Stochastic Casualty Response Planning with Multiple Classes of Patients

Pedram Farghadani-Chaharsooghi
Hossein Hashemi Doulabi
Walter Rei
Michel Gendreau

December 2023

Stochastic Casualty Response Planning with Multiple Classes of Patients

Pedram Farghadani-Chaharsooghi[1,2,*], Hossein Hashemi Doulabi[1,2], Walter Rei[1,3], Michel Gendreau[1,4]

1. Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT)
2. Department of Mechanical, Industrial and Aerospace Engineering, Concordia University
3. Département d'analytique, opérations et technologies de l'information, Université du Québec à Montréal
4. Department of Mathematical and Industrial Engineering, Polytechnique Montréal

Abstract. In this paper, we study the stochastic casualty response planning problem (CRP) in the context of providing treatments to multiple classes of patients with different types of injuries. In this general setting, both patients' demands and hospitals' bed capacity are considered uncertain. To the best of our knowledge, this is the first time that this problem is solved. We propose a novel two-stage stochastic mixed-integer programming model which, in the first stage, determines the location of the Alternative Care Facilities (ACFs) and allocates different resources, such as rescue vehicles, medical equipment, and physicians, to them. In the second stage, this model helps decide how to allocate patients with multiple injuries to either ACFs or hospitals, considering their care itineraries and available resources. Moreover, it recommends potential patient transfers between ACFs and hospitals when required. Furthermore, we introduce an alternative two-stage stochastic model that is more compact than the first. This formulation significantly reduces solution times. We also provide an equivalency proof between the two formulations. As the solution method, we develop both the L-shaped algorithm, a pure cutting-plane method tailored to our stochastic mathematical model, and the Benders Branch-and-Cut (B&C) algorithm. To further enhance the efficiency of these algorithms, we develop a wide range of acceleration techniques, including Benders dual decomposition, Lagrangian dual decomposition, a multi-cut reformulation, Pareto-optimal (PO) cuts, and the inclusion of lower bounding functional valid inequalities. We carry out extensive computational experiments demonstrating that these algorithmic enhancements dramatically improve the performance of the B&C algorithm. The average optimality gap reduces from 7898% in the simple B&C algorithm to 0.92% in the enhanced algorithm. We also present a case study from the 2011 Van earthquake in Turkey to demonstrate the applicability and efficiency of our optimization methods.

Keywords: Catastrophic health events, casualty response planning, multiple classes of patients, stochastic programming, L-shaped algorithm

_____

* Corresponding author: pedram.farghadani-chaharsooghi@mail.concordia.ca

1. Introduction

Catastrophic health events (CHEs) are natural or man-made disasters that result in large numbers of injured individuals and put healthcare systems under pressure to take emergency actions. After CHEs, injured people require immediate medical attention, and their life often depends on a quick reaction from emergency healthcare providers. As a result, efficient resource management in the initial few hours following a disaster is crucial to enhance the casualty survival rate. Casualty response planning refers to the process of collaboration between organizations to efficiently manage limited resources to help survivors in affected areas and transfer them to hospitals. There have been numerous devastating natural disasters in the past few decades, which have motivated many researchers to work on casualty response planning problems. Notably, Turkey's 2023 earthquake claimed $59,259$ lives, Haiti's 2010 earthquake resulted in a staggering $250,000$ deaths, and the 2004 South Asia tsunami took $227,898$ lives, marking these events as some of history's most devastating natural disasters.

In the context of CRP, using Alternative Care Facilities (ACFs) is one of the most common ways to diagnose an overwhelming number of casualties and treat them after a disaster. ACFs are set up as temporary healthcare facilities with the primary objective of delivering essential care during times of heightened demand, such as in the aftermath of a natural disaster. Stadiums, parks, universities, and schools are some instances of locations where ACFs can be established (Caunhye and Nie 2018). The advantages of using ACFs will not realize without a suitable location-allocation plan. For example, during the 1995 Oklahoma City bombing, emergency medical personnel set up an ACF in the form of a triage center. However, they could not assess almost half of the victims because the triage center was initially located too far from the assault site (Larson, Metzger, and Cahn 2006).

The literature on CRP can be categorized into three main groups, scientific papers that have proposed either deterministic, stochastic, or robust optimization methods to solve this type of problem. Deterministic models, while straightforward, have a notable drawback. They may yield infeasible plans or lead to unexpected shortage costs in practice due to ignoring the uncertainty of demands and hospitals' capacity. In contrast, papers in the second category offer more practical reliability by addressing uncertainty through scenarios. In the context of the CRP problem under consideration, the uncertain parameters' actual values become known gradually over time. Nevertheless, healthcare organizations cannot afford to wait until they have gathered all the relevant information before developing an effective plan for responding to casualties. This is because resources, whether renewable or not, may no longer be accessible if they are not procured in advance. Additionally, delaying the acquisition of these resources could lead to increased costs. On the other hand, postponing the operational decision-making process will result in better decisions, considering there will be less uncertainty regarding the parameter values. Consequently, we use a two-stage stochastic model.

A key advantage of robust models is their ability to produce solutions that are robust to worst-case scenarios. Nevertheless, they tend to be overly risk-averse and may not perform well in practical applications. This is because they do not take into account probability distributions and optimization of the expected objective function, which are critical factors for effective response planning in dynamic healthcare scenarios. While we shortly review major works in the area, we refer interested readers to the survey by Farahani et al. (2020) for a more comprehensive literature review.

In the first category, many researchers assumed that all parameters are deterministic and given (e.g., see Apte, Heidtke, and Salmerón (2015), Caunhye, Li, and Nie (2015), Liu, Cui, and Zhang (2019), Setiawan, Liu, and French (2019)). Apte, Heidtke, and Salmerón (2015) developed a mixed-integer programming model to identify the best locations for casualty collection points and to allocate vehicles, resources, and injured people to them, in order to maximize the number of survivors. Caunhye, Li, and Nie (2015) offered a location-allocation optimization model to minimize the total transportation time, which simultaneously locates ACFs, prioritizes casualties, and considers self-evacuees effects. Setiawan, Liu, and French (2019) formulated three mathematical programming models to take into account different levels of coordination for the distribution of relief resources and the allocation of vehicles to minimize the number of victims. Liu, Cui, and Zhang (2019) considered ambulances and helicopters to move casualties. Their model determines the optimal locations for temporary healthcare services and recommends the allocation of medical services to injured people.

Although all the above models have remarkably contributed to the area of CRP, the effect of uncertainty in catastrophic health events cannot be ignored since it significantly increases the complexity of CRPs. Therefore, many researchers in the literature have proposed stochastic optimization (Caunhye and Nie 2018, Alizadeh et al. 2019, Oksuz and Satoglu 2020, Chang et al. 2023) and robust optimization models (Hu et al. 2019, Li, Zhang, and Yu 2020, Sun, Wang, and Xue 2021, Yin et al. 2023). Caunhye and Nie (2018) proposed a three-stage stochastic programming model which determines the location of ACFs in the first stage by considering self-evacuees' movement. Then, in the second and third stages, the model addresses the assignment of casualties to triage and treatment, respectively. Alizadeh et al. (2019) proposed a two-stage robust stochastic optimization model with uncertain levels of injuries. They demonstrated that incorporating uncertainty into strategic-level decision-making leads to better design outcomes for the casualty collection points. Oksuz and Satoglu (2020) developed a two-stage stochastic programming model for locating ACFs in the event of an earthquake. Their stochastic model aimed to minimize establishing ACFs and casualties' transportation costs. They incorporated innovative elements into their model by accounting for triage and α-reliability constraints. α-reliability constraints are employed to guarantee a predefined level of reliability (α) within the solutions. The reliability concept ensures that the solution meets all demands in scenarios that encompass at

least 100α percent of all possible outcomes. Most recently, Chang et al. (2023) developed a strategy to determine the locations of ACFs and efficiently allocate the limited resources of emergency medical services (EMS) for quick casualty transportation to suitable hospitals, thereby enhancing casualty survival rates. They introduced a novel tow-stage simulation-optimization algorithm to tackle the ACF location and EMS resource allocation problem, which combines binary and integer variables. The primary objective was to minimize the expected time for complete casualty delivery from the disaster area to hospitals.

In the context of robust optimization, Hu et al. (2019) paid attention to diurnal population shifts in urban areas and developed a multi-objective robust optimization model. They proposed an accelerated Benders decomposition (BD) algorithm to solve large-size problems. While all previous works covered only the primary disaster, Li, Zhang, and Yu (2020) took secondary shocks into account by offering a scenario-based robust programming model. It is worth mentioning that they assumed that all patients have the same medical needs and belong to the same category. However, CRPs are significantly more complex since patients have multiple types of injuries resulting in different categories of patients with needs for different surgical specialties.Sun, Wang, and Xue (2021) introduced a bi-objective robust optimization model for strategic and operational response within a three-level rescue chain comprising casualty clusters, temporary facilities, and general hospitals. Their approach incorporated uncertainties related to relief supply distribution, demands, and transportation times. To address this model, they used Lagrangian relaxation and e-constraint techniques. Yin et al. (2023) presented a distributionally robust model for complex disaster response, involving facility location, supply inventory, and evacuation planning. Their solution utilized a branch-and-benders-cut algorithm.

Various injuries after CHEs can significantly impact the health and survival of individuals affected by the disaster. For example, earthquakes can cause a variety of injuries, including fractures, lacerations, nerve injuries, and burns (Del Papa et al. 2019). These injuries can be life-threatening and may require immediate medical attention. Additionally, the combination of different injuries can increase the risk of infection, blood loss, and other complications. Lu-Ping et al. (2012) reported that after the Wenchaun earthquake in 2008, 45.8%, 9%, and 1.5% of injured people had two, three, and more than four injuries, respectively. Del Papa et al. (2019) found that 52% of people injured in the 2009 L'Aquila earthquake had multiple injuries. In the 2013 Lushan earthquake in China, this number was 55.7% (Kang et al. 2015). While it is common for patients to have multiple injuries after disasters, none of the studies reviewed earlier have taken this into account. This is an important point to consider in casualty response planning since it can significantly affect the allocation of medical resources and patients to ACFs. Ignoring the variety of injuries and classifying patients only as a high or low priority can result in sending them to ACFs that are not appropriately equipped for any of their specific injuries.

In addition to multiple injuries, there are several important operational details that have yet to be addressed in the literature. For example, different patients have different needs for renewable

and non-renewable resources. Therefore, we have to plan ahead for assigning physicians to ACFs accordingly. For instance, we cannot send a patient with a fracture injury to an ACF that does not have an orthopedic doctor. On the other hand, non-renewable resources such as medical supplies and drugs also play an important role in the functioning of hospitals and ACFs. For example, Kirschner wires (K-wires) are small, thin, metal pins commonly used in orthopedic surgeries. However, their demand may spike during an emergency, and it is impossible to perform orthopedic surgery without them. This highlights the importance of proper resource allocation in CRPs. Additionally, rescue vehicles are crucial for evacuating people from hazardous locations and transporting emergency personnel and supplies. A suitable vehicle allocation after a disaster helps to ensure that emergency medical services personnel and supplies reach the affected population as quickly and efficiently as possible. It can also improve the effectiveness of the emergency response and help save more lives. Although there is a rich literature on CRPs, to the best of our knowledge, no paper has addressed patients with multiple types of injuries and the above mentioned features simultaneously. The main contributions of this research are as follows.

- For the first time, we investigate the stochastic casualty response planning problem concerning treating multiple patient classes with varying injury types. Our approach also involves assigning the appropriate type of physicians, non-renewable resources, and rescue vehicles to ACFs based on each patient's specific injuries.

- We propose a primary two-stage stochastic integer programming model to formulate this problem. Then, we offer an alternative two-stage stochastic formulation, which is more compact than the first. As a result, the second model is significantly more efficient. We also present a proof establishing the equivalence between the two formulations.

- We develop an L-shaped algorithm and a B&C algorithm to solve our most efficient stochastic programming model. We significantly improve the algorithms by using the Benders dual decomposition (BDD), Lagrangian dual decomposition (LDD), a multi-cut reformulation, Pareto-optimal (PO) cuts, and the inclusion of lower bounding functional (LBF) valid inequalities. These enhancements improve the overall performance of the proposed algorithms.

- We present extensive computational results for the CRP when dealing with multiple injuries for each patient. In our computational experiments, we compare the proposed two-stage stochastic programming models and evaluate the performance of the proposed decomposition algorithms and the effect of the developed enhancements.

- We also evaluate the efficiency of our best model and algorithm in an earthquake case study from Turkey.

The structure of this paper is organized as follows. In Section 2, we describe the problem setting. Then, we present the original and improved two-stage stochastic programming models in Sections 3 and 4, respectively. Section 5 details the proposed L-shaped algorithm and explains the

various enhancement techniques we use to improve the algorithm's performance. Our extensive computational results on the CRP are presented in Section 6. We also include a case study in Section 7. Finally, we draw conclusions and discuss future directions in Section 8.

2.   Problem Description

Generally, disaster management includes four main phases: mitigation, preparedness, response, and recovery. Mitigation involves understanding vulnerabilities to potential hazards and implementing protective measures to reduce risk and increase resilience. Preparedness involves evaluating plans to save lives and coordinate response efforts before a disaster strikes. The primary goal is to achieve an adequate state of preparedness for emergency response through initiatives that strengthen the technical and managerial capabilities of governments, organizations, and communities. The response phase aims to provide immediate assistance to victims, distribute relief supplies, and evacuate affected populations to safe areas. The recovery phase occurs after the emergency and focuses primarily on activities such as debris removal, rebuilding damaged structures, and restoring critical infrastructure. Mitigation and preparedness occur before a tragedy, while response and recovery occur thereafter (Farahani et al. 2020).

The problem that we study in this paper deals with the response phase, which occurs immediately post-disaster. In our proposed models, we focus on solving a static version of the problem, where decisions are aggregated over a fixed time horizon (Caunhye and Nie 2018). The time horizon for the decision-making environment begins once the CHEs commence and continues until all affected individuals have been successfully evacuated. The nature of the disaster determines the duration of this time frame. Our study assumes that the critical period, which includes the time during which victims are transported to hospitals or ACFs, is constrained within this predetermined time horizon, which is set at 7 days. Therefore, our primary objective is to design the appropriate medical care supply to meet the overall demand observed during the post-disaster response phase. In this problem, we make three important decisions after a CHE that includes:

1. Determining the location of ACFs,
2. Allocating renewable and non-renewable medical resources such as physicians, rescue vehicles, medications, and medical equipment to ACFs, and
3. Transporting casualties to ACFs and hospitals.

The first two sets of decisions are made immediately after the disaster based on the initial information received about the intensity and geographical dispersion of the disaster. Therefore, we have considered these two sets of decisions in the first stage of our proposed two-stage stochastic programming models. However, the third set of decisions is made after having more information about uncertain parameters, including 1) the number of casualties with different injuries in various areas and 2) the active capacities of hospitals that might be different from their nominal values because of damages to the infrastructure and the normal functioning of hospitals. Therefore, we consider the transportation of casualties to ACFs and hospitals as second-stage decision variables.

We assume that emergency responders will perform a triage and primary survey in the disaster field. This primary survey follows a well-established protocol, represented by the mnemonic A.B.C.D.E., which stands for Airway, Breathing, Circulation, Disability, and Environment. The survey is designed to swiftly identify and address life-threatening conditions while determining the necessary type and extent of medical interventions. For instance, consider a scenario where a victim is alert, able to speak, oriented, and demonstrating movement in all extremities. In such cases, we can confidently conclude that their airway is clear, oxygen is reaching the brain adequately, and there is no significant central neurological injury (Sever, Vanholder, and of ISN Work Group on Recommendations for the Management of Crush Victims in Mass Disasters 2012). Please note that while we briefly discuss the triage and primary survey here, a detailed examination of this process is outside the scope of this article. Our primary focus is on outlining the care itinerary for high-priority and low-priority casualties within the healthcare network.

We categorize patients into two distinct groups: high-priority and low-priority casualties. Each patient group follows a different care itinerary within the healthcare network. High-priority patients are those with breathing problems, significant bleeding, or mental issues that prevent them from following commands. They are prioritized for immediate transportation to hospitals, where they are considered fully treated upon arrival. We assume that hospitals have limited vehicles and bed capacities but are fully equipped with the necessary renewable and non-renewable resources to treat patients with diverse injuries.

Low-priority patients, however, have a different itinerary. They may need to go through more than one ACF before reaching the status of "fully treated." Although Caunhye and Nie (2018) assumed that low-priority patients are only treatable at ACFs, we recognize the need for a more adaptable and flexible approach to disaster response. As such, we have considered a refinement to their model, allowing for low-priority patients to initially receive care at ACFs while retaining the option to refer them to either another ACF or a hospital if their condition necessitates further treatment that cannot be provided at the initial ACF. This adjustment addresses the critical issue of hospital over-occupation, particularly in the immediate aftermath of a catastrophic event, which becomes even more pronounced when accounting for casualties who independently seek treatment at hospitals without assistance from emergency responders. Their actions can disrupt response planning and strain valuable resources that could have been allocated to more severe cases (Runge and Buddemeier 2009). To implement this, first, we transfer low-priority patients to ACFs for the appropriate care and stabilization of their conditions. Their status as fully treated is granted only when the necessary physicians and resources have addressed all of their injuries. This assumption introduces complexity, as ACFs may not always have the essential physicians and medical resources available. In such cases, patients may receive partial treatment at one ACF before being transferred to another ACF or hospital for further care.

Effective management of casualty flows throughout the network is a central aspect of our problem. This includes making decisions about the assignment of patients to specific facilities

based on their facility care itinerary. This intricate process involves coordinating the movement of patients to ensure timely access to the necessary care and resources, even when ACFs may lack certain required physicians and medical supplies. Our research focuses on optimizing these facility care itineraries, addressing the diverse nature of patient conditions and resource availability.

In addition to the previously defined patient classification, we may have patients with multiple injuries, further underscoring the need for a flexible approach to care. Therefore, we may need more than one type of physician. Regarding the allocation of physicians to ACFs in the second decision set, we mainly need four types of physicians in order to cover all common injuries after disasters: (i) Orthopedic physicians for fractures, crushes, and other similar injuries, (ii) Wound specialists for lacerations, contusions, and open wounds; (iii) Neurologists for nerve injuries, and (iv) Plastic physicians or dermatologists for burns (Tanaka et al. 1999, Doocy et al. 2013). Physicians require a certain amount of resources to be operational, such as their medical kit and other essential equipment. An injured person may have any combination of the aforementioned injuries. Regarding the allocation of non-renewable resources to ACFs in the second decision set, we consider four classes of them, including drugs, general surgery instruments, orthopedic surgery instruments, and specialized surgery instruments, to treat all mentioned injuries (World Health Organization 2019). These resources must be supplied by hospitals to ACFs using their limited supply capacities.

The first-stage objective function minimizes the total operational cost of the system, including the establishment cost of ACFs and the transportation cost of non-renewable medical resources from hospitals to ACFs. Moreover, the second-stage objective function minimizes the transportation time of patients and the penalty corresponding to patients that are not fully recovered due to limited resources in ACFs and hospitals and also the limited number of vehicles.

3. Primary Two-Stage Stochastic Integer Programming Model (P1)

We propose a two-stage stochastic integer programming model for the CRP defined in Section 2. As explained earlier in Section 2, first-stage decisions involve the deployment of the main resources that will be used to provide the medical response for the considered horizon. These crucial decisions encompass:

- Opening ACFs in potential locations.
- Assigning non-renewable medical resources from hospitals to ACFs.
- Allocating vehicles of different types and physicians with various specialties to ACFs.

Then, after the realization of uncertain demands and hospital bed capacities, in the second-stage model, we determine the transportation of casualties with varying injuries from different demand points to ACFs and hospitals based on their respective itineraries. High-priority patients receive top priority, ensuring immediate transportation to hospitals, where they are promptly deemed "fully treated" upon arrival. On the other hand, low-priority patients follow a distinct itinerary, which may involve visits to more than one ACF before achieving the "fully treated" status. In the following, we present the first- and second-stage models separately.

3.1. First-Stage Model

We use the following notation to formulate the first-stage model:

Sets:

$\mathcal{I}$: The set of potential locations $i \in \mathcal{I}$ for opening ACFs.

$\mathcal{H}$: The set of existing hospitals $h \in \mathcal{H}$.

$\mathcal{M}$: The set of vehicle types $m \in \mathcal{M}$.

$\mathcal{J}$: The set of possible injuries $j \in \mathcal{J}$.

$\mathcal{J}_i$: The subset of injuries that can be treated at ACF $i \in \mathcal{I}$, where $\mathcal{J}_i \subseteq \mathcal{J}$.

$\mathcal{K}$: The set of non-renewable resources $k \in \mathcal{K}$ such as medications.

Parameters:

f_i : Fixed cost of opening an ACF at location i.

n_m^{ACF} : The available number of vehicles of type m that can be assigned to ACFs.

n_j^{spec} : The available number of physicians with specialty j.

c_{ik}^{Inv} : The inventory capacity for the non-renewable resource k for an ACF located in $i \in \mathcal{I}$.

c_{hk}^{Supply} : The maximum supply capacity of non-renewable resource k from hospital h.

ϱ_k : Procurement cost of non-renewable medical resource k to be purchased in advance for supplying ACFs by hospitals for the emergency response horizon.

First stage variables:

x_i : 1 if an ACF is opened at location i; 0 otherwise.

ϑ_{mi} : The number of vehicles of type m assigned to ACF at location i.

υ_{ij} : The number of physicians with specialty j assigned to ACF at location i (Note that ACF at location $i \in \mathcal{I}$ should be able to treat injury type j, i.e., $j \in J_i$).

r_{ihk} : The number of non-renewable resources k assigned to ACF at location i from hospital h.

Using the above notation, we formulate the first-stage model as follows:

$$\min_{\boldsymbol{x},\boldsymbol{\vartheta},\boldsymbol{v},\boldsymbol{r}} \left(\sum_{i \in \mathcal{I}} f_i x_i + \sum_{i \in \mathcal{I}} \sum_{h \in \mathcal{H}} \sum_{k \in \mathcal{K}} \varrho_k r_{ihk} + Q(\boldsymbol{x},\boldsymbol{\vartheta},\boldsymbol{v},\boldsymbol{r}) \right) \tag{1}$$

Subject to:

$$\sum_{\substack{i \in \mathcal{I}| \\ j \in \mathcal{J}_i}} \upsilon_{ij} \leq n_j^{spec} \qquad\qquad j \in \mathcal{J} \tag{2}$$

$$\sum_{i \in \mathcal{I}} \vartheta_{mi} \leq n_m^{ACF} \qquad\qquad m \in \mathcal{M} \tag{3}$$

$$\sum_{h \in \mathcal{H}} r_{ihk} \leq c_{ik}^{Inv} x_i \qquad\qquad i \in \mathcal{I}, k \in \mathcal{K} \tag{4}$$

$$\sum_{i \in \mathcal{I}} r_{ihk} \leq c_{hk}^{Supply} \qquad\qquad h \in \mathcal{H}, k \in \mathcal{K} \tag{5}$$

$$\vartheta_{mi} \leq n_m^{ACF} x_i \qquad\qquad i \in \mathcal{I}, m \in \mathcal{M} \tag{6}$$

$$\upsilon_{ij} \leq n_j^{spec} x_i \qquad\qquad i \in \mathcal{I}, j \in \mathcal{J}_i \tag{7}$$

$$x_i \in \{0,1\} \qquad\qquad i \in \mathcal{I} \tag{8}$$

$$\vartheta_{mi}, r_{ihk}, \upsilon_{ij} \in \mathbb{Z}^+ \qquad i \in \mathcal{I}, h \in \mathcal{H}, m \in \mathcal{M}, k \in \mathcal{K}, j \in \mathcal{J}_i \tag{9}$$

Objective function (1) includes the total fixed cost of opening ACFs, the procurement cost of non-renewable medical resources, and the second-stage cost $Q(\boldsymbol{x},\boldsymbol{\vartheta},\boldsymbol{v},\boldsymbol{r})$. Here, $Q(\boldsymbol{x},\boldsymbol{\vartheta},\boldsymbol{v},\boldsymbol{r})$ is

the expected second-stage cost that is evaluated in the second-stage problem presented in Section 3.2. We have $Q(\boldsymbol{x}, \boldsymbol{\vartheta}, \boldsymbol{v}, \boldsymbol{r}) = E_{\omega \in \Omega}[Q(\boldsymbol{x}, \boldsymbol{\vartheta}, \boldsymbol{v}, \boldsymbol{r}, \xi(\omega))]$ where $E_{\omega \in \Omega}[.]$ calculates the expected value over outcome $\omega \in \Omega$ and $\xi(\omega)$ is the vector of uncertain parameters including demands and hospitals' bed capacity in outcome ω. Constraint (2) indicates that the number of physicians with any specialty j assigned to ACFs must be less than or equal to the number of available physicians of the same type. Constraint (3) ensures that the total number of any vehicle of type m assigned to different ACFs cannot exceed the corresponding available number of the same type of vehicles. Constraints (4) and (5) imply the restrictions on the inventory and supply capacity levels of ACFs and hospitals, respectively. Also, according to constraints (6) and (7), we can allocate vehicles and non-renewable medical resources only to ACFs that have been established. Constraints (8) and (9) represent integrality constraints for the first-stage variables.

3.2. Second-Stage Model

To formulate the second-stage model, we use the following notation:

Sets:

$\mathcal{L}$	:	The set of demand locations $l \in \mathcal{L}$.
$\mathcal{P}$	:	The set of patient classes. This set is the power set of the injury set $\mathcal{J}$ meaning that each member of this set is a subset of the injury set. Each member of $\mathcal{P}$ represents a class of patients with different injuries. For example, $\{1, 3, 4\} \in \mathcal{P}$ represents a class of patients that have the specific injuries 1, 3, and 4.
$\mathcal{J}'_p$	:	The subset of injuries of patient type $p \in \mathcal{P}$.
Ω	:	The set of all possible outcomes (events) $\omega \in \Omega$.

Parameters:

c_i^{ACF}	:	The bed capacity of ACF at location i.
$c_h^{hospital}(\omega)$:		The bed capacity of hospital h when the outcome ω is observed.
c_m^{rescue}	:	The capacity of the vehicle of type m, measured in terms of transportation time. It represents the total time that a vehicle can spend transporting patients over the considered horizon, multiplied by the number of patients they can carry each time.
$d_{pl}(\omega)$	:	The number of patients type p realized in demand location l when the outcome ω is observed.
α_{pl}	:	The percentage of patients type p who are high-priority and must be transported from demand location l to a hospital immediately.
$d_{pl}^{low}(\omega)$	:	The number of low-priority patients type p realized in demand location l in outcome ω; i.e., $d_{pl}^{low}(\omega) = \lfloor ((1 - \alpha_{pl}) d_{pl}(\omega)) + 0.5 \rfloor \ \forall p \in \mathcal{P}, l \in \mathcal{L}, \omega \in \Omega$.
$d_{pl}^{high}(\omega)$	:	The number of high-priority patients type p realized in demand location l in outcome ω; i.e., $d_{pl}^{high}(\omega) = d_{pl}(\omega) - d_{pl}^{low}(\omega) \ \forall p \in \mathcal{P}, l \in \mathcal{L}, \omega \in \Omega$.
$n_{mh}^{hospital}$	:	The available number of vehicles of type m in hospital h.
t_j	:	The duration of treatment for a single injury type j.
τ_j	:	The available working time of a physician with specialty j in the considered horizon.
π_{ij}	:	Estimated travel time (distance) between point i and point j $(i, j \in \mathcal{J} \cup \mathcal{H} \cup \mathcal{L})$.
g_{jk}	:	The amount of non-renewable medical resource k required for the treatment of a single injury type j.

ρ_{pl}^{low} : Penalty of the unsatisfied low-priority demand type p at demand location l.

ρ_{pl}^{high} : Penalty of the unsatisfied high-priority demand type p at demand location l, where $\rho_{pl}^{low} \ll \rho_{pl}^{high}$.

$\xi(\omega)$: The vector of uncertain parameters including $c_h^{hospital}(\omega)$, $d_{pl}^{high}(\omega)$, and $d_{pl}^{low}(\omega)$ when the outcome ω is observed.

Second stage variables:

$q_{plim}(\omega)$: The number of patients type p moved by a vehicle of type m from demand location l to ACF i when the outcome ω is observed.

$u_{pp'ihm}(\omega)$: The number of patients type p', which was previously classified as type p, moved by a vehicle of type m from ACF i to hospital h when the outcome ω is observed.

$\zeta_{plhm}(\omega)$: The number of patients type p moved by a vehicle of type m from demand location l to hospital h when the outcome ω is observed.

$\eta_{pp'ii'm}(\omega)$: The number of patients type p', which were previously classified as type p, moved by a vehicle of type m from ACF i to ACF i' when the outcome ω is observed.

$\mu_{pl}^{low}(\omega)$: Quantity of unsatisfied low-priority demand type p at demand location l when the outcome ω is observed.

$\mu_{pl}^{high}(\omega)$: Quantity of unsatisfied high-priority demand type p at demand location l when the outcome ω is observed.

$\sigma_{pi}(\omega)$: Patients type p who are fully treated at ACF i when the outcome ω is observed (Note that patient class p should be fully treatable at ACF i).

Based on the given notation, we formulate the second-stage model for each outcome $\omega \in \Omega$ as follows:

$$Q(\boldsymbol{x}, \boldsymbol{\vartheta}, \boldsymbol{v}, \boldsymbol{r}, \xi(\omega)) = \min_{\substack{\boldsymbol{q}, \boldsymbol{u}, \boldsymbol{\eta}, \boldsymbol{\zeta}, \\ \boldsymbol{\mu^{low}}, \boldsymbol{\mu^{high}}, \boldsymbol{\sigma}}} \sum_{p \in \mathcal{P}} \sum_{m \in \mathcal{M}} \left(\sum_{l \in \mathcal{L}} \sum_{i \in \mathcal{I}} \pi_{li} q_{plim}(\omega) + \sum_{\substack{p' \subset p \\ p' \neq \varnothing}} \sum_{i \in \mathcal{I}} \sum_{h \in \mathcal{H}} \pi_{ih} u_{pp'ihm}(\omega) \right.$$

$$+ \sum_{l \in \mathcal{L}} \sum_{h \in \mathcal{H}} \pi_{lh} \zeta_{plhm}(\omega) + \sum_{\substack{p' \subset p \\ p' \neq \varnothing}} \sum_{i \in \mathcal{I}} \sum_{\substack{i' \in \mathcal{I} \\ i' \neq i}} \pi_{ii'} \eta_{pp'ii'm}(\omega) \Bigg)$$

$$+ \sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}} \left(\rho_{pl}^{low} \mu_{pl}^{low}(\omega) + \rho_{pl}^{high} \mu_{pl}^{high}(\omega) \right) \tag{10}$$

Subject to:

$$\sum_{i \in \mathcal{I}} \sum_{m \in \mathcal{M}} q_{plim}(\omega) + \mu_{pl}^{low}(\omega) = d_{pl}^{low}(\omega) \qquad p \in \mathcal{P}, l \in \mathcal{L} \tag{11}$$

$$\sum_{h \in \mathcal{H}} \sum_{m \in \mathcal{M}} \zeta_{plhm}(\omega) + \mu_{pl}^{high}(\omega) = d_{pl}^{high}(\omega) \qquad p \in \mathcal{P}, l \in \mathcal{L} \tag{12}$$

$$\sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}} q_{plim}(\omega) + \sum_{\substack{p' \in \mathcal{P}| \\ p \subset p', p' \neq \varnothing}} \sum_{i' \in \mathcal{I} \setminus \{i\}} \sum_{m \in \mathcal{M}} \eta_{p'pi'im}(\omega) =$$

$$\sigma_{pi}(\omega) + \sum_{\substack{p' \subset p| \\ (p \setminus p') \subseteq \mathcal{I}_i, p' \neq \varnothing}} \sum_{h \in \mathcal{H}} \sum_{m \in \mathcal{M}} u_{pp'ihm}(\omega)$$

$$+ \sum_{\substack{p' \subset p| \\ (p \setminus p') \subseteq \mathcal{I}_i, p' \neq \varnothing}} \sum_{i' \in \mathcal{I} \setminus \{i\}} \sum_{m \in \mathcal{M}} \eta_{pp'ii'm}(\omega) \qquad i \in \mathcal{I}, p \in \mathcal{P} \tag{13}$$

$$\sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}} \zeta_{plhm}(\omega) + \sum_{p \in \mathcal{P}} \sum_{\substack{p' \in \mathcal{P}| \\ p \subset p'}} \sum_{i \in \mathcal{I}} \sum_{m \in \mathcal{M}} u_{p'pihm}(\omega) \leq c_h^{hospital}(\omega) \qquad h \in \mathcal{H} \tag{14}$$

$$\sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}} q_{plim}(\omega) + \sum_{p \in \mathcal{P}} \sum_{\substack{p' \in \mathcal{P}| \\ p \subset p'}} \sum_{i' \in \mathcal{I} \setminus \{i\}} \sum_{m \in \mathcal{M}} \eta_{p'pi'im}(\omega) \leq c_i^{ACF} x_i \qquad i \in \mathcal{I} \qquad (15)$$

$$t_j \Bigg(\sum_{\substack{p \in \mathcal{P}| \\ j \in p}} \sigma_{pi}(\omega) + \sum_{\substack{p \in \mathcal{P}| \\ j \in p}} \sum_{\substack{p' \subset p| \\ j \notin p'}} \sum_{h \in \mathcal{H}} \sum_{m \in \mathcal{M}} u_{pp'ihm}(\omega)$$

$$+ \sum_{\substack{p \in \mathcal{P}| \\ j \in p}} \sum_{\substack{p' \subset p| \\ j \notin p'}} \sum_{\substack{i' \in \mathcal{I}| \\ i' \neq i}} \sum_{m \in \mathcal{M}} \eta_{pp'ii'm}(\omega) \Bigg) \leq \tau_j v_{ij} \qquad i \in \mathcal{I}, j \in \mathcal{J}_i \qquad (16)$$

$$\sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}} \pi_{lh} \zeta_{plhm}(\omega) \leq c_m^{rescue} n_{mh}^{hospital} \qquad m \in \mathcal{M}, h \in \mathcal{H} \qquad (17)$$

$$\sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}} \pi_{li} q_{plim}(\omega) + \sum_{p \in \mathcal{P}} \sum_{\substack{p' \subset p| \\ p' \neq \varnothing}} \sum_{h \in \mathcal{H}} \pi_{ih} u_{pp'ihm}(\omega)$$

$$+ \sum_{p \in \mathcal{P}} \sum_{\substack{p' \subset p| \\ p' \neq \varnothing}} \sum_{i' \in \mathcal{I} \setminus \{i\}} \pi_{ii'} \eta_{pp'ii'm}(\omega) \leq c_m^{rescue} \vartheta_{mi} \qquad m \in \mathcal{M}, i \in \mathcal{I} \qquad (18)$$

$$\sum_{p \in P} \sum_{j \in \{p \cap \mathcal{J}_i\}} \Bigg(g_{jk} \Big(\sigma_{pi}(\omega) + \sum_{\substack{p' \subset p| \\ j \notin p'}} \sum_{h \in \mathcal{H}} \sum_{m \in \mathcal{M}} u_{pp'ihm}(\omega)$$

$$+ \sum_{\substack{p' \subset p| \\ j \notin p'}} \sum_{\substack{i' \in \mathcal{I}| \\ i' \neq i}} \sum_{m \in \mathcal{M}} \eta_{pp'ii'm}(\omega) \Big) \leq \sum_{h \in \mathcal{H}} r_{ihk} \qquad i \in \mathcal{I}, k \in \mathcal{K} \qquad (19)$$

$$q_{plim}(\omega), u_{pp'ihm}(\omega), \zeta_{plhm}(\omega),$$

$$\eta_{pp'ii'm}(\omega), \mu_{pl}^{low}(\omega), \mu_{pl}^{high}(\omega), \sigma_{pi}(\omega) \geq 0 \qquad i \in \mathcal{I}, j \in \mathcal{J}_i, m \in \mathcal{M},$$

$$h \in \mathcal{H}, k \in \mathcal{K}, p \in \mathcal{P}, l \in \mathcal{L} \qquad (20)$$

The objective function (10) represents the second-stage cost when the outcome ω is observed, including 1) the transportation time of casualties between demand locations, ACFs, and hospitals and 2) the penalties corresponding to not serving some patients due to limited resources. Constraints (11) and (12) allocate low- and high-priority patients to ACFs and hospitals, respectively. In these constraints, unmet patient demands may occur, leading to penalties in the objective function.

Constraint (13) indicates that the number of patients type p admitted to ACF i should match the combined count of patients who have either fully recovered within ACF i or partially recovered at ACF i before being transferred, with a different status $p' \subset p$, to hospitals or other ACFs for additional treatment. The left-hand side of this constraint computes the number of patients type p entering ACF i from demand points (i.e., $\sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}} q_{plim}(\omega)$) and other ACFs (i.e., $\sum_{\substack{p' \in \mathcal{P}| \\ p \subset p', p' \neq \varnothing}} \sum_{i' \in \mathcal{I} \setminus \{i\}} \sum_{m \in \mathcal{M}} \eta_{p'pi'im}(\omega)$). On the other hand, the right-hand side of this constraint computes the total number of patients that leave ACF i as fully-recovered patients (i.e., $\sigma_{pi}(\omega)$), or are transferred to hospitals (i.e., $\sum_{\substack{p' \subset p| \\ (p \setminus p') \subseteq \mathcal{J}_i, p' \neq \varnothing}} \sum_{h \in \mathcal{H}} \sum_{m \in \mathcal{M}} u_{pp'ihm}(\omega)$) or other ACFs (i.e., $\sum_{\substack{p' \subset p| \\ (p \setminus p') \subseteq \mathcal{J}_i, p' \neq \varnothing}} \sum_{i' \in \mathcal{I} \setminus \{i\}} \sum_{m \in \mathcal{M}} \eta_{pp'ii'm}(\omega)$) while having a different status $p' \subset p$. The last two terms on the right-hand side of this constraint ensure that those patients being transferred to

hospitals or other ACFs are already treated for at least one of their injuries in ACF i. That is why type-p patients that are not fully treated at ACF i are transferred to hospitals or other ACFs as a new patient class $p' \subset p$.

Constraints (14) and (15) ensure that the bed capacity of hospitals and ACFs is respected. Constraint (16) guarantees that the required treatment time for injury type j at ACF i does not exceed the available time of the corresponding physician at that ACF. Constraints on the vehicle capacity in hospitals and ACFs are formulated by (17) and (18), respectively. On the left-hand side of these constraints, we have computed the total transportation time of patients imposed on the vehicles of ACFs and hospitals, while on the right-hand side, we have the available capacity. Constraint (19) states that the total non-renewable medical resource type k used at ACF i for treating patients must be less than or equal to the amount of this resource supplied by hospitals. Constraint (20) sets the nonnegativity of second-stage decision variables.

Later in our computational results, we will demonstrate a crucial issue with the previously defined second-stage model (P1). Unfortunately, it becomes computationally intractable for realistic-size problems due to its substantial number of variables. Consequently, in the upcoming section, we present a novel and enhanced mathematical model, named model (P2), which significantly outperforms model (P1). In support of this claim, we have established Theorem 1 in the online Appendix EC.1, showing the equivalence between these two mathematical models.

4.   Improved Two-Stage Stochastic Integer Programming Model (P2)

In this section, we present an enhanced version of the model (P1) achieved by reducing the dimensionality of the original model, particularly by minimizing the number of decision variables in the second stage. This reduction is accomplished through the manipulation of patient indices within the variables $u_{pp'ihm}(\omega)$ and $\eta_{pp'ii'm}(\omega)$. Sets, parameters, first-stage variables, and a part of second-stage variables remain unchanged. We define the new second-stage variables used in the model (P2) as follows:

New second-stage variables:

$u_{pihm}(\omega)$:   The number of patients type p moved by a vehicle of type m from ACF i to hospital h when the outcome ω is observed. Here, p refers to the state of the patient while being transferred to the hospital, not its primary state at the ACF.

$\eta_{pii'm}(\omega)$:   The number of patients type p moved by a vehicle of type m from ACF i to ACF i' when the outcome ω is observed. Here, p refers to the state of the patient while being transferred to ACF i', not its primary state at ACF i.

$\sigma_{pji}(\omega)$:   The number of patients type p at ACF i that are first treated for their injury j when the outcome ω is observed. These patients could be treated for other injuries in the same ACF, too, but injury j is the first injury type that they are treated for. This variable is the main key for formulating the improved model presented in this section.

The first-stage model (1)-(9) remain unchanged. In the following, we only present the improved second-stage model.

$$Q(\boldsymbol{x}, \boldsymbol{\vartheta}, \boldsymbol{v}, \boldsymbol{r}, \xi(\omega)) =$$

$$\min_{\substack{\boldsymbol{q}, \boldsymbol{u}, \boldsymbol{\eta}, \boldsymbol{\zeta}, \\ \boldsymbol{\mu}^{low}, \boldsymbol{\mu}^{high}}} \sum_{p \in \mathcal{P}} \sum_{m \in \mathcal{M}} \left(\sum_{l \in \mathcal{L}} \sum_{i \in \mathcal{I}} \pi_{li} q_{plim}(\omega) + \sum_{i \in \mathcal{I}} \sum_{h \in \mathcal{H}} \pi_{ih} u_{pihm}(\omega) + \sum_{l \in \mathcal{L}} \sum_{h \in \mathcal{H}} \pi_{lh} \zeta_{plhm}(\omega) \right.$$

$$\left. + \sum_{i \in \mathcal{I}} \sum_{\substack{i' \in \mathcal{I} \\ i' \neq i}} \pi_{ii'} \eta_{pii'm}(\omega) \right) + \sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}} \left(\rho_{pl}^{low} \mu_{pl}^{low}(\omega) + \rho_{pl}^{high} \mu_{pl}^{high}(\omega) \right) \tag{21}$$

Subject to:

$$\sum_{i \in \mathcal{I}} \sum_{m \in \mathcal{M}} q_{plim}(\omega) + \mu_{pl}^{low}(\omega) = d_{pl}^{low}(\omega) \qquad\qquad p \in \mathcal{P}, l \in \mathcal{L} \tag{22}$$

$$\sum_{h \in \mathcal{H}} \sum_{m \in \mathcal{M}} \zeta_{plhm}(\omega) + \mu_{pl}^{high}(\omega) = d_{pl}^{high}(\omega) \qquad\qquad p \in \mathcal{P}, l \in \mathcal{L} \tag{23}$$

$$\sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}} q_{plim}(\omega) + \sum_{i' \in \mathcal{I} \setminus \{i\}} \sum_{m \in \mathcal{M}} \eta_{pi'im}(\omega) + \sum_{j \in \mathcal{J}_i} \sum_{\substack{p' \in \mathcal{P} | \\ p' \setminus \{j\} = p \\ p' \setminus \{j\} \neq \varnothing}} \sigma_{p'ji}(\omega)$$

$$= \sum_{h \in \mathcal{H}} \sum_{m \in \mathcal{M}} u_{pihm}(\omega) + \sum_{i' \in \mathcal{I} \setminus \{i\}} \sum_{m \in \mathcal{M}} \eta_{pii'm}(\omega) + \sum_{j \in (\mathcal{J}_i \cap p)} \sigma_{pji}(\omega) \qquad i \in \mathcal{I}, p \in \mathcal{P} \tag{24}$$

$$\sum_{j \in (\mathcal{J}_i \cap p)} \sigma_{pji}(\omega) \geq \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}} q_{plim}(\omega) + \sum_{i' \in \mathcal{I} \setminus \{i\}} \sum_{m \in \mathcal{M}} \eta_{pi'im}(\omega) \qquad i \in \mathcal{I}, p \in \mathcal{P} \tag{25}$$

$$\sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}} \zeta_{plhm}(\omega) + \sum_{p \in \mathcal{P}} \sum_{i \in \mathcal{I}} \sum_{m \in \mathcal{M}} u_{pihm}(\omega) \leq c_h^{hospital}(\omega) \qquad h \in \mathcal{H} \tag{26}$$

$$\sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}} q_{plim}(\omega) + \sum_{p \in \mathcal{P}} \sum_{i' \in \mathcal{I} \setminus \{i\}} \sum_{m \in \mathcal{M}} \eta_{pi'im}(\omega) \leq c_i^{ACF} x_i \qquad i \in \mathcal{I} \tag{27}$$

$$\sum_{\substack{p \in \mathcal{P} | \\ j \in p}} t_j \sigma_{pji}(\omega) \leq \tau_j \upsilon_{ij} \qquad i \in \mathcal{I}, j \in \mathcal{J}_i \tag{28}$$

$$\sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}} \pi_{lh} \zeta_{plhm}(\omega) \leq c_m^{rescue} n_{mh}^{hospital} \qquad m \in \mathcal{M}, h \in \mathcal{H} \tag{29}$$

$$\sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}} \pi_{li} q_{plim}(\omega) + \sum_{p \in \mathcal{P}} \sum_{h \in \mathcal{H}} \pi_{ih} u_{pihm}(\omega)$$

$$+ \sum_{p \in \mathcal{P}} \sum_{i' \in \mathcal{I} \setminus \{i\}} \pi_{ii'} \eta_{pii'm}(\omega) \leq c_m^{rescue} \vartheta_{mi} \qquad m \in \mathcal{M}, i \in \mathcal{I} \tag{30}$$

$$\sum_{p \in \mathcal{P}} \sum_{j \in \{p \cap \mathcal{J}_i\}} g_{jk} \sigma_{pji}(\omega) \leq \sum_{h \in \mathcal{H}} r_{ihk} \qquad i \in \mathcal{I}, k \in \mathcal{K} \tag{31}$$

$$q_{plim}(\omega), u_{pihm}(\omega), \zeta_{plhm}(\omega), \eta_{pii'm}(\omega),$$

$$\mu_{pl}^{low}(\omega), \mu_{pl}^{high}(\omega), \sigma_{pji}(\omega) \geq 0 \qquad p \in \mathcal{P}, l \in \mathcal{L}, i \in \mathcal{I}, m \in \mathcal{M}, h \in \mathcal{H},$$

$$k \in \mathcal{K}, j \in \mathcal{J}_i, i' \in \mathcal{I}, i' \neq i \tag{32}$$

The objective function (21) minimizes transportation times and penalties corresponding to not serving patients. Constraint (22) ensures that low-priority patients are allocated to ACFs, and constraint (23) assigns high-priority patients to the hospitals. Constraint (24), together with variables $\sigma_{pji}(\omega)$, are the key components of this model that make it possible to formulate the revised model with significantly fewer variables compared to the primary model (P1). This constraint

works as a balance constraint in the improved model. The left-hand side of constraint (24) computes the number of patients of type p in an ACF that are either received from demand locations (i.e., $\sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}} q_{plim}(\omega)$) and other ACFs (i.e., $\sum_{i' \in \mathcal{I} \setminus \{i\}} \sum_{m \in \mathcal{M}} \eta_{pi'im}(\omega)$), or are internally generated in the same ACF from treating the first injuries of other patients of types $p' \in \mathcal{P}|_{p' \setminus \{j\}=p}$, for $j \in \mathcal{J}_i$. The other side of this constraint either transfers these type p patients to hospitals (i.e., $\sum_{h \in \mathcal{H}} \sum_{m \in \mathcal{M}} u_{pihm}(\omega)$) and other ACFs (i.e., $\sum_{i' \in \mathcal{I} \setminus \{i\}} \sum_{m \in \mathcal{M}} \eta_{pii'm}(\omega)$), or treats them for another injury (i.e., $\sum_{j \in (\mathcal{J}_i \cap p)} \sigma_{pji}(\omega)$).

Constraint (25) indicates that all patients type p sent to ACF i must be treated for at least one injury. The hospital and ACF bed capacities are respected by constraints (26) and (27), respectively. Physicians of different specialties have limited work time during the planning horizon, which is considered by constraint (28). Constraints (29) and (30) formulate the rescue vehicle capacities for hospitals and ACFs, respectively. Constraint (31) indicates the restriction on the availability of non-renewable medical resources. Constraint (32) declares the nonnegativity of second-stage decision variables.

THEOREM 1. *Models (P1) and (P2) are equivalent and, when solved, produce the same first-stage optimal solutions with the same optimal objective values.*

PROOF. We have provided the proof in Appendix EC.1. The main idea of this proof is that we show that any feasible solution of model (P1) can be projected to a feasible solution of model (P2) with the same first-stage solution and the same objective value, and vice versa. Therefore, both models will result in the same first-stage optimal solution and the same optimal objective value.

5. L-shaped Algorithm and Enhancements

In this section, we propose an L-shaped algorithm (Van Slyke and Wets 1969) with various enhancements to efficiently solve the model (P2). In the following, we first present the master problem and subproblem of the L-shaped algorithm. We formulate the master problem (MP) as follows:

$$(\text{MP}) \min_{\boldsymbol{x}, \boldsymbol{\vartheta}, \boldsymbol{v}, \boldsymbol{r}, \theta} \sum_{i \in \mathcal{I}} f_i x_i + \sum_{i \in \mathcal{I}} \sum_{h \in \mathcal{H}} \sum_{k \in \mathcal{K}} \varrho_k r_{ihk} + \theta \tag{33}$$

Subject to:

$$(2) - (9)$$

$$\theta \geq \sum_{\omega \in \Omega} \phi_\omega \theta_\omega \tag{34}$$

$$\theta_\omega \geq 0 \qquad\qquad \omega \in \Omega \tag{35}$$

In MP, θ approximates the second-stage cost and ϕ_ω represents the probability of outcome ω. Optimality cuts are iteratively added after solving subproblems (SPs) as we will discuss later.

Observe that Benders feasibility cuts are not necessary because the feasibility of the primal subproblem is ensured by constraints (22) and (23). The subproblem (SP_ω) is formulated for each outcome $\omega \in \Omega$ as follows:

$$(SP_\omega) \quad \min_{\substack{q,u,\eta,\zeta, \\ \mu^{low},\mu^{high},\sigma}} \sum_{p\in\mathcal{P}}\sum_{m\in\mathcal{M}} \left(\sum_{l\in\mathcal{L}}\sum_{i\in\mathcal{I}} \pi_{li}q_{plim}(\omega) + \sum_{i\in\mathcal{I}}\sum_{h\in\mathcal{H}} \pi_{ih}u_{pihm}(\omega) + \sum_{l\in\mathcal{L}}\sum_{h\in\mathcal{H}} \pi_{lh}\zeta_{plhm}(\omega) \right.$$

$$\left. + \sum_{i\in\mathcal{I}}\sum_{\substack{i'\in\mathcal{I} \\ i'\neq i}} \pi_{ii'}\eta_{pii'm}(\omega) \right) + \sum_{p\in\mathcal{P}}\sum_{l\in\mathcal{L}} \left(\rho_{pl}^{low}\mu_{pl}^{low}(\omega) + \rho_{pl}^{high}\mu_{pl}^{high}(\omega) \right) \tag{36}$$

Subject to:

$$(22) - (26), (29), (32)$$

$$\sum_{p\in\mathcal{P}}\sum_{l\in\mathcal{L}}\sum_{m\in\mathcal{M}} q_{plim}(\omega) + \sum_{p\in\mathcal{P}}\sum_{i'\in\mathcal{I}\setminus\{i\}}\sum_{m\in\mathcal{M}} \eta_{pi'im}(\omega) \leq c_i^{ACF}\hat{x}_i \qquad i\in\mathcal{I} \tag{37}$$

$$\sum_{\substack{p\in\mathcal{P}| \\ j\in p}} t_j\sigma_{pji}(\omega) \leq \tau_j\hat{v}_{ij} \qquad i\in\mathcal{I}, j\in\mathcal{J}_i \tag{38}$$

$$\sum_{p\in\mathcal{P}}\sum_{l\in\mathcal{L}} \pi_{li}q_{plim}(\omega) + \sum_{p\in\mathcal{P}}\sum_{h\in\mathcal{H}} \pi_{ih}u_{pihm}(\omega)$$

$$+ \sum_{p\in\mathcal{P}}\sum_{i'\in\mathcal{I}\setminus\{i\}} \pi_{ii'}\eta_{pii'm}(\omega) \leq c_m^{rescue}\hat{\vartheta}_{mi} \qquad m\in\mathcal{M}, i\in\mathcal{I} \tag{39}$$

$$\sum_{p\in\mathcal{P}}\sum_{j\in\{p\cap\mathcal{J}_i\}} g_{jk}\sigma_{pji}(\omega) \leq \sum_{h\in\mathcal{H}} \hat{r}_{ihk} \qquad i\in\mathcal{I}, k\in\mathcal{K} \tag{40}$$

In the SP, $\hat{x}_i$, $\hat{v}_{ij}$, $\hat{\vartheta}_{mi}$, and $\hat{r}_{ihk}$ are constants that are set to the values of their corresponding variables obtained by the MP. In the remainder of this section, we propose the following eight types of enhancements to accelerate the convergence of our L-shaped algorithm:

- Multiple cuts strategy (Martins de Sá et al. 2015)
- Warm-up strategy (McDaniel and Devine 1977)
- Lower bounding functional valid inequalities (Hashemi Doulabi, Pesant, and Rousseau 2020)
- Strong (Pareto-optimal) cuts (Sherali and Lunday 2013)
- Benders dual decomposition (Rahmaniani et al. 2020)
- Lagrangian dual decomposition (Zou, Ahmed, and Sun 2019)
- Benders B&C scheme (Martins de Sá et al. 2015)
- Strengthened cut generation strategies

5.1. Multiple Cuts and Warm-up Strategies

In this section, we discuss two enhancements used to improve our L-shaped algorithm. The first enhancement involves incorporating multiple Benders cuts into the MP at each iteration. There are two strategies to add optimality cuts. In the first strategy, we add the following aggregated cut after solving all SPs:

$$\theta \geq \sum_{\omega\in\Omega} \phi_\omega \left(\sum_{p\in\mathcal{P}}\sum_{l\in\mathcal{L}} \left(d_{pl}^{low}(\omega)\bar{w}_{\omega pl}^{(1)} + d_{pl}^{high}(\omega)\bar{w}_{\omega pl}^{(2)} \right) - \sum_{h\in\mathcal{H}} c_h^{hospital}(\omega)\bar{w}_{\omega h}^{(5)} \right.$$

$$- \sum_{h \in \mathcal{H}} \sum_{m \in \mathcal{M}} c_m^{rescue} n_{mh}^{hospital} \bar{w}_{\omega mh}^{(9)} - \sum_{i \in \mathcal{I}} c_i^{ACF} x_i \bar{w}_{\omega i}^{(6)} - \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \tau_j v_{ij} \bar{w}_{\omega ij}^{(7)}$$

$$- \sum_{i \in \mathcal{I}} \sum_{m \in \mathcal{M}} c_m^{rescue} \vartheta_{mi} \bar{w}_{\omega im}^{(10)} - \sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} \Big(\sum_{h \in \mathcal{H}} r_{ihk} \Big) \bar{w}_{\omega ik}^{(14)} \bigg) \tag{41}$$

In constraint (41), $\bar{w}_{\omega pl}^{(1)}, \bar{w}_{\omega pl}^{(2)}, \bar{w}_{\omega h}^{(5)}, \bar{w}_{\omega i}^{(6)}, \bar{w}_{\omega ij}^{(7)}, \bar{w}_{\omega mh}^{(9)}, \bar{w}_{\omega im}^{(10)}$, and $\bar{w}_{\omega ik}^{(14)}$ are the optimal dual variables of constraints (22), (23), (26), (27), (28), (29), (30), and (31). In the second strategy, after solving each subproblem, a single optimality cut is added per outcome:

$$\theta_\omega \geq \sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}} \Big(d_{pl}^{low}(\omega) \bar{w}_{\omega pl}^{(1)} + d_{pl}^{high}(\omega) \bar{w}_{\omega pl}^{(2)} \Big) - \sum_{h \in \mathcal{H}} c_h^{hospital}(\omega) \bar{w}_{\omega h}^{(5)}$$

$$- \sum_{h \in \mathcal{H}} \sum_{m \in \mathcal{M}} c_m^{rescue} n_{mh}^{hospital} \bar{w}_{\omega mh}^{(9)} - \sum_{i \in \mathcal{I}} c_i^{ACF} x_i \bar{w}_{\omega i}^{(6)} - \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \tau_j v_{ij} \bar{w}_{\omega ij}^{(7)}$$

$$- \sum_{i \in \mathcal{I}} \sum_{m \in \mathcal{M}} c_m^{rescue} \vartheta_{mi} \bar{w}_{\omega im}^{(10)} - \sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} \Big(\sum_{h \in \mathcal{H}} r_{ihk} \Big) \bar{w}_{\omega ik}^{(14)} \qquad \omega \in \Omega \tag{42}$$

In this research, we apply the second strategy since, in the literature, it is computationally proven to be more efficient (Batun et al. 2011).

The second enhancement, the warm-up strategy, involves solving the MP as a linear program (LP) for a predefined number of iterations denoted by α^{warm}, or until the number of iterations that the gap is not improved exceeds β^{warm} (McDaniel and Devine 1977). Then, the integrality constraints are added to the MP and it is solved as a mixed integer linear programming (MILP).

5.2. Lower Bounding Functional Valid Inequalities

In the following, we present lower bounding functional valid inequalities that improve the lower bound obtained by the master problem. The main idea of LBF valid inequalities is to give the master problem a rough estimation of the second-stage cost based on the information of an average outcome. This approach is in line with the principles of partial Benders decomposition, as shown in the work of Crainic et al. (2021). Our lower bounding functional works based on the following theorem:

THEOREM 2. $\theta \geq Q(\boldsymbol{x}, \boldsymbol{\vartheta}, \boldsymbol{v}, \boldsymbol{r}, \xi(\bar{\omega}))$ is a valid inequality for the MP where $\xi(\bar{\omega})$ represents the vector of average outcome, i.e., $\xi(\bar{\omega}) = \sum_{\omega \in \Omega} \phi_\omega \xi(\omega)$.

PROOF. We have provided the proof in Appendix EC.2. The validity of this theorem relies on Jensen's Inequality (Jensen 1906).

Based on Theorem 2, we add the following constraints to MP to impose the lower bounding functional:

$$\theta \geq \sum_{p \in \mathcal{P}} \sum_{m \in \mathcal{M}} \Bigg(\sum_{l \in \mathcal{L}} \sum_{i \in \mathcal{I}} \pi_{li} q_{plim}(\bar{\omega}) + \sum_{i \in \mathcal{I}} \sum_{h \in \mathcal{H}} \pi_{ih} u_{pihm}(\bar{\omega}) + \sum_{l \in \mathcal{L}} \sum_{h \in \mathcal{H}} \pi_{lh} \zeta_{plhm}(\bar{\omega})$$

$$+ \sum_{i \in \mathcal{I}} \sum_{\substack{i' \in \mathcal{I} \\ i' \neq i}} \pi_{ii'} \eta_{pii'm}(\bar{\omega}) \Bigg) + \sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}} \Big(\rho_{pl}^{low} \mu_{pl}^{low}(\bar{\omega}) + \rho_{pl}^{high} \mu_{pl}^{high}(\bar{\omega}) \Big) \tag{43}$$

$$(22) - (32) \qquad \text{for the average outcome, i.e., } \bar{\omega}, \text{ rather than outcome } \omega \qquad (44)$$

According to the above relations, the incorporation of LBF valid inequalities requires adding a copy of all second-stage variables and their corresponding constraints for an average outcome to the MP. This utilization of LBF valid inequalities aligns with the principles of partial Benders decomposition, which allows us to generate valid inequalities based on artificial scenarios. Importantly, even if only one such scenario is produced, the best bound that can be obtained is with the average scenario, as demonstrated in Crainic et al. (2021).

5.3. Strong (Pareto-optimal) Cuts

When applying BD, it is common to find multiple optimal solutions when solving the SPs, especially when dealing with network optimization problems such as facility location problems (Magnanti and Wong 1981). All optimal solutions to the SPs can be used to produce valid cuts for the master problem. However, if the strongest cuts (i.e., Pareto-optimal cuts) are used, the number of iterations and solution time can be significantly decreased. Sherali and Lunday (2013) suggested solving a pre-emptive priority multiple-objective program. According to this method, a perturbed version of the dual subproblem is created to generate Pareto-optimal cuts, represented as model (45). The modification involves incorporating a perturbation term into the dual objective function using a goal-programming weight (μ) and core points of the feasible solution set (i.e., x_i^0, v_{ij}^0, ϑ_{mi}^0, and r_{ihk}^0). When μ takes a small value, for example 10^{-11} in our case, solving the perturbed model (45) yields an optimal solution that corresponds to a Pareto-optimal solution, as stated in Proposition 1.

$$
\begin{aligned}
\max \sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}} & \left(d_{pl}^{low}(\omega) w_{\omega pl}^{(1)} + d_{pl}^{high}(\omega) w_{\omega pl}^{(2)} \right) - \sum_{h \in \mathcal{H}} c_h^{hospital}(\omega) w_{\omega h}^{(5)} - \sum_{h \in \mathcal{H}} \sum_{m \in \mathcal{M}} c_m^{rescue} n_{mh}^{hospital} w_{\omega mh}^{(9)} \\
& - \sum_{i \in \mathcal{I}} c_i^{ACF} \hat{x}_i w_{\omega i}^{(6)} - \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \tau_j \hat{v}_{ij} w_{\omega ij}^{(7)} - \sum_{i \in \mathcal{I}} \sum_{m \in \mathcal{M}} c_m^{rescue} \hat{\vartheta}_{mi} w_{\omega im}^{(10)} - \sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} \left(\sum_{h \in \mathcal{H}} \hat{r}_{ihk} \right) w_{\omega ik}^{(14)} \\
& + \mu \left(\sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}} \left(d_{pl}^{low}(\omega) w_{\omega pl}^{(1)} + d_{pl}^{high}(\omega) w_{\omega pl}^{(2)} \right) - \sum_{h \in \mathcal{H}} c_h^{hospital}(\omega) w_{\omega h}^{(5)} - \sum_{h \in \mathcal{H}} \sum_{m \in \mathcal{M}} c_m^{rescue} n_{mh}^{hospital} w_{\omega mh}^{(9)} \right. \\
& \left. - \sum_{i \in \mathcal{I}} c_i^{ACF} x_i^0 w_{\omega i}^{(6)} - \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \tau_j v_{ij}^0 w_{\omega ij}^{(7)} - \sum_{i \in \mathcal{I}} \sum_{m \in \mathcal{M}} c_m^{rescue} \vartheta_{mi}^0 w_{\omega im}^{(10)} - \sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} \left(\sum_{h \in \mathcal{H}} r_{ihk}^0 \right) w_{\omega ik}^{(14)} \right) \qquad (45)
\end{aligned}
$$

Subject to:

Constraints for the dual form of the model (21) to (32)

PROPOSITION 1. If we define $opt(\hat{x}, \hat{\vartheta}, \hat{v}, \hat{r})$ and $sec(\hat{x}, \hat{\vartheta}, \hat{v}, \hat{r})$ as the optimal and second-optimal values of the dual sub-problem corresponding to the solution of $\hat{x}, \hat{\vartheta}, \hat{v}$, and $\hat{r}$, and $opt(x^0, \vartheta^0, v^0, r^0)$ as the optimal value of the dual sub-problem associated with x^0, v^0, ϑ^0, and r^0. Then when the goal-programming weight μ meets the condition $0 < \mu < \frac{sec(\hat{x}, \hat{\vartheta}, \hat{v}, \hat{r}) - opt(\hat{x}, \hat{\vartheta}, \hat{v}, \hat{r})}{opt(x^0, \vartheta^0, v^0, r^0)}$, an optimal solution to the perturbed model (45) will also be a Pareto-optimal solution.

PROOF. We have provided the proof in Appendix EC.3.

5.4. Benders Dual Approach

We also develop the BDD cuts proposed by Rahmaniani et al. (2020) based on Lagrangian duality. In this method, instead of substituting the values of the first-stage solution obtained from MP in the subproblem, we impose the solution to the subproblem as coupling constraints. This process involves creating a new MILP model for the subproblem, where copied variables are introduced to mirror those in the MP with equality constraints (50)-(53). Then, we price out those constraints into the objective function using the associated dual multipliers. By applying this relaxation step, the copied variables no longer equal the first-stage solutions. Consequently, we can enforce integrity constraints on these variables (e.g., constraints (56) and (57)). To develop BDD for our problem, we first replace constraints (37)-(40) in the subproblem with constraints (46)-(54).

$$\sum_{p\in\mathcal{P}}\sum_{l\in\mathcal{L}}\sum_{m\in\mathcal{M}}q_{plim}(\omega)+\sum_{p\in\mathcal{P}}\sum_{i'\in\mathcal{I}\setminus\{i\}}\sum_{m\in\mathcal{M}}\eta_{pi'im}(\omega)\leq c_i^{ACF}x_i^{Copy} \qquad \omega\in\Omega,i\in\mathcal{I} \quad (46)$$

$$\sum_{\substack{p\in\mathcal{P}|\\j\in p}}t_j\sigma_{pji}(\omega)\leq\tau_j v_{ij}^{Copy} \qquad \omega\in\Omega,i\in\mathcal{I},j\in\mathcal{J}_i \quad (47)$$

$$\sum_{p\in\mathcal{P}}\sum_{l\in\mathcal{L}}\pi_{li}q_{plim}(\omega)+\sum_{p\in\mathcal{P}}\sum_{h\in\mathcal{H}}\pi_{ih}u_{pihm}(\omega)$$
$$+\sum_{p\in\mathcal{P}}\sum_{i'\in\mathcal{I}\setminus\{i\}}\pi_{ii'}\eta_{pii'm}(\omega)\leq c_m^{rescue}\vartheta_{mi}^{Copy} \qquad \omega\in\Omega,m\in\mathcal{M},i\in\mathcal{I} \quad (48)$$

$$\sum_{p\in\mathcal{P}}\sum_{j\in\{p\cap\mathcal{J}_i\}}g_{jk}\sigma_{pji}(\omega)\leq\sum_{h\in\mathcal{H}}r_{ihk}^{Copy} \qquad \omega\in\Omega,i\in\mathcal{I},k\in\mathcal{K} \quad (49)$$

$$x_i^{Copy}(\omega)=\hat{x}_i \qquad \omega\in\Omega,i\in\mathcal{I} \quad (50)$$

$$v_{ij}^{Copy}(\omega)=\hat{v}_{ij} \qquad \omega\in\Omega,i\in\mathcal{I},j\in\mathcal{J}_i \quad (51)$$

$$\vartheta_{mi}^{Copy}(\omega)=\hat{\vartheta}_{mi} \qquad \omega\in\Omega,m\in\mathcal{M},i\in\mathcal{I} \quad (52)$$

$$r_{ihk}^{Copy}(\omega)=\hat{r}_{ihk} \qquad \omega\in\Omega,i\in\mathcal{I},h\in\mathcal{H},k\in\mathcal{K} \quad (53)$$

$$x_i^{Copy}(\omega),v_{ij}^{Copy}(\omega),\vartheta_{mi}^{Copy}(\omega),r_{ihk}^{Copy}(\omega)\in\mathbb{R}^+ \qquad \omega\in\Omega,i\in\mathcal{I},j\in\mathcal{J}_i,m\in\mathcal{M},h\in\mathcal{H},k\in\mathcal{K} \quad (54)$$

Then we let $\lambda_{i\omega}^{xCopy}$, $\lambda_{ij\omega}^{vCopy}$, $\lambda_{mi\omega}^{\vartheta Copy}$, and $\lambda_{ihk\omega}^{rCopy}$ denote the dual multiplier vector corresponding to the constraints (50)-(53), respectively, and form the corresponding Lagrangian relaxation problem as (55)-(57).

$$\max_{\substack{\lambda^{xCopy},\lambda^{vCopy},\\\lambda^{\vartheta Copy},\lambda^{rCopy}\in\mathbb{R}^n}}\min\sum_{p\in\mathcal{P}}\sum_{m\in\mathcal{M}}\left(\sum_{l\in\mathcal{L}}\sum_{i\in\mathcal{I}}\pi_{li}q_{plim}(\omega)+\sum_{i\in\mathcal{I}}\sum_{h\in\mathcal{H}}\pi_{ih}u_{pihm}(\omega)+\sum_{l\in\mathcal{L}}\sum_{h\in\mathcal{H}}\pi_{lh}\zeta_{plhm}(\omega)\right.$$

$$+\sum_{i\in\mathcal{I}}\sum_{\substack{i'\in\mathcal{I}\\i'\neq i}}\pi_{ii'}\eta_{pii'm}(\omega)\bigg)+\sum_{p\in\mathcal{P}}\sum_{l\in\mathcal{L}}\left(\rho_{pl}^{low}\mu_{pl}^{low}(\omega)+\rho_{pl}^{high}\mu_{pl}^{high}(\omega)\right)-\bigg(\lambda_{i\omega}^{xCopy}(x_i^{Copy}(\omega)-\hat{x}_i)$$

$$+\lambda_{ij\omega}^{vCopy}(v_{ij}^{Copy}(\omega)-\hat{v}_{ij})+\lambda_{mi\omega}^{\vartheta Copy}(\vartheta_{mi}^{Copy}(\omega)-\hat{\vartheta}_{mi})+\lambda_{ihk\omega}^{rCopy}(r_{ihk}^{Copy}(\omega)-\hat{r}_{ihk})\bigg) \quad (55)$$

Subject to:

$(2) - (9), (22) - (26), (29), (32), (38) - (43)$

$$x_i^{Copy}(\omega) \in \{0,1\} \qquad\qquad\qquad\qquad\qquad\qquad\qquad i \in \mathcal{I} \qquad\qquad (56)$$

$$\vartheta_{mi}^{Copy}(\omega), \upsilon_{ij}^{Copy}(\omega), r_{ihk}^{Copy}(\omega) \in \mathbb{Z}^+ \qquad\qquad i \in \mathcal{I}, m \in \mathcal{M}, h \in \mathcal{H}, k \in \mathcal{K}, j \in \mathcal{J}_i \quad (57)$$

Note that, in the model (55)-(57), we simply apply the integrality requirements on x_i^{Copy}, υ_{ij}^{Copy}, ϑ_{mi}^{Copy}, and r_{ihk}^{Copy}. Then, for any values of λ_i^{xCopy}, $\lambda_{ij}^{\upsilon Copy}$, $\lambda_{mi}^{\vartheta Copy}$, and λ_{ihk}^{rCopy}, after solving the inner minimization of the Lagrangian relaxation model (55)-(57), we can add the strengthened cut (58) to the MP. Notably, in BDD, we do not need to optimize the Lagrangian multipliers.

$$\bar{\theta}^\omega \geq \sum_{p \in \mathcal{P}} \sum_{m \in \mathcal{M}} \left(\sum_{l \in \mathcal{L}} \sum_{i \in \mathcal{I}} \pi_{li} \bar{q}_{plim}(\omega) + \sum_{i \in \mathcal{I}} \sum_{h \in \mathcal{H}} \pi_{ih} \bar{u}_{pihm}(\omega) + \sum_{l \in \mathcal{L}} \sum_{h \in \mathcal{H}} \pi_{lh} \bar{\zeta}_{plhm}(\omega) + \sum_{i \in \mathcal{I}} \sum_{\substack{i' \in \mathcal{I} \\ i' \neq i}} \pi_{ii'} \bar{\eta}_{pii'm}(\omega) \right)$$

$$+ \sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}} \left(\rho_{pl}^{low} \bar{\mu}_{pl}^{low}(\omega) + \rho_{pl}^{high} \bar{\mu}_{pl}^{high}(\omega) \right) - \left(\lambda_i^{xCopy}(x_i - \bar{x}_i^{LR}) + \lambda_{ij}^{\upsilon Copy}(\upsilon_{ij} - \bar{\upsilon}_{ij}^{LR}) \right.$$

$$\left. + \lambda_{mi}^{\vartheta Copy}(\vartheta_{mi} - \bar{\vartheta}_{mi}^{LR}) + \lambda_{ihk}^{rCopy}(r_{ihk} - \bar{r}_{ihk}^{LR}) \right) \qquad\qquad (58)$$

In the strengthened cut (58), $\bar{x}_i^{LR}$, $\bar{\upsilon}_{ij}^{LR}$, $\bar{\vartheta}_{mi}^{LR}$, and $\bar{r}_{ihk}^{LR}$ are parameters that have already been determined in the Lagrangian relaxation problem. Rahmaniani et al. (2020) showed that when the optimal solution for the MP is an integer solution, the optimality cut (58) is at most as strong as the general optimality cut (42). However, when dealing with fractional solutions in the MP, the strengthened cut can outperform (42).

We also investigated the utilization of the LDD method. In the next section, we apply the heuristic model that Rahmaniani et al. (2020) used to improve the Lagrangian multipliers' value. In this approach, in dealing with (55)-(57), we first optimize the Lagrangian multiplier values, i.e., λ_i^{xCopy}, $\lambda_{mi}^{\vartheta Copy}$, $\lambda_{ij}^{\upsilon Copy}$, λ_{ihk}^{rCopy}, through the heuristic model (59)-(60) and then generate the cut (58). We refer to the obtained cuts as Lagrangian Cut.

$$\max_{\eta^{Heuristic} \in \mathbb{R}^1, \lambda \in \mathbb{R}^n} \left(\eta^{Heuristic} - \frac{\delta}{2} \|\lambda^{(t-1)} - \lambda\|_2^2 \right) \qquad\qquad (59)$$
$$\text{Subject to:}$$

$$\eta^{Heuristic} \leq SP_{obj}^{\mathcal{V}} + (\bar{y} - \bar{z}^{\mathcal{V}})\lambda \qquad\qquad \mathcal{V} = 1, 2, \ldots, t-1 \qquad (60)$$

The objective function (59) aims to increase the amount of cut lifting minus the distance value. The distance value allows the algorithm to find a new vector λ^t close to the vector obtained from the previous iteration $\lambda^{(t-1)}$. In this model, $\bar{y}$ is the vector of fractional values corresponding to the first-stage variable fixed in the MP. The parameter δ is selected as the step size. In (60), $SP_{obj}^{\mathcal{V}}$ denotes the value of the objective function (55) without the Lagrangian penalties obtained through solving the Lagrangian relaxation model in the $\mathcal{V}^{th}$ iteration. Also, $\bar{z}^{\mathcal{V}}$ represents the value vector of the corresponding first-stage variables. After solving the heuristic model (59)-(60), we have the new Lagrangian multiplier values, build the inner minimization problem (55)-(57), and obtain the relaxed first- and second-stage variables. We then add a new constraint to (60) and solve the revised model (59)-(60) again. This process continues until either the new multiplier values show less improvement than the prior iteration or the maximum iteration limit is reached.

5.5. Benders Branch-and-Cut Scheme

In the classical BD, a single MILP problem is optimized in each iteration. This requires constructing a new branch-and-bound tree each time, leading to potential inefficiencies as previously eliminated candidate solutions are revisited. In this article, whenever we mention "L-shaped," we are actually referring to the classical BD, which is a pure cutting-plane method applied to our stochastic mathematical model. An alternative approach involves constructing a single search tree and generating constraints for both integer and fractional solutions encountered within the tree. This alternative approach leads to the same optimal solution and is commonly known as the Benders branch-and-cut scheme (Rahmaniani et al. 2017). In this article, we will refer to this alternative method as the "branch-and-cut (B&C)" algorithm. We will apply both the "L-shaped" and "B&C" algorithms and compare the results in Sections 6.3 and EC.8.

5.6. Strengthened Cut Generation Strategies

In this section, we introduce two enhanced cut generation strategies aimed at reducing computational effort, particularly when optimizing the Lagrangian dual problem for BDD and LDD cuts, as detailed in Section 5.4. These acceleration techniques, including the warm-up, BDD, and LDD, are exclusively applied at the root node of the B&C tree (where at most half of the time limit, i.e., 12 hours, is dedicated to solving the LP at the root node). However, Modifications are still needed to reduce the computational burden of generating strengthened cuts, as solving a large number of MIP subproblems at each iteration is time-consuming. This is because the method introduced by Rahmaniani et al. (2020) was meant for a single-scenario problem, while we aim to solve a stochastic problem with at least 100 outcomes. It is important to note that, apart from the root node, we use only generalized Benders cuts (42) at all other nodes in the tree.

The first strategy, which we refer to as BC_I, is inspired by the three-phase approach proposed by Rahmaniani et al. (2020). In this strategy, we first quickly obtain valid cuts by solving the LP relaxation of the MP. We do this because, during the early stages of the Benders algorithm, the solutions of the MP tend to have low quality. Subsequently, we follow an iterative approach by solving the LP relaxation of the MP and SPs while incorporating generalized Benders cuts in each iteration. We continue this process for a specific number of iterations ($\alpha^{warm} = 15$) or until the number of iterations with unimproved gaps exceeds a predefined threshold ($\beta^{warm} = 3$). Then, we create strengthened Benders cuts, as explained in Section 5.4 considering the modified version. In the modified version for BDD and LDD, when we encounter a fractional solution in the MP at the root node, we do not solve all SPs to produce strengthened cuts for every outcome. Instead, we tackle only a subset of them in each iteration. For example, in the first iteration, we solve the first $\alpha^{cluster} = 10\%$ of outcomes; in the second iteration, we solve the next $\alpha^{cluster} = 10\%$, and so on. This process is repeated for a limited number of iterations ($\alpha^{BDD} = 10$) or when the B&C gap shows no improvement over three consecutive iterations ($\beta^{BDD} = 3$).

In our second strategy, known as BC_{II}, we employ a unique approach by applying the warm-up process, BDD, and LDD consecutively but on specific outcome clusters. This means that when we encounter a fractional solution in the master problem at the root node, we handle one cluster at a time. For instance, in the first iteration, we focus on the initial $\alpha^{\text{cluster}} = 10\%$ of outcomes, utilizing the warm-up process along with the corresponding generalized Benders cuts. In the subsequent iteration, we move to the next $\alpha^{\text{cluster}} = 10\%$ of outcomes and apply strengthened Benders cuts tailored to this cluster, and so on. This process of addressing each cluster one at a time is repeated until we reach a set number of iterations ($\alpha^{BC_{II}} = 15$) or experience a failure to reduce the gap for three consecutive iterations ($\beta^{BC_{II}} = 3$). So, in essence, we employ the warm-up, BDD, and LDD techniques individually for each cluster rather than consecutively, making it a more step-by-step or "layered" approach. In BC_{II}, the heuristic model (59)-(60) is terminated either if we reach the maximum number of iterations, $\alpha^{Heuristic} = 10$, or if we fail to lift the cut by at least $\gamma^{Heuristic} = 10^{-1}$. We have also selected the following values for the remaining parameters in this model: $\delta = 10^5$ and $\mathcal{V} = 10^2$. We choose a large number for δ since we aim to remain close to what we obtained as Lagrangian multiplier values after dualizing of constraints (50)-(53).

We have also developed two improvement methods, namely Partially Relaxed Subproblems (PRS) and ϵ-optimal cuts, for both strategies (Rahmaniani et al. 2020). In the former method, when we are applying the BDD or LDD, the integrality constraints are only imposed on a subset of first-stage variables, i.e., x_i and ϑ_{mi}. In that case, we produce a weaker optimality cut compared to the Lagrangian cut due to relaxed integrality constraints on variables v_{ij} and r_{ihk}. However, it still provides a stronger cut than a general Benders cut. In the latter method, we stop solving the MIP subproblems in BDD and LDD once we reach the ϵ-optimal solutions. These solutions have objective values that are no more than $\epsilon = 10^{-2}$ units higher than the lower bound.

6.   Computational Results

We carried out the computational tests on a machine with two 2.40 GHz Intel Xeon Gold 6148 processors using C++ and IBM ILOG CPLEX Optimization Studio V22.1 software. One core was used to execute each test instance. In Section 6.1, we describe the process by which the set of CRP instances with uncertain demand and hospital bed capacity was obtained. In Section 6.2, we compare the performance of the two proposed mathematical models, and then we assess the computational efficiency of the different enhancements used in our algorithm in Sections 6.3 and EC.8. The effectiveness of each accelerator is also evaluated in Section 6.4. To respect the length limitation of the journal and given the comprehensive nature of the developed framework, we have included additional sensitivity analysis and managerial implications in the online supplement (please cf. attached).

6.1. Instance Generation

In this section, we explain the details of generating CRP instances with uncertain demand and hospital bed capacity. For generating a major part of our instances, we used the data provided by Oksuz and Satoglu (2020) and Sun, Wang, and Xue (2021). However, since we consider some details in CRPs for the very first time, we needed to generate the newly defined parameters as detailed later in this section. *Parameter Settings*: We set the number of ACFs and demand locations to $\{10, 15, 20, 25, 30\}$ and the number of hospitals to $\{10\}$. We assume there are three types of rescue vehicles with capacities of 1, 2, and 3 patients per trip (Rambha et al. 2021). We uniformly generated the demand locations, ACFs, and hospitals in a square area with a side length of 50 km. For each demand location, we generated the number of patients by a uniform distribution of $[50, 200]$.

Regarding the injuries of patients, we considered four common categories of disaster-related injuries, including fractures/crush (FC), laceration/contusions/open wounds (LCO), peripheral nerve injury (PNI), and burns (BU). The occurrence probabilities of these injuries are 63%, 33%, 2%, and 2% of injuries in disasters, respectively (Tanaka et al. 1999, Doocy et al. 2013). To determine which patient suffers from which injuries, we first determine the number of injuries for each patient utilizing an integer uniform distribution $[1, 4]$. Then, we generate the patients' injuries using the above occurrence probabilities. This results in $2^4 - 1$ categories of patients, including all possible injury subsets except the empty set. Moreover, Sun, Wang, and Xue (2021) supposed that 40% to 45% of patients are high-priority. In our study, we assume that patients with 1, 2, 3, and 4 injuries are categorized as serious patients, with a uniform distribution range of $20 - 30$, $35 - 40$, $45 - 50$, and $55 - 60$, respectively.

For the duration of treatments, we used the means and standard deviations for orthopedics, general surgeries, neurosurgery, and plastic surgeries provided by Costa (2017). The non-renewable resources are also classified into four general groups of drug kits and three other surgical kits for general, orthopedic, and specialized surgeries (World Health Organization 2019). The cost of these kits is $7.5 for drugs and $13.5 for other surgical kits (Sun, Wang, and Xue 2021). All patients require the first group of non-renewable resources, while using the other kits depends on the type of injury. We provide the rest of the parameters in the online Appendix EC.4.

6.2. Performance of Mathematical Models

As discussed in Section 3.2, we provided an improved model (P2) that is an enhanced version of the basic mathematical model (P1) for the class of CRPs with multiple injuries. In Table 1, we have compared the performance of these two mathematical models. In the instances of Table 1, we have set the number of outcomes and hospitals to 10. Each row of this table presents the average value for 5 different instances. The first two columns of Table 1 show instances' information, including the number of ACFs and the number of demand points. Under "*# Vars.*" and "*# Const.*", we have provided the number of variables and constraints of each model. Under the

"*Time (sec)*"column, we report the time required to solve the problem optimally using CPLEX directly. We have reported the three recent columns for both models (P1) and (P2). Finally, the last column reports the improvement in the solution time resulting from using model (P2) instead of model (P1), i.e., $\Delta^{\text{time}} = 100 \times \left(\frac{P1\text{'s time} - P2\text{'s time}}{P1\text{'s time}} \right)$. Table 1 shows that model (P2) significantly outperforms model (P1) in terms of solution time and is, on average, 42.56% faster. This saving in solution time is mainly a result of the reduction in the number of variables by approximately 88% in the model (P2) compared to (P1). Since model (P2) shows significantly stronger performance, we present all numerical results in the following sections based on this model.

Table 1: Comparison of Mathematical Models P1 and P2.

Problem Size		P1 Model			P2 Model			Δ^{time}
# ACF	# De.Loc.	# Vars.	# Const.	Time (sec)	# Vars.	# Const.	Time (sec)	(%)
20	10	4,056,924	8,671	34,578	403,444	11,671	18,303	47.79
	15	4,125,926	10,195	19,050	472,622	13,195	9,320	48.61
	20	4,194,924	11,669	24,422	541,428	14,669	16,326	33.33
	25	4,263,928	13,195	36,833	610,622	16,195	18,749	41.14
	30	4,332,922	14,682	21,363	679,525	17,682	10,924	44.01
25	10	5,902,912	10,074	42,414	549,268	13,824	23,610	42.89
	15	5,983,156	11,479	41,463	628,821	15,229	24,902	37.13
	20	6,063,406	12,973	35,645	709,023	16,723	19,174	43.57
	25	6,143,656	14,486	52,620	789,370	18,236	29,359	44.92
	30	6,223,908	15,997	54,769	869,701	19,747	29,810	44.93
30	10	8,086,392	11,318	54,778	716,425	15,818	26,652	51.38
	15	8,177,885	12,752	67,389	807,439	17,252	39,199	41.59
	20	8,269,391	14,327	60,052	899,490	18,827	37,193	40.38
	25	8,360,882	15,754	62,961	990,455	20,254	38,681	38.00
	30	8,452,387	17,292	63,979	1,082,231	21,792	39,074	38.80
Average		6,175,907	12,991	44,821	716,658	16,741	25,418	42.56

6.3. Performance of Algorithms

The main goal of this subsection is to assess the performance of accelerators and compare the two strengthened cut generation strategies explained in Section 5.6. In Table 2, we have provided the computational results of the Benders B&C algorithm without any other acceleration techniques explained in Section 5 (referred to as BC_0) and the Benders B&C algorithm with acceleration techniques based on the two proposed strengthened cut generation strategies (referred to as BC_I and BC_{II}). In this table, each row presents the average values for five instances. We have considered a time limit of 24 hours for solving instances with these three algorithms. In this table, under "*Nodes No.*" and "*Time (sec)*", we have respectively reported the number of explored nodes in the branch-and-bound tree and also the total solution time. Column "*Gap (%)*" computes the gap between the lower- and the upper-bound, which are presented under columns "*LB*" and "*UB*" respectively.

Since the BC_I has resulted in the best performance, we reported more information for this algorithm. "EC (%)," "PC (%)," "TC (%)," and "SC (%)" give the contribution percentage of establishment cost, procurement cost, transportation cost, and penalty cost in the total cost for the solutions by solving the stochastic model using the BC_I algorithm. "VSS (%)" also computes the value of the stochastic solution in percentage. This value is computed as follows: $VSS = 100 \times \left(\frac{EEV - UB}{EEV} \right)$, where EEV is the expected objective value of the expected value solution computed using the outcome set. To obtain EEV, we first solve a deterministic problem using the expected value for the stochastic parameters, also known as the "mean-value" problem. Then, we save the optimal variables obtained in the first stage, fix them in the stochastic model, and proceed to solve the second stage for the different outcomes.

According to Table 2, BC_0, BC_{II}, and BC_I algorithms have resulted in an average optimality gap of 563.1%, 0.4%, and 0.1%, respectively. These results show that the proposed acceleration techniques significantly contribute to improving the performance of BC_I and BC_{II} compared to the BC_0 algorithm. For the largest instances, the average optimality gap of the BC_0 algorithm has skyrocketed to 7,900%, while we have obtained 2.4% and 1.0% gaps for BC_{II} and BC_I, highlighting the accelerators' effectiveness. Both BC_I and BC_{II} algorithms performed well and found optimal solutions in 21 instance settings. In other instances, BC_I performed better than BC_{II} in 3 cases, whereas BC_{II} outperformed BC_I on a single instance.

The solution time analysis reveals that BC_I achieved the optimal solutions faster when compared to BC_{II} in 57% of the cases. Moreover, both proposed algorithms performed better than the BC_0 approach in terms of solution time across all instances. In addition, the results show that the VSS is approximately 95% for all instances. This highlights the importance of accounting for stochasticity when modeling the CRP problem, as solutions obtained by solving the "mean-value" problem are not good enough for the stochastic problem. To respect the page limit, we have presented the results of the L-shaped algorithm in Online Appendix EC.8. However, it is worth noting that the B&C algorithm outperformed the L-shaped algorithm when using the same combination of accelerators.

6.4. Marginal Impacts of Accelerators

To gain insight into the effectiveness of each enhancement, we use BC_I as our baseline and remove the accelerators one by one, analyzing the results of each removal. The summarized results can be found in Table 3 and 4, as well as in Figure EC.5.1 and Figure EC.5.2 in Section EC.5 of the online Appendix. In these tables, under "$BC_{(I \setminus \{LBF\})}$" and "$BC_{(I \setminus \{PO\})}$", the performance of the best algorithm, i.e., BC_I, without LBF valid inequalities and PO cuts, are given, respectively. Additionally, we report the results for the BC_I algorithm without warm-up and BDD under columns "$BC_{(I \setminus \{Warm-up\})}$", and "$BC_{(I \setminus \{BDD\})}$", respectively. Column "$BC_{(I + \{LDD\})}$" gives information about BC_I algorithm when LDD is added to it. Comparison of the results in Tables 3 and 4 with the gaps reported for the BC_I algorithm in Table 2 show that each accelerator used in the BC_I

Table 2: Computational Results of the B&C Algorithm for the CRP problem with Stochastic Demand and Hospital Bed Capacity.

Problem Size		BC_0					BC_{II}					BC_I								
# ACF	# De.Loc.	Nodes No.	Time (sec)	LB	UB	Gap (%)	Nodes No.	Time (sec)	LB	UB	Gap (%)	Nodes No.	Time (sec)	LB	UB	Gap (%)	EC (%)	PC (%)	TC (%)	SC (%)	VSS (%)
10	10	20,352	8,742	71,294	71,294	0.0	6,861	2,398	71,294	71,294	0.0	2,673	2,082	71,294	71,294	0.0	55.8	21.6	22.5	0.0	95.98
	15	57,222	29,572	84,936	84,936	0.0	3,274	3,102	84,936	84,936	0.0	4,517	3,206	84,936	84,936	0.0	47.1	26.1	26.8	0.0	95.88
	20	77,248	65,904	100,547	112,544	12.7	4,112	4,740	112,396	112,396	0.0	6,703	6,453	112,396	112,396	0.0	44.5	26.8	28.7	0.0	95.25
	25	64,102	86,400	98,700	134,746	37.4	3,803	4,381	134,560	134,560	0.0	2,512	3,674	134,560	134,560	0.0	43.5	25.9	30.7	0.0	94.97
	30	53,991	86,400	93,787	160,863	72.6	5,316	5,911	152,864	152,864	0.0	6,015	9,149	152,864	152,864	0.0	44.0	27.1	28.9	0.0	94.83
15	10	44,817	52,929	60,662	78,111	32.50	4,975	3,620	70,144	70,144	0.0	4,670	3,863	70,144	70,144	0.0	54.2	23.4	22.5	0.0	95.84
	15	49,260	74,733	62,658	91,502	49.90	4,767	7,013	86,515	86,515	0.0	4,887	3,492	86,515	86,515	0.0	45.3	24.8	28.2	1.7	95.70
	20	51,057	86,400	60,357	123,197	105.4	2,638	5,800	106,614	106,614	0.0	7,152	8,848	106,614	106,614	0.0	45.2	26.4	28.4	0.0	95.69
	25	40,426	86,400	64,655	174,057	174.7	3,991	11,442	129,213	129,213	0.0	8,099	11,198	129,213	129,213	0.0	44.7	27.1	28.2	0.0	95.43
	30	36,343	86,400	70,889	195,192	176.4	7,664	13,265	148,799	148,799	0.0	6,218	11,585	148,799	148,799	0.0	45.9	27.9	26.3	0.0	95.25
20	10	44,755	82,091	64,434	111,083	82.1	5,471	8,209	89,341	89,341	0.0	8,027	8,210	89,341	89,341	0.0	52.5	24.9	22.5	0.0	96.46
	15	46,615	86,400	54,250	93,900	83.5	2,866	6,809	81,913	81,913	0.0	3,719	6,333	81,913	81,913	0.0	44.4	26.4	29.2	0.0	96.24
	20	33,866	86,400	62,387	182,208	193.7	4,714	14,799	135,098	135,098	0.0	3,447	11,397	135,098	135,098	0.0	45.1	27.7	27.2	0.0	96.40
	25	37,311	86,400	57,033	183,345	219.5	9,073	17,773	129,052	129,052	0.0	8,083	16,728	129,052	129,052	0.0	45.1	27.3	27.6	0.0	95.28
	30	28,487	86,400	81,606	489,132	493	15,786	39,170	189,248	189,248	0.0	9,713	19,861	189,248	189,248	0.0	44.1	27.9	28.0	0.0	95.58
25	10	38,567	86,076	42,266	103,501	174.4	6,218	11,123	62,969	62,969	0.0	4,639	9,288	62,969	62,969	0.0	49.2	24.6	26.2	0.0	96.47
	15	31,126	86,400	41,776	131,532	224.0	2,489	10,572	81,884	81,884	0.0	4,412	12,310	81,884	81,884	0.0	44.0	26.3	29.7	0.0	96.06
	20	30,653	86,400	45,888	132,961	190.9	7,612	25,843	103,774	103,774	0.0	2,889	20,392	103,774	103,774	0.0	43.8	27.0	29.3	0.0	95.70
	25	28,500	86,400	54,354	504,775	827.9	12,004	51,736	126,730	127,516	0.6	11,691	34,941	127,287	127,287	0.0	43.8	27.7	28.5	0.0	95.19
	30	22,462	86,400	64,169	227,172	255.0	7,684	52,985	142,678	146,703	2.9	10,058	51,805	144,366	146,494	1.5	44.4	28.3	27.2	0.1	95.27
30	10	24,411	86,400	33,432	170,149	410.6	1,994	19,230	65,726	65,726	0.0	2,924	24,869	65,726	65,726	0.0	52.7	22.9	24.4	0.0	96.22
	15	28,085	86,400	35,685	110,018	208.9	4,034	32,332	82,406	82,406	0.0	7,054	34,350	82,406	82,406	0.0	46.7	26.5	26.8	0.0	96.13
	20	25,575	86,400	44,349	158,219	256.5	4,427	40,186	102,028	102,028	0.0	8,453	46,085	101,943	102,028	0.1	45.4	28.1	26.5	0.0	95.81
	25	21,031	86,400	53,541	1,065,915	1,897	10,650	77,125	122,064	126,828	4.0	9,719	57,124	124,152	125,422	1.1	43.5	27.7	28.8	0.0	95.37
	30	17,531	86,400	69,759	5,519,098	7,899	8,961	83,403	155,760	159,507	2.4	7,879	62,579	157,632	159,152	1.0	45.0	28.3	26.8	0.0	95.36
Average		38,152	78,210	62,937	416,378	563.1	6,055	22,119	110,720	111,253	0.4	6,246	19,193	110,965	111,165	0.1	46.4	26.4	27.2	0.07	95.69

Table 3: Reverse Marginal Impact (part 1).

Problem Size		$BC_{I\setminus\{LBF\}}$					$BC_{I\setminus\{PO\}}$					$BC_{I\setminus\{Warm-up\}}$				
# ACF	# De.Loc.	Nodes No.	Time (sec)	LB	UB	Gap (%)	Nodes No.	Time (sec)	LB	UB	Gap (%)	Nodes No.	Time (sec)	LB	UB	Gap (%)
10	10	9,414	4,855	71,294	71,294	0.0	4,570	1,540	71,294	71,294	0.0	3,162	2,488	71,294	71,294	0.0
	15	14,787	10,384	84,936	84,936	0.0	5,302	2,764	84,936	84,936	0.0	2,791	2,762	84,936	84,936	0.0
	20	19,081	13,379	112,396	112,396	0.0	6,272	4,521	112,396	112,396	0.0	2,378	4,509	112,396	112,396	0.0
	25	52,252	20,115	134,560	134,560	0.0	6,292	5,554	134,560	134,560	0.0	6,193	5,872	134,560	134,560	0.0
	30	46,012	42,556	152,864	152,864	0.0	16,040	7,104	152,864	152,864	0.0	2,971	5,601	152,864	152,864	0.0
15	10	13,114	8,614	70,144	70,144	0.0	8,353	4,584	70,144	70,144	0.0	4,697	3,140	70,144	70,144	0.0
	15	25,603	29,093	86,515	86,515	0.0	8,235	4,931	86,515	86,515	0.0	7,961	6,798	86,515	86,515	0.0
	20	22,155	51,963	102,054	106,614	5.1	10,847	7,919	106,614	106,614	0.0	5,171	4,738	106,614	106,614	0.0
	25	32,855	80,137	113,223	130,250	15.9	3,704	7,013	129,213	129,213	0.0	4,953	9,182	129,213	129,213	0.0
	30	22,566	78,381	126,120	153,251	24.0	9,468	13,473	148,799	148,799	0.0	5,056	11,200	148,799	148,799	0.0
20	10	14,635	34,132	89,341	89,341	0.0	8,095	18,346	89,341	89,341	0.0	8,769	11,284	89,341	89,341	0.0
	15	20,265	31,421	81,913	81,913	0.0	2,421	6,654	81,913	81,913	0.0	2,222	5,022	81,913	81,913	0.0
	20	22,264	86,400	113,344	139,149	25.2	9,684	13,144	135,098	135,098	0.0	7,322	12,985	135,098	135,098	0.0
	25	10,998	86,400	91,525	139,287	52.2	19,211	28,912	128,626	129,052	0.3	7,833	25,119	129,052	129,052	0.0
	30	11,256	86,400	135,316	210,898	56.0	14,914	38,802	187,162	190,398	1.8	12,169	35,559	189,248	189,248	0.0
25	10	10,335	40,512	56,180	64,278	18.5	3,413	11,671	62,969	62,969	0.0	5,579	8,481	62,969	62,969	0.0
	15	23,886	67,763	75,700	82,758	10.8	6,959	11,544	81,884	81,884	0.0	9,693	11,922	81,884	81,884	0.0
	20	14,885	86,400	76,449	105,587	39.3	6,753	19,162	103,774	103,774	0.0	9,526	25,453	103,774	103,774	0.0
	25	7,240	86,400	88,808	144,047	62.2	7,569	35,694	127,287	127,287	0.0	9,599	36,465	127,287	127,287	0.0
	30	6,930	86,400	104,718	165,502	58.0	15,194	63,452	143,478	146,690	2.3	7,524	34,335	146,319	146,319	0.0
30	10	22,915	53,806	65,726	65,726	0.0	5,596	28,587	65,726	65,726	0.0	4,120	20,156	65,726	65,726	0.0
	15	23,646	86,400	63,444	87,418	41.7	7,981	45,210	82,406	82,406	0.0	6,101	27,481	82,406	82,406	0.0
	20	5,030	86,400	72,179	129,491	79.4	12,199	46,737	101,565	102,303	0.7	5,214	38,126	102,028	102,028	0.0
	25	4,798	86,400	86,513	247,202	184.3	16,353	84,704	122,573	125,892	2.8	9,717	70,184	122,237	125,995	3.1
	30	3,710	86,400	113,071	740,939	498.1	16,026	82,585	155,554	159,709	2.6	7,811	86,400	151,275	159,218	5.1
Average		18,425	57,244	94,733	143,854	46.8	9,258	23,784	110,668	111,271	0.4	6,341	20,210	110,716	111,184	0.3

Table 4: Reverse Marginal Impact (part 2).

Problem Size		$BC_{I\backslash\{BDD\}}$					$BC_{I+\{LDD\}}$				
# ACF	# De.Loc.	Nodes No.	Time (sec)	LB	UB	Gap (%)	Nodes No.	Time (sec)	LB	UB	Gap (%)
10	10	5,072	1,844	71,294	71,294	0.00	6,169	5,319	71,294	71,294	0.00
	15	6,815	2,605	84,936	84,936	0.00	3,853	6,300	84,936	84,936	0.00
	20	4,579	3,530	112,396	112,396	0.00	6,571	9,270	112,396	112,396	0.00
	25	797	3,381	134,560	134,560	0.00	4,737	15,466	134,560	134,560	0.00
	30	2,057	4,649	152,864	152,864	0.00	6,961	14,015	152,864	152,864	0.00
15	10	3,922	4,427	70,144	70,144	0.00	5,036	13,774	70,144	70,144	0.00
	15	2,999	6,978	86,515	86,515	0.00	1,668	15,937	86,515	86,515	0.00
	20	3,295	5,831	106,614	106,614	0.00	5,017	40,206	106,614	106,614	0.00
	25	3,042	8,818	129,213	129,213	0.00	6,129	44,029	129,213	129,213	0.00
	30	15,529	14,692	148,799	148,799	0.00	4,351	44,741	148,799	148,799	0.00
20	10	12,284	11,795	89,341	89,341	0.00	6,923	43,523	89,341	89,341	0.00
	15	3,355	4,444	81,913	81,913	0.00	1,861	36,722	81,913	81,913	0.00
	20	6,234	11,877	135,098	135,098	0.00	6,521	54,277	135,098	135,098	0.00
	25	4,538	15,468	129,052	129,052	0.00	2,457	52,442	129,052	129,052	0.00
	30	4,280	20,069	189,248	189,248	0.00	7,263	56,968	187,782	189,248	0.76
25	10	4,179	6,725	62,969	62,969	0.00	5,556	49,882	62,969	62,969	0.00
	15	3,660	7,780	81,884	81,884	0.00	5,955	51,754	81,884	81,884	0.00
	20	3,261	12,802	103,774	103,774	0.00	2,881	54,936	103,774	103,774	0.00
	25	7,515	29,719	127,287	127,287	0.00	6,032	61,904	127,287	127,287	0.00
	30	11,297	37,542	146,319	146,319	0.00	9,867	67,092	143,926	146,319	1.70
30	10	925	5,296	65,726	65,726	0.00	4,370	52,372	65,726	65,726	0.00
	15	4,922	20,331	82,406	82,406	0.00	2,280	60,992	82,406	82,406	0.00
	20	11,231	25,750	102,028	102,028	0.00	5,178	62,452	101,710	102,193	0.48
	25	17,447	78,520	121,379	126,280	4.03	7,157	81,956	121,398	125,569	3.54
	30	13,973	67,884	154,876	159,273	2.86	5,476	83,595	152,230	819,356	409.97
Average		6,288	16,510	110,825	111,197	0.28	5,211	43,197	110,553	137,579	16.66

algorithm positively marginalizes the gap. The only accelerator that elongates solution time within our algorithm is BDD. It has raised the average solution time from 16,510 seconds in $BC_{(I \setminus \{BDD\})}$ to 19,193 seconds in BC_I. This increased time could be attributed to the necessity of solving MIP subproblems with each iteration. Additionally, we found that the LBF valid inequalities and PO cuts have the most significant and least significant marginal impacts on the algorithm's efficiency, respectively. Removing LBF valid inequalities dramatically increases solution time and prevents the algorithm from reaching an optimal solution in 15 instances. Furthermore, we have assessed the impact of a combination of enhancements in Online Appendix EC.6.

7. Case Study

Apart from evaluating the generated instances, we also examined the effectiveness of our proposed algorithm in solving a CRP problem derived from an earthquake case study in Turkey with real data. Our study aimed to identify the location of ACFs in Van Province, Turkey, following the earthquake of 2011. Van district in Turkey is located in a seismically active region and, as a result, has experienced numerous earthquakes throughout its history. The eastern Turkish province of Van experienced a catastrophic earthquake on October 23, 2011, which measured 7.2 on the Richter scale and resulted in numerous fatalities. A second harmful tremor (5.6 on the Richter scale) struck near Van City two weeks after the initial earthquake in October. These two earthquakes had a profound impact, as over 600 individuals lost their lives, and $2,500$ suffered injuries (International Federation of Red Cross and Red Crescent Societies 2012). The 2011 Van earthquake has been the subject of numerous scholarly studies, covering a range of aspects such as last-mile relief network design (Noyan, Balcik, and Atakan 2016, Noyan and Kahvecioğlu 2018), humanitarian relief network design (Elçi, Noyan, and Bülbül 2018), shelter location (Kılcı, Kara, and Bozkaya 2015) and assessment planning (Balcik and İhsan Yanıkoğlu 2020). We contribute to this stream of research by studying the location-allocation of ACFs in the Van district after the disaster utilizing a dataset from the related natural disaster. We mainly follow the data scheme presented by Oksuz and Satoglu (2020) and Noyan, Balcik, and Atakan (2016). Below, we outline the key parameters for the case study.

We have 94 demand neighborhoods affected by the earthquake (squares in Figure 1, with size indicating demand) and 31 potential ACF locations (circles in Figure 1, with size indicating capacity). Noyan, Balcik, and Atakan (2016) also grouped these 94 points into 30-point and 60-point clusters using a p-median model, which minimizes the total demand weighted travel time. We also evaluate our model on these two versions. Each neighborhood is classified into one of three damage intensity levels, "destructive," "damaging," and "strong or very strong," based on the earthquake's intensity and distance from the fault line. Each damage intensity level corresponds to different damage state ratios, including "no damage," "slight and medium damage," and "heavy damage and collapse" for buildings in the neighborhood. The base demand for each location was then determined by considering the neighborhood's population, the damage intensity level, the
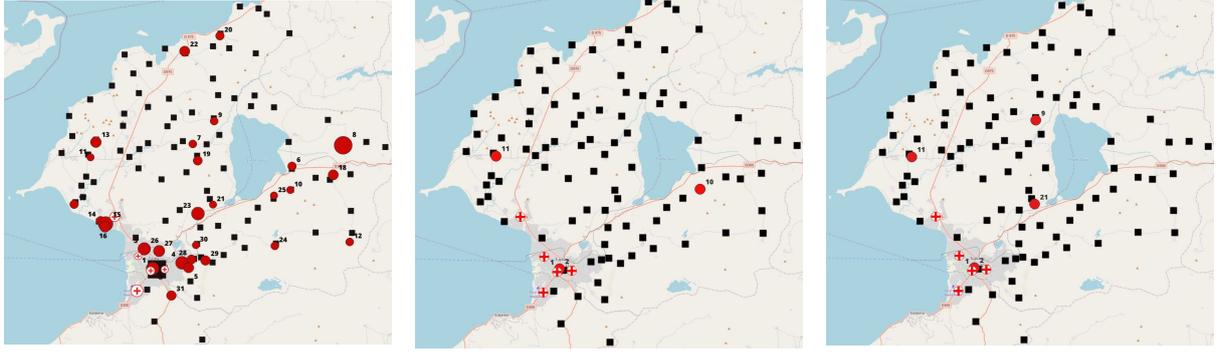
Figure 1   Van's Demands, ACFs, and Hospitals Points.

Figure 2   Van District ACFs (Deterministic).

Figure 3   Van District ACFs (Stochastic).

damage state, and the total number of injuries resulting from the 2011 Van earthquake (see the online Appendix in Noyan, Balcik, and Atakan (2016), for details). To realize demands for each outcome, we multiply the base demands by a deviation factor. These deviation factors are sampled from uniform distributions, which have intervals of $[0.75, 1.10]$, $[0.75, 1.20]$, and $[0.75, 1.30]$ for intensity levels of "destructive," "damaging," and "strong or very strong," respectively.

In the Van district, five primary hospitals are available to assist patients during a disaster (represented with circles with a plus inside in Figure 1, with size indicating capacity). However, in the case of an earthquake, the capacity of these hospitals becomes uncertain. Therefore, we assume their capacity will be reduced by a random percentage between 0% and 21% (Oksuz and Satoglu 2020). To calculate the distance between points a and b, we use the Haversine formula (61), in which ϕ_a and λ_a are the latitude and longitude of point a, respectively, and r represents the Earth's radius, i.e., $6{,}371$ km (Chen et al. 2022). The distance H in kilometers is then multiplied by the network circuity factor for Turkey, i.e., 1.36, provided by Ballou, Rahardja, and Sakai (2002). We obtain the remaining parameters based on what we explained in Section 6.3.

$$H = 2r\sin^{-1}\left(\sqrt{\sin^2\left(\frac{\phi_2 - \phi_1}{2}\right) + \cos(\phi_1)\cos(\phi_2)\sin^2\left(\frac{\lambda_2 - \lambda_1}{2}\right)}\right) \tag{61}$$

The average optimality gap of approximately 0.79% signifies that our proposed algorithm is capable of generating high-quality solutions. Additionally, the substantial values of VSS, averaging at 97.79%, underscore the critical role of stochasticity in disaster management. Let us delve deeper into the solutions. In both deterministic and stochastic models, we established ACFs with specific capacities and the ability to treat certain injuries. In both models, we established the following ACFs:

- ACF #1 (with a capacity of 750, capable of treating the following injuries: fractures/crush (Injury #1), laceration/contusions/open wounds (Injury #2), peripheral nerve injury (Injury #3), and burns (Injury #4)),
- ACF #2 (with a capacity of 570, treating Injuries #2 and #3),
- and ACF #11 (with a capacity of 52, treating injury #2).

However, in the deterministic model, we opened an additional ACF, #10, with a capacity of 70 patients capable of treating injuries #1, #2, and #4 concurrently (Figure 2). In the stochastic model, considering 100 different scenarios with varying patient numbers and hospital bed capacities, we established two more ACFs: ACF #9 (treating injuries #1 and #2) and ACF #21 (specializing in treating Injuries #1 and #4), with capacities of 130 and 100, respectively (Figure 3). Notably, the decision to incorporate two additional ACFs, namely #9 and #21, rather than just one (e.g., ACF #10), results in a total capacity increase of 160 patients compared to the deterministic model. This decision not only allows us to provide care for more patients but also enhances our geographical coverage, leading to reduced patient transport times.

Examining the challenging scenarios that can arise after an earthquake, particularly when we base decisions on the average demand in the deterministic model, we come across a critical situation in scenario #22. This scenario has the highest number of low-priority cases associated with Injuries #1 and #2. Moreover, if first-stage decisions are made based on the average scenario in the deterministic model, this scenario also leads to the highest number of unsatisfied low-priority patients. In scenario #22, we encounter a remarkable 1,596 low-priority patients, with 402 patients falling into class #3, the most heavily populated patient class within this scenario. Upon closer examination of class #3, we discover that it exclusively comprises patients with Injuries #1 and #2. This revelation underscores the necessity of an ACF specializing in treating these two specific injuries, much like ACF #9 in the stochastic model. This level of responsiveness to the unique demands of various scenarios is a notable advantage of the stochastic approach, and it addresses an oversight in the deterministic model where such specialized care was not adequately considered.

Effective collaboration among various organizations is crucial for ensuring the availability of necessary equipment, an often overlooked aspect in post-disaster scenarios. The differences between the deterministic and stochastic models extend beyond the number, capacity, and treatment capability of the established ACFs. For example, the stochastic model significantly increases the number of vehicles of various types assigned to ACFs to ensure all low-priority patients can be attended to in every scenario, including worst-case scenarios. Concerning the number of physicians, while the deterministic model accurately estimated the number of required orthopedic physicians, it significantly underestimated the need for wound specialists and plastic physicians or dermatologists and overestimated the number of required neurologists. This highlights the importance of a collaborative effort among companies and organizations to ensure that the right mix of medical professionals and resources is readily available in post-disaster situations, as the requirements vary considerably between deterministic and stochastic models.

8. Conclusion

In this paper, we addressed the challenge of a casualty response planning problem with patients having multiple injuries. To tackle this problem, we developed two two-stage stochastic integer

programming models to formulate the CRP with uncertain demands and hospital bed capacities. The first model is a basic mathematical model, whereas the second one is a more intricate version that significantly reduces solution times. We proved that the two proposed models are equivalent. As our solution methods, we introduced both the L-shaped algorithm, a pure cutting-plane method tailored to our stochastic mathematical model, and the Benders branch-and-cut algorithm. In order to significantly improve the efficiency of these algorithms, we have incorporated a diverse set of acceleration techniques, including Benders dual decomposition, Lagrangian dual decomposition, a multi-cut reformulation, the integration of Pareto-optimal cuts, and the utilization of lower bounding functional valid inequalities.

Our computational results showed that the improved two-stage stochastic programming model is around 42.5% faster than the initial model. We demonstrated that the proposed acceleration techniques significantly improve the performance of the algorithms and reduce the average optimality gap from 563.1% to 0.1%. Particularly, in some instances, the gap improved from nearly 7,900% to around 1%. We also observed that the average VSS is 95.7%, that demonstrates the importance of addressing uncertainty in casualty response planning problem and highlights the advantage of our proposed two-stage stochastic programming model compared to its corresponding deterministic model. We also provided extensive results to measure the contribution of enhancements separately. To highlight the applicability of our proposed models and algorithms, we ran our best algorithm on a case study from Turkey and demonstrated that we can find near-optimal solution with an average gap of 0.8%.

We suggest two important directions for future research. First, as we have demonstrated the crucial role of road conditions and travel times in post-disaster outcomes, it would be valuable to incorporate debris clearance as an integral component of the problem outlined in this paper. Second, developing robust optimization for the CRP with uncertain demands, stochastic hospital capacities, and multiple injuries would be paramount, considering that information ambiguity is often present in such planning contexts.

References

Alizadeh M, Amiri-Aref M, Mustafee N, Matilal S, 2019 *A robust stochastic casualty collection points location problem. European Journal of Operational Research* 279(3):965–983, URL http://dx.doi.org/https://doi.org/10.1016/j.ejor.2019.06.018.

Apte A, Heidtke C, Salmerón J, 2015 *Casualty collection points optimization: A study for the district of columbia. Interfaces* 45(2):149–165, URL http://dx.doi.org/10.1287/inte.2014.0757.

Balcik B, İhsan Yanıkoğlu, 2020 *A robust optimization approach for humanitarian needs assessment planning under travel time uncertainty. European Journal of Operational Research* 282(1):40–57, URL http://dx.doi.org/https://doi.org/10.1016/j.ejor.2019.09.008.

Ballou RH, Rahardja H, Sakai N, 2002 *Selected country circuity factors for road travel distance estimation. Transportation Research Part A: Policy and Practice* 36(9):843–848, URL http://dx.doi.org/https://doi.org/10.1016/S0965-8564(01)00044-1.

Batun S, Denton BT, Huschka TR, Schaefer AJ, 2011 *Operating room pooling and parallel surgery processing under uncertainty. INFORMS Journal on Computing* 23(2):220–237, URL http://dx.doi.org/10.1287/ijoc.1100.0396.

Caunhye AM, Li M, Nie X, 2015 *A location-allocation model for casualty response planning during catastrophic radiological incidents. Socio-Economic Planning Sciences* 50:32–44, URL http://dx.doi.org/https://doi.org/10.1016/j.seps.2015.02.001.

Caunhye AM, Nie X, 2018 *A stochastic programming model for casualty response planning during catastrophic health events. Transportation Science* 52(2):437–453, URL http://dx.doi.org/10.1287/trsc.2017.0777.

Chang KH, Chen TL, Yang FH, Chang TY, 2023 *Simulation optimization for stochastic casualty collection point location and resource allocation problem in a mass casualty incident. European Journal of Operational Research* 309(3):1237–1262, URL http://dx.doi.org/https://doi.org/10.1016/j.ejor.2023.01.065.

Chen Y, Marković N, Ryzhov IO, Schonfeld P, 2022 *Data-driven robust resource allocation with monotonic cost functions. Operations Research* 70(1):73–94, URL http://dx.doi.org/10.1287/opre.2021.2145.

Costa AdS, 2017 *Assessment of operative times of multiple surgical specialties in a public university hospital. Einstein (Sao Paulo)* 15:200–205.

Crainic TG, Hewitt M, Maggioni F, Rei W, 2021 *Partial benders decomposition: General methodology and application to stochastic network design. Transportation Science* 55(2):414–435, URL http://dx.doi.org/10.1287/trsc.2020.1022.

Del Papa J, Vittorini P, D'Aloisio F, Muselli M, Giuliani AR, Mascitelli A, Fabiani L, 2019 *Retrospective analysis of injuries and hospitalizations of patients following the 2009 earthquake of l'aquila city. International journal of environmental research and public health* 16(10):1675.

Doocy S, Daniels A, Packer C, Dick A, Kirsch TD, 2013 *The human impact of earthquakes: a historical review of events 1980-2009 and systematic literature review. PLoS currents* 5.

Elçi Ö, Noyan N, Bülbül K, 2018 *Chance-constrained stochastic programming under variable reliability levels with an application to humanitarian relief network design. Computers & Operations Research* 96:91–107, URL http://dx.doi.org/https://doi.org/10.1016/j.cor.2018.03.011.

Farahani RZ, Lotfi M, Baghaian A, Ruiz R, Rezapour S, 2020 *Mass casualty management in disaster scene: A systematic review of or&ms research in humanitarian operations. European Journal of Operational Research* 287(3):787–819.

Hashemi Doulabi H, Pesant G, Rousseau LM, 2020 *Vehicle routing problems with synchronized visits and stochastic travel and service times: Applications in healthcare. Transportation Science* 54(4):1053–1072, URL http://dx.doi.org/10.1287/trsc.2019.0956.

Hu QM, Zhao L, Li H, Huang R, 2019 *Integrated design of emergency shelter and medical networks considering diurnal population shifts in urban areas. IISE Transactions* 51(6):614–637, URL http://dx.doi.org/10.1080/24725854.2018.1519744.

International Federation of Red Cross and Red Crescent Societies, 2012 *Emergency appeal operation update turkey: Van earthquake.* Online, URL `https://adore.ifrc.org/Download.aspx?FileId=23186`.

Jensen JLWV, 1906 *Sur les fonctions convexes et les inégalités entre les valeurs moyennes. Acta Mathematica* 30(none):175 – 193, URL `http://dx.doi.org/10.1007/BF02418571`.

Kang P, Tang B, Liu Y, Liu X, Liu Z, Lv Y, Zhang L, 2015 *Medical efforts and injury patterns of military hospital patients following the 2013 lushan earthquake in china: a retrospective study. International journal of environmental research and public health* 12(9):10723–10738.

Kılcı F, Kara BY, Bozkaya B, 2015 *Locating temporary shelter areas after an earthquake: A case for turkey. European Journal of Operational Research* 243(1):323–332, URL `http://dx.doi.org/https://doi.org/10.1016/j.ejor.2014.11.035`.

Larson RC, Metzger MD, Cahn MF, 2006 *Responding to emergencies: Lessons learned and the need for analysis. Interfaces* 36(6):486–501, URL `http://www.jstor.org/stable/20141439`.

Li Y, Zhang J, Yu G, 2020 *A scenario-based hybrid robust and stochastic approach for joint planning of relief logistics and casualty distribution considering secondary disasters. Transportation Research Part E: Logistics and Transportation Review* 141:102029, URL `http://dx.doi.org/https://doi.org/10.1016/j.tre.2020.102029`.

Liu Y, Cui N, Zhang J, 2019 *Integrated temporary facility location and casualty allocation planning for post-disaster humanitarian medical service. Transportation Research Part E: Logistics and Transportation Review* 128:1–16, URL `http://dx.doi.org/https://doi.org/10.1016/j.tre.2019.05.008`.

Lu-Ping Z, Rodriguez-Llanes JM, Qi W, van den Oever B, Westman L, Albela M, Liang P, Gao C, De-Sheng Z, Hughes M, et al., 2012 *Multiple injuries after earthquakes: a retrospective analysis on 1,871 injured patients from the 2008 wenchuan earthquake. Critical Care* 16:1–9.

Magnanti TL, Wong RT, 1981 *Accelerating benders decomposition: Algorithmic enhancement and model selection criteria. Operations Research* 29(3):464–484, URL `http://www.jstor.org/stable/170108`.

Martins de Sá E, Contreras I, Cordeau JF, Saraiva de Camargo R, de Miranda G, 2015 *The hub line location problem. Transportation Science* 49(3):500–518, URL `http://dx.doi.org/10.1287/trsc.2014.0576`.

McDaniel D, Devine M, 1977 *A modified benders' partitioning algorithm for mixed integer programming. Management Science* 24(3):312–319, URL `http://www.jstor.org/stable/2630825`.

Noyan N, Balcik B, Atakan S, 2016 *A stochastic optimization model for designing last mile relief networks. Transportation Science* 50(3):1092–1113, URL `http://dx.doi.org/10.1287/trsc.2015.0621`.

Noyan N, Kahvecioğlu G, 2018 *Stochastic last mile relief network design with resource reallocation. Or Spectrum* 40:187–231.

Oksuz MK, Satoglu SI, 2020 *A two-stage stochastic model for location planning of temporary medical centers for disaster response. International Journal of Disaster Risk Reduction* 44:101426, URL `http://dx.doi.org/https://doi.org/10.1016/j.ijdrr.2019.101426`.

Rahmaniani R, Ahmed S, Crainic TG, Gendreau M, Rei W, 2020 *The benders dual decomposition method. Operations Research* 68(3):878–895, URL `http://dx.doi.org/10.1287/opre.2019.1892`.

Rahmaniani R, Crainic TG, Gendreau M, Rei W, 2017 *The benders decomposition algorithm: A literature review. European Journal of Operational Research* 259(3):801–817, URL `http://dx.doi.org/https://doi.org/10.1016/j.ejor.2016.12.005`.

Rambha T, Nozick LK, Davidson R, Yi W, Yang K, 2021 *A stochastic optimization model for staged hospital evacuation during hurricanes. Transportation Research Part E: Logistics and Transportation Review* 151:102321, URL `http://dx.doi.org/https://doi.org/10.1016/j.tre.2021.102321`.

Runge JW, Buddemeier BR, 2009 *Explosions and radioactive material: A primer for responders. Prehospital Emergency Care* 13(4):407–419, URL `http://dx.doi.org/10.1080/10903120902935371`, pMID: 19731151.

Setiawan E, Liu J, French A, 2019 *Resource location for relief distribution and victim evacuation after a sudden-onset disaster. IISE Transactions* 51(8):830–846, URL `http://dx.doi.org/10.1080/24725854.2018.1517284`.

Sever MS, Vanholder R, of ISN Work Group on Recommendations for the Management of Crush Victims in Mass Disasters R, 2012 *Recommendations for the management of crush victims in mass disasters. Nephrology dialysis transplantation* 27(Suppl_1):i1–i67.

Sherali HD, Lunday BJ, 2013 *On generating maximal nondominated benders cuts. Annals of Operations Research* 210:57–72.

Sun H, Wang Y, Xue Y, 2021 *A bi-objective robust optimization model for disaster response planning under uncertainties. Computers & Industrial Engineering* 155:107213, URL `http://dx.doi.org/https://doi.org/10.1016/j.cie.2021.107213`.

Tanaka H, Oda J, Iwai A, Kuwagata Y, Matsuoka T, Takaoka M, Kishi M, Morimoto F, Ishikawa K, Mizushima Y, et al., 1999 *Morbidity and mortality of hospitalized patients after the 1995 hanshin-awaji earthquake. The American journal of emergency medicine* 17(2):186–191.

Van Slyke RM, Wets R, 1969 *L-shaped linear programs with applications to optimal control and stochastic programming. SIAM Journal on Applied Mathematics* 17(4):638–663, URL `http://dx.doi.org/10.1137/0117061`.

World Health Organization, 2019 *Trauma and emergency surgery kit (TESK).* `https://www.who.int/emergencies/emergency-health-kits/trauma-emergency-surgery-kit-who-tesk-2019`, accessed: 2023-08-08.

Yin Y, Wang J, Chu F, Wang D, 2023 *Distributionally robust multi-period humanitarian relief network design integrating facility location, supply inventory and allocation, and evacuation planning. International Journal of Production Research* 1–26.

Zou J, Ahmed S, Sun XA, 2019 *Stochastic dual dynamic integer programming. Mathematical Programming* 175:461–502.

Electronic Companion — "Stochastic Casualty Response Planning with Multiple Classes of Patients"

Contents

ec1

EC.1. Proof of Theorem 1

In order to prove the equality of the two models, it is sufficient to demonstrate that an algorithm exists, such as Algorithm EC.1.1 and Algorithm EC.1.2, which can transform any feasible solution of model (P1) into a feasible solution of model (P2), and vice versa, while preserving the same objective function value. In these algorithms, we introduce a one-to-one function, $\mathcal{F}(p)$, which maps the value of p to the corresponding set of injuries. For example, $\mathcal{F}(3)$ returns the set injuries $\{1,2\}$, $\mathcal{F}(15)$ returns the injury set $\{1,2,3,4\}$, and $\mathcal{F}(0)$ corresponds to the empty set. To maintain our investigation into model equivalence, we introduce Algorithm EC.1.1, which takes the decision variable values from model (P1) and transforms them into equivalent variables for model (P2).

It is worth highlighting that model (P1) and model (P2) are nearly identical except for three variables: $u_{pp'ihm}(\omega)$, $\eta_{pp'ii'm}(\omega)$, and $\sigma_{pi}(\omega)$ in model (P1), and $u_{pihm}(\omega)$, $\eta_{pii'm}(\omega)$, and $\sigma_{pji}(\omega)$ in model (P2). Due to this similarity, our focus will be solely on these specific variables. This leads us to the subsequent Algorithm EC.1.2, where our attention turns to the decision variable values of model (P2). In Algorithm EC.1.2, we aim to find feasible solutions for model (P1) using the variables obtained from model (P2). In this algorithm we have:

$\mathcal{Z}_{ip}(\omega)$: The number of patients of type p at ACF i when the outcome ω is observed.

$\mathcal{N}_{ip'p}(\omega)$: The number of class p patients previously classified as class p', who are now in a state of readiness for transfer from ACF i to either hospital h or another ACF when the outcome ω is observed.

The objective function of model (P1) consists of objectives (1) and (10), while the objective function of model (P2) comprises objectives (1) and (21). With the understanding that $\sum_{p'\in\mathcal{P}} u_{p'pihm}(\omega) = u_{pihm}(\omega)$ for all $p \in \mathcal{P}, i \in \mathcal{I}, h \in \mathcal{H}, m \in \mathcal{M}, and\ \omega \in \Omega$, and $\sum_{p'\in\mathcal{P}} \eta_{p'pii'm}(\omega) = \eta_{pii'm}(\omega)$ for all $p \in \mathcal{P}, i \in \mathcal{I}, i' \in \mathcal{I}, i' \neq i, m \in \mathcal{M}, and\ \omega \in \Omega$, we can confidently assert that the objective functions of both models are effectively aligned. These insights establish a firm basis for confirming the equality of the objective functions, thus providing a strong foundation for the conversion process between the two models.

Algorithm EC.1.1 Model (P1) to Model (P2).

1: **for** $(\omega = 1$ to $\Omega)$ **do**

2:      $\sigma_{pji}(\omega) := 0$ for $p \in \mathcal{P}, j \in \mathcal{J}_i, i \in \mathcal{I}$

3:      **for** $(i = 1$ to $|\mathcal{I}|)$ **do**

4:          $p^{new} := 0$

5:          **for** $(p = 1$ to $|\mathcal{P}|)$ **do**

6:             $\mathcal{Q} := \varnothing$

7:             **for** $(p' = 1$ to $|\mathcal{P}|)$ **do**

8:                 **if** $(\mathcal{F}(p') \subset \mathcal{F}(p))$ **then**

9:                     $\mathcal{Q} := (\mathcal{F}(p) \setminus \mathcal{F}(p'))$

10:                     $t := 1$

11:                     $p^{new} := p$

12:                     **for** $(\mathcal{Q}_{\{t\}}$ to $\mathcal{Q}_{\{|\mathcal{Q}|\}})$ **do**

13:                        $\sigma_{p^{new}\mathcal{Q}_{\{t\}}i\omega} += \left(\sum\limits_{h \in \mathcal{H}} \sum\limits_{m \in \mathcal{M}} u_{pp'ihm}(\omega) + \sum\limits_{\substack{i' \in \mathcal{J} \\ i' \neq i}} \sum\limits_{m \in \mathcal{M}} \eta_{pp'ii'm}(\omega) \right)$

14:                       $p^{new} := \mathcal{F}^{-1}(\mathcal{F}(p^{new}) \setminus \mathcal{Q}_{\{t\}})$

15:                       $\mathcal{Q}_{\{t\}} := \mathcal{Q}_{\{t+1\}}$

16:                     **end for**

17:                 **end if**

18:             **end for**

19:             $p^{new} := 0$

20:             $\mathcal{Q} := \varnothing$

21:             $p^{new} := p$

22:             $\mathcal{Q} := \mathcal{F}(p)$

23:             $t := 1$

24:             **for** $(\mathcal{Q}_{\{t\}}$ to $\mathcal{Q}_{\{|\mathcal{Q}|\}})$ **do**

25:                 $\sigma_{p^{new}\mathcal{Q}_{\{t\}}i\omega} += \sigma_{pi}(\omega)$

26:                 $p^{new} := \mathcal{F}^{-1}(\mathcal{F}(p^{new}) \setminus \mathcal{Q}_{\{t\}})$

27:                 $\mathcal{Q}_{\{t\}} := \mathcal{Q}_{\{t+1\}}$

28:             **end for**

29:          **end for**

30:      **end for**

31: **end for**

Algorithm EC.1.2 Model (P2) to Model (P1).

1: **for** $(\omega = 1 \text{ to } \Omega)$ **do**

2:     $\mathcal{N}_{ip'p}(\omega) := 0$

3:     $\mathcal{Z}_{ip}(\omega) := 0$

4:     **for** $(i = 1 \text{ to } |\mathcal{I}|)$ **do**

5:        **for** $(p = 1 \text{ to } |\mathcal{P}|)$ **do**

6:           $\mathcal{Z}_{ip}(\omega) := -(\sum\limits_{l \in \mathcal{L}} \sum\limits_{m \in \mathcal{M}} q_{plim}(\omega) + \sum\limits_{\substack{i' \in \mathcal{I} \\ i' \neq i}} \sum\limits_{m \in \mathcal{M}} \eta_{pi'im}(\omega) - \sum\limits_{j \in \mathcal{J}_i} \sigma_{pji}(\omega))$

7:           **for** $(p' = 1 \text{ to } |\mathcal{P}|)$ **do**

8:              **if** $(\mathcal{F}(p) \subset \mathcal{F}(p'))$ **then**

9:                 **for** $(j = 1 \text{ to } \mathcal{J})$ **do**

10:                    **if** $((j \in \mathcal{F}(p')) \, \& \, ((\mathcal{F}(p') \setminus j) = \mathcal{F}(p)))$ **then**

11:                       $\mathcal{N}_{ip'p}(\omega) += \sigma_{p'ji}(\omega) - \mathcal{Z}_{ip}(\omega)$

12:                    **end if**

13:                 **end for**

14:              **end if**

15:              $\sum\limits_{m \in \mathcal{M}} \sum\limits_{h \in \mathcal{H}} u_{p'pihm}(\omega) + \sum\limits_{m \in \mathcal{M}} \sum\limits_{\substack{i' \in \mathcal{I} \\ i' \neq i}} \eta_{p'pii'm}(\omega) = \mathcal{N}_{ip'p}(\omega)$

16:           **end for**

17:           $\sigma_{pi}(\omega) := \sum\limits_{l \in \mathcal{L}} \sum\limits_{m \in \mathcal{M}} q_{plim}(\omega) + \sum\limits_{m \in \mathcal{M}} \sum\limits_{\substack{i' \in \mathcal{I} \\ i' \neq i}} \eta_{pi'im}(\omega) - \sum\limits_{p' \in \mathcal{P}} \mathcal{N}_{ip'p}(\omega)$

18:        **end for**

19:     **end for**

20: **end for**

EC.2. Proof of Theorem 2

The standard L-shaped algorithm begins by solving the restricted master problem given as follows.

$$\min_{\boldsymbol{x},\boldsymbol{\vartheta},\boldsymbol{v},\boldsymbol{r},\theta} \sum_{i \in \mathcal{I}} f_i x_i + \sum_{i \in \mathcal{I}} \sum_{h \in \mathcal{H}} \sum_{k \in \mathcal{K}} \varrho_k r_{ihk} + \theta \qquad \text{(EC.2.1)}$$

Subject to:

$$(2) - (9)$$

$$\theta \geq Q(\boldsymbol{x}, \boldsymbol{\vartheta}, \boldsymbol{v}, \boldsymbol{r}) \qquad \text{(EC.2.2)}$$

In which,

$$Q(\boldsymbol{x}, \boldsymbol{\vartheta}, \boldsymbol{v}, \boldsymbol{r}) = \sum_{\omega \in \Omega} \phi_\omega Q(\boldsymbol{x}, \boldsymbol{\vartheta}, \boldsymbol{v}, \boldsymbol{r}, \xi(\omega)) = E_{\omega \in \Omega}[Q(\boldsymbol{x}, \boldsymbol{\vartheta}, \boldsymbol{v}, \boldsymbol{r}, \xi(\omega))] \qquad \text{(EC.2.3)}$$

We can write Jensen's Inequality (Jensen 1906) for $Q(\boldsymbol{x}, \boldsymbol{\vartheta}, \boldsymbol{v}, \boldsymbol{r}, \xi(\omega))$ as follows.

$$E_{\omega \in \Omega}[Q(\boldsymbol{x}, \boldsymbol{\vartheta}, \boldsymbol{v}, \boldsymbol{r}, \xi(\omega))] \geq Q(\boldsymbol{x}, \boldsymbol{\vartheta}, \boldsymbol{v}, \boldsymbol{r}, \xi(\bar{\omega})) \qquad \text{(EC.2.4)}$$

Where $\xi(\bar{\omega})$ represents the vector of average outcome, i.e., $\xi(\bar{\omega}) = \sum_{\omega \in \Omega} \phi_\omega \xi(\omega)$. Relations (EC.2.3) and (EC.2.4) show that the following relation holds.

$$Q(\boldsymbol{x}, \boldsymbol{\vartheta}, \boldsymbol{v}, \boldsymbol{r}) \geq Q(\boldsymbol{x}, \boldsymbol{\vartheta}, \boldsymbol{v}, \boldsymbol{r}, \xi(\bar{\omega})) \qquad \text{(EC.2.5)}$$

Because we have $\theta \geq Q(\boldsymbol{x}, \boldsymbol{\vartheta}, \boldsymbol{v}, \boldsymbol{r})$ as a part of our formulation, then based on $(EC.2.5)$, $\theta \geq Q(\boldsymbol{x}, \boldsymbol{\vartheta}, \boldsymbol{v}, \boldsymbol{r}, \xi(\bar{\omega}))$ is a valid inequality for our master problem.

EC.3. Proof of Proposition 1

For each outcome $\omega \in \Omega$ the dualization of the subproblem (21)-(32) is as follows.

$$
\begin{aligned}
\max \sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}} &\left(d_{pl}^{low}(\omega) w_{\omega pl}^{(1)} + d_{pl}^{high}(\omega) w_{\omega pl}^{(2)} \right) - \sum_{h \in \mathcal{H}} c_h^{hospital}(\omega) w_{\omega h}^{(5)} - \sum_{h \in \mathcal{H}} \sum_{m \in \mathcal{M}} c_m^{rescue} n_{mh}^{hospital} w_{\omega mh}^{(9)} \\
&- \sum_{i \in \mathcal{I}} c_i^{ACF} \hat{x}_i w_{\omega i}^{(6)} - \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \tau_j \hat{v}_{ij} w_{\omega ij}^{(7)} - \sum_{i \in \mathcal{I}} \sum_{m \in \mathcal{M}} c_m^{rescue} \hat{\vartheta}_{mi} w_{\omega im}^{(10)} - \sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} \left(\sum_{h \in \mathcal{H}} \hat{r}_{ihk} \right) w_{\omega ik}^{(14)}
\end{aligned}
$$
$$\text{(EC.3.1)}$$

Subject to:

$$w_{pl}^{(1)} - w_{ip}^{(3)} - w_{ip}^{(4)} - w_i^{(6)} - \pi_{li} w_{im}^{(10)} \leq \pi_{li} \qquad p \in \mathcal{P}, l \in \mathcal{L}, i \in \mathcal{I}, m \in \mathcal{M} \quad \text{(EC.3.2)}$$

$$w_{ip}^{(3)} - w_h^{(5)} - \pi_{ih} w_{im}^{(10)} \leq \pi_{ih} \qquad p \in \mathcal{P}, i \in \mathcal{I}, h \in \mathcal{H}, m \in \mathcal{M} \quad \text{(EC.3.3)}$$

$$w_{pl}^{(2)} - w_h^{(5)} - \pi_{lh} w_{mh}^{(9)} \leq \pi_{lh} \qquad p \in \mathcal{P}, l \in \mathcal{L}, h \in \mathcal{H}, m \in \mathcal{M} \quad \text{(EC.3.4)}$$

$$w_{ip}^{(3)} - w_{i'p}^{(3)} - w_{i'p}^{(4)} - w_{i'}^{(6)} - \pi_{ii'} w_{im}^{(10)} \leq \pi_{ii'} \qquad p \in \mathcal{P}, i \in \mathcal{I}, i' \in \mathcal{I}, i \neq i', m \in \mathcal{M} \quad \text{(EC.3.5)}$$

$$w_{pl}^{(1)} \leq \rho_{pl}^{low} \qquad p \in \mathcal{P}, l \in \mathcal{L} \quad \text{(EC.3.6)}$$

$$w_{pl}^{(2)} \leq \rho_{pl}^{high} \qquad p \in \mathcal{P}, l \in \mathcal{L} \quad \text{(EC.3.7)}$$

$$w_{ip}^{(3)} - \sum_{\substack{p' \in p| \\ j \notin p' \\ p \setminus \{j\} = p' \\ p \setminus \{j\} \neq 0}} w_{ip'}^{(3)} + w_{ip}^{(4)}$$
$$- \tau_j w_{ij}^{(7)} - \sum_{k \in \mathcal{K}} g_{jk} w_{ik}^{(14)} \leq 0 \qquad p \in \mathcal{P}, i \in \mathcal{I}, j \in \{p \cap \mathcal{J}_i\} \quad \text{(EC.3.8)}$$

$$w_{ip}^{(4)}, w_h^{(5)}, w_i^{(6)}, w_{ij}^{(7)}, w_{mh}^{(9)}, w_{im}^{(10)}, w_{ik}^{(14)} \geq 0 \qquad p \in \mathcal{P}, i \in \mathcal{I}, h \in \mathcal{H}, m \in \mathcal{M}, k \in \mathcal{K}, j \in \mathcal{J}_i \quad \text{(EC.3.9)}$$

We also know that we can write model $(EC.3.1) - (EC.3.9)$ in the following format.

$$
\begin{aligned}
\min \sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}} &\left(-d_{pl}^{low}(\omega) w_{\omega pl}^{(1)} - d_{pl}^{high}(\omega) w_{\omega pl}^{(2)} \right) + \sum_{h \in \mathcal{H}} c_h^{hospital}(\omega) w_{\omega h}^{(5)} + \sum_{h \in \mathcal{H}} \sum_{m \in \mathcal{M}} c_m^{rescue} n_{mh}^{hospital} w_{\omega mh}^{(9)} \\
&+ \sum_{i \in \mathcal{I}} c_i^{ACF} \hat{x}_i w_{\omega i}^{(6)} + \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \tau_j \hat{v}_{ij} w_{\omega ij}^{(7)} + \sum_{i \in \mathcal{I}} \sum_{m \in \mathcal{M}} c_m^{rescue} \hat{\vartheta}_{mi} w_{\omega im}^{(10)} + \sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} \left(\sum_{h \in \mathcal{H}} \hat{r}_{ihk} \right) w_{\omega ik}^{(14)}
\end{aligned}
$$
$$\text{(EC.3.10)}$$

Subject to:

Constraints $(EC.3.2) - (EC.3.9)$ $\qquad \qquad \qquad \qquad \qquad \qquad \qquad \text{(EC.3.11)}$

If $x^0, \vartheta^0, \upsilon^0$, and r^0 are core points within the feasible solution set, we can achieve the Pareto optimal solution using model $(EC.3.12) - (EC.3.14)$.

$$\min \sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}} \left(-d_{pl}^{low}(\omega) w_{\omega pl}^{(1)} - d_{pl}^{high}(\omega) w_{\omega pl}^{(2)} \right) + \sum_{h \in \mathcal{H}} c_h^{hospital}(\omega) w_{\omega h}^{(5)} + \sum_{h \in \mathcal{H}} \sum_{m \in \mathcal{M}} c_m^{rescue} n_{mh}^{hospital} w_{\omega mh}^{(9)}$$
$$+ \sum_{i \in \mathcal{I}} c_i^{ACF} x_i^0 w_{\omega i}^{(6)} + \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \tau_j \upsilon_{ij}^0 w_{\omega ij}^{(7)} + \sum_{i \in \mathcal{I}} \sum_{m \in \mathcal{M}} c_m^{rescue} \vartheta_{mi}^0 w_{\omega im}^{(10)} + \sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} \left(\sum_{h \in \mathcal{H}} r_{ihk}^0 \right) w_{\omega ik}^{(14)}$$
$$(EC.3.12)$$

Subject to:

$$\sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}} \left(-d_{pl}^{low}(\omega) w_{\omega pl}^{(1)} - d_{pl}^{high}(\omega) w_{\omega pl}^{(2)} \right) + \sum_{h \in \mathcal{H}} c_h^{hospital}(\omega) w_{\omega h}^{(5)} + \sum_{h \in \mathcal{H}} \sum_{m \in \mathcal{M}} c_m^{rescue} n_{mh}^{hospital} w_{\omega mh}^{(9)}$$
$$+ \sum_{i \in \mathcal{I}} c_i^{ACF} \hat{x}_i w_{\omega i}^{(6)} + \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \tau_j \hat{\upsilon}_{ij} w_{\omega ij}^{(7)} + \sum_{i \in \mathcal{I}} \sum_{m \in \mathcal{M}} c_m^{rescue} \hat{\vartheta}_{mi} w_{\omega im}^{(10)}$$
$$+ \sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} \left(\sum_{h \in \mathcal{H}} \hat{r}_{ihk} \right) w_{\omega ik}^{(14)} = opt(\hat{x}, \hat{\vartheta}, \hat{\upsilon}, \hat{r})$$
$$(EC.3.13)$$

Constraints $(EC.3.2) - (EC.3.9)$ $\hspace{3cm}$ $(EC.3.14)$

For the sake of illustration, consider $\mathcal{D}$ as a vector containing dual variables $w_{pl}^{(1)}, w_{pl}^{(2)}, w_{ip}^{(3)}, w_{ip}^{(4)}, w_h^{(5)}, w_i^{(6)}, w_{ij}^{(7)}, w_{mh}^{(9)}, w_{im}^{(10)}, w_{ik}^{(14)}$. We establish a function denoted as

$F^\omega(\mathcal{D}, x, \vartheta, \upsilon, r) =$

$\sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}} \left(-d_{pl}^{low}(\omega) w_{\omega pl}^{(1)} - d_{pl}^{high}(\omega) w_{\omega pl}^{(2)} \right)$ $\hspace{1cm}$ $+$ $\hspace{1cm}$ $\sum_{h \in \mathcal{H}} c_h^{hospital}(\omega) w_{\omega h}^{(5)}$ $\hspace{1cm}$ $+$

$\sum_{h \in \mathcal{H}} \sum_{m \in \mathcal{M}} c_m^{rescue} n_{mh}^{hospital} w_{\omega mh}^{(9)}$ $\hspace{0.5cm}$ $+$ $\hspace{0.5cm}$ $\sum_{i \in \mathcal{I}} c_i^{ACF} x_i w_{\omega i}^{(6)}$ $\hspace{0.5cm}$ $+$ $\hspace{0.5cm}$ $\sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \tau_j \upsilon_{ij} w_{\omega ij}^{(7)}$ $\hspace{0.5cm}$ $+$

$\sum_{i \in \mathcal{I}} \sum_{m \in \mathcal{M}} c_m^{rescue} \vartheta_{mi} w_{\omega im}^{(10)} + \sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} \left(\sum_{h \in \mathcal{H}} r_{ihk} \right) w_{\omega ik}^{(14)}$.

We can categorize all feasible solutions of the constraints $(EC.3.2) - (EC.3.9)$ into three groups:

$\mathcal{D}'$: These solutions are neither optimal for model $(EC.3.10) - (EC.3.11)$ nor feasible for model $(EC.3.12) - (EC.3.14)$.

$\mathcal{D}^+$: These solutions are optimal for model $(EC.3.10) - (EC.3.11)$, but they are not optimal for model $(EC.3.12) - (EC.3.14)$.

$\mathcal{D}^*$: These solutions are optimal for both models $(EC.3.10) - (EC.3.11)$ and $(EC.3.12) - (EC.3.14)$.

The equality $F(\mathcal{D}^+, \hat{x}, \hat{\vartheta}, \hat{\upsilon}, \hat{r}) = F(\mathcal{D}^*, \hat{x}, \hat{\vartheta}, \hat{\upsilon}, \hat{r})$ arises from both function values being equivalent to the optimal objective value within model $(EC.3.10) - (EC.3.11)$. In the case of $F(\mathcal{D}^+, x^0, \vartheta^0, \upsilon^0, r^0)$ and $F(\mathcal{D}^*, x^0, \vartheta^0, \upsilon^0, r^0)$, the former surpasses the latter due to $\mathcal{D}^*$ emerging as the optimal solution for model $(EC.3.12) - (EC.3.14)$, while $\mathcal{D}^+$ does not achieve optimality, i.e., $F(\mathcal{D}^+, x^0, \vartheta^0, \upsilon^0, r^0) > F(\mathcal{D}^*, x^0, \vartheta^0, \upsilon^0, r^0)$. Moreover, $F(\mathcal{D}', \hat{x}, \hat{\vartheta}, \hat{\upsilon}, \hat{r}) \geq \sec(\hat{x}, \hat{\vartheta}, \hat{\upsilon}, \hat{r})$, as $\mathcal{D}'$ does not represent an optimal solution for model $(EC.3.10) - (EC.3.11)$.

When considering $\mathcal{D}^+$, it is obvious that $F(\mathcal{D}^+, \hat{x}, \hat{\vartheta}, \hat{\upsilon}, \hat{r}) + \mu F(\mathcal{D}^+, x^0, \vartheta^0, \upsilon^0, r^0) > F(\mathcal{D}', \hat{x}, \hat{\vartheta}, \hat{\upsilon}, \hat{r}) + \mu F(\mathcal{D}', x^0, \vartheta^0, \upsilon^0, r^0)$ due to the equality $F(\mathcal{D}^+, \hat{x}, \hat{\vartheta}, \hat{\upsilon}, \hat{r}) = F(\mathcal{D}', \hat{x}, \hat{\vartheta}, \hat{\upsilon}, \hat{r})$, along with the fact that $F(\mathcal{D}^+, x^0, \vartheta^0, \upsilon^0, r^0) > F(\mathcal{D}', x^0, \vartheta^0, \upsilon^0, r^0)$ and $\mu > 0$. As a result, $\mathcal{D}^+$ does not achieve optimality in the subsequent model, which is equivalent to model $(47) - (48)$.

$$\min \sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}} \left(-d_{pl}^{low}(\omega) w_{\omega pl}^{(1)} - d_{pl}^{high}(\omega) w_{\omega pl}^{(2)} \right) + \sum_{h \in \mathcal{H}} c_h^{hospital}(\omega) w_{\omega h}^{(5)} + \sum_{h \in \mathcal{H}} \sum_{m \in \mathcal{M}} c_m^{rescue} n_{mh}^{hospital} w_{\omega mh}^{(9)}$$

$$+\sum_{i\in\mathcal{I}}c_i^{ACF}\hat{x}_i w_{\omega i}^{(6)}+\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\tau_j\hat{\upsilon}_{ij}w_{\omega ij}^{(7)}+\sum_{i\in\mathcal{I}}\sum_{m\in\mathcal{M}}c_m^{rescue}\hat{\vartheta}_{mi}w_{\omega im}^{(10)}+\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}}\Big(\sum_{h\in\mathcal{H}}\hat{r}_{ihk}\Big)w_{\omega ik}^{(14)}$$

$$+\mu\Bigg(\sum_{p\in\mathcal{P}}\sum_{l\in\mathcal{L}}\Big(-d_{pl}^{low}(\omega)w_{\omega pl}^{(1)}-d_{pl}^{high}(\omega)w_{\omega pl}^{(2)}\Big)+\sum_{h\in\mathcal{H}}c_h^{hospital}(\omega)w_{\omega h}^{(5)}+\sum_{h\in\mathcal{H}}\sum_{m\in\mathcal{M}}c_m^{rescue}n_{mh}^{hospital}w_{\omega mh}^{(9)}$$

$$+\sum_{i\in\mathcal{I}}c_i^{ACF}x_i^0 w_{\omega i}^{(6)}+\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\tau_j\upsilon_{ij}^0 w_{\omega ij}^{(7)}+\sum_{i\in\mathcal{I}}\sum_{m\in\mathcal{M}}c_m^{rescue}\vartheta_{mi}^0 w_{\omega im}^{(10)}+\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}}\Big(\sum_{h\in\mathcal{H}}r_{ihk}^0\Big)w_{\omega ik}^{(14)}\Bigg)$$

$$(EC.3.15)$$

Subject to:

Constraints $(EC.3.2)$ to $(EC.3.9)$ $\qquad\qquad (EC.3.16)$

When considering $\mathcal{D}'$, it is evident that $\sec(\hat{x},\hat{\vartheta},\hat{\upsilon},\hat{r})>F(\mathcal{D}^*,\hat{x},\hat{\vartheta},\hat{\upsilon},\hat{r})+\mu F(\mathcal{D}^*,x^0,\vartheta^0,\upsilon^0,r^0)$ due to the equality $F(\mathcal{D}^*,x^0,\vartheta^0,\upsilon^0,r^0)=\mathrm{opt}(x^0,\vartheta^0,\upsilon^0,r^0)$, $F(\mathcal{D}^*,\hat{x},\hat{\vartheta},\hat{\upsilon},\hat{r})=\mathrm{opt}(\hat{x},\hat{\vartheta},\hat{\upsilon},\hat{r})$, and the condition $0<\mu<\frac{\sec(\hat{x},\hat{\vartheta},\hat{\upsilon},\hat{r})-\mathrm{opt}(\hat{x},\hat{\vartheta},\hat{\upsilon},\hat{r})}{\mathrm{opt}(x^0,\vartheta^0,\upsilon^0,r^0)}$. Since $F(\mathcal{D}',\hat{x},\hat{\vartheta},\hat{\upsilon},\hat{r})+\mu F(\mathcal{D}',x^0,\vartheta^0,\upsilon^0,r^0)>\sec(\hat{x},\hat{\vartheta},\hat{\upsilon},\hat{r})$, it follows that $F(\mathcal{D}',\hat{x},\hat{\vartheta},\hat{\upsilon},\hat{r})+\mu F(\mathcal{D}',x^0,\vartheta^0,\upsilon^0,r^0)>F(\mathcal{D}^*,\hat{x},\hat{\vartheta},\hat{\upsilon},\hat{r})+\mu F(\mathcal{D}^*,x^0,\vartheta^0,\upsilon^0,r^0)$. Consequently, $\mathcal{D}'$ also does not represent an optimal solution for model $(47)-(48)$.

In the case where a solution does not meet optimality criteria for model $(EC.3.12)-(EC.3.14)$, it consequently fails to attain optimality in model $(47)-(48)$ as well. This implies that a solution achieving optimality in model $(47)-(48)$ is an optimal solution in model $(EC.3.12)-(EC.3.14)$, representing a Pareto-optimal solution.

EC.4. Parameter Settings

In this section, the remaining parameters utilized to generate our instances are presented and listed in Table EC.4.

Table EC.4.1: Parameters Used for Instance Generation.

Parameters	Value
Probability of occurrence of outcome ω	$\phi^\omega=\frac{1}{|\Omega|}$
The bed capacity of ACF i	$c_i^{ACF}=U[400,1200]$
Fixed cost of opening an ACF at point i	$f_i=c_i^{ACF}\times 1.3\times 25$
The bed capacity of hospital h when the outcome ω is observed	$c_h^{hospital}(\omega)=U[1000,2000]$
The inventory capacity for non-renewable resource k at ACF i	$c_{ik}^{Inv}=c_i^{ACF}\times U[7,9]$
The maximum supply capacity of non-renewable resource k from hospital h	$c_{hk}^{Supply}=c_h^{hospital}(\omega)\times U[14,16]$
Shortage cost of the unsatisfied high-priority demand type p at location l	$\rho_{pl}^{high}=\$150,000$
Shortage cost of the unsatisfied low-priority demand type p at location l	$\rho_{pl}^{low}=\frac{\rho_{pl}^{high}}{2}$

EC.5. Accelerators' Marginal Impact (Single Version)

This section summarizes the main algorithm's performance in terms of solution time and gap percentage when we remove each one of the accelerators or add the LDD. However, due to a high gap in the algorithm's performance without the LBF in most cases or with the LDD in the
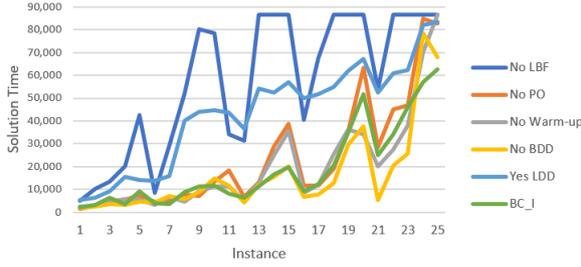
Figure EC.5.1   Solution Times for Marginal Impacts.
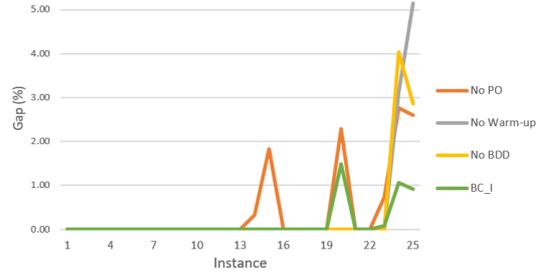
Figure EC.5.2   Gaps for Marginal Impact.

largest-sized example, we have excluded their results from Figure EC.5.2 to preserve the figure's scale. For solution times, please refer to Figure EC.5.1, and for gap percentages, please refer to Figure EC.5.2.

EC.6.   Accelerators' Impact

We need to analyze the individual contributions of each strategy within the combination. After identifying the specific roles of each strategy, we can categorize them into groups and combine enhancements only from different groups. By Pareto-optimal cuts, we are improving the quality of each cut. Almost the same thing happens for the BDD and LDD. In those accelerators, again, we are trying to generate strengthened (high-quality) cuts. Therefore, if we put both Pareto-optimal cuts and BDD together, they both aim to enhance the quality of the cuts we add to the MP. On the other hand, partial Benders decomposition diversifies the cuts by adding LBF valid inequalities in addition to optimality cuts. By incorporating LBF valid inequalities, the emphasis shifts toward enhancing the diversity of the overall lower bound rather than solely concentrating on cut optimization. In the warm-up process, we are generating a set of cuts rapidly. So, if we do not use the warm-up process at the very beginning steps, there will be no cuts; therefore, there will be no information regarding the SPs. By solving the linear relaxation of the MP and generating cuts quickly, the warm-up process provides a starting point for the search, which would otherwise be a blind search of the feasible domain of the MP. In a nutshell, we have the following classes:

- Strengthening the cuts (PO cuts, BDD, and LDD)
- Diversifying the cuts (LBF valid inequalities)
- Improving the quality of the MP (Warm-up)

There are three different combinations of strategies from various classes. In Table EC.6.2, we have four main columns. In the first one, we present the problem information, including the number of ACFs and demand locations; then, in the second column, we evaluate the combination of warm-up, LDD, and LBF valid inequalities. The third and fourth columns combine warm-up and LBF valid inequalities with BDD and PO cuts, respectively. For each algorithm, we provided the number of evaluated nodes in the search tree, the solution time, the lower and upper bound, and the gap. Our analysis revealed that the two algorithms, $BC_{(WarmUp+BDD+LBF)}$

Table EC.6.2: Impact of Different Combinations for the Enhancements.

Problem Size		$BC_{(WarmUp+LDD+LBF)}$					$BC_{(WarmUp+BDD+LBF)}$					$BC_{(WarmUp+PO+LBF)}$				
# ACF	# De.Loc.	Nodes No.	Time (sec)	LB	UB	Gap (%)	Nodes No.	Time (sec)	LB	UB	Gap (%)	Nodes No.	Time (sec)	LB	UB	Gap (%)
10	10	6,974	7,147	71,294	71,294	0.00	4,570	1,535	71,294	71,294	0.00	5,072	1,775	71,294	71,294	0.00
	15	3,467	7,281	84,936	84,936	0.00	5,302	2,570	84,936	84,936	0.00	6,815	2,749	84,936	84,936	0.00
	20	10,422	10,206	112,396	112,396	0.00	6,272	4,785	112,396	112,396	0.00	4,579	3,493	112,396	112,396	0.00
	25	4,655	13,400	134,560	134,560	0.00	6,292	5,415	134,560	134,560	0.00	797	3,754	134,560	134,560	0.00
	30	2,111	15,907	152,864	152,864	0.00	16,040	6,785	152,864	152,864	0.00	2,057	4,866	152,864	152,864	0.00
15	10	9,191	23,562	70,144	70,144	0.00	8,353	5,080	70,144	70,144	0.00	3,922	4,451	70,144	70,144	0.00
	15	4,257	20,994	86,515	86,515	0.00	8,235	4,927	86,515	86,515	0.00	2,999	7,116	86,515	86,515	0.00
	20	5,956	38,658	106,614	106,614	0.00	10,847	7,695	106,614	106,614	0.00	3,295	5,442	106,614	106,614	0.00
	25	6,469	47,246	129,213	129,213	0.00	3,704	6,888	129,213	129,213	0.00	3,042	8,784	129,213	129,213	0.00
	30	7,485	47,985	148,799	148,799	0.00	9,468	12,254	148,799	148,799	0.00	15,529	14,335	148,799	148,799	0.00
20	10	6,611	50,970	89,341	89,341	0.00	8,095	18,033	89,341	89,341	0.00	12,284	11,257	89,341	89,341	0.00
	15	2,713	45,820	81,913	81,913	0.00	2,421	5,874	81,913	81,913	0.00	3,355	3,745	81,913	81,913	0.00
	20	8,194	52,630	135,098	135,098	0.00	9,684	13,134	135,098	135,098	0.00	6,234	13,233	135,098	135,098	0.00
	25	11,638	60,283	129,049	129,400	0.28	18,116	31,036	128,069	129,538	1.21	4,538	16,099	129,052	129,052	0.00
	30	20,788	65,284	186,354	189,248	1.62	15,700	41,793	187,371	190,568	1.80	4,280	21,340	189,248	189,248	0.00
25	10	4,384	48,712	62,969	62,969	0.00	3,413	13,028	62,969	62,969	0.00	4,179	7,597	62,969	62,969	0.00
	15	3,993	49,962	81,884	81,884	0.00	6,959	13,093	81,884	81,884	0.00	3,660	8,319	81,884	81,884	0.00
	20	11,694	53,187	103,774	103,774	0.00	6,753	15,672	103,774	103,774	0.00	3,261	12,422	103,774	103,774	0.00
	25	6,708	66,920	125,870	127,877	1.60	7,569	32,113	127,287	127,287	0.00	7,515	23,200	127,287	127,287	0.00
	30	9,558	74,476	142,147	146,589	3.27	10,849	66,497	143,008	146,690	2.63	11,297	38,854	146,319	146,319	0.00
30	10	1,743	52,610	65,726	65,726	0.00	5,596	20,728	65,726	65,726	0.00	925	5,936	65,726	65,726	0.00
	15	6,396	62,358	82,406	82,406	0.00	7,981	37,651	82,406	82,406	0.00	4,922	22,669	82,406	82,406	0.00
	20	5,870	66,098	101,074	102,214	1.16	12,093	50,627	101,565	102,303	0.73	11,231	23,049	102,028	102,028	0.00
	25	9,974	86,400	118,216	128,532	8.77	12,022	83,913	121,968	125,892	3.26	22,388	77,403	121,629	126,280	3.81
	30	8,935	86,400	151,794	160,853	5.80	15,136	81,232	155,541	159,740	2.68	14,488	67,580	155,061	159,273	2.73
Average		7,207	46,180	110,198	111,406	0.90	8,859	23,294	110,610	111,299	0.49	6,507	16,379	110,843	111,197	0.26

and $BC_{(WarmUp+PO+LBF)}$, performed similarly regarding the gap, but both outperformed the algorithm with LDD. However, in terms of solution time, $BC_{(WarmUp+PO+LBF)}$ outperformed $BC_{(WarmUp+BDD+LBF)}$, likely due to its use of linear SPs instead of MIP ones.

EC.7. Sensitivity Analysis and Managerial Implications

In this section, we evaluate how the flexibility of the ACF in treatment, the distances between demand points, hospitals, and ACFs, and variations in demand can affect the objective function. We also explore ways to manage chaotic situations after a disaster by making informed decisions about ACFs. To perform the sensitivity analysis, we have a region with 20 demand points, 20 potential locations for ACFs, 5 hospitals, and 3 types of vehicles. We uniformly select the demand and hospital capacity from the range of $[100, 200]$ and $[500, 1000]$, respectively. All other parameters used for evaluating the algorithms' performance in Section 6 remain the same.
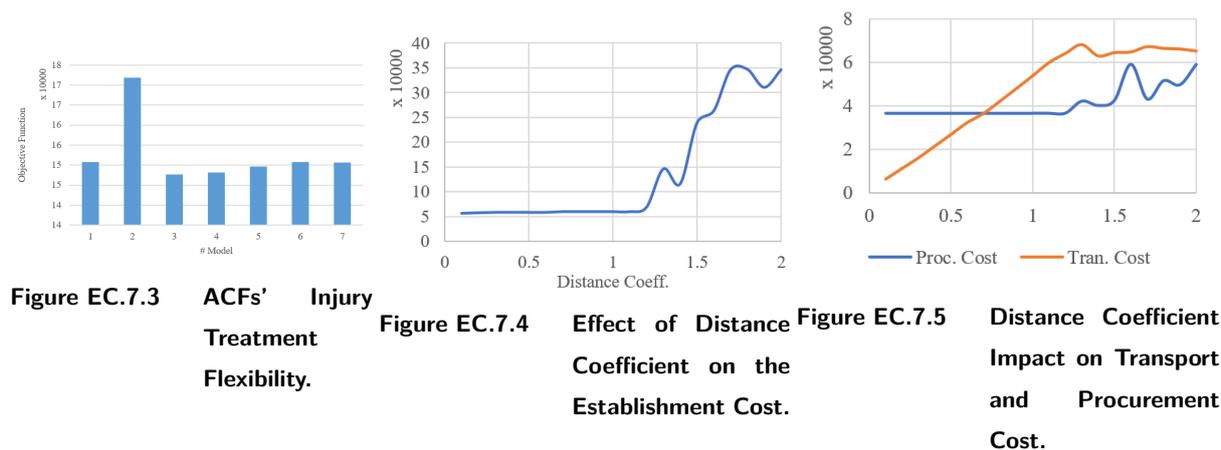
– *Importance of ACFs' Flexibility*

In our initial assumption, we randomly assigned the set of injuries that each ACF could treat (model #1 in Figure EC.7.3). However, we were curious to explore the impact of the following alternative scenarios:

- Model #2 in Figure EC.7.3, where ACFs were classified into different classes, and each class could only treat one type of injury.
- Model #3 in Figure EC.7.3, where all ACFs were allowed to treat all injuries.
- Model #4 in Figure EC.7.3, where each ACF was randomly assigned a set of injuries, but all ACFs could treat injury #1.
- Model #5 in Figure EC.7.3, where each ACF was randomly assigned a set of injuries, but all ACFs could treat injury #2.
- Model #6 in Figure EC.7.3, where each ACF was randomly assigned a set of injuries, but all ACFs could treat injury #3.
- Model #7 in Figure EC.7.3, where each ACF was randomly assigned a set of injuries, but all ACFs could treat injury #4.

Figure EC.7.3 illustrates that the minimum objective function is achieved when we give maximum freedom to the ACFs by enabling them to treat all patients. However, focusing on only one type of injury at each ACF (model #2) may lead to a terrible decision. Additionally, suppose we cannot treat all injuries in all ACFs due to a lack of physicians. In that case, it is better to randomly assign all injuries to ACFs but ensure that all ACFs treat the injury with the highest probability of occurrence (e.g., injury #1 in our case).

– *Sensitivity Analysis on Distances*

When the distance coefficient between points is less than 1.3, indicating that the points are close to each other or the roads are in good condition, the model may choose to transfer patients from the demand locations to ACFs to stabilize their situation. This is done before transporting them from ACFs to hospitals for complete treatment. This approach results in

Figure EC.7.3 **ACFs'** **Injury Treatment Flexibility.**

Figure EC.7.4 **Effect of Distance Coefficient on the Establishment Cost.**

Figure EC.7.5 **Distance Coefficient Impact on Transport and Procurement Cost.**

a reduced number of ACFs, but it also leads to an increase in transportation costs (as shown in Figure EC.7.5). When the distance coefficient exceeds 1.3, which is a common occurrence after any disaster due to road blockages from debris, damage, or other obstacles, both establishment and procurement costs increase (as shown in Figure EC.7.4 and Figure EC.7.5). The procurement cost increases because we need to equip more ACFs with the required non-renewable resources. Therefore, as we increase the number of ACFs and treat more patients, we also need to procure more non-renewable resources. Additionally, the increasing trend in total costs by increasing the distance coefficient from 0.1 to 2 underscores the importance of debris clearance after disasters. Blocked or damaged roads can make it difficult or impossible for emergency responders and aid workers to reach affected communities, hindering their ability to provide timely assistance to those in need.

– *Sensitivity Analysis on Demand's Variations*

To understand how changes in demand affect the objective function, it is essential to know the capacities of the ACFs and the types of injuries they can treat. You can find this information in Table EC.7.3. For instance, ACF #1 can accommodate 497 patients and only treat nerve injuries (#3) and burns (#4). Among all the ACFs, seven have a capacity of over 1000 patients, making them the most expensive to establish. However, ACFs #2 and #6 are unable to treat fractures (#1) and lacerations (#2), which are common injuries after an earthquake, affecting approximately 96% of patients. This makes them inefficient, and the model does not recommend selecting them as part of its optimal solution. Nonetheless, if the demand increases for any reason, the model will establish large but effective ACFs to ensure no unmet demands occur. The rise in establishment costs shown in Figure EC.7.4 results directly from this decision.

Based on the data presented in Figure EC.7.5, we can see that once the demand for hospital beds reaches 1.5 times its initial value, the transportation cost stabilizes, indicating no significant changes. This is because the hospitals' bed capacity is full, and we cannot transfer any more patients from the ACFs. To treat as many patients as possible in ACFs, we establish more of them and provide more non-renewable resources. However, if the demand

Table EC.7.3   ACFs' Capacities and Treatment Abilities.

#ACF	Capacity	Treatable Injuries				#ACF	Capacity	Treatable Injuries			
		#1	#2	#3	#4			#1	#2	#3	#4
1	497			×	×	11	401	×			
2	1082			×	×	12	572			×	
3	416	×	×		×	13	739	×	×	×	×
4	715	×	×	×	×	14	471	×	×	×	
5	1101	×				15	1196	×		×	×
6	1091			×		16	421			×	
7	1157			×	×	17	594		×	×	
8	461	×	×	×		18	974		×		×
9	1152	×	×	×	×	19	1046	×	×		×
10	638	×	×	×	×	20	789	×			

coefficient exceeds 2, even the ACFs have no available beds, resulting in a steady procurement cost beyond that point.

EC.8.   L-Shaped Algorithm Performance

In this section, we assess the performance of the L-shaped algorithm with and without the inclusion of LBF valid inequalities. It is important to highlight that both versions of the algorithm incorporate the generation of PO cuts and the warm-up process. The results for the CRP solved using the L-shaped algorithm are presented in Table EC.8.4. Each row in this table represents the average results across five instances. As discussed in Section 6, the key distinction between the L-shaped and B&C algorithms lies in how they handle the implementation of the master problem. The L-shaped algorithm, which is a pure cutting-plane method applied to our stochastic mathematical model, solves the master problem optimally at each iteration, while the B&C algorithm requires solving the master problem only once.

Table EC.8.4 presents the instance information in the "Problem Size" column, followed by the performance results for the L-shaped algorithm without and with the application of the LBF, listed under "L-shaped Algorithm without LBF" and "L-shaped Algorithm with LBF," respectively. The table also includes additional columns such as "Iter.," "Time (sec)," "LB," "UB," and "Gap (%)," presenting the number of iterations, solution time, lower bound, upper bound, and optimal gap in percentage, respectively, for each algorithm. The "# ACF" and "# De.Loc." columns provide the count of potential locations for the ACFs and the number of demand locations, respectively.

While the the inclusion of LBF valid inequalities within the L-shaped algorithm led to a significant reduction in the optimality gap and the required number of iterations, from approximately 722% to nearly 478% and from 170 to 37, respectively, the remaining gap still falls outside the acceptable range. Consequently, we chose to proceed with the B&C algorithm for the subsequent sections of this paper.

Table EC.8.4: Computational Results of the L-Shaped Algorithm for the CRP with Stochastic Demand and Hospital Bed Capacity.

Problem Size		L-shaped Algorithm without LBF					L-shaped Algorithm with LBF				
# ACF	# De.Loc.	Iter.	Time (sec)	LB	UB	Gap (%)	Iter.	Time (sec)	LB	UB	Gap (%)
10	10	118	31,651	70,937	71,408	0.67	24	6,501	70,955	71,352	0.57
	15	165	52,910	83,894	85,007	1.32	22	20,268	84,657	84,914	0.30
	20	197	61,351	109,528	116,237	6.50	44	48,111	111,973	112,658	0.61
	25	230	82,553	125,755	156,037	24.43	33	18,783	133,927	134,644	0.53
	30	236	86,400	131,754	189,684	45.47	51	30,581	152,403	153,188	0.51
15	10	163	66,833	67,844	73,866	8.84	45	33,884	69,680	70,468	1.07
	15	189	76,518	79,395	115,183	48.33	33	27,985	86,138	86,644	0.59
	20	194	86,400	87,910	206,204	143.06	41	45,071	105,777	110,158	4.12
	25	196	86,400	99,072	203,782	107.13	60	74,365	127,714	135,059	5.67
	30	188	86,400	113,773	372,035	229.89	54	70,857	147,322	173,758	17.86
20	10	171	83,708	81,175	190,021	137.21	51	64,163	88,361	94,837	6.66
	15	190	86,400	73,362	96,632	31.06	26	37,361	81,463	81,907	0.54
	20	167	86,400	101,853	392,503	289.21	41	76,871	132,324	843,526	530.06
	25	172	86,400	92,637	887,438	868.07	48	83,757	127,519	253,845	96.10
	30	169	86,400	134,911	1,463,184	976.90	52	86,400	184,122	1,811,799	896.17
25	10	157	81,332	57,259	102,664	78.46	19	41,402	61,679	64,004	3.30
	15	150	86,400	65,684	133,010	106.01	24	57,574	81,535	81,916	0.47
	20	159	86,400	75,021	392,523	426.89	35	75,965	101,121	1,537,159	1415.06
	25	157	86,400	90,058	811,981	800.93	34	83,613	124,002	2,295,585	1733.92
	30	139	86,400	105,158	1,632,509	1441.97	38	86,400	142,597	2,570,940	1723.28
30	10	146	86,400	56,300	92,383	64.40	26	48,779	64,778	103,855	60.10
	15	146	86,400	60,945	285,785	371.37	28	77,372	80,299	800,762	921.07
	20	153	86,400	74,639	939,314	1178.10	32	83,797	99,874	942,987	827.60
	25	139	86,400	87,624	4,629,815	5216.96	33	86,400	120,730	1,867,882	1453.81
	30	152	86,400	114,394	6,097,654	5460.21	35	86,400	154,719	3,613,042	2270.51
Average		170	80,226	89,635	789,474	722.54	37	58,106	109,427	723,876	478.82