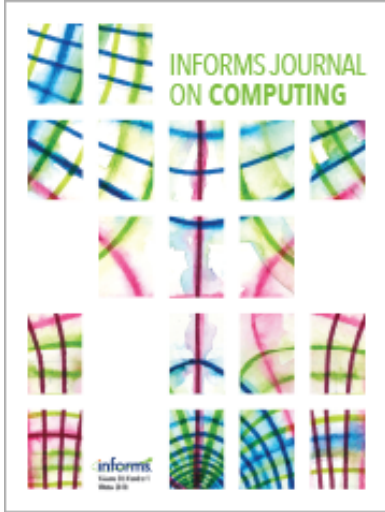




INFORMS Journal on Computing

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

On the Value of Risk-Averse Multistage Stochastic Programming in Capacity Planning

Xian Yu, Siqian Shen

To cite this article:

Xian Yu, Siqian Shen (2024) On the Value of Risk-Averse Multistage Stochastic Programming in Capacity Planning. INFORMS Journal on Computing

Published online in Articles in Advance 23 Oct 2024

. <https://doi.org/10.1287/ijoc.2023.0396>

Full terms and conditions of use: <https://pubsonline.informs.org/Publications/Librarians-Portal/PubsOnLine-Terms-and-Conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2024, INFORMS

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes. For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

On the Value of Risk-Averse Multistage Stochastic Programming in Capacity Planning

Xian Yu,^{a,*} Siqian Shen^b
^aDepartment of Integrated Systems Engineering, The Ohio State University, Columbus, Ohio 43210; ^bDepartment of Industrial and Operations Engineering, University of Michigan, Ann Arbor, Michigan 48109

*Corresponding author

Contact: yu.3610@osu.edu,  <https://orcid.org/0000-0003-1059-5303> (XY); siqian@umich.edu,  <https://orcid.org/0000-0002-2854-163X> (SS)

Received: November 4, 2023

Revised: November 8, 2023; April 20, 2024;
September 7, 2024; September 25, 2024

Accepted: September 28, 2024

Published Online in Articles in Advance:
October 23, 2024

<https://doi.org/10.1287/ijoc.2023.0396>

Copyright: © 2024 INFORMS

Abstract. We consider a risk-averse stochastic capacity planning problem under uncertain demand in each period. Using a scenario tree representation of the uncertainty, we formulate a multistage stochastic integer program to adjust the capacity expansion plan dynamically as more information on the uncertainty is revealed. Specifically, in each stage, a decision maker optimizes capacity acquisition and resource allocation to minimize certain risk measures of maintenance and operational cost. We compare it with a two-stage approach that determines the capacity acquisition for all the periods up front. Using expected conditional risk measures, we derive a tight lower bound and an upper bound for the gaps between the optimal objective values of risk-averse multistage models and their two-stage counterparts. Based on these derived bounds, we present general guidelines on when to solve risk-averse two-stage or multistage models. Furthermore, we propose approximation algorithms to solve the two models more efficiently, which are asymptotically optimal under an expanding market assumption. We conduct numerical studies using randomly generated and real-world instances with diverse sizes, to demonstrate the tightness of the analytical bounds and efficacy of the approximation algorithms. We find that the gaps between risk-averse multistage and two-stage models increase as the variability of the uncertain parameters increases and decrease as the decision maker becomes more risk averse. Moreover, a stagewise-dependent scenario tree attains much higher gaps than a stagewise-independent counterpart, whereas the latter produces tighter analytical bounds.

History: Accepted by Andrea Lodi, Area Editor for Design & Analysis of Algorithms–Discrete.

Funding: This work of Dr. X. Yu was partially supported by the U.S. National Science Foundation Division of Information and Intelligent Systems [Grant 2331782].

Supplemental Material: The software that supports the findings of this study is available within the paper and its Supplemental Information (<https://pubsonline.informs.org/doi/suppl/10.1287/ijoc.2023.0396>) as well as from the IJOC GitHub software repository (<https://github.com/INFORMSJoC/2023.0396>). The complete IJOC Software and Data Repository is available at <https://informsjoc.github.io/>.

Keywords: multistage stochastic integer programming • risk-averse optimization • dynamic time-consistent risk measure • approximation algorithms • capacity planning

1. Introduction

The capacity planning problem, aiming to determine the optimal level of capacity acquisition and allocation, is a classic optimization problem solved in a broad spectrum of applications, including capacity expansion in supply chains (Aghezzaf 2005), production capacity planning for semiconductor manufacturing (Swaminathan 2000, Hood et al. 2003), electric vehicle (EV) charging station expansion (Bayram et al. 2015), and so on. In these applications, demand fluctuates spatially and temporally, and as the cost of resource allocation in each stage depends on the demand input, it is also random depending on the distribution of demand over time. To adapt to new demand, service providers need to expand the existing capacities in an efficient way. Therefore, estimating and utilizing uncertain demand in the decision processes of capacity planning is crucial for cost reduction and quality-of-service improvement.

In this paper, we focus on a finite time horizon, in which the uncertain demand in each period is modeled by a random vector. A decision maker optimizes when and how to expand the capacity and how to allocate resources to meet the demand to minimize a certain risk measure of the maintenance and operational cost over multiple periods. We compare two modeling frameworks: two-stage (TS) stochastic programming and multistage (MS) stochastic programming. In the two-stage model, we determine capacity acquisition for all time periods up front as here-and-now decisions and decide optimal resource allocation as wait-and-see recourse decisions based on

realized demand and acquired capacities. In the multistage model, the uncertain demand values are revealed gradually over time and the capacity-acquisition decisions are adapted to this process.

Both two-stage and multistage decision frameworks are commonly used in the stochastic capacity planning literature. We refer interested readers to Luss (1982), Davis et al. (1987), and Sabet et al. (2020) for comprehensive reviews on capacity planning problems under uncertainty. To handle general parameter uncertainty, robust optimization and stochastic programming are two of the popular approaches. Paraskevopoulos et al. (1991) and Aghezzaf (2005) discuss robust capacity planning problems, and Eppen et al. (1989), Fine and Freund (1990), and Berman et al. (1994) consider stochastic programming approaches using scenarios to model the uncertain parameters. Most of these approaches are based on the two-stage decision-making framework. There is a limited number of papers considering multistage stochastic capacity planning problems. Both Ahmed and Sahinidis (2003) and Singh et al. (2009) formulate a multistage stochastic mixed-integer programming model for capacity-planning problems, for which Ahmed and Sahinidis (2003) propose a linear programming (LP) relaxation-based approximation scheme with asymptotic optimality and Singh et al. (2009) apply variable splitting and Dantzig–Wolfe decomposition to tackle the problem.

Stochastic programming approaches mainly focus on minimizing the total cost on average. However, minimizing the expected cost does not necessarily avoid the rare occurrences of undesirably high cost, and in a situation in which it is important to maintain reliable performance, we aim to evaluate and control the risk. In particular, coherent risk measures (Artzner et al. 1999) have been used in many risk-averse stochastic programs as they satisfy several natural and desirable properties. Schultz and Tiedemann (2006), Shapiro et al. (2009), Ahmed (2006), and Miller and Ruszczyński (2011) extend two-stage stochastic programs with risk-neutral, expectation-based objective functions to risk-averse ones. It becomes nontrivial to model multistage risk-averse stochastic programs because the risk can be measured based on the cumulative cost, a total sum of the individual risk from each stage, or the conditional risk in a nested way. When the risk is calculated in a nested way via dynamic risk measures, a desirable property is called time consistency (Ruszczyński 2010), which ensures consistent risk preferences over time.

Pflug and Ruszczyński (2005) propose a class of multiperiod risk measures, which can be calculated by solving a stochastic dynamic linear optimization problem, and they analyze its convexity and duality structure. Homem-de Mello and Pagnoncelli (2016) extend the above risk measures to expected conditional risk measures (ECRMs) and prove some appealing properties. First, ECRMs, originally defined for each stage separately, can be rewritten in a nested form. Second, any risk-averse multistage stochastic programs with ECRMs using a conditional value at risk (CVaR) measure can be recast as a risk-neutral multistage stochastic program with additional variables and constraints that can be efficiently solved using existing decomposition algorithms. Third, ECRMs are time consistent following the inherited optimality property proposed in Homem-de Mello and Pagnoncelli (2016). In this paper, we show that ECRMs also follow the definition of time consistency proposed in Ruszczyński (2010) in Online Appendix A. Because of these properties, we base our analysis on the dynamic time-consistent ECRMs using the CVaR measure, but we note that all the analysis and results derived in this paper also can be applied to other CVaR-based risk measures, such as static CVaR or period-wise CVaR, which are not time consistent.

A natural question is then about the performance of the risk-averse two-stage and multistage stochastic programs. Specifically, we are interested in comparing the optimal objective values and computational effort between solving risk-averse two-stage and multistage models for stochastic capacity planning problems. Noting that the multistage models have larger feasible regions and, thus, will always have lower costs, we aim to bound the gap between optimal objective values of the two-stage and multistage models given specific risk measures of the cost and characteristics of the uncertainty. Huang and Ahmed (2009) are the first to show analytical lower bounds for the value of multistage stochastic programming (VMS) compared with the two-stage approach for capacity planning problems with an expectation-based objective function. They developed an asymptotically optimal approximation algorithm (AA) by exploiting a decomposable substructure in the problem. Our work differs from Huang and Ahmed (2009) in the following aspects. First, instead of using expectation as the objective function, we consider a dynamic time-consistent risk measure (i.e., ECRMs). Although the resulting risk-averse problem can be reformulated as a risk-neutral counterpart, the risk-neutral results derived from Huang and Ahmed (2009) cannot be directly applied here because of a new structure induced by ECRM-related decision variables and constraints. Instead, we derive new bounds on the VMS under this risk-averse setting. Second, we propose both lower and upper bounds to the gaps between ECRM-based optimal objective values of two-stage and multistage models and provide an example to illustrate the tightness of the lower bound, whereas Huang and Ahmed (2009) only provided lower bounds without a tightness guarantee. Third, our approximation algorithms improve the one developed in Huang and Ahmed (2009) by deriving closed-form solutions of the

substructure problem in the risk-averse setting and strengthening the upper bound iteratively, whereas the one in Huang and Ahmed (2009) relies on solving the substructure problem in each time. We further prove that the approximation schemes are asymptotically optimal when the demand increases over time.

There is rich literature on proposing decomposition algorithms to solve risk-neutral and risk-averse multistage stochastic programs. Pereira and Pinto (1991) are the first to develop stochastic dual dynamic programming (SDDP) algorithms for efficiently computing multistage stochastic linear programs under stagewise-independent scenario trees, which construct an under-approximation for the value function in each stage iteratively. We refer interested readers to Philpott and Guan (2008), Girardeau et al. (2014), Shapiro (2011), Shapiro et al. (2013), and Guigues (2016) for studies on the convergence of the SDDP algorithm under different problem settings. Among them, Shapiro et al. (2013) and Guigues (2016) extend the SDDP algorithm to solve a risk-averse multistage stochastic program. Stochastic dual dynamic integer programming (SDDiP), first proposed by Zou et al. (2019), is an extension of SDDP to handle the nonconvexity arising in multistage stochastic integer programs based on a similar computational framework. The main goal of this paper is to bound the gap between two-stage and multistage stochastic programs and provide general guidelines on model choices. Using the substructure explored when deriving the bounds, we also propose approximation algorithms to solve risk-averse two-stage and multistage stochastic capacity planning problems more efficiently. Note that a common assumption made in the aforementioned SDDP-type algorithms is that the underlying stochastic process is stagewise independent. However, our analysis and proposed approximation algorithms are not restricted by this stagewise independence assumption. Whereas our analytical results can be applied to a more general stochastic process, using a stage-wise independent scenario tree, we compare the performance of our approximation algorithm with the SDDiP algorithm in Section 5.1.5.

The remainder of the paper is organized as follows. In Section 2, we present the formulations of risk-averse two-stage and multistage capacity planning problems under ECRM. In Section 3, we derive two lower bounds and an upper bound for the gaps between the optimal ECRM-based objective values of two-stage and multistage stochastic capacity planning models. In Section 4, we propose approximation algorithms with performance guarantees. In Section 5, we conduct numerical studies using instances with diverse uncertainty patterns to demonstrate the tightness of our bounds and the performance of the approximation algorithms. Section 6 concludes the paper and states future research directions. Throughout the paper, we use bold symbols to denote vectors/matrices and use $[n]$ to denote the set $\{1, 2, \dots, n\}$.

2. Problem Formulations

Consider a general class of capacity planning problems, in which we have $1, \dots, T$ time periods; $1, \dots, M$ facilities; and $1, \dots, N$ customer sites. Let c_{tij} be the cost of serving one unit of demand from customer site j in facility i at period t ; f_{ti} be the cost of maintaining one unit of resources in facility i at period t ; h_{ti} be the number of customers that one unit of resources can serve in facility i at period t ; and d_{tj} be the demand at customer site j at period t for all $i \in [M]$, $j \in [N]$, and $t \in [T]$.

Define decision variables $x_{ti} \in \mathbb{Z}_+$ as the units of resources we invest in facility i at period t , and decision variables $y_{tij} \in \mathbb{R}_+$ as the amount of demand from customer site j served by facility i at period t . The deterministic capacity planning problem is given by

$$\min \sum_{t=1}^T \sum_{i=1}^M f_{ti} \sum_{\tau=1}^t x_{\tau i} + \sum_{t=1}^T \sum_{i=1}^M \sum_{j=1}^N c_{tij} y_{tij} \quad (1a)$$

$$\text{s.t.} \quad \sum_{i=1}^M y_{tij} = d_{tj}, \quad \forall j = 1, \dots, N, t = 1, \dots, T, \quad (1b)$$

$$\sum_{j=1}^N y_{tij} \leq h_{ti} \sum_{\tau=1}^t x_{\tau i}, \quad \forall i = 1, \dots, M, t = 1, \dots, T, \quad (1c)$$

$$x_{ti} \in \mathbb{Z}_+, \quad \forall i = 1, \dots, M, t = 1, \dots, T, \quad (1d)$$

$$y_{tij} \in \mathbb{R}_+, \quad \forall i = 1, \dots, M, j = 1, \dots, N, t = 1, \dots, T. \quad (1e)$$

The objective function (1a) minimizes the total maintenance and operational cost over all time periods, where $\sum_{\tau=1}^t x_{\tau i}$ represents the number of units of resources available in facility i at period t . Constraints (1b) require all the demands to be satisfied in each period. Constraints (1c) indicate that, in each period, we can only serve customer demand within the current capacity of each facility.

In the remainder of this paper, we work with a generic capacity planning model in vector forms below:

$$\min_{\substack{x_1, \dots, x_T \\ y_1, \dots, y_T}} \sum_{t=1}^T f_t^T \sum_{\tau=1}^t x_\tau + \sum_{t=1}^T c_t^T y_t \quad (2a)$$

$$\text{s.t.} \quad A_t y_t = d_t, \quad \forall t = 1, \dots, T, \quad (2b)$$

$$B_t y_t \leq \sum_{\tau=1}^t x_\tau, \quad \forall t = 1, \dots, T, \quad (2c)$$

$$x_t \in \mathbb{Z}_+^M, \quad y_t \in \mathbb{R}_+^{MN \times 1}, \quad \forall t = 1, \dots, T,$$

where $x_t \in \mathbb{Z}_+^M$ represents the capacity acquisition decisions (in number of units) for the resources and $y_t \in \mathbb{R}_+^{MN \times 1}$ represents the operational-level allocation of resources in period t . Matrices $A_t \in \mathbb{R}^{N \times MN}$ and $B_t \in \mathbb{R}^{M \times MN}$ correspond to the coefficients in Constraints (1b) and (1c), respectively (i.e., A_t is a 0–1 matrix corresponding to Constraints (1b), and B_t is a matrix with entries 0 or $1/h_{ti}$ corresponding to Constraints (1c)). Note that the difference between Model (2) and models (7)–(10) in Huang and Ahmed (2009) is that, instead of considering one-time acquisition costs of resources, we take into account the period-wise maintenance costs for all acquired resources, which is more realistic in applications such as EV charging station maintenance (Caldwell 2022), power system equipment maintenance (Rehman 2020), etc.

2.1. Scenario Tree Representation

In Model (2), the data we acquire at each period t is the demand $d_t \in \mathbb{R}_+^N$ for all $t = 1, \dots, T$. We consider that the data series $\{d_2, \dots, d_T\}$ evolve according to a known probability distribution, and $d_1 \in \mathbb{R}_+^N$ is deterministic (see similar assumptions made in, e.g., Huang and Ahmed 2009, Shapiro et al. 2009, Zou et al. 2019). In practice, d_1 can be derived and forecasted as the average of historical demand. Consider a discrete distribution and assume that the number of realizations is finite, where such an approximation can be constructed by Monte Carlo sampling if the probability distribution is instead continuous and the resultant problem is called sample average approximation (Kleywegt et al. 2002).

In a multistage stochastic setting, the uncertain demand is revealed gradually, where we need to decide both capacity acquisition x_t and resource allocation y_t in each stage t based on the currently realized demand d_t . Correspondingly, the decision-making process can be described as follows:

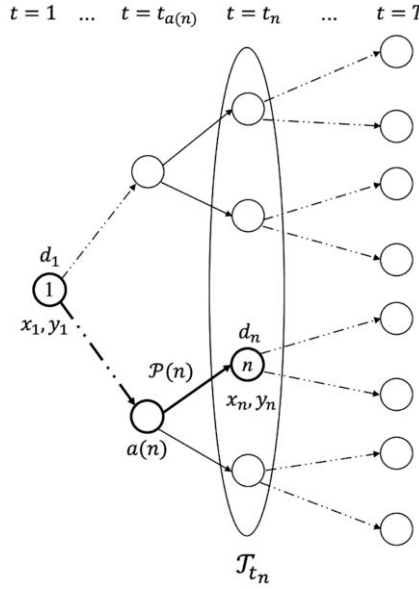
$$\underbrace{\text{decision } (x_1, y_1)}_{\text{Stage 1}} \rightarrow \underbrace{\text{observation } (d_2) \rightarrow \text{decision } (x_2, y_2)}_{\text{Stage 2}} \rightarrow \text{observation } (d_3) \rightarrow \dots \rightarrow \text{decision } (x_{T-1}, y_{T-1}) \rightarrow \underbrace{\text{observation } (d_T) \rightarrow \text{decision } (x_T, y_T)}_{\text{Stage } T}.$$

To facilitate formulating the stochastic programs, we introduce a scenario tree representation of the uncertainty and decision variables (Shapiro et al. 2009). Let $\mathcal{T}$ be the set of all nodes in the scenario tree associated with the underlying stochastic process (see Figure 1). Each node n in period $t > 1$ has a unique parent node $a(n)$ in period $t - 1$, and the set of children nodes of node n is denoted by $\mathcal{C}(n)$. The set $\mathcal{T}_t$ denotes all the nodes corresponding to time period t , and t_n is the time period corresponding to node n . Specifically, all nodes in the last period $\mathcal{T}_T$ are referred to as leaf nodes and denoted by $\mathcal{L}$. The set of all nodes on the path from the root node to node n is denoted by $\mathcal{P}(n)$. Each node n is associated with a (unconditional) probability p_n , which is the probability of the realization of the t_n -period data sequence $\{d_m\}_{m \in \mathcal{P}(n)}$. The probabilities of the nodes in each period sum up to one, that is, $\sum_{n \in \mathcal{T}_t} p_n = 1, \quad \forall t = 1, \dots, T$, and the probabilities of all children nodes sum up to the probability of the parent node, that is, $\sum_{m \in \mathcal{C}(n)} p_m = p_n, \quad \forall n \notin \mathcal{L}$. If n is a leaf node, that is, $n \in \mathcal{L}$, then $\mathcal{P}(n)$ corresponds to a scenario that represents a joint realization of the problem parameters over all periods. For each leaf node $n \in \mathcal{L}$, we denote the probability of the scenario $\mathcal{P}(n)$ as p_n . We denote the capacity acquisition decision variable at node n by x_n and the resource allocation decision variable at node n by y_n . Figure 1 depicts a scenario tree representation of T -period uncertainty and its related notation.

2.2. ECRMs

Next, we introduce the definition and key properties of coherent risk measures. Consider a probability space $(\Xi, \mathcal{F}, P)$ and let $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots \subset \mathcal{F}_T$ be sub-sigma-algebras of $\mathcal{F}$ such that each $\mathcal{F}_t$ corresponds to the information

Figure 1. Illustration of a Scenario Tree and Its Related Notation



available up to (and including) period t with $\mathcal{F}_1 = \{\emptyset, \Xi\}$, $\mathcal{F}_T = \mathcal{F}$. Let Z_t denote a space of $\mathcal{F}_t$ -measurable functions from Ξ to $\mathbb{R}$, for example $d_t \in Z_t$, $\forall t = 1, \dots, T$, and let $Z := Z_1 \times \dots \times Z_T$.

Definition 1 (Artzner et al. 1999). A conditional risk measure $\rho_t^{d_{[t-1]}} : Z_t \rightarrow Z_{t-1}$ is coherent if it satisfies the following four properties: (i) monotonicity: if $Z_1, Z_2 \in Z_t$ and $Z_1 \geq Z_2$, then $\rho_t^{d_{[t-1]}}(Z_1) \geq \rho_t^{d_{[t-1]}}(Z_2)$; (ii) convexity: $\rho_t^{d_{[t-1]}}(\gamma Z_1 + (1 - \gamma)Z_2) \leq \gamma \rho_t^{d_{[t-1]}}(Z_1) + (1 - \gamma)\rho_t^{d_{[t-1]}}(Z_2)$ for all $Z_1, Z_2 \in Z_t$ and all $\gamma \in [0, 1]$; (iii) translation invariance: if $W \in Z_{t-1}$ and $Z \in Z_t$, then $\rho_t^{d_{[t-1]}}(Z + W) = \rho_t^{d_{[t-1]}}(Z) + W$; and (iv) positive homogeneity: if $\gamma \geq 0$ and $Z \in Z_t$, then $\rho_t^{d_{[t-1]}}(\gamma Z) = \gamma \rho_t^{d_{[t-1]}}(Z)$. Here, $Z_1 \geq Z_2$ if and only if $Z_1(\xi) \geq Z_2(\xi)$ for almost everywhere $\xi \in \Xi$.

For notational simplicity, denote the period-wise cost $g_t(x_{1:t}, y_t) = f_t^\top \sum_{\tau=1}^t x_\tau + c_t^\top y_t$ by g_t . We consider a multi-period risk function $\mathbb{F}$ as a mapping from Z to $\mathbb{R}$ below:

$$\mathbb{F}(g_1, \dots, g_T) = g_1 + \rho_2(g_2) + \mathbb{E}_{d_{[2]}}[\rho_3^{d_{[2]}}(g_3)] + \mathbb{E}_{d_{[3]}}[\rho_4^{d_{[3]}}(g_4)] + \dots + \mathbb{E}_{d_{[T-1]}}[\rho_T^{d_{[T-1]}}(g_T)], \quad (3)$$

where $\rho_t^{d_{[t-1]}}$ is a conditional risk measure mapping from Z_t to Z_{t-1} to represent the risk given the information available up to (including) period $t - 1$, that is, $d_{[t-1]} = (d_1, \dots, d_{t-1})$. This class of multiperiod risk measures is called ECRMs, first introduced by Homem-de Mello and Pagnoncelli (2016). We show that the ECRMs (3) are time-consistent following the definition in Ruszczyński (2010) in Online Appendix A.

Using the tower property of expectations (Wasserman 2004), the multiperiod risk function ECRM (3) can be written in a nested form below:

$$\mathbb{F}(g_1, \dots, g_T) = g_1 + \rho_2(g_2) + \mathbb{E}_{d_2}[\rho_3^{d_{[2]}}(g_3) + \mathbb{E}_{d_3|d_{[2]}}[\rho_4^{d_{[3]}}(g_4) + \dots + \mathbb{E}_{d_{T-1}|d_{[T-2]}}[\rho_T^{d_{[T-1]}}(g_T)] \dots]], \quad (4)$$

where $\mathbb{E}_{d_t|d_{[t-1]}}$ represents the expectation with respect to the conditional probability distribution of d_t given realization $d_{[t-1]}$.

For our problem, we consider a special class of single-period coherent risk measures—a convex combination of conditional expectation and CVaR (for the definition of CVaR, please refer to Rockafellar and Uryasev 2000), that is, for $t = 2, \dots, T$,

$$\rho_t^{d_{[t-1]}}(g_t) = (1 - \lambda_t)\mathbb{E}[g_t | d_{[t-1]}] + \lambda_t \text{CVaR}_{\alpha_t}^{d_{[t-1]}}[g_t], \quad (5)$$

where $\lambda_t \in [0, 1]$ is a parameter that adjusts between optimizing on average and risk control and $\alpha_t \in (0, 1)$ is the confidence level. Notice that this risk measure is more general than CVaR, and it includes CVaR as a special case when $\lambda_t = 1$.

Following the results by Rockafellar and Uryasev (2002), CVaR can be expressed as the following optimization problem:

$$\text{CVaR}_{\alpha_t}^{d_{[t-1]}}[g_t] := \inf_{\eta_t \in \mathbb{R}} \left\{ \eta_t + \frac{1}{1 - \alpha_t} \mathbb{E}[(g_t - \eta_t)_+ | d_{[t-1]}] \right\}, \quad (6)$$

where $[a]_+ := \max\{a, 0\}$, and η_t is an auxiliary variable. The minimum of the right-hand side of Equation (6) is attained at $\eta_t^* = \text{VaR}_{\alpha_t}[g_t] := \inf\{v : \mathbb{P}(g_t \leq v) \geq \alpha_t\}$, and thus, CVaR is the mean of the upper $(1 - \alpha_t)$ -tail distribution of g_t , that is, $\mathbb{E}[g_t | g_t > \eta_t^*]$. Selecting a large α_t value makes CVaR sensitive to rare but very high costs (we fix $\alpha_t = 0.95$ and vary λ_t to reflect different risk attitudes in our numerical experiments). To linearize $[g_t - \eta_t]_+$ in (6), we replace it by a variable u_t with two additional constraints: $u_t \geq 0$, $u_t \geq g_t - \eta_t$.

2.3. Risk-Averse Two-Stage and Multistage Stochastic Programs

Combining (4)–(6), the objective function of our risk-averse multistage capacity planning model is specified as

$$\begin{aligned} z_R^{MS} = & \min_{\substack{(x_1, y_1) \in X_1(d_1), \\ \eta_2 \in \mathbb{R}}} g_1(x_1, y_1) + \lambda_2 \eta_2 + \mathbb{E}_{d_2} \left[\min_{\substack{(x_2, y_2, u_2) \in X_2(x_1, \eta_2, d_2), \\ \eta_3 \in \mathbb{R}}} \left\{ \frac{\lambda_2}{1 - \alpha_2} u_2 + (1 - \lambda_2) g_2(x_{1:2}, y_2) + \lambda_3 \eta_3 + \dots \right. \right. \\ & + \mathbb{E}_{d_{T-1} | d_{[1, T-2]}} \left[\min_{\substack{(x_{T-1}, y_{T-1}, u_{T-1}) \in X_{T-1}(x_{1:T-2}, \eta_{T-1}, d_{T-1}), \\ \eta_T \in \mathbb{R}}} \left\{ \frac{\lambda_{T-1}}{1 - \alpha_{T-1}} u_{T-1} + (1 - \lambda_{T-1}) g_{T-1}(x_{1:T-1}, y_{T-1}) + \lambda_T \eta_T \right. \right. \\ & \left. \left. + \mathbb{E}_{d_T | d_{[1, T-1]}} \left[\min_{(x_T, y_T, u_T) \in X_T(x_{1:T-1}, \eta_T, d_T)} \left\{ \frac{\lambda_T}{1 - \alpha_T} u_T + (1 - \lambda_T) g_T(x_{1:T}, y_T) \right\} \right] \dots \right] \right] \right], \end{aligned} \quad (7)$$

where the auxiliary variable $\eta_t \in \mathbb{R}$ is a function of $d_1, \dots, d_{t-1}$, that is, η_t is a $(t-1)$ -stage variable similar to x_{t-1} for all $t = 2, \dots, T$ and the auxiliary variable $u_t \in \mathbb{R}$ is a t -stage variable to represent the excess t -stage cost of the above η_t . Sets $X_1(d_1) = \{(x_1, y_1) \in \mathbb{Z}_+^M \times \mathbb{R}_+^{MN \times 1} : A_1 y_1 = d_1, B_1 y_1 \leq x_1\}$ and $X_t(x_{1:t-1}, \eta_t, d_t) = \{(x_t, y_t, u_t) \in \mathbb{Z}_+^M \times \mathbb{R}_+^{MN \times 1} \times \mathbb{R} : A_t y_t = d_t, B_t y_t \leq \sum_{\tau=1}^t x_\tau, u_t + \eta_t \geq g_t(x_{1:t}, y_t)\}$, $\forall t = 2, \dots, T$ are the feasible regions in each stage.

Plugging $g_t(x_{1:t}, y_t) = f_t^T \sum_{\tau=1}^t x_\tau + c_t^T y_t$ into Equation (7) and using scenario node-based notation, the risk-averse multistage model (7) can be written in the following extensive form:

$$z_R^{MS} = \min_{\substack{x_n, y_n, n \in \mathcal{T} \\ \eta_n, n \notin \mathcal{L}, u_n, n \neq 1}} \sum_{n \in \mathcal{T}} p_n \left(\tilde{f}_n^T \sum_{m \in \mathcal{P}(n)} x_m + \tilde{c}_n^T y_n + \tilde{\lambda}_n \eta_n + \tilde{\alpha}_n u_n \right) \quad (8a)$$

$$\text{s.t. } A_{t_n} y_n = d_n, \quad \forall n \in \mathcal{T}, \quad (8b)$$

$$B_{t_n} y_n \leq \sum_{m \in \mathcal{P}(n)} x_m, \quad \forall n \in \mathcal{T}, \quad (8c)$$

$$u_n + \eta_{a(n)} \geq f_{t_n}^T \sum_{m \in \mathcal{P}(n)} x_m + c_{t_n}^T y_n, \quad \forall n \neq 1, \quad (8d)$$

$$x_n \in \mathbb{Z}_+^M, y_n \in \mathbb{R}_+^{MN \times 1}, \quad \forall n \in \mathcal{T}, u_n \geq 0, \quad \forall n \neq 1,$$

where $\tilde{f}_n = f_{t_n}$ if $n = 1$ and $\tilde{f}_n = (1 - \lambda_{t_n}) f_{t_n}$ otherwise, $\tilde{c}_n = c_{t_n}$ if $n = 1$ and $\tilde{c}_n = (1 - \lambda_{t_n}) c_{t_n}$ otherwise, $\tilde{\lambda}_n = 0$ if $n \in \mathcal{L}$ and $\tilde{\lambda}_n = \lambda_{t_{n+1}}$ otherwise, and $\tilde{\alpha}_n = 0$ if $n = 1$ and $\tilde{\alpha}_n = \frac{\lambda_{t_n}}{1 - \alpha_{t_n}}$ otherwise.

Note that, in the above multistage stochastic program, we have both capacity expansion x_n and allocation y_n decisions corresponding to each node n in the scenario tree, η_n -variables are defined for all nodes n except for the leaf nodes $\mathcal{L}$, and u_n -variables are defined for nonroot nodes $n \neq 1$. When $\lambda_t = 0$, $\forall t = 2, \dots, T$, the risk measure ρ_t becomes the expectation and (8) reduces to a risk-neutral multistage stochastic program.

In a two-stage stochastic program, we decide the capacity expansion plan for all time periods in the first stage regardless of the parameter realizations, and in the second stage, we decide the capacity allocation plan based on the realized demand in each period. Thus, we do not allow any flexibility in the capacity-acquisition plan with respect to the demand realizations; that is, x_n should take the same values for all the nodes in the same stage. This gives us the following extensive formulation of the two-stage capacity

planning model:

$$z_R^{TS} = \min_{\substack{x_n, y_n, n \in \mathcal{T} \\ \eta_n, n \notin \mathcal{L}, u_n, n \neq 1}} \sum_{n \in \mathcal{T}} p_n \left(\tilde{f}_n^T \sum_{m \in \mathcal{P}(n)} x_m + \tilde{c}_n^T y_n + \tilde{\lambda}_n \eta_n + \tilde{\alpha}_n u_n \right) \quad (9a)$$

$$\text{s.t.} \quad (8b)–(8d) \quad (9b)$$

$$x_m = x_n, \quad \forall m, n \in \mathcal{T}_t, t = 1, \dots, T,$$

$$x_n \in \mathbb{Z}_+^M, y_n \in \mathbb{R}_+^{MN \times 1}, \quad \forall n \in \mathcal{T}, u_n \geq 0, \quad \forall n \neq 1.$$

In the following proposition, we first show that the optimal objective values of Models (8) and (9) both increase as the decision maker becomes more risk averse. All the proofs throughout this paper are presented in Online Appendix B.

Proposition 1. Both z_R^{MS} and z_R^{TS} increase as λ increases.

We define the difference between the optimal objective values of the risk-averse two-stage and multistage formulations as the *value of risk-averse multistage stochastic programming*: $VMS_R = z_R^{TS} - z_R^{MS}$. From the extensive formulations, we observe that the risk-averse two-stage Model (9) is the multistage Model (8) with additional Constraints (9b). We, thus, conclude that $VMS_R \geq 0$ as the multistage model provides more flexibility in the capacity acquisition decisions with respect to the uncertain parameter realizations. Because of Constraints (9b), x_n variables can be indexed by the time stage t instead, and the two-stage Model (9) involves $M \times T$ integer variables, whereas the multistage Model (8) involves $M \times |\mathcal{T}|$ integer variables ($|\mathcal{T}|$ is the total number of nodes in the scenario tree $\mathcal{T}$). For any nontrivial scenario tree, $|\mathcal{T}| \gg T$, and thus, solving the multistage Model (8) requires much more computational effort than solving the two-stage counterpart (9). If VMS_R is high, then it may be worth solving a more computationally expensive multistage model; otherwise, a two-stage model is sufficient to provide a good enough solution.

3. Value of Multistage Risk-Averse Stochastic Programming in Capacity Planning

Computing VMS_R exactly requires us to solve the multistage model, which is computationally expensive. In this section, we provide lower and upper bounds on VMS_R without the need to solve the multistage model. These bounds can serve as a priori estimates of VMS_R to analyze the trade-off between risk-averse two-stage and multistage models and can be used to design guidelines on model choices. Next, we first examine a substructure of our problem and provide its analytical optimal solutions in Section 3.1. Based on that, we derive lower and upper bounds for VMS_R in Sections 3.2 and 3.3, respectively. Using the derived bounds, we design a flowchart to help decide which model to solve in Section 3.4.

3.1. Analytical Solutions of the Substructure Problem

We examine an important substructure of Models (8) and (9) once we fix (y, u) -variables. We denote the resultant problems with known (y_n^*, u_n^*) values as **SP-RMS** (y_n^*, u_n^*) and **SP-RTS** (y_n^*, u_n^*) , which are defined as follows:

$$\text{SP-RMS}(y_n^*, u_n^*): \min_{\substack{x_n, n \in \mathcal{T} \\ \eta_n, n \notin \mathcal{L}}} \sum_{n \in \mathcal{T}} p_n \left(\tilde{f}_n^T \sum_{m \in \mathcal{P}(n)} x_m + \tilde{\lambda}_n \eta_n \right) \quad (10a)$$

$$\text{s.t.} \quad \sum_{m \in \mathcal{P}(n)} x_m \geq B_{t_n} y_n^*, \quad \forall n \in \mathcal{T}, \quad (10b)$$

$$\eta_{a(n)} \geq \tilde{f}_{t_n}^T \sum_{m \in \mathcal{P}(n)} x_m + \tilde{c}_{t_n}^T y_n^* - u_n^*, \quad \forall n \neq 1, \quad (10c)$$

$$x_n \in \mathbb{Z}_+^M, \quad \forall n \in \mathcal{T},$$

and

$$\text{SP-RTS}(\mathbf{y}_n^*, u_n^*) : \min_{\substack{\mathbf{x}_n, n \in T \\ \eta_n, n \notin \mathcal{L}}} \sum_{n \in T} p_n \left(\tilde{\mathbf{f}}_n^T \sum_{m \in \mathcal{P}(n)} \mathbf{x}_m + \tilde{\lambda}_n \eta_n \right) \quad (11)$$

s.t. (10b)–(10c)

(9b) (Two-stage constraints for $\mathbf{x}$)

$$\mathbf{x}_n \in \mathbb{Z}_+^M, \quad \forall n \in T.$$

Here, we denote the optimal objective values of Models (10) and (11) as $Q^M(\mathbf{y}_n^*, u_n^*)$, $Q^T(\mathbf{y}_n^*, u_n^*)$, respectively. The next proposition demonstrates the analytical forms of the optimal solutions to **SP-RMS**($\mathbf{y}_n^*, u_n^*$) and **SP-RTS**($\mathbf{y}_n^*, u_n^*$), and we present a detailed proof in Online Appendix B.

Proposition 2. *Given any $(\mathbf{y}^*, u^*)$ values, the optimal solutions of (10) and (11) have the following analytical forms:*

$$\begin{aligned} \mathbf{x}_1^{MS} &= \lceil \mathbf{B}_{t_1} \mathbf{y}_1^* \rceil, \quad \mathbf{x}_n^{MS} = \max_{m \in \mathcal{P}(n)} \lceil \mathbf{B}_{t_m} \mathbf{y}_m^* \rceil - \max_{m \in \mathcal{P}(a(n))} \lceil \mathbf{B}_{t_m} \mathbf{y}_m^* \rceil, \quad \forall n \neq 1, \\ \eta_n^{MS} &= \max_{m \in \mathcal{C}(n)} \left\{ \mathbf{f}_{t_m}^T \sum_{l \in \mathcal{P}(m)} \mathbf{x}_l^{MS} + \mathbf{c}_{t_m}^T \mathbf{y}_m^* - u_m^* \right\}, \quad \forall n \notin \mathcal{L}, \\ \mathbf{x}_1^{TS} &= \lceil \mathbf{B}_{t_1} \mathbf{y}_1^* \rceil, \quad \mathbf{x}_n^{TS} = \max_{m \in \mathcal{P}(n)} \lceil \max_{l \in T_{t_m}} \mathbf{B}_{t_l} \mathbf{y}_l^* \rceil - \max_{m \in \mathcal{P}(a(n))} \lceil \max_{l \in T_{t_m}} \mathbf{B}_{t_l} \mathbf{y}_l^* \rceil, \quad \forall n \neq 1, \\ \eta_n^{TS} &= \max_{m \in \mathcal{C}(n)} \left\{ \mathbf{f}_{t_m}^T \sum_{l \in \mathcal{P}(m)} \mathbf{x}_l^{TS} + \mathbf{c}_{t_m}^T \mathbf{y}_m^* - u_m^* \right\}, \quad \forall n \notin \mathcal{L}, \end{aligned}$$

and correspondingly, we have $Q^M(\mathbf{y}_n^*, u_n^*) = \sum_{n \in T} p_n (\tilde{\mathbf{f}}_n^T \sum_{m \in \mathcal{P}(n)} \mathbf{x}_m^{MS} + \tilde{\lambda}_n \eta_n^{MS})$, $Q^T(\mathbf{y}_n^*, u_n^*) = \sum_{n \in T} p_n (\tilde{\mathbf{f}}_n^T \sum_{m \in \mathcal{P}(n)} \mathbf{x}_m^{TS} + \tilde{\lambda}_n \eta_n^{TS})$.

Remark 1. From Proposition 2, one can easily verify that $Q^T(\mathbf{y}_n^*, u_n^*) - Q^M(\mathbf{y}_n^*, u_n^*) \geq 0$ as $\sum_{m \in \mathcal{P}(n)} \mathbf{x}_m^{TS} = \max_{m \in \mathcal{P}(n)} \lceil \max_{l \in T_{t_m}} \mathbf{B}_{t_l} \mathbf{y}_l^* \rceil \geq \max_{m \in \mathcal{P}(n)} \lceil \mathbf{B}_{t_m} \mathbf{y}_m^* \rceil = \sum_{m \in \mathcal{P}(n)} \mathbf{x}_m^{MS}$. Next, we use Proposition 2 to construct lower and upper bounds on the VMS_R in Theorem 1, Corollary 1, and Theorem 2 as well as to design approximation algorithms to solve Model (8) in Algorithm 1.

3.2. Lower Bound on VMS_R for Risk-Averse Capacity Planning

We now describe two lower bounds on the VMS_R for the risk-averse multistage and two-stage capacity planning Models (8) and (9) based on the analysis of the substructure problems. Note that these lower bounds are useful when they are sufficiently large, indicating that there is a need to solve the much harder multistage model.

Theorem 1. *Let $\{\mathbf{y}_n^*\}_{n \in T}$, $\{u_n^*\}_{n \in T \setminus \{1\}}$ be the decisions in an optimal solution to the two-stage Model (9), and let $\mathbf{x}^{MS}$, $\boldsymbol{\eta}^{MS}$, $\mathbf{x}^{TS}$, $\boldsymbol{\eta}^{TS}$ follow the definitions in Proposition 2, which are constructed by $\{\mathbf{y}_n^*\}_{n \in T}$, $\{u_n^*\}_{n \in T \setminus \{1\}}$. Then,*

$$\text{VMS}_R \geq \text{VMS}_R^{\text{LB}} := \sum_{n \in T \setminus \{1\}} p_n (1 - \lambda_{t_n}) \mathbf{f}_{t_n}^T \left(\max_{m \in \mathcal{P}(n)} \lceil \max_{l \in T_{t_m}} \mathbf{B}_{t_l} \mathbf{y}_l^* \rceil - \max_{m \in \mathcal{P}(n)} \lceil \mathbf{B}_{t_m} \mathbf{y}_m^* \rceil \right) + \sum_{n \in T \setminus \mathcal{L}} p_n \lambda_{t_{n+1}} (\eta_n^{TS} - \eta_n^{MS}). \quad (12)$$

Later, in Example 1, we show that VMS_R^{LB} (Equation (12)) is tight. Next, we derive another lower bound $\text{VMS}_R^{\text{LB1}}$ that is not necessarily tight but more computationally tractable, which utilizes the optimal solutions to the LP relaxation of the two-stage Model (9). Because we consider the LP relaxation of the two-stage model, we do not round up when constructing the two-stage solutions $\mathbf{x}_n^{TS}$.

Corollary 1. Let $\{y_n^{TSLP}\}_{n \in T}$, $\{u_n^{TSLP}\}_{n \in T \setminus \{1\}}$ be the decisions in an optimal solution to the LP relaxation of the two-stage Model (9), and let

$$\begin{aligned} x_1^{MS} &= \lceil B_{t_1} y_1^{TSLP} \rceil, x_n^{MS} = \max_{m \in \mathcal{P}(n)} \lceil B_{t_m} y_m^{TSLP} \rceil - \max_{m \in \mathcal{P}(a(n))} \lceil B_{t_m} y_m^{TSLP} \rceil, \forall n \neq 1, \\ \eta_n^{MS} &= \max_{m \in \mathcal{C}(n)} \left\{ f_{t_m}^\top \sum_{l \in \mathcal{P}(m)} x_l^{MS} + c_{t_m}^\top y_m^{TSLP} - u_m^{TSLP} \right\}, \forall n \notin \mathcal{L}, \\ x_1^{TS} &= B_{t_1} y_1^{TSLP}, x_n^{TS} = \max_{m \in \mathcal{P}(n)} \max_{l \in T_{t_m}} B_{t_l} y_l^{TSLP} - \max_{m \in \mathcal{P}(a(n))} \max_{l \in T_{t_m}} B_{t_l} y_l^{TSLP}, \forall n \neq 1, \\ \eta_n^{TS} &= \max_{m \in \mathcal{C}(n)} \left\{ f_{t_m}^\top \sum_{l \in \mathcal{P}(m)} x_l^{TS} + c_{t_m}^\top y_m^{TSLP} - u_m^{TSLP} \right\}, \forall n \notin \mathcal{L}. \end{aligned}$$

Then,

$$\begin{aligned} \text{VMS}_R \geq \text{VMS}_R^{\text{LB1}} &:= f_{t_1}^\top (B_{t_1} y_1^{TSLP} - \lceil B_{t_1} y_1^{TSLP} \rceil) + \sum_{n \in T \setminus \{1\}} p_n (1 - \lambda_{t_n}) f_{t_n}^\top \left(\max_{m \in \mathcal{P}(n)} \max_{l \in T_{t_m}} B_{t_l} y_l^{TSLP} - \max_{m \in \mathcal{P}(n)} \lceil B_{t_m} y_m^{TSLP} \rceil \right) \\ &\quad + \sum_{n \in T \setminus \mathcal{L}} p_n \lambda_{t_n+1} (\eta_n^{TS} - \eta_n^{MS}). \end{aligned}$$

Note that, different than VMS_R^{LB} , $\text{VMS}_R^{\text{LB1}}$ is not guaranteed to be always nonnegative. When $\text{VMS}_R^{\text{LB1}} < 0$, we can simply replace it with a trivial lower bound zero, that is, $\text{VMS}_R \geq \max\{\text{VMS}_R^{\text{LB1}}, 0\}$. Computing $\text{VMS}_R^{\text{LB1}}$ and VMS_R^{LB} requires solving the two-stage Model (9) and its LP relaxation, respectively, both of which are more computationally tractable than solving the multistage counterpart.

3.3. Upper Bound on VMS_R for Risk-Averse Capacity Planning

We now describe an upper bound on the VMS_R based on the optimal solution of the LP relaxation of the multistage Model (8). Because we consider the LP relaxation of the multistage model, we do not round up when constructing the multistage solutions x_n^{MS} . Note that this upper bound is useful when it is small enough, indicating that the two-stage model is sufficient to provide a good enough solution.

Theorem 2. Let $\{y_n^{\text{MSLP}}\}_{n \in T}$, $\{u_n^{\text{MSLP}}\}_{n \in T \setminus \{1\}}$ be the decisions in an optimal solution to the LP relaxation of multistage Model (8), and let

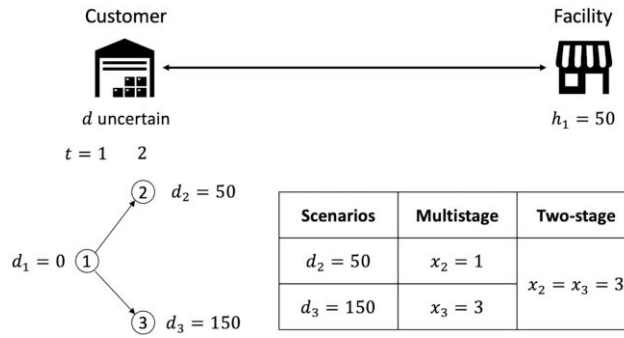
$$\begin{aligned} x_1^{MS} &= B_{t_1} y_1^{\text{MSLP}}, x_n^{MS} = \max_{m \in \mathcal{P}(n)} B_{t_m} y_m^{\text{MSLP}} - \max_{m \in \mathcal{P}(a(n))} B_{t_m} y_m^{\text{MSLP}}, \forall n \neq 1, \\ \eta_n^{MS} &= \max_{m \in \mathcal{C}(n)} \left\{ f_{t_m}^\top \sum_{l \in \mathcal{P}(m)} x_l^{MS} + c_{t_m}^\top y_m^{\text{MSLP}} - u_m^{\text{MSLP}} \right\}, \forall n \notin \mathcal{L}, \\ x_1^{TS} &= \lceil B_{t_1} y_1^{\text{MSLP}} \rceil, x_n^{TS} = \max_{m \in \mathcal{P}(n)} \lceil \max_{l \in T_{t_m}} B_{t_l} y_l^{\text{MSLP}} \rceil - \max_{m \in \mathcal{P}(a(n))} \lceil \max_{l \in T_{t_m}} B_{t_l} y_l^{\text{MSLP}} \rceil, \forall n \neq 1, \\ \eta_n^{TS} &= \max_{m \in \mathcal{C}(n)} \left\{ f_{t_m}^\top \sum_{l \in \mathcal{P}(m)} x_l^{TS} + c_{t_m}^\top y_m^{\text{MSLP}} - u_m^{\text{MSLP}} \right\}, \forall n \notin \mathcal{L}. \end{aligned}$$

Then,

$$\begin{aligned} \text{VMS}_R \leq \text{VMS}_R^{\text{UB}} &:= f_{t_1}^\top (\lceil B_{t_1} y_1^{\text{MSLP}} \rceil - B_{t_1} y_1^{\text{MSLP}}) + \sum_{n \in T \setminus \{1\}} p_n (1 - \lambda_{t_n}) f_{t_n}^\top \left(\max_{m \in \mathcal{P}(n)} \lceil \max_{l \in T_{t_m}} B_{t_l} y_l^{\text{MSLP}} \rceil - \max_{m \in \mathcal{P}(n)} B_{t_m} y_m^{\text{MSLP}} \right) \\ &\quad + \sum_{n \in T \setminus \mathcal{L}} p_n \lambda_{t_n+1} (\eta_n^{TS} - \eta_n^{MS}). \end{aligned} \tag{13}$$

The following example illustrates the usefulness of the derived lower and upper bounds.

Example 1. We consider a special example with one facility ($h_1 = 50$) and one customer site (see Figure 2). The underlying scenario tree has two periods, and the root node has two branches with equal probability. Let the customer demand be $d_1 = 0$ in node 1 (stage 1), $d_2 = 50$ in node 2 (stage 2), and $d_3 = 150$ in node 3 (stage 2). Denote f as the unit maintenance cost for one stage and c as the operational cost, respectively. Solving a two-stage model

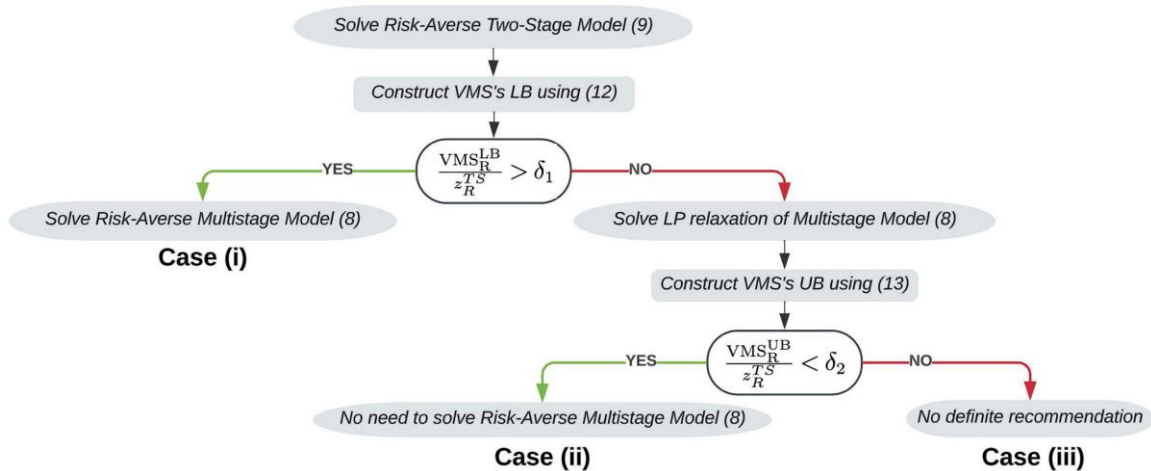
Figure 2. An Instance to Illustrate the Gap Between Multistage and Two-Stage Capacity Planning Models

yields a total cost of $z_R^{TS} = 3f + 100c + 50\lambda c$, and a multistage model gives a total cost of $z_R^{MS} = 2f + 100c + \lambda f + 50\lambda c$. Thus, $VMS_R = z_R^{TS} - z_R^{MS} = (1 - \lambda)f$. According to Equations (12) and (13), we have $VMS_R^{LB} = (1 - \lambda)f$, $VMS_R^{UB} = (1 + \lambda)f$, where the lower bound VMS_R^{LB} is actually tight. We refer interested readers to Online Appendix C for detailed computations of these costs and bounds.

3.4. General Guidelines on Model Choices

The bounds in Theorems 1 and 2 are dependent on the input data and optimal solutions to the two-stage model and LP relaxation of the multistage model. Once the decision maker computes these bounds, they can compare the relative bounds with some pre-given thresholds δ_1 and δ_2 (e.g., $\delta_1 = 10\%$, $\delta_2 = 30\%$) and then decide if the two-stage optimal solution is good enough. We outline a general procedure in Figure 3. If the relative VMS_R^{LB} is sufficiently large (i.e., $\frac{VMS_R^{LB}}{z_R^{TS}} > \delta_1$), then it indicates that the multistage model can reduce the optimal objective value by at least δ_1 compared with the two-stage model, and thus, one should solve the multistage model. On the other hand, if the relative VMS_R^{UB} is sufficiently small (i.e., $\frac{VMS_R^{UB}}{z_R^{TS}} < \delta_2$), then we suggest adopting the optimal solutions to the two-stage model without bearing the additional computational effort of solving the multistage model. In the third case, if the lower bound VMS_R^{LB} is not large enough and the upper bound VMS_R^{UB} is not small enough, then there is no definite recommendation. Later, in Section 5, we show that most of our instances fall into either case (i) or case (ii). Note that these two parameters δ_1 and δ_2 are user-defined and can represent the decision maker's trade-off between optimality and tractability.

Example 2. Using the same setting in Example 1 and letting $\lambda = 0.5$, $\delta_1 = 10\%$, $\delta_2 = 30\%$, $f = 1,000$, $c = 10$, we have $z_R^{TS} = 4,250$ and $VMS_R^{LB} = 500$. Because $\frac{VMS_R^{LB}}{z_R^{TS}} = \frac{500}{4,250} > \delta_1 = 10\%$, we solve the risk-averse multistage model. In this case, the unit capacity expansion cost f is relatively high and the true $VMS_R = (1 - \lambda)f$ is also high, indicating the value of the multistage model. On the other hand, if the unit capacity expansion cost is relatively low, for example,

Figure 3. (Color online) A Flowchart to Decide Which Model to Solve

$f = 100$, then $z_R^{TS} = 1,550$ and $VMS_R^{LB} = 50$, $VMS_R^{UB} = 150$. Because $\frac{VMS_R^{LB}}{z_R^{TS}} = \frac{50}{1,550} < \delta_1 = 10\%$ and $\frac{VMS_R^{UB}}{z_R^{TS}} = \frac{150}{1,550} < \delta_2 = 30\%$, we use the optimal solution of the two-stage model without solving the multistage model.

4. Approximation Algorithms

Up to this point, we have discussed how to bound the gap between risk-averse two-stage and multistage models based on their optimal objective values. However, when we enter case (i), it is inherently hard to solve a risk-averse multistage model as we show in Theorem 3. In this section, we propose approximation algorithms to solve the risk-averse multistage Model (8) more efficiently, which can be also applied to solve the risk-averse two-stage Model (9).

Theorem 3. *The deterministic capacity planning problem (1) and its risk-averse multistage and two-stage counterparts (8) and (9) are NP-hard.*

Motivated by the computational intractability of Models (8) and (9), we proceed to introduce approximation algorithms that can solve the risk-averse multistage and two-stage models efficiently by utilizing the decomposition structure we investigate in Section 3.1. Next, we describe the main idea of the algorithm for the risk-averse multistage Model (8) as follows: we first solve the LP relaxation of Model (8) to obtain a feasible solution $(\mathbf{y}_n^{MSLP}, u_n^{MSLP})$, which is fed into the substructure problem **SP-RMS** $(\mathbf{y}_n^{MSLP}, u_n^{MSLP})$ to obtain an optimal solution $(\mathbf{x}_n^1, \eta_n^1)$. Note that we can always find a feasible solution after plugging $(\mathbf{y}_n^{MSLP}, u_n^{MSLP})$ into **SP-RMS** because we can make x and η sufficiently large to satisfy Constraints (10b) and (10c). We then solve Model (8) with fixed $(\mathbf{x}_n, \eta_n) = (\mathbf{x}_n^1, \eta_n^1)$ to derive an optimal solution $(\mathbf{y}_n^1, u_n^1)$, which together with $(\mathbf{x}_n^1, \eta_n^1)$, constitutes a feasible solution and, thus, an upper bound to Model (8). This upper bound can be strengthened iteratively by repeating the process, and we denote the feasible solution produced at the end of Algorithm 1 by $(\mathbf{x}_n^H, \eta_n^H, \mathbf{y}_n^H, u_n^H)$. The detailed steps are described in Algorithm 1. We show the monotonicity of the upper bounds derived in Algorithm 1 in Proposition 3 and show that the optimality gap of Algorithm 1 can be upper bounded in Proposition 4, which eventually leads to an approximation ratio stated in Theorem 4.

Algorithm 1 (Approximation Algorithm for Risk-Averse Multistage Capacity Planning)

- 1: Solve the LP relaxation of the risk-averse multistage capacity planning problem (8) and let $(\mathbf{x}_n^{MSLP}, \eta_n^{MSLP}, \mathbf{y}_n^{MSLP}, u_n^{MSLP})_{n \in \mathcal{T}}$ be an optimal solution. If $\mathbf{x}_n^{MSLP}$ is integral for all $n \in \mathcal{T}$, stop and return $(\mathbf{x}_n^{MSLP}, \eta_n^{MSLP}, \mathbf{y}_n^{MSLP}, u_n^{MSLP})_{n \in \mathcal{T}}$.
- 2: Initialize $k = 0$ and $(\mathbf{x}_n^0, \eta_n^0, \mathbf{y}_n^0, u_n^0)_{n \in \mathcal{T}} = (\mathbf{x}_n^{MSLP}, \eta_n^{MSLP}, \mathbf{y}_n^{MSLP}, u_n^{MSLP})_{n \in \mathcal{T}}$.
- 3: **while** $\|\mathbf{x}^k - \mathbf{x}^{k-1}\| \geq \epsilon, \|\eta^k - \eta^{k-1}\| \geq \epsilon, \|\mathbf{y}^k - \mathbf{y}^{k-1}\| \geq \epsilon, \|u^k - u^{k-1}\| \geq \epsilon$ **do**
- 4: Solve problem **SP-RMS** $(\mathbf{y}_n^k, u_n^k)$ and let $\mathbf{x}_n^{k+1}, \eta_n^{k+1}$ denote the corresponding optimal solutions. We have the analytical form of the optimal solutions as $\mathbf{x}_1^{k+1} = \lceil \mathbf{B}_{t_1} \mathbf{y}_1^k \rceil$, $\mathbf{x}_n^{k+1} = \max_{m \in \mathcal{P}(n)} [\mathbf{B}_{t_m} \mathbf{y}_m^k] - \max_{m \in \mathcal{P}(a(n))} [\mathbf{B}_{t_m} \mathbf{y}_m^k]$, $\forall n \neq 1$, $\eta_n^{k+1} = \max_{m \in \mathcal{C}(n)} \{f_{t_m}^\top \sum_{l \in \mathcal{P}(m)} \mathbf{x}_l^{k+1} + \mathbf{c}_{t_m}^\top \mathbf{y}_m^k - u_m^k\}$, $\forall n \notin \mathcal{L}$.
- 5: Solve the following problem for each $n \in \mathcal{T}$, $n \neq 1$ independently

$$\begin{aligned} \min_{\mathbf{y}_n, u_n} \quad & \tilde{\mathbf{c}}_n^\top \mathbf{y}_n + \tilde{\alpha}_n u_n \\ \text{s.t.} \quad & \mathbf{B}_{t_n} \mathbf{y}_n \leq \sum_{m \in \mathcal{P}(n)} \mathbf{x}_m^{k+1}, \end{aligned} \quad (14a)$$

$$\mathbf{A} \mathbf{y}_n = \mathbf{d}_n, \quad (14b)$$

$$u_n - \mathbf{c}_{t_n}^\top \mathbf{y}_n \geq f_{t_n}^\top \sum_{m \in \mathcal{P}(n)} \mathbf{x}_m^{k+1} - \eta_{a(n)}^{k+1}, \quad (14c)$$

$$\mathbf{y}_n \in \mathbb{R}_+, u_n \geq 0,$$

and when $n = 1$, we solve Problem (14) without the variables u_n and Constraints (14c). Let $\mathbf{y}_n^{k+1}, u_n^{k+1}$ be the optimal solutions.

- 6: Update $k = k + 1$.
- 7: **end while**
- 8: Return $(\mathbf{x}_n^H, \eta_n^H, \mathbf{y}_n^H, u_n^H)_{n \in \mathcal{T}} := (\mathbf{x}_n^{k+1}, \eta_n^{k+1}, \mathbf{y}_n^{k+1}, u_n^{k+1})_{n \in \mathcal{T}}$.

Throughout this section, we use $z_{T,R}^{MS}$ to denote the objective value of the multistage model (8) with the total number of stages being T . With some abuse of notation, we use $(\mathbf{x}_n^*, \eta_n^*, \mathbf{y}_n^*, u_n^*)$ to denote an optimal solution to the multistage model (8) and use $z_{T,R}^{MS}(\mathbf{x}_n, \eta_n, \mathbf{y}_n, u_n)$ to denote the objective value of Model (8) with decision variable values $(\mathbf{x}_n, \eta_n, \mathbf{y}_n, u_n)$.

Proposition 3. The objective value at the end of each iteration in Algorithm 1 provides an upper bound to the optimal objective value of Model (8), and it satisfies $z_{T,R}^{MS}(\mathbf{x}_n^*, \eta_n^*, \mathbf{y}_n^*, u_n^*) \leq z_{T,R}^{MS}(\mathbf{x}_n^{k+1}, \eta_n^{k+1}, \mathbf{y}_n^{k+1}, u_n^{k+1}) \leq z_{T,R}^{MS}(\mathbf{x}_n^k, \eta_n^k, \mathbf{y}_n^k, u_n^k)$, $\forall k \geq 1$.

Proposition 4. The optimality gap can be bounded above by a quantity only dependent on the maintenance cost, that is, $z_{T,R}^{MS}(\mathbf{x}_n^H, \eta_n^H, \mathbf{y}_n^H, u_n^H) - z_{T,R}^{MS}(\mathbf{x}_n^*, \eta_n^*, \mathbf{y}_n^*, u_n^*) \leq \sum_{t=1}^T \sum_{i=1}^M f_{ti}$.

Theorem 4. Algorithm 1 has an approximation ratio of

$$1 + \frac{M \sum_{t=1}^T f_{t,\max}}{M_{\min} \sum_{t=1}^T f_{t,\min} + \sum_{t=1}^T c_{t,\min} \min_{n \in \mathcal{T}_t} \left\{ \sum_{j=1}^N d_{n,j} \right\}},$$

where $h_{\max} = \max_{i=1}^M \{h_{ti}\}$, $f_{t,\max} = \max_{i=1}^M \{f_{ti}\}$, $f_{t,\min} = \min_{i=1}^M \{f_{ti}\}$, $c_{t,\min} = \min_{i \in [M], j \in [N]} c_{tij}$, and $M_{\min} = \left\lceil \frac{\sum_{j=1}^N d_{1j}}{h_{\max}} \right\rceil$ measures at least how many units of resources we need to cover the first stage demand.

Corollary 2. Assume that $f_{t,\max} = O(1)$, $f_{t,\min} = O(1)$, $c_{t,\min} = O(1)$, $\min_{n \in \mathcal{T}_t} \sum_{j=1}^N d_{n,j} = O(t)$ when $t \rightarrow \infty$. Then, Algorithm 1 is asymptotically optimal, that is, $\lim_{T \rightarrow \infty} \frac{z_{T,R}^{MS}(\mathbf{x}_n^H, \eta_n^H, \mathbf{y}_n^H, u_n^H)}{z_{T,R}^{MS}(\mathbf{x}_n^*, \eta_n^*, \mathbf{y}_n^*, u_n^*)} = 1$.

Note that the assumptions in Corollary 2 are not particularly restrictive. They only require that the unit maintenance and operational costs can be bounded above by a value that does not depend on the time stage, and the demand grows at least linearly when time increases (the result still holds when the demand grows faster than linearly, e.g., when $\min_{n \in \mathcal{T}_t} \sum_{j=1}^N d_{n,j} = O(t^2)$). The last condition can be achieved if we have an expanding market and the minimum of the demand is increased for each subsequent year (e.g., $\min_{n \in \mathcal{T}_t} \sum_{j=1}^N d_{n,j} = \tilde{d}(1 + 2(t-1))$ with $\tilde{d}$ being the nominal demand in the first stage). We test different demand patterns in Section 5.2. Detailed proofs of Propositions 3 and 4, Theorem 4, and Corollary 2 are given in Online Appendix B.

One can also tailor Algorithm 1 to solve the risk-averse two-stage Model (9) by modifying step 4 to use the analytical solutions of $\mathbf{SP-RTS}(\mathbf{y}_n^k, u_n^k)$ as we discuss in Proposition 2. We omit the details in the interest of brevity.

5. Computational Results

We test the risk-averse two-stage and multistage models on two types of networks: a randomly generated grid network in which we vary the parameter settings extensively and a real-world network based on the U.S. map with 49 candidate facilities and 88 customer sites (Daskin 2011). Specifically, we conduct sensitivity analysis and report results based on the synthetic data to illustrate the tightness of the analytical bounds and the efficacy and efficiency of the proposed approximation algorithms in Section 5.1. We also conduct a case study on EV charging station capacity planning based on the U.S. map to display the solution patterns under different settings of uncertainties in Section 5.2. We use Gurobi 10.0.0 coded in Python 3.11.0 for solving all mixed-integer programming models, and the computational time limit is set to one hour. Our numerical tests are conducted on a Macbook Pro with 8 GB RAM and an Apple M1 Pro chip. The source code and data files can be found in Yu and Shen (2024).

5.1. Result Analysis on Synthetic Data

We first introduce the experimental design and setup in Section 5.1.1 and report sensitivity analysis results in Section 5.1.2. Then, we examine the tightness of the analytical bounds in Section 5.1.3 and the performance of the approximation algorithms in Section 5.1.4, respectively. Finally, in Section 5.1.5, we report the computational time for solving the two risk-averse models.

5.1.1. Experimental Design and Setup. We randomly sample M potential facilities and N customer sites on a 100×100 grid, and in the default setting, we have the number of stages (T) being 3, the number of facilities (M) being 5, the number of customer sites (N) being 10, and the number of branches in each nonleaf node (C) being 2. The risk attitude parameters are set to $\lambda_t = 0.5$, $\alpha_t = 0.95$, $\forall t = 2, \dots, T$ at default. The operational costs between facilities and customer sites are calculated by their Manhattan distances times the unit travel cost. We set the per stage maintenance costs $f_{ti} = 6 \times 10^4$, and all the facilities have the same unit capacity $h_{ti} = h = 10^3$ for all $t = 1, \dots, T$, $i = 1, \dots, M$. For each customer site $j = 1, \dots, N$ and stage $t = 1, \dots, T$, we uniformly sample the demand mean from $U(1000(2t-1), 5000(2t-1))$ and then multiply each mean by a fixed number ($\sigma = 0.8$ at default) to generate its demand standard deviation. Lastly, we sample demand data following a truncated normal distribution with the generated mean and standard deviation, whereas negative demand values are deleted. We consider two types of scenario trees:

- Stagewise dependent (SD): at every stage $t = 1, \dots, T - 1$, every node $n \in \mathcal{T}_t$ is associated with a different set of children nodes $\mathcal{C}(n)$, that is, $\mathcal{C}(n) \neq \mathcal{C}(m)$, $\forall n, m \in \mathcal{T}_t$.
- Stagewise independent (SI): at every stage $t = 1, \dots, T - 1$, every node $n \in \mathcal{T}_t$ is associated with an identical set of children nodes $\mathcal{C}(n)$, that is, $\mathcal{C}(n) = \mathcal{C}(m)$, $\forall n, m \in \mathcal{T}_t$.

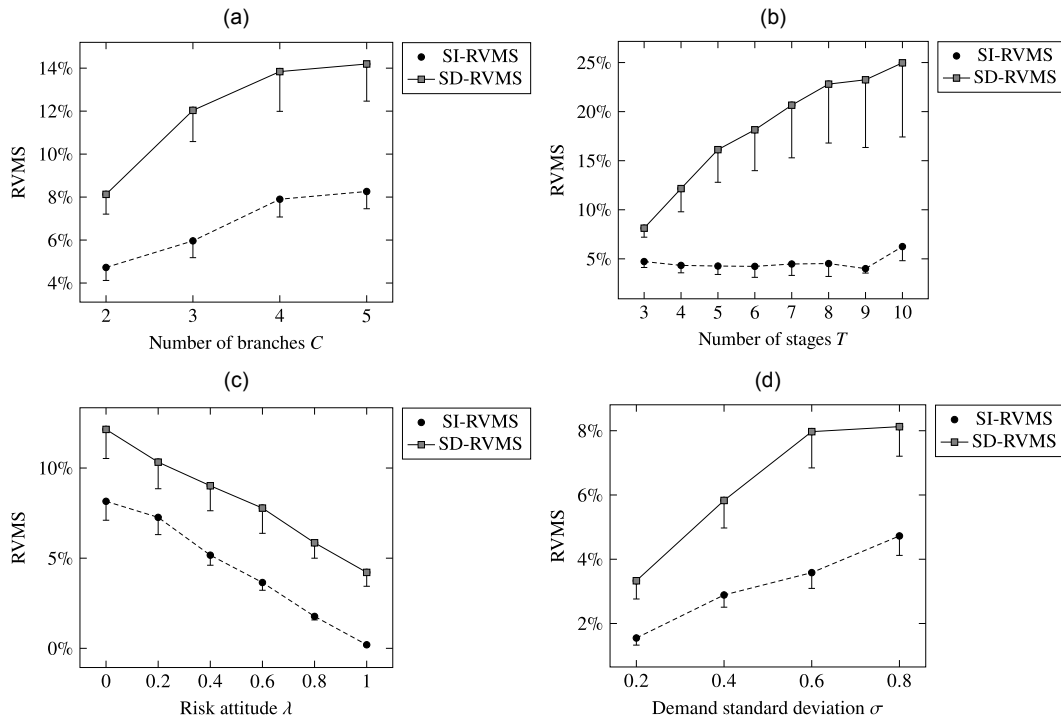
Here, SD represents the most general case in which in each stage t , we have at most C^t different realizations of the uncertainty d_t . On the other hand, SI assumes that the stochastic process $(d_1, d_2, \dots, d_T)$ is stagewise independent, and thus, we have at most C different realizations of the uncertainty in each stage t .

5.1.2. Sensitivity Analysis on VMS. To compare the two-stage and multistage models, we define the relative value of risk-averse multistage stochastic programming as $RVMS = \frac{z_R^{TS} - z_R^{MS}}{z_R^{TS}}$ and a lower bound on the RVMS as $RVMS_{LB} = \frac{VMS_{LB}^{TS}}{z_R^{TS}}$ using VMS_{LB}^{TS} computed by Equation (12). To illustrate what types of demand scenarios make the multistage model more valuable, we compare the above two types of scenario trees to evaluate RVMS, namely, SD-RVMS, and SI-RVMS, respectively. We first vary the number of branches C from 2 to 5, the number of stages T from 3 to 6, the risk attitude λ from 0 to 1, and the standard deviation σ from 0.2 to 0.8 to see how RVMS changes with respect to different parameter settings.

The corresponding results are presented in Figure 4, in which we plot the mean of RVMS over 100 independently generated instances, and $RVMS_{LB}$ is represented by the lower end of the error bars.

From the figure, when we increase the demand variability, such as the number of branches, the number of stages, and the standard deviation, RVMS values with stagewise-dependent scenario trees increase drastically (see Figure 4(a), (b), and (d)). A higher RVMS indicates that the multistage model is much more valuable than the two-stage counterpart. Comparing different types of scenario trees, SD scenario trees always gain much higher RVMS than the stagewise-independent ones, and adding more stages makes no significant changes in SI-RVMS as can be seen in Figure 4(b). This is because, in SI scenario trees, the number of realizations in each stage does not depend on the number of stages, and having a deeper scenario tree would not necessarily increase the demand variability. Following these observations, if the stochastic capacity planning problem has a large number of branches, stages, or high standard deviation with a stagewise-dependent scenario tree, then it is worth solving a multistage model because the gap between a two-stage formulation and a multistage formulation is very high. Notably, from Figure 4(c), RVMS decreases approximately linearly with respect to the risk

Figure 4. Statistics of RVMS over 100 Instances with Different Numbers of Branches C , Stages T , Risk Attitudes λ , and Demand Standard Deviations σ



Notes. (a) Different numbers of branches C . (b) Different numbers of stages T . (c) Different risk attitudes λ . (d) Different standard deviations σ .

Table 1. Percentage of Instances in Different Cases with SD Scenario Trees and Varying Numbers of Branches C When $\delta_1 = 10\%$, $\delta_2 = 30\%$

Branches C	Case (i)	Case (ii)	Case (iii)
2	20/100	73/100	7/100
3	56/100	25/100	19/100
4	69/100	8/100	23/100
5	79/100	2/100	19/100

attitude parameter λ , suggesting that, as we become more risk averse, the performances of the two models tend to be closer. This can be understood from the following facts: (i) a two-stage model restricts all capacity-expansion decisions to be identical in the same stage, and (ii) a risk-averse multistage model aims to reduce the upper tail of the cost distribution in each stage. As a result, the risk-averse multistage model tends to reduce the dispersion of capacity expansion decisions in each stage as we become more risk averse, which has similar effects to two-stage models. Because of this, if a decision maker is extremely risk averse (e.g., $\lambda \approx 1$), given the small gaps between the two models, we suggest solving a risk-averse two-stage model without bearing the additional computational effort. Moreover, the $RVMS_{LB}$ obtained by all instances are pretty close to the true $RVMS$, and stagewise-independent scenario trees obtain tighter lower bounds than stagewise-dependent ones. We take a closer look at these lower bounds in Section 5.1.3.

Following the flowchart in Figure 3 and setting $\delta_1 = 10\%$, $\delta_2 = 30\%$, we have the following three cases: case (i) $RVMS_{LB} > 10\%$; case (ii) $RVMS_{LB} \leq 10\%$ and $RVMS_{UB} < 30\%$; and case (iii) $RVMS_{LB} \leq 10\%$ and $RVMS_{UB} \geq 30\%$. We record the number of instances in each case with SD scenario trees and varying numbers of branches and stages in Tables 1 and 2. From these tables, in most of the instances, we have either case (i) or case (ii). As we increase the number of branches or stages, most cases shift to case (i) from case (ii), which indicates a higher need for solving multistage models. We also present the results when we set $\delta_1 = 5\%$, $\delta_2 = 20\%$, and change other parameters in Online Appendix D.

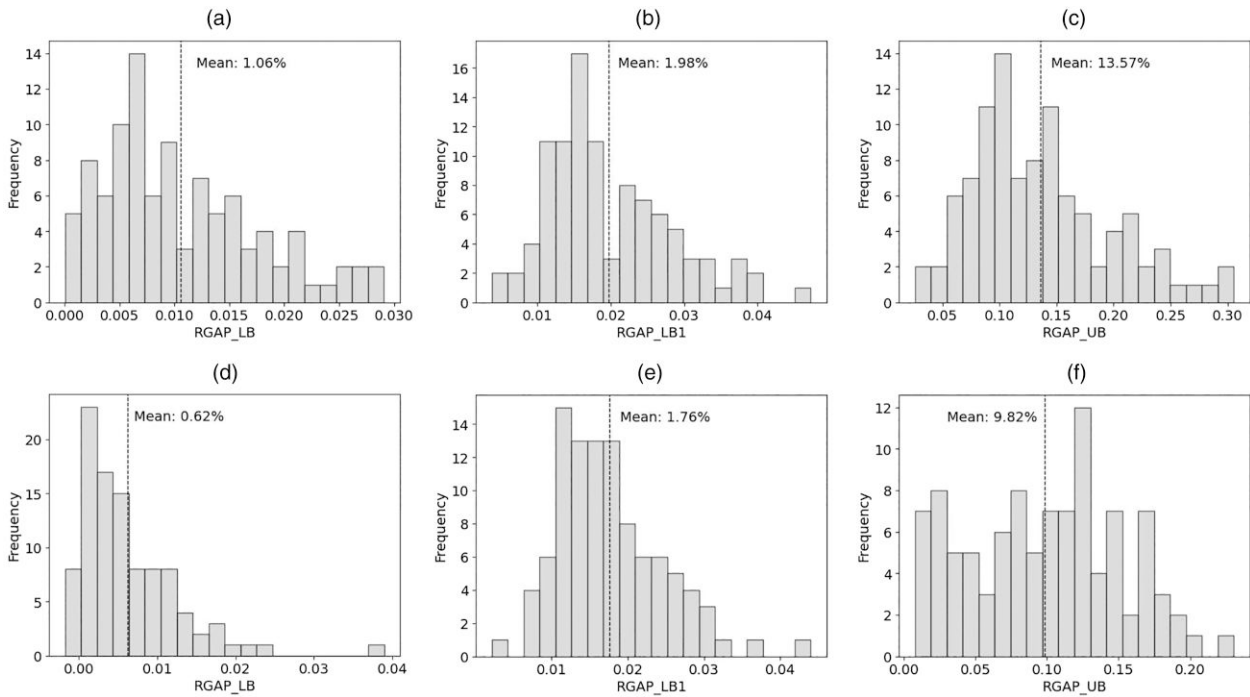
5.1.3. Tightness of the Analytical Bounds on the Synthetic Data Set. To examine the tightness of the analytical bounds derived in Theorems 1 and 2 and Corollary 1, we define the relative gaps of the analytical bounds as $RGAP_{LB} = RVMS - RVMS_{LB}$, $RGAP_{LB1} = RVMS - RVMS_{LB1}$, $RGAP_{UB} = RVMS_{UB} - RVMS$, respectively. A lower $RGAP$ means that the analytical bound recovers the true VMS better. Using the default setting and the two scenario trees defined in Section 5.1.1, we plot the histograms of $RGAP_{LB}$, $RGAP_{LB1}$, $RGAP_{UB}$ over 100 independently generated instances in Figure 5. From the figure, $RGAP_{LB}$ with SI scenario trees obtains the lowest $RGAP$ mean (0.62%), and in most instances, $RGAP_{LB}$ is within 1%. Stagewise-dependent scenario trees obtain a slightly higher mean of $RGAP_{LB}$. This is because $RVMS$ in stagewise-dependent scenario trees is much higher than the stagewise-independent counterparts (see Figure 4), and as a result, the optimal solutions of the two-stage models are farther away from the optimal ones of the multistage models. As can be seen from the proof of Theorem 1, the tightness of the lower bounds depends on the gap between these two optimal solutions. On the other hand, VMS_R^{LB1} produces a slightly higher $RGAP$ with a mean of 1.98% in the SD scenario trees and 1.76% in the SI scenario trees. Compared with the tightness of lower bounds, VMS_R^{UB} is much looser with an $RGAP$ mean of 13.57% in SD scenario trees and 9.82% in SI scenario trees.

5.1.4. Performance of the Approximation Algorithm on the Synthetic Data Set. Next, we vary the number of branches C from 2 to 5, number of stages T from 3 to 6, number of facilities M from 5 to 20, and number of customer sites N from 10 to 40 to see how the empirical approximation ratio changes with respect to different

Table 2. Percentage of Instances in Different Cases with SD Scenario Trees and Varying Numbers of Stages T When $\delta_1 = 10\%$, $\delta_2 = 30\%$

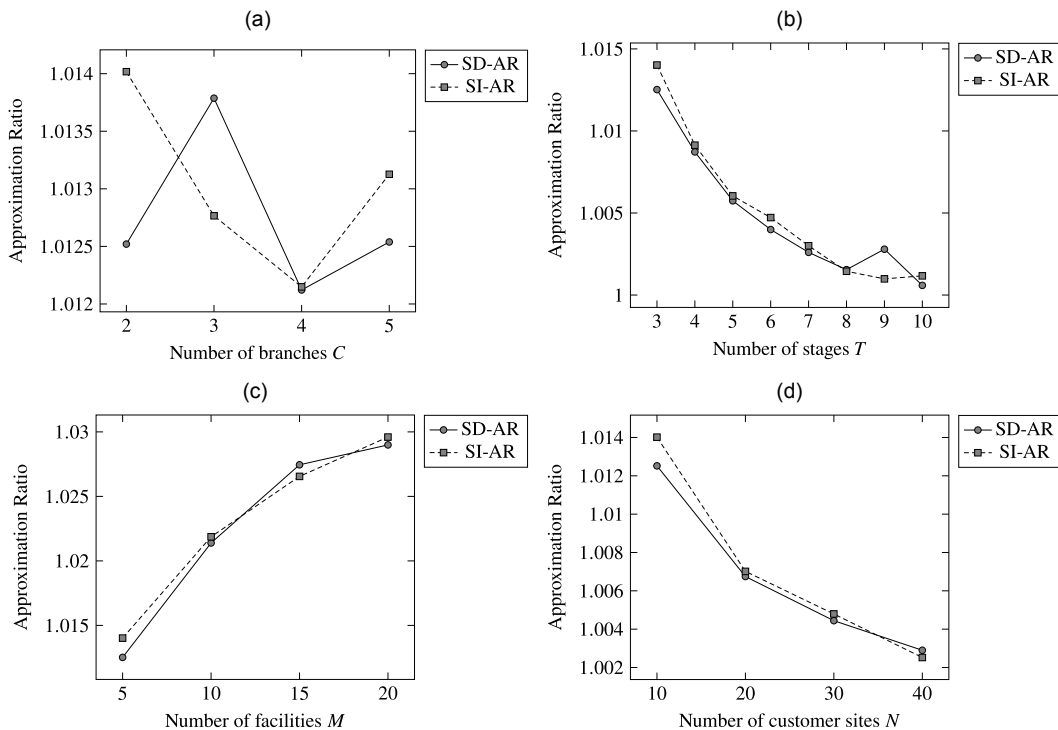
Stages T	Case (i)	Case (ii)	Case (iii)
3	20/100	73/100	7/100
4	40/100	35/100	25/100
5	85/100	5/100	10/100
6	95/100	0/100	5/100
7	98/100	0/100	2/100
8	100/100	0/100	0/100

Figure 5. Tightness of Analytical Bounds with SD and SI Scenario Trees



Notes. (a) RGAP_{LB} with SD trees. (b) RGAP_{LB1} with SD trees. (c) RGAP_{UB} with SD trees. (d) RGAP_{LB} with SI trees. (e) RGAP_{LB1} with SI trees. (f) RGAP_{UB} with SI trees.

Figure 6. Statistics of the Approximation Ratios over 100 Instances with Different Numbers of Branches C , Stages T , Risk Attitudes λ , Demand Standard Deviations σ , Facilities M , and Customer Sites N



Notes. (a) Different numbers of branches. (b) Different numbers of stages. (c) Different numbers of facilities. (d) Different numbers of customer sites.

Table 3. Computational Time Comparison with SI Scenario Trees and Different Numbers of Branches C

C	z_R^{TS} via Gurobi		z_R^{MS} via Gurobi		z_R^{MS} via AA			z_R^{MS} via SDDiP		
	Time (s)	Obj. (\$)	Time (s)	Obj. (\$)	Time (s)	Obj. (\$)	Gap	Time (s)	Obj. (\$)	Gap
2	1.46	13,482K	3.97	13,330K	0.13	13,592K	1.97%	90.06	13,270K	−0.45%
3	6.40	21,714K	3,600 (0.02%)	20,137K	0.31	20,474K	1.67%	679.13	20,098K	−0.19%
4	3.55	24,363K	3,600 (0.11%)	22,322K	0.43	22,712K	1.75%	2,102.27	22,272K	−0.22%
5	10.01	24,142K	3,600 (0.19%)	22,520K	0.66	22,907K	1.72%	3,430.90	22,467K	−0.24%

Note. Bold values are the smallest values for the computational time.

parameter settings. The results are presented in Figure 6, in which we plot the mean of empirical approximation ratios (i.e., $\frac{z_R^{MS}(x_n^H, \eta_n^H, y_n^H, u_n^H)}{z_R^{MS}(x_n^H, \eta_n^H, y_n^H, u_n^H)}$) over 100 independently generated instances.

From the figure, the approximation ratios decrease gradually when we increase the number of branches C and the number of stages T (see Figure 6(a) and (b)). Moreover, from Figure 6(c) and (d), the approximation ratios are clearly positively related to the number of facilities M and negatively impacted by the number of customer sites N , which is because increasing N would increase the total demand and the number of units of resources needed in the first stage ($M_{\min}$) and our approximation ratio has an upper bound $1 + \frac{M \sum_{t=1}^T f_{t, \max}}{M_{\min} \sum_{t=1}^T f_{t, \min} + \sum_{t=1}^T c_{t, \min} \min_{n \in T_t} \{\sum_{j=1}^N d_{n,j}\}}$. In all of these instances, the approximation algorithm achieves an approximation ratio of at most 1.03.

5.1.5. Computational Time Comparison. We end this section by comparing the computational time of two-stage and multistage models solved via Gurobi and multistage models solved by our AAs (Algorithm 1). We also use the Python package (Ding et al. 2019) to solve the multistage model via SDDiP algorithm (Zou et al. 2019) and terminate the algorithm after 10 iterations when the objective values become stable. Because the SDDiP algorithm assumes that the underlying stochastic process is stagewise independent, we only consider SI scenario trees when evaluating their performance in this section. We note that our approximation algorithms can be also applied to solve SD scenario trees, and we present the computational time comparison between our approximation algorithm and directly solving the problem using Gurobi 10.0.0 under SD scenario trees in Online Appendix D. We fix the risk parameters $\lambda = 0.5$, $\alpha = 0.95$ and the standard deviation $\sigma = 0.8$, varying the number of branches C from 2 to 5 in Table 3, number of stages T from 3 to 6 in Table 4, respectively. The percentages shown in parenthesis are the optimality gaps produced by Gurobi within the one-hour time limit. We also display the gaps of the optimal objective value between our approximation algorithm/SDDiP and Gurobi in the column “Gap.” Note that we solve step 5 for all node $n \in T$ as a whole in Algorithm 1, and one may further speed up the approximation algorithms by utilizing parallel computing techniques. From the tables, we observe that the approximation algorithms scale very well with respect to the problem size, whereas Gurobi cannot solve larger scale risk-averse multistage models to optimality within the one-hour time limit. Compared with multistage models, two-stage models are less computationally expensive and, therefore, can be preferred over multistage models when their objective gaps are relatively small. Whereas our approximation algorithms provide an upper bound on the multistage model by producing a feasible solution, the SDDiP algorithm returns a lower bound on the multistage model by constructing under-approximations of the value functions. As the problem size increases, the SDDiP algorithm fails to produce a high-quality lower bound within several hours (with the highest gap being 17.02%), whereas our approximation algorithms can always find a high-quality upper bound within several seconds (with all the gaps below 2%).

Table 4. Computational Time Comparison with SI Scenario Trees and Different Numbers of Stages T

T	z_R^{TS} via Gurobi		z_R^{MS} via Gurobi		z_R^{MS} via AA			z_R^{MS} via SDDiP		
	Time (s)	Obj. (\$)	Time (s)	Obj. (\$)	Time (s)	Obj. (\$)	Gap	Time (s)	Obj. (\$)	Gap
3	1.46	13,482K	3.97	13,330K	0.13	13,592K	1.97%	90.06	13,270K	−0.45%
4	8.49	26,484K	3,600 (0.18%)	26,212K	0.67	26,568K	1.36%	847.48	26,101K	−0.42%
5	3,600 (0.05%)	44,003K	3,600 (0.14%)	43,205K	1.07	43,623K	0.97%	1,028.06	43,063K	−0.33%
6	83.80	63,440K	3,600 (0.17%)	62,624K	6.97	63,177K	0.88%	1,108.69	62,467K	−0.25%
7	640.56	138,372K	3,600 (0.13%)	112,508K	5.80	113,207K	0.62%	5,556.12	93,360K	−17.02%
8	3,600 (0.01%)	201,522K	3,600 (0.14%)	150,002K	14.64	150,718K	0.48%	13,859.29	125,095K	−16.60%

Note. Bold values are the smallest values for the computational time.

Table 5. Demand Patterns

Pattern	Distribution at stage t
I. Constant mean, constant standard deviation	$\mathcal{N}(\tilde{d}, \tilde{\sigma})$
II. Constant mean, increasing standard deviation	$\mathcal{N}(\tilde{d}, \tilde{\sigma} \cdot (\sigma + 2(t-1)))$
III. Increasing mean, constant standard deviation	$\mathcal{N}(\tilde{d} \cdot (1 + 2(t-1)), \tilde{\sigma})$
IV. Increasing mean, increasing standard deviation	$\mathcal{N}(\tilde{d} \cdot (1 + 2(t-1)), \tilde{\sigma} \cdot (\sigma + 2(t-1)))$

5.2. Case Study on Real-World EV Charging Station Capacity Planning

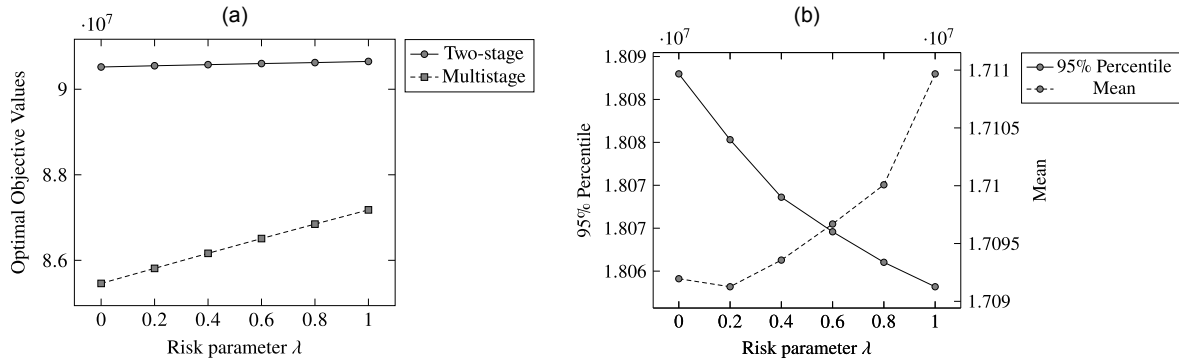
5.2.1. Experimental Design and Setup. We consider the 49-node and 88-node data sets described in Daskin (2011) with $M = 49$, $N = 88$, $T = 5$, $C = 2$, where each stage is one year. The maintenance cost for each charger is set to \$100 per year (EV Connect 2022). The capacity of each charger h_{it} is set to 6×360 , assuming that each charger can charge six EVs from empty per day (U.S. Department of Transportation 2022). The resource allocation costs are set equal to the great-circle distance times the travel cost per mile per unit of demand, that is, $c_{ij} = \text{dist}(i, j) \times 0.00001$. We then collect the population data in each city and multiply them by 6% times 120 days, assuming that 6% of the population currently own EVs (Smith 2022) and charge their EVs every three days (BloombergNEF 2021), which gives nominal demand $\tilde{d}$ in each customer site in the beginning year of the planning horizon. We consider four demand patterns (described in column “Pattern” in Table 5), all of which follow truncated normal distributions. In patterns III and IV, the nominal demand $\tilde{d}$ is increased with a rate of 200% for each subsequent year, which reaches $6\% \times (1 + 2 \times 4) = 54\%$ of EV market share by the end of the fifth year, matching President Biden’s goal of having 50% of all new vehicle sales be electric (The White House 2023); in patterns II and IV, the standard deviation is increased with the same rate.

5.2.2. Result Comparison with Different Demand Patterns. With the baseline setting $C = 2$, $\sigma = 0.8$, $\lambda_t = 0.5$, $\alpha_t = 0.95$, and SD scenario trees, we present the optimal solutions and cost breakdown of two-stage and multistage models under different demand patterns in Table 6. We use Gurobi to solve the risk-averse TS and MS models directly, and we compare these two models with the multistage models solved via AA. Columns “ $|x_1|$,” “ $|x_2|$,” “ $|x_3|$,” “ $|x_4|$,” and “ $|x_5|$ ” display the average number of chargers installed in each stage across all scenarios. Columns “Maintenance (\$)” and “Operational (\$)” show the maintenance and operational cost without considering the risk parameters, and column “Obj. (\$)” presents the overall risk-averse optimal objective values, in which we mark the lowest ones among the three models in bold. The last column “RVMS/AR” shows the relative VMS or the approximation ratios. Comparing the three models, multistage models solved by Gurobi always achieve the least maintenance cost because they have more flexibility in deciding the capacity expansion plans. In terms of operational cost, the multistage models solved by approximation algorithms always achieve the minimum among the three. In all of the cases, our approximation algorithm aligns with the optimal solutions and objective values of the multistage models very well and achieves an approximation ratio of at most 1.00004.

We also find that the demand patterns with increasing standard deviation always obtain higher RVMS_R compared with the ones with constant standard deviation, which agrees with our findings in Section 5.1.2. In pattern I, most chargers are invested in the first stage because the demand mean is constant and we do not need to build many more chargers in later stages. Under this setting, we achieve an RVMS of 5.19%. In patterns III and IV in

Table 6. Optimal Solutions and Costs of Two-Stage and Multistage Models Under Different Demand Patterns

Pattern	Model	$ x_1 $	$ x_2 $	$ x_3 $	$ x_4 $	$ x_5 $	Maintenance (\$)	Operational (\$)	Obj. (\$)	Time (s)	RVMS/AR
I	TS	149,502	33,681	3,845	16,686	20,869	\$94,801K	\$2,368K	\$97,336K	8.90	5.19%
	MS	149,489	29,303	961	7,018	6,509	\$88,808K	\$2,492K	\$92,284K	8.94	
	AA	149,489	29,315	962	7,018	6,509	\$88,813K	\$2,491K	\$92,288K	58.72	1.00004
II	TS	149,523	261,128	265,695	389,223	488,955	\$385,661K	\$9,891K	\$396,742K	8.68	14.82%
	MS	149,469	253,004	199,195	279,369	266,388	\$318,207K	\$9,062K	\$337,928K	57.36	
	AA	149,494	252,997	199,198	279,371	266,388	\$318,218K	\$9,062K	\$337,939K	56.65	1.00003
III	TS	149,531	399,975	385,524	490,802	595,216	\$508,095K	\$12,256K	\$521,440K	8.81	12.02%
	MS	149,469	386,866	305,341	384,100	374,120	\$435,315K	\$11,303K	\$458,753K	750.35	
	AA	149,495	386,858	305,345	384,101	374,120	\$435,326K	\$11,303K	\$458,763K	59.20	1.00002
IV	TS	149,528	1,082,360	2,149,745	4,075,429	6,453,154	\$2,613,033K	\$62,620K	\$2,683,077K	9.00	18.61%
	MS	149,470	1,057,952	1,800,916	3,158,538	3,797,729	\$2,049,671K	\$55,468K	\$2,183,774K	67.52	
	AA	149,495	1,057,946	1,800,916	3,158,537	3,797,730	\$2,049,681K	\$55,467K	\$2,183,784K	56.94	1.000005

Figure 7. Optimal Objective Value and Stagewise Cost Comparison Between Two-Stage and Multistage Models with Varying Risk Attitude λ 

Notes. (a) Optimal objective values in two models. (b) Stagewise cost in multistage models.

which the demand mean is increased at a rate of 200%, more chargers are installed in later stages, and the RVMS increases to 12.02% and 18.61%, respectively. Note that this gap can be further increased with more branches C and higher standard deviations σ .

5.2.3. Effect of the Risk Parameter. Finally, we discuss the effect of the risk attitude on the optimal solutions. We display the optimal objective values of two-stage and multistage models with varying risk attitude λ in Figure 7(a) and the 95% percentile and mean of the stagewise cost in multistage models with varying risk attitude λ in Figure 7(b). From Figure 7(a), the optimal objective values of both two-stage and multistage models increase as we become more risk averse, agreeing with our results in Proposition 1. Moreover, the gap between the two models reduces as we increase λ as shown in Section 5.1.2, Figure 4(c). From Figure 7(b), when we increase λ , the mean of the stagewise cost increases, whereas the 95% percentile decreases in most cases, representing a switch from a risk-neutral to risk-averse attitude.

6. Conclusion

We consider a general class of multiperiod capacity planning problems under uncertain demand in each period. We compare a multistage stochastic programming model in which the capacities of facilities can be determined dynamically throughout the uncertainty realization process with a two-stage model in which decision makers have to fix capacity acquisition at the beginning of the planning horizon. Using ECRMs, we bound the gaps between the optimal objective values of risk-averse two-stage and multistage models from below and above and provide an example to show that one of the lower bounds is tight. We also propose an approximation algorithm to solve the risk-averse multistage models more efficiently, which is asymptotically optimal in an expanding market. Our numerical tests indicate that the RVMS increases as the uncertainty variability increases and decreases as the decision maker becomes more risk averse. Moreover, stagewise-dependent scenario trees attain higher RVMS than the stagewise-independent counterparts. On the other hand, the analytical lower bounds can recover the true gaps very well. Moreover, the approximation ratios are pretty close to one in the case study based on a real-world network.

There are several interesting directions to investigate for future research. The risk measure we use in this paper is the ECRM, of which the risk is measured separately for each stage. There are other risk measure choices that can be applied here, such as the nested risk measures. More research can be done to explore the relationship between these two models under different choices of risk measures. Moreover, this paper assumes that the uncertainty has a known distribution, whereas it is more realistic to assume that the probability distribution of the uncertainty belongs to an ambiguity set. Therefore, a distributionally robust optimization framework can be considered for two-stage and multistage capacity planning problems. As a middle ground between two-stage and multistage decision frameworks, Basciftci et al. (2019) study an adaptive two-stage stochastic programming framework, in which each component of the decision policy has its own revision point, before which the decisions are determined until this revision point, and after which they are revised for adjusting to the uncertainty. They analyze this approach on a capacity expansion planning problem and derive bounds on the value of the proposed adaptive two-stage stochastic programming. Similarly, we can obtain a relaxation of the multistage

stochastic program by ignoring the nonanticipativity constraints at some nonroot stages, and such a relaxation provides a lower bound to the multistage model. In this paper, we focus on the comparison between risk-averse two-stage and multistage stochastic programs and leave the exploration of general adaptive stochastic programs as future research.

Acknowledgments

The authors are grateful for the constructive feedback from the editorial team and two anonymous reviewers.

References

- Aghezzaf E (2005) Capacity planning and warehouse location in supply chains with uncertain demands. *J. Oper. Res. Soc.* 56(4):453–462.
- Ahmed S (2006) Convexity and decomposition of mean-risk stochastic programs. *Math. Programming* 106(3):433–446.
- Ahmed S, Sahinidis NV (2003) An approximation scheme for stochastic integer programs arising in capacity expansion. *Oper. Res.* 51(3):461–471.
- Artzner P, Delbaen F, Eber JM, Heath D (1999) Coherent measures of risk. *Math. Finance* 9(3):203–228.
- Basciftci B, Ahmed S, Gebräel N (2019) Adaptive two-stage stochastic programming with an application to capacity expansion planning. Preprint, submitted June 8, <https://arxiv.org/abs/1906.03513>.
- Bayram IS, Tajer A, Abdallah M, Qaraqe K (2015) Capacity planning frameworks for electric vehicle charging stations with multiclass customers. *IEEE Trans. Smart Grid* 6(4):1934–1943.
- Berman O, Ganz Z, Wagner JM (1994) A stochastic optimization model for planning capacity expansion in a service industry under uncertain demand. *Naval Res. Logist.* 41(4):545–564.
- BloombergNEF (2021) EV drivers save big when they charge smart. *Online* (October 26), <https://about.bnef.com/blog/ev-drivers-save-big-when-they-charge-smart/>.
- Caldwell T (2022) The five pillars of EV charging station maintenance and management. *Forbes Online* (July 29), <https://www.forbes.com/sites/forbestechcouncil/2022/07/29/the-five-pillars-of-ev-charging-station-maintenance-and-management/?sh=5d7ee8d931fa>.
- Connect EV (2022) Estimated cost guide to operating and maintaining EV charging stations. *evconnect* (June 20), <https://www.evconnect.com/blog/ev-charging-station-infrastructure-costs>.
- Daskin MS (2011) *Network and Discrete Location: Models, Algorithms, and Applications* (John Wiley & Sons, Hoboken, NJ).
- Davis MH, Dempster M, Sethi S, Vermes D (1987) Optimal capacity expansion under uncertainty. *Adv. Appl. Probab.* 19(1):156–176.
- Ding L, Ahmed S, Shapiro A (2019) A python package for multi-stage stochastic programming. *Optimization Online* (May 7), <https://optimization-online.org/2019/05/7199/>.
- Eppen GD, Martin RK, Schrage L (1989) OR practice—A scenario approach to capacity planning. *Oper. Res.* 37(4):517–527.
- Fine CH, Freund RM (1990) Optimal investment in product-flexible manufacturing capacity. *Management Sci.* 36(4):449–466.
- Girardeau P, Leclerc V, Philpott AB (2014) On the convergence of decomposition methods for multistage stochastic convex programs. *Math. Oper. Res.* 40(1):130–145.
- Guigues V (2016) Convergence analysis of sampling-based decomposition methods for risk-averse multistage stochastic convex programs. *SIAM J. Optim.* 26(4):2468–2494.
- Homem-de Mello T, Pagnoncelli BK (2016) Risk aversion in multistage stochastic programming: A modeling and algorithmic perspective. *Eur. J. Oper. Res.* 249(1):188–199.
- Hood SJ, Bermon S, Barahona F (2003) Capacity planning under demand uncertainty for semiconductor manufacturing. *IEEE Trans. Semiconductor Manufacturing* 16(2):273–280.
- Huang K, Ahmed S (2009) The value of multistage stochastic programming in capacity planning under uncertainty. *Oper. Res.* 57(4):893–904.
- Kleywegt AJ, Shapiro A, Homem-de Mello T (2002) The sample average approximation method for stochastic discrete optimization. *SIAM J. Optim.* 12(2):479–502.
- Luss H (1982) Operations research and capacity expansion problems: A survey. *Oper. Res.* 30(5):907–947.
- Miller N, Ruszczyński A (2011) Risk-averse two-stage stochastic linear programming: Modeling and decomposition. *Oper. Res.* 59(1):125–132.
- Paraskevopoulos D, Karakitsos E, Rustem B (1991) Robust capacity planning under uncertainty. *Management Sci.* 37(7):787–800.
- Pereira MV, Pinto LM (1991) Multi-stage stochastic optimization applied to energy planning. *Math. Programming* 52(1–3):359–375.
- Pflug GC, Ruszczyński A (2005) Measuring risk for income streams. *Comput. Optim. Appl.* 32(1–2):161–178.
- Philpott AB, Guan Z (2008) On the convergence of stochastic dual dynamic programming and related methods. *Oper. Res. Lett.* 36(4):450–455.
- Rehman A (2020) The importance of equipment maintenance plan for electrical power systems. *AllumiaX* (September 28), <https://www.allumiax.com/blog/the-importance-of-equipment-maintenance-plan-for-electrical-power-systems>.
- Rockafellar RT, Uryasev S (2000) Optimization of conditional value-at-risk. *J. Risk* 2(3):21–42.
- Rockafellar RT, Uryasev S (2002) Conditional value-at-risk for general loss distributions. *J. Banking Finance* 26(7):1443–1471.
- Ruszczyński A (2010) Risk-averse dynamic programming for Markov decision processes. *Math. Programming* 125(2):235–261.
- Sabet E, Yazdani B, Kian R, Galanakis K (2020) A strategic and global manufacturing capacity management optimisation model: A scenario-based multi-stage stochastic programming approach. *Omega* 93:102026.
- Schultz R, Tiedemann S (2006) Conditional value-at-risk in stochastic programs with mixed-integer recourse. *Math. Programming* 105(2–3):365–386.
- Shapiro A (2011) Analysis of stochastic dual dynamic programming method. *Eur. J. Oper. Res.* 209(1):63–72.
- Shapiro A, Dentcheva D, Ruszczyński A (2009) *Lectures on Stochastic Programming: Modeling and Theory* (SIAM, Philadelphia).
- Shapiro A, Tekaya W, da Costa JP, Soares MP (2013) Risk neutral and risk averse stochastic dual dynamic programming method. *Eur. J. Oper. Res.* 224(2):375–391.

- Singh KJ, Philpott AB, Wood RK (2009) Dantzig-Wolfe decomposition for solving multistage stochastic capacity-planning problems. *Oper. Res.* 57(5):1271–1286.
- Smith S (2022) What percentage of US car sales are electric? *Way*, <https://www.way.com/blog/what-percentage-of-us-car-sales-are-electric/>.
- Swaminathan JM (2000) Tool capacity planning for semiconductor fabrication facilities under demand uncertainty. *Eur. J. Oper. Res.* 120(3):545–558.
- The White House (2023) Fact Sheet: Biden-Harris administration announces new private and public sector investments for affordable electric vehicles. *The White House* (April 17), <https://www.whitehouse.gov/briefing-room/statements-releases/2023/04/17/>.
- U.S. Department of Transportation (2022) Electric vehicle charging speeds. *U.S. Department of Transportation*, <https://www.transportation.gov/rural/ev/toolkit/ev-basics/charging-speeds>.
- Wasserman L (2004) *All of Statistics: A Concise Course in Statistical Inference*, vol. 26 (Springer, New York).
- Yu X, Shen S (2024) On the value of risk-averse multistage stochastic programming in capacity planning. <http://dx.doi.org/10.1287/ijoc.2023.0396.cd>, <https://github.com/INFORMSJoC/2023.0396>.
- Zou J, Ahmed S, Sun XA (2019) Stochastic dual dynamic integer programming. *Math. Programming* 175(1–2):461–502.