Advances in the decision-making of set covering models under uncertainty

Lakmali Weerasena[1] · Chathuri Aththanayake[1] · Damitha Bandara[2]

Abstract
The Set Covering Problem (SCP) has been an extensively studied NP-hard problem in the field of combinatorial optimization since 1970. Over the past five decades, a significant amount of research has led to the development of a diverse set of covering models to support decision-making in various areas. However, the SCPs related to real-world applications are often too complex to solve using existing algorithms due to uncertain problem parameters. Thus, given the diversity of new developments, there is a pressing need to know both the current solution approaches and the advanced strategies for studying the uncertain SCP. This study summarizes the various modeling and solution approaches to the SCP when the model parameters are uncertain. Further, this study discusses some promising future research directions of the uncertain SCP that will impact new investigations of decisions on complex and competitive real-world issues.

Keywords Decision-making · NP-problems · Optimization models · Location models · Set covering problem · Uncertain optimization

1 Introduction

The Set Covering Problem (SCP) has been an extensively studied NP-hard problem in combinatorial optimization since 1970. The SCP can be described by a collection of m items $E = \{e_1, e_2, \ldots, e_m\}$ with the index set $I = \{1, 2, \ldots, m\}$ and a family of n subsets $S = \{S_1, S_2, \ldots, S_n\}$ of E with the index set $J = \{1, 2, \ldots, n\}$. The goal of the SCP is to find a collection of subsets $\mathcal{C} \subseteq S$ involving a minimum cost to cover all of these elements,

✉ Lakmali Weerasena
lakmali-weerasena@utc.edu

Chathuri Aththanayake
swc213@mocs.utc.edu

Damitha Bandara
damitha-bandara@utc.edu

1   Department of Mathematics, University of Tennessee at Chattanooga, 615 McCallie Avenue, Chattanooga, TN 37403-2598, USA

2   Gary W. Rollins College of Business, 615 McCallie Avenue, Chattanooga, TN 37403, USA

with the collection $\mathcal{C}$ being referred to as a cover. The earliest studies on the SCP were conducted by Lemke et al. (1971) and Toregas et al. (1971). The mathematical formulation of the SCP, along with other notation essential for this formulation, is provided below:

$$\min z(x) = \sum_{j \in J} c_j x_j \tag{1a}$$

$$\text{s.t.} \sum_{j \in J_i} a_{ij} x_j \geq b_i, \quad \forall i \in I \tag{1b}$$

$$x_j \in \{0, 1\}, \quad \forall j \in J, \tag{1c}$$

$x_j = 1$ if the set S_j for $j \in J$ is selected to cover an item e_i for $i \in I$; 0 otherwise
$c_j :=$ cost of set S_j for $j \in J$
$a_{ij} :=$ constraint coefficient for $i \in I$, $j \in J$
$J_i = \{j \in J : a_{ij} = 1\} :=$ the index set of sets that can cover items e_i for $i \in I$
$b_i :=$ the demand value of constraint $i \in I$

If $c_j = 1$ for all $j \in J$, the model (1a–1c) is equivalent to finding a cover with a minimum number of subsets in the collection $\mathcal{C}$.

The SCP is important in pedagogical and practical areas. In the linear-relaxed version of the SCP, where the integrality requirement is relaxed to linear constraints, the integrality gap is bounded by at most $\log m$ (Chvatal, 1979; Slavík, 1996; Grossman & Wool, 1997). Therefore, studying the SCP provides insight into the use of approximation algorithms in solving NP-hard problems and, as such, is a prominent example in teaching approximation algorithms. The simplicity of the mathematical model of the SCP aids in visualizing the importance of the approximation methods needed for effective educational use.

The initial covering model by Toregas et al. (1971) focused on locating emergency service facilities optimally. Since then, numerous generalized covering models have been developed, including the first type, known as the SCP, where all elements must be covered by at least one set for a specific objective function. When $b_i = 1$ for $i \in I$, we refer to the mathematical formulation of this model as the classical SCP (1a–1c). In contrast, the second type optimizes the number of covered items under specific constraints.

The deterministic approaches to solving the SCP are widely discussed and applied in studies, including literature reviews (Schilling, 1993; Farahani et al., 2012; Wang et al., 2021) of various deterministic covering models. Schilling (1993) provided the first comprehensive review of covering problems, while Farahani et al. (2012) reviewed covering models, solutions, and their applications. Bélanger et al. (2019)'s review focused on early work studying static ambulance location problems. In addition, this study concentrated on new techniques for addressing tactical and operational decisions, including a summary of the interaction between the two. The most recent survey on covering models conducted by Wang et al. (2021) focused on mathematical covering models and their applications in emergency facility location problems. After excluding the studies that did not meet their inclusion criteria, they found 87 related papers from three databases (Sc-opus, Google Scholar, and Science Citation Index). However, these reviews did not explicitly focus on the model (1a–1c); rather they discussed classical SCP and maximal covering problems together. Further, the surveys conducted by Schilling (1993) and Farahani et al. (2012) focused on only deterministic approaches, while the study conducted by Wang et al. (2021) focused on deterministic and probabilistic covering models only considering emergency facility location, relocation, and dispatching problems.

1.1 Classification of uncertainty

More recently, a significant number of new developments have been proposed in decision-making under uncertainty focusing on the classical SCP. Applying exact or heuristic methods to make decisions related to real-world applications of the SCP often faces difficulty when there is uncertainty in the input data. These uncertainties are classified as probabilistic, stochastic, and robust based on the uncertainty of particular input data or combination of input data. Probabilistic uncertainties involve quantifying unpredictability using probability distributions (Aly & White, 1978), while stochastic uncertainties incorporate randomness or variability, often with time-dependent variables (Powell, 2019) and robust uncertainties address situations where the exact values or distributions of uncertain parameters are unknown, but their potential range of values is considered (Pereira & Averbakh, 2013). However, a literature review on the SCP from the perspective of parameter uncertainty has not yet been conducted, emphasizing the need for such a review. Therefore, a broad and deep examination of classical SCP from the perspective of an uncertainty approach supporting decision-making, which has yet to be conducted, merits study.

In this survey, we restrict the search only to focus on the model (1a–1c). More specifically, our study focused on the existing models studying the uncertain input parameters a_{ij}, b_i, c_j; the decision variable x_j for $i \in I$, $j \in J$ of the model (1a–1c); and satisfying the covering constraints with the chance constraint $P(\sum_{j \in J_i} a_{ij} x_j \geq 1) \geq \alpha_i$ where α_i is a threshold level.

1.2 Survey scope and method

We searched for studies in ScienceDirect, Google Scholar, and Connected Papers based on the keywords ("Set Covering Problems", "Probabilistic", "Stochastic" and "Robust"), using Boolean operators, article type, and subject area. The subject areas that we focused on include "Computer Science", "Decision Science", "Engineering", "Environmental Science" and "Mathematics". The initial search resulted in papers focused on the uncertainty approach of SCP. Then, we extended the search process by including the studies cited in these papers. We excluded the studies that did not introduce a mathematical model and those that focused on the uncertainty of equivalent mathematical models such as maximal covering models. This process resulted in 16 studies that considered uncertainty counterparts of SCP in the model (1a–1c) which are summarized in Table 1. The number of studies published every 10 years since 1970 is listed in Table 2. Based on Table 2, the number of publications supporting decision-making under uncertain SCP has slowly increased over the past 50 years. Most of the studies have focused on the uncertainty of constraints' coefficients or chance constraints. Further, we observed that the uncertain version of SCPs received significant attention among researchers after 2000, meaning this topic is relatively new.

1.3 Real-world applications devoted to decision-making in the uncertain SCP

We discovered 16 studies, seven directly addressing the theoretical aspect of uncertainty in SCP, while nine discussed real-world scenarios of uncertainty in SCP. In the SCP formulation, a_{ij} ensures the covering of item e_i by the set S_j with certainty and b_i ensures the covering of each item at least by one set S_j. As evident in the real-world scenario studies examined below, we observe that when the problem parameters are uncertain, there is no guarantee that we make decisions satisfying these conditions with 100% accuracy.

Table 1 Solution techniques

Uncertainty of constraints' coefficients a_{ij} for $i \in I$, $j \in J$	
Solution technique	References
Probabilistic formulation	Aly and White (1978)
Probabilistic formulation	Benveniste (1982)
0–1 Programming model	Hwang (2002)
Heuristic algorithm	Cabeza et al. (2004)
0–1 Programming model	Fischetti and Monaci (2012)
Compact mixed-integer linear programming	Degel and Lutter (2013)
Non-compact integer linear model formulations	Lutter et al. (2017)
Chance-constrained program by sampling based approach	Wu and Kucukyavuz (2019a)
Mixed-integer programming model	Wu and Kucukyavuz (2019b)
Uncertainty of objective function's coefficients c_j for $j \in J$	
Min–max strategy	Pereira and Averbakh (2013)
Uncertainty of demand values b_i for $i \in I$	
Chance-constrained 0–1 integer programming	ReVelle and Hogan (1989)
Chance-constrained 0–1 integer programming	Marianov and Revelle (1994)
Enumeration algorithms and branch-and-bound algorithms	Beraldi and Ruszczyński (2002a)
Beam search heuristic strategy	Beraldi and Ruszczyński (2005)
Mixed-integer programming reformulation	Saxena et al. (2010)
Compact mixed-integer linear programming	Ding et al. (2020)

1. Uncertainty of constraints' coefficients a_{ij} of the SCP

   (a) Aly and White (1978) discussed locating emergency service facilities by considering that the location of incidents such as accidents, fires, or customers are random variables, with customers being considered as items and sites as sets (that is a_{ij} is a random variable).

   (b) Hwang (2002) conducted a study on logistic system design with the aim of optimizing the performance of logistic system. In this study, customers are taken as the items, and warehouse or distribution centers (W/D) are considered as sets. It is assumed that the probability of each demand point being covered is not less than a specified threshold value.

   (c) Cabeza et al. (2004) discussed the probabilistic approach of SCP from the perspective of biodiversity, specifically focusing on selecting reserve networks that represent biodiversity efficiently. In this study, the sites and species represent the sets and items respectively for the modeling. It is assumed that a minimum probability level is required to represent the species considered.

   (d) Fischetti and Monaci (2012) and Ahmed and Papageorgiou (2013) focused on the uncertainty of a_{ij} for $i \in I$, $j \in J$ as a binary random variable indicating whether set S_j covers the item e_i depending on the probability of the disappearance of the decision variable x_j for $j \in J$.

   (e) Degel and Lutter (2013) applied robust SCP to an emergency medical service facility problem. Items are taken as the demand nodes and the sets as the location of the emergency service facility site.

Table 2 The number of publications of uncertain SCPs since 1970

Year	Uncertainty of			References
	Constraints' coefficients	Objective function's coefficients	Demand Values	
1970s	1	–	–	Aly and White (1978)
1980s	1	–	1	Benveniste (1982); ReVelle and Hogan (1989)
1990s	–	–	1	Marianov and Revelle (1994)
2000s	2	–	2	Hwang (2002), Cabeza et al. (2004), Beraldi and Ruszczyński (2002b, 2005)
2010s	5	1	1	Fischetti and Monaci (2012), Degel and Lutter (2013), Lutter et al. (2017), Wu and Kucukyavuz (2019a, b), Pereira and Averbakh (2013), Saxena et al. (2010)
2020s	–	–	1	Ding et al. (2020)
Total	9	1	6	

2. Uncertainty of satisfying demand values b_i

   (a) ReVelle and Hogan (1989) and Marianov and Revelle (1994) discussed an application of probabilistic SCP modeling in queuing theory for emergency vehicle location problems. Items are considered as the demand points and sets as eligible sites for facilities. These two studies are from a different perspective, where the uncertainty of demand values is addressed by a concept referred to as the "busy fraction". Various versions of these busy fractions were introduced by researchers based on the real-world scenario.

   (b) Ding et al. (2020) discussed uncertainty in the SCP when the coverage of each item is required by two types of facilities. Items are seen as demand nodes and two types of facility locations are represented in two different sets.

The review is organized as follows. Sections 2–4 discuss these 16 models in detail, focusing on three main uncertain categories of the SCP. First, we review the literature on uncertain input parameters a_{ij} for $i \in I$, $j \in J$. Second, we review the literature on the studies ensuring that the probability of meeting the coverage constraints $\sum_{j \in J_i} a_{ij} x_j \geq 1$, for $i \in I$ above a certain threshold level α. Third, we review the literature on uncertain input parameters b_i for $i \in I$. When introducing the models, we use ":=" to define new notation and "=" for representing equations. Further, the notation introduced in the classical model (1a–1c) is commonly used in the models described in this review. If the additional notation is used to describe a particular model, we will include them above each model. Section 5 concludes the study and discusses promising future research directions for the uncertain SCP, which will impact new investigations of complex and competitive real-world issues.

2 Uncertainty of constraints' coefficients

The first category of SCP models focuses on the uncertainty of constraint coefficient a_{ij} for $i \in I$, $j \in J$ and their existing applications. Initially, Soyster (1973) introduces the uncertainty of the constraint coefficients in linear models. Since there is no guarantee that the classical covering constraint (1b) can be satisfied with a_{ij} uncertainty, the majority of the models reviewed in this section utilize the chance-constrained method for modeling, one of the primary approaches used to solve optimization problems under various uncertainties. Employing this process to formulate an optimization model ensures that the probability of meeting the coverage constraints is above a specified level.

2.1 Specialized application-focused models

In this section, we discuss four studies that specifically target uncertain input parameters a_{ij}, where $i \in I$ and $j \in J$. These models have been developed with a focus on particular applications.

The pioneering work on probabilistic covering by Aly and White (1978) explored the formulation of a service location problem. Focusing on emergency health services, they considered the incident location as a random variable. In their model, items were treated as customers, and sets were defined as sites. The notation of the model is provided below:

$k :=$ number of facilities available

$c_j :=$ the cost of locating a facility at site j

$t_{ij} :=$ response time from an emergency facility at site j to an incident in region i

$\lambda_i :=$ upper bound on the response time from location j to an incident in region i

$\gamma_i :=$ required service (aspiration) level, $0 \leq \gamma_i \leq 1$

$P(t_{ij} \leq \lambda_i) :=$ probability of the time $t_{ij} \leq$ to upper bound on the response time λ_i

$x_j = 1$ if a facility is to be located at j; 0 otherwise

$\theta(x) = \{j | x_j = 1\} :=$ set of sites where a facility has been located

Their formulation is given by model (2a–2c), which is a non-linear integer programming problem:

$$\min z(x) = \sum_{j \in J} x_j \tag{2a}$$

$$\text{s.t } P(t_{ij} \leq \lambda_i) \leq \gamma_i, \quad \text{for some} \quad j \in \theta(x), \quad \forall i \in I \tag{2b}$$

$$x_j \in \{0, 1\}, \quad \forall j \in J. \tag{2c}$$

The distance traveled from the emergency unit to the incident location is assumed to be random, characterizing the traveling time t_{ij} as a random variable. This study introduces a cumulative distribution function to determine $P(t_{ij} \leq \lambda_i)$. Assuming the availability of at least one unit at site j, the variable a_{ij} signifies coverage by the unit at site j for an incident in subregion i. The model is solved by defining the input variable $a_{ij} = 1$ if $P(t_{ij} \leq \lambda_i) \geq \gamma_i$ and 0 otherwise. In this model, the objective function (2a) minimizes the number of selected sites. Chance constraint (2b) implies that any incident in $i \in I$ must be covered for some $j \in \theta(x)$. With this information, the model (2a–2c) is reformulated as in model (3a–3c), which is equivalent to the classical SCP (1a–1c) and, thus, can be solved using its solution methods:

$$\min z(x) = \sum_{j \in J} c_j x_j \tag{3a}$$

$$\text{s.t.} \sum_{j \in J_i} a_{ij} x_j \geq 1, \quad \forall i \in I \tag{3b}$$

$$x_j \in \{0, 1\}, \quad \forall j \in J. \tag{3c}$$

The study's theoretical approach is used to locate emergency service facilities in an urban environment. The solution approach involves partitioning the emergency service region (such as a city or country) into m rectangular subregions, guided by specific assumptions. Potential locations for new sites are then identified, taking into account existing facilities to mitigate additional fixed costs.

Benveniste (1982) further discusses the derivation of the probability distribution of the travel time t_{ij} proposed by Aly and White (1978). This author continues using the same model given by (2a–2c). The chance constraints $P(t_{ij} \leq \lambda_i) \leq \gamma_i$ for some $j \in \theta(x)$ are linearized, deriving the term $P(t_{ij} \leq \lambda_i)$ as the proportion of incidents in subregion i covered by facility j. Accordingly, the input parameter $a_{ij} = 1$ if $P(t_{ij} \leq \lambda_i) \geq \gamma_i$ and 0 if $P(t_{ij} \leq \lambda_i) < \gamma_i$.

Hwang (2002) proposes a logistics system that includes plants, warehouses or distribution centers (W/D), and customers. A 0–1 programming model is developed to find the minimum number of W/D among potential sites so that the probability of each demand point being covered is not less than a specified value. The formulation employs a stochastic variant of the SCP, treating items as retailers and sets as supply centers. The notation for the model is provided below:

$x_j = 1$ if the facility is located at point j; 0 otherwise

$c_i :=$ logistic cost incurred for node i

$S_j(XS_j, YS_j) := $ possible location on supply centers
$R_i(XR_i, YR_i) := $ location of retailers
$\text{Dist}(R_i, S_j) = (|XS_j - XR_i|^p + |YS_j - YR_i|^p)^{1/p} := $ distance between R_i and S_j
$D_i := $ demand at R_i
$\alpha_i := $ demand change (increasing or decreasing) rate
$\text{Tim}(R_i, S_j) := $ travel time between R_i and S_j
$F_{ij}(R_i, S_j) := $ logistic cost incurred between R_i and S_j
$F_{ij} = c_i \, \text{Dist}(S_j, R_i) D_i \exp(\alpha_i \text{Tim}(S_j, R_i))$
$A_i := $ required service level
$r_i := $ critical service level
$a_{ij} = 1$ if $P(F_{ij} \leq A_i) \geq r_i$; 0 otherwise

If $p = 1$, then $\text{Dist}(S_j, R_i)$ becomes rectilinear distance, and if $p = 2$, it becomes Euclidean distance. This formulation is given by model (4a–4c), which is a binary integer programming problem:

$$\min z(x) = \sum_{j \in J} x_j \tag{4a}$$

$$\text{s.t.} \sum_{j \in J} a_{ij} x_j \geq 1, \quad \forall i \in I \tag{4b}$$

$$x_j \in \{0, 1\}, \quad \forall j \in J. \tag{4c}$$

The objective function (4a) minimizes the total number of facilities. Constraint (4b) ensures that all the retailers are at least covered by one supply center. This model is then applied to a logistic system design. It is assumed that a W/D center is always available in model (4a–4c). In the second stage of this study, the model (4a–4c) is applied to solve a vehicle routine problem using an improved version of a genetic algorithm. GUI-type programming, an integrated VRP-solver based on a genetic algorithm, is developed to solve the problem. The proposed method has proven effective in addressing logistics system design for multi-warehouse/distribution centers.

Cabeza et al. (2004) discuss probabilistic SCP in relation to biodiversity persistence. Their study develops two reserve selection approaches by considering both habitat models and spatial reserve design. In the formulation, the items are treated as sites and sets as species. The notation of the model is provided below:

$N := $ total number of sites
$p_{ij} := $ probability of finding a species j in site i
$S := $ set of selected sites
$I_i = 1$ for $i \in S$; 0 otherwise
$b := $ boundary length penalty (when $b = 0$, the problem becomes the classical version)
$L' := $ ratio of boundary length of the selected reserve system to the total area
$p_j := $ probability of having at least one occurrence of species j in any site
$c_i := $ cost of including site i
$T_j := $ the minimum probability level that is required to represent species j

Their formulation is given by model (5a–5d).

$$\min z(I) = \sum_{i \in S} I_i c_i + bL' \tag{5a}$$

$$\text{s.t. } p_j \geq T_j, \quad \forall j \in J \tag{5b}$$

$$p_j = 1 - \left(\prod_{i \in S} (1 - p_{ij}) \right), \quad \forall j \in J \tag{5c}$$

$$I_i \in \{0, 1\}, \quad \forall i \in S. \tag{5d}$$

The objective function (5a) minimizes the combination of the number of areas and the boundary length required to represent all target species. Constraint (5b) gives the minimum level of target probability for each species, while constraint (5c) defines that p_j equals the probability of having at least one occurrence of species j in any site i. A case study with a dataset of 26 butterfly species from Creuddyn Peninsula in north Wales illustrates the model. The researchers introduced two heuristic iterative algorithms for comparing the model (5a–5d): a forward algorithm adding sites and a backward algorithm starting with all sites and then excluding them one by one. Results indicate that although the backward algorithm outperforms the forward one, it does not guarantee finding the exact optimal solution.

2.2 Specialized generic models

In this section, we discuss three generic studies that specifically target uncertain input parameters a_{ij} with probabilistic and stochastic uncertainties.

Fischetti and Monaci (2012) introduce a stochastic variant of the SCP, called the Uncertain SCP (USCP). The notation of the model is provided below:

$N = \{1, 2, \ldots, n\} :=$ set of columns
$M = \{1, 2, \ldots, m\} :=$ set of rows
$\bar{P}_i \in [0, 1] :=$ minimum required probability for row i to be covered by at least one selected column j
$P_j \in [0, 1] :=$ disappearing probability of column j
$a_i :=$ coefficients in ith row for $i \in M$
$c_j :=$ costs associated with a column j

The model assumes that the input parameter a_{ij} for $i \in M$, $j \in N$ follows a Bernoulli distribution. Their formulation is given by model (6a–6c), which is a non-linear integer programming problem with a chance constraint.

$$\min z(x) = \sum_{j \in N} c_j x_j \tag{6a}$$

$$\text{s.t. } P\{a_i^T x \geq 1\} \geq \bar{P}_i, \quad \forall i \in M \tag{6b}$$

$$x_j \in \{0, 1\}, \quad \forall j \in N. \tag{6c}$$

The objective function (6a) minimizes the number of selected sets, and constraint (6b) ensures the coverage of item e_i by at least one selected set S_j with the minimum required level of probability $\bar{P}_i$. Following a modeling technique proposed by Haight et al. (2000) and letting $w_j = -\ln p_j$ ($j \in N$) and $\bar{W}_i = -\ln(1 - \bar{P}_i)$, Fischetti and Monaci (2012) propose a new model (7a–7c) which is equivalent to the model (6a–6c). Their new model is a binary linear

integer programming problem given below:

$$\min z(x) = \sum_{j \in N} c_j x_j \tag{7a}$$

$$\text{s.t.} \sum_{j \in J_i} w_j x_j \geq \bar{W}_i, \quad \forall i \in M \tag{7b}$$

$$x_j \in \{0, 1\}, \quad \forall j \in N. \tag{7c}$$

where $J_i = \{j \in N : a_{ij} = 1\}$ for $i \in M$. In this model, the objective function (7a) minimizes the number of selected sets. Constraint (7b) gives the linearized version of constraint (6b). Note the challenging nature of constraints (7b), which take the form of a knapsack-type constraint. The researchers propose a cutting plane model (8a–8c) as a noncompact-integer linear programming solution, modifying (7b) to (8b):

$$\min z(x) = \sum_{j \in N} c_j x_j \tag{8a}$$

$$\text{s.t.} \sum_{j \in J_i} x_j - \sum_{j \in S} x_j \geq 1, \quad S \subseteq J_i : \sum_{j \in S} w_j < \bar{W}_i, \quad i \in M \tag{8b}$$

$$x_j \in \{0, 1\}, \quad j \in N. \tag{8c}$$

where $S \subseteq J_i$ such that $\sum_{j \in S} w_j < \bar{W}$ and $\sum_{j \in S} x_j^* > \sum_{j \in J_i} x_j^* - 1$. The set of constraints (8b) ensures that row $i \in M$ must be covered by a subset of columns which has a small probability of disappearing. A cutting plane algorithm was proposed to solve model (8a–8c). These researchers implemented the cutting plane algorithm in C language, and a CPLEX solver is used for optimization. Further, the performance of the algorithm was evaluated using all NETLIB (2013) instances.

Degel and Lutter (2013) generalize the assumption that the a_{ij} follows a Bernoulli distribution proposed by Fischetti and Monaci (2012) which results from a known column disappearing probability p_j for $j \in J$. They propose individual and independent coefficient disappearing probabilities p_{ij} for $i \in I$, $j \in J$, and define the generalized USCP (GUSCP). The notation of their model is provided below:

$\mathcal{N}_i = \{j \in J \,|\, i \text{ can be covered by } j\} := \text{neighborhood of a given row } i$
$a_{ij} = 1$ if $j \in \mathcal{N}_i$; 0 otherwise
$p_{ij} \in [0, 1] := \text{probability of disappearing the coefficient } a_{ij} \text{ for } i \in I \text{ and } j \in J$
$p_{ij} := \text{individual and independent coefficient disappearing probabilities of } a_{ij}$
$y_j = 1$ if column j is selected; 0 otherwise
$\alpha \in (0, 1] := \text{minimum coverage probability level}$
$\bar{p}_{ij} := \text{nominal value}$
$\hat{p}_{ij} := \text{worst case deviation of } \bar{p}_{ij}$
$\bar{p}_{ij} + \hat{p}_{ij} := \text{worst case scenario}$
$\Gamma_i := \text{robust-}\alpha \text{ cover of row } i$
$c_j := \text{costs associated with column } j$
$\mathcal{C}(y^*) = \{j \in J \,|\, y_j^* = 1\}$ for all $i \in I$

Γ_i is the solution $y^* \in \{0, 1\}^n$ with $P_{\Gamma_i}\left(\sum_{j \in J} a_{ij} y_j^* \geq 1\right) \geq \alpha$. Their formulation is given by model (9a–9c), which is a nonlinear integer programming problem:

$$\min z(y) = \sum_{j \in J} c_j y_j \tag{9a}$$

$$\text{s.t. } P\left(\sum_{j \in J} a_{ij} y_j \geq 1\right) \geq \alpha, \quad \forall i \in I \tag{9b}$$

$$y_j \in \{0, 1\}, \quad \forall j \in J. \tag{9c}$$

Assuming independence, the probability of covering items is represented as $P\left(\sum_{j \in J} a_{ij} y_j^* \geq 1\right) = 1 - \prod_{j \in C(y^*)} p_{ij}$. The model (9a–9c) can be transformed into a linear integer model similar to (7a–7c). In real-world scenarios, the exact value of p_{ij} is unknown. To accommodate this uncertainty, it is assumed that p_{ij} lies within the interval $[\bar{p}_{ij} - \hat{p}_{ij}, \bar{p}_{ij} + \hat{p}_{ij}] \subseteq [0, 1]$. A robust formulation for the GUSCP is introduced, represented by model (10a–10c):

$$\min z(y) = \sum_{j \in J} c_j y_j \tag{10a}$$

$$\text{s.t. } P_{\Gamma_i}\left(\sum_{j \in J} a_{ij} y_j \geq 1\right) \geq \alpha, \quad \forall i \in I \tag{10b}$$

$$y_j \in \{0, 1\}, \quad \forall j \in J. \tag{10c}$$

In this model, the objective function (10a) minimizes the number of selected sets. Chance constraint (10b) gives the minimum coverage probability on the condition that at most Γ_i realizations of p_{ij} are equal to the worst case and $n - \Gamma_i$ other realizations of p_{ij} equal to the nominal value. Constraint (10c) defines the domain of the decision variable y_j. Additionally, the chance constraint (10b) is in non-linear form, thus, it has been reformulated into a linearized form, creating a mixed-integer linear programming problem (MILP) represented by model (11a–11f). This formulation introduces additional non-negative variables ζ_{ij} and η_i to solve the Robust Uncertain Set Covering Problem (RUSCP):

$$\min z(y) = \sum_{j \in J} c_j y_j \tag{11a}$$

$$\text{s.t. } \sum_{j \in J} w_{ij} y_j + \sum_{j \in J} \zeta_{ij} + \gamma_i \eta_i \leq \ln(1 - \alpha), \quad \forall i \in I \tag{11b}$$

$$\zeta_{ij} + \eta_i \geq (w_{ij}' - w_{ij}) y_j, \quad \forall i \in I; \quad \forall j \in J \tag{11c}$$

$$\zeta_{ij} \geq 0, \quad \forall i \in I; \quad \forall j \in J \tag{11d}$$

$$\eta_i \geq 0, \quad \forall i \in I \tag{11e}$$

$$y_i \in \{0, 1\}, \quad \forall j \in J. \tag{11f}$$

where

$$w_{ij}' = \begin{cases} \ln(\bar{p}_{ij} + \hat{p}_{ij}) & \text{if } \bar{p}_{ij} + \hat{p}_{ij} > 0 \\ \ln(1 - \alpha) & \text{if } \bar{p}_{ij} + \hat{p}_{ij} = 0 \end{cases} \forall i \in I \text{ and } j \in J$$

$$w_{ij} = \begin{cases} \ln(\bar{p}_{ij}) & \text{if } \bar{p}_{ij} > 0 \\ \ln(1 - \alpha) & \text{if } \bar{p}_{ij} = 0 \end{cases} \forall i \in I \text{ and } j \in J$$

Objective function (11a) minimizes the number of selected sets. Constraints (11b–11e) represent the reformulation of chance constraint (10b), and constraint (11f) defines the domain of the decision variable y_j. The importance of the proposed MILP model is explained using an emergency medical service facility problem by considering demand nodes as items and sites as sets.

Lutter et al. (2017) focus on an extension of the RUSCP formulation given by model (11a–11f), proposing two non-compact integer linear formulations to solve the RUSCP by converting chance constraint (10b) into a linear form. We provide the additional notation required to define these two new models below:

$\bar{N}_i = \{j \in J | 1 - \bar{p}_{ij} > 0\} :=$ set of all facility location sites being able to cover demand node $i \in I$ with positive nominal probability $1 - \bar{p}_{ij}$

$\tilde{w}_{ij} = w'_{ij} - w_{ij} \geq 0$

The first non-compact reformulation (RUSCP-NCG) is given by the model (12a–12c), and constraint (12b) provides the linearized version of constraint (10b):

$$\min z(y) = \sum_{j \in J} c_j y_j \tag{12a}$$

$$\text{s.t.} \sum_{j \in \bar{N}_i} w_{ij} y_j + \sum_{j \in S} \tilde{w}_{ij} y_j \leq \ln(1 - \alpha), \quad \forall i \in I; \quad \forall S \subseteq \bar{N}_i : \quad |S| \leq \Gamma_i \tag{12b}$$

$$y_j \in \{0, 1\}, \quad \forall j \in J. \tag{12c}$$

The second non-compact formulation, RUSCP-NCS, focuses on the subsets that fail to satisfy the Γ_i-robust α-covering condition. The definition for the Γ_i-robust α-covering condition can be found in Definition 1.

Definition 1 (Γ_i-robust α-cover) Let $i \in I, \Gamma_i \in \mathbb{N}_0, \Gamma = (\Gamma_i)_{i \in I}, \alpha \in [0, 1)$ and let p_{ij} have realization in $[\bar{p}_{ij}, \bar{p}_{ij} + \hat{p}_{ij}] \subseteq [0, 1]$ for all $j \in J$. Define the worst-case coverage probability for a set $\mathcal{C} \subseteq J$ by

$$P_{\Gamma_i}\left(\sum_{j \in \mathcal{C}} a_{ij} \geq 1\right) := 1 - \max_{\{\mathcal{U} \subseteq \mathcal{C} : |\mathcal{U}| \leq \Gamma_i\}} \left\{ \prod_{j \in \mathcal{U}} (\bar{p}_{ij} + \hat{p}_{ij}) \cdot \prod_{j \in \mathcal{C} \setminus \mathcal{U}} \bar{p}_{ij} \right\}.$$

A Γ_i-robust α-cover $\mathcal{C} \subseteq J$ of the i-th demand node has a worst-case coverage probability $P_{\Gamma_i}(\sum_{j \in \mathcal{C}} a_{ij} \geq 1)$ greater or equal to α. A set $\mathcal{C} \subseteq J$ is called a Γ_i-robust α-cover if $\mathcal{C}$ is a Γ_i-robust α-cover for each row $i \in I$.

The second formulation is given by the model (13a–13d) and constraints (13b–13c) provide the linearized version of constraint (10b):

$$\min z(y) = \sum_{j \in J} c_j y_j \tag{13a}$$

$$\text{s.t.} \sum_{j \in \bar{N}_i} y_j \geq 1, \quad \forall i \in I \tag{13b}$$

$$\sum_{j \in \bar{N}_i \setminus S} y_j \geq 1, \quad \forall i \in I; \quad \forall S \subseteq \bar{N}_i : \sum_{j \in S} w_{ij} + \max_{\{\mathcal{U} \subseteq S : |\mathcal{U}| \leq \Gamma_i\}} \sum_{j \in \mathcal{U}} \tilde{w}_{ij} > \ln(1 - \alpha) \tag{13c}$$

$$y_j \in \{0, 1\}, \quad \forall j \in J. \tag{13d}$$

These authors compare the proposed two non-compact formulations on the basis of large sets of unicost and non-unicost data sets and a modified version of SCP instances from the ORLIB (1990).

2.3 Specialized uncertain SCP models with desired coverage

In this section, we discuss two studies that specifically address the desired coverage of the SCP while considering uncertain input parameters a_{ij} with associated probabilistic uncertainties.

Wu and Kucukyavuz (2019a) discuss the chance-constrained combinatorial optimization problem and consider its application to solve the probabilistic approach of partial SCP. The notation of the model is provided below:

$V_1 = \{1, 2, \ldots, n\} :=$ index set for sets
$V_2 = \{1, 2, \ldots, m\} :=$ index set for items
$x_i = 1$ if component i is selected; 0 otherwise
$b_i :=$ objective coefficient of x_i
$\epsilon \in [0, 1] :=$ risk level
$\tau :=$ number of covered items in V_2
$\sigma(x) :=$ the random variable represents the number of covered items in V_2 for given x.
$\mathcal{B}(x) :=$ random event of interest for a given x
$\sigma(x) \geq \tau :=$ desired covering event $\mathcal{B}(x)$ for given x.

Uncertainty of a_{ij} in the classical SCP model (1a–1c) is the focus in their study for all $i \in V_2$ and $j \in V_1$. Their probabilistic SCP formulation is given by model (14a–14c):

$$\min z(x) = \sum_{i \in V_i} b_i x_i \tag{14a}$$

$$\text{s.t. } P\big(\sigma(x) \geq \tau\big) \geq 1 - \epsilon \tag{14b}$$

$$x \in \{0, 1\}^n. \tag{14c}$$

In this model, the objective function (14a) minimizes the total cost of the sets selected from V_1 while guaranteeing a certain degree of coverage of the items in V_2. Chance constraint (14b) defines the probability that the selected subsets covering a given number τ of items in V_2 is at least $1 - \epsilon$. Finally, constraint (14c) defines the domain of decision vector x. If $\tau = m$ number of items, then model (14a–14c) is equivalent to the probabilistic SCP model, which has a chance constraint. However, since $\tau \leq m$ is addressed in this study, it follows the structure of the Partial Probabilistic SCP (PPSCP). An experimental data set from a human sexual contact network is used to illustrate the proposed algorithm. V_1 and V_2 denote the groups of different genders in the data set, and models were implemented in C++ with a CPLEX optimizer.

Wu and Kucukyavuz (2019b) extend the PPSCP model (14a–14c) using an oracle to reformulate it into a MILP. A dynamic program technique is described to compute $P\big(\sigma(x) \geq \tau\big)$ in their study.

$\bar{A}_{i,j}$, a decision variable representing a dynamic programming recursion for $A(x, i, j)$, is defined for $1 \leq j \leq i, i \in V_2$ as:

$$A(x, i, j) = \begin{cases} A(x, i-1, j)(1 - P(x, i)), & j = 0 \\ A(x, i-1, j)(1 - P(x, i)) + A(x, i-1, j-1)P(x, i), & 0 < j < i \\ A(x, i-1, j-1)P(x, i), & j = i \end{cases}$$

A proposed compact MILP formulation is given in model (15a–15h):

$$\min z(x) = \sum_{i \in V_1} b_i x_i \tag{15a}$$

$$\text{s.t. } \bar{A}_{0,0} = 1 \tag{15b}$$

$$\bar{A}_{i,j} = \bar{A}_{i-1,j}(1 - P(x,i)), \quad \forall i \in V_1; \quad j = 0 \tag{15c}$$

$$\bar{A}_{i,j} = \bar{A}_{i-1,j}(1 - P(x,i)) + \bar{A}_{i-1,j-1}P(x,i), \quad \forall i \in V_1; \quad 0 < j < i \tag{15d}$$

$$\bar{A}_{i,j} = \bar{A}_{i-1,j-1}P(x,i), \quad \forall i \in V_1; \quad j = i \tag{15e}$$

$$\sum_{j=\tau}^{m} \bar{A}_{m,j} \geq 1 - \epsilon \tag{15f}$$

$$x \in \{0,1\}^n \tag{15g}$$

$$\bar{A}_{i,j} \in \mathbb{R}_+, \quad 0 \leq j \leq i \leq m. \tag{15h}$$

The objective function (15a) minimizes the total cost selection of subsets of items. Constraint (15b) is the boundary condition of the dynamic programming, and constraints (15c–15e) are the dynamic programming recursive functions. Constraint (15f) is the goal function of the model. These researchers demonstrated the effectiveness of their proposed method by implementing it in C++ with a CPLEX optimizer.

3 Uncertainty of objective cost coefficients

The third category of SCP models reviewed here focuses on the uncertainty of cost coefficient c_j for $j \in J$. In this section, we highlight a unique study that addresses this uncertainty.

Pereira and Averbakh (2013) study the uncertainty of cost coefficients in the objective function of the SCP, assuming that an interval estimate value is known for each cost coefficient. According to the authors, their study is the first which applies robust techniques to address this uncertainty. The notation of the model is provided below:

$M = \{1, \ldots, m\} := \text{set of rows}$
$N = \{1, \ldots, n\} := \text{set of columns}$
$\Gamma := \text{set of all coverings}$
$X \in \Gamma := \text{a subset of covering and also a } |N|\text{-dimensional characteristic vector of covering}$
$c_j^- := \text{lower bound of the cost coefficient } c_j$
$c_j^+ := \text{upper bound of the cost coefficient } c_j$
$[c_j^-, c_j^+] := \text{uncertain interval of the cost coefficient } c_j$
$S := \text{Cartesian product of the uncertain intervals } [c_j^-, c_j^+] \text{ for } j \in N$
$s := \text{scenario in the set } S$
$c_j^s := \text{cost corresponds to scenario } s$
$x_j = 1 \text{ if } j \in X; 0 \text{ otherwise}$
$s(X) := \text{scenario induced by } X$
$c_j^{s(X)} = c_j^+ \text{ if } j \in X; c_j^- \text{ otherwise}$
$\theta := \text{free variable}$

In their notation the classical SCP for any $X \in \Gamma$ and fixed $s \in S$ is given by

$$\min F(s, X) = \sum_{j \in X} c_j^s \tag{16a}$$

$$\text{s.t } X \in \Gamma. \tag{16b}$$

The objective function (16a) minimizes the cost of the selected cover in this model while the constraint set (16b) contains all possible covers. Let $R(s, X)$ be the regret for X under scenario s. Then, the authors define $R(s, X) = F(s, X) - F^*(X)$ where $F^*(X)$ is the optimum objective value of model (18a–18b). The min–max regret strategy is used to address the uncertainty of the cost coefficients. Thus, the min–max robust deviation version of model (16a–16b) is defined as the robust SCP, which is referred to as the ROB.SETCOVER; its formulation is given below:

$$\min Z(X) = \max_{s \in S} R(s, X) \tag{17a}$$

$$\text{s.t } X \in \Gamma. \tag{17b}$$

The objective function (17a) minimizes the worst-case regret for the covering X while the constraint set (17b) contains all possible covers. These researchers reformulate model (17a–17b) as follows:

$$\min_{X \in \Gamma} \left(\sum_{j \in N} c_j^+ x_j - \min_{Y \in \Gamma} \left(\sum_{j \in N} c_j^{s(X)} y_j \right) \right) \tag{18a}$$

$$\text{s.t } X \in \Gamma. \tag{18b}$$

Here $y_j = 1$ if $j \in Y$, and 0 otherwise. By introducing a free variable θ, they provide an equivalent formulation to model (18a–18b) as follows:

$$\min z(x) = \sum_{j \in N} c_j^+ x_j - \theta \tag{19a}$$

$$\text{s.t } \theta \leq \sum_{j \in N} c_j^- y_j + \sum_{j \in N} y_j(c_j^+ - c_j^-)x_j, \quad \text{forall,} \quad Y \in \Gamma \tag{19b}$$

$$\sum_{j \in N} a_{ij} x_j \geq 1, \quad \forall i \in M \tag{19c}$$

$$x_j \in \{0, 1\}, \quad \forall j \in N. \tag{19d}$$

The objective function (19a) along with constraint (19b) minimizes the worst-case regret for the covering X. In other words, the objective function (19a) finds the covering $X \in \Gamma$ with the smallest maximum regret. Constraints (19c) and (19d) are equivalent to $X \in \Gamma$. Since model (19a–19d) generates an exponential number of constraints for constraint (19b), it is challenging to solve with optimization solvers. Therefore, Benders decomposition and the Branch-Cut approach are applied to simplify the model. The researchers present numerical results considering three algorithms, two using Benders decomposition, with the third using the Branch-Cut approach in conjunction with several heuristic techniques.

4 Uncertainty of satisfying demand values

The fourth category of SCP models reviewed here focuses on the uncertainty of demand values b_i for $i \in I$ and their applications. We also include the studies finding generalized b_i values to ensure that the probability of meeting the coverage constraints is above a specified threshold level. In past studies, emergency vehicle location problems were studied from a deterministic perspective, where items are considered as customers and sets as vehicles. However, the availability of vehicles to satisfy a specific demand is not guaranteed all the time since they could be engaged in previous calls. This issue motivated the studies focusing on probabilistic location SCP. Most models in this category utilize the concept of a "busy fraction" of vehicles developed using queuing theory. This busy fraction is defined as the ratio of time the vehicles are busy responding to calls. Since there is no guarantee that classical covering constraint (1b) can be satisfied with the uncertainty of b_i, these models are also considered under the chance-constrained category on the reliability of service availability.

4.1 Uncertainty in demand values: a queuing theory approach with a busy fraction

In this section, we explore two studies that employ a specific formula to estimate the uncertainty of b_i, and utilize the queuing theory approach to solve these models. ReVelle and Hogan (1989) study the probabilistic location SCP. The notation of their formulation is provided below:

$J :=$ set of eligible facility sites (indexed by j)
$I :=$ set of demand nodes (indexed by i)
$t_\mu :=$ shortest time from potential facility site j to demand node i
$S :=$ standard for coverage (for either time or distance)
$f_k :=$ frequency of calls for service at demand node k, in calls per day
$\bar{t} :=$ the average duration of a call (hours)
$M_i :=$ set of demand nodes within S of node i
$N_i = \{j \mid t_\mu \leq S\}$ (the set of nodes j located within the time or distance standard of demand node i)
$F_i = \dfrac{\bar{t} \sum_{k \in M_i} f_k}{24}$ (the denominator provides the daily hours of service availability and the numerator provides the total number of calls for service at all the demand nodes ($k \in M_i$) per day)
$b_i :=$ the smallest integer satisfying $1 - (\frac{F_i}{b_i})^{b_i} \geq \alpha$
$x_j = 1$ if a facility is located at node j; 0 otherwise.

The formulation for the Probabilistic Location SCP (PLSCP) is given in (20a–20c):

$$\min z(x) = \sum_{j \in J} x_j \tag{20a}$$

$$\text{s.t.} \sum_{j \in N_i} x_j \geq b_i, \ \forall i \in I \tag{20b}$$

$$x_j \in \{0, 1\}, \ \forall j \in J. \tag{20c}$$

where b_i is the smallest integer satisfying $1 - (\frac{F_i}{b_i})^{b_i} \geq \alpha$. These researchers introduced an estimation of system-wide busy fraction in Equation (21):

$$q_i = \frac{\bar{t} \sum_{k \in M_i} f_k}{24 \sum_{j \in N_i} x_j}. \tag{21}$$

The objective function (20a) minimizes the number of facility locations. Constraint (20b) enforces the assignment of the minimum number of facility locations to guarantee reliability level α of node i. Here, the items are considered as demand zones, and sets are considered as potential facility sites.

Marianov and Revelle (1994) further study the PLSCP. In the previous studies, it was assumed that the availability of servers is independent of one another. However, this assumption may be violated practically; hence, the study focuses on an adjustment of the assumption of independence of server availability, presenting a new formulation. It is shown that the probability of at least one server being available equals 1-probability of all the servers within S being busy. As this model is an extension of the formulation presented in ReVelle and Hogan (1989), we will only introduce the additional notation necessary to define the new formulation. The required notation is provided below:

$\lambda_i :=$ arrival rate
$\frac{1}{\mu_i} :=$ single server's mean service time
$s = \sum_{j \in N_i} x_j :=$ number of servers
$\rho_i = \frac{i \sum_{k \in M_i} f_k}{24} = \frac{\lambda_i}{\mu_i}$
$p_i = \frac{\lambda_i}{\mu_i * s} :=$ probability of a server being busy in the region

Thus, using Eq. (21) and ReVelle and Hogan (1989)'s assumption of the binomial distribution, the probability of at least one server being busy can be calculated as:

$$1 - P \text{ [all servers of node } i \text{ are busy]} = 1 - \left(\frac{\rho_i}{\sum_{j \in N_i} x_j} \right)^{\sum_{j \in N_i} x_j}. \tag{22}$$

ReVelle and Hogan (1989) determine the deterministic equivalent form of Eq. (22) as $\sum_{j \in N_i} x_j \geq b_i$ where b_i is the smallest integer satisfying $1 - \left(\frac{\rho_i}{b_i} \right)^{b_i} \geq \alpha$. The principal difference between Marianov and Revelle (1994)'s model and the PLSCP model (20a–20c) introduced by ReVelle and Hogan (1989) is how the parameter b_i is calculated. The implementation of the model is exemplified using a real-world scenario of the emergency vehicle location problem considering items as demand points and sets as eligible sites for facilities. Further, this concept is applied to the maximum location SCP (ReVelle & Marianov, 1991; Borrás & Pastor, 2002) by introducing different versions of busy fractions and mathematical formulations.

4.2 Uncertainty in demand values: a probabilistic approach

In this section, we explore three studies that replace the demand value vector $(b_1, \ldots, b_m)$ with a binary random vector $\epsilon \in \{0, 1\}^m$. Probabilistic approaches have been utilized to solve these models.

Beraldi and Ruszczyński (2002b) initially introduce the uncertainty of demand values b_i. They discussed probabilistic SCP, where the demand values of each covering constraint in the classical SCP (1a–1c) are replaced by a binary random vector. The set of rows is indexed by i and the set of columns by j in their formulation. The notation of the model is provided below:

$\epsilon :=$ a binary vector in $\{0, 1\}^m$
$T :=$ 0–1 matrix with m rows, n columns
$p \in (0, 1) :=$ pre-specified reliability level

Their formulation is given by the model (23a–23c), which is a nonlinear binary integer programming problem:

$$\min z(x) = c^T x \tag{23a}$$

$$\text{s.t. } P\Big(Tx \geq \epsilon\Big) \geq p \tag{23b}$$

$$x \in \{0, 1\}^n. \tag{23c}$$

The objective function (23a) minimizes the cost of selected sets. Constraint (23b) ensures the probabilistic version of the covering constraint is at least satisfied by the probability p. They solve this model using two different approaches, the complete and the hybrid approaches. Two steps are applied in the complete approach, the first finding the p-efficient points using either forward or backward enumeration methods and the second using various tailored solution methods to determine the optimal solution for the probabilistic SCP. The meaning of the p-efficient point condition can be found in Definition 2.

Definition 2 A point $v \in \{0, 1\}^m$ is called a p-efficient point of the probability distribution function F, if $F(v) \geq p$ and there is no binary point $y \leq v$, $y \neq v$ such that $F(y) \geq p$.

The hybrid solution approach generates only the required p-efficient points, avoiding their complete enumeration. After finding the p-efficient points, each problem can be represented as a classical SCP. Then its optimal solution is obtained using three methods, the Forward Branch-and-Bound method, the Backward Branch-and-Bound method, and greedy heuristics. The performance of these methods and their efficiency is discussed using test problems under the three categories of small, medium, and large.

Beraldi and Ruszczyński (2005) conduct an extensive study on model (23a–23c). A Beam Search heuristic strategy was proposed to solve the model. The Beam Search strategy is a modified version of the classical Branch-and-Bound algorithm that employs a Breadth-First Search approach. It narrows down the search space by considering only a limited number of promising nodes, resulting in reduced memory requirements. Beam Search has found applications in various domains, including speech recognition (Lowerre, 1976), scheduling (Sabuncuoglu & Bayiz, 1999), and engineering design (Deb & Kumar, 2007). A computational experiment was carried out to compare the performance of methods proposed in Beraldi and Ruszczyński (2002b). This comparison involved the classical Branch and Bound method and the introduced Beam Search strategy. The results indicate that the Beam Search technique outperforms the classical Branch and Bound method in solving stochastic integer problems under probabilistic constraints.

Saxena et al. (2010) introduce a new formulation to model (23a–23c), the notation of which is provided below:

$N :=$ row index set
$M :=$ column index set
$L :=$ number of blocks
$M_1, \ldots, M_L :=$ partitions of M
$\{\xi^1, \ldots, \xi^L\} :=$ a set such that ξ^t is a 0–1 random M_t vector $\forall t \in \{1, 2, \ldots, L\}$
$z_i :=$ a binary value in $\{0, 1\}$
$F_t(z^t) = P(\xi^t \leq z^t)$
$S_t :=$ set of binary vectors which are either p-efficient or dominate a p-efficient point of F_t
$I_t :=$ set of p-efficient point of F_t
$\eta_t = \ln F_t(z^t), \forall t \in \{1, 2, \ldots, L\}$

Their formulation is given by model (24a–24g), which is a mixed-integer programming problem:

$$\min z(x) = cx \tag{24a}$$

$$\text{s.t. } Ax \geq z \tag{24b}$$

$$\sum_{t \in L} \eta_t \geq \ln p \tag{24c}$$

$$\eta_t \leq (\ln F_t(v)) \left(1 - \sum_{i \in M_t, v_i = 0} z_i\right), \quad \forall v \in S_t; \quad \forall t \in \{1, 2, \ldots, L\} \tag{24d}$$

$$1 \leq \sum_{i \in M_t, v_i = 0} z_i, \quad \forall v \in I_t; \quad \forall t \in \{1, 2, \ldots, L\} \tag{24e}$$

$$z_i \in \{0, 1\}, \quad \forall i \in M \tag{24f}$$

$$x_j \in \{0, 1\}, \quad \forall j \in N. \tag{24g}$$

Objective function (24a) minimizes the cost of selected sets, and constraints (24b–24f) represent the probabilistic covering constraint (23b) in the model (23a–23c). An example is used to illustrate MIP reformulation, and a large test bed is used to discuss computational efficiency. A CPLEX solver is used to solve the instances, with the computational results showing that the proposed procedure efficiently solves the probabilistic SCP.

4.3 Uncertainty in demand values from multiple facilities

Another area of interest related to uncertainty in demand value is the study of multiple coverages for each item with probabilistic conditions. In this section, we highlight a unique analysis that addresses a specific variant of the covering problem: the robust set covering problem with probabilistic and cooperative covering by two types of facilities.

Ding et al. (2020) investigate the robust uncertain two-level cooperative SCP when the coverage of each item is required by two types of facilities denoted by y and z. This study combines the concepts of robust, probabilistic, and cooperative covering by introducing Γ-*robust two-level-cooperative α-cover* constraints. Two mathematical models have been proposed. The notation of the model is provided below:

$I = \{1, 2, \ldots, m\} :=$ set of demand nodes indexed by k
$J = \{1, 2, \ldots, n_1\} :=$ set of y-facility location sites indexed by j
$K = \{1, 2, \ldots, n_2\} :=$ set of z-facility location sites indexed by k
$c_j^1 :=$ costs of building y-facility located at site j
$c_j^2 :=$ costs of building z-facility located at site k
$a_{ij} \in \{0, 1\} := y$-facility location $j \in J$ is able to cover the demand node i
$b_{ik} \in \{0, 1\} := z$-facility location $k \in K$ is able to cover the demand node i

Since the model has two types of facilities, the deterministic model is a little different from the classical SCP model (1a–1c). Further, this study considered the uncertainty of the a_{ij} and b_{ik} values, which represent whether a location $j \in J$ or $k \in K$ is able to cover the demand node i. Hence, a_{ij} and b_{ik} are considered independent random binary variables.

It considered that $a_{ij} = 1$ with a probability of $(1 - p_{ij})$ and $a_{ij} = 0$ with probability p_{ij}. Likewise, $b_{ik} = 1$ with a probability of $(1 - q_{ik})$ and $b_{ik} = 0$ with probability q_{ik}. Their

formulation is given by model (25a–25d), which is a binary non-linear integer programming problem for two facility types. This model is referred to as the Two-Level Cooperative SCP (TLCSCP):

$$\min \sum_{j \in J} c_j^1 y_j + \sum_{k \in K} c_k^2 z_k \tag{25a}$$

$$\text{s.t.} \left(\sum_{j \in J} a_{ij} y_j \right) \left(\sum_{k \in K} b_{ik} z_k \right) \geq 1, \quad \forall i \in I \tag{25b}$$

$$y_j \in \{0, 1\}, \quad \forall j \in J \tag{25c}$$

$$z_k \in \{0, 1\}, \quad \forall k \in K. \tag{25d}$$

The objective function (25a) minimizes the building cost of two types of facilities. Constraint (25b) ensures that each demand node covers at least one y-facility and z-facility simultaneously. Then, the generalized probabilistic model is developed based on TLCSCP model (25a–25d), which is in non-linear form.

The notation of the model is provided below:

$$\delta = \frac{2\gamma + \alpha\beta - \alpha\gamma - \sqrt{\alpha(4\beta\gamma + \alpha\beta^2 + \alpha\gamma^2 - 2\alpha\beta\gamma)}}{2\gamma},$$

$\beta, \gamma \in [0, 1]$ are constants with $\beta + \gamma = 1$
$\alpha \in [0, 1) :=$ coverage level of the two facility types as covering at least one y-facility and z-facility simultaneously
$\hat{p}_{ij} \geq 0$ worst case deviation values
$\hat{q}_{ik} \geq 0$ worst case deviation values
$p_{ij} = [\bar{p}_{ij}, \bar{p}_{ij} + \hat{p}_{ij}] \subseteq [0, 1], \bar{p}_{ij} \geq 0$ a nominal value
$q_{ik} = [\bar{q}_{ik}, \bar{q}_{ik} + \hat{q}_{ik}] \subseteq [0, 1], \bar{q}_{ik} \geq 0$ a nominal value

The authors converted the constraint (25b) to two linear constraints $\sum_{j \in J} a_{ij} y_j \geq 1, \forall i \in I$ and $\sum_{k \in K} b_{ik} z_k \geq 1, \forall i \in I$, and then two-level cooperative constraint $P\left(\sum_{j \in J} a_{ij} y_j \geq 1, \sum_{k \in K} b_{ik} z_k \geq 1 \right) \geq \alpha, \forall i \in I$ is introduced.

By assuming that the variables y_j and z_k are independent variables, they obtained

$$\left(1 - \prod_{j \in J} p_{ij}^{y_j} \right) \cdot \left(1 - \prod_{k \in K} q_{ik}^{z_k} \right) \geq \alpha, \quad \forall i \in I \tag{26}$$

where constraint (26) is in non-linear form and its relaxation leads to the following linear approximation formulation. Their first formulation is given by model (27a–27f), which is an

integer linear programming program:

$$\min \sum_{j \in J} c_j^1 y_j + \sum_{k \in K} c_k^2 z_k \tag{27a}$$

$$\text{s.t.} \sum_{j \in J} \ln(p_{ij}) y_j \le \ln(1 - \alpha), \quad \forall i \in I \tag{27b}$$

$$\sum_{k \in K} \ln(q_{ij}) z_k \le \ln(1 - \alpha), \quad \forall i \in I \tag{27c}$$

$$\beta \sum_{j \in J} \ln(p_{ij}) y_j + \gamma \sum_{k \in K} \ln(q_{ik}) z_k \le \ln\left[\left(1 - \frac{\alpha}{1 - \delta}\right)^\beta \eta_i^\gamma\right], \quad \forall i \in I \tag{27d}$$

$$y_j \in \{0, 1\}, \quad \forall j \in J \tag{27e}$$

$$z_k \in \{0, 1\}, \quad \forall k \in K \tag{27f}$$

The second model is obtained based on the concept of Γ-robust two-level-cooperative α-cover provided in Definition 1. The set of constraints (27a–27f) are reformulated into constraint (28b–28k) using Definition 1. The second formulation is given by model (28a–28m), which is a compact mixed-integer linear programming problem:

$$\min \sum_{j \in J} c_j^1 y_j + \sum_{k \in K} c_k^2 z_k \tag{28a}$$

$$\text{s.t.} \sum_{j \in J} \ln(\bar{p}_{ij}) y_j + \sum_{j \in J} \zeta_{ij}^1 + \Gamma_i \eta_i^1 \le \ln(1 - \alpha); \quad \forall i \in I \tag{28b}$$

$$\sum_{k \in K} \ln(\bar{q}_{ik}) z_k + \sum_{k \in K} \zeta_{ik}^2 + \Gamma_i \eta_i^2 \le \ln(1 - \alpha); \quad \forall i \in I \tag{28c}$$

$$\beta \left[\sum_{j \in J} \ln(\bar{p}_{ij}) y_{ij} + \sum_{j \in J} \zeta_{ij}^1 + \Gamma_i \eta_i^1\right] + \tag{28d}$$

$$\gamma \left[\sum_{k \in K} \ln(\bar{q}_{ik}) z_k + \sum_{k \in K} \zeta_{ik}^2 + \Gamma_i \eta_i^2\right] \le \ln\left[1 - \frac{\alpha}{1 - \delta}^\beta \eta_i^\gamma\right]; \quad \forall i \in I \tag{28e}$$

$$\zeta_{ij}^1 + \eta_i^1 \ge \ln(\bar{p}_{ij} + \hat{p}_{ij}) - \ln(\hat{p}_{ij}) y_j; \quad \forall i \in I, \quad j \in J \tag{28f}$$

$$\zeta_{ik}^2 + \eta_i^2 \ge \ln(\bar{q}_{ik} + \hat{q}_{ik}) - \ln(\hat{q}_{ik}) z_k; \quad \forall i \in I, \quad k \in K \tag{28g}$$

$$\zeta_{ij}^1 \ge 0; \quad \forall i \in I, \quad j \in J \tag{28h}$$

$$\zeta_{ik}^2 \ge 0; \quad \forall i \in I, \quad k \in K \tag{28i}$$

$$\eta_i^1 \ge 0; \quad \forall i \in I \tag{28j}$$

$$\eta_i^2 \ge 0; \quad \forall i \in I \tag{28k}$$

$$y_j \in \{0, 1\}; \quad \forall j \in J \tag{28l}$$

$$z_k \in \{0, 1\}; \quad \forall k \in K. \tag{28m}$$

Computational experiments and analysis are conducted by implementing problems in MATLAB R2016a, and a CPLEX solver is used to solve the problems. Computational results demonstrate that the compact mixed-integer linear programming model can efficiently solve the uncertain SCP when each item requires coverage from two types of facilities.

5 Conclusion and possible future research avenues

5.1 Summary of the survey

The world and its needs change constantly and dramatically; thus, uncertainty appears almost everywhere. This situation is also true in the mathematical sciences. Here, we focused on the uncertain counterparts of a pedagogically and practically significant problem known as the SCP. The SCP has been extensively studied since 1970. Early studies mainly aimed to develop efficient methods for solving the deterministic SCP. However, since 2020, researchers have shifted their focus towards addressing the inherent uncertainty associated with the SCP by developing new methodologies. Due to the considerable development of generalized covering models, Schilling (1993) provided a classification of two types of these models as mentioned in the introduction. In our study, we specifically targeted the uncertain variant of the SCP defined in the model (1a–1c). This variant involves the requirement of covering all items of the SCP using at least one set while simultaneously minimizing the value of the objective function.

We identified 16 models, with chance-constraint concepts being the most commonly used in them along with probabilistic and robust optimization techniques. The uncertain input parameters a_{ij}, b_i, c_j for $i \in I$, $j \in J$ have been studied in these models. Among these uncertainties, the ability to include a particular item e_i in one specific set S_j is most frequently discussed. In this review, we described the complete mathematical models, including all variables and model parameters, and we highlighted the application if the model was developed explicitly focusing on a particular practical problem.

Further, we outlined the solution technique utilized to solve each model. Since the SCP is an NP-hard optimization problem, the optimal global solution employing the Branch-and-Bound or Branch-and-Cut algorithm will work only for small-scale test problems. Some studies grouped test problems as small, medium, or large based on the number of items and the number of sets utilized to construct the test problem since there is no generally accepted measurement for grouping test problems. The models become more complicated with the integration of the uncertain concept. Thus, heuristic and approximation algorithms were developed to efficiently solve the uncertain models computationally. Most of the application models presented in this study explored issues related to applications for emergency services. However, we also found several models focusing on reserve design and logistic system design, among others. We also included tables summarizing our findings. Specifically, Table 1 displays the solution technique utilized by each model, while Table 2 displays the number of studies conducted during each decade between 1970 and 2020.

Progress has been made, and the SCP with uncertain problem parameters has been well-explored over the past five decades in decision science. Still, this study discovered some crucial future research questions and further computationally efficient theoretical developments to the SCP with uncertain problem parameters when we make decisions on complex and competitive real-world issues.

5.2 Future research directions

This section presents promising and specific future research directions focusing on three research avenues: (1) improved robust optimization methods when an estimated interval containing the nominal cost value of each set is known, but the nominal value is unknown; (2) theoretical investigations when a random integer vector replaces the demand values of

each covering constraint; (3) innovative designs when multiple objectives appear in the SCP model with uncertain parameters.

1. Although the primary goal of the SCP is to identify a least-cost collection of sets to cover all items, the nominal cost values of sets are unknown even though an estimated interval containing the nominal cost value for the set is known. We found only one study conducted by Pereira and Averbakh (2013), which utilized Robust Optimization (RO) techniques to address this issue. Their proposed solution approach generates an exponential number of constraints, and three solution techniques are developed using Benders decomposition and the Branch-Cut approaches. RO became widely used during the last decade; thus, we believe there is potential for new research focusing on RO and straightforward generalized covering models. For example, recently, (Coco et al., 2022) use model (19a–19d) to study the solutions of the min-max regret weighted SCP (min–max regret WSCP) and the min–max regret maximum benefit SCP (min–max regret MSCP). The deterministic model MSCP has additional conditions compared to the classical SCP, specifically $\sum_{j \in M} w_j x_j \leq T$ where M is set of columns, w_j is the weight of the columns j and T is the maximum capacity. To improve computing capabilities, research on how to solve large problems in a reasonable amount of time using this robust SCP formulation is beneficial. Several solution methods based on RO have been developed for similar types of NP-hard optimization problems such as the knapsack problem and the traveling salesman problem.

2. This survey found studies focused on the uncertainty of demand values b_i in four areas: (1) cover each item at least once with a reliability level; (2) find the minimum number of demand points to achieve a predefined reliability coverage level; (3) replace the demand values of each covering constraint with a binary random vector; and (4) cover each item with two types of facilities. In addition to these investigations, several real-world applications of the SCP—including vehicle routing, crew scheduling, and logical analysis of data (Marsten & Shepardson, 1981; ReVelle & Hogan, 1989; Marchiori & Steenbeek, 2000b; Daskin & Stern, 1981; McDonnell et al., 2002; Weerasena et al., 2014; Kohl & Karisch, 2004; Hammer & Bonates, 2006; Bettinelli et al., 2014; Marchiori & Steenbeek, 2000; Saxena & Arora, 1981; Bandara et al., 2012)—require more than one set to cover each item (multiple covers), meaning a random integer vector replaces the demand values of each covering constraint in the model. These problems are formulated as SCPs with generalized coverage constraints. Studies on the generalized SCP with multiple objective functions have been proposed by Weerasena (2020) and Weerasena et al. (2022). The feasible set for this generalized case can be written as the set $\left\{x \in \{0,1\}^n : \sum_{j \in J} a_{ij} x_j \geq b_i \text{ for } i \in I, \text{ where } b_i \text{ is a random integer}\right\}$. With this extension, a straightforward uncertain feasible region can be represented by the set

$$\left\{x \in \{0,1\}^n : P\left(\sum_{j \in J} a_{ij} x_j \geq b_i\right) \geq \alpha_i \text{ for } i \in I, \text{ where } b_i \text{ is a random integer}\right\}$$

where α_i is the minimum reliability level of item e_i and P is a probability function. Even though the binary vector is transferred to an integer vector, solving the new model is more challenging due to multiple covers. Thus, an important direction for future research is to investigate modeling approaches to solve the SCP when a random integer vector replaces the demand values of each covering constraint in the model.

3. This survey discovered decision-making in SCPs with uncertain counterparts only with single objective optimization models. Some application areas of SCPs include location/allocation science (emergency medical vehicle allocations, facility locations) and conservation biology (designing reserve systems for managing wildlife habitats and populations). Naturally, such applications require decision-making with conflicting objective

functions (centering different service areas or species) that need optimization with an uncertain counterpart. When solving models involving multiple conflicting objective functions, the optimization stage provides a set of all Pareto points (or efficient points). The typical decision-making process for particular application problems is subjective and primarily driven by the expert opinion of the specific application area. Cases like this would benefit significantly from a mathematically driven approach for multi-objective optimization with uncertain counterparts. The uncertain features and multiple objective functions are more challenging to address; therefore, it requires innovative optimization methods to solve large-scale SCP models. While the classical SCP with multiple objective functions has received attention since 2000, (Jaszkiewicz, 2003; Prins et al., 2006; Florios & Mavrotas, 2014; Weerasena et al., 2017; Weerasena & Wiecek, 2019; Weerasena, 2020; Weerasena et al., 2022), thus far research is limited on finding Pareto solutions for the uncertain SCP with multiple objective functions. Therefore, another important future research direction is to investigate solution approaches to SCP consisting of multiple conflicting objective functions and uncertain counterparts.

5.3 Concluding remarks

The SCP is important in the combinatorial optimization literature due to its diverse applicability to real-world issues focusing on location, science, and other related areas. This study reviewed the uncertain variants of the SCPs with applications. Progress has been made over the past five decades in decision science, focusing on the uncertain SCP. Still, this study discovered several crucial future research questions to investigate with the decision-making process of the set covering models under uncertainty. We discussed three promising future research directions to support decisions on complex and competitive real-world issues based on current accomplishments.

Declarations

Conflict of interest None.

References

Ahmed, S., & Papageorgiou, D. J. (2013). Probabilistic set covering with correlations. *Operations Research, 61*(2), 438–452.

Aly, A. A., & White, J. A. (1978). Probabilistic formulation of the emergency service location problem. *Journal of the Operational Research Society, 29*(12), 1167–1179.

Bandara, D., Mayorga, M., & McLay, M. (2012). Optimal dispatching strategies for emergency vehicles to increase patient survivability. *International Journal of Operational Research, 15*(2), 195–214.

Bélanger, V., Ruiz, A., & Soriano, P. (2019). Recent optimization models and trends in location, relocation, and dispatching of emergency medical vehicles. *European Journal of Operational Research, 272*(1), 1–23.

Benveniste, R. (1982). A note on the set covering problem. *Journal of Operational Research Society, 33*, 261–265.

Beraldi, P., & Ruszczyński, A. (2002). A branch and bound method for stochastic integer problems under probabilistic constraints. *Optimization Methods and Software, 17*(3), 359–382.

Beraldi, P., & Ruszczyński, A. (2002). The probabilistic set-covering problem. *Operations Research, 50*(6), 956–967.

Beraldi, P., & Ruszczyński, A. (2005). Beam search heuristic to solve stochastic integer problems under probabilistic constraints. *European Journal of Operational Research, 167*(1), 35–47.

Bettinelli, A., Ceselli, A., & Righini, G. (2014). A branch-and-price algorithm for the multi-depot heterogeneous-fleet pickup and delivery problem with soft time windows. *Mathematical Programming Computation, 6*(2), 171–197.

Borrás, F., & Pastor, J. T. (2002). The ex-post evaluation of the minimum local reliability level: An enhanced probabilistic location set covering model. *Annals of Operations Research, 111*(1), 51–74.

Cabeza, M., Araújo, M. B., Wilson, R. J., Thomas, C. D., Cowley, M. J., & Moilanen, A. (2004). Combining probabilities of occurrence with spatial reserve design. *Journal of Applied Ecology, 41*(2), 252–262.

Chvatal, V. (1979). A greedy heuristic for the set-covering problem. *Mathematics of Operations Research, 4*(3), 233–235.

Coco, A. A., Santos, A. C., & Noronha, T. F. (2022). Robust min–max regret covering problems. *Computational Optimization and Applications, 83*(1), 111–141.

Daskin, M. S., & Stern, E. H. (1981). A hierarchical objective set covering model for emergency medical service vehicle deployment. *Transportation Science, 15*(2), 137–152.

Deb, K., & Kumar, A. (2007). Light beam search based multi-objective optimization using evolutionary algorithms. In *2007 IEEE congress on evolutionary computation* (pp. 2125–2132). IEEE.

Degel, D., & Lutter, P. (2013). *A robust formulation of the uncertain set covering problem*. Technical report, Ruhr-University Bochum.

Ding, S., Zhang, Q., & Yuan, Z. (2020). An under-approximation for the robust uncertain two-level cooperative set covering problem. In *2020 59th IEEE conference on decision and control (CDC)* (pp. 1152–1157). IEEE.

Farahani, R. Z., Asgari, N., Heidari, N., Hosseininia, M., & Goh, M. (2012). Covering problems in facility location: A review. *Computers & Industrial Engineering, 62*(1), 368–407.

Fischetti, M., & Monaci, M. (2012). Cutting plane versus compact formulations for uncertain (integer) linear programs. *Mathematical Programming Computation, 4*(3), 239–273.

Florios, K., & Mavrotas, G. (2014). Generation of the exact pareto set in multi-objective traveling salesman and set covering problems. *Applied Mathematics and Computation, 237*, 1–19.

Grossman, T., & Wool, A. (1997). Computational experience with approximation algorithms for the set covering problem. *European Journal of Operational Research, 101*(1), 81–92.

Haight, R. G., Revelle, C. S., & Snyder, S. A. (2000). An integer optimization approach to a probabilistic reserve site selection problem. *Operations Research, 48*(5), 697–708.

Hammer, P. L., & Bonates, T. O. (2006). Logical analysis of data: An overview: From combinatorial optimization to medical applications. *Annals of Operations Research, 148*(1), 203–225.

Hwang, H.-S. (2002). Design of supply-chain logistics system considering service level. *Computers & Industrial Engineering, 43*(1–2), 283–297.

Jaszkiewicz, A. (2003). Do multiple-objective metaheuristics deliver on their promises? A computational experiment on the set-covering problem. *IEEE Transactions on Evolutionary Computation, 7*(2), 133–143.

Kohl, N., & Karisch, S. E. (2004). Airline crew rostering: Problem types, modeling, and optimization. *Annals of Operations Research, 127*(1–4), 223–257.

Lemke, C. E., Salkin, H., & Spielberg, K. (1971). Set covering by single-branch enumeration with linear-programming subproblems. *Operations Research, 19*(4), 998–1022.

Lowerre, B. T. (1976). *The harpy speech recognition system*. Pittsburgh: Carnegie Mellon University.

Lutter, P., Degel, D., Büsing, C., Koster, A. M., & Werners, B. (2017). Improved handling of uncertainty and robustness in set covering problems. *European Journal of Operational Research, 263*(1), 35–49.

Marchiori, E., & Steenbeek, A. (2000). An evolutionary algorithm for large scale set covering problems with application to airline crew scheduling. In *41st Annual symposium on real-world applications of evolutionary computation, workshops* (pp. 370–384). Berlin: Springer.

Marchiori, E., & Steenbeek, A. (2000b). An evolutionary algorithm for large scale set covering problems with application to airline crew scheduling. In *Workshops on real-world applications of evolutionary computation* (pp 370–384). Springer.

Marianov, V., & Revelle, C. (1994). The queuing probabilistic location set covering problem and some extensions. *Socio-Economic Planning Sciences, 28*(3), 167–178.

Marsten, R. E., & Shepardson, F. (1981). Exact solution of crew scheduling problems using the set partitioning model: Recent successful applications. *Networks, 11*(2), 165–177.

McDonnell, M. D., Possingham, H. P., Ball, I. R., & Cousins, E. A. (2002). Mathematical methods for spatially cohesive reserve design. *Environmental Modeling and Assessment, 7*(2), 107–114.

NETLIB. (2013). http://www.netlib.org/lp/data/.

ORLIB. (1990). http://people.brunel.ac.uk/~mastjjb/jeb/info.html.

Pereira, J., & Averbakh, I. (2013). The robust set covering problem with interval data. *Annals of Operations Research, 207*(1), 217–235.

Powell, W. B. (2019). A unified framework for stochastic optimization. *European Journal of Operational Research, 275*(3), 795–821.

Prins, C., Prodhon, C., & Calvo, R. W. (2006). Two-phase method and Lagrangian relaxation to solve the bi-objective set covering problem. *Annals of Operations Research, 147*(1), 23–41.

ReVelle, C., & Hogan, K. (1989). The maximum availability location problem. *Transportation science, 23*(3), 192–200.

ReVelle, C., & Marianov, V. (1991). A probabilistic fleet model with individual vehicle reliability requirements. *European Journal of Operational Research, 53*(1), 93–105.

Sabuncuoglu, I., & Bayiz, M. (1999). Job shop scheduling with beam search. *European Journal of Operational Research, 118*(2), 390–412.

Saxena, A., Goyal, V., & Lejeune, M. A. (2010). MIP reformulations of the probabilistic set covering problem. *Mathematical Programming, 121*(1), 1–31.

Saxena, R. R., & Arora, S. (1981). Exact solution of crew scheduling problems using the set partitioning model: Recent successful applications. *Optimization, 11*(2), 165–177.

Schilling, D. A. (1993). A review of covering problems in facility location. *Location Science, 1*, 25–55.

Slavík, P. (1996). A tight analysis of the greedy algorithm for set cover. In *Proceedings of the twenty-eighth annual ACM symposium on theory of computing* (pp. 435–441).

Soyster, A. L. (1973). Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research, 21*(5), 1154–1157.

Toregas, C., Swain, R., ReVelle, C., & Bergman, L. (1971). The location of emergency service facilities. *Operations Research, 19*(6), 1363–1373.

Wang, W., Wu, S., Wang, S., Zhen, L., & Qu, X. (2021). Emergency facility location problems in logistics: Status and perspectives. *Transportation Research Part E: Logistics and Transportation Review, 154*, 102465.

Weerasena, L. (2020). Algorithm for generalised multi-objective set covering problem with an application in ecological conservation. *International Journal of Mathematical Modelling and Numerical Optimisation, 10*(2), 167–186.

Weerasena, L., & Wiecek, M. M. (2019). A tolerance function for the multiobjective set covering problem. *Optimization Letters, 13*(1), 3–21.

Weerasena, L., Shier, D., & Tonkyn, D. (2014). A hierarchical approach to designing compact ecological reserve systems. *Environmental Modeling & Assessment, 19*(5), 437–449.

Weerasena, L., Wiecek, M. M., & Soylu, B. (2017). An algorithm for approximating the pareto set of the multiobjective set covering problem. *Annals of Operations Research, 248*(1), 493–514.

Weerasena, L., Ebiefung, A., & Skjellum, A. (2022). Design of a heuristic algorithm for the generalized multi-objective set covering problem. *Computational Optimization and Applications*, 1–35.

Wu, H.-H., & Kucukyavuz, S. (2019a). Chance-constrained combinatorial optimization with a probability oracle and its application to probabilistic partial set covering. *SIAM Journal on Optimization*, 1–37.

Wu, H.-H., & Kucukyavuz, S. (2019). Probabilistic partial set covering with an oracle for chance constraints. *SIAM Journal on Optimization, 29*(1), 690–718.