Transportation Research Part E 186 (2024) 103531

Contents lists available at ScienceDirect

Transportation Research Part E

journal homepage: www.elsevier.com/locate/tre

A machine learning based sample average approximation for supplier selection with option contract in humanitarian relief

Shaolong Hu [a,b], Zhijie Sasha Dong [c,*], Rui Dai [d]

[a] School of Economics and Management, Southwest Jiaotong University, Chengdu 610031, China
[b] Key Laboratory of Service Science and Innovation of Sichuan Province, Chengdu 610031 China
[c] Department of Construction Management, Cullen College of Engineering, University of Houston University of Houston, Houston, TX 77204, United States
[d] Norfolk Southern, 650 W Peachtree St NE, Atlanta, GA 30308, United States

ARTICLE INFO

Keywords:
Emergency logistics
Supplier selection
Option contract
Stochastic programming
Sample average approximation

ABSTRACT

The humanitarian relief plays an important role in reducing the impact of disasters and avoiding humanitarian crises. As one of the essential activities, selecting a series of proper suppliers is particularly helpful for a successful and efficient disaster response to provide victims with necessary supplies. To optimize the benefits of the relief agency, victims, and suppliers, this paper proposes a supplier selection problem with consideration of the option contract, in which all-unit quantity and incremental quantity discounts are integrated. The problem is formulated as a multi-objective stochastic programming model with the objectives of minimizing the cost of the relief agency and maximizing the profit of suppliers, which are two conflict objectives, and also reducing the shortage risk for victims. Moreover, a machine learning based sample average approximation (SAA) is designed to solve the proposed model in large-scale cases. Specifically, stratified random sampling is integrated into K-means++ to improve quality of samples. The numerical analysis demonstrates that the proposed strategy can achieve a win–win situation for the relief agency, victims, and suppliers. It also justifies the efficiency of applying the machine learning method to enhance SAA for solving large-scale stochastic programming models.

1. Introduction

Over the last decades, the rate and impact of natural disasters have remarkably increased, affecting millions of people's lives, resulting in thousands of deaths, and causing a substantial economic loss worldwide. Recent events, such as Hurricane Harvey and Irma in 2017 and the Indonesia earthquakes and tsunami in 2018, have caused catastrophic damages to the disaster-affected areas (Amadeo, 2018; Pinelli et al., 2018). Therefore, humanitarian relief after a disaster is critical to reduce the impact of disasters and avoid the humanitarian crisis. Lack of relief supplies would result in human suffering and loss of life. Hence, a stable and responsive humanitarian relief supply chain needs to be developed to ensure the efficient supply of relief products in the disaster-affected area (Hu and Dong, 2019). In order to design an efficient relief supply chain, humanitarian agencies need to cooperate with suppliers (e.g., manufacturers and retailers) to guarantee a stable supply before and after a disaster occurs by making an effective plan to select the proper suppliers to purchase relief supplies on the preparedness stage.

* Corresponding author.
E-mail addresses: shaolong.hu@swjtu.edu.cn (S. Hu), sasha@central.uh.edu (Z.S. Dong).

https://doi.org/10.1016/j.tre.2024.103531
Received 1 March 2023; Received in revised form 5 April 2024; Accepted 7 April 2024
Available online 22 April 2024
1366-5545/

Similar to a commercial supply chain, a relief supply chain contains the relief agency (purchaser) and the suppliers. Relief and commercial supply chains have some similarities, but some fundamental differences exist in terms of such as strategic goals, customer and demand characteristics. We focus on these two aspects. The strategic goal of humanitarian relief is to minimize economic costs and shortage risks simultaneously (Sabbaghtorkan et al., 2020; Hu et al., 2022). As demand is highly uncertain, a large amount of procurement may lead to financial loss when demand is low for a long time; while shortage risks (i.e., human suffering and loss of life) will be enormous when demand is greater than procurement quantity. A main strategic goal in commercial supply chain is to maximize profit. We thus design an objective function to consider the suppliers' benefits that can motivate them to take a more active part in humanitarian relief (Shokr and Torabi, 2017; Dabbagh et al., 2018). In addition, demand characteristics are another fundamental difference. As the location, occurrence time and severity level of disasters are unpredictable, demand quantities and types are highly uncertain. We propose a scenario-based stochastic programming model to cope with uncertainties.

Relief agencies usually entail high costs associated with supplier selection, such as minimum commitment quantity or standby fee (Department of Homeland Security., 2014); whether disasters strike or not. The option contract is an effective approach to dealing with the above situations. Option is a kind of derivatives, and it has been used as an effective tool to avoid risks and reduce uncertainty in finance (Liang et al., 2012). In humanitarian relief, an option contract allows the relief agency to get the right to buy relief supplies at a fixed price within reserved quantity by prepaying the premium before the expiration date. This right is not the obligation (Aghajani et al., 2020). In other word, the relief agency may not exercise the option contract. This may be happening for non-seasonal disasters, e.g., earthquakes, because it may happen only once in many years. The option contract endows the relief agency with the flexibility to manage large amount of relief supplies with few funds to reduce delay and shortage risks of supply. Therefore, the option contract is a perfect fit to improve humanitarian relief.

The option contract includes two critical elements: the option price and the exercise price (Arani et al., 2016). The option price is an allowance paid by the relief agency to a supplier for reserving one unit of the supplier's capacity before a disaster occurs. The relief agency also has the right to exercise the option contract after the actual demand is realized and purchase the supplies with the exercise price after a disaster strikes. Moreover, the option contract usually applies quantity discounts, allowing the purchaser to allocate large orders to multiple suppliers with different quantity discount types (Zheng et al., 2019). In this paper, we consider the all-unit quantity discount and the incremental quantity discount. The all-unit quantity discount refers to a discount that lowers a purchaser's wholesale price on every unit purchased, while the incremental quantity discount only applies to items exceeding a certain order quantity. All-unit quantity discount ensures the purchasers buy items at the lowest cost, while the incremental quantity discount holds the profit for suppliers.

Over the past few years, studies in the area of supplier selection in humanitarian relief have considered reducing the purchasing cost and inventory-related risk for purchasers by using option contracts. However, suppliers incur additional costs when stockpiling supplies for relief agencies. The supplier has the motivation not to keep relief supplies if profit is insufficient from the option contract. We thus propose to protect the supplier's benefits during the supplier selection process. In order to create a sustainable humanitarian supply chain, it is essential to consider and address the objectives of every participant involved in this chain, i.e., the financial implications for relief agencies, the vulnerabilities related to shortages faced by those in need, and the profitability concerns of relief suppliers. This paper proposes a multi-objective two-stage stochastic programming (SP) model with the objectives of minimizing the total expected cost and unmet demand, and maximizing the total expected profit for suppliers. The proposed model ensures the profits to attract the suppliers.

In order to ensure the applicability of the proposed model in large-scale cases, a machine learning (ML) based sample average approximation (SAA) is proposed. SAA is an approach for solving stochastic optimization problems by using Monte Carlo simulation (Pagnoncelli et al., 2009). However, SAA is time-consuming and inefficient when solving large-scale SP problems, which should generate samples with a large number of scenarios. A ML based computational framework is proposed in this paper to generate optimized samples effectively for SAA for solving large-scale SP problems. Specifically, K-means++ (Kapoor and Singhal, 2017) and stratified random sampling are employed to group scenarios and select scenarios from each cluster. Our experimental results demonstrate that the proposed method can find provably near-optimal solutions using only moderate computation time.

In sum, the key differences of this paper can be outlined as follows:

- Introduction of a pioneering option contract tailored for humanitarian relief. This contract integrates two distinct quantity discount models, aiming to harmonize potential conflicts between relief agencies and suppliers while optimizing both relief supplier selection, order allocation and distribution of relief supplies.
- Development of a multi-objective SP model specifically designed for supplier selection. This model aims to optimize various objectives, including reducing anticipated costs for relief agencies, mitigating shortages for those in need, and ensuring supplier profits amidst uncertain relief demand.
- Introduction of a machine learning-based SAA algorithm. Specifically, the algorithm leverages stratified random sampling and K-means++ to generate high-quality samples, ultimately enhancing computational efficiency of SAA in addressing these complex models.

The paper is organized as follows. In Section 2, we investigate gaps in supplier selection and SAA by reviewing the relevant literature. Section 3 describes the problem and presents a two-stage stochastic programming model. In Section 4, a solution methodology that integrates machine learning techniques and SAA is proposed. Section 5 conducts the numerical analysis to verify the applicability of the proposed model and algorithm. Finally, we conclude our work and discuss future research directions in Section 6.

2. Literature review

This section presents the most relevant literature in three areas: 1) supplier selection, 2) option contract in humanitarian relief, and 3) SAA method, which is the methodology used in this paper.

2.1. Supplier selection

Supplier selection is one of the important processes in supply chain management for the purchaser to identify, evaluate, and contract with suppliers. The topic related to supplier selection in humanitarian relief has attracted many researchers' attention. Interested readers can refer to Hu et al. (Hu et al., 2022); which presented a comprehensive review of recent works about supplier selection in disaster operations management and proposed future directions in this area. In the relief supply chain, purchasers are usually government related agencies, e.g., the Ministry of Civil Affairs in China and the Federal Emergency Management Agency in the United States, and non-government organizations, e.g., the International Federation of Red Cross and Red Crescent Societies. They select suitable suppliers that require the assessment of alternative suppliers based on different criteria such as agreement cost, procurement price, reserve capacity, bonus, commitment quantity, production capacity, delivery time, and supply resilience (Balcik and Ak, 2014; Hammami et al., 2014; Torabi et al., 2015; Hu et al., 2017; Olanrewaju et al., 2020; Kaur and Singh, 2021; Thevenin et al., 2022). Among them; agreement cost, procurement price, and reserve capacity are usually different with suppliers, which are mostly regarded for supplier selection. These three characteristics are integrated into the proposed model as well.

Overall, the above criteria adopted for supplier selection focus on the interests of relief agencies, such as minimizing the cost of procurement, reducing the risk of relief supply shortage, and optimizing inventory and distribution of relief supplies. Thus, the research objectives are mainly to minimize agreement cost with suppliers, procurement cost, transportation cost, holding cost, and penalty cost for relief agencies (Balcik and Ak, 2014; Hammami et al., 2014; Torabi et al., 2015; Hu et al., 2017; Olanrewaju et al., 2020; Kaur and Singh, 2021; Thevenin et al., 2022). Disasters are unpredictable that results in uncertainties in victims' demand and supply capacity of suppliers, etc. Studies on supplier selection usually integrated these uncertainties to optimize such as supplier selection and order allocation decisions (Chen et al., 2022; Shokr et al., 2022; Pamucar et al., 2023; Wang et al., 2023). Relief supplier plays a crucial role in humanitarian relief. When studying supplier selection problems, relief agencies should consider the suppliers' benefits as well to motivate them to take a more active part and achieve a win–win solution ultimately for disaster response. Shokr and Torabi (2017) studied an enhanced reverse auction framework for relief procurement management. In the bid construction phase, they build a mixed integer programming model with the objectives of maximizing profit under limitations in the maximum available budget and allowable delivery times. Dabbagh et al. (2018) studied a multi-attribute reverse auction framework under uncertainty to the procurement of relief items. In bid construction phase; they built a bi-objective fuzzy model that balances the objectives of minimizing lead time and maximizing profit. These studies focused on the optimization of pricing, supply, and transportation decisions from the perspective of suppliers. To the best of our knowledge, how to balance the conflict between the purchasers' cost and suppliers' profit is lacking.

2.2. Option contract

The option contract is originated from financial derivatives, which conveys the right but not the obligation to buy (call option) or sell (put option) an underlying asset during a fixed period at a predetermined exercise or strike price. In recent years, research on option contracts in the context of humanitarian relief has been increasing (Liu et al., 2019; Chen et al., 2022). Wang et al. (2015) determined that pre-purchasing relief supplies from a supplier with an option contract are superior to both pre-purchasing with a buyback contract and instant purchasing with a return contract. Shamsi et al. (2018) developed an option contract for vaccine dose procurement from two suppliers with the objectives of minimizing the procurement and social costs using the SIR epidemic model. Hu et al. (2019) explored the characteristics of the put option contract and proved that it could provide coordination of the relief supply chain. Aghajani et al. (2020) proposed a novel two-period option contract integrated with supplier selection and inventory prepositioning. Ghavamifar et al. (2022) proposed a hybrid relief procurement contract in which the option contract and quantity flexible contract are used to coordinate the supply of a relief item. They proved and demonstrated that option contract improves performance of relief supply chain from two aspects, i.e., decreasing economic costs in supplying relief products; managing risks from demand uncertainty, supply unreliability and price volatility. The previous works on the option contract for humanitarian relief did not consider quantity discount. All-unit quantity discount can help purchasers to save procurement cost, and incremental quantity discount is commonly applied in the commercial supply chain to incent suppliers to participate in supply activities positively (Bohner and Minner, 2017; Nourmohamadi Shalke et al., 2018). To explore the impacts of both discount types on the benefits of relief agencies and suppliers, we propose an option contract model with all-unit and incremental quantity discounts that can provide mutual benefits to both relief agencies and suppliers, as well as conduct the sensitivity analysis for the rate changes of the discount.

2.3. Sample average approximation

SAA is a Monte Carlo simulation-based approach, which is frequently employed to solve SP problems. Existing studies on SAA can be categorized into two groups. The first group is SAA with a fixed sample size. As one of the most popular SAA studies, Kleywegt et al. (2002) presented SAA to solve stochastic discrete optimization problems and evaluates convergence rates, stopping rules, and computational complexity. Verweij et al. (2003) used SAA for stochastic routing problems, in which the computational time to solve the sample problems grows linearly with the sample size. Nur et al. (2021) combine SAA with an enhanced progressive hedging algorithm for biofuel supply chain network design. The secondary category is SAA with a variable sample size. Homem-De-Mello (2003) integrated a variable sample technique into SAA to replace the objective function at each iteration. Pasupathy (2010) proposed a variable sample method to balance sample size and error tolerance in SAA. Royset (2013) developed a discrete-time optimal-control problem by adaptively selecting sample sizes. According to these studies, one can see that compared to the fixed sample size, to some extent, changing the sample size at each iteration can reduce the loss of information incurred by sampling.

Due to the uncertainty in disaster operations management, SAA is a popular solution algorithm to solve the SP model that is developed to handle uncertainty in studies of humanitarian relief. Wang et al. (2022) proposed SAA for EMS location-allocation problem under uncertainties, and integrated a two-phase denders decomposition solution scheme to accelerate computation. Aslan and Çelik (2019) proposed a two-stage stochastic program to support the pre-disaster decisions of warehouse location and item prepositioning under the uncertain relief item demand and availability of roads, and use the SAA scheme as the heuristic solution. Klibi et al. (2018) developed a scenario-based approach with the SP model for the design of relief networks, including three phases: scenario generation, design generation, and design evaluation, and solved the SP model using SAA method.

2.4. *Our contributions*

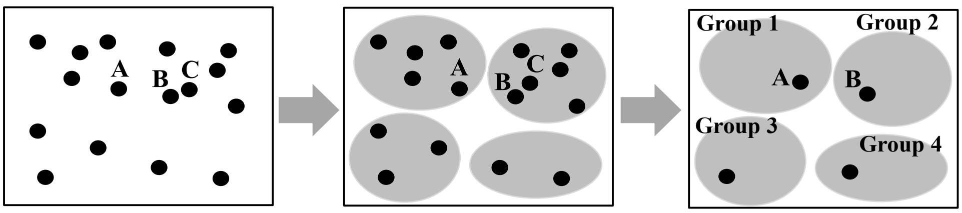
From the modeling perspective, this study addresses a stochastic supplier selection problem considering option contract with two quantity discounts, in which suppliers' profit is subject to a certain value. Although quantity discount has been widely applied in practice, it is never integrated into option contract for supplier selection in humanitarian relief. This study can explore the impacts of quantity discount on the benefits of relief agencies and suppliers. To motivate suppliers to take a more active part in humanitarian relief, the relief agency's cost and suppliers' profit are simultaneously considered in the multi-objective SP model. This strategy can ensure profit for suppliers and achieve a win–win situation under uncertain relief demand in humanitarian relief. Therefore, this work is not a trivial extension to existing studies of supplier selection.

From the methodological perspective, our work is most relevant to Emelogu et al. (2016). The existing SAA studies can achieve good performance on small-scale SP problems, but they are computationally expensive for solving large-scale ones. This is because they only focused on increasing sample size, which cannot guarantee to identify and select the most representative scenarios. To solve this challenge, Emelogu et al. (2016) presented an approach to generate samples. They utilized different clustering techniques (e.g., K-means) to group similar scenarios, then selected one scenario from each group. Compared to the classical SAA, their approach can generate samples more efficiently as shown in Fig. 1. This is because scenario A and scenarios B and C are clustered into two different groups, and scenarios B and C will be in different samples, as shown in Fig. 1 (*b*) and 1 (*c*). However, their approach may be inefficient when the size of groups is different. This is because they only selected one scenario from each group as the representative scenario, even for large groups. A large group contains more information than a small one, so that it might need to select more scenarios to represent it. Therefore, motivated by the analysis mentioned above, our first contribution in solution methodology is to employ a stratified sampling technique to determine scenario number that should be selected in each group. In addition, above-mentioned solution methodologies use sample average function. We propose another strategy that is still one scenario from each cluster, but weight those scenarios with a probability proportional to the cluster's size.

3. Modeling

3.1. *Problem description*

Within the scope of this study, we focus on a scenario involving a solitary relief agency responsible for procuring relief supplies. Additionally, a range of potential suppliers is available for selection as the sources of these much-needed relief provisions. In the pre-disaster phase, the relief agency makes supplier selection decision and signs option contracts with selected suppliers. In option contract, option quantity and its corresponding quantity discount are determined. The option contract requires the relief agency to pay suppliers an option price per unit to reserve an option quantity of relief supplies. In return, the selected suppliers promise to deliver the reserved relief supplies with an exercise price per unit. In the post-disaster phase, the relief agency determines exercise quantity and delivers relief supplies to affected locations for different scenarios. Exercise quantity from a selected supplier should be less than its



(*a*) Full of scenarios (*b*) Scenario grouping (*c*) Scenario selection

Fig. 1. Process of generating samples by using clustering technique (Emelogu et al., 2016).

option quantity. Especially, exercise quantity will be zero if no disaster occurs. Delivered quantity from each supplier should be less than its exercise quantity.

In addition, we integrate quantity discounts into option and exercise prices, in other words, these prices vary with procurement quantity. We discretize quantities and introduce a set representing quantity intervals (i.e., breaks). All-unit quantity discount means that the discount is applied to all procurements, whereas the incremental discount only applies to those quantities which exceed the interval. We introduce an instance to explain the above description as follows. We suppose a supplier offers quantity discounts (assuming all-unit and incremental quantity discounts are the same) with two intervals, i.e., [0, 100], and (100, 200]. The start and end of one interval are denoted as a price break quantity. For example, let option prices be 2, 1.8, and exercise prices are 7, 6.3, corresponding to two intervals.

Case 1: Assuming demand is 50 for the first scenario with a probability of 0.5, 80 for the second scenario with a probability of 0.5, and all demand is satisfied; Option quantity should be 80 that is in the first interval; exercise quantities should be 50 and 80 that are also in the first interval as well. For an all-unit quantity discount, the total cost of the relief agency is $80 \cdot 2 + (50 \cdot 7 \cdot 0.5 + 80 \cdot 7 \cdot 0.5) = 615$; for an incremental quantity discount, the total cost of the relief agency is $80 \cdot 2 + (50 \cdot 7 \cdot 0.5 + 80 \cdot 7 \cdot 0.5) = 615$.

Case 2: Assuming demand is 50 for the first scenario with a probability of 0.5, 150 for the second scenario with a probability of 0.5, and all demand is satisfied; Option quantity should be 150 and is in the second interval; exercise quantities should be 50 and 150 that are in the first interval for the first scenario, in the second interval for the second scenario. For all-unit quantity discount, the total cost of the relief agency is $150 \cdot 1.8 + (50 \cdot 7 \cdot 0.5 + 150 \cdot 6.3 \cdot 0.5) = 917.5$; for incremental quantity discount, the total cost of the relief agency is $(100 \cdot 2 + 50 \cdot 1.8) + [50 \cdot 7 \cdot 0.5 + (100 \cdot 7 + 50 \cdot 6.3) \cdot 0.5] = 972.5$.

The above two cases reveal that all-unit quantity discount outperforms incremental quantity in saving costs for relief agencies when procurement quantity is not in the first interval. Therefore, negotiating for the all-unit quantity discount is beneficial for relief agencies. However, from the perspective of suppliers, it can obtain more profit by offering the incremental quantity discount because this discount only applies to items exceeding a certain order quantity. This type of discount can incent suppliers that build a close relationship with relief agencies for disaster response. Therefore, both types of discounts are integrated into the option contract.

3.2. Multi-objective stochastic programming model

The sets, parameters and decision variables are described in Table 1. Two objective functions are designed by considering different

Table 1
Model notations.

Notations	Explanation
Sets	
I	Set of candidate supplier, indexed by $i \in I$
J	Set of demand locations, indexed by $j \in J$
A	Set of quantity intervals (breaks), indexed by $a \in A$
S	Set of disaster scenarios, indexed by $s \in S$
Parameters	
$D_{j,s}$	Demand of supplies at location j in scenario s
ω_s	Probability of occurrence of scenario s
C_i	Available capacity of supplier i
$H_{i,j}$	Distance between supplier i and location j
F_i	Unit agreement cost with supplier i
$O_{a,i}$	Unit option price of relief supplies at quantity interval a offered by supplier i
$E_{a,i}$	Unit exercise price of relief supplies at quantity interval a offered by supplier i
$[LO_{a,i}, UO_{a,i}]$	Lower and upper quantity breakpoints (for option price) associated with quantity interval a offered by each supplier i
$[LE_{a,i}, UE_{a,i}]$	Lower and upper quantity breakpoints (for exercise price) associated with quantity interval a offered by each supplier i
T	Unit transportation cost of moving supplies from facilities or suppliers to demand locations
P	Unit penalty cost for unsatisfied demand at demand locations
G_i	Unit production cost of relief supplies by supplier i
First-Stage Decision Variables	
x_i	Binary decision variable if a supplier i is selected
$y_{a,i}$	Binary decision variable if executed all-unit quantity discount purchasing supplies at quantity interval a with supplier i
$y'_{a,i}$	Binary decision variable if executed incremental quantity discount purchasing supplies at quantity interval a with supplier i
$z_{a,i}$	Option quantity of relief supplies at quantity interval a from supplier i (all-unit quantity discount)
$z'_{a,i}$	Option quantity of relief supplies at quantity interval a from supplier i (incremental quantity discount)
Second-Stage Decision Variables	
$e_{a,i,s}$	Binary decision variable if executed all-unit quantity discount purchasing supplies at quantity interval a with supplier i in scenario s
$e'_{a,i,s}$	Binary decision variable if executed incremental quantity discount purchasing supplies at quantity interval a with supplier i in scenario s
$q_{a,i,s}$	Exercise quantity of relief supplies at quantity interval a from supplier i in scenario s (all-unit quantity discount)
$q'_{a,i,s}$	Exercise quantity of relief supplies at quantity interval a from supplier i in scenario s (incremental quantity discount)
Auxiliary Variable	
$g_{i,j,s}$	Quantity of relief supplies transported from supplier i to demand location j in scenario s
$u_{j,s}$	Unsatisfied demand at location j in scenario s
$oi_{a,i}$	Option cost at quantity interval a from supplier i (obtained incremental quantity discount)
$ei_{a,i,s}$	Exercise cost at quantity interval a from supplier i in scenario s (obtained incremental quantity discount)

goals of the relief agency and suppliers. The major task of the relief agency is to satisfy victims' demand for relief supplies under a limited budget. As demand is highly uncertain, a large amount of procurement may lead to financial loss when demand is low for a long time. At the same time, shortage risks (i.e., human suffering and loss of life) will be enormous when demand is greater than procurement quantity. Therefore, the relief agency's goal is to avoid large financial losses and shortage risks simultaneously. The objective function (1) of this model is designed to minimize the total expected cost, which includes economic cost and penalty cost. Financial loss is reduced by minimizing economic cost, which includes fixed agreement cost, option cost, exercise cost, and transportation cost. Fixed agreement cost is used to represent those costs associated with managing a contract, such as overhead and coordination costs. Option cost is incurred by reserving pre-disaster relief supplies. Exercise cost is incurred by exercising post-disaster relief supplies. In option and exercise cost functions, the first term is option and exercise cost when the all-unit quantity discount is selected. As an all-unit quantity discount is for every purchased unit, the cost is equal to quantity times price. The second term is option and exercise cost when incremental quantity discount is selected, which is given in constraints (9)-(10) and (14)-(15), respectively. Transportation cost captures the cost of delivering relief supplies from suppliers to affected locations. Shortage risks are decreased by minimizing the penalty cost. It is the most common measurement of shortage risks by capturing the monetary cost when the demand for relief supplies is unsatisfied (Aghajani et al., 2020; Balcik and Ak, 2014).

$$f_1 = \min \quad fc + oc + \sum_s \omega_s \cdot (ec_s + tc_s + pc_s) \tag{1}$$

More details of these costs are as follows:

Fixed agreement cost

$$fc = \sum_i F_i \cdot x_i$$

Option cost

$$oc = \sum_a \sum_i (O_{a,i} \cdot z_{a,i} + oi_{a,i})$$

Exercise cost

$$ec_s = \sum_a \sum_i (E_{a,i} \cdot q_{a,i,s} + ei_{a,i,s}), \quad \forall s$$

Transportation cost

$$tc_s = \sum_i \sum_j T \cdot H_{i,j} \cdot g_{i,j,s}, \quad \forall s$$

(i) Penalty cost

$$pc_s = \sum_j P \cdot u_{j,s}, \quad \forall s$$

The suppliers' goal is to obtain profit. The objective function (2) of this model maximizes the total expected profit, which includes income from the option contract (i.e., *oc* and *ec*), and the cost of producing relief supplies.

$$f_2 = \max \quad oc + \sum_s \omega_s \cdot \left(ec_s - \sum_a \sum_i G_i \cdot \left(q_{a,i,s} + q'_{a,i,s} \right) \right) \tag{2}$$

Constraints of this model are as follows:

Constraint (3) restricts that the option quantity of relief supplies is less than the capacity of selected suppliers. These two types of option quantity cannot be nonzero at the same time because a supplier would offer one type of discount.

$$\sum_a \left(z_{a,i} + z'_{a,i} \right) \leqslant x_i \cdot C_i, \quad \forall i \tag{3}$$

Regardless of which type of option quantity is selected, constraints (4) and (5) restrict that exercise quantity is less than option quantity; and constraints (6) and (7) guarantee that option quantity is consistent with the selected supplier and its quantity interval.

$$\sum_a q_{a,i,s} \leqslant \sum_a z_{a,i}, \quad \forall i, s \tag{4}$$

$$\sum_a q'_{a,i,s} \leqslant \sum_a z'_{a,i}, \quad \forall i, s \tag{5}$$

$$LO_{a,i} \cdot y_{a,i} \leqslant z_{a,i} \leqslant UO_{a,i} \cdot y_{a,i}, \quad \forall a, i \tag{6}$$

$$LO_{a,i} \cdot y'_{a,i} \leqslant z'_{a,i} \leqslant UO_{a,i} \cdot y'_{a,i}, \quad \forall a, i \tag{7}$$

Constraint (8) restricts that 1) relief supplies can only be ordered from the selected supplier; 2) only one quantity interval can be selected; and 3) for all-unit and incremental quantity discounts, only one type of discount can be executed.

$$\sum_a \left(y_{a,i} + y'_{a,i}\right) \leqslant x_i, \quad \forall i \tag{8}$$

As the incremental quantity discount only applies to supplies exceeding a certain order quantity, constraints (9) and (10) calculate option cost when the incremental quantity discount is selected. Constraint (9) calculates option cost within quantity interval 1. For quantity interval a is larger than 1, constraint (10) calculates option cost for each quantity interval.

$$oi_{1,i} = O_{1,i} \cdot z'_{1,i}, \quad \forall i \tag{9}$$

$$oi_{a,i} = \sum_{a' \in \{1,\dots,a-1\}} O_{a',i} \cdot \left(UO_{a',i} - LO_{a',i}\right) \cdot y'_{a,i} + O_{a,i} \cdot \left(z'_{a,i} - UO_{a-1,i} \cdot y'_{a,i}\right), \quad \forall a \geqslant 2, i \tag{10}$$

Constraints (11)-(15) are similar to constraints (6)-(10). Constraints (11) and (12) guarantee that exercise quantity from every supplier is consistent with the selected supplier and its quantity interval for each scenario. Constraint (13) also restricts that 1) relief supplies can only be ordered from the selected supplier; 2) only one quantity interval can be selected; and 3) for all-unit and incremental quantity discounts, only one discount can be executed. Constraints (14) and (15) calculate exercise cost when the incremental quantity discount is selected.

$$LE_{a,i} \cdot e_{a,i,s} \leqslant q_{a,i,s} \leqslant UE_{a,i} \cdot e_{a,i,s}, \quad \forall a, i, s \tag{11}$$

$$LE_{a,i} \cdot e'_{a,i,s} \leqslant q'_{a,i,s} \leqslant UE_{a,i} \cdot e'_{a,i,s}, \quad \forall a, i, s \tag{12}$$

$$\sum_a \left(e_{a,i,s} + e'_{a,i,s}\right) \leqslant x_i, \quad \forall i, s \tag{13}$$

$$ei_{1,i,s} = E_{1,i} \cdot q'_{1,i,s}, \quad \forall i, s \tag{14}$$

$$ei_{a,i,s} = \sum_{a' \in \{1,\dots,a-1\}} E_{a',i} \cdot \left(UE_{a',i} - LE_{a',i}\right) \cdot e'_{a,i,s} + E_{a,i} \cdot \left(q'_{a,i,s} - UE_{a-1,i} \cdot e'_{a,i,s}\right), \quad \forall a \geqslant 2, i, s \tag{15}$$

Constraint (16) computes transported quantity from selected suppliers to demand locations. Constraint (17) computes the unsatisfied demand. Constraints (18) $-$ (20) define decision variables.

$$\sum_j g_{i,j,s} = \sum_a \left(q_{a,i,s} + q'_{a,i,s}\right), \quad \forall i, s \tag{16}$$

$$u_{j,s} = D_{j,s} - \sum_i g_{i,j,s}, \quad \forall j, s \tag{17}$$

$$x_i, y_{a,i}, y'_{a,i}, e_{a,i,s}, e'_{a,i,s} \in \{0, 1\}, \quad \forall a, i, s \tag{18}$$

$$z_{a,i}, z'_{a,i}, q_{a,i,s}, q'_{a,i,s}, g_{i,j,s}, u_{j,s} \in \mathrm{Z}^+, \quad \forall a, i, j, s \tag{19}$$

$$oi_{a,i}, ei_{a,i,s} \geqslant 0, \quad \forall a, i, s \tag{20}$$

Haimes et al. were the first to present the epsilon (ε)-constraint method to solve the multi-objective optimization problem (Haimes et al., 1971). In this method, only one objective will be kept while the rest will be transformed into constraints. The most important objective function is usually optimized while the other objectives are transformed into constraints. In humanitarian relief, one critical goal is to ensure the efficient supply of relief products so that shortage risks due to a lack of relief supplies can be mitigated. Moreover, a large number of pre-positioning relief supplies is financially prohibitive for many relief agencies due to budgetary limitations and uncertain demand. The existing studies thus usually optimize relief decisions (e.g., pre-positioning, procurement, and supplier selection) to save costs and reduce shortage risks for relief agencies (Aghajani et al., 2020; Balcik and Ak, 2014). Therefore, we set objective function f_1 is optimized, and the f_2 is transformed into constraints. The model can be reformulated as follows.

$$f_1 = \min \quad fc + oc + \sum_s \omega_s \cdot (ec_s + tc_s + pc_s)$$

s.t.

$$f_2 \geqslant \varepsilon$$

Constraints (2) ~ (20)

The method works by pre-defining a virtual grid in the objective space and solving different single objective problems constrained to each grid cell. However, it cannot guarantee to reach to efficient Pareto solutions. The augmented $\varepsilon-$constraint method is developed to prevent the generation of weakly efficient Pareto solution (Rezaei-Malek et al., 2016; Jabbarzadeh et al., 2020). The formulation of the augmented $\varepsilon-$constraint method for the proposed model is shown as follows.

$$f_1 = \min \quad fc + oc + \sum_s \omega_s \cdot (ec_s + tc_s + pc_s) + (\varphi \cdot \gamma)$$

s.t.

$$f_2 - \gamma = \varepsilon$$

$$\gamma > 0$$

Constraints (2) ~ (20)

Where φ is a sufficiently small number (usually between 10^{-3} and 10^{-6}) and the augmented term $\varphi \cdot \gamma$ ensures obtaining just an efficient solution for each epsilon vector. In addition, the way to calculate value of ε is as:

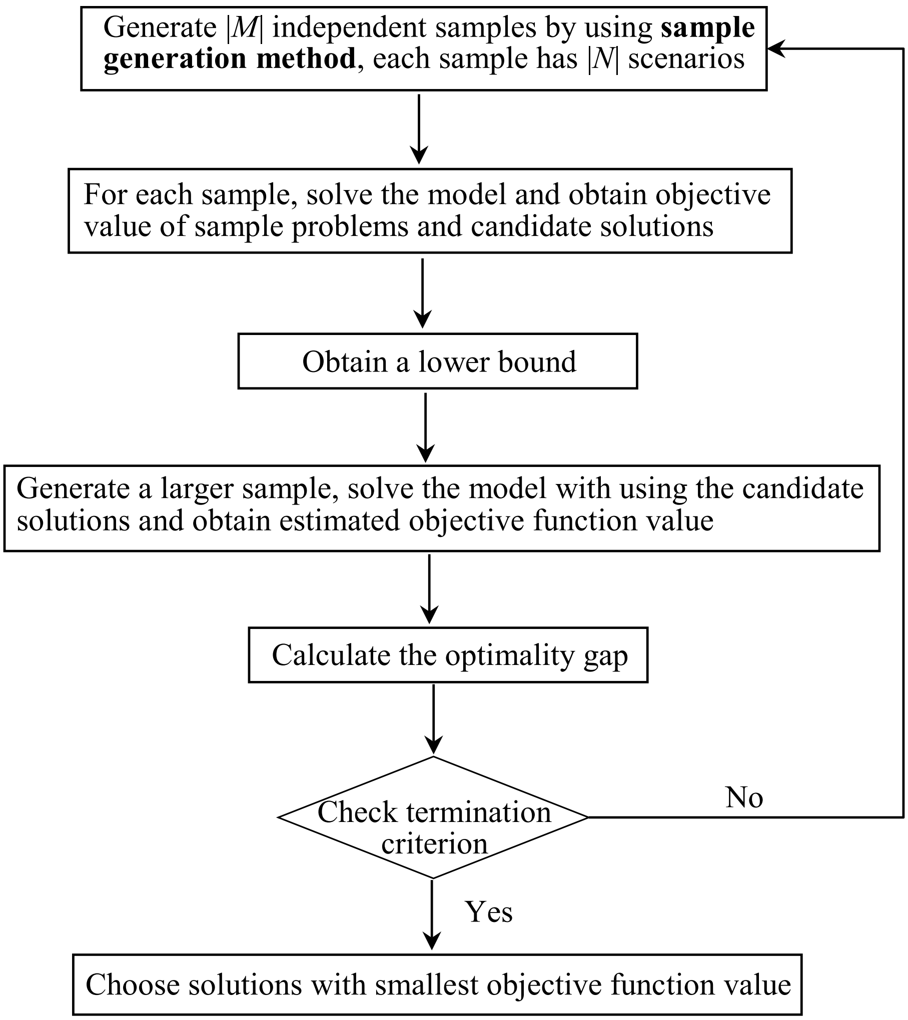


Fig. 2. The framework of ML-based SAA.

$$\varepsilon = f_2^{\max} - \frac{f_2^{\max} - f_2^{\min}}{q} \times l, \qquad l = 0, ..., q-1$$

The different values of ε are calculated by dividing the range of constrained objectives f_2 to q equal intervals. $f_2^{\max}$ and $f_2^{\min}$ can be calculated by optimizing the constrained objective function f_2.

4. Solution methodology

Sample average approximation (SAA) possesses good convergence properties and well-developed statistical methods for validating solutions and conducting error analysis (Birge and Louveaux, 2011). As these advantages, SAA has been successfully utilized to solve large-scale stochastic programming models. In the basic SAA framework, this paper presents a sample generation method to generate samples that can enhance performance of SAA for the proposed stochastic programming model.

In this section, we propose a ML-based SAA for large-scale problems. In our approach, a sample generation method that utilizes K-means++ and stratified sampling techniques is presented to generate samples. K-means++ is employed to reduce the size of stochastic parameter ($D_{j,s}$) into small number of clusters. Stratified sampling is employed to selected most representative scenarios for each cluster. Note that multiple scenarios may be selected for some clusters. Then, the expected value function f_1 is approximated by the corresponding sample average function. For each sample, we solve the proposed model and obtain objective value of sample problems and candidate solutions. At last, the procedure is stopped until optimality gap is less than the specified threshold. Optimality gap is difference between average objective value of sample problems and that of a larger sample problem. The framework of ML-based SAA is shown in Fig. 2.

4.1. Basic SAA

The basic idea of SAA is to randomly generate samples based on the distribution of the data and then approximate the expected value function by the corresponding sample average function. The procedure is stopped until the prespecified termination criterion is satisfied. The implementation procedure of SAA is described below.

- **Step 1:** Generate $|M|$ (i.e., number of samples) independent samples according to the probability distribution, and each sample has $|N|$ (i.e., sample size) scenarios.
- **Step 2:** For $m \in M$, solve the model

$$\widehat{g}^N = \min \quad fc + oc + \frac{1}{|N|}\sum_{s \in N}(ec_s + tc_s + pc_s) + (\varphi \cdot \gamma)$$

s.t.

$$f_2 - \gamma = \varepsilon$$

$$\gamma > 0$$

Constraints (3) ~ (20)

where $\widehat{g}_1^N$, $\widehat{g}_2^N$, ..., $\widehat{g}_{|M|}^N$ are objective function values of sample problems, and $(\widehat{x}_1^N, \widehat{y}_1^N, \widehat{y}_1^{'N}, \widehat{z}_1^N, \widehat{z}_1^{'N})$, $(\widehat{x}_2^N, \widehat{y}_2^N, \widehat{y}_2^{'N}, \widehat{z}_2^N, \widehat{z}_2^{'N})$, ..., $\left(\widehat{x}_{|M|}^N, \widehat{y}_{|M|}^N, \widehat{y}_{|M|}^{'N}, \widehat{z}_{|M|}^N, \widehat{z}_{|M|}^{'N}\right)$ are candidate solutions. Note that $(\widehat{x}, \widehat{y}, \widehat{y}^{'}, \widehat{z}, \widehat{z}^{'})$ are the first-stage decision variables.

- **Step 3:** Obtain a lower bound

$$\overline{g}_M^N = \frac{1}{|M|} \cdot \sum_{m \in M} \widehat{g}_m^N$$

The expected value $\widehat{g}_N$ of each sample problem would not be larger than the optimal value f_1 of the original problem. Since $\overline{g}_M^N$ is an unbiased estimator of $E\left[\overline{g}_M^N\right]$ and $E\left[\overline{g}_M^N\right] \leqslant f_1$, we can say that $\overline{g}_M^N$ provides a lower statistical bound for the original problem.

- **Step 4:** Generate a sample with $|N^{'}|$ scenarios, where $|N^{'}| >> |N|$. For $m \in M$, solve the model

$$\widehat{g}^{N^{'}}\left(\widehat{x}_m^N, \widehat{y}_m^N, \widehat{y}_m^{'N}, \widehat{z}_m^N, \widehat{z}_m^{'N}\right) = \min \quad fc + oc + \frac{1}{|N^{'}|}\sum_{s \in N^{'}}(ec_s + tc_s + pc_s) + (\varphi \cdot \gamma)$$

s.t.

$$f_2 - \gamma = \varepsilon$$

$$\gamma > 0$$

$$(x, y, y', z, z') = (\widehat{x}_m^N, \widehat{y}_m^N, \widehat{y}_m^{'N}, \widehat{z}_m^N, \widehat{z}_m^{'N})$$

Constraints (3) ~ (20)

where $\widehat{g}^{N'}(\widehat{x}_m^N, \widehat{y}_m^N, \widehat{y}_m^{'N}, \widehat{z}_m^N, \widehat{z}_m^{'N})$ is the estimated objective function value of a feasible solution $(\widehat{x}_m^N, \widehat{y}_m^N, \widehat{y}_m^{'N}, \widehat{z}_m^N, \widehat{z}_m^{'N})$.

- **Step 5:** Calculate the optimality gap

$$GAP = \frac{\widehat{g}^{N'}(\widehat{x}^*, \widehat{y}^*, \widehat{y}^{'*}, \widehat{z}^*, \widehat{z}^{'*}) - \overline{g}_M^N}{\overline{g}_M^N} \times 100 \backslash \%$$

if $GAP \leqslant \eta$(i.e., the specified threshold), go to **Step 6**; otherwise go to **Step 1.**

- **Step 6:** Choose $(\widehat{x}^*, \widehat{y}^*, \widehat{y}^{'*}, \widehat{z}^*, \widehat{z}^{'*})$ as the optimal solution, which satisfies

$$(\widehat{x}^*, \widehat{y}^*, \widehat{y}^{'*}, \widehat{z}^*, \widehat{z}^{'*}) \in \text{argmin}\left\{ \begin{array}{c} \widehat{g}^{N'}(\widehat{x}, \widehat{y}, \widehat{y}', \widehat{z}, \widehat{z}') : (\widehat{x}, \widehat{y}, \widehat{y}', \widehat{z}, \widehat{z}') \in \\ \left\{ (\widehat{x}_1^N, \widehat{y}_1^N, \widehat{y}_1^{'N}, \widehat{z}_1^N, \widehat{z}_1^{'N}), (\widehat{x}_2^N, \widehat{y}_2^N, \widehat{y}_2^{'N}, \widehat{z}_2^N, \widehat{z}_2^{'N}), \dots, \left(\widehat{x}_{|M|}^N, \widehat{y}_{|M|}^N, \widehat{y}_{|M|}^{'N}, \widehat{z}_{|M|}^N, \widehat{z}_{|M|}^{'N}\right) \right\} \end{array} \right\}$$

4.2. Sample generation method

The idea of the SAA is to generate a series of random samples of realizations of the random vector, and the corresponding sample average function approximates the involved expected value function. SAA possesses the computational advantage because it solves the sample problems with a significant reduction in the number of scenarios compared to the original problem. SAA estimators converge exponentially fast to their true counterparts with the sample size increase under certain regularity conditions. However, for large-scale problems, the large number of scenarios restricts the computational efficiency and convergence of SAA. If the sample size is small, SAA cannot converge easily. If the sample size is large, SAA needs to spend much more time solving the problem.

In order to balance the conflict between the size and quality of samples, we need a new method to generate relatively small samples with the most representative scenarios. Obviously, the process of randomly generating samples (see **Step 1** in Section 4.1) cannot guarantee to the selection of the most representative scenarios. Therefore, based on the framework of K-means++ presented by Emelogu et al. (Emelogu et al., 2016), we propose a **sample generation method** that integrates stratified random sampling into K-means++ to improve **Step 1** of SAA. Firstly, K-means++ is employed to group scenarios. Then, stratified sampling is applied to determine the scenario number for each cluster. Last, we randomly select scenarios in each cluster.

K-means++ is one of the most popular unsupervised learning algorithms which can group a given data into a certain number of clusters. It contains three phases. Initial k core scenarios (i.e., centers) are selected in the first phase. The first core scenario is randomly selected, while the remaining (k-1) core scenarios are chosen via a lottery process based on the distance between candidate scenarios and the selected core scenarios. Compared to K-means, instead of random selection centers, K-means++ provides a procedure for choosing a good starting point that can guarantee better quality and convergence properties (Arthur and Vassilvitskii, 2007). In the second phase, each scenario in the demand data set is assigned to its nearest core scenario, and the scenarios assigned to the same core scenario form a cluster. Then, the mean of demand for each cluster is calculated. This process is implemented repeatedly until all the core scenarios do not change. In the third phase, stratified random sampling is implemented. For each cluster, the number of scenarios should be selected, which is calculated based on variance, and scenarios are randomly selected according to the probability distribution.

Stratified random sampling is one of the most majorly used methods to estimate the density of a biological population within a defined geographical area (Cochran, 1977). Based on this method, the number of scenarios selected from each cluster can be calculated as follows.

$$N_k = \frac{|g_k| \cdot \sigma_k}{\sum_k |g_k| \cdot \sigma_k} \cdot |N|, \qquad \forall k$$

Suppose that there are $|K|$ clusters with g_k scenarios in cluster $k \in K$. The standard deviation of g_k is σ_k. $|g_k|$ denotes the number of scenarios in cluster k. There are $g_k \subset S$ and $|S| = \sum_k |g_k|$. Note that $|N|$ denotes the sample size (i.e., the number of scenarios in each sample). The standard deviation is a measurement of variance for scenarios. If a cluster has a small variance, a few number of scenarios can represent this cluster. On the contrary, more scenarios should be selected if a cluster has a large variance.

The implementation procedure of the **sample generation method** is shown as follows.

- **Step 1:** Determine $|K|$, the number of clusters; randomly select a scenario from demand data $D_{j,s}$ as the first core scenario, denoted by $\mu_{j,1}$.
- **Step 2:** Lottery process: determine if a scenario is selected as a new core scenario with its probability

$$p_s = \frac{\min\{\|D_{j,s} - \mu_{j,k}\|\}}{\sum_s \min\{\|D_{j,s} - \mu_{j,k}\|\}}, \quad \forall s$$

Where $\|D_{j,s} - \mu_{j,k}\| = \sqrt{\sum_j (D_{j,s} - \mu_{j,k})^2}$. Note that if a scenario s has been selected as a core scenario, it will not be selected again because $\min(\|D_{j,s} - \mu_{j,k}\|) = 0$.

- **Step 3:** Repeat **Step 2** until $|K|$ core scenarios are selected.
- **Step 4:** Assign scenarios to the nearest core scenarios, and obtain $\{g_1, g_2, \dots, g_{|K|}\}$.
- **Step 5:** Calculate the mean of each cluster

$$\mu_{j,k} = \frac{1}{|g_k|} \sum_{s \in g_k} D_{j,s}, \quad \forall j, k$$

- **Step 6:** Repeat **Steps 4 and 5** until convergence
- **Step 7:** Calculate the number of scenarios that should be selected

$$N_k = \frac{|g_k| \cdot \sigma_k}{\sum_k |g_k| \cdot \sigma_k} \cdot |N|, \quad \forall k$$

$$\sigma_k = \sqrt{\frac{1}{|g_k|} \cdot \sum_j \sum_{s \in g_k} \left(D_{j,s} - \frac{\sum_{s \in g_k} D_{j,s}}{|g_k|} \right)^2}, \quad \forall k$$

- **Step 8:** Select scenarios from each cluster using simple random sampling.
- **Step 9:** Repeat **Steps 8** until $|M|$ independent samples are generated.

5. Numerical analysis

5.1. Input data

The network for the case study is from Hu and Dong (Hu and Dong, 2019); and we present a brief description here. The southeastern

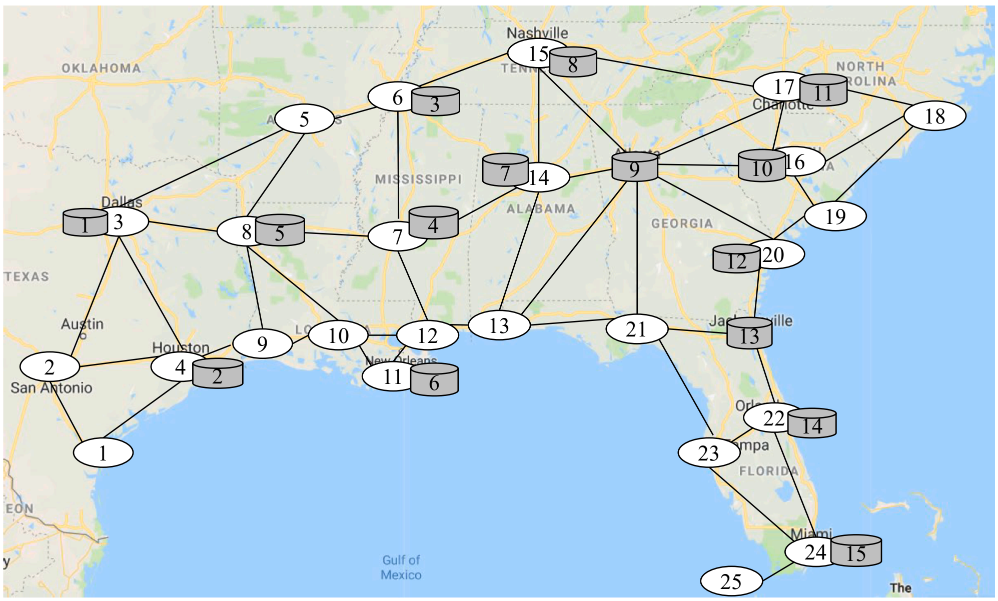


Fig. 3. Case study network.

coastal areas of the United State are usually hit by hurricanes from May to October every year. Drinking water is considered as the relief supply. Federal Emergency Management Agency will select suppliers with option contract to supply water during hurricane seasons. There are 25 demand locations (J), as shown in Fig. 3. Table 2 shows the index and name of demand locations. Set I contains 15 candidate suppliers in total. Note that we consider a supplier in both Atlanta and Jacksonville because of their critical geographical locations, although there is no demand at these two locations. The network is composed of 27 nodes and 45 links, and the distances between each pair of nodes are computed using Google Maps. As shown in Fig. 3, links such as 1–2, 1–4 and 2–3 means that these nodes are directly connected. All candidate suppliers can supply to all demand locations by several links. For instance, candidate supplier 1 can supply demand location 1 via links 1–2 and 2–3. Therefore, $H_{1,1}$ is sum of distance of links 1–2 and 2–3.

Categories, and parameters related to suppliers are shown in Table 3. We assume sizes of suppliers based on population and geographical importance of demand location. We thus assume that there is a large supplier at Dallas, Houston, Atlanta and Miami; there is a medium supplier at Memphis, New Orleans, Birmingham, Nashville, Charlotte and Orlando; the rest are small suppliers. We also consider a large supplier For Atlanta because of its important geographical location. One unit of drinking water is 10 gallons, the unit retail price is \$6.5, and the unit transportation cost is \$0.003 (Rawls and Turnquist, 2010). Set A has five order quantity intervals. The setting for option and exercise price is based on the following four assumptions: 1) summation of option and exercise price is less than retail price (Wang et al., 2015); 2) option price takes smaller proportion of retail price than exercise price (Aghajani et al., 2020); 3) larger size of suppliers (i.e., suppliers with larger capacity) can offer low option and exercise price (per unit); 4) option and exercise price will decrease with the increase in the order quantity. Therefore, we first set option and exercise prices at quantity interval 1. Then, based on these prices and given discounts, other option and exercise prices are set. We assume suppliers offer a discount even when procurement quantity is at quantity interval 1. For all sizes of suppliers, $O_{1,i}$ is set to be $6.5 \cdot 0.2 = 1.3$. For small, medium and large suppliers, $E_{1,i}$ is set to be $6.5 \cdot 0.76 = 4.94$, $6.5 \cdot 0.73 = 4.75$ and $6.5 \cdot 0.7 = 4.55$, respectively. $O_{1,i} + E_{1,i} < 6.5$ is satisfied. Discounts are generated by the uniform distribution. Take quantity interval 5 and supplier 4 as an example, $O_{5,4} = 1.3 \cdot (1\text{-}U[0.08, 0.12])$.

Unit agreement cost is assumed to be 30 % of available capacity, and the unity penalty cost is assumed to be four times the purchase price. Data for demands and probabilities with scenarios, and lower and upper quantities can be found in the online version of this paper.

5.2. Results analysis

5.2.1. Tradeoff between cost and profit (Pareto analysis)

We let q equal 20, which means 20 Pareto optimal points are generated. The complete Pareto front is presented in Fig. 4. There are four Pareto optimal points at the left of the Pareto front that show similar total expected cost and expected profit. Specifically, the total expected cost varies from \$3,776,978 to \$3,777,372, and the total expected profit is from \$871,210 to \$879,849. When l is larger than 15, ε is less than 879,848, and the total expected profit would not decrease constantly. This is because the model would procure a certain number of relief supplies to avoid a large shortage risk, no matter how ε small is. For these four Pareto optimal points, the penalty cost ranges from \$1,148,592 to \$1,158,773. In addition, minimizing the cost of relief agencies and maximizing suppliers' profit are observed to be two conflicting objectives. As total expected profit expands, total expected cost also increases. However, the penalty cost has been greatly decreased. Profit grows by approximately 380 %, and cost only increases by 77 % due to a greatly reduced penalty cost (from \$1,158,773 to 0). The reason is that as profit grows, the relief agency procures and delivers more relief supplies to decrease unmet demand. This implies that taking the suppliers' profit into account is beneficial for victims.

According to the definition of the objectives, the changes in the cost of the relief agency and the suppliers' profit depend on the discount adoption. As explained before, the all-unit quantity discount benefits the relief agency that procures the relief products, and the incremental unit quantity discount can increase suppliers' profit. Different combinations of these two discounts make the shift of Pareto optimal points. Fig. 5 shows change trend of suppliers' number for both quantity discounts with Pareto optimal points. Increasing the number of suppliers using the incremental unit quantity discount can increase the profit of suppliers. In extreme cases ($l \leqslant 7$), all selected suppliers using the incremental unit quantity discount. On the contrary, increasing the number of suppliers using the all-unit quantity discount can save cost for the relief agency. When $l \geqslant 8$, the profit is less than that of $l \leqslant 7$, the number of suppliers using the all-unit quantity discount increases. In extreme cases ($l \geqslant 11$), all selected suppliers using the all-unit quantity discount. In practice, the relief agency should select one Pareto optimal point to implement that guarantees the tradeoff between cost and profit. Specifically,

Table 2
Index and name of demand locations.

Index	Location	Index	Location	Index	Location
1	Corpus Christi, TX	10	Baton Rouge, LA	18	Wilmington, NC
2	San Antonio, TX	11	New Orleans, LA	19	Charleston, SC
3	Dallas, TX	12	Biloxi, MS	20	Savannah, GA
4	Houston, TX	13	Mobile, AL	21	Tallahassee, FL
5	Little Rock, AR	14	Birmingham, AL	22	Orlando, FL
6	Memphis, TN	15	Nashville, TN	23	Tampa, FL
7	Jackson, MS	16	Columbia, SC	24	Miami, FL
8	Shreveport, LA	17	Charlotte, NC	25	Key West, FL
9	Beaumont, TX				

Table 3
Categories, and parameters related to suppliers.

Descriptor	Supplier number	C_i	$O_{5,i}$	$E_{5,i}$	G_i
Small	4, 5, 10, 12	U[50000, 100000]	1.3*(1-U[0.08, 0.12])	4.94*(1-U[0.08, 0.12])	6.5*(1-U[0.26, 0.3])
Medium	3, 6, 7, 8, 11, 14	U[150000, 200000]	1.3*(1-U[0.13, 0.17])	4.75*(1-U[0.13, 0.17])	6.5*(1-U[0.31, 0.35])
Large	1, 2, 9, 13, 15	U[250000, 300000]	1.3*(1-U[0.18, 0.22])	4.55*(1-U[0.18, 0.22])	6.5*(1-U[0.36, 0.4])

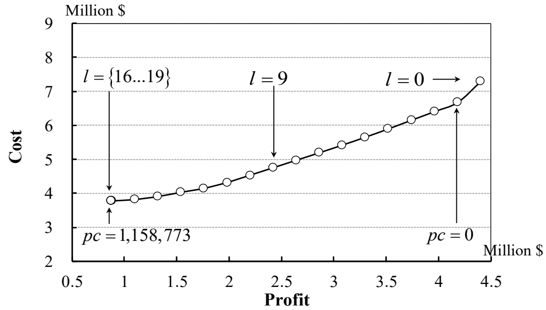


Fig. 4. Tradeoff between cost and profit (Pareto front).

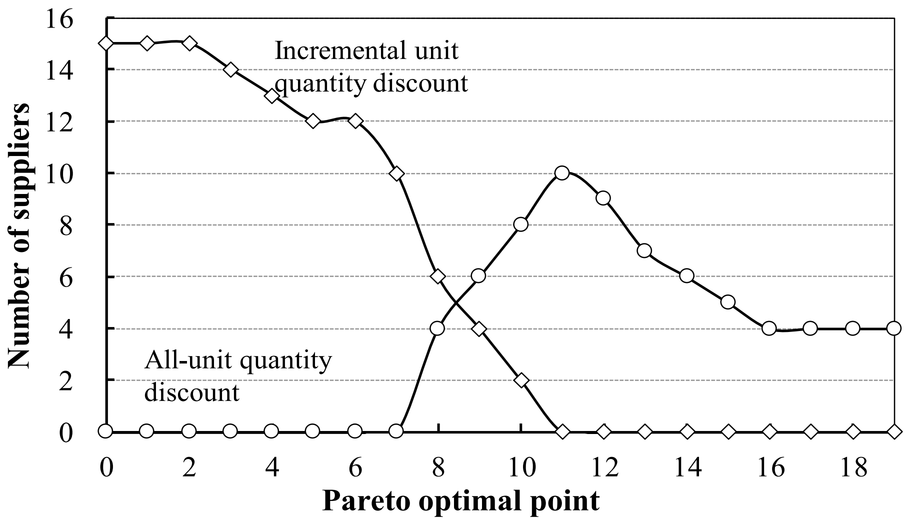


Fig. 5. Change trend of suppliers' number for both quantity discounts with Pareto optimal points.

the relief agency selects: 1) the incremental unit quantity discount for all selected suppliers when $f_2 \geqslant 2{,}859{,}505$; 2) the all-unit quantity discount for all selected suppliers when $f_2 \leqslant 1,979,658$. For other situations, both quantity discounts can be selected. Given the relationship between the combination of discounts and the shift of Pareto optimal points, the relief agency can allocate the benefits to suppliers under its budget limit. This can motivate suppliers to be actively engaged in humanitarian relief. In the next experiments, we let l for ε equals to 9, which will fix the value of f_2, i.e., the total expected profit of suppliers. The formulation of a multi-objective stochastic programming model balances the conflicting objectives of minimizing the cost for the relief agency and maximizing the profit for suppliers under uncertain demand, which enriches the literature on humanitarian relief.

5.2.2. Comparison of discount strategies

We conduct an experiment to compare the performance of different discount strategies. Four reference strategies are compared to the proposed discount strategy. These strategies are described as follows.

- **Strategy 1**: no quantity discount. In this strategy, we let $O_{a,i} = O_{1,i}$ and $E_{a,i} = E_{1,i}$.
- **Strategy 2**: only all-unit quantity discount. This strategy applies the discount to all procurement, i.e.,$y'_{a,i} = 0, \quad e'_{a,i,s} = 0, \quad \forall a,i,s.$
- **Strategy 3**: only incremental unit quantity discount. It applies the discount to the quantities exceeding the interval, i.e., $y_{a,i} = 0, \quad e_{a,i,s} = 0, \quad \forall a,i,s.$

· **Strategy 4**: consistent discount types selected in the first and second stage. In other word, for a supplier, if one discount type is selected in the first stage, it should be the same in the second stage. So in this strategy, we let $\sum_a e_{a,i,s} \leqslant \sum_a y_{a,i}, \ \sum_a e'_{a,i,s} \leqslant \sum_a y'_{a,i}, \quad \forall i, s$.

The option cost takes up the most significant percentage of the total expected cost in the proposed strategy, followed by exercise cost, fixed agreement cost, transportation cost, and penalty cost. Especially procurement cost (i.e., option and exercise costs) accounts for 81 % of the total expected cost. The reason is that the model has to reserve a large number of relief supplies by guaranteeing profit constraint, which contributes to a considerable reduction in shortage risks (only two scenarios have unmet demand). There is a total of ten suppliers have been selected. Fig. 6 shows location of selected suppliers. Six of them with all-unit quantity discounts are selected, and the rest are incremental unit quantity discounts. As cities in coastal areas were usually affected by hurricanes, six suppliers are located in Houston, New Orleans, Savannah, Jacksonville, Orlando and Miami. The supplier at Atlanta is selected due to its critical geographical locations. Only one small supplier at Savannah has been selected. It is more desirable to save costs because larger suppliers offer more significant quantity discounts.

The costs of the proposed strategy and the four reference strategies are reported in Table 4. Strategy 1 has the highest total cost and penalty cost. Compared to other discount strategies, there is only a slight change in procurement cost (option and exercise costs) of Strategy 1. Fewer relief supplies are procured due to higher procurement prices when no quantity discount is applied. Thus, this is the reason why Strategy 1 has the highest penalty cost. It concludes that it would be more desirable if suppliers could offer quantity discounts.

Strategy 2 has the lowest penalty cost, which means it outperforms other strategies to reduce shortage risks. The reason is that Strategy 2 has to largely increase the option and exercise quantity to ensure the profit for suppliers when only the all-unit quantity discount is applied. This indicates that using the all-unit quantity discount is particularly beneficial for victims. While in Strategy 3, a higher profit rate can be obtained when only the incremental quantity discount is applied. Fewer procurement quantities can also ensure the profit, but result in a larger penalty cost. In addition, there are slight changes in the costs of both hybrid discount strategies (i.e., the proposed strategy and Strategy 4), although supplier selection and procurement decisions differ. Hybrid discount strategies present a lower total expected cost than the single discount strategy (Strategy 2 and Strategy 3). The total expected cost of a single discount strategy indicates that applying a discount strategy to both stages is worse than the hybrid discount strategies. The results demonstrate the efficacy of hybrid discount strategies by providing relief agencies flexibility in the discount selection. By integrating suppliers' profits and quantity discounts, relief agencies can enhance the cost-efficiency of supplier selection processes. Moreover, this approach mitigates the shortage risks of relief supplies, which is crucial for effective disaster response.

5.2.3. *Impact of changing option or/and exercise prices*

In this subsection, we conduct three experiments to investigate the impact of the changing unit option and exercise price. First, we consider the change in both option and exercise prices. As shown in Section 5.1, we assume for all sizes of suppliers, $O_{1,i}$ is set to be 20 % of retail price; for small, medium, and large suppliers, $E_{1,i}$ is set to be 76 %, 73 %, and 70 % of the retail price, respectively. We

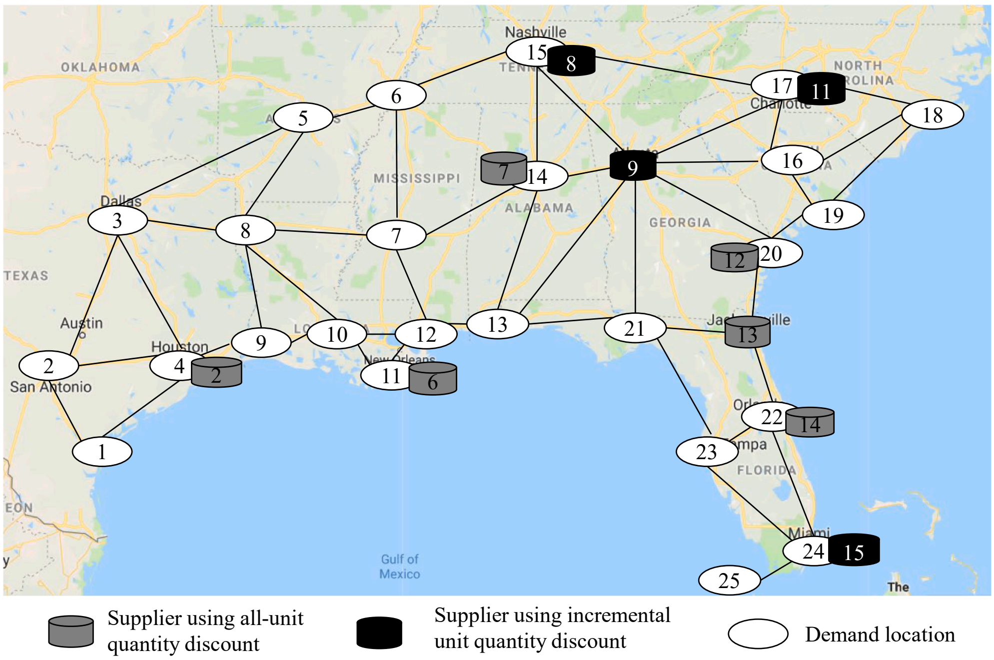


Fig. 6. Location of selected suppliers.

Table 4
Costs of different discount strategies.

Costs	Proposed Strategy	Strategy 1	Strategy 2	Strategy 3	Strategy 4
Total expected cost	4,755,232	4,801,661	4,772,310	4,795,915	4,755,319
Fixed agreement cost	619,182	530,982	697,907	542,115	619,182
Option cost	2,323,640	2,298,665	2,493,347	2,172,252	2,310,436
Exercise cost	1,527,459	1,540,267	1,376,710	1,659,169	1,541,176
Transportation cost	154,075	141,627	149,959	167,645	153,649
Penalty cost	130,876	290,120	54,388	254,734	130,876

change these percentages as follows: for all sizes of suppliers, $O_{1,i}$ is assumed to be 10 %, 30 %, 40 %, and 50 %, respectively; for small, medium, and large suppliers, $E_{1,i}$ is assumed to be [86 %, 83 %, 80 %], [66 %, 63 %, 60 %], [56 %, 53 %, 50 %], [46 %, 43 %, 40 %], respectively. Table 5 shows the results in which the variation of costs is calculated by comparing costs of the base value of option and exercise prices.

Table 5 shows that 20&70 + can obtain the minimal total expected cost, followed by 10&80+, 30&60+, 40&50+, and 50&40 + . 10&80 + has the lowest penalty cost, which means it outperforms other settings in reducing shortage risks. The lower option price is very attractive for relief agencies to procure more relief supplies. In addition, although option cost for 10&80 + decreases by 41.7 %, procurement quantity increases because option price is reduced by 50 %. To secure the profit, all suppliers are selected to apply the incremental quantity discount for the price combination 10&80 + . On the contrary, as the option price increases, the penalty cost dramatically increases, and all suppliers are selected to apply the all-unit quantity discount. The reason is that the growing rate of option cost is less than that of the option price. Thus, it results in the reduction of procurement quantity. The relief agency should negotiate with suppliers for a low option price. It would benefit victims without primarily increasing costs for the relief agency especially when option price is lower than 20 % of retail price. In addition, the relief agency should select the incremental unit quantity discount for all selected suppliers for in setting of 10&80+; 2) the all-unit quantity discount for all selected suppliers in settings of 30&60+, 40&50 + and 50&40 + .

Next, we fix the option price of relief supplies for each interval to set $O_{a,i}$ equal to the same value for all intervals. Specifically, we let $O_{a,i} = O_{1,i}$, $O_{a,i} = O_{2,i}$, $O_{a,i} = O_{3,i}$, $O_{a,i} = O_{4,i}$, $O_{a,i} = O_{5,i}$, respectively, which means the option price decreases. The experiment results are summarized in Table 6. For the experiments that fix the option price, the total expected cost shows a slight change in response to the decrease of the option price. When the fixed option price is high (i.e., $O_{a,i} = O_{1,i}$, $O_{2,i}$, $O_{3,i}$), penalty cost increases at least by 13 %. This indicates that a high option price would be disadvantageous for victims. While as the fixed option price becomes low (i.e., $O_{a,i} = O_{4,i}$, $O_{5,i}$), neither the total cost nor the penalty cost is changed. However, because the profit rate decreases as the option price reduce, the number of suppliers using the incremental quantity discount increases to ensure profit. As shown in Table 6, this number increases from 1 to 7.

Last, we fix the exercise price. Similar to the above experiments, we let $E_{a,i} = E_{1,i}$, $E_{a,i} = E_{2,i}$, $E_{a,i} = E_{3,i}$, $E_{a,i} = E_{4,i}$, $E_{a,i} = E_{5,i}$, respectively, which means the exercise price decreases. For the experiments that fix the exercise price, Table 7 demonstrates a very similar pattern of the total expected cost to the decline in the exercise price. Also, more suppliers with incremental quantity discounts are selected according to the decrease in the exercise price.

5.2.4. *Sensitivity analysis*

The experiments in this subsection examine the influence of changing the value of parameters on costs and supplier selection decisions. Regarding all suppliers' capacity, its value is changed by −25 %, +25 %, +50 %, and + 75 %. When all suppliers' capacity decreases by 50 %, the profit constraint cannot be satisfied. Option contract reduces relief agencies' risk by paying an option cost. However, the reduced risk is transferred to the suppliers. When the suppliers are smaller entities, they are likely to take low risks. Thus, we let small suppliers' capacity decrease by 25 %, 50 %, 75 %, and 100 %. Smaller capacity means that suppliers keep a lower inventory for the relief agency. In addition, the sensitivity analysis focuses on unit agreement cost and unit penalty cost, and their values are changed from −50 %, −25 %, +25 %, and + 50 % of the base values. For different suppliers' capacities, the results are presented in Table 8. For different unit agreement and unit penalty costs, the results are presented in Table 9.

(1) Suppliers' capacity

Table 5
Results of option and exercise prices change.

Costs (unit: %)	10&80+	30&60+	40&50+	50&40+
Total expected cost	+1.0	+1.7	+5.3	+7.9
Fixed agreement cost	+9.0	−17.7	−32.8	−44.2
Option cost	−41.6	+15.8	+23.4	+29.9
Exercise cost	+64.2	−26.4	−41.4	−54.7
Transportation cost	+20.9	+0.2	−4.9	+13.1
Penalty cost	−43.2	+172.3	+423.4	+590.0
Supplier selection				
Number of suppliers using all-unit quantity discount	0	8	7	5
Number of suppliers using incremental quantity discount	13	0	0	0

Table 6
Results of fixed option price.

Costs (unit: %)	$O_{a,i} = O_{1,i}$	$O_{a,i} = O_{2,i}$	$O_{a,i} = O_{3,i}$	$O_{a,i} = O_{4,i}$	$O_{a,i} = O_{5,i}$
Total expected cost	+0.14	+0.11	+0.02	+0.01	0
Fixed agreement cost	−6.78	−5.00	−2.87	0	0
Option cost	+7.65	+7.69	+5.69	+2.45	−5.50
Exercise cost	−12.43	−11.64	−8.43	−3.73	+8.36
Transportation cost	+2.52	−3.58	−3.55	+0.23	+0.22
Penalty cost	+43.62	+31.00	+15.91	0	0
Supplier selection					
Number of suppliers using all-unit quantity discount	9	10	10	5	3
Number of suppliers using incremental quantity discount	1	2	1	5	7

Table 7
Results of fixed exercise price.

Costs (unit: %)	$E_{a,i} = E_{1,i}$	$E_{a,i} = E_{2,i}$	$E_{a,i} = E_{3,i}$	$E_{a,i} = E_{4,i}$	$E_{a,i} = E_{5,i}$
Total expected cost	0	0	0	0	0
Fixed agreement cost	0	0	0	0	0
Option cost	−3.38	−3.87	−1.78	+0.16	+2.31
Exercise cost	+5.19	+5.83	+2.71	−0.23	−3.53
Transportation cost	−0.45	+0.60	−0.05	−0.15	+0.10
Penalty cost	0	0	0	0	0
Supplier selection					
Number of suppliers using all-unit quantity discount	8	8	6	5	3
Number of suppliers using incremental quantity discount	2	2	4	5	7

Table 8 reports that the total expected cost increases as all suppliers' capacities decrease, and the penalty cost significantly increases by over 60 %. In addition, more suppliers are selected to ensure sufficient supply compared to the result of the base value. Another phenomenon is that over half (9 in 13) of suppliers with incremental quantity discounts are selected, which is larger than the result (4 in 10) of the base value. The reason is that more profit can be achieved using the incremental quantity discount.

The total expected cost decreases, and penalty cost is greatly decreased by over 50 % as suppliers' capacities increase. The number of suppliers decreases if the suppliers' capacity is augmented. The rise in suppliers' capacity lets each supplier provide more relief products. It will reduce the fixed agreement cost since fewer suppliers are selected. Hence, the total cost is decreased in response to the increase in the suppliers' capacity. In addition, as more relief supplies are procured, the model tends to use the all-unit quantity discount to save costs. As a result, one can see that the number of suppliers using the incremental quantity discount is reduced.

Letting small suppliers' capacity decrease is to simulate a situation where smaller entities may be risk conservative. An extreme case is that small suppliers would not reserve relief supplies for the relief agency, i.e., small suppliers' capacity decreases by 100 %. The results report that a decrease in small suppliers' capacity does not significantly affect total expected cost and supplier section selections. However, it should notice that compared to the result of the base value (4 large suppliers, 5 medium suppliers, and 1 small supplier), the only small supplier have not been selected when small suppliers' capacity decreases. It results in the increase of penalty cost as the model procures fewer relief supplies from the remaining 9 suppliers. In situations where a large number of small suppliers are involved, a decrease in small suppliers' capacity may incur a significant shortage. Compared to the strategy of selecting much suppliers with small capacity, selecting less suppliers with large capacity can save more costs and satisfy more victims.

(2) Unit agreement cost

Unit agreement cost is an important parameter affecting procurement and supplier selection decisions. Table 8 indicates that the total expected cost increases/decreases as the value of the unit agreement cost is increased/decreased. Changing unit agreement cost leads to a significant variation in the penalty cost. When unit agreement cost is low, the relief agency can reduce the unmet demand by

Table 8
Results of different suppliers' capacity.

	All suppliers' capacity $C_{1\text{-}15}$				Small suppliers' capacity $C_{4,5,10,12}$			
Costs (unit: %)	−25	+25	+50	+75	−25	−50	−75	−100
Total expected cost	+5.1	−3.0	−5.2	−6.6	+0.02	+0.02	+0.02	+0.01
Fixed agreement cost	+20.9	−11.4	−25.4	−34.6	−2.5	−2.5	−2.5	−2.5
Option cost	−7.2	+6.5	+7.0	+6.9	−2.6	−2.4	−0.1	−3.4
Exercise cost	+10.5	−10.2	−10.8	−10.4	+3.6	+3.3	−0.1	+4.9
Transportation cost	+25.8	−0.8	−8.4	−12.4	+2.1	+2.6	+1.9	+1.8
Penalty cost	+61.8	−51.6	−56.5	−61.5	+13.0	+13.0	+13.0	+13.0
Supplier selection								
Number of suppliers using all-unit quantity discount	4	6	6	6	5	5	4	6
Number of suppliers using incremental quantity discount	9	2	1	0	4	4	5	3

Table 9
Results of different unit agreement and unit penalty costs.

Costs (unit: %)	Unit agreement cost F_i −50	−25	+25	+50	Unit penalty cost P −50	−25	+25	+50
Total expected cost	−7.0	−3.4	+3.2	+6.1	−2.1	−0.9	+0.6	+0.9
Fixed agreement cost	−43.6	−18.3	+17.5	+35.6	−18.1	−9.6	+9.0	+12.7
Option cost	+7.8	+6.3	−4.5	−7.0	−12.2	−8.3	+5.6	+8.1
Exercise cost	−10.7	−8.7	+6.3	+9.6	+16.1	+11.4	−7.6	−11.1
Transportation cost	−2.3	−2.9	+0.9	+3.1	−2.0	+2.9	−2.8	−2.4
Penalty cost	−58.4	−43.2	+38.1	+63.7	+39.5	+22.6	−28.9	−37.5
Supplier selection								
Number of suppliers using all-unit quantity discount	10	8	4	3	0	3	11	11
Number of suppliers using incremental quantity discount	2	5	6	6	10	6	2	1

selecting more suppliers with the all-unit quantity discount and procuring more relief supplies. This implies that negotiating a low unit agreement would be critical for relief agencies to save the total cost and reduce shortage risks.

(3) Unit penalty cost

As Table 9 shows, when the unit penalty cost increases, the penalty cost largely reduces with a slight increase in the total expected cost. It means that when a higher penalty is applied, procured and delivered quantity through the model tends to be increased, which results in lower unmet demand. In addition, almost all suppliers (11 in 13 or 11 in 12) take the all-unit quantity discount, which is much larger than the result (6 in 10) of the base value of the unit penalty cost. The reason is that the all-unit quantity discount is very important for the relief agency to save costs as procurement quantity increases. This result indicates that those suppliers who offer the all-unit quantity are more likely to be selected due to cost-effectiveness when the unit penalty cost increases. The results of sensitivity analysis underscore the importance of flexibility in contracting and the need for adaptive strategies in humanitarian relief.

5.3. Performance of ML-based SAA

In this section, we generate different data sets to verify the efficiency of the proposed solution methodology for large-scale problems. The algorithms are coded in Matlab R2022a, and Gurobi 9.5.1 is called as an optimization solver (https://www.gurobi.com/) to solve sample problems. All experiments are run on a desktop with Intel Core i9 3.10 GHz processor and 32.0 GB RAM. When Gurobi is called to solve the original problems, we set the optimality gap to be 1 % (or maximum time limitation to be 3,600 s) and obtain the upper bound.

We consider 20, 40, and 60 demand locations, and 100, 200, 300, 400, and 500 scenarios. Coordinates of demand locations are generated by (U[1, 3000], U[1, 1500]), and disaster affected population $population_j$ for each demand location j is generated by U[5000,400000]. Distance H is Euclidean distance. Based on the method presented by Dalal and Üster (Dalal and Üster, 2018); the demand can be generated as follows. We define a scenario by two attributes: occurrence location and disaster intensity. For occurrence location, we assume that disasters can hit all locations. For disaster intensity, we also consider hurricanes, which are usually categorized based on a five-point Saffir–Simpson scale (5 is the most severe). For instance, scenario 1 can denote a category 1 hurricane hit location 1. Moreover, we consider two different hurricanes that can hit two locations simultaneously. Thus, scenario 2 can denote a category 2 hurricane hit location 1, and a category 3 hurricane hit location 2. Demand is affected population which depends on the total population of affected areas, landfall location, and impact degree by the hurricane. The distance factor ($DF_{i,s}$) measures the fraction of the population of an area j is affected by the distance ($d_{j,j'}$) to landfall location j'. For instance, if a location j is over 100 far from the occurrence location when a category 1 or 2 hurricane hits, the affected fraction of the population is zero.

$$DF_{j,s} = \begin{cases} 1, & \text{if } d_{j,j'} \leqslant 100 \text{ and } Catg = \{1,2,3,4,5\} \\ 0.6, & \text{if } 100 < d_{j,j'} \leqslant 300 \text{ and } Catg = \{3,4,5\} \\ 0.3, & \text{if } 300 < d_{j,j'} \leqslant 500 \text{ and } Catg = \{4,5\} \\ 0 & \text{if } d_{j,j'} > 500 \end{cases}$$

Moreover, the hurricane category also should be considered because severe hurricanes have a more considerable impact on the fraction of the population. Therefore,

$$v_{j,s} = W_1 \cdot DF_{j,s} + W_2 \cdot \frac{Catg}{5} \Gamma_{(DF_{j,s}>0)}$$

Where $v_{j,s} \in (0,1)$ and $Catg = \{1,2,3,4,5\}$ denote hurricane category. W_1 and W_2 are weight, we let $W_1 = 0.2$, $W_2 = 0.8$. $\Gamma_{(DF_{j,s}>0)}$ is an indicator function based on the value of $DF_{j,s}$, i.e., $\Gamma = 1$ if $DF_{j,s} > 0$, otherwise $\Gamma = 0$. Therefore $D_{j,s} = v_{j,s} \cdot population_j$. To simplify the expression for algorithms, the following abbreviations are used to represent the presented algorithms.

- ***SKR**: SAA_K-means++_Random sampling (select one scenario for each cluster)*, which is presented by Emelogu et al. (Emelogu et al., 2016).

- ***SKRC***: *SAA_**K-m**eans++_**R**andom sampling (select one scenario for each cluster)_Scenario probability proportional to the **C**luster's size*, which is presented in Appendix A.
- ***SKS***: *SAA_**K-m**eans++_**S**tratified sampling technique (selected variable number of scenarios for each cluster)*, which is presented in section 4.

We consider a small sample size, i.e., $N = 20$; number of samples M is 10. We also consider two types of large sample size because the performance of algorithms in finding a better upper bound is usually better with a larger sample size. We thus set $N = 30$ for 100 and 200 scenarios cases and $N = 50$ for 300–500 scenarios cases. To compare the objective function value obtained by ***Gurobi***, we let $N^{'} = 100, 200, 300, 400, 500$ in ***SKR, SKRC*** and ***SKS***. Let Obj_{Gurobi}, Obj_{SKR}, Obj_{SKRC} and Obj_{SKS} be the best upper bound that is obtained respectively by ***Gurobi***, ***SKR, SKRC*** and ***SKS***. We have $Gap_{SKR} = \left(Obj_{SKR} - Obj_{Gurobi}\right) \cdot 100\%/Obj_{Gurobi}$, $Gap_{SKRC} = \left(Obj_{SKRC} - Obj_{Gurobi}\right) \cdot 100\%/Obj_{Gurobi}$ and $Gap_{SKS} = \left(Obj_{SKS} - Obj_{Gurobi}\right) \cdot 100\%/Obj_{Gurobi}$. The results are highlighted (in bold) if ***SKR*** and ***SKS*** found a better upper bound within less time than ***Gurobi***. The performances of ***Gurobi***, ***SKR*** and ***SKS*** on different cases are shown in Table 10. Take 20 locations problems as examples, the distribution of objective function value with algorithms is shown in Fig. 7.

Compared to ***Gurobi***, when the scenario number becomes large, ***SKR*** performs significantly better for most cases in saving computational time, except for the cases of 100 scenarios. ***SKS*** also demonstrates a great advantage in saving computational time for large-scale cases. More specifically, when the number of scenarios is larger than 100, ***SKS*** respectively saved 80 %, 63 %, and 72 % of the computational time on average, as the number of demand locations increases. ***SKS*** cannot guarantee convergence faster than ***SKR*** in all cases. For example, when the numbers of demand locations are 20 and 40, ***SKS*** is 18 % and 9 % faster than ***SKR*** on average; while when the number of demand locations is 60, ***SKS*** is 8 % slower than ***SKR*** on average. However, ***SKS*** can always guarantee to obtain a better upper bound. Compared to the upper bound obtained by ***Gurobi***, a gap of ***SKR*** is 2.82 %, 0.28 %, and 1.67 % on average, respectively, while the gap of ***SKS*** is only 0.47 %, −1.6 %, and −0.49 % on average, respectively.

The average gap of ***SKRC*** for 20 locations cases is 0.16 %; for 40 locations cases, that is −2.12 %; for 60 locations cases, that is −1.35 %. The results reveal that ***SKRC*** is the best in finding a better upper bound for all cases. However, ***SKRC*** consumes more time than ***SKS*** in large cases. For (40,40,500) case, ***SKS*** requires less than 500 s on average; ***SKRC*** consumes over 1000 s on average. This is because, in ***SKRC,*** more time is required to solve sample problems than ***SKS***.

Fig. 7 elucidates the superior performance of ***SKRC*** and ***SKS*** over ***SKR*** in achieving a more favorable upper bound. As depicted in Fig. 7, the objective function values for ***SKRC*** and ***SKS*** are consistently lower than those for ***SKR*** across a majority of sample problems. This indicates that ***SKRC*** and ***SKS*** are adept at generating superior samples compared to ***SKR***. The rationale behind this lies in the fact that a larger cluster inherently encapsulates more information than a smaller one. Consequently, more information, such as selecting additional scenarios or assigning larger probabilities, is required to effectively represent a large cluster in samples. For instance, when we group cases with $(I, J, S) = (20{,}20{,}100)$, the smallest cluster contained only 2 scenarios, while the largest cluster comprised 34 scenarios. ***SKR***, unfortunately, selected only one scenario from each cluster, even in the case of larger clusters. In the case of ***SKS***, the standard deviation was employed to gauge variance during the steps of stratified random sampling. For clusters exhibiting substantial variances, a greater number of scenarios were selected. Additionally, it is crucial to note that in ***SKRC***, we weighted the selected scenarios proportionally to the cluster's size. This means that scenarios from larger clusters were assigned higher probabilities.

Moreover, As the sample size increases, the variance of objective function values of ***SKR***, ***SKRC,*** and ***SKS*** would all become smaller. Therefore, the increase in sample size can be a useful strategy to improve the performance of these algorithms. ***SKRC*** produces the smallest variance of objective function values, even for a small sample size. This implies that for ***SKRC***, increasing sample size may have no significant improvement in finding a better upper bound but consume more computational time. Lastly, ***SKS*** that cannot guarantee all generated samples are good enough since the objective function values of some samples are obviously larger than others. This is because simple random sampling is applied in ***SKS,*** and this procedure cannot guarantee to the selection of the most representative scenarios. This motivates further improvement of ***SKS*** in the future. The use of machine learning-based optimization methods provides relief agencies with tools to better handle the complexities and uncertainties inherent in humanitarian relief.

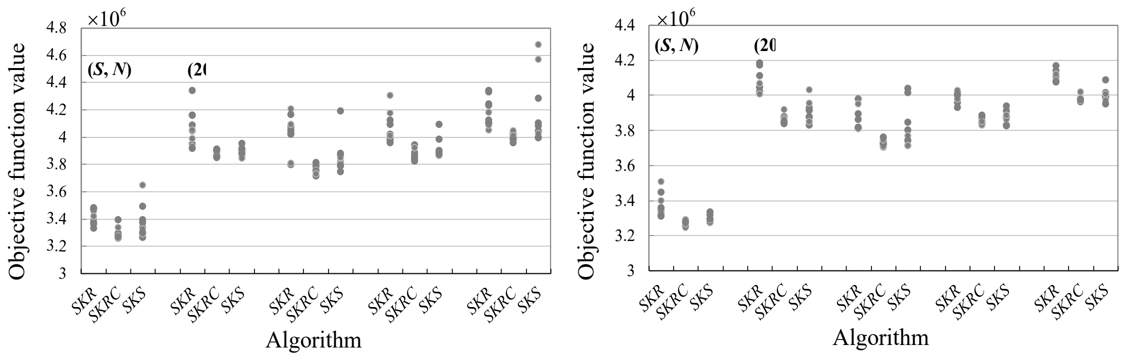
6. Conclusions

This paper addresses supplier selection in the humanitarian relief considering the tradeoff of the benefits between the relief agency and suppliers based on the option contract. The proposed option contract integrates two quantity discounts, i.e., all-unit quantity discount and incremental quantity discount, to not only reduce the expected cost for the relief agency and shortage risks for victims, but also ensure profit to attract suppliers. This problem is formulated as a multi-objective SP model to handle the uncertainty of disaster relief demand. A ML based SAA is developed to solve the proposed model in large-scale cases. The case study demonstrates that 1) the Pareto frontier indicates the solutions minimizing the expected cost of the relief agency and maximizing the expected profit of suppliers that can be implemented by the decision-maker considering different balances; 2) sensitivity analysis provides insights to relief agencies to capture the best solution to select proper suppliers and apply corresponding quantity discounts in response to the changes in parameters; 3) the computational comparison results justify the efficacy of the SAA enhanced by the ML approach for solving large-scale problems.

Several managerial insights were summarized as follows. First, our analysis demonstrated the importance of integrating suppliers' profit and quantity discount in enhancing the cost-efficiency of supplier selection for the relief agency and mitigating shortage risks for victims. This work provided relief agencies with an efficient approach to examining the combination of all-unit and incremental quantity discounts to capture the total minimal expected cost and shortage risks and not the violate profit constraint. Second, applying

Table 10
Performance of ***Gurobi***, ***SKR, SKRC*** and ***SKS*** in different cases.

(*I*, *J*, *S*)	***Gurobi*** Upper bound	CPU time	*N*	***SKR*** GAP	CPU time	***SKRC*** GAP	CPU time	***SKS*** GAP	CPU time
(20,20,100)	3,249,942	24 s	20	2.52 %	48 s	0.25 %	40 s	0.43 %	42 s
			30	1.90 %	71 s	−0.16 %	71 s	0.73 %	70 s
(20,20,200)	3,829,136	201 s	20	2.30 %	74 s	0.52 %	83 s	0.39 %	65 s
			30	4.57 %	95 s	0.25 %	90 s	**−0.03 %**	69 s
(20,20,300)	3,708,284	503 s	20	2.32 %	104 s	0.14 %	87 s	0.94 %	73 s
			50	2.67 %	144 s	**−0.13 %**	185 s	0.08 %	153 s
(20,20,400)	3,811,578	1064 s	20	3.80 %	138 s	0.31 %	132 s	1.36 %	109 s
			50	3.04 %	209 s	0.45 %	332 s	0.30 %	160 s
(20,20,500)	3,961,362	1447 s	20	2.24 %	192 s	**−0.05 %**	173 s	0.78 %	115 s
			50	2.81 %	224 s	**−0.03 %**	364 s	**−0.26 %**	185 s
(40,40,100)	6,313,838	24 s	20	4.60 %	89 s	0.53 %	97 s	1.30 %	90 s
			30	2.21 %	106 s	−0.04 %	161 s	0.29 %	108 s
(40,40,200)	6,291,648	359 s	20	3.44 %	136 s	0.31 %	138 s	2.56 %	131 s
			30	3.26 %	160 s	0.52 %	161 s	0.53 %	139 s
(40,40,300)	9,387,274	290 s	20	2.73 %	212 s	**−0.36 %**	217 s	0.62 %	224 s
			50	2.10 %	281 s	−0.06 %	470 s	0.14 %	266 s
(40,40,400)	11,174,691	3600 s	20	**−8.80 %**	361 s	**−10.45 %**	382 s	**−10.12 %**	678 s
			50	**−8.57 %**	599 s	**−10.04 %**	699 s	**−10.25 %**	404 s
(40,40,500)	10,700,152	3600 s	20	0.75 %	1177 s	**−0.50 %**	1637 s	**−0.59 %**	422 s
			50	1.02 %	1380 s	**−1.10 %**	1210 s	**−0.45 %**	488 s
(60,60,100)	8,813,229	78 s	20	2.23 %	158 s	0.02 %	124 s	1.67 %	148 s
			30	3.15 %	187 s	0.38 %	247 s	1.07 %	173 s
(60,60,200)	9,010,566	538 s	20	3.22 %	262 s	**−0.15 %**	210 s	0.71 %	224 s
			30	2.52 %	308 s	0.20 %	398 s	0.68 %	274 s
(60,60,300)	9,173,366	1713 s	20	4.54 %	338 s	**−0.38 %**	342 s	1.00 %	343 s
			50	3.05 %	493 s	**−0.21 %**	767 s	0.26 %	443 s
(60,60,400)	11,373,538	3600 s	20	**−1.98 %**	462 s	**−4.33 %**	481 s	**−3.88 %**	845 s
			50	**−1.12 %**	584 s	**−4.56 %**	1642 s	**−4.07 %**	796 s
(60,60,500)	12,218,023	3600 s	20	0.66 %	676 s	**−2.10 %**	2928 s	**−0.56 %**	570 s
			50	0.39 %	798 s	**−2.33 %**	1075 s	**−1.75 %**	1020 s



(a) Small sample size (b) Large sample size

Fig. 7. Distribution of objective function value with algorithms (20 locations).

the all-unit quantity discount helps save costs for relief agencies and reduce shortage risks for victims. Applying the incremental quantity discount can increase the suppliers' profit, which is why the number of suppliers using the incremental quantity discount increases when option or exercise price reduce. Third, negotiating for large capacity, low option price, and low unit agreement cost benefits relief agencies and victims. Different option prices can impact the number of suppliers selected, which facilitates relief agencies in managing the number of suppliers. Last, ***SKRC*** and ***SKS*** outperform ***SKR*** and ***Gurobi*** in finding better solutions within less time for different problems. This expands applicability of the proposed model in large-scale problems.

There are several limitations of the research which can be the directions for future research. First, disaster occurrence time has been ignored in this paper. For example, hurricanes often occur from June to September, and inventory is idle until hurricane season. Second, this study has focused on the water supply for responding to hurricanes, which might not be suitable for other disasters/ suppliers. Third, once disasters are occurring, quickly transport relief supplies to disaster-affected areas for victims is also important. Last, in the proposed sample generation method, randomly selecting scenarios cannot guarantee that the most representative samples will be chosen. Therefore, one can integrate occurrence time into supplier selection and/or the option contract is meaningful to reduce option quantity (Hu et al., 2022); tailor the option contract for a specific type of disaster and supplies that may be applicable for more general applications (Hu et al., 2022; Moshtari et al., 2021); develop an effective distribution plan to allocate vehicles and schedule routes for quickly deliver relief supplies (Minh and Noi, 2023; Jin et al., 2022); create scenario selection methods to improve scenario generation process.

CRediT authorship contribution statement

Shaolong Hu: Writing – original draft, Software, Methodology, Conceptualization. **Zhijie Sasha Dong:** Investigation, Data curation, Writing – original draft. **Rui Dai:** Writing – review & editing, Visualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was mianly supported by National Science Foundation [grant number CISE CCF - 1948159]. This research was supported in pamainly supported by the rt by the National Natural Science Foundation of China [grant number 72101214]; Sichuan Science and Technology Program [grant number 2022NSFSC1115]. We also would like to sincerely thank anonymous reviewers for their constructive feedbacks on this paper.

Appendix

In this section, one more algorithm is designed for comparison, that is ***SKRC*: *SAA_K-means++_Random sampling (select one scenario for each cluster)_Scenario probability proportional to the Cluster's size*** and is discussed in Appendix A. By instead of using sample average function to approximate the expected value function in ***SKR*** and ***SKS,*** we weight the selected scenarios with a probability proportional to the cluster's size in ***SKRC***. In addition, we present two different problems to validate the performance of ***SKR***, ***SKRC*** and ***SKS***. In Appendix B, we provide a pre-positioning problem. In Appendix C, a supplier selection and pre-positioning problem is provided. The computer configuration and termination criteria of the algorithms are the same as the previous description in Section 5.3.

Appendix A. *SKRC*: *SAA_K-means++_random sampling (select one scenario for each cluster)_Scenario probability proportional to the Cluster's size*

- **Step 1–6: Same to *SKS*.**
- **Step 7:** Select one scenario from each cluster by using simple random sampling, and calculate its probability using

$$\omega_k' = \frac{|g_k|}{|S|}, \qquad \forall k$$

Where ω_k' is the probability proportional to the size of cluster k.

- **Step 8:** Repeat **Steps 7** until $|M|$ independent samples N are generated.
- **Step 9:** For $m \in M$, solve the model

$$\widehat{g}^N = \min \quad fc + oc + \omega_k' \cdot \sum_{s \in N} (ec_s + tc_s + pc_s) + (\varphi \cdot \gamma)$$

s.t.

$$f_2 - \gamma = \varepsilon$$

$$\gamma > 0$$

Constraints (2) ~ (20)

where, $\widehat{g}_1^N, \widehat{g}_2^N, \ldots, \widehat{g}_{|M|}^N$ are objective function values of sample problems, and $\widehat{x}_1^N, \widehat{x}_2^N, \ldots, \widehat{x}_{|M|}^N$ are candidate solutions. Note that $\widehat{x}$ denote the first-stage decision variables.

- **Step 10:** For $m \in M$, solve the model (original problem)

$$\widehat{g}^{N'}(\widehat{x}) = \min \quad fc + oc + \omega_s \cdot \sum_{s \in N'} (ec_s + tc_s + pc_s) + (\varphi \cdot \gamma)$$

s.t.

$$f_2 - \gamma = \varepsilon$$

$$\gamma > 0$$

Constraints (2) ~ (20)

where $\hat{g}^{N'}(\hat{x})$ is the estimated objective function value of a feasible solution $\hat{x}$.

- **Step 11:** Calculate the optimality gap, compared to Gurobi

$$GAP = \frac{\hat{g}^{N'}(\hat{x}^*) - f_1}{f_1} \times 100 \backslash \%$$

where f_1 is the objective function value obtained by ***Gurobi***.

Appendix B. Pre-positioning problem

Based on the study of Rawls and Turnquist (Rawls and Turnquist, 2010); a pre-positioning problem is proposed, and data is also from their paper. Pre-positioning of relief supplies is a basic problem in humanitarian logistics. The goal is to determine the location and quantities of one/multiple types of relief supplies to be pre-positioned under uncertainty. A two-stage stochastic programming model is provided as follows.

Table B1
Model notations in pre-positioning problem.

Notation	Explanation
Sets	
I	Set of locations, indexed by $i,j \in I$
L	Set of size categories, indexed by $l \in L$
A	Set of relief items, indexed by $a \in A$
S	Set of disaster scenarios, indexed by $s \in S$
Parameters	
$D_{a,j,s}$	Demand for item a at location j in scenario s
P_s	Probability of occurrence of scenario s
U_l	Capacity of l size facility
CF_l	Unit setup cost of l size facility
$H_{i,j}$	Distance from location i to location j
V_a	Volume of a unit of item a
CP_a	Unit purchase cost of item a
CT_a	Unit transportation cost of item a
CH_a	Unit holding cost of item a
G_a	Unit penalty cost of item a
First-Stage Decision Variables	
$x_{i,l}$	Binary decision variable if a size l facility is opened at location i
$y_{a,i}$	Procurement quantities of item a at location i
Second-Stage Decision Variables	
$q_{a,i,j,s}$	Amount of item a shipped from facility i to location j in scenario s
$z_{a,i,s}$	Surplus quantity of item a at facility i in scenario s
$w_{a,j,s}$	Shortage quantity of item a at location j in scenario s

The objective function (B1) minimizes the overall cost consisting of setup cost and procurement cost in the first stage, and the expected transportation cost, holding cost, and penalty cost under uncertain demands in the second stage. Constraint (B2) restricts that procurement quantities would not exceed facilities' storage capacity. Constraint (B3) restricts that only one size category of facility can be opened at most for each location. Constraint (B4) calculates the surplus quantity of relief supplies at each facility in each scenario. Constraint (B5) calculates the shortage quantity at each location in each scenario. Constraints (B6) and (B7) define the integrity of the variables.

$$f = \min \sum_i \sum_l CF_l \cdot x_{i,l} + \sum_a \sum_i CP_a \cdot y_{a,i} + \sum_s P_s \cdot \left(\sum_a \sum_i \sum_j CT_a \cdot H_{i,j} \cdot q_{a,i,j,s} + \sum_a \sum_i CH_a \cdot z_{a,i,s} + \sum_a \sum_j G_a \cdot w_{a,j,s} \right) \tag{B1}$$

$$\sum_a y_{a,i} \cdot V_a \leqslant \sum_l U_l \cdot x_{i,l}, \qquad \forall i \tag{B2}$$

$$\sum_l x_{i,l} \leqslant 1, \qquad \forall i \tag{B3}$$

$$z_{a,i,s} = y_{a,i} - \sum_j q_{a,i,j,s}, \qquad \forall a, i, s \tag{B4}$$

$$w_{a,j,s} = D_{a,j,s} - \sum_i q_{a,i,j,s}, \qquad \forall a, j, s \tag{B5}$$

$$x_{i,l} = \{0, 1\}, \qquad \forall i, l \tag{B6}$$

$$y_{a,i}, q_{a,i,j,s}, z_{a,i,s}, w_{a,j,s} \geqslant 0, \qquad \forall a, i, j, s \tag{B7}$$

Table B.2 shows the performance of ***Gurobi***, ***SKR, SKRC,*** and ***SKS*** in different cases for the proposed pre-positioning problem. The results report that as the number of locations and scenarios increases, compared to ***Gurobi***, ***SKRC*** and ***SKS*** can find good solutions with less time. Especially for (60,60,500) case, ***SKRC*** and ***SKS*** can both produce better upper bound than ***Gurobi*** with less time. The average gap of ***SKR***, ***SKRC***, and ***SKS*** for the cases of 20 locations is 6.14 %, 0.38 %, and 0.68 %, respectively. For the cases of 40 locations, that is 3.17 %, 0.51 % and 1.62 %; for the cases of 60 locations, that is 2.71 %, 0.1 % and 0.48 %. The results reveal that ***SKRC*** can produce the best upper bound for all cases, followed by ***SKS***, then ***SKR***.

Table B2
Performance of Gurobi, SKR, SKRC and SKS in different cases.

(*I*, *J*, *S*)	***Gurobi*** Upper bound	CPU time	*N*	***SKR*** GAP	CPU time	***SKRC*** GAP	CPU time	***SKS*** GAP	CPU time
(20,20,100)	6.070×10^7	7	20	2.43 %	42 s	0.84 %	41 s	0.66 %	40 s
			30	2.04 %	47 s	0.36 %	46 s	2.55 %	47 s
(20,20,200)	6.841×10^7	28	20	3.74 %	77 s	0.24 %	78 s	0.88 %	77 s
			30	4.89 %	83 s	0.09 %	82 s	0.43 %	81 s
(20,20,300)	6.836×10^7	95	20	9.50 %	113 s	0.56 %	108 s	0.66 %	111 s
			50	8.31 %	130 s	0.11 %	136 s	0.05 %	126 s
(20,20,400)	6.906×10^7	204	20	1.65 %	149 s	0.19 %	145 s	0.19 %	150 s
			50	14.83 %	167 s	0.07 %	160 s	0.26 %	160 s
(20,20,500)	7.107×10^7	183	20	4.66 %	185 s	0.76 %	179 s	0.44 %	185 s
			50	9.30 %	204 s	0.57 %	200 s	0.69 %	200 s
(40,40,100)	1.243×10^8	63	20	5.27 %	150 s	0.62 %	148 s	3.66 %	144 s
			30	4.46 %	165 s	0.20 %	171 s	4.14 %	165 s
(40,40,200)	1.209×10^8	2322	20	1.49 %	277 s	0.09 %	282 s	0.74 %	281 s
			30	1.81 %	300 s	0.18 %	303 s	2.84 %	286 s
(40,40,300)	1.817×10^8	3600	20	1.14 %	401 s	0.69 %	418 s	1.82 %	411 s
			50	8.24 %	451 s	0.93 %	462 s	0.99 %	456 s
(40,40,400)	1.841×10^8	1892	20	2.36 %	535 s	1.10 %	526 s	0.84 %	540 s
			50	3.97 %	617 s	0.80 %	573 s	1.05 %	576 s
(40,40,500)	1.897×10^8	3600	20	0.52 %	661 s	0.58 %	646 s	0.02 %	675 s
			50	2.45 %	737 s	**−0.09 %**	730 s	0.07 %	729 s
(60,60,100)	1.710×10^8	956	20	5.41 %	324 s	0.26 %	335 s	0.65 %	322 s
			30	0.82 %	370 s	0.31 %	381 s	1.91 %	360 s
(60,60,200)	1.700×10^8	3600	20	2.34 %	638 s	0.42 %	617 s	0.77 %	573 s
			30	3.69 %	665 s	0.62 %	655 s	0.93 %	603 s
(60,60,300)	1.713×10^8	3600	20	1.76 %	914 s	0.35 %	891 s	0.81 %	879 s
			50	1.47 %	1115 s	0.24 %	994 s	0.48 %	950 s
(60,60,400)	1.999×10^8	3600	20	1.06 %	1192 s	0.39 %	1137 s	0.54 %	1144 s
			50	5.57 %	1444 s	0.17 %	1307 s	0.42 %	1249 s
(60,60,500)	2.202×10^8	3600	20	0.69 %	1435 s	**−0.92 %**	1420 s	**−0.49 %**	1490 s
			50	4.33 %	1651 s	**−0.84 %**	1624 s	**−1.26 %**	1614 s

Appendix C. Supplier selection and pre-positioning problem

Hu and Dong (Hu and Dong, 2019) presented a joint decision problem of pre-positioning of relief supplies and supplier selection that is to flexible deal with highly uncertainty of demand. In the preparedness phase, the relief agency purchases relief items according to the pre-specified commitment quantity and stock these items at warehouses. The pre-disaster procurement price mainly depends on the order quantity. After a disaster occurs, the relief agency makes an order with required lead time, and cooperated suppliers will produce relief supplies and deliver them to affected regions. This procurement price mainly depends on lead time. The relief agency can also purchase supplies from the supplier's physical inventory that is originally used for satisfying its regular business demand. The relief agency pays extra money as compensation because suppliers may lose their regular customers as well as some profits. They introduced a risk parameter to represent how much supplies will still stay usable after disaster occurrences, and the survival rate of pre-positioned stocks depends on the location and disaster. The data is from their paper. The model is provided as follows.

Table C1
Model notations in supplier selection and pre-positioning problem.

Sets	
I	Set of candidate locations for facility and supplier, indexed by $i, k \in I$
J	Set of demand locations, indexed by $j \in J$
N	Set of facility size categories;$n \in N$
L	Set of supplier size categories;$l \in L$
A	Set of order quantity intervals;$a \in A$
T	Set of lead time intervals; $t \in T$.
S	Set of disaster scenarios;$s \in S$
Parameters	
$D_{j,s}$	Demand of supplies at location j in scenario s
ω_s	Probability of occurrence of scenario s
$H_{i,j}$	Distance between location i and location j
SC_n	Storage capacity of each facility of size category n
CF_n	Unit fixed cost incurred by opening a facility of size category n
PC_l	Production capacity of each supplier of size category l
PI_l	Supplier physical inventory (SPI) of each supplier of size category l
CQ_l	Commitment quantity of each supplier of size category l
$[LQ_{l,a}, UQ_{l,a}]$	Lower and upper quantity breakpoints associated with quantity interval a offered by each supplier of size category l
$CP_{l,a}$	Procurement price associated with quantity interval a offered by each supplier of size category l
$CP'_{l,t}$	Procurement price associated with the lead time interval t offered by each supplier of size category l
δ	Procurement compensation for using SPI per unit
$\alpha_{i,s}$	Proportion of supplies remaining usable after a disaster at location i in scenario s
CT	Unit transportation cost of moving supplies from facilities or suppliers to demand locations
CH	Unit holding cost for surplus at facilities
CU	Unit penalty cost for unsatisfied demand at demand locations
M	A very large positive number
First-Stage Decision Variables	
$x_{n,i}$	Binary decision variable if a facility of size category n is built at location i
$y_{l,k}$	Binary decision variable if a supplier of size category l is selected at location k
$z_{l,a,k}$	Binary decision variable if the agreement with supplier k of size category l and executed by purchasing supplies at quantity interval a
$p_{i,l,a,k}$	Procurement quantity of facility i from supplier k of size category l with quantity interval a before a disaster strikes
Second-Stage Decision Variables	
$q_{i,j,t,s}$	Quantity of supplies of lead time interval t, transported from facility i to demand location j in scenario s
$q'_{l,k,j,t,s}$	Procurement quantity of supplies of lead time interval t, purchased from the SPI of supplier k of size category l, transported to demand location j in scenario s (i.e., postQ_1)
$q''_{l,k,j,t,s}$	Procurement quantity of supplies of lead time interval t, supplier k of size category l produced and transported supplies to demand location j in scenario s (i.e., postQ_2)
Auxiliary Decision Variables	
$e_{n,i}$	Inventory at facility n of size category l
$b_{k,t,s}$	Surplus quantity of supplies produced by supplier k in scenario s at the end of lead time interval t
$h_{i,s}$	Surplus inventory at facility i in scenario s after responding to the disaster
$u_{j,t,s}$	Unsatisfied demand at location j in scenario s at the end of lead time interval t

The objective function (C1) of this model minimizes the total expected cost, which includes fixed cost of establishing facility (Eq (C2)), procurement cost (Eqs (C3)-(C5)), transportation cost (Eqs (C6)-(C7)), holding cost (Eq (C8)), as well as penalty cost for the shortage of supplies (Eq (C9)). Constraints (C10) and (C11) limit the number of open facilities and the number of selected suppliers at each location to one at most. Constraint (C12) guarantees that the total pre-disaster procurement quantity from every supplier is consistent with supplier size category and quantity interval. Constraint (C13) restricts only one quantity interval at most can be chosen for the selected suppliers. Constraint (C14) computes inventory at each facility before disasters occur. Constraint (C15) ensures that the amount of supplies stocked at open facilities does not exceed the facility capacity. Constraint (C16) restricts that purchases of relief supplies can only be made from selected suppliers and have to follow the agreement on the commitment quantity. Constraint (C17) computes the surplus of supplies at each facility from pre-disaster procurement after responding to one disaster. Constraint (C18) limits postQ_1 by SPI. Constraint (C19) initializes the amount of produced supplies to zero, and therefore constraint (C20) initializes postQ_2 to 0 also. Constraint (C21) presents the relationship of the surplus quantity of produced supplies between the end of lead time interval t and the end of lead time interval t-1. Constraint (C22) computes the unsatisfied demand. Constraints (C23) – (C26) define decision variables.

$$f = \min \quad fc + pc + tc + \sum_s \omega_s \cdot \left(pc'_s + pc''_s + tc'_s + hc_s + uc_s\right) \tag{C1}$$

$$fc = \sum_n \sum_i CF_n \cdot x_{n,i} \tag{C2}$$

$$pc = \sum_i \sum_l \sum_a \sum_k CP_{l,a} \cdot p_{i,l,a,k} \tag{C3}$$

$$pc_s^{'} = \sum_l \sum_k \sum_j \sum_t (1+\delta) \cdot CP_{l,1} \cdot q_{l,k,j,t,s}^{'}, \quad \forall s \tag{C4}$$

$$pc_s^{''} = \sum_l \sum_k \sum_j \sum_t CP_{l,t}^{'} \cdot q_{l,k,j,t,s}^{''}, \quad \forall s \tag{C5}$$

$$tc = \sum_i \sum_l \sum_a \sum_k CT \cdot H_{i,k} \cdot p_{i,l,a,k} \tag{C6}$$

$$tc_s^{'} = \sum_i \sum_j \sum_t CT \cdot H_{i,j} \cdot q_{i,j,t,s} + \sum_l \sum_k \sum_j \sum_t CT \cdot H_{k,j} \cdot \left(q_{l,k,j,t,s}^{'} + q_{l,k,j,t,s}^{''} \right), \quad \forall s \tag{C7}$$

$$hc_s = \sum_i CH \cdot h_{i,s}, \quad \forall s \tag{C8}$$

$$uc_s = \sum_j \sum_t CU \cdot u_{j,t,s}, \quad \forall s \tag{C9}$$

$$\sum_n x_{n,i} \leqslant 1, \quad \forall i \tag{C10}$$

$$\sum_l y_{l,k} \leqslant 1, \quad \forall k \tag{C11}$$

$$LQ_{l,a} \cdot z_{l,a,k} \leqslant \sum_i p_{i,l,a,k} \leqslant UQ_{l,a} \cdot z_{l,a,k}, \quad \forall l,a,k \tag{C12}$$

$$\sum_a z_{l,a,k} \leqslant y_{l,k}, \quad \forall l,k \tag{C13}$$

$$\sum_n e_{n,i} = \sum_l \sum_a \sum_k p_{i,l,a,k}, \quad \forall i \tag{C14}$$

$$e_{n,i} \leqslant x_{n,i} \cdot SC_n, \quad \forall n,i \tag{C15}$$

$$\sum_l y_{l,k} \cdot CQ_l \leqslant \sum_i \sum_l \sum_a p_{i,l,a,k} \leqslant M \cdot \sum_l y_{l,k}, \quad \forall k \tag{C16}$$

$$h_{i,s} = \sum_n \alpha_{i,s} \cdot e_{n,i} - \sum_j \sum_t q_{i,j,t,s}, \quad \forall i,s \tag{C17}$$

$$\sum_l \sum_j \sum_t q_{l,k,j,t,s}^{'} \leqslant \sum_l \alpha_{k,s} \cdot y_{l,k} \cdot PI_l, \quad \forall k,s \tag{C18}$$

$$b_{k,1,s} = 0, \quad \forall k,s \tag{C19}$$

$$q_{l,k,j,1,s}^{''} = 0, \quad \forall l,k,j,s \tag{C20}$$

$$b_{k,t,s} = b_{k,t-1,s} + \sum_l \alpha_{k,s} \cdot y_{l,k} \cdot PC_l - \sum_l \sum_j q_{l,k,j,t,s}^{''}, \quad t \geqslant 2, \forall k,s \tag{C21}$$

$$u_{j,t,s} = D_{j,s} - \sum_i q_{i,j,t,s} - \sum_l \sum_k q_{l,k,j,t,s}^{'} - \sum_l \sum_k q_{l,k,j,t,s}^{''}, \quad \forall j,t,s \tag{C22}$$

$$x_{n,i} \in \{0,1\}, \quad \forall n,i \tag{C23}$$

$$y_{l,k} \in \{0,1\}, \quad \forall l,k \tag{C24}$$

$$z_{l,a,k} \in \{0,1\}, \quad \forall l,a,k \tag{C25}$$

$$e_{n,i}, p_{i,l,a,k}, q_{i,j,t,s}, q_{l,k,j,t,s}^{'}, b_{k,t,s}, h_{i,s}, u_{j,t,s} \in \mathrm{Z}^+, \quad \forall n,l,a,i,j,k,t,s \tag{C26}$$

Table C.2 shows the performance of ***Gurobi***, ***SKR, SKRC,*** and ***SKS*** in different cases for the proposed supplier selection and pre-positioning problem. For cases with $(I, J, S) = (40{,}40{,}300 \sim 500)$ and $N = 50$, the sample problems become hard to be solved by

Gurobi. We thus set its optimality gap to 5 %. Nevertheless, the algorithms still outperform ***Gurobi*** in finding good solutions with less time. The results report that as the number of locations and scenarios increases, compared to ***Gurobi***, ***SKRC*** and ***SKS*** can find better solutions with less time. For large-scalar cases, i.e., $(I, J, S) = (30,30,400 \sim 500)$ and $(40,40,200 \sim 500)$, ***SKRC*** and ***SKS*** can both produce better upper bound than ***Gurobi*** and ***SKR*** with less time. The average gap of ***SKR***, ***SKRC***, and ***SKS*** for the cases of 20 locations is respectively 1.06 %, 0.10 %, 0.12 %. For the cases of 30 locations, that is −0.48 %, −1.77 % and −1.58 %; for the cases of 40 locations, that is −15.38 %, −17.32 % and −17.23 %. The results reveal that ***SKRC*** can produce the best upper bound for all cases, followed by ***SKS***, then ***SKR***. However, ***SKRC*** consumes more time than ***SKS*** in most cases.

Table C2
Performance of Gurobi, SKR, SKRC and SKS in different cases.

(I, J, S)	***Gurobi***		N	***SKR***		***SKRC***		***SKS***	
	Upper bound	CPU time		GAP	CPU time	GAP	CPU time	GAP	CPU time
(20,20,100)	1.355×10^7	63	20	2.39 %	138 s	0.38 %	112 s	0.18 %	95 s
			30	0.41 %	139 s	**−0.08 %**	138 s	0.89 %	177 s
(20,20,200)	1.682×10^7	295	20	2.06 %	264 s	0.24 %	243 s	0.37 %	189 s
			30	1.46 %	296 s	**−0.23 %**	293 s	0.24 %	241 s
(20,20,300)	1.591×10^7	737	20	3.81 %	441 s	**−0.10 %**	401 s	0.34 %	298 s
			50	1.77 %	768 s	**−0.31 %**	1170 s	0.43 %	591 s
(20,20,400)	1.639×10^7	2115	20	0.71 %	490 s	**−0.03 %**	554 s	0.20 %	399 s
			50	1.35 %	653 s	0.01 %	799 s	**−0.05 %**	716 s
(20,20,500)	1.725×10^7	2874	20	0.97 %	480 s	0.05 %	599 s	0.33 %	619 s
			50	1.06 %	970 s	0.10 %	1406 s	0.12 %	757 s
(30,30,100)	2.020×10^7	175	20	2.37 %	398 s	0.12 %	398 s	0.51 %	342 s
			30	1.05 %	677 s	0.28 %	677 s	0.54 %	360 s
(30,30,200)	2.039×10^7	1334	20	1.78 %	511 s	0.15 %	511 s	0.16 %	498 s
			30	1.87 %	713 s	**−0.12 %**	713 s	0.16 %	569 s
(30,30,300)	3.118×10^7	3600	20	1.07 %	755 s	**−0.08 %**	755 s	0.42 %	702 s
			50	1.44 %	1350 s	**−0.25 %**	1350 s	**−0.31 %**	1278 s
(30,30,400)	3.647×10^7	3600	20	−2.69 %	1046 s	**−3.64 %**	1046 s	**−3.15 %**	930 s
			50	−2.50 %	1465 s	**−3.72 %**	1465 s	**−3.68 %**	1760 s
(30,30,500)	4.138×10^7	3600	20	−4.63 %	1427 s	**−5.28 %**	1427 s	**−5.16 %**	1162 s
			50	−4.56 %	1934 s	**−5.20 %**	1934 s	**−5.29 %**	1745 s
(40,40,100)	2.565×10^7	1970	20	4.16 %	578 s	0.46 %	588 s	0.34 %	579 s
			30	1.04 %	1315 s	**−0.04 %**	823 s	0.08 %	827 s
(40,40,200)	3.393×10^7	3600	20	−18.14 %	913 s	**−20.30 %**	856 s	**−20.30 %**	736 s
			30	−17.72 %	1329 s	**−20.64 %**	1009 s	**−20.43 %**	1029 s
(40,40,300)	4.954×10^7	3600	20	−16.17 %	1399 s	**−17.24 %**	1521 s	**−17.26 %**	1247 s
			50	−14.19 %	1280 s	**−17.47 %**	2762 s	**−17.35 %**	1199 s
(40,40,400)	5.762×10^7	3600	20	−18.85 %	1330 s	**−20.68 %**	1610 s	**−20.70 %**	1319 s
			50	−19.49 %	1533 s	**−20.84 %**	2307 s	**−20.60 %**	1568 s
(40,40,500)	7.024×10^7	3600	20	−27.70 %	1760 s	**−28.21 %**	1867 s	**−28.12 %**	1695 s
			50	−26.71 %	2012 s	**−28.23 %**	2748 s	**−27.93 %**	1962 s

Mendeley data

Data for parameters of the proposed model and algorithms can be found on https://data.mendeley.com/preview/kcjftcdxyz?a=602bac27-ef73-4f25-8a3d-b7a346f48251.

References

Aghajani, M., Torabi, S.A., Heydari, J., 2020. A novel option contract integrated with supplier selection and inventory prepositioning for humanitarian relief supply chains. Socioecon. Plann. Sci. 100780.

Amadeo, K. (2018). Hurricane Harvey facts, damage and costs. Available at: https://www.lamar.edu/_files/documents/resilience-recovery/grant/recovery-and-resiliency/hurric2.pdf (accessed data April 23, 2022).

Arani, H.V., Rabbani, M., Rafiei, H., 2016. A revenue-sharing option contract toward coordination of supply chains. Int. J. Prod. Econ. 178, 42–56.

Arthur, D., Vassilvitskii, S. (2007). K-meanspp: the advantages of careful seeding. In: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms. Society for Industrial and Applied Mathematics, Louisiana, New Orleans, pp. 1027–1035.

Aslan, E., Çelik, M., 2019. Pre-positioning of relief items under road/facility vulnerability with concurrent restoration and relief transportation. IISE Transactions 51 (8), 847–868.

Balcik, B., Ak, D., 2014. Supplier selection for framework agreements in humanitarian relief. Prod. Oper. Manag. 23 (6), 1028–1041.

Birge, J.R., Louveaux, F., 2011. Introduction to stochastic programming. Springer Science & Business Media.

Bohner, C., Minner, S., 2017. Supplier selection under failure risk, quantity and business volume discounts. Comput. Ind. Eng. 104, 145–155.

Chen, Y., Zhao, Q., Huang, K., Xi, X., 2022. A Bi-objective optimization model for contract design of humanitarian relief goods procurement considering extreme disasters. Socioecon. Plann. Sci. 81, 101214.

Cochran, W.G., 1977. Sampling Techniques. John Wiley & Sons, New York.

Dabbagh, R., Khoshsirat, M., Bozorgi-Amiri, A., 2018. A multi-attribute reverse auction framework under uncertainty to the procurement of relief items. Int. J. Eng. 31 (12), 2068–2076.
Dalal, J., Üster, H., 2018. Combining worst case and average case considerations in an integrated emergency response network design problem. Transp. Sci. 52 (1), 171–188.
Department of Homeland Security. (2014). FEMA's logistics supply chain management system may not be effective during a catastrophic disaster. Available at: https://www.oig.dhs.gov/assets/Mgmt/2014/OIG_14-151_Sep14.pdf (accessed data June 25, 2019).
Emelogu, A., Chowdhury, S., Marufuzzaman, M., Bian, L., Eksioglu, B., 2016. An enhanced sample average approximation method for stochastic optimization. Int. J. Prod. Econ. 182, 230–252.
Ghavamifar, A., Torabi, S.A., Moshtari, M., 2022. A hybrid relief procurement contract for humanitarian logistics. Transportation Research Part e: Logistics and Transportation Review 167, 102916.
Haimes, Y. Y., Ladson, L. & Wismer, D. A. (1971). Bicriterion Formulation of Problems of Integrated System Identification and System Optimization. 10017–2394. IEEE Institute of Electrical and Electronics Engineers (IEEE), New York, NY, pp. 296-297.
Hammami, R., Temponi, C., Frein, Y., 2014. A scenario-based stochastic model for supplier selection in global context with multiple buyers, currency fluctuation uncertainties, and price discounts. Eur. J. Oper. Res. 233 (1), 159–170.
Homem-De-Mello, T., 2003. Variable-sample methods for stochastic optimization. ACM Transactions on Modeling and Computer Simulation (TOMACS) 13 (2), 108–133.
Hu, S., Dong, Z.S., 2019. Supplier selection and pre-positioning strategy in humanitarian relief. Omega 83, 287–298.
Hu, S., Dong, Z.S., Lev, B., 2022. Supplier selection in disaster operations management: Review and research gap identification. Socioecon. Plann. Sci. 101302.
Hu, S.L., Han, C.F., Meng, L.P., 2017. Stochastic optimization for joint decision making of inventory and procurement in humanitarian relief. Comput. Ind. Eng. 111, 39–49.
Hu, Z., Tian, J., Feng, G., 2019. A relief supplies purchasing model based on a put option contract. Comput. Ind. Eng. 127, 253–262.
Jabbarzadeh, A., Azad, N., Verma, M., 2020. An optimization approach to planning rail hazmat shipments in the presence of random disruptions. Omega 96, 102078.
Jin, Z., Wang, X., Sun, J., Xu, Q., 2022. Optimization of fleet structure and investment evaluation–the cargo owner's fleet perspective. Maritime Business Review 7 (3), 239–254.
Kapoor, A., Singhal, A., 2017. A comparative study of K-Means, K-Means++ and Fuzzy C-Means clustering algorithms. In: 2017 3rd International Conference on Computational Intelligence & Communication Technology (CICT). IEEE, pp. 1–6.
Kaur, H., Singh, S.P., 2021. Multi-stage hybrid model for supplier selection and order allocation considering disruption risks and disruptive technologies. Int. J. Prod. Econ. 231, 107830.
Kleywegt, A.J., Shapiro, A., Homem-de-Mello, T., 2002. The sample average approximation method for stochastic discrete optimization. SIAM J. Optim. 12 (2), 479–502.
Klibi, W., Ichoua, S., Martel, A., 2018. Prepositioning emergency supplies to support disaster relief: a case study using stochastic programming. INFOR: Information Systems and Operational Research 56 (1), 50–81.
Liang, L., Wang, X., Gao, J., 2012. An option contract pricing model of relief material supply chain. Omega 40 (5), 594–600.
Liu, Y., Tian, J., Feng, G., Hu, Z., 2019. A relief supplies purchasing model via option contracts. Comput. Ind. Eng. 137, 106009.
Minh, C.C., Noi, N.V., 2023. Optimising truck arrival management and number of service gates at container terminals. Maritime Business Review 8 (1), 18–31.
Moshtari, M., Nezih, A., Heikkila, J., Gonçalves, P., 2021. Procurement in humanitarian organizations: Body of knowledge and practitioner's challenges. Int. J. Prod. Econ. 233, 108017.
Nourmohamadi Shalke, P., Paydar, M.M., Hajiaghaei-Keshteli, M., 2018. Sustainable supplier selection and order allocation through quantity discounts. International Journal of Management Science and Engineering Management 13 (1), 20–32.
Nur, F., Aboytes-Ojeda, M., Castillo-Villar, K.K., Marufuzzaman, M., 2021. A two-stage stochastic programming model for biofuel supply chain network design with biomass quality implications. IISE Transactions 53 (8), 845–868.
Olanrewaju, O.G., Dong, Z.S., Hu, S., 2020. Supplier selection decision making in disaster response. Comput. Ind. Eng. 143, 106412.
Pagnoncelli, B.K., Ahmed, S., Shapiro, A., 2009. Sample average approximation method for chance constrained programming: theory and applications. J. Optim. Theory Appl. 142 (2), 399–416.
Pamucar, D., Torkayesh, A.E., Biswas, S., 2023. Supplier selection in healthcare supply chain management during the COVID-19 pandemic: a novel fuzzy rough decision-making approach. Ann. Oper. Res. 328 (1), 977–1019.
Pasupathy, R., 2010. On choosing parameters in retrospective-approximation algorithms for stochastic root finding and simulation optimization. Oper. Res. 58 (4-part-1), 889–901.
Pinelli, J.P., Roueche, D., Kijewski-Correa, T., Plaz, F., Prevatt, D., Zisis, I., Moravej, M., 2018. Overview of damage observed in regional construction during the passage of Hurricane Irma over the State of Florida. In: Forensic Engineering 2018: Forging Forensic Frontiers. American Society of Civil Engineers, Reston, VA, pp. 1028–1038.
Rawls, C.G., Turnquist, M.A., 2010. Pre-positioning of emergency supplies for disaster response. Transp. Res. B Methodol. 44 (4), 521–534.
Rezaei-Malek, M., Tavakkoli-Moghaddam, R., Cheikhrouhou, N., Taheri-Moghaddam, A., 2016. An approximation approach to a trade-off among efficiency, efficacy, and balance for relief pre-positioning in disaster management. Transportation Research Part e: Logistics and Transportation Review 93, 485–509.
Royset, J.O., 2013. On sample size control in sample average approximations for solving smooth stochastic programs. Comput. Optim. Appl. 55 (2), 265–309.
Sabbaghtorkan, M., Batta, R., He, Q., 2020. Prepositioning of assets and supplies in disaster operations management: Review and research gap identification. Eur. J. Oper. Res. 284 (1), 1–19.
Shamsi, N.G., Torabi, S.A., Shakouri, H.G., 2018. An option contract for vaccine procurement using the SIR epidemic model. Eur. J. Oper. Res. 267 (3), 1122–1140.
Shokr, I., Jolai, F., Bozorgi-Amiri, A., 2022. A collaborative humanitarian relief chain design for disaster response. Comput. Ind. Eng. 172, 108643.
Shokr, I., Torabi, S.A., 2017. An enhanced reverse auction framework for relief procurement management. Int. J. Disaster Risk Reduct. 24, 66–80.
Thevenin, S., Ben-Ammar, O., Brahimi, N., 2022. Robust optimization approaches for purchase planning with supplier selection under lead time uncertainty. Eur. J. Oper. Res. 303 (3), 1199–1215.
Torabi, S.A., Baghersad, M., Mansouri, S.A., 2015. Resilient supplier selection and order allocation under operational and disruption risks. Transportation Research Part e: Logistics and Transportation Review 79, 22–48.
Verweij, B., Ahmed, S., Kleywegt, A.J., Nemhauser, G., Shapiro, A., 2003. The sample average approximation method applied to stochastic routing problems: a computational study. Comput. Optim. Appl. 24 (2–3), 289–333.
Wang, X., Li, F., Liang, L., Huang, Z., Ashley, A., 2015. Pre-purchasing with option contract and coordination in a relief supply chain. Int. J. Prod. Econ. 167, 170–176.
Wang, W., Wang, S., Zhen, L., Qu, X., 2022. EMS location-allocation problem under uncertainties. Transportation Research Part e: Logistics and Transportation Review 168, 102945.
Wang, D., Yang, K., Yang, L., Dong, J., 2023. Two-stage distributionally robust optimization for disaster relief logistics under option contract and demand ambiguity. Transportation Research Part e: Logistics and Transportation Review 170, 103025.
Zheng, Q., Zhou, L., Fan, T., Ieromonachou, P., 2019. Joint procurement and pricing of fresh produce for multiple retailers with a quantity discount contract. Transportation Research Part e: Logistics and Transportation Review 130, 16–36.