)$$

$$\forall i \in D, \omega \in S.$$

Because the number of non self-evacuees transferred to a facility for triage cannot exceed its available triage capacity,

$$X_{ij}(\omega) \leq aY_j \quad \forall i \in D, j \in \Pi_i(\omega) \cap N, \omega \in S,$$

$$X_{ij}(\omega) \leq u_j \quad \forall i \in D, j \in \Pi_i(\omega) \cap E \setminus \Lambda(\omega), \omega \in S, \text{ and}$$

$$X_{ij}(\omega) \leq \max \left\{ u_j - \sum_{i' \in D} \alpha_{i'j}(\omega) d_{i'}(\omega), 0 \right\} \quad \forall i \in D, j \in \Pi_i(\omega) \cap \Lambda(\omega), \omega \in S.$$

This implies that

$$\begin{aligned} & \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega)\right] d_i(\omega) - \sum_{j \in \Pi_i(\omega) \cap E \setminus \Lambda(\omega)} u_j - \sum_{j \in \Pi_i(\omega) \cap \Lambda(\omega)} \max \left\{ u_j \right. \\ & \left. - \sum_{i' \in D} \alpha_{i'j}(\omega) d_{i'}(\omega), 0 \right\} \leq \sum_{j \in \Pi_i(\omega) \cap N} X_{ij}(\omega) \leq \sum_{j \in \Pi_i(\omega) \cap N} aY_j \quad \forall i \in D, \omega \in S, \end{aligned}$$

proving that inequalities (4.33) are valid inequalities for our deterministic equivalent.

The set $\Omega_i(\omega) = \{i' \in D \setminus \{i\} : \Pi_{i'}(\omega) \cap \Pi_i(\omega) \neq \emptyset \wedge \Pi_{i'}(\omega) \cap \Pi_i(\omega) \not\subseteq \Gamma(\omega)\}$ is the set of demand points whose covering-time polygons intersect with that of demand point i . We say that two polygons intersect when they overlap and there is at least one existing ACF (with available capacity) or potential new ACF location in the overlapping region. This means that the overlapping region of the two polygons must satisfy the two conditions of 1) not being empty and 2)

not containing only overwhelmed self-evacuee destinations (which the propositional logic $\Pi_{i'}(\omega) \cap \Pi_i(\omega) \neq \emptyset \wedge \Pi_{i'}(\omega) \cap \Pi_i(\omega) \not\subseteq \Gamma(\omega)$ represents). The set $\Theta_i(\omega) = \bigcup_{i' \in \Omega_i(\omega) \cup \{i\}} \Pi_{i'}(\omega)$ includes all ACFs present in the covering-time polygon of demand point i or in any covering-time polygon intersecting with that of i .

From demand satisfaction constraints

$$\sum_{j \in \Pi_i(\omega)} X_{ij}(\omega) = \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega)\right] d_i(\omega) \quad \forall i \in D, \omega \in S,$$

we can infer that

$$\begin{aligned} \sum_{i' \in \Omega_i(\omega) \cup \{i\}} \sum_{j \in \Theta_i(\omega)} X_{i'j}(\omega) &= \sum_{i' \in \Omega_i(\omega) \cup \{i\}} \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{i'j}(\omega)\right] d_{i'}(\omega) \\ \forall i \in \{\bar{i} \in D : \Omega_{\bar{i}}(\omega) \neq \emptyset\}, \omega \in S. \end{aligned}$$

This means that

$$\begin{aligned} &\sum_{i' \in \Omega_i(\omega) \cup \{i\}} \sum_{j \in \Theta_i(\omega) \cap N} X_{i'j}(\omega) + \sum_{i' \in \Omega_i(\omega) \cup \{i\}} \sum_{j \in \Theta_i(\omega) \cap E \setminus \Lambda(\omega)} X_{i'j}(\omega) + \sum_{i' \in \Omega_i(\omega) \cup \{i\}} \sum_{j \in \Theta_i(\omega) \cap \Lambda(\omega)} X_{i'j}(\omega) \\ &= \sum_{i' \in \Omega_i(\omega) \cup \{i\}} \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{i'j}(\omega)\right] d_{i'}(\omega) \quad \forall i \in \{\bar{i} \in D : \Omega_{\bar{i}}(\omega) \neq \emptyset\}, \omega \in S. \end{aligned}$$

Obviously, the total number of non self-evacuees allocated for triage from the set of demand points $\Omega_i(\omega) \cup \{i\}$ to new ACFs must not be more than the total capacity afforded by new ACFs inside $\Theta_i(\omega)$. Therefore,

$$\sum_{i' \in \Omega_i(\omega) \cup \{i\}} \sum_{j \in \Theta_i(\omega) \cap N} X_{i'j}(\omega) \leq \sum_{j \in \Theta_i(\omega) \cap N} aY_j \quad \forall i \in \{\bar{i} \in D : \Omega_{\bar{i}}(\omega) \neq \emptyset\}, \omega \in S.$$

The same capacity restriction applies to existing ACFs, meaning that the inequalities

$$\begin{aligned} \sum_{i' \in \Omega_i(\omega) \cup \{i\}} \sum_{j \in \Theta_i(\omega) \cap E \setminus \Lambda(\omega)} X_{i'j}(\omega) &\leq \sum_{j \in \Theta_i(\omega) \cap E \setminus \Lambda(\omega)} u_j \quad \forall i \in \{\bar{i} \in D : \Omega_{\bar{i}}(\omega) \neq \emptyset\}, \omega \in S \text{ and} \\ \sum_{i' \in \Omega_i(\omega) \cup \{i\}} \sum_{j \in \Theta_i(\omega) \cap \Lambda(\omega)} X_{i'j}(\omega) &\leq \sum_{j \in \Theta_i(\omega) \cap \Lambda(\omega)} \max \left\{ u_j - \sum_{i' \in D} \alpha_{i'j}(\omega) d_{i'}(\omega), 0 \right\} \end{aligned}$$

$$\forall i \in \{\bar{i} \in D : \Omega_{\bar{i}}(\omega) \neq \emptyset\}, \omega \in S$$

are valid for our deterministic equivalent. Therefore,

$$\begin{aligned} & \sum_{i' \in \Omega_i(\omega) \cup \{i\}} \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{i'j}(\omega) \right] d_{i'}(\omega) - \sum_{j \in \Theta_i(\omega) \cap E \setminus \Lambda(\omega)} u_j - \sum_{j \in \Theta_i(\omega) \cap \Lambda(\omega)} \max \left\{ u_j \right. \\ & \left. - \sum_{i' \in D} \alpha_{i'j}(\omega) d_{i'}(\omega), 0 \right\} \leq \sum_{i' \in \Omega_i(\omega) \cup \{i\}} \sum_{j \in \Theta_i(\omega) \cap N} X_{i'j}(\omega) \leq \sum_{j \in \Theta_i(\omega) \cap N} aY_j \\ & \forall i \in \{\bar{i} \in D : \Omega_{\bar{i}}(\omega) \neq \emptyset\}, \omega \in S. \end{aligned}$$

This proves that inequalities (4.34) are valid inequalities for our deterministic equivalent.

Our triage capacity constraints are

$$\begin{aligned} & \sum_{i \in D} X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} W_{j'j}(\omega) \leq aY_j \quad \forall j \in N, \omega \in S, \\ & \sum_{i \in D} X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} W_{j'j}(\omega) \leq u_j \quad \forall j \in E \setminus \Lambda(\omega), \omega \in S, \text{ and} \\ & \sum_{i \in D} X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} W_{j'j}(\omega) - \sum_{j' \in N \cup E} W_{jj'}(\omega) \leq u_j - \sum_{i \in D} \alpha_{ij}(\omega) d_i(\omega) \quad \forall j \in \Lambda(\omega), \omega \in S. \end{aligned}$$

Summing up the left-hand sides of these inequalities, we obtain

$$\sum_{i \in D} \sum_{j \in N \cup E} X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} \sum_{j \in N \cup E} W_{j'j}(\omega) - \sum_{j \in \Lambda(\omega)} \sum_{j' \in N \cup E} W_{jj'}(\omega) \quad \forall \omega \in S,$$

which reduces to

$$\sum_{i \in D} \sum_{j \in N \cup E} X_{ij}(\omega) \quad \forall \omega \in S.$$

Summing up the right-hand sides of our triage capacity constraints, we obtain

$$\sum_{j \in N} aY_j + \sum_{j \in E} u_j - \sum_{i \in D} \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega) d_i(\omega) \quad \forall \omega \in S.$$

Obviously,

$$\sum_{i \in D} \sum_{j \in N \cup E} X_{ij}(\omega) \leq \sum_{j \in N} aY_j + \sum_{j \in E} u_j - \sum_{i \in D} \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega) d_i(\omega) \quad \forall \omega \in S.$$

Since

$$\sum_{i \in D} \sum_{j \in N \cup E} X_{ij}(\omega) = \sum_{i \in D} \sum_{j \in \Pi_i(\omega)} X_{ij}(\omega) = \sum_{i \in D} \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega) \right] d_i(\omega) \quad \forall \omega \in S,$$

$$\sum_{i \in D} \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega) \right] d_i(\omega) \leq \sum_{j \in N} aY_j + \sum_{j \in E} u_j - \sum_{i \in D} \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega) d_i(\omega) \quad \forall \omega \in S.$$

Therefore,

$$\sum_{j \in N} aY_j \geq \sum_{i \in D} d_i(\omega) - \sum_{j \in E} u_j \quad \forall \omega \in S,$$

showing that inequalities (4.35) are valid inequalities for our deterministic equivalent.

Our treatment capacity constraints are

$$\sum_{i \in D} f_i(\beta_\omega) X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} \phi_{j'}^{SE}(\beta_\omega) W_{j'j}(\omega) + \sum_{l \in N \cup E \cup H} Z_{0lj}(\beta_\omega) - \sum_{j' \in N \cup E} Z_{0jj'}(\beta_\omega) \leq bY_j$$

$$\forall j \in N, \omega \in S, \beta_\omega \in I_\omega,$$

$$\sum_{i \in D} f_i(\beta_\omega) X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} \phi_{j'}^{SE}(\beta_\omega) W_{j'j}(\omega) + \sum_{l \in N \cup E \cup H} Z_{0lj}(\beta_\omega) - \sum_{j' \in N \cup E} Z_{0jj'}(\beta_\omega) \leq v_j$$

$$\forall j \in E \setminus \Lambda(\omega), \omega \in S, \beta_\omega \in I_\omega, \text{ and}$$

$$\begin{aligned} & \sum_{i \in D} f_i(\beta_\omega) X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} \phi_{j'}^{SE}(\beta_\omega) W_{j'j}(\omega) - \sum_{j' \in N \cup E} \phi_{j'}^{SE}(\beta_\omega) W_{jj'}(\omega) + \sum_{l \in N \cup E \cup H} Z_{0lj}(\beta_\omega) \\ & - \sum_{j' \in N \cup E} Z_{0jj'}(\beta_\omega) \leq v_j - \sum_{i \in D} f_i^{SE}(\beta_\omega) \alpha_{ij}(\omega) d_i(\omega) \quad \forall j \in \Lambda(\omega), \omega \in S, \beta_\omega \in I_\omega. \end{aligned}$$

Summing up the left-hand sides, we obtain

$$\sum_{i \in D} \sum_{j \in N \cup E} f_i(\beta_\omega) X_{ij}(\omega) + \sum_{l \in H} \sum_{j \in N \cup E} Z_{0lj}(\beta_\omega) \quad \forall \omega \in S, \beta_\omega \in I_\omega.$$

Summing up the right-hand sides, we obtain

$$\sum_{j \in N} bY_j + \sum_{j \in E} v_j - \sum_{i \in D} \sum_{j \in \Lambda(\omega)} f_i^{SE}(\beta_\omega) \alpha_{ij}(\omega) d_i(\omega) \quad \forall \omega \in S, \beta_\omega \in I_\omega.$$

Obviously,

$$\begin{aligned} & \sum_{i \in D} \sum_{j \in N \cup E} f_i(\beta_\omega) X_{ij}(\omega) + \sum_{l \in H} \sum_{j \in N \cup E} Z_{0lj}(\beta_\omega) \leq \sum_{j \in N} bY_j + \sum_{j \in E} v_j \\ & - \sum_{i \in D} \sum_{j \in \Lambda(\omega)} f_i^{SE}(\beta_\omega) \alpha_{ij}(\omega) d_i(\omega) \quad \forall \omega \in S, \beta_\omega \in I_\omega. \end{aligned}$$

Taking into consideration demand satisfaction constraints, we obtain

$$\sum_{i \in D} f_i(\beta_\omega) \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega) \right] d_i(\omega) + \sum_{l \in H} \sum_{j \in N \cup E} Z_{0lj}(\beta_\omega) \leq \sum_{j \in N} bY_j + \sum_{j \in E} v_j$$

$$-\sum_{i \in D} \sum_{j \in \Lambda(\omega)} f_i^{SE}(\beta_\omega) \alpha_{ij}(\omega) d_i(\omega) \quad \forall \omega \in S, \beta_\omega \in I_\omega. \quad (4.37)$$

From our hospital capacity constraints

$$\sum_{j \in N \cup E} Z_{1jk}(\beta_\omega) - \sum_{j \in N \cup E} Z_{0kj}(\beta_\omega) \leq m_k \quad \forall k \in H, \omega \in S, \beta_\omega \in I_\omega,$$

we can see that

$$\sum_{j \in N \cup E} \sum_{k \in H} Z_{1jk}(\beta_\omega) - \sum_{j \in N \cup E} \sum_{k \in H} Z_{0kj}(\beta_\omega) \leq \sum_{k \in H} m_k \quad \forall \omega \in S, \beta_\omega \in I_\omega.$$

This implies that

$$\sum_{j \in N \cup E} \sum_{k \in H} Z_{0kj}(\beta_\omega) \geq \sum_{j \in N \cup E} \sum_{k \in H} Z_{1jk}(\beta_\omega) - \sum_{k \in H} m_k \quad \forall \omega \in S, \beta_\omega \in I_\omega.$$

Moreover,

$$\sum_{j \in N \cup E} \sum_{k \in H} Z_{0kj}(\beta_\omega) \geq 0 \quad \forall \omega \in S, \beta_\omega \in I_\omega.$$

Therefore,

$$\sum_{j \in N \cup E} \sum_{k \in H} Z_{0kj}(\beta_\omega) \geq \max \left\{ \sum_{j \in N \cup E} \sum_{k \in H} Z_{1jk}(\beta_\omega) - \sum_{k \in H} m_k, 0 \right\} \quad \forall \omega \in S, \beta_\omega \in I_\omega.$$

From the constraints

$$\sum_{i \in D} \left[1 - f_i(\beta_\omega) \right] X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} \left[1 - \phi_{j'}^{SE}(\beta_\omega) \right] W_{j'j}(\omega) - \sum_{k \in H} Z_{1jk}(\beta_\omega) = 0$$

$\forall j \in N \cup E \setminus \Lambda(\omega), \omega \in S, \beta_\omega \in I_\omega$ and

$$\begin{aligned} & \sum_{i \in D} \left[1 - f_i(\beta_\omega) \right] X_{ij}(\omega) + \sum_{j' \in \Lambda(\omega)} \left[1 - \phi_{j'}^{SE}(\beta_\omega) \right] W_{j'j}(\omega) - \sum_{j' \in N \cup E} \left[1 - \phi_{j'}^{SE}(\beta_\omega) \right] W_{jj'}(\omega) \\ & - \sum_{k \in H} Z_{1jk}(\beta_\omega) + \sum_{i \in D} \left[1 - f_i^{SE}(\beta_\omega) \right] \alpha_{ij}(\omega) d_i(\omega) = 0 \quad \forall j \in \Lambda(\omega), \omega \in S, \beta_\omega \in I_\omega, \end{aligned}$$

if we perform left-hand sides and right-hand sides summation and use the demand

satisfaction constraints, we obtain

$$\begin{aligned} & \sum_{j \in N \cup E} \sum_{k \in H} Z_{1jk}(\beta_\omega) = \sum_{i \in D} \left[1 - f_i(\beta_\omega) \right] \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega) \right] d_i(\omega) \\ & + \sum_{i \in D} \sum_{j \in \Lambda(\omega)} \left[1 - f_i^{SE}(\beta_\omega) \right] \alpha_{ij}(\omega) d_i(\omega) \quad \forall \omega \in S, \beta_\omega \in I_\omega. \end{aligned}$$

Therefore,

$$\begin{aligned} \sum_{j \in N \cup E} \sum_{k \in H} Z_{0kj}(\beta_\omega) &\geq \max \left\{ \sum_{i \in D} \left[1 - f_i(\beta_\omega) \right] \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega) \right] d_i(\omega) \right. \\ &\quad \left. + \sum_{i \in D} \sum_{j \in \Lambda(\omega)} \left[1 - f_i^{SE}(\beta_\omega) \right] \alpha_{ij}(\omega) d_i(\omega) - \sum_{k \in H} m_k, 0 \right\} \quad \forall \omega \in S, \beta_\omega \in I_\omega. \end{aligned} \quad (4.38)$$

Combining (4.38) with (4.37), we obtain valid inequalities (4.36). $\square$

Inequalities (4.33) - (4.36), when added to the master problem, ensure that the location solution 1) provides enough capacity to triage all self-evacuees and treat all low-priority casualties and 2) has strong restrictions on capacity provision to triage non self-evacuees. Allocation of non self-evacuees for triage occurs in covering-time polygons. The number of non self-evacuees at a demand point i in initial damage scenario ω is $\left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega) \right] d_i(\omega)$. The number of non self-evacuees which can be allocated to existing ACFs that are not self-evacuee destinations inside the covering-time polygon of i in ω is $\sum_{j \in \Pi_i(\omega) \cap E \setminus \Lambda(\omega)} u_j$ and that which can be allocated to self-evacuee destinations inside the covering-time polygon of i in ω is $\sum_{j \in \Pi_i(\omega) \cap \Lambda(\omega)} \max \left\{ u_j - \sum_{i' \in D} \alpha_{i'j}(\omega) d_{i'}(\omega), 0 \right\}$. Therefore, the minimum number of non self-evacuees that new ACFs inside the covering-time polygon of i have to triage in scenario ω is $\left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega) \right] d_i(\omega) - \sum_{j \in \Pi_i(\omega) \cap E \setminus \Lambda(\omega)} u_j - \sum_{j \in \Pi_i(\omega) \cap \Lambda(\omega)} \max \left\{ u_j - \sum_{i' \in D} \alpha_{i'j}(\omega) d_{i'}(\omega), 0 \right\}$. The capacity provided by new ACFs inside the covering-time polygons of i must be enough to cover that minimum number of self-evacuees in every initial damage scenario for there to exist a feasible non self-evacuee triage allocation plan. Inequalities (4.33) express this condition for every demand point.

If demand point i has a covering-time polygon that intersects with at least one other covering-time polygon (two covering-time polygons are said to intersect when they both contain at least one common potential new ACF location or at least one common existing ACF with available capacity), this condition, though necessary, is weak. Let us illustrate this with a simple example. The triangles in Figure 4.2 are numbered potential ACF locations, each having a triage capacity of 5. The condition (4.33) is represented by inequalities $Y_1 + Y_2 + Y_3 \geq \frac{10}{5}$, $Y_2 + Y_3 + Y_4 + Y_5 + Y_6 \geq \frac{20}{5}$, and $Y_2 + Y_4 + Y_7 + Y_8 + Y_9 \geq \frac{15}{5}$. A solution where we choose locations 2, 3, 4, 5, and 7 satisfies (4.33). However, the total triage capacity it provides is $5 \times 5 = 25$, which is 20 short of the total demand. It is therefore impossible to find a feasible triage allocation plan if we only open ACFs 2, 3, 4, 5, and 7. We can eliminate more such infeasible locations by adding the condition (4.34) to ensure that the total new ACF capacity inside a system of intersecting covering-time polygons is enough to triage the minimum total non self-evacuee demand in that system. In our example, condition (4.34) yields $Y_1 + Y_2 + Y_3 + Y_4 + Y_5 + Y_6 + Y_7 + Y_8 + Y_9 \geq \frac{10 + 20 + 15}{5} = 9$. It is easy to see that there exists a feasible allocation plan if we choose all nine ACFs.

For self-evacuees, there are no covering-time polygon restrictions. If the total triage capacity, $\sum_{j \in N} aY_j$, provided by new ACFs is enough to triage the minimum number of casualties $\sum_{i \in D} d_i(\omega) - \sum_{j \in E} u_j$, we can always find a feasible triage allocation plan for self-evacuees. This is expressed in inequalities (4.35). For low-priority treatment as well, there are no covering-time polygon restrictions. The minimum number of low-priority casualties that new ACFs must treat

FIGURE 4.2: Example for triage satisfaction in intersecting polygons

is calculated using the same principle of maximal allocation at existing ACFs, with the additional term $\max \left\{ \sum_{i \in D} \left[1 - f_i(\beta_\omega) \right] \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega) \right] d_i(\omega) + \sum_{i \in D} \sum_{j \in \Lambda(\omega)} \left[1 - f_i^{SE}(\beta_\omega) \right] \alpha_{ij}(\omega) d_i(\omega) - \sum_{k \in H} m_k, 0 \right\}$, which is the number of low-priority patients that must be transferred out of hospitals to ACFs to liberate hospital capacity when the total number of high-priority casualties in β_ω exceeds the total hospital capacity available. The condition for low-priority treatment is expressed in inequalities (4.36).

4.4.2 Flow Cover Inequalities

The linear relaxation of our deterministic equivalent is an important consideration in Benders' decomposition. Tightening the polyhedron of the linear relaxation results in stronger cuts [92, 94]. Flow cover inequalities, a family of inequalities first

formulated by [95], are especially effective at tightening the relaxation for our case.

A set $\Delta_i(\omega) \subseteq \Pi_i(\omega) \cap N$ is called a cover if $a|\Delta_i(\omega)| > \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega)\right]d_i(\omega)$.

With $\lambda_i(\omega) = a|\Delta_i(\omega)| - \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega)\right]d_i(\omega)$, the following are valid flow cover inequalities for our model:

$$\begin{aligned} & \sum_{j \in \Pi_i(\omega) \cap N} X_{ij}(\omega) + \sum_{j \in \Delta_i(\omega)} \max\{a - \lambda_i(\omega), 0\}(1 - Y_j) \\ & \leq \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega)\right]d_i(\omega) + \sum_{j \in \Pi_i(\omega) \cap N \setminus \Delta_i(\omega)} \left[a - \lambda_i(\omega)\right]Y_j \\ & \forall i \in \{\bar{i} \in D : \Pi_{\bar{i}}(\omega) \cap E \setminus \Lambda(\omega) \neq \emptyset \vee A_{\bar{i}}(\omega) \neq \emptyset\}, \omega \in S \end{aligned} \quad (4.39)$$

and

$$\begin{aligned} & \sum_{j \in \Delta_i(\omega)} X_{ij}(\omega) + \sum_{j \in \Delta_i(\omega)} \max\{a - \lambda_i(\omega), 0\}(1 - Y_j) \leq \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega)\right]d_i(\omega) \\ & \forall i \in \{\bar{i} \in D : \Pi_{\bar{i}}(\omega) \cap E \setminus \Lambda(\omega) = \emptyset \wedge A_{\bar{i}}(\omega) = \emptyset\}, \omega \in S, \end{aligned} \quad (4.40)$$

where $A_{\bar{i}}(\omega) = \{j \in \Lambda(\omega) \cap \Pi_{\bar{i}}(\omega) : \sum_{i \in D} \alpha_{ij}(\omega)d_i(\omega) < u_j\}$.

Based on the works of [95], the inequality (4.39) for a demand point i and an

initial damage scenario ω is a valid inequality for the convex hull of the set

$\mathcal{P}_i(\omega) = \left\{ \left(\mathbf{x}_i(\omega), \mathbf{y} \right) \in \mathbb{R}^{|\Pi_i(\omega) \cap N|} \times \mathbb{B}^{|\Pi_i(\omega) \cap N|} : \sum_{j \in \Pi_i(\omega) \cap N} X_{ij}(\omega) \leq \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega)\right]d_i(\omega), 0 \leq X_{ij}(\omega) \leq aY_j, \forall j \in \Pi_i(\omega) \cap N \right\}$, where $\mathbf{x}_i(\omega)$ and $\mathbf{y}$

are vectors for $X_{ij}(\omega)$ and Y_j , respectively, where $j \in \Pi_i(\omega) \cap N$. The inequality

is therefore also valid for our deterministic equivalent. Based on the same paper,

when $a > \lambda_i(\omega)$, the inequality (4.40) for i and ω is a facet of the convex hull

of the set $\mathcal{Q}_i(\omega) = \left\{ \left(\mathbf{x}_i(\omega), \mathbf{y} \right) \in \mathbb{R}^{|\Pi_i(\omega) \cap N|} \times \mathbb{B}^{|\Pi_i(\omega) \cap N|} : \sum_{j \in \Pi_i(\omega) \cap N} X_{ij}(\omega) = \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega)\right]d_i(\omega), 0 \leq X_{ij}(\omega) \leq aY_j, \forall j \in \Pi_i(\omega) \cap N \right\}$.

The constraint $\sum_{j \in \Pi_i(\omega) \cap N} X_{ij}(\omega) = \left[1 - \sum_{j \in \Lambda(\omega)} \alpha_{ij}(\omega)\right]d_i(\omega)$ in $\mathcal{Q}_i(\omega)$ is only

valid for our model if the covering-time polygon of demand point i in initial damage

scenario ω satisfies two conditions: 1) it contains no existing ACFs that are not self-evacuee destinations and 2) it either does not contain any self-evacuee destinations or all its self-evacuee destinations are overwhelmed. When these two conditions are valid, covering-time-bound new ACFs are the only facilities eligible to triage non self-evacuees at i . In the flow cover inequalities (4.40), the set of demand points $\{i \in D : \Pi_i(\omega) \cap E \setminus \Lambda(\omega) = \emptyset \wedge A_i(\omega) = \emptyset\}$ satisfies the two conditions. The set $\Pi_i(\omega) \cap E \setminus \Lambda(\omega)$ is that of all existing ACFs that are not self-evacuee destinations in the covering-time polygon of demand point i in initial damage scenario ω and $A_i(\omega)$ is the set of existing ACFs that are not overwhelmed in the covering-time polygon of i in scenario ω . When both sets are empty, the propositional logic $\Pi_i(\omega) \cap E \setminus \Lambda(\omega) = \emptyset \wedge A_i(\omega) = \emptyset$ is true and our two conditions are satisfied. Although the facets (4.40) do not necessarily remain facets for the whole model's convex hull, they tighten the model relaxation significantly, especially for cases where these facets are valid for almost all demand points because, for instance, self-evacuees have overwhelmed almost all existing ACFs. Such cases are very likely to happen in CHEs.

4.4.3 Dual Subproblem

Typically, network problems such as ours have more than one optimal subproblem solution, and hence optimality cut, for the same location plan. Among these cuts, some are stronger than others. Our algorithm chooses cuts that are maximal nondominated. Suppose there are two cuts, $R + \sum_{j \in N} M_j Y_j \geq O$ and $R' + \sum_{j \in N} M'_j Y_j \geq O$, for the same location plan. [93] refer to the second cut

as maximal nondominated if $R' \leq R$ and $M'_j \leq M_j, \forall j \in N$, with at least one of these $|N| + 1$ inequalities being strict. A maximal nondominated cut is not necessarily pareto-optimal. Pareto-optimality requires that $R' + \sum_{j \in N} M'_j Y_j \leq R + \sum_{j \in N} M_j Y_j$ be true for all possible location plans, with a strict inequality for at least one of them. [92] have established a well-known method for generating pareto-optimal cuts, but we opt for maximal nondominated cuts because their generation needs less computational effort compared to the Magnanti-Wong method which requires the derivation of core points and the solution of additional linear programming models. By contrast, a simple perturbation of the dual subproblem yields maximal nondominated cuts.

In addition to perturbing the dual subproblem, we also decompose it such that we have one model for each initial damage scenario. This allows the algorithm to generate more cuts per iteration and to, therefore, have faster improving solutions. This method, proposed by [96], takes advantage of the block-angular property of stochastic programming models for decomposition. One must be careful to balance the master problem size and the number of cuts per iteration. The size of the master problem, an integer problem, increases faster with more cuts per iteration and this may lead to excessive solution times. There are many ways to vary the number of cuts per iteration to obtain the right balance, such as by decomposition with respect to both initial damage scenarios and triage scenarios or by decomposition into bundles of scenarios instead of individual ones. If ϵ is the perturbation factor and $\bar{Y}_j$ is the input location solution from the master problem for new ACF j , the formulation of the perturbed dual subproblem for

initial damage scenario ω is:

$$\begin{aligned}
\min \quad & \sum_{j \in N} a[\bar{Y}_j + \epsilon] \Psi_j^{C3}(\omega) + \sum_{\beta_\omega \in I_\omega} \sum_{j \in N} b[\bar{Y}_j + \epsilon] \Psi_j^{C6}(\beta_\omega) + \sum_{\beta_\omega \in I_\omega} \sum_{j \in N \cup E \setminus \Lambda(\omega)} \epsilon \Psi_j^{C9}(\beta_\omega) \\
& - \sum_{i \in D \setminus D'(\omega)} \sum_{j \in \Pi_i(\omega) \cap N} \left[1 - \alpha_i(\omega) \right] d_i(\omega) [1 - \bar{Y}_j - \epsilon] \Psi_{ij}^{C13}(\omega) + \sum_{i \in D'(\omega)} \sum_{j \in \Pi_i(\omega) \cap N} \\
& \left[\sum_{j' \in \Pi_i(\omega) \cap N \setminus j} \left[1 - \alpha_i(\omega) \right] d_i(\omega) [\bar{Y}_{j'} + \epsilon] - \left[1 - \alpha_i(\omega) \right] d_i(\omega) [1 - \bar{Y}_j - \epsilon] \right] \Psi_{ij}^{C13}(\omega) \\
& + (1 + \epsilon) \left[\sum_{i \in D} \left[1 - \alpha_i(\omega) \right] d_i(\omega) \Psi_i^{C1}(\omega) + \sum_{j \in \Lambda(\omega)} \max \left\{ 0, \sum_{i \in \Gamma_j(\omega)} \alpha_i(\omega) d_i(\omega) - u_j \right\} \Psi_j^{C2}(\omega) \right. \\
& + \sum_{j \in E \setminus \Lambda(\omega)} u_j \Psi_j^{C4}(\omega) + \sum_{j \in \Lambda(\omega)} \left[u_j - \sum_{i \in \Gamma_j(\omega)} \alpha_i(\omega) d_i(\omega) \right] \Psi_j^{C5}(\omega) + \sum_{\beta_\omega \in I_\omega} \sum_{j \in E \setminus \Lambda(\omega)} v_j \Psi_j^{C7}(\beta_\omega) \\
& + \sum_{\beta_\omega \in I_\omega} \sum_{j \in \Lambda(\omega)} \left[v_j - \sum_{i \in \Gamma_j(\omega)} f_i^{SE}(\beta_\omega) \alpha_i(\omega) d_i(\omega) \right] \Psi_j^{C8}(\beta_\omega) \\
& - \sum_{\beta_\omega \in I_\omega} \sum_{j \in \Lambda(\omega)} \left[\sum_{i \in \Gamma_j(\omega)} \left[1 - f_i^{SE}(\beta_\omega) \right] \alpha_i(\omega) d_i(\omega) \right] \Psi_j^{C10}(\beta_\omega) \\
& \left. + \sum_{\beta_\omega \in I_\omega} \sum_{k \in H} g_k \Psi_k^{C11}(\beta_\omega) + \sum_{\beta_\omega \in I_\omega} \sum_{k \in H} m_k \Psi_k^{C12}(\beta_\omega) \right] \\
\text{s.t.} \quad & \Psi_i^{C1}(\omega) + \theta_j^{S2} \left[\Psi_j^{C3}(\omega) + \sum_{\beta_\omega \in I_\omega} f_i(\beta_\omega) \Psi_j^{C6}(\beta_\omega) + \sum_{\beta_\omega \in I_\omega} \left[1 - f_i(\beta_\omega) \right] \Psi_j^{C9}(\beta_\omega) \right] \\
& + \theta_j^{S3}(\omega) \left[\Psi_j^{C4}(\omega) + \sum_{\beta_\omega \in I_\omega} f_i(\beta_\omega) \Psi_j^{C7}(\beta_\omega) + \sum_{\beta_\omega \in I_\omega} \left[1 - f_i(\beta_\omega) \right] \Psi_j^{C9}(\beta_\omega) \right] \\
& + \theta_j^{S4}(\omega) \left[\Psi_j^{C5}(\omega) + \sum_{\beta_\omega \in I_\omega} f_i(\beta_\omega) \Psi_j^{C8}(\beta_\omega) + \sum_{\beta_\omega \in I_\omega} \left[1 - f_i(\beta_\omega) \right] \Psi_j^{C10}(\beta_\omega) \right] \\
& + \theta_j^{S1}(\omega) \theta_j^{S2} \theta_i^{S5}(\omega) \sum_{j \in \Pi_i(\omega) \cap N} \Psi_{ij}^{C13}(\omega) + \theta_j^{S1}(\omega) \theta_j^{S2} \left[1 - \theta_i^{S5}(\omega) \right] \Psi_{ij}^{C13}(\omega) \\
& \geq -Pr_\omega t_{ij}(\omega) \quad \forall i \in D, j \in N \cup E, \beta_\omega \in I_\omega
\end{aligned} \tag{4.41}$$

$$\begin{aligned}
& \Psi_{j'}^{C2}(\omega) - \Psi_{j'}^{C5}(\omega) - \sum_{\beta_\omega \in I_\omega} \phi_{j'}^{SE}(\beta_\omega) \Psi_{j'}^{C8}(\beta_\omega) - \sum_{\beta_\omega \in I_\omega} \left[1 - \phi_{j'}^{SE}(\beta_\omega) \right] \Psi_{j'}^{C10}(\beta_\omega) \\
& + \theta_j^{S2} \left[\Psi_j^{C3}(\omega) + \sum_{\beta_\omega \in I_\omega} \phi_{j'}^{SE}(\beta_\omega) \Psi_j^{C6}(\beta_\omega) + \sum_{\beta_\omega \in I_\omega} \left[1 - \phi_{j'}^{SE}(\beta_\omega) \right] \Psi_j^{C9}(\beta_\omega) \right] \\
& + \theta_j^{S3}(\omega) \left[\Psi_j^{C4}(\omega) + \sum_{\beta_\omega \in I_\omega} \phi_{j'}^{SE}(\beta_\omega) \Psi_j^{C7}(\beta_\omega) + \sum_{\beta_\omega \in I_\omega} \left[1 - \phi_{j'}^{SE}(\beta_\omega) \right] \Psi_j^{C9}(\beta_\omega) \right]
\end{aligned}$$

$$\begin{aligned}
& + \theta_j^{S4}(\omega) \left[\Psi_j^{C5}(\omega) + \sum_{\beta_\omega \in I_\omega} \phi_{j'}^{SE}(\beta_\omega) \Psi_j^{C8}(\beta_\omega) + \sum_{\beta_\omega \in I_\omega} [1 - \phi_{j'}^{SE}(\beta_\omega)] \Psi_j^{C10}(\beta_\omega) \right] \\
& \geq -Pr_\omega t_{j'j}(\omega) \quad \forall j' \in \Lambda(\omega), j \in N \cup E, \beta_\omega \in I_\omega
\end{aligned} \tag{4.42}$$

$$\begin{aligned}
& \theta_{j'}^{S2} \Psi_{j'}^{C6}(\beta_\omega) - \theta_j^{S2} \Psi_j^{C6}(\beta_\omega) + \theta_{j'}^{S3}(\omega) \Psi_{j'}^{C7}(\beta_\omega) - \theta_j^{S3}(\omega) \Psi_j^{C7}(\beta_\omega) + \theta_{j'}^{S4}(\omega) \Psi_{j'}^{C8}(\beta_\omega) \\
& - \theta_j^{S4}(\omega) \Psi_j^{C8}(\beta_\omega) \geq -Pr_\omega Pr_{\beta|\omega} q_0 t_{jj'}(\omega) \quad \forall j' \in N \cup E, j \in N \cup E, \beta_\omega \in I_\omega
\end{aligned} \tag{4.43}$$

$$\begin{aligned}
& \theta_j^{S2} \Psi_j^{C6}(\beta_\omega) + \theta_j^{S3}(\omega) \Psi_j^{C7}(\beta_\omega) + \theta_j^{S4}(\omega) \Psi_j^{C8}(\beta_\omega) + \Psi_k^{C11}(\beta_\omega) - \Psi_k^{C12}(\beta_\omega) \\
& \geq -Pr_\omega Pr_{\beta|\omega} q_0 t_{kj}(\omega) \quad \forall k \in H, j \in N \cup E, \beta_\omega \in I_\omega
\end{aligned} \tag{4.44}$$

$$\begin{aligned}
& \Psi_k^{C12}(\beta_\omega) - \theta_j^{S2} \Psi_j^{C9}(\beta_\omega) - \theta_j^{S3}(\omega) \Psi_j^{C9}(\beta_\omega) - \theta_j^{S4}(\omega) \Psi_j^{C10}(\beta_\omega) \\
& \geq -Pr_\omega Pr_{\beta|\omega} q_1 t_{jk}(\omega) \quad \forall k \in H, j \in N \cup E, \beta_\omega \in I_\omega
\end{aligned} \tag{4.45}$$

$$\text{All } \Psi_i^{C1}(\omega), \Psi_j^{C2}(\omega), \Psi_j^{C9}(\beta_\omega), \Psi_j^{C10}(\beta_\omega) \in \mathbb{R}, \text{ Other variables } \geq 0, \tag{4.46}$$

where $\theta_j^{S1}(\omega) = 1$ if ACF $j \in \Pi_i(\omega)$ and 0 otherwise, $\theta_j^{S2} = 1$ if $j \in N$ and 0 otherwise, $\theta_j^{S3}(\omega) = 1$ if $j \in E \setminus \Lambda(\omega)$ and 0 otherwise, $\theta_j^{S4}(\omega) = 1$ if $j \in \Lambda(\omega)$ and 0 otherwise, $\theta_i^{S5}(\omega) = 1$ if $i \in D'(\omega)$ and 0 otherwise, where $D'(\omega) = \{i \in D : \Pi_i(\omega) \cap E \setminus \Lambda(\omega) \neq \emptyset \vee A_i(\omega) \neq \emptyset\}$, and $\theta_{ij}^{S6}(\omega) = 1$ if $j \in \Delta_i(\omega)$ and $i \in D$ and 0 otherwise. The superscript Cn of a dual variable indicates that the variable corresponds to the n^{th} set of constraints of the deterministic equivalent (with location variables set as parameters). Decision variables $\Psi_i^{C13}(\omega)$ are variables associated with flow cover inequalities. The dual has an interesting structure. Constraints (4.41) impose bounds on reduced costs for triage allocation from demand points to ACFs, while constraints (4.42) are for triage allocation of excess self-evacuees to ACFs. Constraints (4.43), (4.44), and (4.45) bound reduced costs for ACF-to-ACF, hospital-to-ACF, and ACF-to-hospital treatment

allocation, respectively.

4.4.4 Algorithm Steps

Below, we describe the steps of our algorithm. Note the omission of transpose symbols from vector multiplications to keep notations simple.

Initialization. Set $r = 0$, $U \leftarrow \infty$, and $L \leftarrow -\infty$. Solve the master problem $\max \left\{ \sum_{\omega \in S} O_{\omega} : \sum_{j \in N} Y_j \leq V, (4.33) - (4.36), \forall \omega \in S, Y_j \in \{0, 1\}, \forall j \in N \right\}$. Store the binary variable solution for every new ACF j in $\bar{Y}_j^{(r)}$.

Step 1. Solve the perturbed dual subproblem

$$\begin{aligned} \min \quad & (1 + \epsilon) \mathbf{B}_{\omega} \Phi_{\omega} + \sum_{j \in N} \mathbf{C}_{\omega j} \Xi_{\omega j} (\bar{Y}_j^{(r)} + \epsilon) \\ \text{s.t.} \quad & \mathbf{F}_{\omega} \begin{pmatrix} \Phi_{\omega} \\ \Xi_{\omega 1} \\ \vdots \\ \Xi_{\omega |N|} \end{pmatrix} \geq \mathbf{G}_{\omega} \end{aligned}$$

for each initial damage scenario $\omega \in S$. ϵ is a small positive number used to create perturbation, and Φ_{ω} and $\Xi_{\omega j}$ are vectors of decision variables. The remaining notations $\mathbf{B}_{\omega}$, $\mathbf{C}_{\omega j}$, $\mathbf{F}_{\omega}$, and $\mathbf{G}_{\omega}$ are matrices/vectors containing parameters. For every initial damage scenario $\omega \in S$, if the perturbed dual subproblem is feasible and bounded, go to Step 1a. If it is unbounded, set $T_{\omega}^{(r)} \leftarrow -\infty$ and go to Step 1b. If it is infeasible, stop. The whole problem is infeasible.

Step 1a. Store the optimal objective function value in $T_{\omega}^{(r)}$ and optimal dual variable values in vectors $\bar{\Phi}_{\omega}^{(r)}$ and $\bar{\Xi}_{\omega j}^{(r)}$. Add optimality cut $\mathbf{B}_{\omega} \bar{\Phi}_{\omega}^{(r)} + \sum_{j \in N} \mathbf{C}_{\omega j} \bar{\Xi}_{\omega j}^{(r)} Y_j \geq O_{\omega}$ to the master problem.

Step 1b. Solve the modified dual subproblem

$$\begin{aligned}
 & \min \text{ Dummy} \\
 & \text{s.t.} \quad (1 + \epsilon) \mathbf{B}_\omega \Phi_\omega + \sum_{j \in N} \mathbf{C}_{\omega j} \Xi_{\omega j} (\bar{Y}_j^{(r)} + \epsilon) = -1 \\
 & \quad \mathbf{F}_\omega \begin{pmatrix} \Phi_\omega \\ \Xi_{\omega 1} \\ \vdots \\ \Xi_{\omega |N|} \end{pmatrix} \geq 0 \\
 & \quad \text{Dummy} = 0,
 \end{aligned}$$

where *Dummy* is a dummy variable. Store the optimal dual variable values in vectors $\bar{\Phi}_\omega^{(r)}$ and $\bar{\Xi}_{\omega j}^{(r)}$. Add feasibility cut $\mathbf{B}_\omega \bar{\Phi}_\omega^{(r)} + \sum_{j \in N} \mathbf{C}_{\omega j} \bar{\Xi}_{\omega j}^{(r)} Y_j \geq 0$ to the master problem.

Step 2. If $\sum_{\omega \in S} T_\omega^{(r)} > L$, $L \leftarrow \sum_{\omega \in S} T_\omega^{(r)}$, where L is the current best objective value. Set $r = r + 1$. Solve the master problem. Store the binary variable solution for every new ACF j in $\bar{Y}_j^{(r)}$, and the optimal values of O_ω in $\bar{O}_\omega^{(r)}$, for each $\omega \in S$. If $\sum_{\omega \in S} \bar{O}_\omega^{(r)} < U$, $U \leftarrow \sum_{\omega \in S} \bar{O}_\omega^{(r)}$. If $U - L < \tau$, where τ is a small tolerance value, stop, else return to Step 1.

4.5 Case Study

We implement our model and algorithm in the ShakeOut Scenario [see 97], a probable earthquake situation put together in a major collaborative effort by the U.S. Geological Survey and California Geological Survey, involving more than 300 experts and contributors. The ShakeOut Scenario report (SSR) details the impact

of a 7.8-magnitude earthquake on the southernmost 300 km of the San Andreas Fault, between the Salton Sea and Lake Hughes, occurring at 10 a.m. on a weekday. Damage and casualty estimates are based on rigorous ground motion analyses and physics-based computer simulations, combined with the Federal Emergency Management Agency's loss estimation software HAZUS-MH and expert opinions through 13 special studies and 6 expert panels.

Casualties spread over eight counties, with more than 99% of these casualties in Los Angeles, Orange, Riverside, and San Bernardino. For simplicity, we restrict our region of study to these four counties. We place demand points at the population centers of gravity of the subcounties of these counties. These subcounties are county subdivisions provided by the U.S. Census Bureau. To prevent demand points from covering too large areas, we further subdivide large subcounties and place demand points at the resulting subdivisions' population centers of gravity instead.

For each county, the SSR provides the expected number of casualties resulting from four sources, namely, ground motion incidents, damage to steel-frame buildings, fire, and transportation incidents. We use these data to estimate casualty demands for our first initial damage scenario. Assuming a uniform distribution of casualties, the demand in a subcounty is calculated using the formula: $\text{Subcounty Demand} = (\text{Subcounty Population} / \text{County Population}) \times \text{Number of Casualties in County}$. We also assume that subcounty subdivisions have equal populations and use the same principle to estimate their demands. In addition to data on the expected number of casualties, the SSR also provides

the ratios of serious injuries (high-priority) to non-serious injuries (low-priority) for every county. We use these ratios to form triage scenarios, taking the ratio of high-priority to low-priority casualties in a county to be the same for all its subdivisions.

Damages are not known with certainty. According to the SSR, aftershock possibilities can raise damages by up to 30%. Based on this information, we build a second initial damage scenario with 30% higher number of casualties at every demand point and, for each initial damage scenario, a second triage scenario with 30% of low-priority casualties converted to high-priority.

Among casualties, 95% are rescued by fellow victims, with no emergency responder interference. These rescues constitute the self-evacuees in our first initial damage scenario. Self-evacuees are expected to move to designated trauma centers, a list of which is provided by California Emergency Medical Services Authority [see 98]. Trauma centers are hospitals which are seismically sound and can remain self-sufficient after an earthquake to provide high-priority care. We assume that a self-evacuee moves to the trauma center that is closest (by travel time) to his/her demand point. The treatment capacity at a trauma center is 22% of its bed capacity, based on the 2007 California average trauma bed occupancy rate of 78% [see 99]. Because California's Assembly Bill No. 394 requires a 1:1 nurse-to-patient ratio for trauma cases and a 1:4 ratio for triage [100], a trauma center can triage $(\text{Number of Nurses} - \text{Number of Beds}) \times 4$ casualties. There are no other existing ACFs because we assume medical facilities that are not trauma centers to be severely damaged and, as a result, not functional.

Some of the data we need are not available in the SSR or other reports or articles. We use fictitious substitutes for them. The SSR mentions that transportation times are 21% higher than usual after the earthquake. We use this knowledge to calculate transportation times for our first initial damage scenario. In our second initial damage scenario, we use a percentage of 25%. Pre-disaster transportation times are obtained using the ArcGIS 9.3 Network Analyst tool. We also use a percentage of self-evacuees of 97% for our second initial damage scenario and assume that 20% (the California average discharge rate of patients [see 101]) of trauma patients at hospitals can be moved to ACFs to continue their recovery.

Suppose emergency responders have 50 ACFs, each of triage capacity 1000 and treatment capacity 1000 to deploy at public schools with, for the sake of a base case, student capacities greater than or equal to 2500. In the ShakeOut Scenario, schools experience minimal damage and are even suggested as possible Red Cross shelters. To ensure feasibility, we also use the biggest school in every covering-time (set to 15 minutes) polygon that is without any school with student capacity greater than or equal to 2500. The resulting total number of potential ACF locations is 200. Figure 4.3 shows the region under consideration with its demand points, possible ACF locations, and hospitals (designated trauma centers). The shaded subcounties indicate the omission, for simplicity, of subcounties with close to zero casualties. Existing ACFs are placed at hospital locations and have no treatment capacity allocated for low-priority casualties. Let priority weights be $q_0 = 2$ and $q_1 = 10$.

FIGURE 4.3: Region under study

4.5.1 Optimal Plan

Figure 4.4 illustrates the optimal response plan for the first initial damage scenario and triage scenario in San Bernardino subcounty, which is the subcounty with the highest casualty count. The covering-time polygons of the three demand points are also shown. Outbound triage allocation arrows from hospitals indicate that they are overwhelmed with self-evacuees and have to transfer excesses out. Furthermore, after triage, all low-priority casualties at hospitals must be allocated to new ACFs for treatment, given that hospitals do not reserve any low-priority treatment capacity. No transfer of low-priority casualties occurs between new ACFs because each has a triage capacity equal to its treatment capacity, meaning that there is sufficient capacity to treat all triaged casualties. Non self-evacuees are transported over short distances to in-polygon ACFs due to the imposed covering time. Finally, it is important to note the interchange of casualties between San Bernardino and other subcounties, showing the collaborative nature of the response plan.

4.5.2 Algorithm Performance

The need for an algorithm stems from the fact that the solution time increases steeply with the number of possible ACF locations. Table 4.1, shows the rapid increase in the solution time as $|N|$, the number of possible ACF locations, is increased by using schools of smaller student capacities as possible locations. Our models are solved with IBM-ILOG CPLEX 12.2 on a 3 GHz dual-core processor

FIGURE 4.4: Illustration of the optimal plan

with 4 GB of RAM. When $|N| = 429$, the model takes almost 3 hours to solve, which is unacceptable in emergency situations.

TABLE 4.1: Variation of the optimal value and solution time with $|N|$

Student capacity	$ N $	Optimal value (min)	Solution time (sec)
≥ 3000	133	1,446,070	123.0
≥ 2500	200	1,241,900	738.2
≥ 2000	308	1,140,570	2,333.6
≥ 1500	429	1,088,550	10,114.4

Some side experimentations indicate that $\epsilon = 10^{-6}$ is a suitable perturbation factor. For insights on algorithm effectiveness, we compare three configurations of our algorithm, which we call A, B, and C. Configuration A is our algorithm without our master problem modification and flow cover inequalities. This configuration is effectively a classical multi-cut Benders' decomposition algorithm with perturbed

dual subproblems. Perturbation and multiple cuts are proven acceleration techniques in the literature. In configuration B, we add our derived valid inequalities to the master problem of configuration A to observe the improvement they yield. In configuration C, we add our flow cover inequalities to configuration B to test their contribution. Configuration C is basically our complete algorithm. Table 4.2 shows the full list of the best solutions obtained with time. Figure 4.5 offers a pictorial representation of the behaviours of configurations B and C in early time periods so as to compare their abilities to generate good solutions fast. In these time periods, configuration A is unable to find any initial feasible solution.

FIGURE 4.5: Algorithm performance

In the $|N| = 308$ and $|N| = 429$ cases, configuration A is unable to improve any bounds within one hour, mainly due to weak feasibility cuts from unbounded dual subproblems. For $|N| = 133$, the initial feasible solution is obtained after 677.8

TABLE 4.2: Solutions over time for different configurations of our algorithm

N = 133					
Configuration A		Configuration B		Configuration C	
Time (sec)	Solution (min)	Time (sec)	Solution (min)	Time (sec)	Solution (min)
[677.8, 704.4)	1,637,780	[2.4, 4.5)	1,794,200	[1.2, 2.4)	1,934,650
[704.4, 706.5)	1,535,740	[4.5, 15.4)	1,489,980	[2.4, 4.4)	1,482,390
[706.5, 707.2)	1,496,670	[15.4, 190.2)	1,453,240	[2.4, 4.4)	1,482,390
[707.2, 769.4)	1,464,050	[190.2, 227.0)	1,450,500	[4.4, 7.3)	1,462,860
[769.4, 789.2)	1,458,820	[227.0, 393.7)	1,447,580	7.3	1,446,070
[789.2, 1049.1)	1,447,580	393.7	1,446,070		
1049.1	1,446,070				
N = 200					
Configuration A		Configuration B		Configuration C	
Time (sec)	Solution (min)	Time (sec)	Solution (min)	Time (sec)	Solution (min)
[2024.4, 2070.3)	1,476,400	[4.9, 13.6)	1,803,510	[2.7, 9.5)	1,937,040
[2070.3, 2073.0)	1,456,980	[13.6, 24.7)	1,455,980	[9.5, 14.9)	1,438,350
[2073.0, 2074.7)	1,420,780	[24.7, 88.1)	1,336,330	[9.5, 14.9)	1,438,350
[2074.7, 2080.5)	1,373,600	[88.1, 105.2)	1,331,850	[14.9, 25.5)	1,359,720
[2080.5, 2083.6)	1,367,960	[105.2, 459.0)	1,281,180	[25.5, 28.8)	1,331,430
[2083.6, 2142.6)	1,300,130	[459.0, 553.8)	1,278,000	[28.8, 47.9)	1,260,720
[2142.6, 2196.4)	1,299,000	[553.8, 589.2)	1,273,180	[47.9, 75.6)	1,247,810
[2196.4, 2258.4)	1,286,110	[589.2, 738.6)	1,265,680	[75.6, 84.9)	1,246,550
[2258.4, 2393.3)	1,279,820	[738.6, 1079.5)	1,262,340	[84.9, 89.2)	1,243,140
[2393.3, 2401.5)	1,277,450	[1079.5, 1498.3)	1,260,380	[89.2, 205.8)	1,242,520
[2401.5, 2520.3)	1,275,710	[1498.3, 3600.0]	1,252,010	[205.8, 211.2)	1,242,430
[2520.3, 2583.4)	1,260,120			[211.2, 252.8)	1,242,400
[2583.4, 3600.0]	1,251,160			252.8	1,241,900
N = 308					
Configuration A		Configuration B		Configuration C	
Time (sec)	Solution (min)	Time (sec)	Solution (min)	Time (sec)	Solution (min)
No initial feasible solution		[7.0, 25.9)	2,017,890	[6.9, 39.1)	1,922,730
		[25.9, 49.2)	1,496,230	[39.1, 60.6)	1,334,180
		[49.2, 102.0)	1,357,190	[60.6, 95.4)	1,301,550
		[102.0, 113.5)	1,343,560	[95.4, 105.7)	1,201,330
		[113.5, 131.3)	1,310,930	[105.7, 183.3)	1,154,820
		[131.3, 138.5)	1,253,310	[183.3, 232.1)	1,149,760
		[138.5, 221.0)	1,196,510	[232.1, 381.7)	1,147,580
		[221.0, 277.7)	1,191,770	[381.7, 520.0)	1,142,870
		[277.7, 940.4)	1,179,440	[520.0, 683.9)	1,142,130
		[940.4, 1754.2)	1,176,720	[683.9, 1057.6)	1,141,500
		[1754.2, 3279.5)	1,172,360	[1057.6, 1115.3)	1,141,290
		[3279.5, 3453.3)	1,165,110	[1115.3, 1126.3)	1,140,990
		[3453.3, 3600.0]	1,159,490	[1126.3, 1183.8)	1,140,640
				1183.8	1,140,570
N = 429					
Configuration A		Configuration B		Configuration C	
Time (sec)	Solution (min)	Time (sec)	Solution (min)	Time (sec)	Solution (min)
No initial feasible solution		[15.8, 101.1)	2,025,160	[13.6, 126.3)	1,731,930
		[101.1, 213.5)	1,491,060	[126.3, 210.0)	1,328,500
		[213.5, 302.7)	1,413,000	[210.0, 271.3)	1,279,700
		[302.7, 446.8)	1,247,300	[271.3, 416.9)	1,174,990
		[446.8, 493.7)	1,239,980	[416.9, 531.6)	1,112,990
		[493.7, 602.0)	1,225,590	[531.6, 734.6)	1,110,190
		[602.0, 790.0)	1,162,220	[734.6, 859.1)	1,104,380
		[790.0, 3600.0]	1,120,030	[859.1, 952.7)	1,101,960
				[952.7, 1049.8)	1,095,640
				[1049.8, 3600.0]	1,088,610

seconds and for $|N| = 200$, it is obtained after 2024.4 seconds. In configuration B, solutions improve drastically, primarily due to the reduction of feasibility cuts via our master problem modification. Configuration B finds the optimal solution for the $|N| = 133$ case within 393.7 seconds, whilst for the remaining cases, solutions improve fast with time, reaching close (2.9% from the optimal in the worst case) to optimal on the hour mark.

The addition of flow cover inequalities contributes further improvement, with generally faster solution improvement over time. The $|N| = 133$, $|N| = 200$, and $|N| = 308$ cases for configuration C find optimal solutions within 7.3 seconds, 252.8 seconds, and 1183.8 seconds, respectively. For the $|N| = 429$ case, the solution on the hour mark is almost optimal (0.005% higher).

4.5.3 ACF Budgeting and Capacity Sensitivity Analysis under Fixed Total Resource

Suppose emergency responders have a total resource equalling 50 ACFs, each with a triage capacity of 1500 and a treatment capacity of 1500. They can choose to locate more but smaller, or less but bigger ACFs, maintaining the condition, assumed in our model, that all new ACFs have the same capabilities. This means that if responders want to change the budget by δ (can be a positive or negative number), every ACF will have a triage capacity (and treatment capacity) equal to $\frac{50}{50 + \delta} \times 1500$. In addition to changing the budget, responders can vary the

triage-to-treatment capacity ratio, by converting triage capacity to treatment capacity and vice versa. Figure 4.6 shows the change in the optimal objective value for various budgeting and capacity planning possibilities in a fixed total resource situation.

FIGURE 4.6: Sensitivity analysis on the budget and triage-to-treatment capacity ratio for fixed total resource

A bigger budget of smaller ACFs tends to improve the optimal solution, with a decreasing rate. With increasing budget, more locations can be chosen, but at the expense of some casualties having to travel longer because of smaller capacities. Due to prioritization, the choice of more locations is favourable as it allows the placement of ACFs close to hospitals or other ACFs to shorten the travel times of high-priority and low-priority casualties for treatment.

The triage-to-treatment capacity ratio also affects the transfer of low-priority casualties for treatment. Decreasing the ratio means that more triaged casualties can be afforded treatment without having to transfer to other ACFs. However,

when the ratio is too low, longer travel times for triage and high-priority treatment outweigh the travel time saved on low-priority treatment allocation, leading to worse optimal plans.

4.5.4 ACF Budgeting and Capacity Sensitivity Analysis under Variable Total Resource

The total available resource is also often modifiable during a disaster. Responders can increase the number of ACFs without capacity alterations by requesting for more emergency response teams from other agencies, states, or countries. If they wish to decrease it, they can reallocate ACF personnel to other activities, such as helping at hospitals and search-and-rescue. Figure 4.7 shows how the optimal value varies with the ACF budget.

FIGURE 4.7: Sensitivity analysis on the budget of available ACFs for variable total resource

Having more ACFs at one's disposal improves the optimal objective value, but with decreasing marginal improvement. A limit is reached beyond which the optimal value remains constant, no matter how many more ACFs available. Knowledge of this limit helps avoid an over-supply of ACFs and therefore maintain better resource planning. In our case study, we observe that responders do not need to waste resources, which could be better employed doing other tasks, to set up more than 90 ACFs.

4.5.5 Conversion of High-Priority Treatment Capacity to Triage Capacity at Hospitals

Figure 4.8 shows how the optimal value changes when high-priority treatment capacity is converted to triage capacity at all hospitals. We show the variation for Loma Linda on a separate graph because its scale of change is larger than the rest, making changes for other hospitals appear insignificant when incorporated in the same graph. Every removed unit of high-priority treatment capacity adds four units of triage capacity because of the 1:4 nurse-to-patient triage ratio.

For hospitals operating at full treatment capacity utilization in at least one triage scenario, removing treatment capacity worsens the plan for two main reasons. First, it means that high-priority casualties who were previously afforded treatment have to be allocated to other less preferable hospitals. Second, in case of overwhelming numbers of self-evacuees, the additional triage capacity forces

FIGURE 4.8: Sensitivity analysis on the conversion of high-priority treatment capacity to triage at hospitals

hospitals to triage more self-evacuees (since they cannot refuse triage when capacity is available), with the likely result being that these self-evacuees have to travel longer to obtain treatment. Had these hospitals not been forced to triage these self-evacuees, they would have sent the latter to new ACFs that are close to hospitals with available treatment capacity for triage. This would have yielded transportation times that are higher for triage but lower for treatment. Because of prioritization weights, the benefit of lower transportation times for treatment

tends to outweigh the cost of longer travels for triage.

For those hospitals which have an overwhelming number of self-evacuees and under-utilized treatment capacity, converting some excess treatment capacity to triage capacity also forces more self-evacuee triage but with the benefit of treating those found to be high-priority with the remaining treatment capacity. The downside is that low-priority self-evacuees tend to have higher transportation times for treatment. The cost-benefit interplay can improve the optimal plan for some hospitals (such as Mission Medical Center), but not others (such as Henry Mayo Newhall Memorial Hospital).

For those hospitals with an under-utilized triage capacity and high-priority treatment capacity, such as Northridge Medical Center, UCLA Medical Center, and Cedars-Sinai Medical Center, conversion of the excess high-priority treatment capacity to triage does not affect the optimal plan since both the triage capacity and the treatment capacity are already to spare. Beyond the excess treatment capacity, these hospitals behave similar to the full utilization case.

4.5.6 Conversion of High-Priority Treatment Capacity to Low-Priority Treatment Capacity at Hospitals

The general trend when hospitals convert high-priority treatment capacity to low-priority treatment capacity is an improvement in the optimal plan, as shown in Figure 4.9. Because California's Assembly Bill No. 394 requires a 1:4 nurse-to-patient ratio for all emergency room patients that are not trauma or ICU cases,

we assume that a unit of high-priority treatment capacity is converted to four units of low-priority treatment capacity. For a hospital with under-utilized high-priority treatment capacity, using the excess treatment capacity for the treatment of low-priority casualties benefits the optimal plan by allowing more triaged low-priority casualties to receive treatment without having to transfer to other ACFs. For hospitals that are at full treatment capacity utilization, removing units of treatment capacity increases the highly-weighted transportation times of high-priority casualties for treatment. However, this is outweighed by the benefits of shorter low-priority casualty transportation times due to the 1:4 ratio and the fact that, in the ShakeOut Scenario, the vast majority (more than 95% on average) of casualties are low-priority in all scenarios. Due to the large number of low-priority casualties at hospitals, nearby ACFs have inadequate capacity to treat all, leaving many to travel for long times to obtain treatment. Therefore, with additional low-priority treatment capacity at hospitals, the optimal plan is spared these long travel times (four for each high-priority treatment capacity converted), even if it means that high-priority casualties have to travel longer for treatment.

4.5.7 Conversion of Triage Capacity to Low-Priority Treatment Capacity at Hospitals

Conversion of triage capacity to low-priority treatment capacity at hospitals exhibits close to the same trends as for ACFs, as shown in Figure 4.10. Our hospitals tend to give the best optimal plan when they can convert triage to low-priority treatment in such a way that they can treat all triaged casualties. Although this

FIGURE 4.9: Sensitivity analysis on the conversion of high-priority treatment capacity to low-priority treatment capacity at hospitals

will lead to longer transportation times for high-priority treatment, the much larger number of low-priority casualties counter-balances the effect. In addition, with a lower triage capacity, fewer self-evacuees are forced to undergo triage, meaning that they can be sent to other ACFs where they would obtain both triage and treatment. Northridge Medical Center shows anomalous behaviour because it triages and treats both self-evacuees and non self-evacuees. In the beginning, the plan gets better as its excess triage capacity is converted. Removal of more triage

capacity results in non self-evacuees having to seek other less favourable destinations for triage. Beyond the point where we remove all triage capacity for non self-evacuees, the optimal plan behaves similar to other hospitals.

FIGURE 4.10: Sensitivity analysis on the conversion of triage capacity to low-priority treatment capacity at hospitals

4.5.8 Insights

Our sensitivity analyses show that when excess capacity is available at a hospital, it is better to convert it into a combination of triage and low-priority treatment. The lines of Figure 4.10 show generally lower minimum optimal values compared to those of Figures 4.8 and 4.9. The minimum for a hospital occurs at the point where its treatment capacity is sufficient to attend to all its triaged low-priority casualties. From our sensitivity analyses, it would seem that our plan is low-priority-centric, given that improving low-priority treatment allocation results in a better optimal plan, even when it sacrifices triage allocation and high-priority treatment allocation. This is specific to the ShakeOut Scenario, where the number of low-priority casualties far exceeds that of high-priority casualties. Even then, responders can generate plans that give greater preference to high-priority casualties by setting $q_1 \gg q_0$.

A limitation of our model is that it tends to be disadvantageous to self-evacuees. When self-evacuee destinations receive self-evacuees, they are unaware of the demand points from which these self-evacuees originate. Therefore, the time self-evacuees have spent travelling from their demand points to these facilities is unaccounted for. This is, in general, a worrying issue in disaster casualty response planning. A way to tackle this issue is to have emergency preparedness directives that: 1) guide people in their self-evacuation movements and 2) encourage training of emergency responders to take control of disaster situations quicker (before people have time to self-evacuate).

Chapter 5

Conclusions

In this chapter, we summarize our findings in this thesis and use the work done as a platform for future research suggestions.

5.1 Summary

In this thesis, we aim to build effective tactical location-allocation plans for casualty response during CHEs. These events are so large in scale that they overwhelm healthcare capabilities, even when the healthcare system triggers contingency capacity surges that are usually sufficient to deal with regular disasters. Effective tactical response plans for such events need to address three major issues, namely surge capacity, effective treatment allocation, and self-evacuee movement. We use ACFs to provide community-based surge solutions and incorporate triage and the movement of self-evacuees in an original general framework for effective treatment

allocation. This framework is used to build location-allocation optimization models.

For an overview of previous work on our theme of study, we review the literature on optimization models for casualty response to disasters. We separate the review into three categories: 1) models that consider the location of ACFs, 2) models that plan for triage, and 3) models that carry out casualty response planning without ACFs or triage. We find that most models that consider only the location of ACFs are p -center-like models, whereas models that plan for triage deal primarily with dynamic treatment planning, with the possibility of casualty priorities changing over time. For models that carry out casualty response planning without ACFs or triage, we find a sizeable portion linked to pandemic and bioterrorism incidents.

We first apply the framework to build a deterministic location-allocation model for casualty response following catastrophic radiological incidents, which are events whereby the release of radioactive material leads to significant consequences to people, the environment, and facilities. We characterize radiation effects using three zones: the inner zone, the intermediate zone, and the outer zone. The inner zone contains only casualties with severe ARS casualties/blast injuries and mild ARS. The intermediate and outer zones contain casualties with mild ARS and light external contamination. Severe ARS/blast injuries require specialized treatment at specialized hospitals whereas mild ARS casualties are treated at nonspecialized hospitals. ACFs of two types, type I and type II, are used for casualty triage and for the treatment of light external contamination. Type II ACFs are reserved for inner zone casualties to ensure their rapid transition to hospitals and type I ACFs

deal with the other two zones. In this setting, self-evacuees originate from the intermediate and outer zones and move to the existing ACFs or hospitals nearest to them for treatment. These self-evacuees can overwhelm facilities, in which case excess self-evacuees are transferred out to ACFs with available capacity for triage.

Using the above-described framework, we formulate a mixed integer linear programming model to locate ACFs, allocate casualties for triage, and allocate triaged casualties for treatment to minimize the total weighted casualty transportation time. The addition of weights enables casualty prioritization based on affliction seriousness. We apply the model to two RDD case studies based on U.S. Department of Homeland Security's NPS 11. From the optimal plans, we observe the importance of ACFs in adding surge capacity and improving efficiency. We draw further insights via sensitivity analyses on the budget of type I ACFs and nonspecialized medical center triage capacities. The various response planning rules of thumb we come up with are: 1) If excess resources are available, limit the budget of type I ACFs to a certain value and reassign the remaining resources to other duties, such as search-and-rescue or capacity building at new or existing facilities, 2) if total resources are fixed, find the right balance between the choice of more locations and the increasing transportation times due to smaller capacities, 3) if a hospital is far from radiation zones, not close to any facility that receives self-evacuees, and is not expected to receive any self-evacuees, use most of its resources for treatment, and 4) for an overwhelmed medical center, allocate a triage capacity where it can treat all its triaged casualties.

With the above model as a basis, we incorporate data uncertainties and the time

sensitivity of information to create a location-allocation model for general CHE situations. Data uncertainties are characterized using discrete probabilistically distributed scenarios. The model uses a three-stage stochastic programming methodology. In the first stage, it locates ACFs prior to any scenario realization. In the second stage, it allocates casualties to ACFs or hospitals for triage, given initial damage scenarios involving uncertainties in casualty demands, transportation times, and numbers of self-evacuees. Allocation decisions involve sending non self-evacuees to nearby ACFs or hospitals for triage and moving excess self-evacuees from hospitals to other possible triage destinations. In the third stage, it allocates triaged casualties to ACFs or hospitals for treatment, given triage scenarios and initial damage scenarios. After triage, casualties are separated into either low-priority or high-priority, the ratio of these two depending on triage scenarios. Low-priority casualties are treated at ACFs, whereas high-priority casualties need hospital capabilities. Our methodology allows for robust location of ACFs and allocation decisions that are flexible to new information becoming available with time.

Our model solution times increase exponentially with the number of binary location variables. Large solution times are unacceptable in disaster response, which is why we propose an algorithm based on Benders' decomposition to obtain good solutions fast. We first formulate a modified master model to reduce feasibility cuts. Furthermore, we incorporate flow cover inequalities that tighten the polyhedron of the linear relaxation of our model, leading to stronger cuts. Going further, the algorithm uses perturbed dual subproblem formulations to reduce the effect of

degeneracy via maximal nondominated cuts. We implement the model and algorithm in the case study of an earthquake situation in the southernmost part of the San Andreas Fault. Our algorithm proves to be very effective, compared to classical Benders' decomposition, generating better solutions and reaching optimality in shorter times. For further insights on the model, we perform sensitivity analyses on ACF budget and capacities, as well as hospital capacities. We show that increasing the budget of ACFs improves the plan, but with decreasing marginal improvement. On hospitals, we find that when excess capacity is available at a hospital, it is better to convert it into a combination of triage and low-priority treatment. The best solution for a hospital seems occur at the point where its treatment capacity is sufficient to attend to all its triaged low-priority casualties.

5.2 Future Work

Casualty demands and priorities can change with time. For instance, a casualty initially labelled as low-priority can turn into high-priority if he/she does not receive treatment within a time limit. That time limit is of course dependent on many factors, such as the nature of the injury and the characteristics of the casualty (for example, age and health). In this thesis, we consider only static models. By using dynamic ones, we can take into account the evolution of casualty characteristics with time, as well as mobile ACFs and new casualties discovered with time.

Moreover, our tactical plans can open doors to future research on operational planning. By operational planning, we mean more detailed studies on how plans ought to be carried out. This involves various activities such as manpower planning and scheduling, vehicle routing, ambulance positioning, queueing of casualties in facilities, and facility-specific capacity planning. Another important operational detail concerns how ACFs are set up. Our model makes no mention of where resources for ACFs are coming from and how they are set up.

Having a central planner that generates an optimal plan and carries it out is idealistic, considering that, in reality, many planners have their own plans and are sometimes even competing for resources to carry out their own plans in the best possible way. It would be useful to adopt a game-theoretic approach to study how the interaction would unfold among various planners with different, sometimes conflicting, objectives. And this can lead to suggestions on how best to manage resources amid the competition.

Bibliography

- [1] EM-DAT, The International Disaster Database, Center for Research on the Epidemiology of Disasters. Natural disaster trends. 2013. URL <http://www.emdat.be/natural-disasters-trends>. Accessed: 27th June 2013.
- [2] E. Toner, R. E. Waldhorn, C. Franco, A. Norwood, B. Courtney, K. Rambhia, M. Watson, and T. V. Inglesby. The next challenge in health-care preparedness: catastrophic health events. Center for Biosecurity of UPMC. 2010. URL <http://www.upmc-biosecurity.org/website/resources/publications/2010/pdf/2010-01-29-prepreport.pdf>. Accessed: 15th September 2012.
- [3] D. Hanfling. Equipment, supplies, and pharmaceuticals: How much might it cost to achieve basic surge capacity? *Acad. Emerg. Med.*, 13(11):1232–1237, 2006.
- [4] United States House of Representative Committee on Oversight and Government Reform. Hospital emergency surge capacity: not ready for the “predictable surprise”. 2008. URL <http://cdrp.net/amts/docs/HospitalEmergencySurgeCapacityReport.pdf>. Accessed: 27th June 2013.

-
- [5] J. L. Hick, D. Hanfling, J. L. Burstein, C. DeAtley, D. Barbisch, G. M. Bogdan, and S. Cantrill. Health care facility and community strategies for patient care surge capacity. *Ann. Emerg. Med.*, 44(3):253–261, 2004.
- [6] C. Lam, R. Waldhorn, E. Toner, T. V. Inglesby, and T. O’Toole. The prospect of using alternative medical care facilities in an influenza pandemic. *Biosecur. Bioterror.*, 4(4):384–390, 2006.
- [7] S. Phillips. Current status of surge research. *Acad. Emerg. Med.*, 13(11):1103–1108, 2006.
- [8] J. L. Hick, J. A. Barbera, and G. D. Kelen. Refining surge capacity: conventional, contingency, and crisis capacity. *Disaster Med. Public Health Preparedness*, 3(1):59–67, 2009.
- [9] K. V. Iserson and J. C. Moskop. Triage in medicine, part I: concept, history, and types. *Ann. Emerg. Med.*, 49(3):275–281, 2007.
- [10] E. R. Frykberg. Principles of mass casualty management following terrorist disasters. *Ann. Surg.*, 239(3):319–321, 2004.
- [11] N. A. Bostick, I. Subbarao, F. M. Burkle Jr, E. B. Hsu, J. H. Armstrong, and J. J. James. Disaster triage systems for large-scale catastrophic events. *Disaster Med. Public Health Preparedness*, 2(1):35–39, 2008.
- [12] E. B. Lerner, R. B. Schwartz, P. L. Coule, E. S. Weinstein, D. C. Cone, R. C. Hunt, S. M. Sasser, J. M. Liu, N. G. Nudell, I. S. Wedmore, J. Hammond, E. M. Bulger, J. P. Salomone, T. L. Sanddal, D. Markenson, and R. E. O’Connor. Mass casualty triage: an evaluation of the data and development

- of a proposed national guideline. *Disaster Med. Public Health Preparedness*, 2(1):25–34, 2008.
- [13] J. W. Runge and B. R. Buddemeier. Explosions and radioactive material: a primer for responders. *Prehosp. Emerg. Care*, 13(4):407–419, 2009.
- [14] J. G. Cushman, H. L. Pachter, and H. L. Beaton. Two New York City hospitals’ surgical response to the September 11, 2001, terrorist attack in New York City. *J. Trauma*, 54(1):147–155, 2003.
- [15] J. P. G. de Ceballos, F. Turégano-Fuentes, D. Perez-Diaz, M. Sanz-Sanchez, C. Martin-Llorente, and J. E. Guerrero-Sanz. 11 March 2004: The terrorist bomb explosions in Madrid, Spain – an analysis of the logistics, injuries sustained and clinical management of casualties treated at the closest hospital. *Crit. Care*, 9(1):104–111, 2005.
- [16] C. J. Aylwin, T. C. König, N. W. Brennan, P. J. Shirley, G. Davies, M. S. Walsh, and K. Brohi. Reduction in critical mortality in urban mass casualty incidents: analysis of triage, surge, and resource use after the London bombings on July 7, 2005. *Lancet*, 368(9554):2219–2225, 2006.
- [17] R. C. Larson, M. D. Metzger, and M. F. Cahn. Responding to emergencies: lessons learned and the need for analysis. *Interfaces*, 36(6):486–501, 2006.
- [18] B. Courtney, E. Toner, R. Waldhorn, C. Franco, K. Rambhia, A. Norwood, T. V. Inglesby, and T. O’Toole. Healthcare coalitions: the new foundation for national healthcare preparedness and response for catastrophic health emergencies. *Biosecur. Bioterror.*, 7(2):153–163, 2009.

- [19] L. Cooper. Location-allocation problems. *Oper. Res.*, 11(3):331–343, 1963.
- [20] M. L. Brandeau and S. S. Chiu. An overview of representative problems in location research. *Management Sci.*, 35(6):645–674, 1989.
- [21] H. D. Sherali, T. B. Carter, and A. G. Hobeika. A location-allocation model and algorithm for evacuation planning under hurricane/flood conditions. *Transport. Res. B-Meth.*, 25(6):439 – 452, 1991.
- [22] A. M. Campbell and P. C. Jones. Prepositioning supplies in preparation for disasters. *Eur. J. Oper. Res.*, 209(2):156 – 165, 2011.
- [23] S. Rahman and D. K. Smith. Use of location-allocation models in health service development planning in developing nations. *Eur. J. Oper. Res.*, 123(3):437 – 452, 2000.
- [24] L. A. Zadeh. Optimality and non-scalar-valued performance criteria. *IEEE Trans. Automat. Contr.*, 8(1):59–60, 1963.
- [25] R. T. Marler and J. S. Arora. Survey of multi-objective optimization methods for engineering. *Struct. Multidisc. Optim.*, 26(6):369–395, 2004.
- [26] R. T. Marler and J. S. Arora. The weighted sum method for multi-objective optimization: new insights. *Struct. Multidisc. Optim.*, 41(6):853–862, 2010.
- [27] A. M. Caunhye, X. Nie, and S. Pokharel. Optimization models in emergency logistics: a literature review. *Socio-Econ. Plan. Sci.*, 46(1):4–13, 2012.
- [28] T. Drezner. Location of casualty collection points. *Environ. Plann. C.*, 22(6):899–912, 2004.

- [29] T. Drezner, Z. Drezner, and S. Salhi. A multi-objective heuristic approach for the casualty collection points location problem. *J. Oper. Res. Soc.*, 57(6):727–734, 2006.
- [30] H. Jia, F. Ordóñez, and M. Dessouky. A modeling framework for facility location of medical services for large-scale emergencies. *IIE Trans.*, 39(1):41–55, 2007.
- [31] W. Yi and L. Özdamar. A dynamic logistics coordination model for evacuation and support in disaster response activities. *Eur. J. Oper. Res.*, 179(3):1177–1193, 2007.
- [32] R. Huang, S. Kim, and M. B. C. Menezes. Facility location for large-scale emergencies. *Ann. Oper. Res.*, 181(1):271–286, 2010.
- [33] Q. Gong and R. Batta. A queue-length cutoff model for a preemptive two-priority $M/M/1$ system. *SIAM J. Appl. Math.*, 67(1):99–115, 2006.
- [34] N. T. Argon, S. Ziya, and R. Righter. Scheduling impatient jobs in a clearing system with insights on patients triage in mass casualty incidents. *Probab. Eng. Inform. Sc.*, 22(3):301–332, 2008.
- [35] A. Kilic, M. C. Dincer, and M. A. Gokce. Determining optimal treatment rate after a disaster. *Forthcoming in J. Oper. Res. Soc.*, 2013.
- [36] W. J. Sacco, D. M. Navin, R. K. Waddell, K. E. Fiedler, W. B. Long, and R. F. Buckman. A new resource-constrained triage method applied to victims of penetrating injury. *J. Trauma*, 63(2):316–325, 2007.

-
- [37] D. T. Wilson, G. I. Hawe, G. Coates, and R. S. Crouch. A multi-objective combinatorial model of casualty processing in major incident response. *Eur. J. Oper. Res.*, 230(3):643–655, 2013.
- [38] P. Sun, L. Yang, and F. de Véricourt. Selfish drug allocation for containing an international influenza pandemic at the onset. *Oper. Res.*, 57(6):1320–1332, 2009.
- [39] S. Wang, F. de Véricourt, and P. Sun. Decentralized resource allocation to control an epidemic: a game theoretic approach. *Math. Biosci.*, 222(1):1–12, 2009.
- [40] G. S. Parnell, C. M. Smith, and F. I. Moxley. Intelligent adversary risk analysis: a bioterrorism risk management model. *Risk Anal.*, 30(1):32–48, 2010.
- [41] R. Patel, I. M. Longini, and M. E. Halloran. Finding optimal vaccination strategies for pandemic influenza using genetic algorithms. *J. Theor. Biol.*, 234(2):201–212, 2005.
- [42] N. P. Rachaniotis, T. K. Dasaklis, and C. P. Pappis. A deterministic resource scheduling model in epidemic control: a case study. *Eur. J. Oper. Res.*, 216(1):225–231, 2012.
- [43] G. S. Zaric, D. M. Bravata, J. E. Cleophas-Holty, K. M. McDonald, D. K. Owens, and M. L. Brandeau. Modeling the logistics of response to anthrax bioterrorism. *Med. Decis. Making*, 28(3):332–350, 2008.

-
- [44] E. K. Lee, S. Maheshwary, J. Mason, and W. Glisson. Large-scale dispensing for emergency response to bioterrorism and infectious-disease outbreak. *Interfaces*, 36(6):591–607, 2006.
- [45] Z. Shen, M. M. Dessouky, and F. Ordóñez. A two-stage vehicle routing model for large-scale bioterrorism emergencies. *Networks*, 54(4):255–269, 2009.
- [46] Q. Gong and R. Batta. Allocation and reallocation of ambulances to casualty clusters in a disaster relief operation. *IIE Trans.*, 39(1):27–39, 2007.
- [47] A. Jotshi, Q. Gong, and R. Batta. Dispatching and routing of emergency vehicles in disaster mitigation using data fusion. *Socio-Econ. Plan. Sci.*, 43(1):1–24, 2009.
- [48] G. Zayas-Cabán, M. E. Lewis, M. Olson, and S. Schmitz. Emergency medical service allocation in response to large-scale events. *IIE Trans. Healthc. Syst. Eng.*, 3(1):57–68, 2013.
- [49] F. Fiedrich, F. Gehbauer, and U. Rickers. Optimized resource allocation for emergency response after earthquake disasters. *Safety Sci.*, 35(1):41–57, 2000.
- [50] G. Barbarosoğlu, L. Özdamar, and A. Çevik. An interactive approach for hierarchical analysis of helicopter logistics in disaster relief operations. *Eur. J. Oper. Res.*, 140(1):118–133, 2002.
- [51] H. O. Mete and Z. B. Zabinsky. Stochastic optimization of medical supply location and distribution in disaster management. *Int. J. Prod. Econ.*, 126(1):76–84, 2010.

- [52] M. Najafi, K. Eshghi, and W. Dullaert. A multi-objective robust optimization model for logistics planning in the earthquake response phase. *Transport. Res. E-Log.*, 49(1):217–249, 2013.
- [53] International Atomic Energy Agency and Nuclear Energy Agency. The International Nuclear and Radiological Event Scale user’s manual, Vienna. 2008. URL http://www-pub.iaea.org/mtcd/publications/pdf/ines-2009_web.pdf. Accessed: 15th September 2012.
- [54] F. T. Harper, S. V. Musolino, and W. B. Wentz. Realistic radiological dispersal device hazard boundaries and ramifications for early consequence management decisions. *Health Phys.*, 93(1):1–16, 2007.
- [55] L. A. Il’in and O. A. Pavlovskii. Radiological consequences of the Chernobyl accident and the measures implemented to mitigate them. *Sov. Atom. Energy*, 65(2):667–677, 1988.
- [56] The White House. Homeland Security Presidential Directive 18: medical countermeasures against weapons of mass destruction. 2007. URL <http://www.fas.org/irp/offdocs/nspd/hspd-18.html>. Accessed: 11th September 2012.
- [57] The White House. Homeland Security Presidential Directive 21: public health and medical preparedness. 2007. URL <http://www.fas.org/irp/offdocs/nspd/hspd-21.html>. Accessed: 11th September 2012.

- [58] T. Sullivan. Federal response assets for a radioactive dispersal device incident. Brookhaven National Laboratory. 2009. URL <http://www.bnl.gov/isd/documents/45492.pdf>. Accessed: 15th September 2012.
- [59] J. Medalia. “Dirty bombs”: technical background, attack prevention and response, issues for Congress. Congressional Research Service. 2011. URL <http://www.fas.org/sgp/crs/nuke/R41890.pdf>. Accessed: 15th September 2012.
- [60] F. A. Mettler and G. L. Voelz. Major radiation exposure – what to expect and how to respond. *New Engl. J. Med.*, 346(20):1554–1561, 2002.
- [61] C. N. Coleman, C. Hrdina, J. L. Bader, A. Norwood, R. Hayhurst, J. Forsha, K. Yeskey, and A. Knebel. Medical response to a radiologic/nuclear event: integrated plan from the Office of the Assistant Secretary for Preparedness and Response, Department of Health and Human Services. *Ann. Emerg. Med.*, 53(2):213–222, 2009.
- [62] A. B. Wolbarst, A. L. Wiley, J. B. Nemhauser, D. M. Christensen, and W. R. Hendee. Medical response to a major radiologic emergency: a primer for medical and public health practitioners. *Radiology*, 254(3):660–677, 2010.
- [63] C. M. Hrdina, C. N. Coleman, S. Bogucki, J. L. Bader, R. E. Hayhurst, J. D. Forsha, D. Marcozzi, K. Yeskey, and A. R. Knebel. The “RTR” medical response system for nuclear and radiological mass-casualty incidents: a functional TRiage-TReatment-TRansport medical response model. *Prehosp. Disaster Med.*, 24(3):167–178, 2009.

- [64] Office of the Assistant Secretary for Preparedness and Response, Department of Health and Human Services. Radiological dispersal device playbook, 2001. URL <http://www.phe.gov/Preparedness/planning/playbooks/rdd/Pages/intro.aspx>. Accessed: 13th September 2012.
- [65] D. M. Bravata, K. M. McDonald, H. Szeto, W. M. Smith, C. Rydzak, and D. K. Owens. A conceptual framework for evaluating information technologies and decision support systems for bioterrorism preparedness and response. *Med. Decis. Making*, 24(2):192–206, 2004.
- [66] M. L. Brandeau, J. H. McCoy, N. Hupert, J. E. Holty, and D. M. Bravata. Recommendations for modeling disaster responses in public health and medicine: a position paper of the society for medical decision making. *Med. Decis. Making*, 29(1):438–460, 2009.
- [67] Conference of Radiation Control Program Directors. Handbook for responding to a radiological dispersal device. 2006. URL <http://hps.org/homeland/documents/RDD-Handbook-ForWeb.pdf>. Accessed: 20th September 2012.
- [68] Centers for Disease Control and Prevention, U.S. Department of Health and Human Services. Acute radiation syndrome: a fact sheet for physicians. 2005. URL <http://www.bt.cdc.gov/radiation/pdf/arsphysicianfactsheet.pdf>. Accessed: 20th September 2012.
- [69] The Merck Manual. Radiation exposure and contamination. 2009. URL http://www.merckmanuals.com/professional/injuries_poisoning/

[radiation_exposure_and_contamination/radiation_exposure_and_contamination.html#v1113494](#). Accessed: 21st September 2012.

- [70] J. K. Waselenko, T. J. MacVittie, W. F. Blakely, N. Pesik, A. L. Wiley, W. E. Dickerson, H. Tsu, D. L. Confer, C. N. Coleman, T. Seed, P. Lowry, J. O. Armitage, and N. Dainiak. Medical management of the acute radiation syndrome: recommendations of the Strategic National Stockpile Radiation Working Group. *Ann. Intern. Med.*, 140(12):1037–1051, 2004.
- [71] J. T. Bushberg, K. L. Miller, M. Hartman, R. Derlet, V. Ritter, and E. M. Leidholdt. Emergency department management of radiation casualties. Radiological Emergency Medical Preparedness & Management Subcommittee of the National Health Physics Society Ad Hoc Committee on Homeland Security. 2005. URL http://oep.berkeley.edu/pdf/OtherResources/BAII_BBerg.pdf. Accessed: 20th September 2012.
- [72] K. L. Koenig, C. J. Boatright, J. A. Hancock, F. J. Denny, D. S. Teeter, C. A. Kahn, and C. H. Schultz. Health care facility-based decontamination of victims exposed to chemical, biological, and radiological materials. *Am. J. Emerg. Med.*, 26(1):71–80, 2008.
- [73] Department of Homeland Security. National planning scenarios: executive summaries, 2005. URL <http://cees.tamtu.edu/covertheborder/TOOLS/NationalPlanningSen.pdf>. Accessed: 26th September 2012.

- [74] Western Regional Climate Center, Desert Research Institute. Wind direction for 1994-2012. 2012. URL <http://www.wrcc.dri.edu/htmlfiles/westwinddir.html>. Accessed: 26th September 2012.
- [75] New York State Climate Office. The climate of New York. URL http://nysc.eas.cornell.edu/climate_of_ny.html. Accessed: 27th September 2013.
- [76] M. F. Goodchild. The aggregation problem in location-allocation. *Geogr. Anal.*, 11(3):240–255, 1979.
- [77] A. S. Fotheringham, P. J. Densham, and A. Curtis. The zone definition problem in location-allocation modeling. *Geogr. Anal.*, 27(1):60–77, 1995.
- [78] G. Trotter. *Surviving health care: a manual for patients and their families*, chapter What you need to know about disasters, pages 279–293. Cambridge University Press, New York, 2010.
- [79] Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services. Hospital surge model version 1.3. 2010. URL <http://archive.ahrq.gov/prep/hospurgemodel/description/>. Accessed: 7th November 2012.
- [80] American Hospital Association. AHA healthcare data viewer. 2012. URL <http://www.ahadataviewer.com>. Accessed: 4th October 2012.
- [81] U.S. News and World Report. Best hospitals in Los Angeles, CA. 2012. URL <http://health.usnews.com/best-hospitals/area/los-angeles-ca>. Accessed: 4th October 2012.

- [82] U.S. News and World Report. Best hospitals in New York, NY. 2013.
URL <http://health.usnews.com/best-hospitals/area/new-york-ny/specialty>. Accessed: 28th September 2013.
- [83] P. A. Maningas, M. Robison, and S. Mallonee. The EMS response to the Oklahoma City bombing. *Prehospital Disaster Med.*, 12(2):9–14, 1997.
- [84] C. Martinez and D. Gonzalez. The World Trade Center attack: doctors in the fire and police services. *Crit. Care*, 5(6):304–306, 2001.
- [85] T. C. Chan, J. Killeen, W. Griswold, and L. Lenert. Information technology and emergency medical care during disasters. *Acad. Emerg. Med.*, 11(11):1229–1236, 2004.
- [86] V. Hristidis, S. C. Chen, T. Li, S. Luis, and Y. Deng. Survey of data management and analysis in disaster situations. *J. Syst. Software*, 83(10):1701–1714, 2010.
- [87] M.T. Melo, S. Nickel, and F. Saldanha da Gama. Facility location and supply chain management - A review. *Eur. J. Oper. Res.*, 196(2):401 – 412, 2009.
- [88] T. Okumura, K. Uzuki, A. Fukuda, S. Himatsum, S. Inohara, A. Kohama, and N. Takasu. The Tokyo subway sarin attack: disaster management. *Acad. Emerg. Med.*, 5(6):613–628, 1998.
- [89] J. F. Benders. Partitioning procedures for solving mixed-variables programming problems. *Numer. Math.*, 4(1):238–252, 1962.

- [90] G. Laporte and F. V. Louveaux. The integer L-shaped method for stochastic integer programs with complete recourse. *Oper. Res. Lett.*, 13(3):133–142, 1993.
- [91] W. K. K. Haneveld and M. H. van der Vlerk. Stochastic integer programming: general models and algorithms. *Ann. Oper. Res.*, 85(1):39–57, 1999.
- [92] T. L. Magnanti and R. T. Wong. Accelerating Benders decomposition: algorithmic enhancement and model selection criteria. *Oper. Res.*, 29(3):464–484, 1981.
- [93] H. D. Sherali and B. J. Lunday. On generating maximal nondominated Benders cuts. *Ann. Oper. Res.*, 210(1):57–72, 2013.
- [94] A. M. Geoffrion and G. W. Graves. Multicommodity distribution system design by Benders decomposition. *Management Sci.*, 20(5):822–844, 1974.
- [95] M. W. Padberg, T. J. van Roy, and L. A. Wolsey. Valid linear inequalities for fixed charge problems. *Oper. Res.*, 33(4):842–861, 1985.
- [96] J. R. Birge and F. V. Louveaux. A multicut algorithm for two stage stochastic linear programs. *Eur. J. Oper. Res.*, 34(3):384–392, 1988.
- [97] L. M. Jones, R. Bernknopf, D. Cox, J. Goltz, K. Hudnut, D. Mileti, S. Perry, D. Ponti, K. Porter, M. Reichle, H. Seligson, K. Shoaf, J. Treiman, and A. Wein. The ShakeOut Scenario. U.S. Geological Survey and California Geological Survey. 2008. URL <http://pubs.usgs.gov/of/2008/1150/of2008-1150small.pdf>. Accessed: 5th July 2013.

- [98] California Emergency Medical Services Authority. Trauma centers in California. 2013. URL <http://www.emsa.ca.gov/systems/Trauma/default.asp>. Accessed: 5th July 2013.
- [99] Office of Statewide Health Planning and Development. Hospital annual utilization data. 2007. URL http://www.oshpd.ca.gov/HID/Products/Hospitals/Utilization/Text_pdf_files/Hospital/Annual_Utilization/Hospital07_util_data_FINAL.xls. Accessed: 13th September 2012.
- [100] J. Spetz, J. A. Seago, J. Coffman, E. Rosenoff, and E. O’Neil. Minimum nurse staffing ratios in California acute care hospitals. California Workforce Initiative. 2000. URL <http://www.chcf.org/~media/MEDIA%20LIBRARY%20Files/PDF/M/PDF%20MinNurseStaffingRatios.pdf>. Accessed: 5th July 2013.
- [101] California Healthcare Foundation. California hospitals: buildings, beds, and business. 2013. URL <http://www.chcf.org/~media/MEDIA%20LIBRARY%20Files/PDF/C/PDF%20CaliforniaHospitals2013.pdf>. Accessed: 5th July 2013.