Satisficing Awakens: Models to Mitigate Uncertainty

Patrick Jailet
Department of Electrical Engineering and Computer Science, Operations Research Center, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, jaillet@mit.edu

Sanjay Dominik Jena
Singapore-MIT Alliance for Research and Technology, Singapore, 138602, sanjay@smart.mit.edu

Tsan Sheng Ng
Department of Industrial & Systems Engineering, National University of Singapore, Singapore 117576, isentsa@nus.edu.sg

Melvyn Sim
NUS Business School, National University of Singapore, Singapore 119245, dscsimm@nus.edu.sg

Satisficing, as an approach to decision-making under uncertainty, aims at achieving solutions that satisfy the problem’s constraints as well as possible. Mathematical optimization problems that are related to this form of decision-making include the P-model of Charnes and Cooper (1963), where satisficing is the objective, as well as chance-constrained and robust optimization problems, where satisficing is articulated in the constraints. In this paper, we first propose the R-model, where satisficing is the objective, and where the problem consists in finding the most “robust” solution, feasible in the problem’s constraints when uncertain outcomes arise over a maximally sized uncertainty set. We then study the key features of satisficing decision making that are associated with these problems and provide the complete functional characterization of a satisficing decision criterion. As a consequence, we are able to provide the most general framework of a satisficing model, termed the S-model, which seeks to maximize a satisficing decision criterion in its objective, and the corresponding satisficing-constrained optimization problem that generalizes robust optimization and chance-constrained optimization problems. Next, we focus on a tractable probabilistic S-model, termed the T-model whose objective is a lower bound of the P-model. We show that when probability densities of the uncertainties are log-concave, the T-model can admit a tractable concave objective function. In the case of discrete probability distributions, the T-model is a linear mixed integer program of moderate dimensions. We also show how the T-model can be extended to multi-stage decision-making and present the conditions under which the problem is computationally tractable. Our computational experiments on a stochastic maximum coverage problem strongly suggest that the T-model solutions can be highly effective, thus allaying misconceptions of having to pay a high price for the satisficing models in terms of solution conservativeness.

History: January 28, 2016
1. Introduction

Parametric uncertainty in mathematical programming problems is ubiquitous in many real-world problems. When uncertain parameters are revealed, the solution obtained may become infeasible and the actual objective value attained may be significantly worse than the “optimal” value achieved based on solving optimization models assuming deterministic parameter values (Ben-Tal et al. 2004). A common approach to deal with uncertainty in an optimization problem is to incorporate attitudes of risks, when probability distributions are available, or ambiguity, whenever they are not. Such models have been extensively studied in stochastic programming (see, e.g., Prékopa 1995, Birge and Louveaux 1997, Prékopa 2003), robust optimization (see, e.g., Ben-Tal and Nemirovski 1999, Bertsimas et al. 2011) and distributionally robust optimization (see, e.g., Delage and Ye 2010, Goh and Sim 2010, Wiesemann et al. 2014).

The concept of satisficing, a portmanteau of the terms ‘satisfy’ and ‘suffice’ first introduced by Simon (1959), addresses uncertainty with the aims of achieving feasibility in an uncertain environment. In many real world problems under uncertainty, the goal is not necessarily to maximize or minimize objective functions such as profits or costs. Instead, decision-makers may be more interested in obtaining solutions that can “satisfice” the constraints of the problem, in some sense, as well as possible. For instance, it is reasonable in a project management problem with uncertain activity completion times to ensure that the project can be completed on schedule and within the allocated budget (see, e.g., Goh and Hall 2013).

Satisficing as an objective in decision making has mostly been explored from the economic perspective (see, e.g. Güth 2010, Stüttgen et al. 2012). Charnes and Cooper (1963) were the first to incorporate the principal idea of satisficing in the mathematical framework of success probability maximization, which has been termed the P-model. In simplified form, the P-model can be stated as:

$$\max \ln \mathbb{P}(A(\tilde{z})x \geq b(\tilde{z}))$$

subject to $x \in X$, (1)

where $x$ are decision variables of dimension $N$ defined on a feasible set $X$ and $\tilde{z}$ is a $K$ dimensional random vector that influences the entries of the function maps $A: \mathbb{R}^K \mapsto \mathbb{R}^{M \times N}$ and $b: \mathbb{R}^K \mapsto \mathbb{R}^M$. We refer to the problem’s constraints as the set of randomly perturbed linear constraints, $A(z)x \geq b(z)$, where $z$ is a random outcome of $\tilde{z}$ (under the probability measure $\mathbb{P}$). A satisficing decision criterion evaluates how well a solution $x$ would remain feasible in the problem’s constraints under uncertainty. In this regard, the P-model is an optimization problem that maximizes a log-probability satisficing decision criterion, $\nu_P: \mathbb{R}^N \mapsto \mathbb{R} \cup \{-\infty\}$, given by

$$\nu_P(x) = \ln \mathbb{P}(A(\tilde{z})x \geq b(\tilde{z})).$$

(2)
Note that the decision criterion is based on a log-probability function, instead of directly a probability function. Using the convention $\ln 0 = -\infty$, if $x$ is always infeasible in the problem’s constraints, then $\nu_P(x) = -\infty$. Moreover, if the probability function is log-concave in $x$, which may arise in useful instances, then the objective function of Problem (1) would be concave.

The P-model is closely related to the more popular chance-constrained optimization problem pioneered by Charnes and Cooper (1959, 1963) where the satisficing decision criterion (2) is imposed in the constraints. The following gives a simplified definition of a chance-constrained optimization model:

$$\begin{align*}
\min & \quad c'x \\
\text{s.t.} & \quad \ln P(A(z)x \geq b(z)) \geq \Delta \\
& \quad x \in \mathcal{X}.
\end{align*}$$

(3)

In contrast to the P-model, the objective in (3) is a deterministic cost function, where $c \in \mathbb{R}^N$ defines the objective function coefficients, and the satisficing decision criterion of the P-model is now subject to a lower bound parameter, $\Delta \in \mathbb{R}$. In other words, the satisficing decision criterion constraint above enforces how well the problem’s constraints $A(z)x \geq b(z)$ must be satisfied under uncertainty, and among feasible solutions $x \in \mathcal{X}$ that fulfill this decision criterion constraint, select one that is cost-minimizing. We refer readers to Prékopa (2003) and Henrion (2004) for an excellent introduction to the models, algorithms and theory of chance-constrained programming. Chance-constrained programming has also found wide-ranging applications in important finance and engineering problems, for instance in renewable energy planning under uncertain loads and intermittent renewable supplies (see e.g., Van et al. 2014, Bremer et al. 2015).

The P-model and the chance-constrained optimization problem are known to be generally intractable, because evaluating the probability typically demands high dimensional integration and is a computationally excruciating procedure. It is also noteworthy that the log-probability satisficing decision criterion (2) is concave for the special case when $A(\cdot)$ is a constant (see, e.g. Prékopa 2003, Theorem 2.5) and the vector $b(\bar{z})$ is affinely dependent on $\bar{z}_1, \ldots, \bar{z}_K$, which are independently distributed random variables with log-concave density functions. However, notwithstanding the convexity of the resultant optimization problems in (1) and (3), the evaluation of the log-probability satisficing decision criterion (2) can itself still be computationally challenging even when $M = 1$ and the random variables are iid uniformly distributed (Nemirovski and Shapiro 2006). Hence, more restrictive conditions are required to obtain computationally scalable results. For more general distributions, one may use sample average approximation approaches, such as Monte Carlo methods (see, e.g. Shapiro, A. 2003) to approximate the objective function. Unfortunately, even small problems with relatively simple structure can require hundreds of samples (Shapiro and Homem-de-Mello 2000, Pagnoncelli et al. 2009) to achieve a desired level of accuracy. When the problem size is large or the variability of the uncertain parameters is high, the sample size required is likely
to become prohibitively large. Furthermore, a sample-based optimization model for solving these problems requires a large number of binary variables in its formulation, making it computationally expensive.

Robust optimization is a more tractable alternative to chance-constrained optimization and has become an important approach in addressing practical optimization problems with data uncertainty. A robust optimization problem can generally be formulated as follows,

$$\begin{align*}
\min & \quad c'x \\
\text{s.t.} & \quad A(z)x \geq b(z) \quad \forall z \in \mathcal{U}(\Gamma) \\
& \quad x \in \mathcal{X}.
\end{align*}$$

(4)

In the above, the vector \( z \) denotes a realization of \( \tilde{z} \) from an uncertainty set \( \mathcal{U}(\Gamma) \subseteq \mathcal{W} \), with \( \mathcal{W} \subseteq \mathbb{R}^K \) being the support of \( \tilde{z} \), and where \( \Gamma \in \mathbb{R}_+ \) is a user specified parameter that relates to the level of uncertainty that must be tolerated. In particular, \( \mathcal{U}(\cdot) \) is designed such that \( \mathcal{U}(\alpha_1) \subseteq \mathcal{U}(\alpha_2) \subseteq \mathcal{W} \) for all \( 0 \leq \alpha_1 \leq \alpha_2 \). The parameter \( \Gamma \) is sometimes referred to as a “budget of uncertainty”, and is a simple and powerful tool that allow decision-makers to adjust and customize the robust optimization models according to their degree of aversion to uncertainty. Such an approach can provide some flexibility in preventing linear programming solutions from being overly conservative (see, e.g., Ben-Tal and Nemirovski 1999, Bertsimas et al. 2004, and references therein). When \( \mathcal{U}(\Gamma) \) is assumed to be a conic representable convex set, and \( A \) and \( b \) are affine in \( z \), problem (4) can yield tractable and computationally attractive formulations (see, e.g. Ben-Tal et al. 2004), provided of course that \( \mathcal{X} \) is also tractable. For instance, when \( \mathcal{U}(\Gamma) \) is described as norm-based sets \( \mathcal{U}(\Gamma) = \{ z \in \mathcal{W} \mid ||z|| \leq \Gamma \} \), the problem can be reformulated as a linear programming (LP) model for the \( ||\cdot||_1 \) and \( ||\cdot||_{\infty} \) norms (Bertsimas et al. 2004), as well as for the D-norm introduced by Bertsimas and Sim (2004), provided \( \mathcal{X} \) is polyhedral. For the Euclidean norm \( ||\cdot||_2 \), the resulting problem is a second-order cone problem (SOCP) (Chen et al. 2007), provided \( \mathcal{X} \) is SOCP-representable.

Similar to the case of the P-model, we can also obtain a robust satisficing decision criterion, associated with the robust optimization problem (4), in a form of a function \( \nu_R : \mathbb{R}^N \mapsto \mathbb{R} \cup \{-\infty\} \) given by

$$\nu_R(x) = \max_{\alpha \geq 0} \{ \alpha \mid A(z)x \geq b(z) \quad \forall z \in \mathcal{U}(\alpha) \}.$$  

(5)

Because the family of uncertainty sets \( \mathcal{U}(\alpha) \) is non-decreasing in \( \alpha \in \mathbb{R}_+ \), the constraints of the robust optimization problem in (4) is the same as \( \nu_R(x) \geq \Gamma \). Consequently, the robust satisficing decision criterion (5) determines, for a given \( x \), the largest possible uncertainty set \( \mathcal{U}(\nu_R(x)) \) so that the problem’s constraints would remain feasible for all realization of uncertain outcomes within the set. Using the convention \( \max \emptyset = -\infty \), if \( x \) is infeasible in the problem’s constraints for all \( z \in \mathcal{W} \), then \( \nu_R(x) = -\infty \).
Analogous to the P-model, we propose the R-model in which the robust satisficing decision criterion (5) is the objective as follows

\[
\begin{align*}
\max \alpha \\
\text{s.t. } A(z)x &\geq b(z) \quad \forall z \in U(\alpha) \\
x &\in X \\
\alpha &\geq 0.
\end{align*}
\] (6)

Speaking intuitively, the R-model determines the most “robust” solution that would remain feasible in the problem’s constraints when z arise over a maximally sized uncertainty set. It is also easy to see that the solution of the R-model is almost as efficient as solving a standard robust optimization problem. In particular, the problem can be solved via a binary search in \(\alpha\), where in each step, the remaining problem with a fixed \(\alpha\) is equivalent to a robust optimization problem with a structure as in (4).

Some recent works (Brown and Sim 2008, Chen and Sim 2009, Brown et al. 2012) have proposed decision criteria related to satisficing via a dual characterization of convex risk measures in the context of financial investments with specified target returns. An extension of such a framework for the case of multiple objectives is considered by Lam et al. (2013). The authors proposed a multiple objectives shortfall-aware (MOS) criterion to evaluate the level of compliance in achieving a set of targets under uncertainty jointly, and developed efficient heuristics based on optimizing the MOS criterion. The authors also report that the MOS optimization solutions are very competitive in performance compared to the P-model in their numerical studies.

In this work, we propose a new and generalized representation of satisficing decision criteria, based on the key characteristics of satisficing, which can then be applied in various optimization paradigms including the above-mentioned models. In relation to this, we propose the S-model, where the objective is to maximize a satisficing decision criterion, and also the satisficing constrained optimization problem, where the satisficing decision criterion is enforced as a constraint. In particular, we motivate and focus on a class of tractable probabilistic S-model, denoted as the T-model, which is a lower bound of the P-model and inherits the computational benefits of the R-model. As a disclaim, we use the term “tractable” more loosely in this paper to include moderate sized optimization problems such as mixed integer programs (MIP), which can be solved to optimality by general purpose solvers in many practical situations. While He et al. (2015) propose a special case of our T-model to address a queuing network problem with discrete random service times, our framework provides a “standard form” approach that can be adopted for a wide variety of satisficing problems based on both continuous and discrete random variables. We also provide computational studies for a stochastic maximum coverage problem, showing that the T-model outperforms other benchmarks based on sampling average approximations.
Structure of the Paper. This paper is organized as follows. In §2 we elucidate the key features of satisficing decision making and provide the functional representation of a satisficing decision criterion. Among others, we also propose the general framework of a satisficing model or the S-model, in which the satisficing decision criterion is maximized. In §3, we focus on a tractable class of S-models that incorporate available probability information, and refer to these as T-models. Both continuous and discrete probability distributions are discussed, each leading to different implementations. In the case when probability density functions of the uncertainties are log-concave, we propose a solution approach based on piecewise-linear concave approximations. In the case of discrete probability distributions, we provide a linear MIP formulation of the T-model. Finally, we consider extensions of the T-model to multi-stage settings where decisions are adaptable in the uncertainties revealed. §4 provides computational studies on a stochastic coverage problem, confirming that the previous theoretical developments may result in highly tractable models. Extensive computational experiments show that the implementations of the T-model are highly scalable and have a significant computational advantage in terms of computing time and solution quality when compared to standard sample based approaches. We conclude the paper in §5.

Notation. Given \( N \in \mathbb{N} \), we use \([N]\) to denote the set of running indices, \( \{1, \ldots, N\} \). We generally use bold faced characters such as \( \mathbf{x} \in \mathbb{R}^N \) and \( \mathbf{A} \in \mathbb{R}^{M \times N} \) to represent vectors and matrices. We use \( x_i \) to denote the \( i \)-th element of vector \( \mathbf{x} \). We let \( \mathbf{e}_i \) be the unit vector, i.e., the \( i \)-th column of the identity matrix. We use the tilde sign to denote an uncertain or random parameter such as \( \tilde{z} \) without necessarily associating it with a particular probability distribution. For a set \( \mathcal{U} \subseteq \mathbb{R}^K \), \( P(\tilde{z} \in \mathcal{U}) \) represents the probability of \( \tilde{z} \) being in the set \( \mathcal{U} \) evaluated on the distribution \( P \). We use the convention, \( \ln 0 = \max \emptyset = -\infty \) and \( \min \emptyset = \infty \).

2. Satisficing Decision Criteria and Models

In this section, we formalize the notion of satisficing decision criteria that would encompass, as special cases, the log-probability and the robust satisficing decision criteria in (2) and (5) respectively. Based on the premise of a satisficing approach to decision-making under uncertainty as motivated in the previous section, we first define the two key characteristics of a satisficing decision criterion as follows.

**Definition 1.** Let \( \mathbf{A} \) and \( \mathbf{b} \) be defined as in (1), and let \( \mathcal{W} \subseteq \mathbb{R}^K \) be the support of the random vector \( \tilde{z} \). A function \( \nu: \mathbb{R}^N \mapsto \mathbb{R} \cup \{-\infty\} \) is a satisficing decision criterion if and only if it has the following two properties. For all \( \mathbf{x}, \mathbf{y} \in \mathbb{R}^N \),

1. (Satisficing dominance) If \( \mathbf{A}(\tilde{z})\mathbf{y} \geq \mathbf{b}(\tilde{z}) \) implies \( \mathbf{A}(\tilde{z})\mathbf{x} \geq \mathbf{b}(\tilde{z}) \) for all \( \tilde{z} \in \mathcal{W} \), then \( \nu(\mathbf{x}) \geq \nu(\mathbf{y}) \).
2. (Infeasibility) If there does not exist \( \tilde{z} \in \mathcal{W} \) such that \( \mathbf{A}(\tilde{z})\mathbf{x} \geq \mathbf{b}(\tilde{z}) \), then \( \nu(\mathbf{x}) = -\infty \).
Note that the *satisficing dominance* property ensures that if \( x \) is feasible whenever \( y \) is feasible, then \( x \) should be no less preferable than \( y \). As a consequence, if \( x \) and \( y \) are always feasible across all uncertain outcomes \( z \in \mathcal{W} \), then they should have the highest preference and \( \nu(x) = \nu(y) \) because \( \nu(x) \geq \nu(y) \) and \( \nu(y) \geq \nu(x) \). Therefore, solutions that are always feasible are equally preferred.

In contrast, the *infeasibility* property requires that a solution that never achieves feasibility in any possible outcome would not be an acceptable solution. It is easy to verify that the decision criteria \( \nu_P \) in (2) and \( \nu_R \) in (5) are indeed satisficing decision criteria according to the above definition. On the other hand, consider a decision criterion \( \nu_E : \mathbb{R}^N \mapsto \mathbb{R} \) based on penalizing expected shortfalls defined as follows,

\[
\nu_E(x) = \mathbb{E}\left( \sum_{j \in \mathcal{M}} \min \left\{ \left[ A(\tilde{z})x - b(\tilde{z}) \right]_j, 0 \right\} \right).
\]

Although it appears to be related to satisficing decision making, it is not a satisficing decision criterion as defined in Definition 1, because it violates both properties. There is also an important distinction of this satisficing decision criterion with the different types of satisficing-related decision criteria proposed in Brown and Sim (2008), Chen and Sim (2009), Brown et al. (2012) and Lam et al. (2013). While those criteria are consistent with the infeasibility property, they are not necessarily consistent with the satisficing dominance property. Instead, these criteria would satisfy a weaker form of satisficing property that requires solutions that are always feasible to have the highest preference, which is implied by the satisficing dominance property.

Based on Definition 1, we now provide a general representation of any satisficing decision criterion \( \nu \) in the following result.

**Theorem 1.** Consider a function \( \nu : \mathbb{R}^N \mapsto \mathbb{R} \cup \{-\infty\} \) defined as

\[
\nu(x) = \max_{\alpha \in \mathcal{S}} \{ \rho(\alpha) \mid A(z)x \geq b(z) \ \forall z \in \mathcal{U}(\alpha) \} \tag{7}
\]

for some function \( \rho : \mathcal{S} \mapsto \mathbb{R} \cup \{-\infty\} \) on domain \( \mathcal{S} \subseteq \mathbb{R}^P \), and for some family of nonempty uncertainty sets \( \mathcal{U}(\alpha) \subseteq \mathcal{W} \) defined for all \( \alpha \in \mathcal{S} \). Then the function \( \nu \) is a satisficing decision criterion. Moreover, any satisficing decision criterion can be represented in a form given by (7) with \( \mathcal{S} \subseteq \mathbb{R}^N \).

**Proof:** We first establish that the function in (7) is indeed a satisficing decision criterion. Let \( x, y \in \mathbb{R}^N \) such that \( A(z)y \geq b(z) \) implies \( A(z)x \geq b(z) \) for all \( z \in \mathcal{W} \). Since \( \mathcal{U}(\alpha) \subseteq \mathcal{W} \) for all \( \alpha \in \mathcal{S} \), we have \( A(z)x \geq b(z) \) for all \( z \in \mathcal{U}(\alpha) : A(z)y \geq b(z) \). We then have \( \nu(y) \leq \nu(x) \) by definition of (7). This establishes the satisficing dominance property. Assume now that \( x \) is infeasible in the constraints \( A(z)x \geq b(z) \) for any \( z \in \mathcal{W} \), then Problem (7) will always be infeasible for all \( \alpha \in \mathcal{S} \). Hence, we have \( \nu(x) = -\infty \), which shows consistency with the infeasibility property.
We next show that any satisficing decision criterion can be represented in a form given by (7) in which \( P = N \). Consider a satisficing decision criterion, \( \bar{\nu} : \mathbb{R}^N \mapsto \mathbb{R} \cup \{-\infty\} \). Define

\[
S = \{ \alpha \in \mathbb{R}^N \mid \exists z \in W \text{ such that } A(z)\alpha \geq b(z) \},
\]

\[
U(\alpha) = \{ z \in W \mid A(z)\alpha \geq b(z) \},
\]

and \( \rho(\alpha) = \bar{\nu}(\alpha) \). Note that \( S \subseteq \mathbb{R}^N \) and \( U(\alpha) \neq \emptyset \) for all \( \alpha \in S \). For a given \( x \in \mathbb{R}^N \), define \( \nu(x) \) as

\[
\nu(x) = \max_{\alpha \in S} \bar{\nu}(\alpha) \quad \text{s.t.} \quad A(z)x \geq b(z) \forall z \in U(\alpha)
\]

(8)

Note that \( \nu(x) \) is of the form represented in (7). We now show that \( \nu(x) = \bar{\nu}(x) \). Given a solution \( \alpha \in S \) of Problem (8), then for all \( z \in W \) such that \( A(z)\alpha \geq b(z) \), we must have, by definition of (8), \( x \) feasible in \( A(z)x \geq b(z) \). Hence, it follows from the property of satisficing dominance of \( \bar{\nu} \) that \( \bar{\nu}(x) \geq \bar{\nu}(\alpha) \) for all \( \alpha \) feasible in Problem (8). Consequently, \( \bar{\nu}(x) \) is an upper bound of Problem (8). For the case when there exists \( z \in W \) such that \( A(z)x \geq b(z) \), we observe that the bound is achievable since \( \alpha = x \) is feasible in Problem (8). In which case, we would have \( \nu(x) = \bar{\nu}(x) \). For the case when there does not exist \( z \in W \) such that \( A(z)x \geq b(z) \), then \( \bar{\nu}(x) = -\infty \). Consequently, Problem (8) would be infeasible for all \( \alpha \in S \). Hence, we would have \( \nu(x) = -\infty \) as well. \( \square \)

Based on the complete characterization and representation of the satisficing decision criterion in Theorem 1, we can now propose a general format of the satisficing model or the S-model where satisficing is the objective as follows:

\[
\max_{\alpha \in S} \rho(\alpha) \\
\text{s.t. } A(z)x \geq b(z) \forall z \in U(\alpha) \\
x \in \mathcal{X} \\
\alpha \in S
\]

(9)

where \( \rho(\alpha), U(\alpha) \) and \( S \subseteq \mathbb{R}^P \), are specific problem-dependent representation choices.

Note that the R-model has a form similar to the S-model, but restricted to having \( \alpha \in \mathbb{R}_+ \) and \( \rho(\alpha) = \alpha \). The S-model, with its more general family of uncertainty sets \( U(\alpha) \) indexed over \( S \), and its more general function \( \rho(\alpha) \) for representing its level of satisficing, offers greater flexibility. Less intuitively, but a consequence of the proof of Theorem 1, this flexibility is also rich enough to incorporate the P-model as a special case. In line with the chance-constrained optimization problem (3) and the robust optimization problem (4), we also propose the general format of the satisficing-constrained optimization problem as follows

\[
\min c^T x \\
\text{s.t. } \rho(\alpha) \geq \sum \quad A(z)x \geq b(z) \forall z \in U(\alpha) \\
x \in \mathcal{X} \\
\alpha \in S
\]

(10)
where a lower bound $\Sigma$ is now enforced on the satisficing decision criterion, over which the cost objective is minimized.

In spite of the popularity of some satisficing-constrained optimization models such as robust optimization and chance-constrained optimization problems, from the modeling perspective, we believe that the S-model (9) has an advantage over the satisficing-constrained optimization problem (10). In particular, the former removes the burden of having the decision maker to articulate her satisficing level, $\Sigma$, which may be an abstract concept that she is unable to relate to naturally and prohibiting her from specifying its value correctly. Although the S-model does not minimize a cost function, we can always relegate the cost function to a budget constraint, $c'x \leq \beta$ as part of the feasible set $X$ of the S-model. The key parameter to specify in the S-model would then be the budget $\beta$, which, in our opinion, is more intuitive for the decision maker to specify as compared to the parameter $\Sigma$.

By careful design of the adjustable uncertainty set $U(\alpha)$, and the function $\rho(\alpha)$, many interesting models, both from a solution quality and a tractability point of view, can be proposed. In particular, in relation to the P-model and the chance-constrained optimization problem, we can let $\rho(\alpha) = \ln \mathbb{P}(\tilde{z} \in U(\alpha))$, which relates to the probability that the adjustable uncertainty set $U(\alpha)$ allows feasibility of any random outcomes of $\tilde{z}$. Correspondingly, we propose the probabilistic S-model,

$$\begin{align*}
\max \ln \mathbb{P}(\tilde{z} \in U(\alpha)) \\
\text{s.t. } A(z)x & \geq b(z) \quad \forall z \in U(\alpha) \\
x & \in X, \\
\alpha & \in S,
\end{align*}$$

(11)

and the probabilistic satisficing-constrained model,

$$\begin{align*}
\min c'x \\
\text{s.t. } \ln \mathbb{P}(\tilde{z} \in U(\alpha)) & \geq \Delta \\
A(z)x & \geq b(z) \quad \forall z \in U(\alpha) \\
x & \in X, \\
\alpha & \in S.
\end{align*}$$

(12)

In general, the above two frameworks can be viewed as conservative approximations of the P-model (1) and the chance-constrained optimization problem (3), respectively. The approximation is exact when $U(\alpha) = \{z \in W \mid A(z)\alpha \geq b(z)\}$, and $S = \{\alpha \in \mathbb{R}^N \mid \exists z \in W \text{ such that } A(z)\alpha \geq b(z)\}$ (this follows directly from the proof of Theorem 1). Nevertheless, the key to computational tractability is to properly design the adjustable uncertainty set, $U(\alpha)$ so that calculating $\ln \mathbb{P}(\tilde{z} \in U(\alpha))$ would be computationally less demanding compared to that of the log-probability satisficing decision criterion (2).

From a modeling perspective, any poor and conservative approximation of the probabilistic satisficing constrained optimization problem (12) could lead to having good solutions being eliminated.
from the feasible set, resulting in significant cost increases in the objective, which will be unappealing to the decision maker. In contrast, this is less likely to be an issue in the probabilistic S-model, since the satisficing criterion is the objective. Furthermore, a poor and conservative approximation of the probabilistic S-model implies that the actual probability of feasibility may be significantly greater than the objective value reflected in Problem (11), which may not necessarily be detrimental to the decision maker. This has also been articulated in Chen and Sim (2009), Goh and Hall (2013), Lam et al. (2013) and He et al. (2015). By careful design of $\mathcal{U}(\alpha)$, we will focus on the tractable probabilistic S-model and show in our extensive numerical study that this approach can yield high quality solutions that can outperform the solutions of other competing approaches.

3. A Tractable Probabilistic S-Model

In this section, we will focus on a class of tractable probabilistic S-model, but our analysis can easily be extended to solving the probabilistic satisficing-constrained optimization problem. To provide explicit formulations, we assume that the uncertain parameters $\tilde{z}_k$, $k \in [K]$ are independently, but not necessarily identically, distributed random variables with support $\mathcal{W}_k$ so that $\mathcal{W} = \times_{k=1}^{K} \mathcal{W}_k$. We will also focus on the case where each entry in the constraint coefficients $A(z)$ and $b(z)$ are affine functions in $z$. Such an assumption is rather common in the literature (see, e.g., Ben-Tal et al. 2004, Chen et al. 2007) and is effective in capturing data characteristics such as correlations.

To be precise, denote $a_{ij}(z)$ as the $(i,j)^{th}$ entry of $A(z)$ and $b_i(z)$ as the $i^{th}$ entry in $b(z)$. The affine factor model is defined as:

$$a_{ij}(z) = a_{ij}^0 + \sum_{k \in [K]} a_{ij}^k z_k \forall i \in [M], \forall j \in [N], \quad b_i(z) = b_i^0 + \sum_{k \in [K]} b_i^k z_k \forall i \in [M]. \quad (13)$$

In view of (13), we will, whenever appropriate, also refer to the uncertain parameters $\tilde{z}_k$, $k \in [K]$, as random factors in our model of uncertainty. To further obtain a tractable model, the probabilistic S-model is implemented as follows. Define $\underline{\alpha}_k$, $\overline{\alpha}_k$, with $\underline{\alpha}_k \leq \overline{\alpha}_k$, $\underline{\alpha}_k, \overline{\alpha}_k \in \mathcal{W}_k$ for all $k \in [K]$. Define also $\underline{\alpha} = (\underline{\alpha}_1, \ldots, \underline{\alpha}_K)$, $\overline{\alpha} = (\overline{\alpha}_1, \ldots, \overline{\alpha}_K)$, and $\alpha = (\underline{\alpha}, \overline{\alpha})$, where $\alpha$ are the adjustable uncertainty set parameters in Problem (9). We then have $\alpha \in \mathcal{S} \subseteq \mathbb{R}^P$ where $P = 2K$ and $\mathcal{S} = \{(\underline{\alpha}, \overline{\alpha}) \in \mathbb{R}^{2K} : \underline{\alpha} \leq \overline{\alpha}, \underline{\alpha}, \overline{\alpha} \in \mathcal{W}\}$. The family of adjustable uncertainty sets $\mathcal{U}(\alpha)$ in (11) is then defined as:

$$\mathcal{U}(\alpha) = \mathcal{U}(\underline{\alpha}, \overline{\alpha}) = \left\{ z \in \mathbb{R}^K : z \in [\underline{\alpha}, \overline{\alpha}] \right\}. $$

Under the assumption of stochastic independence and the “box” typed sets $\mathcal{U}(\alpha)$, the objective function of the probabilistic S-model can then be evaluated as:

$$\ln \mathbb{P}(\tilde{z} \in \mathcal{U}(\underline{\alpha}, \overline{\alpha})) = \ln \prod_{k \in [K]} \mathbb{P}(\underline{\alpha}_k \leq \tilde{z}_k \leq \overline{\alpha}_k) = \sum_{k \in [K]} \ln \mathbb{P}(\underline{\alpha}_k \leq \tilde{z}_k \leq \overline{\alpha}_k).$$
Note that a computational attractiveness of the above is that it does not require high dimensional integration to be performed. The resulting probabilistic S-model formulation, which we call the T-model, is then as follows:

$$\begin{align*}
\max & \sum_{k \in [K]} \ln \mathbb{P}(\alpha_k \leq z_k \leq \alpha_k) \\
\text{s.t.} & \quad A(z)x \geq b(z) \quad \forall z \in [\alpha, \alpha] \\
& \quad x \in X \\
& \quad \alpha \leq \alpha, \alpha, \alpha \in W.
\end{align*}$$ (14)

Generally, we can regard the T-model as a conservative approximation to the P-model that eliminates the cumbersome procedure of high-dimensional integration. Nevertheless, there are instances where the solutions of these two models would coincide. This is shown in the following result.

**Theorem 2.** The solutions of the P-model (1) are equivalent to those of the T-model (14) if each constraint \(a_i(z)x \geq b_i(z), i \in [M],\) is affected by at most one random factor \(\tilde{z}_{\kappa(i)},\) where \(\kappa : [M] \to [K]\) is a function that identifies the random factor \(\tilde{z}_k\) for the \(i^{th}\) constraint.

**Proof:** We focus on the nontrivial case when the objective of the P-model (1) is finite. Since each constraint is affected by at most one random factor, we consider the following restricted affine factor model based on (13):

$$a_{ij}(z) = a_{ij}^0 + a_{ij}^{\kappa(i)}z_{\kappa(i)} \quad \forall i \in [M], j \in [N], \quad b_i(z) = b_i^0 + b_i^{\kappa(i)}z_{\kappa(i)} \quad \forall i \in [M].$$

Consider any feasible solution \(x \in X\) such that \(\mathbb{P}(A(z)x \geq b(z)) > 0.\) Let

$$M_k^0 = \left\{ i \in [M] \mid \sum_{j \in [N]} a_{ij}^k x_j = b_i^k, \kappa(i) = k \right\},$$

$$M_k^+ = \left\{ i \in [M] \mid \sum_{j \in [N]} a_{ij}^k x_j > b_i^k, \kappa(i) = k \right\}, \quad M_k^- = \left\{ i \in [M] \mid \sum_{j \in [N]} a_{ij}^k x_j < b_i^k, \kappa(i) = k \right\},$$

for \(k \in [K].\) Observe that under the restricted affine uncertainty, the robust counterpart for the \(i^{th}\) constraint \(a_i(z)x \geq b_i(z), \forall z \in [\alpha, \alpha], i \in [M],\) of Problem (14) is equivalent to

$$\sum_{j \in [N]} a_{ij}^0 x_j + \min_{\omega_{\kappa(i)} \leq \tilde{z}_{\kappa(i)} \leq \alpha_{\kappa(i)}} \left( \sum_{j \in [N]} a_{ij}^{\kappa(i)} x_j - b_i^{\kappa(i)} \right) z_{\kappa(i)} \geq b_i^0,$$

or equivalently, we have for all \(k \in [K],\)

$$\sum_{j \in [N]} a_{ij}^0 x_j \geq b_i^0 \quad \forall i \in M_k^0, \quad z_k = \alpha_k \geq \frac{b_i^0 - \sum_{j \in [N]} a_{ij}^0 x_j}{\sum_{j \in [N]} a_{ij}^k x_j - b_i^k} \quad \forall i \in M_k^+, \quad z_k = \alpha_k \leq \frac{b_i^0 - \sum_{j \in [N]} a_{ij}^0 x_j}{\sum_{j \in [N]} a_{ij}^k x_j - b_i^k} \quad \forall i \in M_k^-.$$
Note that since \( P(\mathbf{A}(\tilde{z})\mathbf{x} \geq \mathbf{b}(\tilde{z})) > 0 \), it implies that the constraints from the set \( \bigcup_{k \in [K]} \mathcal{M}_k^0 \) are all feasible. Based on the notation \( \max \emptyset = -\infty \) and \( \min \emptyset = \infty \), observe that
\[
\ln P(\mathbf{A}(\tilde{z})\mathbf{x} \geq \mathbf{b}(\tilde{z}))
= \ln P \left( \sum_{j \in [N]} \left( a_{ij}^0 + a_{ij}^k \tilde{z}_{k(i)} \right) x_j \geq b_i^0 + b_i^k \tilde{z}_{k(i)}, \, \forall i \in \mathcal{M}_k^+ \cup \mathcal{M}_k^- \cup \mathcal{M}_k^0, \, k \in [K] \right)
= \sum_{k \in [K]} \ln P \left( \max_{i \in \mathcal{M}_k^+} \left( \frac{a_{ij}^0 - \sum_{j \in [N]} a_{ij}^k x_j}{\sum_{j \in [N]} a_{ij}^k x_j - b_i^k} \right) \leq \tilde{z}_k \leq \min_{i \in \mathcal{M}_k^-} \left( \frac{b_i^0 - \sum_{j \in [N]} a_{ij}^k x_j}{\sum_{j \in [N]} a_{ij}^k x_j - b_i^k} \right) \right)
= \max_{\tilde{\alpha}, \tilde{\alpha} \in \mathbb{W}} \left\{ \sum_{k \in [K]} \ln P \left( \tilde{\alpha}_k \leq \tilde{z}_k \leq \tilde{\alpha}_k \right) \mid \tilde{\alpha}_k \geq \sum_{j \in [N]} a_{ij}^k x_j - b_i^k \tilde{\alpha}_k, \forall i \in \mathcal{M}_k^+, \, \tilde{\alpha}_k \leq \sum_{j \in [N]} a_{ij}^k x_j - b_i^k \forall i \in \mathcal{M}_k^-, \, k \in [K] \right\}
= \max_{\tilde{\alpha}, \tilde{\alpha} \in \mathbb{W}} \left\{ \sum_{k \in [K]} \ln P \left( \tilde{\alpha}_k \leq \tilde{z}_k \leq \tilde{\alpha}_k \right) \mid \mathbf{A}(\tilde{z})\mathbf{x} \geq \mathbf{b}(\tilde{z}) \forall \tilde{z} \in [\tilde{\alpha}, \tilde{\alpha}] \right\}.
\]

\( \square \)

We now turn our attention to reformulations, or robust counterparts, of the T-model (14) that are more amenable to computation.

**Theorem 3.** The T-model (14) is equivalent to the following explicit nonlinear optimization problem:
\[
\max_{\tilde{\alpha}, \tilde{\alpha} \in \mathbb{W}} \sum_{k \in [K]} \ln P \left( \tilde{\alpha}_k \leq \tilde{z}_k \leq \tilde{\alpha}_k \right)
\text{s.t.} \quad \sum_{j \in [N]} a_{ij}^0 x_j + \sum_{k \in [K]} v_{ik} \geq b_i^0 \quad \forall i \in [M]
= \sum_{j \in [N]} a_{ij}^k x_j - b_i^k \tilde{\alpha}_k \quad \forall i \in [M], \, k \in [K]
\]
(15)

\[
\sum_{j \in [N]} a_{ij}^k x_j - b_i^k \tilde{\alpha}_k \quad \forall i \in [M], \, k \in [K]
\]
\[\mathbf{x} \in \mathbb{X}, \mathbf{v} \in \mathbb{R}^{M \times K}, \quad \tilde{\alpha} \leq \tilde{\alpha}, \quad \tilde{\alpha}, \tilde{\alpha} \in \mathbb{W}.\]

**Proof:** Under the general affine factor model in (13), the robust counterpart for each of the \( i^{th} \) constraint \( \mathbf{a}_i(z)^T \mathbf{x} \geq b_i(z), \forall z \in [\tilde{\alpha}, \tilde{\alpha}], \, i \in [M] \), of Problem (14) can be written as:
\[
\sum_{j \in [N]} a_{ij}^0 x_j + \min_{\tilde{\alpha} \leq \tilde{\alpha} \leq \tilde{\alpha}} \left\{ \sum_{k \in [K]} \left( \sum_{j \in [N]} a_{ij}^k x_j - b_i^k \right) z_k \right\} \geq b_i^0.
\]
Since the minimization in the above operates on a linear function in \( z \), with \( z \in [\tilde{\alpha}, \tilde{\alpha}] \), the above is equivalent to:
\[
\sum_{j \in [N]} a_{ij}^0 x_j + \sum_{k \in [K]} \min_{\tilde{\alpha} \leq \tilde{\alpha} \leq \tilde{\alpha}} \left\{ \left( \sum_{j \in [N]} a_{ij}^k x_j - b_i^k \right) z_k \right\} \geq b_i^0,
\]
(16)
which results in the desired formulation. \( \square \)
We next present an important condition, which may arise in practical situations, and which has significant consequences on the formulation and tractability of the T-model.

**Definition 2.** A T-model is **monotone** with respect to the uncertain parameters $\tilde{z}$ if there exists a partition $\overline{K}, \underline{K} \subseteq [K]$, i.e., $\overline{K} \cap \underline{K} = \emptyset$, $\overline{K} \cup \underline{K} = [K]$ such that for all $k \in \overline{K}$

$$\sum_{j \in [N]} a_{ij} \alpha_{kj} \leq b_{ki} \forall i \in [M], \alpha_{k} \in \mathcal{X}$$

and for all $k \in \underline{K}$

$$\sum_{j \in [N]} a_{ij} \alpha_{kj} \geq b_{ki} \forall i \in [M], \alpha_{k} \in \mathcal{X}.$$  

Speaking intuitively, the conditions of monotone T-models arise if each uncertain parameter, $\tilde{z}_k$, $k \in [K]$ always has the same direction of influence (either more or less of its value) in a manner that reduces the feasibility of the problem’s constraints, regardless of the solution $\alpha \in \mathcal{X}$. For instance, in a problem of fulfilling uncertain customers’ demands by a company, increasing a customer’s demand will always have a negative effect on the company’s ability to satisfy all the demands collectively. Our next result shows that monotone T-models enable simpler formulations than that in (15).

**Proposition 1.** *If the T-model is monotone, Problem (15) can be further simplified as follows*

$$\max \sum_{k \in \overline{K}} \ln P(\tilde{z}_k \leq \alpha_k) + \sum_{k \in \underline{K}} \ln P(\tilde{z}_k \geq \alpha_k)$$

s.t. $\sum_{j \in [N]} \left( a_{ij}^0 + \sum_{k \in \overline{K}} a_{ij}^k \alpha_k + \sum_{k \in \underline{K}} a_{ij}^k \overline{\alpha}_k \right) x_j \geq b_{i}^0 + \sum_{k \in \overline{K}} b_{i}^k \alpha_k + \sum_{k \in \underline{K}} b_{i}^k \overline{\alpha}_k \forall i \in [M]$ \hspace{1cm} (17)

$x \in \mathcal{X}, \alpha, \overline{\alpha} \in \mathcal{W}$.

We omit the proof as it follows trivially from Theorem 3. Note that the constraint $\alpha \leq \overline{\alpha}$ is redundant and hence removed.

As we will see in §3.3, monotone T-models also have significant computational advantages when addressing multi-stage satisficing problems. Nevertheless, the T-model remains computationally challenging. Specifically,

1. the objective function is not necessarily concave;
2. the terms $x_j \alpha_k$ and $x_j \overline{\alpha}_k$, $j \in [N], k \in [K]$ are bilinear.

Consequently, we will explore various ways to linearize the problem and therefore enable the T-model to be solved via state-of-the-art general purpose MIP solvers such as CPLEX or Gurobi.
3.1. Case of log concave densities

For certain classes of continuously distributed random variables the T-model objective function in (14) remains concave in \((\alpha, \bar{\alpha})\). This is the case, for instance, for the class of random variables with log-concave densities, which include commonly-used distributions such as the exponential, uniform and normal distributions. Indeed, this is a consequence of the well known results of Prékopa (1980) and, from which, we obtain the following corollary.

**Proposition 1** (See Prékopa 1980, Theorem 9) Suppose \(\tilde{z}_k\) is a continuously distributed random variable with log-concave density function \(f_k(z) : W_k \mapsto \mathbb{R}_+\). Then the function \(F_k(\hat{\delta}, \tilde{\delta}) : D_k \mapsto \mathbb{R}\),

\[
F_k(\hat{\delta}, \tilde{\delta}) = \ln \mathbb{P}(\hat{\delta} \leq \tilde{z}_k \leq \tilde{\delta})
\]

is a concave function of \((\hat{\delta}, \tilde{\delta})\) on domain \(D_k = \{(\hat{\delta}, \tilde{\delta}) \in W_k^2 : \hat{\delta} < \tilde{\delta}\}\).

Notwithstanding the fact that the constraint functions are bilinear, there are useful situations where Problem (15) will become tractable. Most notably, this is the case when the uncertainty occurs only at the right-hand side, so that \(a_{kj}^r = 0, \forall i \in \{M\}, j \in \{N\}, k \in \{K\}\). Another tractable situation arises when \(x_j\) is discrete, in which case, the bilinear terms \(x_j\alpha_k\) and \(x_j\pi_k\) can be linearized using standard mixed integer programming techniques. For instance, if \(x_j \in \{0, 1\}\), the bilinear term \(x_j\alpha_k\) can be replaced by a new decision variable \(r_{jk}\) that satisfies the following linear inequalities:

\[-\Theta(1 - x_j) \leq r_{jk} - \alpha_k \leq \Theta(1 - x_j), \text{ and } -\Theta x_j \leq r_{jk} \leq \Theta x_j,
\]

for a sufficiently large constant \(\Theta\).

**Density Curve Cuts.** Although the objective function in (14) is nonlinear, an important advantage of maximizing a concave objective function is that it can be solved efficiently in practice via piecewise linear approximations of arbitrary accuracy. For any \(D' \subseteq D_k\), we note that each term \(\ln \mathbb{P}(\alpha_k \leq \tilde{z}_k \leq \pi_k)\) in the objective function of (14) can be written as follows:

\[
\ln \mathbb{P}(\alpha_k \leq \tilde{z}_k \leq \pi_k) = \max \left\{ p_k \in \mathbb{R} \mid p_k \leq F_k(\alpha_k, \pi_k) \right\} = \max \left\{ p_k \in \mathbb{R} \mid p_k \leq \min_{(\hat{\delta}, \tilde{\delta}) \in D_k} F_k(\hat{\delta}, \tilde{\delta}) + F_k^1(\hat{\delta}, \tilde{\delta})(\pi_k - \tilde{\delta}) + F_k^2(\hat{\delta}, \tilde{\delta})(\alpha_k - \hat{\delta}) \right\} \leq \max \left\{ p_k \in \mathbb{R} \mid p_k \leq F_k(\hat{\delta}, \tilde{\delta}) + F_k^1(\hat{\delta}, \tilde{\delta})(\pi_k - \tilde{\delta}) + F_k^2(\hat{\delta}, \tilde{\delta})(\alpha_k - \hat{\delta}) \forall (\hat{\delta}, \tilde{\delta}) \in D'_k \right\}
\]

where the functions \(F_k^1\), and \(F_k^2\) on the domain \(D_k\) are defined as

\[
F_k^1(\hat{\delta}, \tilde{\delta}) = \frac{\partial}{\partial \hat{\delta}} F_k(\hat{\delta}, \tilde{\delta}) = \frac{f_k(\tilde{\delta})}{\mathbb{P}(\hat{\delta} \leq \tilde{z}_k \leq \tilde{\delta})}, \text{ and } F_k^2(\hat{\delta}, \tilde{\delta}) = \frac{\partial}{\partial \tilde{\delta}} F_k(\hat{\delta}, \tilde{\delta}) = -\frac{f_k(\tilde{\delta})}{\mathbb{P}(\hat{\delta} \leq \tilde{z}_k \leq \tilde{\delta})}.
\]
The second equality follows from the first-order necessary and sufficient conditions of concave functions, and by noting that for any given \((\alpha_k, \sigma_k) \in D_k\) the minima is trivially achieved by choosing \((\delta, \overline{\delta}) = (\alpha_k, \sigma_k)\). This thus establishes the equivalent reformulation of \(F_k(\alpha_k, \sigma_k)\) as a point-wise minimum of a set of affine functions.

Hence, a relaxation of the problem can be achieved by replacing the original objective function by the last construct as in the inequality above for some \(D'_k\), for each \(k \in K\). We term the augmented inequalities:

\[
p_k \leq F_k(\delta, \overline{\delta}) + F_1(\delta, \overline{\delta})(\sigma_k - \overline{\delta}) + F_2(\delta, \overline{\delta})(\alpha_k - \delta) \forall (\delta, \overline{\delta}) \in D'_k
\]

as density curve cuts, which are affine in the decision variables, and are therefore implementable on general purpose solvers. In practice, the problem can be solved in an iterative cut generation procedure (if all variables are continuous) or in a branch-and-cut procedure (if some of the variables are integer) (see, e.g., Wolsey 1998), adding cuts (18) whenever they are violated. From a computational point of view, this approach is likely to be particularly attractive when the number of violated cuts is small.

### 3.2. Case of discrete distributions

We now focus on T-models based on random variables with discrete distributions. In particular, we model \(z_k\) on the discrete support \(\mathcal{W}_k = \{\zeta_1^k, \zeta_2^k, \ldots, \zeta_{L(k)}^k\}\), with strictly positive probability mass functions \(P(\zeta_k = \zeta_\ell^k) = p_\ell^k \forall \ell \in [L(k)]\). We also define \(\lambda_0^k = 0\) and \(\lambda_\ell^k = \sum_{t \in [t]} p_\ell^k\) for all \(k \in [K], \ell \in [L(k)]\). Without loss of generality, we assume that the outcomes \(\zeta_\ell^k\) are ranked in nondecreasing values in \(\ell = 1, \ldots, L(k)\). A specification of the T-model is developed as follows.

First, we define the adjustable uncertainty set parameters \(\alpha\) in (9). Let \(\overline{\alpha}_k = (\overline{\alpha}_1^k, \ldots, \overline{\alpha}_{L(k)}^k)\), \(\underline{\alpha}_k = (\underline{\alpha}_1^k, \ldots, \underline{\alpha}_{L(k)}^k)\), \(\overline{\alpha} = (\overline{\alpha}_1, \ldots, \overline{\alpha}_K)\), \(\underline{\alpha} = (\underline{\alpha}_1, \ldots, \underline{\alpha}_K)\), and \(\alpha = (\underline{\alpha}, \overline{\alpha})\). Hence, we have \(P = 2 \times \sum_{k \in [K]} L(k)\). In the following, \(\overline{\alpha}_k^\ell\) and \(\underline{\alpha}_k^\ell\) are modeled as binary variables that take value 1 if \(\zeta_\ell^k\) are the selected lower and upper bounds, respectively, of the interval that defines the discrete outcomes in the adjustable uncertainty set for random variable \(z_k\), and 0 otherwise. We define:

\[
S = \left\{ \alpha \in \{0, 1\}^P \mid \sum_{\ell \in [L(k)]} \underline{\alpha}_k^\ell = 1, \sum_{\ell \in [L(k)]} \overline{\alpha}_k^\ell = 1, \sum_{\ell \in [L(k)]} \ell (\overline{\alpha}_k^\ell - \underline{\alpha}_k^\ell) \geq 0 \forall k \in [K] \right\},
\]

where \(S\) refers to the domain of the \(\alpha\) variables in (9). The intuition for the choice of \(\alpha\) above is to enable a partial ordering on the adjustable uncertainty sets \(\mathcal{U}(\alpha)\) based on ‘counting’ the total number of outcomes \(\zeta_\ell^k\) (and weighted by respective probabilities \(p_\ell^k\)) contained in intervals indicated by \(\underline{\alpha}_k^\ell\) and \(\overline{\alpha}_k^\ell\). The inequality \(\sum_{\ell \in [L(k)]} \ell (\overline{\alpha}_k^\ell - \underline{\alpha}_k^\ell) \geq 0\), together with the fact that \(\zeta_\ell^k\) is
a nondecreasing sequence in $\ell$, ensures that the outcome values corresponding to $\alpha_k$ are at least as large as those of $\overline{\alpha}_k$. Next, we define the adjustable uncertainty sets: 

$$
U(\alpha) = \left\{ z \in \mathcal{W} \mid \sum_{\ell \in [L(k)]} \zeta^\ell_k \alpha^\ell_k \leq z_k \leq \sum_{\ell \in [L(k)]} \zeta^\ell_k \overline{\alpha}^\ell_k, \forall k \in [K] \right\}. \tag{19}
$$

It can be observed that the sets $U(\alpha)$ as defined above are entirely analogous to the adjustable uncertainty sets of the previous case of §3.1 using continuous intervals, with the exception that here $U(\alpha)$ describes discrete sets.

Consolidating the above, the resulting T-model can then be written as:

$$
\begin{align*}
\text{max} & \ln \mathbb{P}(\tilde{z} \in U(\alpha)) \\
\text{s.t.} \quad & A(z)x \geq b(z) \quad \forall z \in U(\alpha) \\
& \sum_{\ell \in [L(k)]} \alpha^\ell_k = 1, \quad \sum_{\ell \in [L(k)]} \overline{\alpha}^\ell_k = 1, \quad \forall k \in [K] \\
& \sum_{\ell \in [L(k)]} \ell(\overline{\alpha}^\ell_k - \alpha^\ell_k) \geq 0 \quad \forall k \in [K] \\
& \alpha_k, \overline{\alpha}_k \in \{0,1\}^{L(k)} \quad \forall k \in [K], \quad x \in \mathcal{X}.
\end{align*} \tag{20}
$$

The following result provides the robust counterpart model to (20), which can then be further reformulated as a linear MIP model.

**Theorem 4.** Under discrete distributions, the T-model (20) is equivalent to the following reformulation:

$$
\begin{align*}
\text{max} & \sum_{k \in [K]} s_k \\
\text{s.t.} \quad & s_k \leq \ln(\gamma) - 1 + \sum_{\ell \in [L(k)]} \frac{1}{\gamma} (\lambda^\ell_k \alpha^\ell_k - \lambda^{\ell-1}_k \alpha^{\ell-1}_k) \forall \gamma \in \mathcal{C}, \, k \in [K] \\
& \sum_{j \in [N]} a^0_{ij} x_j + \sum_{k \in [K]} v_{ik} \geq b^0_i \quad \forall i \in [M] \\
& v_{ik} \leq \sum_{j \in [N]} \sum_{\ell \in [L(k)]} \left( a^\ell_{ij} x_j - b^\ell_i \right) \zeta^\ell_k \alpha^\ell_k \quad \forall i \in [M], \, k \in [K] \\
& v_{ik} \leq \sum_{j \in [N]} \sum_{\ell \in [L(k)]} \left( a^\ell_{ij} x_j - b^\ell_i \right) \zeta^\ell_k \overline{\alpha}^\ell_k \quad \forall i \in [M], \, k \in [K] \\
& \sum_{\ell \in [L(k)]} \alpha^\ell_k = 1, \quad \sum_{\ell \in [L(k)]} \overline{\alpha}^\ell_k = 1, \quad \forall k \in [K] \\
& \sum_{\ell \in [L(k)]} \ell(\overline{\alpha}^\ell_k - \alpha^\ell_k) \geq 0 \quad \forall k \in [K] \\
& x \in \mathcal{X}, \quad v \in \mathbb{R}^{M \times K}, \quad s \in \mathbb{R}^K \\
& \alpha_k, \overline{\alpha}_k \in \{0,1\}^{L(k)} \quad \forall k \in [K],
\end{align*} \tag{21}
$$

where

$$
\mathcal{C}_k = \left\{ \lambda^\ell_k - \lambda^{\ell-1}_k \mid \ell, \overline{\ell} \in [L(k)], \overline{\ell} \geq \ell \right\} \quad \forall k \in [K].
$$
Proof: Note that from the uncertainty set $U(\alpha)$, the lower and upper limits of $\tilde{z}_k$, $k \in [K]$ are 
\[ \sum_{\ell \in [L(k)]} \zeta_k^\ell \alpha_k^\ell \text{ and } \sum_{\ell \in [L(k)]} \zeta_k^\ell \overline{\alpha}_k^\ell, \] respectively. Note also that
\[ \ln P\left( \sum_{\ell \in [L(k)]} \zeta_k^\ell \alpha_k^\ell \leq \tilde{z}_k \leq \sum_{\ell \in [L(k)]} \zeta_k^\ell \overline{\alpha}_k^\ell \right) = \ln \left( \sum_{\ell \in [L(k)]} \left( \lambda_k^\ell \overline{\alpha}_k^\ell - \lambda_k^{\ell-1} \alpha_k^\ell \right) \right). \]
Since $\ln(\delta)$ is a concave function on domain $\delta > 0$, it follows that
\[ \ln(\delta) \leq \ln(\gamma) + \frac{1}{\gamma}(\delta - \gamma) = \ln(\gamma) - 1 + \frac{\delta}{\gamma} \]
for all $\gamma > 0$ and that equality is achieved trivially at $\gamma = \delta$. Moreover, we observe that
\[ \sum_{\ell \in [L(k)]} \left( \lambda_k^\ell \overline{\alpha}_k^\ell - \lambda_k^{\ell-1} \alpha_k^\ell \right) \in C_k, \]
since
\[ \sum_{\ell \in [L(k)]} \overline{\alpha}_k^\ell = 1, \sum_{\ell \in [L(k)]} \alpha_k^\ell = 1, \sum_{\ell \in [L(k)]} \ell(\overline{\alpha}_k^\ell - \alpha_k^\ell) \geq 0. \]
Hence, we have
\[ \ln \left( \sum_{\ell \in [L(k)]} \left( \lambda_k^\ell \overline{\alpha}_k^\ell - \lambda_k^{\ell-1} \alpha_k^\ell \right) \right) \leq \ln(\gamma) - 1 + \sum_{\ell \in [L(k)]} \frac{1}{\gamma} \left( \lambda_k^\ell \overline{\alpha}_k^\ell - \lambda_k^{\ell-1} \alpha_k^\ell \right) \]
for all $\gamma > 0$ and note that equality is achieved when
\[ \gamma = \sum_{\ell \in [L(k)]} \left( \lambda_k^\ell \overline{\alpha}_k^\ell - \lambda_k^{\ell-1} \alpha_k^\ell \right) \in C_k. \]
The remaining constraints follow trivially from Theorem 3. \(\Box\)

Formulation (21) contains bilinear terms $x_j \alpha_k^\ell$ and $x_j \overline{\alpha}_k^\ell$, which can further be linearized as in earlier cases. The resulting problem is then a linear MIP and can be solved by general purpose solvers. Although $|C_k|$ is at most $\frac{1}{2}L(k)(L(k) + 1)$, it may still be impractical to introduce the entire first set of constraints of Problem (21). Nevertheless, as discussed earlier, we can also introduce these constraints as cuts and solve the MIP in a Branch-and-Cut fashion.

Finally, if the T-model is also monotone, then it can further be simplified as follows.
Theorem 5. Under discrete distributions, if the T-model (20) is monotone, then it has the following formulation:

\[
\begin{align*}
\max \sum_{k \in \mathcal{K}} \sum_{\ell \in [L(k)]]} \ln (\lambda_k^\ell \overline{\alpha}_k^\ell) + & \sum_{k \in \mathcal{K}} \sum_{\ell \in [L(k)]]} \ln \left(1 - \lambda_k^{\ell-1}\right) \alpha_k^\ell \\
\text{s.t.} \sum_{j \in [N]} \left( a_{ij}^0 + \sum_{k \in \mathcal{K}} \sum_{\ell \in [L(k)]} a_{ij}^k c_k^\ell \alpha_k^\ell + \sum_{k \in \mathcal{K}} \sum_{\ell \in [L(k)]} a_{ij}^k \zeta_k^\ell \alpha_k^\ell \right) x_j & \geq b_i^0 + \sum_{k \in \mathcal{K}} \sum_{\ell \in [L(k)]} b_i^k \zeta_k^\ell \overline{\alpha}_k^\ell + \sum_{k \in \mathcal{K}} \sum_{\ell \in [L(k)]} b_i^k \zeta_k^\ell \alpha_k^\ell \forall i \in [M] \\
\sum_{\ell \in [L(k)]]} \overline{\alpha}_k^\ell = 1, \sum_{\ell \in [L(k)]]} \alpha_k^\ell = 1, & \forall k \in [K]
\end{align*}
\]

\[
\alpha_k, \overline{\alpha}_k \in \{0, 1\}^{L(k)} \forall k \in [K], \ x \in X.
\]

Proof: The formulation follows from Proposition 1 and noting that

\[
\ln \mathbb{P} \left( \tilde{z}_k \leq \sum_{\ell \in [L(k)]]} \zeta_k^\ell \alpha_k^\ell \right) = \ln \left( \sum_{\ell \in [L(k)]]} \lambda_k^\ell \overline{\alpha}_k^\ell \right) = \sum_{\ell \in [L(k)]]} \ln (\lambda_k^\ell) \overline{\alpha}_k^\ell
\]

and

\[
\ln \mathbb{P} \left( \tilde{z}_k \geq \sum_{\ell \in [L(k)]]} \zeta_k^\ell \alpha_k^\ell \right) = \ln \left( 1 - \sum_{\ell \in [L(k)]]} \lambda_k^{\ell-1} \alpha_k^\ell \right) = \sum_{\ell \in [L(k)]]} \ln \left(1 - \lambda_k^{\ell-1}\right) \alpha_k^\ell,
\]

since \( \overline{\alpha}_k^\ell \) and \( \alpha_k^\ell \) are binary variables and that

\[
\sum_{\ell \in [L(k)]]} \overline{\alpha}_k^\ell = 1, \sum_{\ell \in [L(k)]]} \alpha_k^\ell = 1,
\]

for all \( k \in [K] \). \( \square \)

3.3. Adjustable T-model

We now discuss the adjustable T-model that can be used for multi-stage decision making, in which uncertainties are unfolded in stages, and decisions are made accordingly in response to the information available at each stage. We consider a \((T + 1)\)-stage problem. For notational convenience, for any proper subset \( \mathcal{T} \subset [K] \), we denote \( \tilde{z}_T = (\tilde{z}_k)_{k \in \mathcal{T}} \) and \( z_T = (z_k)_{k \in \mathcal{T}} \). In the first stage, the decision \( x^0 \in \mathbb{R}^{N_0} \) is made before any uncertainty is realized. In the subsequent \( T \) stages, the decisions made are \( x^1(\tilde{z}_{\mathcal{T}_1}), \ldots, x^T(\tilde{z}_{\mathcal{T}_T}) \) respectively, where the recourse decision \( x^t \) at stage \( t + 1 \) is a measurable function \( x^t : \mathbb{R}^{||\mathcal{T}_t||} \rightarrow \mathbb{R}^{N_t} \) that maps from the realization of the uncertain parameters \( \tilde{z}_{\mathcal{T}_t} \) to the appropriate action in \( \mathbb{R}^{N_t} \). We assume that the subsets \( \mathcal{T}_1, \ldots, \mathcal{T}_T \) are appropriately defined so that all recourse decisions are feasible in the relevant non-anticipativity requirements. Next, define also

\[
A(z) = [A^0(z) \ A^1(z) \ldots \ A^T(z)], \ x(z) = (x^0, x^1(z_{\mathcal{T}_1}), \ldots, x^T(z_{\mathcal{T}_T}))
\]
of appropriate dimensions so that

\[ A(z)x(z) = A^0(z)x^0 + \sum_{t \in [T]} A^t(z)x^t(z_{\tau_t}). \]

We formulate the adjustable T-model as follows

\[
\begin{align*}
\max & \sum_{k \in [K]} \ln P(\alpha_k \leq z_k \leq \bar{\alpha}_k) \\
\text{s.t.} & \quad A(z)x(z) = b(z) \quad \forall z \in [\alpha, \bar{\alpha}] \\
& \quad x(z) \in \mathcal{X} \quad \forall z \in \mathcal{W} \\
& \quad x^t \in \mathcal{R}(|\mathcal{T}_t|, N_t) \quad \forall t \in [T] \\
& \quad \alpha \leq \bar{\alpha}, \ \alpha, \bar{\alpha} \in \mathcal{W},
\end{align*}
\]

(23)

where \( \mathcal{R}(m, n) \) denotes the family of all measurable functions that map from \( \mathbb{R}^m \) to \( \mathbb{R}^n \). However, solving the adjustable T-model is generally intractable because, among other difficulties, we are optimizing over arbitrary recourse functions. In fact, there are limited cases where approximations are considered in the literature. For instance, in the case of fixed recourse, i.e., \( A^t(z), t \in [T] \) are constant-valued matrices (so we write \( A^t(z) = A^t \)), a common approach in the literature of robust optimization is to restrict the recourse function \( x^t, t \in [T] \) to affine functions, i.e., \( x^t \in \mathcal{L}(|\mathcal{T}_t|, N_t) \), where

\[ \mathcal{L}(m, n) = \left\{ y \in \mathcal{R}(m, n) \mid y(\zeta) = y^0 + \sum_{k \in [m]} y^k \zeta_k, \text{for some } y^0, \ldots, y^m \in \mathbb{R}^n \right\}. \]

Hence, under fixed recourse assumptions, the affinely adjustable T-model is given by

\[
\begin{align*}
\max & \sum_{k \in [K]} \ln P(\alpha_k \leq z_k \leq \bar{\alpha}_k) \\
\text{s.t.} & \quad A^0(z)x^0 + \sum_{t \in [T]} A^t x^t(z_{\tau_t}) \geq b(z) \forall z \in [\alpha, \bar{\alpha}] \\
& \quad x(z) \in \mathcal{X} \quad \forall z \in \mathcal{W} \\
& \quad x^t \in \mathcal{L}(|\mathcal{T}_t|, N_t) \quad \forall t \in [T] \\
& \quad \alpha \leq \bar{\alpha}, \ \alpha, \bar{\alpha} \in \mathcal{W},
\end{align*}
\]

(24)

By inspection, the affinely adjustable T-model is almost identical in structure as T-model (14), except for the semi-infinite constraints \( x(z) \in \mathcal{X}, \forall z \in \mathcal{W} \), which has tractable formulations if \( \mathcal{X} \) is polyhedral with modest number of facets such as the nonnegative orthant. However, if \( \mathcal{X} \) is, for instance, a discrete set, the affine approximation of the recourse would not be useful for obtaining a tractable formulation.

Due to the difficulties of incorporating recourse functions in general, a simple approach would be to ignore them and solve T-model (14). In other words, all decisions are regarded as first-stage decisions. Incidentally, this approach turns out to be optimal if the adjustable T-model is monotone, as shown in the following result.
Theorem 6. If the adjustable T-model (23) is monotone (see Definition 2), then it is equivalent to solving the T-model (14).

Proof: Let $\hat{z}_k = \overline{\alpha}_k, \forall k \in \overline{K}$ and $\hat{z}_k = \underline{\alpha}_k, \forall k \in \underline{K}$. By inspecting the inequalities in (17), we observe that under the monotonicity condition, we have

$$A(\hat{z})x - b(\hat{z}) \leq A(z)x - b(z) \quad \forall z \in [\underline{\alpha}, \overline{\alpha}],$$

for all $x \in \mathcal{X}$. Therefore, any $(\underline{\alpha}, \overline{\alpha})$ feasible in Problem (14) is also feasible in Problem (23). It suffices to show that for any solution $(x(\cdot), \underline{\alpha}, \overline{\alpha})$ feasible in Problem (23), we can find a vector, $\hat{x} \in \mathcal{X}$ so that the solution $(\hat{x}, \underline{\alpha}, \overline{\alpha})$ is also feasible in Problem (14). Indeed, letting $\hat{x} = x(\hat{z}) \in \mathcal{X}$, we have

$$A(z)x(z) \geq b(z) \quad \forall z \in [\underline{\alpha}, \overline{\alpha}]$$

$$\Rightarrow A(\hat{z})x(\hat{z}) \geq b(\hat{z})$$

$$\Rightarrow A(\hat{z})\hat{x} \geq b(\hat{z})$$

$$\Rightarrow A(z)\hat{x} \geq b(z) \quad \forall z \in [\underline{\alpha}, \overline{\alpha}].$$

The result follows as both problems have the same objective function. □

As in the case of solving adjustable robust optimization problems, (see, e.g., Ben-Tal et al. 2004), the adjustable T-models should be implemented in a folding or rolling horizon manner. That is, after the uncertainties in each stage is revealed, the information is updated and the problem is reformulated and solved again for the remaining stages. However only the first-stage solution obtained is implemented for that stage. This process is repeated until the end of the horizon.

Notwithstanding the approximations, the T-model is computationally attractive and has shown to work remarkably well in the empirical tests of He et al. (2015). Likewise, in our computational studies, we will also demonstrate the efficacy of T-model for a two-stage optimization problem to allay fears and misconceptions that we would pay a high price for the satisficing models in terms of having very conservative and uncompetitive solutions.

4. Computational Study: A Stochastic Maximum Coverage Problem

In this section, we illustrate the advantages of the T-model by means of computational experiments on a facility location problem. Facility location problems are among the most relevant and well-studied problems in Operations Research. Accounting for the presence of uncertainty is important since facility location problems typically involve long-term and expensive capital investments that cannot be easily reversed once decisions are made. We refer to Snyder (2006) for a comprehensive review of facility location problems under uncertainty. The particular problem considered here is a stochastic maximum coverage problem (SMCP), in which a set of facilities has to be selected so that the probability that all uncertain customer demands can be met is maximized. As each facility may only serve a predefined subset of the customers, this problem is also similar to the stochastic version
of the set covering problem, which finds applications, for example, in the selection of suppliers that best provides a number of required commodities.

The SMCP is a two-stage stochastic program described as follows. In the first stage, facilities are selected; in the second stage, once the exact customer demands are observed, the demand is allocated from the selected facilities to the customers. Let $\mathcal{I}$ denote the set of candidate facility locations and $\mathcal{J}$ the set of customers. Further, let $\mathcal{I}_j \subseteq \mathcal{I}$ be the subset of facilities that may serve customer $j \in \mathcal{J}$, and, analogously, $\mathcal{J}_i \subseteq \mathcal{J}$ the subset of customers that may be served by facility $i \in \mathcal{I}$. Let $c_i$ and $f_i$ be the capacity and construction costs, respectively, for facility $i \in \mathcal{I}$, $\bar{z}_j$ the uncertain demand of customer $j \in \mathcal{J}$ and $\beta$ the total budget available for facility constructions. Furthermore, let binary variables $x_i$ be 1 if facility $i$ has been selected, and 0 otherwise. The second stage demand allocation from facility $i$ to customer $j$ is denoted by $y_{ij}(\bar{z})$, where $y_{ij}(z) : \mathbb{R}^{|\mathcal{I}|} \mapsto \mathbb{R}$ is a measurable function that depends on the realization of the uncertain customers’ demand quantities $\bar{z}$. As in §3.3, function $y_{ij}(\cdot)$ is optimized over the family of all measurable functions $\mathcal{R}(|\mathcal{J}|,1)$ that map from $\mathbb{R}^{|\mathcal{J}|}$ to $\mathbb{R}$. We attempt to solve the P-model variant of this problem, which can be stated as follows:

\[
\max \ln \mathbb{P} \left( \sum_{i \in \mathcal{I}_j} y_{ij} (\bar{z}) \geq \bar{z}_j \ \forall j \in \mathcal{J} \right)
\]

s.t. \[
\sum_{j \in \mathcal{J}_i} y_{ij}(z) \leq c_i x_i \quad \forall z \in \mathcal{W}, i \in \mathcal{I}
\]
\[
\sum_{i \in \mathcal{I}} f_i x_i \leq \beta
\]
\[
y_{ij}(z) \geq 0 \quad \forall z \in \mathcal{W}, i \in \mathcal{I}, j \in \mathcal{J}_i
\]
\[
y_{ij} \in \mathcal{R}(|\mathcal{J}|,1) \quad \forall i \in \mathcal{I}, j \in \mathcal{J}_i
\]
\[
x_i \in \{0,1\} \quad \forall i \in \mathcal{I},
\]

where the objective function maximizes the probability that the uncertain demand can be met. The first set of constraints are the facility capacity constraints, while the second constraint restricts the construction costs to the available budget. Solving such a two-stage stochastic problem exactly is computationally intractable. In the following numerical studies, we obtain satisficing solutions to the SMCP based on the T-model, and compare the solution performances to the P-model solutions. Below, we first outline the different models and discuss the problem instances and computational settings used in the following experiments.

**T-models for the SMCP.** We evaluate two implementations of the T-model: first, the case of log-concave probability densities (see §3.1); second, the case of discrete probability distributions (see §3.2). In the SMCP considered (see Problem (25) above), the uncertainties appear only on the right hand side of the constraint inequalities. That is, $\sum_{i \in \mathcal{I}_j} y_{ij} \geq \bar{z}_j \ \forall j \in \mathcal{J}$. Furthermore, the
uncertain customer demands $\tilde{z}_j$ for all $j \in J$ are independently distributed by assumption. The problem therefore satisfies the conditions of a monotone T-model (see Definition 2). For the case of log-concave densities, we formulate our model on the results of Proposition 1 and Theorem 6:

$$(\text{Model T-1}) \quad \max \sum_{j \in J} \ln P(\tilde{z}_j \leq \alpha_j)$$

subject to:

- $\sum_{i \in I} y_{ij} \geq \alpha_j \quad \forall j \in J$
- $\sum_{i \in I} y_{ij} \leq c_i x_i \quad \forall i \in I$
- $\sum_{i \in I} f_i x_i \leq \beta$
- $y_{ij} \geq 0 \quad \forall i \in I, j \in J$
- $x_i \in \{0, 1\} \quad \forall i \in I$
- $\alpha_j \in W.$

The problem can be solved via Branch-and-Cut, where the density curve cuts (18) (as defined in §3.1) are added whenever violated in the branch-and-bound tree.

For the T-model based on discrete distributions, its application for the SMCP can be motivated as follows. We assume that, for each customer $j \in J$, a random sample of size $L$ demand outcomes is available. These sample outcomes, defined as $\zeta^\ell_j$, are sorted in non-decreasing order over $\ell \in [L]$ for each customer $j \in J$. This is used to define the support $W_j$ for the random customer demand $\tilde{z}_j$. Based on the formulation (22) for monotone T-models in Theorem 5, the resulting T-model for the SMCP with discrete distributions is then given by:

$$(\text{Model T-2}) \quad \max \sum_{j \in J} \sum_{\ell \in [L]} \ln \left( \frac{\ell}{L} \right) \hat{\alpha}_j^\ell$$

subject to:

- $\sum_{i \in I} y_{ij} \geq \sum_{\ell \in [L]} \zeta^\ell_j \hat{\alpha}_j^\ell \quad \forall j \in J$
- $\sum_{j \in J} y_{ij} \leq c_i x_i \quad \forall i \in I$
- $\sum_{i \in I} f_i x_i \leq \beta$
- $\sum_{\ell \in [L]} \hat{\alpha}_j^\ell = 1 \quad \forall j \in J$
- $y_{ij} \geq 0 \quad \forall i \in I, j \in J$
- $x_i \in \{0, 1\} \quad \forall i \in I$
- $\hat{\alpha}_j^\ell \in \{0, 1\} \quad \forall \ell \in [L], j \in J.$

**Benchmarks.** We consider two sample average approximation (SAA) optimization models, that assume as input a sample of $L$ scenarios of the demand observations. The first corresponds to the P-model objective as in Problem (25), maximizing the number of feasible scenarios in the given sample. The model involves one demand satisfaction constraint for each scenario $\ell \in [L]$, using the
specific demand realizations \( \hat{\zeta}_j^\ell \) for each customer \( j \in \mathcal{J} \). Note that, in contrast to the T-model, \( \hat{\zeta}_j^\ell \) may not necessarily be nondecreasing over \( \ell \in [L] \). Binary variables \( p^\ell \) take value 1 if demand scenario \( \ell \in [L] \) is covered by the set of selected facilities, and 0 otherwise. The model can be stated as:

\[
\text{(Model P-1)} \quad \max \frac{1}{L} \sum_{\ell \in [L]} p^\ell
\]

\[
\text{s.t.} \quad \sum_{i \in \mathcal{I}_j} y_{ij}^\ell \geq \hat{\zeta}_j^\ell - \Theta_\ell (1 - p^\ell) \quad \forall j \in \mathcal{J}, \ell \in [L]
\]

\[
\sum_{j \in \mathcal{J}_i} y_{ij}^\ell \leq c_i x_i \quad \forall i \in \mathcal{I}, \ell \in [L]
\]

\[
\sum_{i \in \mathcal{I}} f_i x_i \leq \beta
\]

\[
\sum_{j \in \mathcal{J}} y_{ij}^\ell \leq c_j x_j \quad \forall i \in \mathcal{I}, \ell \in [L]
\]

\[
x_i \in \{0, 1\} \quad \forall i \in \mathcal{I}
\]

\[
p^\ell \in \{0, 1\} \quad \forall \ell \in [L],
\]

where \( \Theta_\ell \) are constants assumed to be of sufficiently large value. In particular, we set \( \Theta_\ell \) to their tightest values: \( \Theta_\ell = \max_{j \in \mathcal{J}} \left\{ \hat{\zeta}_j^\ell \right\} \). Conceivably, due to the additional binary variables \( p^\ell \), the model will be difficult to solve when the number of scenarios \( L \) is large. Nevertheless, for reasonably sized instances, Model P-1 can be implemented and solved directly using general-purpose MIP solvers.

The second SAA model used for comparison aims at minimizing the expected demand shortfall. Let \( s_j^\ell \) be the demand shortfall for the demand of customer \( j \) at scenario \( \ell \). The problem is formulated as:

\[
\text{(Model E)} \quad \min \frac{1}{L} \sum_{\ell \in [L]} \sum_{j \in \mathcal{J}} s_j^\ell
\]

\[
\text{s.t.} \quad \sum_{i \in \mathcal{I}} y_{ij}^\ell + s_j^\ell \geq \hat{\zeta}_j^\ell \quad \forall j \in \mathcal{J}, \ell \in [L]
\]

\[
\sum_{j \in \mathcal{J}_i} y_{ij}^\ell \leq c_i x_i \quad \forall i \in \mathcal{I}, \ell \in [L]
\]

\[
\sum_{i \in \mathcal{I}} f_i x_i \leq \beta
\]

\[
\sum_{j \in \mathcal{J}} y_{ij}^\ell \leq c_j x_j \quad \forall i \in \mathcal{I}, \ell \in [L]
\]

\[
y_{ij}^\ell \geq 0 \quad \forall i \in \mathcal{I}, j \in \mathcal{J}, \ell \in [L]
\]

\[
s_j^\ell \geq 0 \quad \forall j \in \mathcal{J}, \ell \in [L]
\]

\[
x_i \in \{0, 1\} \quad \forall i \in \mathcal{I}
\]

Like Model P-1, Model E can be implemented and solved directly using general purpose MIP solvers. Note that a computational advantage of Model E over Model P-1 is that it avoids the use of additional binary variables for modeling the objective function.
Problem Instances and Computational Settings. We generate the problem instances for the SMCP as follows. The number of customers $|J|$ is taken from $|J| \in \{100, 250, 500, 1000, 2000\}$; the number of candidate facility locations is chosen accordingly from $|I| \in \{0.5|J|, |J|, 2|J|\}$. The arcs between customers and facilities have been generated randomly such that each customer is connected to $A\%$ of the $|I|$ facilities, where the parameter $A \in \{20, 30, 40\}$ can be regarded as a surrogate for the level of facility-customer connectivity assumed in the problem. We assume that the uncertain customer demands $\tilde{z}_j$, $j \in J$ follow a continuous normal distribution with mean $\mu_j$ and standard deviation $\sigma_j = 0.5\mu_j$, where $\mu_j$ is chosen randomly according to a uniform distribution between 1 and 100. If the random outcome of a customer demand has been smaller than 0, the customer is assumed to not have any demand in that particular scenario. Facility capacities are set sequentially to 500, 750 and 1,000 units, i.e., $c_{3i'} = 500$, $c_{3i'+1} = 750$ and $c_{3i'+2} = 1,000$ for the running index $i'$. As the average customer demand is about 50 units, a total construction budget of $50 \times |J|$ would, on average, ensure that all customer demands can be met, if each customer could be served by any facility (i.e., $A = 100$). However, as demands are uncertain and the number of service arcs between facilities and customers are restricted, we allow for an additional budget of $B\%$, i.e., we set $\beta = (1 + B/100) \times 50 \times |J|$. If not otherwise stated, we set $B = 5$.

All mathematical models have been implemented in C/C++ using the IBM CPLEX 12.6.1 Callable Library. The code has been compiled and executed on openSUSE 11.3. Each problem instance has been run on a single Intel Xeon X5650 processor (2.67GHz), limited to 24GB of RAM and a maximum of 12 hours computing time.

4.1. Performance Study

The performance of the T-models is now evaluated and compared to those of Model P-1 and Model E. Model T-1 is solved using a Branch-and-Cut approach, implemented with user cuts in CPLEX. We initialize the model with 20 density curve cuts for the demand distribution function of each customer $j \in J$, distributed uniformly on the interval $[-3.1\mu_j, 3.1\mu_j]$. We evaluate the performance of the algorithms by measuring the objective of the P-model in (25), termed in the following as the success rate, as well as the expected demand shortfall, which is often of interest in the literature. These performance measures are computed using Monte Carlo simulation on 100,000 independent random demand samples.

Table 1 summarizes the results for Model T-1. Each row corresponds to the problem instances with $|J|$ customers and therefore represents the results for 12 instances each (3 different numbers of candidate facility locations $\times$ 4 different levels of connectivity $A$). We report the average success rate and demand shortfall (in units of 10), the average computing time (in minutes), as well as the number of instances for which the computation exceeded the time limit (and returned the best
| $|\mathcal{F}|$ | success rate % | demand shortfall | time (minutes) | # cut iter. | # cuts |
|---|---|---|---|---|---|
| 100 | 84.41 | 2.0 | 22.7 | 98.0 | 2,742.9 |
| 250 | 82.76 | 4.2 | 2.7 | 83.5 | 4,283.3 |
| 500 | 99.93 | 0.0 | 3.4 | 38.2 | 4,190.1 |
| 1000 | 96.60 | 1.2 | 30.7 | 16.8 | 5,672.4 |
| 2000 | 95.97 | 2.2 | 426.9 | 14.9 | 8,926.0 |
| all | 92.06 | 1.9 | 98.5 | 49.5 | 5203.9 |

Table 1  Stochastic optimization performance study of the T-1 on all instances, reporting the average success rate (%), average demand shortfall (in 10 units), average computing time (in minutes), as well as the average number of cut iterations performed and cuts added.

We also report the average number of cut callbacks that verify whether further density curve cuts should be added to the model, and the average number of cuts actually added. Model T-1 has found feasible solutions for all problem instances and obtained high average success rates and low demand shortfalls. The approach seems quite scalable: in all the numerical experiments the time limit of 12 hours has been exceeded for only 3 large problem instances, yet the solutions found for those instances were of high quality. As expected, the average number of cuts added to the models increases with the number of random customer demands. The number of cut iterations necessary to prove optimality, however, decreases as the number of customers increases.

An analysis of the results showed that on small instances CPLEX tends to improve its solution gradually, each time resulting in a series of new cuts, while for large instances less solutions were explored before finding the best solution.

For Model T-2, which is based on the data sample of customer demands, the solver usually finds high quality solutions with small optimality gaps after short computing times and then spends large amounts of time trying to close that gap. We therefore terminate the optimization process as soon as the solver proves optimality within 0.5%.

<table>
<thead>
<tr>
<th>$L$</th>
<th>Model T-2</th>
<th>Model P-1</th>
<th>Model E</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>succ. rate %</td>
<td>short fall ns</td>
<td>succ. rate %</td>
</tr>
<tr>
<td>5</td>
<td>88.73</td>
<td>350.9</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>85.39</td>
<td>699.8</td>
<td>62.5</td>
</tr>
<tr>
<td>15</td>
<td>88.41</td>
<td>351.0</td>
<td>75.3</td>
</tr>
<tr>
<td>50</td>
<td>86.64</td>
<td>525.4</td>
<td>54.9</td>
</tr>
</tbody>
</table>

Table 2  Performance comparison for different sample sizes $L$ on all 60 instances: success rate, average demand shortfall (in 10 units), average computing time (in minutes) and number of instances where each method has not found a feasible solution “# ns”. As $L$ increases, Models P-1 and E find feasible solutions for less instances, decreasing the average success rate and increasing the demand shortfall. In contrast, Model T-2 remains relatively stable.

Table 2 compares the performance of Model T-2 and the SAA approaches (i.e., Model P-1 and Model E) based on a single run for different sample sizes $L$ from 5 to 50 randomly generated
scenarios. For large $L$ some approaches may either run out of time or out of memory. It is therefore not guaranteed that a particular solution approach returns a feasible solution after 12 hours computing time. In our experiments, Model T-2 always found feasible solutions for all instances; the SAA approaches did not. As $L$ increases, the SAA models become more difficult to solve and do not find feasible solutions for many instances (column “# ns”). This directly impacts the average success rates and the shortfall, which deteriorate as less instances are solved. In contrast, Model T-2 provides consistently high success rates and seems to be quite scalable in terms of number of scenarios, finding feasible solutions for at least 56 out of the 60 problem instances. Its average computing times remain significantly low, which suggests that one may solve the T-2 with much larger samples.

4.2. Robustness Study Via Multiple Replications

We next study the sample-to-sample solution performance variation of the models considered. 10 independent replications of the experiments were performed, each with a sample of 5 randomly-generated demand scenarios used for solving the Models T-2, P-1 and E. Table 3 shows the average success rates and demand shortfalls, as well as their minimum and maximum average values. Instead of reporting the average over all instances, the presented results only take into account those instances for which feasible solutions were found. For Model T-1, we report only the average results, since the sample-to-sample variation, caused by the different number of instances considered in each replication, is negligible.

| $|J|$ | succ. rate | short % fall | Model T-2 | succ. rate | short % fall | Model P-1 | succ. rate | short % fall | Model E | succ. rate | short % fall | # ns |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 100 | 10.4 | 85.64 | 0.3 | 89.67 | 2.1 | 0.2 | 85.03 | 1.9 | 0.1 | 83.03 | 0.0 | | |
| [9, 11] | [70.2, 87.3] | [1.7, 3.4] | [0, 2] | [80.4, 87.3] | [1.7, 2.9] | [0, 1] | [72.2, 88.0] | [1.6, 3.4] | [0, 1] |
| 250 | 12 | 82.76 | 0.7 | 80.44 | 4.4 | 0.0 | 81.11 | 4.7 | 0.0 | 82.51 | 4.2 | 0.0 | |
| [12, 12] | [75.4, 82.7] | [4.2, 4.9] | [0, 0] | [69.1, 82.8] | [4.2, 9.1] | [0, 0] | [82.3, 82.8] | [4.2, 4.3] | [0, 0] |
| 500 | 12 | 99.93 | 0.0 | 98.29 | 0.2 | 0.0 | 96.66 | 1.0 | 0.0 | 95.69 | 1.4 | 0.0 | |
| [12, 12] | [91.8, 99.9] | [0.0, 0.7] | [0, 0] | [91.7, 99.9] | [0.0, 2.7] | [0, 0] | [86.9, 99.9] | [0.0, 4.6] | [0, 0] |
| 1000 | 6.4 | 96.60 | 0.2 | 96.58 | 1.2 | 0.0 | 89.65 | 4.9 | 5.6 | 90.52 | 4.4 | 2.5 | |
| [3, 8] | [96.5, 96.7] | [1.2, 1.2] | [0, 0] | [82.6, 95.1] | [1.8, 9.0] | [4, 9] | [80.8, 96.7] | [1.2, 9.9] | [1, 4] |
| 2000 | 0.1 | 94.62 | 0.5 | 94.51 | 3.0 | 2.6 | 91.96 | 4.7 | 11.9 | 88.69 | 7.1 | 8.4 | |
| [0, 1] | [94.5, 94.5] | [3.0, 3.0] | [1, 4] | [92.9, 92.0] | [4.7, 4.7] | [11.12] | [88.7, 88.7] | [7.1, 7.1] | [8, 9] |
| all | 40.9 | 90.70 | 0.3 | 88.77 | 2.1 | 2.8 | 88.05 | 2.9 | 17.6 | 87.67 | 2.9 | 11.3 | |
| [37, 43] | [86.0, 91.3] | [1.9, 2.4] | [1, 4] | [82.2, 89.8] | [2.3, 5.4] | [16, 21] | [84.9, 90.6] | [1.9, 4.5] | [9, 13] |

Table 3 Performance comparison of average [minimum, maximum] results among 10 replications with sample size $L = 5$ for instances for which all methods found feasible solutions: success rate, average demand shortfall (in 10 units) and number of instances for which each method has not found a feasible solution “# ns”.

For all the 10 replications, the number of instances for which the three approaches found feasible solutions was relatively stable, between 37 and 43. Model T-1 clearly outperforms the other
approaches with respect to average success rates and average demand shortfall. A further analysis has shown that, on average, it provides a strictly better success rate for almost half of the 60 instances, while it performs worse on less than two instances on average. Model T-2 finds feasible solutions for almost all instances throughout the 10 replications. The average success rates and demand shortfalls of Model P-1 and Model E are slightly better than those of the T-2 for small problem instances. For medium and large instances, Model T-2 performs better.

Even though the average success rates are not significantly different across the four solution approaches, the T-models provide a higher success rate for most of the instances. Among the 10 replications and on the same set of instances for which all methods have found a feasible solution, Model T-1 has provided a better coverage than Model P-1 for 29.1 instances on average, while it has been worse for only 1.4 instances. Compared to Model E, these averages have been 26.6 and 1.6 instances, respectively. Model T-2 achieves a higher success rate than the Mode P-1 for 16.7 instances and is worse-off for 3.3. Compared to Model E, it has been better for 14.3 instances and worse for 3.8 instances. Further, it is worthwhile to note that Model T-2 is remarkably fast in solution speed when compared to the other SAA models. On average, it solved a problem in 0.7 minutes, about 40 times faster than Model P-1 (on average 28.6 minutes) and Model E (on average 33.1 minutes).

4.3. Scalability of the Model T-2

We now investigate the performance of the Model T-2 with respect to the problem size and sample size, under multiple replications. Table 4 shows the average, minimum and maximum of the average success rates of the different models. For Models P-1 and E, only results for sample size \( L = 5 \) is reported, as these models were unable to handle large sample sizes in our computational studies (see Table 2). The T-2 results are shown for different sample sizes, from 5 to 1000. The average success rates and sample-to-sample variation tends to decrease with higher sample sizes. With \( L = 1,000 \) scenarios, all replications yield the same average success rates. In fact, further analysis showed that each instance returns the same success rates in each of the 10 replications.

The average computing times of Model T-2 across all instances are 57, 58, 53, 35, 36 and 82 minutes when using 5, 15, 50, 100, 500 and 1,000 demand scenarios, respectively. The results hence strongly suggest that Model T-2 is highly scalable in the size of the sample used.

4.4. Effect of Investment Budget

We now repeat the computational studies by varying the investment budget parameter \( \beta \) in Problem (25). This type of analysis can be useful for decision-makers in calibrating and justifying investment budget requests. Figure 1 shows the average success rates, as well as the intervals spanned by the
minimum and maximum success rates across the 10 optimization runs for a problem instance with 500 facilities, 1000 customers and $A = 30$. In each run, a sample of size $L = 5$ demand scenarios was generated and used for Models T-2, P-1 and E. The investment budget $\beta$ is varied by introducing increments $B$ of up to 10% of the baseline value used. Throughout all replications, the average computing times for Models P-1, E, and T-2 were 97.5, 25.5 and 0.2 minutes, respectively. Even though the average solution times are far below the time limit of 12 hours, it is to be noted that in several runs, Model P-1 exceeded the time limit without finding a feasible solution.

![Figure 1](image-url) Success rate for Models T-2, P-1 and E. The graph illustrates the average success rate and the interval between the minimum and maximum success rates across 10 replications (using sample size $L = 5$) for the problem instance with 500 facilities and 1,000 customers.
In the solutions for all three models, the success rates mostly improve within the first 5% of the budget increase. Beyond that, the improvements diminish rapidly. The large performance intervals depicted in Figure 1 also indicate that Model P-1 and Model E are quite unstable in terms of their success rates, having huge sample-to-sample variation in the performance. In contrast, Model T-2 was remarkably stable, finding solutions with highest success rates in each of the replications, independent of the provided demand scenarios. Furthermore, it is interesting to note that higher levels of investment budgets are required for Models P-1 and E to achieve the same success rate as Model T-2. This is especially pronounced when higher average success rates are desired. Figure 2 illustrates the same analysis for the average demand shortfall. Similarly, Model T-2 presents the lowest demand shortfalls, while the shortfalls for the other two models, in particular Model P-1, appears significantly more unstable. In conclusion, the results presented in both Figures 1 and 2 strongly suggest that the T-model solutions dominate those of Model P-1 and Model E in terms of average success rate, average expected shortfall and sample-to-sample variation across a large number of investment budget levels.

**Figure 2** Demand shortfalls for Models T2, P-1 and E. The graph illustrates the average demand shortfall and the interval between the minimum and maximum demand shortfall across 10 replications (using sample size \( L = 5 \)) for the problem instance with 500 facilities and 1,000 customers.

**Summary of Results**

We summarize the results of the computational studies on the Stochastic Maximum Coverage Problem, comparing the two T-models (T-1 and T-2) with Model P-1, based on maximizing probability of meeting demands, and Model E, based on minimizing expected demand shortfalls, as follows:
- Model T-1 finds high quality solutions for all instances with an average success rate of 92.0% (see Table 1). The average rates of the Model T-2, Model P-1 and Model E are 88.7%, 64.6% and 70.1%, respectively, with \( L = 5 \) samples. When up to \( L = 50 \) samples are used, the rates of the T-2 remain stable, whereas those of Model P-1 and Model E drop to 20.4% and 26.5%, respectively. Computing times for Model T-2 are about 5-10 times lower than those required by Models P-1 and E.

- When experiments are replicated (with \( L = 5 \)), and results are compared only on those instances for which all methods found feasible solutions, Model T-2 tends to provide slightly better success rates and less variation than the SAA models (see Table 3).

- Model T-2 scales well to a sample size of \( L = 1000 \). The more samples are used, the higher tends to be the average success rates and the lower the variation among different replications of the experiments. With 1,000 samples, all instances yield the same success rate among all 10 replications performed.

- For the same available budget and the same computing resources, the T-models provide a higher average success rate with minimal variation (see Figures 1 and 2).

5. Concluding Remarks

We have proposed a new generalization of satisficing decision criteria that comes from two key properties of satisficing preferences, and which appears in various mathematical optimization problems under uncertainty. In relation to the representation of the satisficing decision criterion, we have introduced mathematical modeling frameworks for optimization under uncertainty, encompassing broad classes of problems such as probability maximization problems, robust optimization problems, and chance-constrained optimization problems. In particular, we have proposed the S-model, where satisficing is the objective, and also satisficing-constrained optimization problems, where satisficing performance is enforced as a constraint.

We have next proposed a class of tractable probabilistic S-models, termed the T-models, based on “box”-type uncertainty sets, which can be regarded as safe and tractable approximations of the P-model of Charnes and Cooper (1963). We then developed different implementations of the T-model, and showed that in the cases of log-concave probability densities, the T-model objective is tractable and concave. In the case of discrete probability distributions, we showed that the T-model can be reformulated as a moderate-sized MIP. We further considered T-models in multi-stage problems, and showed that if the T-models are monotone, the resulting problem can be simplified significantly.

Finally, we have presented extensive computational studies on a two-stage stochastic maximum demand coverage problem, for which we have shown that the T-models consistently demonstrate
good solution quality in both success rate and expected demand shortfalls, as compared to SAA models that maximize probability and minimize expected shortfalls. Hence, although the T-models are lower bounds of the P-model, they are not at all conservative across different performance measures. Furthermore, we have shown that the T-models are highly scalable in problem size, and highly efficient in computational speed, as compared to the other models.

Acknowledgments  This research was partially supported by the National Research Foundation (NRF) Singapore through the Singapore-MIT Alliance for Research and Technology (SMART) Center for Future Mobility (FM). The authors would like to thank Calcul Québec and Compute Canada for providing the computing infrastructure used for the experiments.

References


