Marla, Lavanya; Rikun, Alexander; Stauffer, Gautier; Pratsini, Eleni

Article

Robust modeling and planning: Insights from three industrial applications

Operations Research Perspectives

Provided in Cooperation with:
Elsevier

Operations Research Perspectives 7 (2020) 100150

Contents lists available at ScienceDirect

Operations Research Perspectives

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

Robust modeling and planning: Insights from three industrial applications

Lavanya Marla[a,*], Alexander Rikun[b], Gautier Stauffer[c], Eleni Pratsini[d]

[a] University of Illinois at Urbana-Champaign, United States
[b] Analysis Group, United States
[c] Center of Excellence in Supply Chain Innovation and Transportation, Kedge Business School Talence, France
[d] Accenture Technology, Switzerland

ABSTRACT

Optimization under uncertainty has been a well-studied field, with significant interest generated in this field in the past four decades. This paper is both practical and expository – its purpose is to: discuss the process of generating robust solutions, highlight issues that arise in *practice,* and discuss ways to address such issues. For illustrative purposes, we study three different, commonly adopted, approaches for optimization under uncertainty (chance-constrained programming, robust optimization and conditional value at risk); and apply these approaches to three real-world application-based case studies. Our case studies are chosen to span a variety of problem characteristics. For each case study, we discuss the applicability of each of the three approaches, practical issues that arose during application, and robustness and further characteristics of the subsequent solutions. We point out associated advantages and limitations, and illustrate the gap between the theoretical and actual performance of these approaches for each case study. We also discuss how some of the discovered limitations can be overcome using extensions of the approaches or through a better understanding of the data. We conclude by summarizing common and generalizable insights obtained across the three case studies. Our findings suggest the effectiveness of solutions is dependent on: the methods, the size of the problem, the underlying pattern of uncertainty in data, and the metrics of interest. While we provide some guidelines to identify the most suitable approach to a given problem, our experience matches theory to suggest that under carefully tuned parameters accompanied by simulation, the different approaches can generate results that are similar and provide comparable tradeoffs between the mean and robustness metric. However, this could also require considerable tuning requiring experience, and we provide some guidelines to achieve such results. This illustrates that generating high quality robust solutions is both an art and a science.

1. Introduction

Real world systems are routinely and inevitably subject to uncertainty. Solutions built without considering possible future uncertainty, usually using the average values of parameter realizations (the corresponding model is usually referred to the *nominal problem*), are often subject to the well-known flaws of averages: such solutions, when implemented in practice can be sub-optimal, and even infeasible when the uncertainty in the system is actually realized, resulting in significantly higher costs. Therefore, one should proactively plan for future uncertainty and associated costs [1–3]. The susceptibility of a system to uncertainty depends on the level of uncertainty, the timeline of the decision, and the magnitude of the costs incurred to “recover from non-favorable events”. Methodologies for optimization under uncertainty allow decision makers and managers to attain a solution that is less sensitive to variability by incorporating future uncertainty a priori: this allows to move from solutions that essentially minimize costs induced by average scenarios, to solution that accounts for future *true costs associated with all events under consideration,* and to deal in particular with potential *solution infeasibility*.

In this work, we refer to a *robust* plan as one that captures the uncertainty of alternative future scenarios and generate solutions that address future uncertainty with an acceptable quality for all scenarios [1]. Note that, while we use the term ‘scenarios’ this does not always refer to an explicit enumeration of scenarios, but a possible space of parameter realizations. The impact of capturing uncertainty, and consequently, the impact of providing robust solutions, can be enormous. In the past two decades, therefore, there has been an exponential growth in developing uncertainty modeling approaches in the context of numerous applications, including finance [4–8], revenue management [9–14], queueing theory [15–18], transportation [19–23], project management [24], scheduling [25], network optimization [26,27], inventory and supply chain management [28,29] and energy grids [30–33].

In modeling uncertainty in real-world problems, decision-makers face multiple challenges, such as:

- How does one define robustness? How can we model it in different contexts?
- How applicable are classic approaches for optimization under

* Corresponding author.
E-mail addresses: lavanyam@illinois.edu (L. Marla), alexander.rikun@gmail.com (A. Rikun), gautier.stauffer@kedgebs.com (G. Stauffer), eleni.pratsini@gmail.com (E. Pratsini).

https://doi.org/10.1016/j.orp.2020.100150
Received 26 August 2019; Received in revised form 26 January 2020; Accepted 5 March 2020
Available online 14 March 2020
2214-7160/

uncertainty?

- What type of knowledge of the underlying data uncertainty is required for a particular approach to be applied?
- How does more information/data about a system help generate better models and find more robust solutions?
- How can we assess/evaluate ‘true’ or ‘realized’ performance of a solution – both in terms of the objective function and the robustness of the solution under uncertainty?
- Can we a priori know which approach can yield the best solution under uncertainty for a given problem?
- Practically, given a limited budget for investment in software packages for mathematical modeling or coding/implementation time in-house, what types of models and/or frameworks should a company invest in for the long term?

The objective of this work is to help managers and decision-makers to address the above questions, by illustrating the application of three main paradigms for modeling uncertainty, through three applications. The three major paradigms we consider are: (a) chance-constrained programming (CCP), from the field of stochastic programming, (b) the robust optimization (RO) method of Bertsimas and Sim, from the framework of worst-case-based robust optimization, and (c) Conditional Value-at-Risk (CVaR), a risk-mitigation approach. While there have been a number of new approaches recently, as we discuss in Section 2, many are based on bridging the above three approaches. Hence, we believe that focusing our attention on, and taking lessons from practical applications of, these three fundamental paradigms should be relevant to the decision-makers, even in contexts where he might consider hybrid methods. Our focus is on real-world settings of single-stage (static problems without explicit recourse), linear and mixed-integer problems, that is, problems whose nominal form is “*Minimize cx over all x in* R^n *such that* $Ax \leq b$*, where some of the components of x are restricted to be integers”,* but where *b, c* and/or *A* are subject to uncertainty. We consider settings from corporate portfolio optimization, pharmaceutical supply chain management and aircraft routing, which we discuss in further detail in Sections 3, 4 and 5. We illustrate our insights on these three case studies, chosen for the different manifestations of uncertainty and different problem sizes – each allowing for experimentation of different methods, depending on applicability. All these case studies are drawn from real-world data, however, some of the exact quantities and cost parameters are masked for confidentiality purposes. For each case study, we discuss modeling, tractability and solution quality issues for each method, and provide recommendations to facilitate the deployment of these methods, and robust planning in general, in practice.

1.1. Contributions and outline

The contributions of our work are as follows.

We present a discussion of the key issues in choosing an approach for modeling uncertainty, namely: (i) a method's definition of uncertainty, including its managerial and statistical interpretability; (ii) metrics of robustness used by the method and their measurement; (iii) tractability issues arising in the generation of robust solutions; and (iv) evaluation of the performance metrics of robustness.

Our results should help practitioners in addressing the following challenges: (i) capturing uncertainty in a mathematical form using various approaches, by understanding the specific way each approach makes tradeoffs between the objective function and its specific robustness metrics; (ii) modeling uncertainty when uncertain parameters are poorly defined; and (iii) setting robustness parameters in these approaches for real implementation in order to better control performance and tractability.

Our work also provides recommendations to decision-makers on how investments should be targeted. Our experience shows that building robust solutions involves substantial investment in time and money in order to collect data, develop a simulation platform and investigate the applicability of robust planning methods; specifically, in that order.

In particular, we find that *simulation frameworks* are fundamental to investigate robustness of solutions, and are an essential and very beneficial initial investment for the managers. Because each method has different metrics of robustness (that might differ significantly from the manager's conceptualization of robustness), solutions from these approaches often require an independent manner of performance evaluation, more aligned with the manager's or organization's Key Performance Indices (KPIs), typically through simulation, to truly measure performance along both cost and robustness objectives.

We also find that it is difficult to *a priori* predict which method can produce the solution with the best performances for any general problem. We demonstrate that the applicability and success of each of the methods depends greatly on the problem type, size, and the available data. Therefore, while we do not recommend any one method as the ‘best’; we provide guidelines to using these approaches, and discussions of extended methods that can address some limitations of these core approaches. We also recommend *practical* investigation of ongoing theoretical developments that bridge the strengths of these approaches.

Outline: The remainder of this paper is structured as follows. In Section 2, we discuss background related to the three key methodologies of interest. In Sections 3, 4 and 5 we discuss each of the individual case-studies, and discuss our findings from applying the three methods to these approaches. Wherever applicable, we also provide pointers to alternative methods that address some of the issues encountered with each approach. In Section 6, we summarize our lessons from all three case-studies and conclude with general insights.

2. Background

Optimization under uncertainty was first considered in the 1950s with Dantzig [34] and Beale [35] and has since stimulated a large body of research [36]. The last three decades have seen a significantly increased growth in this field, motivated by the increased realized costs from the practical implementation of deterministic solutions, the availability of data, and increased computational power. While a number of approaches to modeling uncertainty to obtain robust solutions exist and many have been recently proposed, they belong to two main families: the first assuming the exact probability distributions of the uncertain parameters (distribution-based), and the second free of such specific distributional assumptions (distribution-free). Some approaches, as we discuss below, attempt to bridge the two families. Nevertheless, in practice, three important approaches have been most employed, namely: chance-constrained programming, robust optimization, and conditional value at risk.

Chance Constrained Programming (CCP) belongs to the family of approaches typically referred to as stochastic programming or distribution-based approaches. Stochastic programming, first pioneered by Dantzig [34], aims to find a policy that is feasible for all (or almost all) possible data instances, while maximizing the expectation of some function of the decisions and the random variables. Formulating stochastic programs require exact knowledge of, or assumptions on, the data distributions underlying uncertain parameters. The most widely applied and studied stochastic programming models are two-stage linear programs, referred to as *stochastic programming with recourse*, which allow for some decisions to be made before the realizations of uncertain parameters are known, and others after the uncertain parameters are realized. We refer the reader to Birge and Louveaux [37] and Kall and Mayer [38] for details.

Chance-constrained programming was originally developed by Charnes and Cooper [39,40] to deal with uncertainty on the constraints $Ax \leq b$ in the nominal problem (the objective function could actually be turned into a constraint as well and studied through this framework too). This approach regulates the level of robustness of the solutions via a maximum requirement on the *probability of violation* (α) of the constraint(s) in the system $Ax \leq b$ that are subject to uncertainty. More precisely, given a subsystem of the form $A'x \leq b'$, a user-specified probability α, and distribution assumptions on the uncertain parameters A' and b', a chance constraint (or set of chance-constraints) is a constraint of the form $\mathbb{P}(A'x \leq b') \geq 1 - \alpha$. The probability of constraint violation α is the *robustness measure* of the approach.

The notion of constraint violation is intuitive and easy to explain to a manager or a stakeholder. However, other than scenario generation methods via sampling, requiring extensive data [41,42], there is no single standard way of converting the probabilistic chance-constraint $\mathbb{P}(A'x \leq b') \geq 1 - \alpha$ into an equivalent deterministic formulation for all distributions of A' and b'. This conversion needs to be performed on a case-by-case basis, with different assumptions valid for different problems. Methods for making the probabilistic chance-constraint deterministic have been developed for specific problem settings, specific distributions of uncertain parameters, or for specific parameters in the mathematical program [40,41,43,44]. For instance, CCP is easy to formulate when the uncertainty is on the right-hand-side of a constraint (parameter b) [45]. In this case, full distribution information is not required and information about uncertainty quantiles is sufficient [20,40]. When uncertainty is in the left-hand-side parameters, a formulation using joint probability distributions of parameters may be needed [43,46,47], which is often difficult to formulate in the absence of large quantities of data, and may also prove intractable. Specific versions of these formulations are modeled as joint probabilistic constraints [48], taking advantage of special structures [49,50] and deterministic equivalents or approximations constructed via sample average approximations [51–53], strong valid inequalities [44,54], and efficient points of the distribution [55–58].

*Conditional Value-at-Risk CVaR*is a *risk measure* that is an extension of the well-known Value-at-Risk (VaR) measure. The latter originates from banking and insurance and quantifies the maximum loss if we exclude the α worst scenarios fraction, when implementing a solution (typically a portfolio in banking). That is, given the distribution of the data uncertainty, $VaR(\alpha, x)$ represent the (1-α) quantile of the loss function, if implementing solution x. $CVaR(\alpha, x)$ is in contrast the expected loss at the α-th probability level, that is, the conditional expected loss given that the loss exceeds $VaR(\alpha, x)$. Specifically, the approach considers a risk function $f(x, \zeta)$ defined with respect to decisions x and a system of input parameters ζ that are subject to uncertainty. For each x, the risk $f(x, \zeta)$ is a random variable having a distribution induced by that of ζ. The $CVaR$ is a function of the decision x and the α-th quantile of protection defined by the user, and is given by: $CVaR(\alpha,\ x) = \mathbb{E}(f(x, \zeta)|f(x, \zeta) \geq VaR(\alpha, x))$. CVaR can be obviously adapted to deal with profit/revenue maximization. In this case, VaR represents the minimum profit/revenue if we exclude the α worst scenarios fraction and $CVaR(\alpha,\ x) = \mathbb{E}(f(x, \zeta)|f(x, \zeta) \leq VaR(\alpha, x))$.

For a profit/revenue maximization problem, typical formulations involving *CVaR* either focus on (i) maximizing the *CVaR* value while imposing a minimum expected return, or (ii) enforcing the *CVaR* to be greater or equal to a user-specified critical value *c.v.* (which represents the practitioner's risk appetite) while maximizing the expected return. A larger expected value of the percentile risk (meaning larger *CVaR*) indicates a more robust solution. *CVaR*, unlike *VaR*, is a coherent measure of risk [59] and as such exhibits nice mathematical properties. In particular, Krokhmal, Palmquist and Uryasev [60] and Rockafellar and Uryasev [61] show that convex formulations exist and that *CVaR* minimization can be formulated as a linear program, via scenario-based constraints obtained through sampling from the data distribution. The core of the approach is to calculate the *VaR* and *CVaR* simultaneously while optimizing the objective [60]. A challenge with this approach is to determine the appropriate critical value that represents the practitioner's risk appetite to use in the formulation, because *c.v.* somewhat reflects C*VaR* itself. The determination of the 'right' *c.v.* is thus typically done iteratively (for example, to avoid infeasibility due to a too-high *c.v.*). *CVaR* has been applied as a risk-averse, distribution-based approach to many settings, with a focus on portfolio optimization and supply chain management [9–11,62–64].

Robust optimization (RO) was first suggested by Soyster [65] to handle data uncertainty in convex optimization problems without knowledge of underlying probability distributions. Soyster's model minimizes the maximum possible loss, that is, assumes each uncertain parameter in a convex programming problem to equal its worst-case value within an uncertainty set. Ben-Tal and Nemirovski [66,67] extended this idea into a powerful approach to deal with ellipsoidal uncertainty sets and parameterize the allowed maximum number of parameter deviations to consider for worst-case analysis. Bertsimas and Sim [68] exploited this approach further for solving linear and mixed-integer programs under uncertainty, and significantly enhanced the underlying theory in this context. The premise underlying [66,67] and [68] is that in general, it is highly unlikely that all parameters will assume their worst case values simultaneously. Bertsimas and Sim introduced a *robustness measure* Γ*,* which quantifies the number of parameters in the constraint that can deviate *simultaneously* to their worst-case values without affecting feasibility of the solution. Planning for a larger Γ, thus, indicates a plan with a greater degree of robustness. The strength of Bertsimas and Sim's approach is the robust counter part of a nominal problem that is a linear (or mixed-integer) program remains a linear (or mixed-integer) program, albeit possibly with a larger number of variables and constraints. The most basic formulation is associated with a nominal problem $\min c^T x\, s.t. Ax \leq b$, with uncertainty in the A parameters, such that each parameter a_{ij} assumes a realization $\tilde{a}_{ij}$, in the uncertainty set $[a_{ij} - \hat{a}_{ij}, a_{ij} + \hat{a}_{ij}]$, for some parameter $\hat{a}_{ij}$, where the underlying distribution of a_{ij} is unknown. The robust version of this formulation, with the Γ capturing the number of parameters that can assumed to realize their worst-case values simultaneously, is re-written (for integer values of Γ) as:

$$\min c^T x$$
$$s.\ t.\ \sum_j a_{ij}x_j + \max_{\{S_i | S_i \subseteq J_i, |S_i| = \Gamma_i\}} \{\hat{a}_{ij}|x_j|)\ \} \leq b_i \quad \forall \quad i$$

For details of the full formulation for non-integer Γ and its linearization, we refer the reader to Bertsimas and Sim [68]. Also, because the definition of the robustness measure Γ can be non-intuitive, the authors provide a method for relating Γ to the probability of constraint violation (similar to α in the CCP). Note that the solutions obtained from this method can be, obviously, largely influenced by the choice of the robustness measure Γ, as well as the choice of uncertainty sets $[a_{ij} - \hat{a}_{ij}, a_{ij} + \hat{a}_{ij}]$.

The original works of Ben-Tal and Nemirovski [69], Bertsimas and Sim [68], Ben-Tal et al. [70]; and more recently, Brown and Bertsimas [71] and Bandi and Bertsimas [16], also provide methods to construct good uncertainty sets, to appropriately match the robustness parameter Γ with the probability of constraint violation.

Some extensions: As we will also demonstrate through the case-studies, the CCP and the RO suffer from two key limitations. First, setting the value of the risk measure of the robust approach of interest – defined in terms of number of parameters (Γ), constraint violation (α) or tail risk (*CVaR*) – may not always be intuitive to the manager. Instead, the manager may be more comfortable specifying the target robustness in terms of the key performance indices – such as, maximizing the level of robustness within a specified loss in objective. Specifically, the method aims to maximize the probability of constraint satisfaction possible (called α) or maximize the number of parameters (called Δ) protected, both within the manager's specified budget δ. Second, the robustness metrics may not directly correlate with the `true' level of robustness achieved with respect to the manager's KPIs, usually measured through simulation. We discuss in detail two extensions of CCP and RO whose purpose is to overcome these two difficulties. The first method is called Extended Chance-Constrained Programming (ECCP); and the second is called the Extended RO ()$\Delta - EV$ model.

Extended CCP (ECCP): In practice, a manager's perspective towards robustness is not always to achieve a specified level of robustness but rather to find a solution that `controls' uncertainty while at the same time optimizing another quantity, e.g. minimizing costs or maximizing revenue. Managers may find it difficult to specify the robustness in terms of a probability of violation, especially for multiple constraints with uncertain parameters. They may be more comfortable with setting a 'robustness budget' δ – that is, the extent to which the key performance metric's value may be decreased in order to obtain a solution which is more robust than the nominal solution. Marla [72] and Marla, Vaze and Barnhart [73] propose to adapt the CCP model in this direction and suggest introducing a constraint that restricts the loss in the performance metric value by a cost budget δ while maximizing the robustness measure α. Specifically tailored to linear programs, their robust formulation sets the protection level α as a variable, while finding the most protected solution within a robustness budget δ compared to an optimal solution x^* to the nominal problem: $\min c^T x \; s.t. Ax \leq b$. The *ECCP* formulation therefore is: $\max \alpha, \quad s.\ t.\ \ c^T x \leq c^T x^* + \delta; \mathbb{P}(Ax \leq b) \geq \alpha$. For details of the linearization of the formulation under right-hand-side uncertainty and special set-partitioning constraints, we refer the reader to Marla [72] and Marla, Vaze and Barnhart [73].

$\Delta - EV$ *(Extended RO) model:* This model builds upon the uncertainty set specifications of the robust optimization approach of Bertsimas and Sim, that is, each parameter a_{ij} assumes a realization $\tilde{a}_{ij}$, in the uncertainty set $[a_{ij} - \hat{d}_{ij}, a_{ij} + \hat{d}_{ij}]$. Marla [72] suggests that it may be difficult for the manager to *a priori* specify a *number* of coefficients Γ to be protected, because it can be unintuitive, and besides it can be cumbersome for large-scale problems with many constraints, because it is difficult to define a robust measure Γ for *each* individual constraint. The selection of an overall robustness budget with respect to the cost may be more intuitive for the decision maker to specify and also eliminates the need to define a priori the robust measure Γ for each constraint. They propose an alternate model, in the same spirit as ECCP for CCP, tailored to large-scale *binary integer* programs, that sets an allowable 'robustness budget' δ with respect to the cost of the optimal solution x^* to the nominal problem and maximize the number of coefficients Δ protected within that budget [72,73]. This formulation therefore is: $\max \Delta, \ \ s.\ t.\ c^T x \leq c^T x^* + \delta; \sum_j a_{ij} x_j + \max_{\{\sum_j z_{ij} \leq \Delta\}} \{\hat{d}_{ij} |z_{ij}|) \} \leq b_i \ \forall \ i$. For details of the linearization of this formulation, we refer the reader to Marla [72] and Marla, Vaze and Barnhart [73]. The *robustness measure* of this method is the 'budget' δ.

The methods discussed so far each have different capabilities and limitations. The RO approach (and also the $\Delta - EV$ model for mixed-integer programs) is applicable to a large variety of problems due to low data requirements and model simplicity; whereas CCP (and ECCP) and CVaR require more information about uncertainty as well as more complex formulations. However, if additional information (qualitative or quantitative, partial or full) about the underlying uncertainty is available, there are limited ways in which RO can capture it (some very recent methods are presented in [74]). Additionally, each approach is different in the way it trades off the objective function with its robustness metric, indicating qualitative differences in the robustness of solutions generated.

There is a growing body of research [75] that has very recently begun to connect the three fundamental types of approaches that we discuss in this work. For example, the interfaces between these three approaches have resulted in a proliferation of work on areas such as distributionally robust optimization [76–80], and an exploration of the connections between CCP, RO and CVaR in works such as Chen et al. [81].

A number of these approaches also exhibit similar tradeoffs between model interpretability, statistical understanding and formulation tractability as those we discuss in this work. Hence, the general insights that we discuss here extends in general to a broader, evolving class of approaches.

3. Case Study 1 – Corporate Portfolio Optimization

Problem setting. A manager at a major corporation must decide how to allocate a sales and marketing budget B among twenty-six business units. It is possible to only allocate the budget once a year. For each business unit i, sixteen quarters of past historical investments x_i and corresponding quarterly revenues r_i are available. Also, for each business unit, minimum and maximum feasible investment amounts l_i and u_i, based on historical investments and business constraints, are known. The manager is interested in two key-performance indicators (KPIs): (i) high mean return, and (ii) low variance of return, which we use as the robustness metric.

Nominal Model. The causal relationship between investment and the corresponding revenue is typically seen to follow an S-curve [82,83]. The manager's observations, existence of limited data, and modeling considerations, all dictate that we can approximate the investment-return relationship to be linear in the portion of the S-curve between l_i and u_i [84]. That is, we describe the revenue r_i by $r_i = a_i + b_i x_i$ where a_i and b_i are parameters that describe the return for an investment x_i. a_i and b_i are estimated using least squares regression, and the corresponding estimates $\overline{a_i}$ and $\overline{b_i}$ and the associated covariance matrix C_i and standard deviation matrix G_i ($C_i = G_i G_i'$) are found. (Note that we find no significant correlation among different business units i.) The nominal model is described by (1)-(4).

$$\max \sum_{i \in I} \overline{r_i} = \sum_{i \in I} (\overline{a_i} + \overline{b_i} x_i) \tag{1}$$

$$s.\ t.\ \sum_{i \in I} x_i \leq B \tag{2}$$

$$l_i \leq x_i \leq u_i \quad \forall\ i \in I \tag{3}$$

$$x_i \geq 0 \quad \forall\ i \in I \tag{4}$$

Motivation for modeling uncertainty. The nominal solution recommends investment in decreasing order of deterministic (average) return rate, within the specified bounds. It is however, well known that diversification in a portfolio can decrease the risk associated with the investment. Further, after simulating the mean-variance tradeoff of the expected revenue (using a normal distribution centered in $\bar{a}$ and $\bar{b}$ with variance-covariance matrix C), the manager deemed it was not robust enough, due to the high variance (Fig. 1).

Robust Models. We model uncertainty using the RO approach, the CCP and the CVaR approaches. The Δ-EV model and ECCP model are not applicable to this problem as they were developed for cases when the basic RO model is a binary integer program and the basic CCP model is a linear program, respectively. These formulations are in Appendix A.

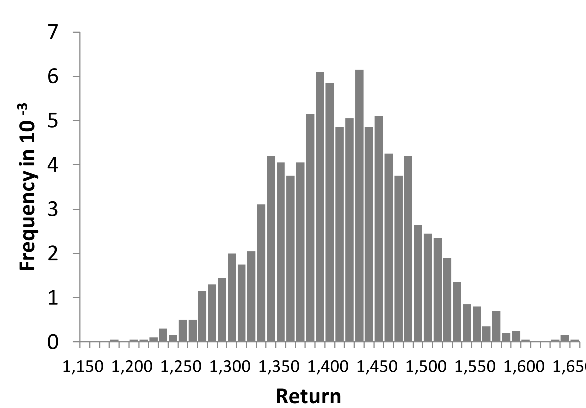

Fig. 1. Frequency curve for nominal portfolio return indicating high variance in return.

3.1. RO applied to corporate portfolio optimization

Uncertainty model: For this approach, we define for each parameter a range of uncertainty around the average estimate, as $[a_i - \hat{a}_i, a_i + \hat{a}_i]$, $[b_i - \hat{b}_i, b_i + \hat{b}_i]$ using the covariance $C_i = G_i G_i' \ \forall i$; derived from the past 16 quarters of investments. We choose $\hat{a}_i$ to be the standard deviation of a_i, computed from G_i. The rationale behind using the standard deviation is discussed in more detail in [86]. As described in the original work [68], the robust formulation yields a larger linear program compared to the nominal model. The linearized formulation is in Appendix A.1.

Conservatism of the approach: We use the guidelines for specifying the budget of uncertainty Γ, provided by Bertsimas and Sim [68], which mathematically connect Γ to an upper bound on the probability of constraint violation. Using Γ, the robust program aims to approximate the variance of the distribution of the portfolio return. However, translating the business performance indicator to the metric Γ is non-intuitive. The actual (realized) probability of violation was always much lower than that predicted by the bound in [68], rendering the approach highly conservative and risk-averse. This is because the bound is derived from analyzing the worst case configuration of symmetric distributions (which assign high mass of probability to the extreme values in the range of uncertainty). This confirms observations made by Sakamoto [26], and Bryant [86] for linear and integer programs. The weakness of the bound, especially for integer programs, makes it difficult to choose a suitable Γ a priori, to relate the manager's metric specified in terms of probability of constraint violation, without incurring conservatism in the objective.

Tractability: Because the realized probability of constraint violation (as evaluated by simulation) is poorly related to Γ, choosing Γ via the bound presented in [68] is not effective. In this problem instance, because we have a limited number of uncertain parameters, we test for all possible integer values of Γ, and test solution robustness via simulation. This strategy can degrade tractability for larger instances but is reasonable for this problem size.

Addressing solution conservatism: We solve the corporate portfolio optimization problem using the RO approach, using $\bar{a}$ and $\bar{b}$ as estimated, and with $\hat{a}$ to equal G and separately to equal $G/3$; and the RO parameter Γ estimated from the bound in [68]. We then evaluate the solutions using simulation, assuming a normal distribution centered at $\bar{a}$ and $\bar{b}$ with standard deviation given by $G_i \forall i$. Our empirical results show that the RO approach produces solutions (evaluated using simulation) with better tradeoffs between the mean and standard deviation of profit when the uncertainty range input to the model is for a smaller range $G/3$ (Fig. 2 right) than the *actually* realized range G (Fig. 2 left).

In the case of this problem, in practice, the manager confirmed that the original solution obtained using the standard deviation G was too conservative. The solutions derived from using $G/3$ instead did produce

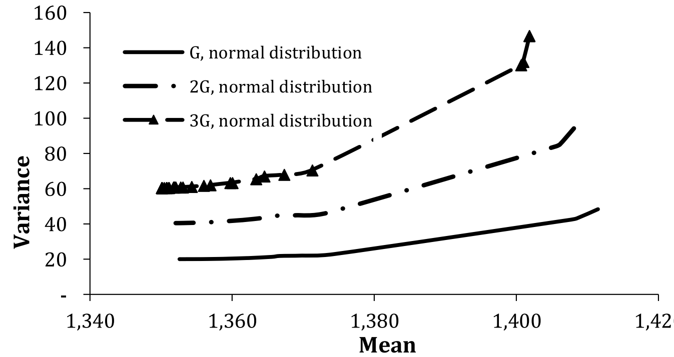

Fig. 3. Sensitivity of the mean-variance to uncertainty range in the RO approach.

a satisfactory solution to the manager as there is a substantial improvement in variance without significant cuts in expected return (see Fig. 2).

Similarly, we evaluate the solutions found with G as the uncertainty range using simulation, under hypothetical scenarios drawn from distributions with the same mean but standard deviations specified by $2G$ and $3G$, to understand the conservatism. Plots of the mean-variance tradeoff for these different distributions are shown in Fig. 3. We find that the increase in the robustness metric (reduction in variance) with decrease in mean is greatest for the curve "$3G$-normal distribution", that is, when data turns out to be, in fact, from a distribution with standard deviation $3G$. On the other hand, for a small drop in variance, the curve "G-normal distribution" has a much larger drop in the mean; that is, this solution is conservative when the data turns out to be from the distribution with standard deviation G. This behavior is again confirmed for various distributions in Fig. 11. This behavior is more evident for light-tailed distributions, such as the normal distribution (observed in historical data for this problem), than in the case of heavy-tailed distributions (tested in our experiments through simulation), because the values at the boundaries of uncertainty exploited in the formula provided in [68] are more likely to be realized.

3.2. Chance-Constrained Programming of Charnes and Cooper

Uncertainty model: The mean and variance of the uncertain parameters, derived from the historical data available, are required to model uncertainty. We assume that a_i and b_i are drawn from a bivariate normal distribution with means $\bar{a}_i$ and $\bar{b}_i$ and covariance C_i (hence implicitly assuming that the business units are independent). The form of the chance-constraint is then: $\mathrm{P}(a_i + b_i x_i \geq c.\,v.) \geq \alpha$where $c.v.$ is a minimum return from the portfolio specified by the manager. The chance-constraint is converted into the deterministic constraint (5) as

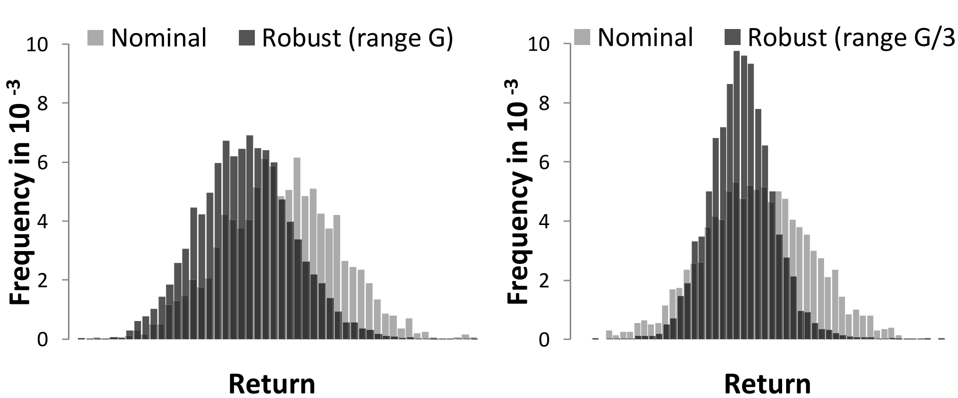

Fig. 2. Robust and nominal solution comparison: (i) with same uncertainty range G as input into optimization (left) (ii) with one-third uncertainty range G/3 as input into optimization (right).

follows.

$$\sum_i \bar{a}_i + \bar{b}_i x_i + \Phi^{-1}(1-\alpha)\sqrt{\sum_i < 1, x^T > C_i < 1, x >} \geq c.\ v. \tag{5}$$

Defining model parameters: The parameter α or the probability of constraint violation, and the threshold *c.v.* of minimum return, are to be specified by the manager. In this case, it is easy to translate the manager's KPI into a numerical value α, as the probability of violation is a measure that is fairly intuitive for the manager to interpret. However, there are challenges involved in setting α and *c.v.* a priori as some values might lead to infeasibility.

Low solution conservatism: The input percentile violation α, and the average return of the portfolio from the optimization approximate well the values obtained through simulation and do not exhibit solution conservatism, unlike the solutions of the robust optimization approach. One could argue that this is because the model used for optimization is closer to the model used for simulation in this later case. However, though we assume a normal distribution of uncertainty for the model, we were still able to obtain solutions that are robust, and that exhibit low conservatism, even when tested under other slightly different hypothetical distributions via simulation. Qualitatively, the solutions are comparable to the most robust solution obtained using the robust optimization approach, with input range *G/3. However the parameters were easier to tune and required less effort from the manager*, which shows that using approximate distributions can add value when used with caution. This also indicates that some equivalence can be achieved between these two approaches, but with careful tuning of multiple parameters representing uncertainty descriptions as well as robustness parameters.

3.3. CVaR (Conditional Value-at-Risk)

Uncertainty model: The CVaR constraint here is modeled through Sample Average Approximation (SAA) over multiple scenarios. Specifically, we capture uncertainty using scenario-based constraints added to the nominal formulation. The constraints, for each scenario $j = 1,\ldots,M$, are described in (6)-(8), with α being the quantile of value-at-risk we want to protect for. A detailed knowledge of the underlying distributions for uncertain parameters is required to generate the required number of scenarios, as described in Rockefellar and Uryasev [87]. The idea is to maximize the expected return while ensuring that CVaR is above a certain critical value (for a chosen α, defined as $c.v.(\alpha)$). Because we want to ensure a minimum return even in the lower tail of the distribution, we define $\beta = VaR(\alpha)$ (the value-at-risk, modeled as a variable), and use constraint (7) to measure the average value of return in the tail. Constraint (8) sets the minimum return in the tail to be at least a minimum value $c.v.(\alpha)$.

$$z_j \geq 0 \quad \forall j = 1, \ldots, M \tag{6}$$

$$z_j \geq \beta - \sum_i (\widetilde{a}_{ij} + \widetilde{b}_{ij} x_i) \ \forall j = 1, \ldots, M \tag{7}$$

$$\beta + \frac{1}{(1-\alpha)M}\sum_{i=1}^{M} z_j \geq c.\ v.\ (\alpha) \ \forall j = 1, \ldots, M \tag{8}$$

Defining model parameters: We specify a tail probability α of the profit function, the critical value $c.v.(\alpha)$, and generate scenarios based on the underlying distributions; to generate constraints that are added into the linear program. Translating the business indicator into a numerical value for the critical value is not straightforward, as the value of CVaR is not intuitive to translate to simpler metrics of interest to the manager.

Tractability considerations: The number of scenarios (constraints) needed to converge to the value of CVaR is not known *a priori* – therefore, we have to iterate while increasing the number of scenarios. Tractability may thus be a challenge because we might have to experiment with the number of scenarios, increasing them iteratively before finding the right value. Fig. 4 demonstrates the shows that as the number of scenarios (constraints) increase, the CVaR function value for the expected return begins to flatten, and may need many more scenarios before convergence; impeding tractability. Moreover, the convergence of the return from the CVaR formulation is highly non-linear, indicating the complexity in achieving a convergent estimate of the return.

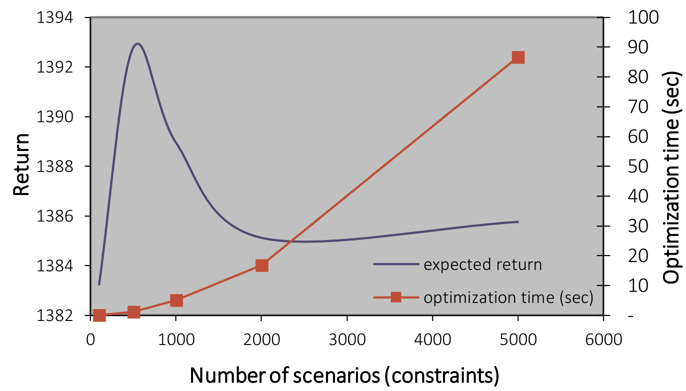

Fig. 4. Accuracy-Tractability Tradeoff for determining true CVaR value.

3.4. Insights

Translating the manager's robustness KPI (here, the variance of the portfolio return) into an input of one of the three models (Γ, α, *and/or c.v.*) is not straightforward and trial and error was required. It is easier in the CCP model because the risk measure (probability of violation) is more intuitive to the manager. Nevertheless, generating robust solutions is an iterative, interactive process, whatever model is chosen. Indeed iteration among values of model parameters is required to achieve the 'right' tradeoff between expected return and volatility/robustness, which may be differently defined by each organization or manager. Interaction between the analyst and manager is crucial in this process, as the manager decides if the nominal or robust approach solutions are 'good enough', with sometimes KPI's that (s)he cannot even describe precisely.

In this case study, the robust solutions produced by the different methods result in the diversification of the portfolio and the generation of solutions that were all considered "reasonable" by the manager (in contrast with the solution to the nominal problem). It demonstrates that incorporating uncertainty a priori adds value. The solutions, however, were very similar in performance, because the uncertainty is relatively small, though not insignificant.

The RO approach is the easiest to use, because of its limited data requirements and its (relatively) automatic mathematical reformulation. We do not recommended to use the probability bound of constraint violation proposed in [68] directly though, as the bound is obtained from analyzing extreme distributions and it usually leads to over-conservative solutions for non-extreme ones (as was the case in this case study).

The CCP and CVaR approaches require assumptions about the underlying data distributions to generate scenarios and/or to derive close form formulae. We observe that solutions using these approaches are robust even under small perturbations of the data distributions, and exhibit less conservatism than the basic RO approach that only uses range information on uncertainty. Therefore, when reasonable assumptions with regard to the underlying data distributions may be appropriate, these two approaches might be preferred (the CCP being easier to implement in this context). However, the analyst should use these corresponding solutions with caution because despite the fact that they are robust to small perturbations of the data distribution, they might become more volatile with larger perturbations. In contrast, for this application (and possibly other problems of small size), is does not hurt much computationally to try all possible integer choices for Γ. Combining this with scaling down the uncertainty sets by a factor $\gamma \geq 1$, for a limited, yet meaningful, number of values for γ, should allow to overcome conservatism of the RO approach and if no reasonable assumption can be made on the underlying data distributions, this approach should probably be preferred.

4. Case Study 2 – Pharmaceutical Supply Chain Optimization

Problem setting. The supply chain planner at a major pharmaceutical company must design the optimal configuration of the strategic supply chain for the next ten years [88]. The configuration cannot be modified during the operational period. The company manufactures 17 broad classes of products using different technologies at different manufacturing plants. Allowed changes to the existing network include closing or opening a plant, improving the technology used at a plant, moving a product from one plant to another, or in some cases, adding or discontinuing a product. The products are subject to inspection by the Food and Drug Administration (FDA), which can 'fail' a batch of products based on sampled testing. For each product-plant combination, there is a *hazard rate* of failed inspections based on the technology used and the production location. The KPIs of interest are to maximize the *realized* profit, and to minimize the probability of failed inspections (this is the robustness metric used for simulation).

Nominal model. Let P be the set of products to be produced, S be the set of locations available for production, T be the set of technologies to be used for producing products $p \in P$; E be the set of discrete time periods into which time horizon is divided. Let $H(t, s, e)$ be the hazard rate, which is subject to uncertainty, and equals the probability that an inspection of technology t at site s during period e results in a failure; $\bar{H}(t, s, e)$ be the expected value of $H(t, s, e)$, that is, the mean hazard rate of failed inspections; $R(p, e)$ be the revenue generated by producing product p during period e; $C(p, t, s, e)$ be the cost of producing p using technology t at site s during period e; CV be the critical value of the revenue at risk, estimated by statistical methods; and $x(p, t, s, e)$ be decision variables that take on value 1 if product p is produced using technology t at site s during period e. The nominal model formulation (9)-(12) maximizes the expected profit, subject to the constraint that the expected revenue that is at risk being limited by a critical threshold CV; all products being produced; and all x variables being binary.

$$\max_{x \in X} \left(\sum_{p,t,s,e} R(p, t, s, e)x(p, t, s, e) - \sum C(p, t, s, e)x(p, t, s, e) \right) \tag{9}$$

$$s.\ t. \sum_{p,t,s,e} \bar{H}(t, s, e)R(p, e)x(p, t, s, e) \leq CV \tag{10}$$

$$\sum_{t,s} x(p, t, s, e) \leq 1 \quad \forall \ p, \ e \tag{11}$$

$$x(p, t, s, e) \in \{0, 1\} \quad \forall \ p, \ t, \ s, \ e \tag{12}$$

Motivation for modeling uncertainty. $H(t, s, e)$ values are estimated from historical data using Bayesian statistical methods and the true realizations are inherently subject to *uncertainty*. Especially, in this data, it is found that with small changes in the values $H(t, s, e)$, the configuration of the supply chain that is optimal as obtained by solving the nominal problem (9)-(12) changes drastically. Since we are interested in a long-term decision, the analyst tests the performance of the model with hazard rates that might change during the ten-year period. We found that the optimal supply chain configuration obtained by solving the nominal problem generated an expected profit of $61,000 (numbers have been scaled); however, if all hazard rates increase by even 2%, the optimal profit (objective function) drops by 40% and the number of product types produced drops from 17 to 14. This indicates very high sensitivity, which is caused by the fact that investments with high mean return also have high variability. The robust formulations are in Appendix B.

4.1. The RO approach

The range of uncertainty in the hazard rates is set to equal the standard deviation of the hazard rate $\bar{H}(t, s, e)$. According to the uncertainty model of this approach, let $\hat{H}(t, s, e)$ be the range of uncertainty around the mean hazard rate . From historical data, the standard deviation in $H(t, s, e)$ is found to be 0.04 units. According to the RO approach, the budget of uncertainty Γ protects against the case when *any* Γ of the hazard rates attain their worst-case values, and with the remaining parameters staying at their averages. The model finds a solution that maximizes profit for that scenario. The linearized formulation is provided in Appendix B.1. As in Case study 1, the choice of Γ was first guided by the bound provided in [68], via the definition of a probability of constraint violation.

Conservatism of the RO approach. Because the true underlying distributions of the risk parameters are unknown, to test robustness, we simulate scenarios where hazard rates originate from several types of distributions, all with the same standard deviation (used as the uncertainty range) of 0.04 (Table 1 Theoretical and actual probabilities of violation from the RO approach).

In each of these cases, the value of the robustness budget Γ computed using the bound in [68] generates solutions that are far more conservative and have a much lower probability of violation than was used as input. Correspondingly, they also have a much lower mean profit, resulting in solutions that are robust but far more conservative than the manager prefers. It appears here again overly conservative to set the Γ according to the bound specified in [68].

Setting parameters: In this setting, it is impractical to try all possible values of Γ as we suggested in case study 1. But this raises the following questions: How do we relate Γ to the amount of protection the manager wants, and more importantly, to the KPIs of robustness if the bound in [68] is inappropriate? Is there a more intuitive robustness measure (instead of Γ) that the manager can adopt? How do we address the issue of conservatism of the solution without enumerating all choices of Γ (and over several rescaled uncertainty sets)?

Addressing solution conservatism of the RO approach: One way to control conservatism is to tighten the bound presented in [68]. Gallay [89] discusses that the mathematical bound relating Γ to the probability of constraint violation in [68] can be modified by using the number of basic variables N' rather than the total number of variables N; generating a tighter bound as shown in Fig. 5. This is because only the parameters that are *active* contribute to the uncertainty in the solution. It is difficult to estimate N' before solving the model, but the nominal problem solution and the rank of the matrix can be used to approximate N' [89].

Alternatively, (or in combination), we can control conservatism by controlling the range of uncertainty; specifically by choosing a convex subset of the original uncertainty range. For example, in Fig. 6 (left), we

Table 1
Theoretical and actual probabilities of violation from the RO approach.

Hypothetical distribution of underlying uncertainty	Theoretical bound for probability of violation (%)	Actual probability of violation from simulation (%)
Truncated normal	5	0
Triangular	5	0
Uniform	5	0
Beta (0.5,0.5)	5	0.003
Beta (0.2,0.2)	5	0.005
Beta (0.01,0.01)	5	0.08
Extreme discrete (bi-modal, 50% at each extreme)	5	0.11

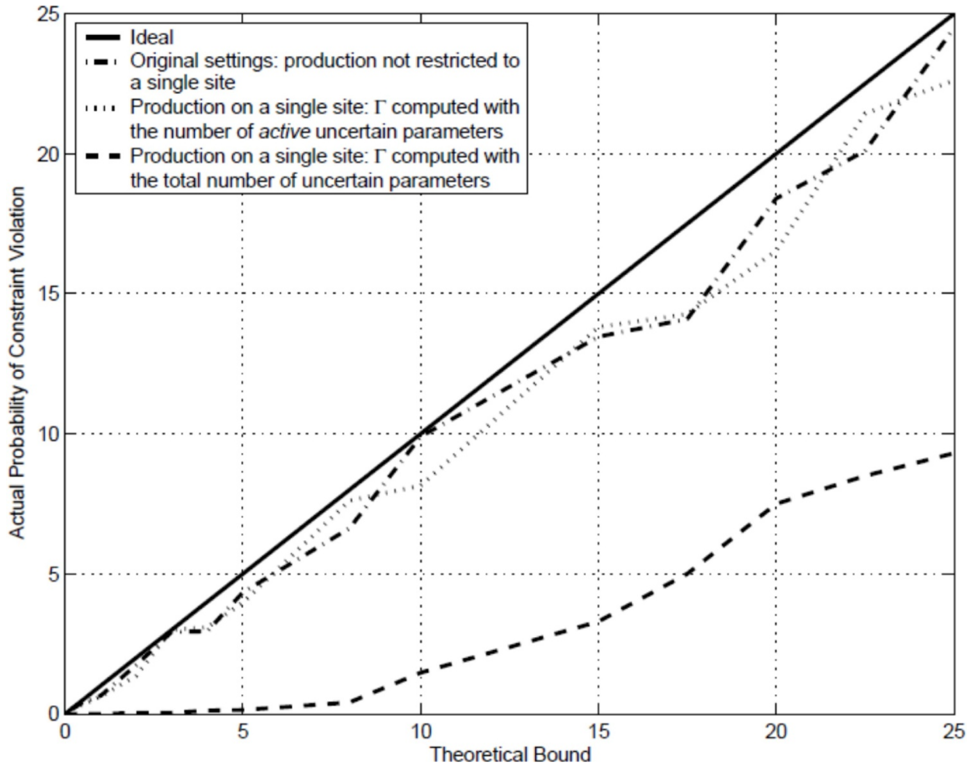

Fig. 5. Controlling solution conservatism using the number of active uncertain parameters.

see that different mean-variance tradeoff curves for the return are generated with uncertainty ranges of 0.04 (the original uncertainty range) and 0.02 (a convex subset of the uncertainty range). Note that the top right point on each curve denotes the nominal solution, with $61,000 of profit and 315 units of variance. Solving for various values of Γ with a range of 0.04, we trace the red curve, decreasing variance as well as the mean with increasing Γ.

Observe that the change in the mean-variance relationship is highly non-linear, and robustness is increased by decreasing the variance, with the tradeoff being the reduction in the mean. However observe that the reduction in the mean can be quite over-conservative with reduction in variance, and may not be acceptable to the decision-maker. The black curve traces, for uncertainty range 0.02, the RO model solved for the same values of Γ. This allows, for the same values of Γ, a different and less-conservative change in the mean relative to the reduced variance and allows the decision-maker a more acceptable level of conservatism. While we show here a smaller set of values of uncertainty ranges such as half the convex subset, this is generalizable to all convex subsets, such as ranges of 0.03 or 0.01. This shows how, by changing the range of uncertainty input to the model, we can modify the tradeoff curve such that the conservatism in objective value (but also the robustness) is decreased.

Controlling conservatism can be further enhanced using minimal (qualitative) knowledge of type of distributions (e.g. heavy-tailed vs. light-tailed). If the distribution is light tailed, the probability of parameters realizing values at the bounds of their uncertainty ranges is low. Setting the uncertainty range to a convex subset of the observed range can result in a less conservative solution, than in the case of a heavier-tailed distribution of uncertainty. These strategies provide a way to tune level of robustness with an iterative process of optimization and

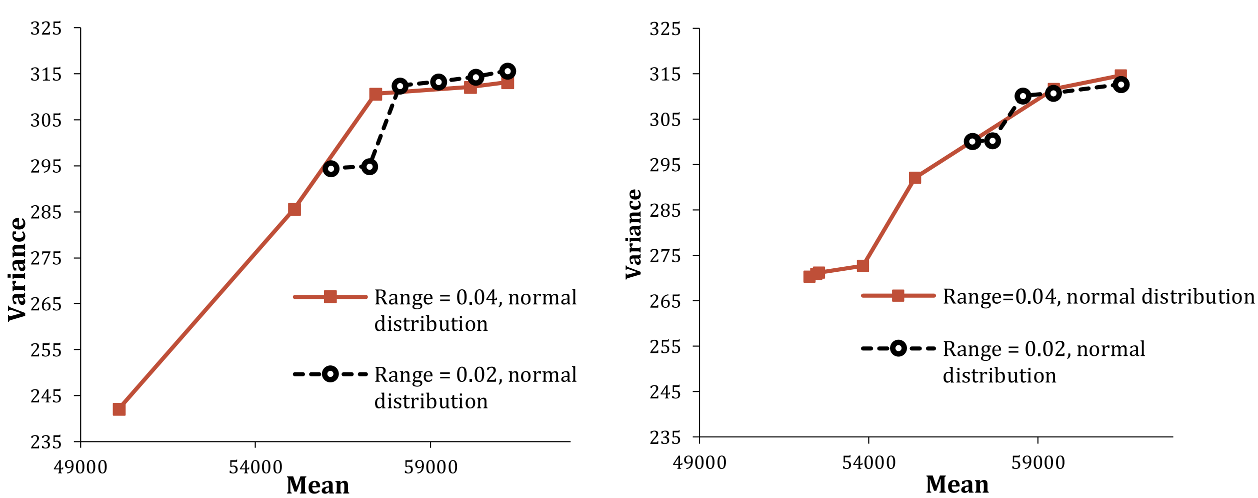

Fig. 6. Mean-variance tradeoff of profit: Less conservative tradeoff curve generated using modified (smaller) uncertainty sets. RO model solutions (left) and Δ-EV model solutions (right). In Figure left, Γ varies from 0, 8.18, 20, 31.78, 51, 80, and 120; and δ in Figure right takes values 199, 1199, 4199, 6199, and 10,199.

simulation; and generate solutions that are robust in a less conservative way.

4.2. Alternative approach: the $\Delta - EV$ model

The $\Delta - EV$ model is applicable in this case study for the following reasons: a) the weakness of bounds specified in Bertsimas and Sim [68] make it difficult to specify the robustness parameter Γ a priori; b) it is non-intuitive for the manager to specify the *number of coefficients* to protect against extreme values. Instead, Marla and Barnhart's $\Delta - EV$ model allows the manager to specify the maximum loss in the objective δ that he is willing to take in order to obtain a robust solution; and maximizes the number of coefficients that can realize their worst-case values under the robustness budget δ; and c) This allows control over the level of conservatism through the allowable change in mean profit.

Uncertainty model: The uncertainty model remains consistent with the definition of uncertainty ranges defined by the RO approach. Each uncertain parameter is assumed to vary within a range around its nominal value, and a subset of parameters realize their worst-case values. However the size of the subset is a variable determined by the model.

The $\Delta - EV$ formulation chooses the solution that protects the highest number of uncertain parameters at their worst-case, while ensuring that the objective function is within a difference of δ from its nominal value. It overcomes the limitation of having to specify the value of Γ a priori.

Because this model is also based on a subset of uncertain parameters realizing their worst-case values, it can exhibit conservatism similar to that seen in the RO approach. The conservatism is again higher when the underlying distributions have thin tails rather than heavy tails. We recommend that the conservatism can be controlled by choosing uncertainty sets that are convex subsets of the original uncertainty set, as seen in Fig. 6. Note that this observation is similar to that in Case Study 1, in Fig. 2 and Fig. 3

The tradeoffs between mean and variance of the profit made by the RO and $\Delta - EV$ approaches differ slightly from each other, as we discuss in the following sections.

4.3. Chance-Constrained Programming

The chance-constrained formulation for (9)–(12) enforces the constraint that the hazard of the return (revenue-at-risk) is less than a critical value, with probability α. α is pre-specified by the manager, and the goal is to maximize profit under the probabilistic constraint (13). The intuitive nature of the CCP constraint makes it easier for the manager to specify a value of α, as in Case Study 1. In the CCP model, constraints (6) are substituted with constraints (13), which contains multiple uncertain parameters $H(t, s, e)$ (assumed to be uncorrelated).

$$\mathbb{P}\left(\sum_{p,t,s,e} H(t,s,e)R(p,e)x(p,t,s,e) \le CV\right) \ge \alpha \tag{13}$$

Typically, such constraints are handled using scenario approximations [41], by sampling a large enough set of realizations of the $H(t, s, e)$ parameter from the (joint) probability distribution. Each of these scenarios is translated into a constraint, all of which together equal (13). However, a very large number of scenarios are required, because of the large dimensionality of $H(t, s, e)$; and this causes tractability issues for the integer program. Hence, we resort to approximations for (13), as described in (14)–(15). Specifically, instead of sampling scenarios, we add an additional constraint (15), which assumes that the βth quantile realizations of each of the uncertain parameters $H(t, s, e)$, (denoted by $F^{-1}_{H(t,s,e)}(\beta)$) occur simultaneously. Thus, the left-hand-side of (15) approximately captures the risk from the kind of realizations of $H(t, s, e)$ we want to protect against. This does not capture all possible scenarios of realizations, but is reasonable for this specific problem, as sensitivity was found specifically for cases where hazard rates increase to some quantiles of their distributions. $F^{-1}_{CV,CCP}(\alpha)$ in the right-hand-side of (15) represents the quantile α of the critical value CV against which we protect.

$$\sum_{p,t,s,e} \bar{H}(t,s,e)R(p,e)x(p,t,s,e) \le CV \tag{14}$$

$$\sum_{p,t,s,e} F^{-1}_{H(t,s,e)}(\beta)R(p,e)x(p,t,s,e) \le F^{-1}_{CV,CCP}(\alpha) \tag{15}$$

Because of the approximate constraint, it is difficult to determine α *priori* what protection level α the manager should target. We encounter cases of infeasibility for values of α such as 99 or 97%, as there exists no solution that can provide as high a protection level. Therefore multiple trials are required to find a feasible solution, and moreover, it is also difficult to predict the tradeoff between the objective (profit) and the protection level α specified. We address this issue using the ECCP method.

4.4. Extended Chance-Constrained Programming (ECCP)

In the ECCP model [72,73], instead of specifying a probability of constraint violation, the manager is asked to specify a loss in profit δ (s) he is willing to accept (called robustness budget), with the formulation maximizing the probability that revenue-at-risk does not exceed the critical value. Because we specify a budget δ we avoid the infeasibilities associated in the CCP with a priori specifying α in the CCP approach. We find that fewer iterations are required to control the tradeoff between the mean and variance of the profit, because it is controlled via the budget δ specified by the manager. The formulation solves in the same order of time as a single iteration of the original CCP approach.

4.5. Conditional-Value-at-Risk (CVaR)

The CVaR model minimizes the expected revenue at risk, when the realized revenue at risk is greater than the VaR corresponding to the αth level of protection. We convert the CVaR constraint (16) to a deterministic linear program, similar to the corporate portfolio case.

$$CVaR(\alpha, x) = \mathbb{E}\left(\sum_{p,t,s,e} \bar{H}(t,s,e)R(p,e)x(p,t,s,e) \,\middle|\, \sum_{p,t,s,e} \bar{H}(t,s,e)R(p,e)x(p,t,s,e) \ge VaR(\alpha, x)\right) \tag{16}$$

Accuracy-tractability tradeoff: Because uncertain parameters are not correlated and there is a fair certainty about their distributions, we are able to generate a set of scenarios describing possible realizations. The number of scenarios required to capture uncertainty in all uncertain parameters is in the tens of thousands. The deterministic CVaR formulation has one constraint per scenario. However, similar to the corporate portfolio case study, we observe a trade-off between accuracy and tractability. In fact, for more than 400 scenarios, the CVaR formulation proves intractable for realistic instances of this problem, and the solutions with fewer scenarios are unreliable. Therefore CVaR is not considered an alternative for this problem and we do not discuss it in this paper.

4.6. Insights

The various methods (the RO, Δ-EV, CCP and ECCP approaches) generate robust solutions with a lower variance than the nominal solution, as was required by the manager. Fig. 7 shows the different solutions obtained using the four approaches applied to this problem. The nominal solution is the point where mean = 61,000 and variance = 315. Using these different approaches, we find robust solutions

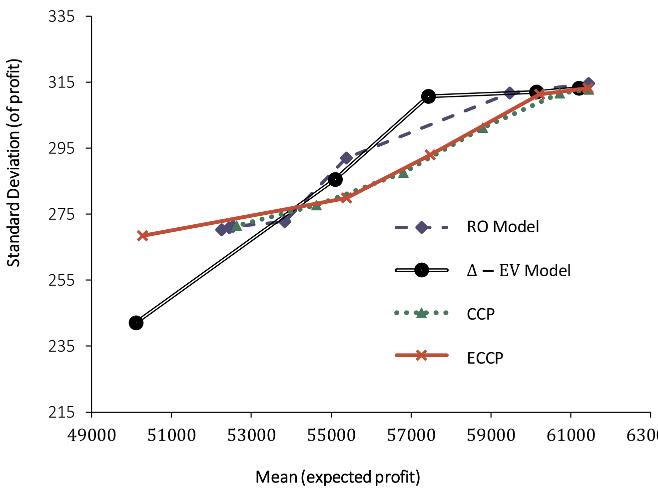

Fig. 7. Mean-variance trade-offs made by different approaches.

that have a lower variance than the nominal solution. For each of these solutions, a decrease in the mean profit is incurred (the price of robustness).

By tuning the robustness parameters Γ or α or robustness budgets δ (for the $\Delta - EV$ and *ECCP*), each method can generate a series of solutions, described by the curves shown in Fig. 7, each for various values of the robustness parameter corresponding to that approach. The points on the curves represent specific solutions obtained using these methods. Applied appropriately, using the right parameter values, solutions can be tuned further to obtain the level of conservatism (or the objective function-robustness tradeoff) desired by the manager.

However, each modeling approach trades off objective function (profit) and robustness metric (variance) differently. As seen in Fig. 7, there is a smoother tradeoff using *CCP, ECCP* and more step-like tradeoff curve using the RO and $\Delta - EV$ approaches. This is seen when the data is drawn from multiple distributions, such as the normal, uniform, or discrete uniform distribution, as seen in Fig. 12 (Appendix B.4). The 'steps' in the tradeoff curve [68] are because of the metric Γ, which is defined in terms of the range of uncertainty input into the model. Our first observation is that each of the methods traverses a significantly different tradeoff curve between the mean and the standard deviation. For example, in the range [54,000, 61,000] for the mean (where 61,000 is the nominal solution), the *CCP* and *ECCP* models have significantly less conservative tradeoffs than the RO and the $\Delta - EV$ models – that is, the drop in mean required for each unit of reduction in variance is lower for the *CCP* and *ECCP* than for the RO and the $\Delta - EV$ models. The conservatism in the tradeoffs of the RO and $\Delta - EV$ models can be controlled by using input ranges that are convex subsets of the standard deviation, as seen in Figs. 13 and 14, compared to Fig. 12, at points close to the nominal solution value. This indicates that the methods fundamentally trade off the objective function (mean return) and robustness metric significantly differently.

Our second observation, therefore, is that obtaining solutions with an acceptable objective function-robustness metric tradeoff is an iterative process, as it involves changing the parameter values, uncertainty ranges, or probability of violations, without tight bounds or guidelines. Moreover, the manager's input is required at each stage to determine the appropriateness of the tradeoff.

Third, among the four models tested, the alternative models ($\Delta - EV$, *ECCP*) can make robust modeling more intuitive, by allowing the manager to control the level of conservatism in objective directly instead of testing of via constraint protection parameters. However, it is difficult to predict *a priori* which method is 'better'. In addition, this is also difficult to gauge *a posteriori* unless we are able to evaluate the solutions using simulation and compare the true performance of the solutions.

5. Case Study 3 – Aircraft Routing

Problem setting. Given a published schedule to be operated by an airline, the airline schedule planner's routing problem is the following. The goal is to find a *feasible* sequence of flight legs, called aircraft routings or rotations, to be operated by each aircraft so that maintenance restrictions on aircraft are satisfied [90]. Each routing (corresponding to an aircraft) consists of a sequence of flights followed by A-checks at a maintenance station, which are to be performed every 60 h of flying [91]. Each flight is required to be operated by exactly one aircraft and there is a limited fleet size available. Details of the data and formulations are in Marla, Vaze and Barnhart [73].

KPIs. Uncertainty arises from the fact that delays occur in the network, and delayed flights can further cause delays in downstream flights. This can disrupt the aircraft rotations, render them infeasible if flights are excessively delayed, and potentially result in flight cancelations. The primary robustness metric is the propagated delay, that is, the delay that is experienced by a flight that is delayed because of the previous flight arriving late. Multiple robustness metrics, representing the interests of different stakeholders, are of interest: the 15-minute, 30-minute, 45-minute and 90-minute on-time performance of flights, the number and percentage of passengers missing connections, and the percentage of flight cancelations are of particular interest.

Nominal model. In the nominal problem we use a composite variable modeling approach [92] where each variable is a *feasible* maintenance route for an aircraft. Composite variables are used because of the ability to capture complex maintenance constraints (involving flying and elapsed time) through the variable specification rather than through additional model constraints. Thus each composite variable includes a set of flights to be operated by an aircraft followed by aircraft maintenance. It takes on value 1 if that route is chosen to be operated and 0 otherwise. The nominal problem formulation's objective is to choose a *feasible* selection of routes, with constraints being (i) the need to operate all flights (set-partitioning constraint), (ii) balance aircraft flow (network-flow constraint) and (iii) not using more aircraft than available (capacity constraint). For details, refer [93] and [92]. The number of variables (routes) is of the order of tens of thousands to hundreds of thousands for the instances of interest. It is key to observe here, that like most network-based problems, this problem is likely to have *multiple optimal* (and feasible) solutions given nearly any objective function, due to the nature of networks and set-partitioning constraints. This feature plays an important role in our discussion in this section.

Motivation for modeling uncertainty. On solving the nominal formulation, which is geared towards feasibility, multiple feasible solutions can be generated. Each of the solutions has an equal probability of being used by the airline, but can differ significantly from other solutions in terms of delays and passenger service levels. Fig. 8 shows the performance of multiple feasible routings that are all alternatives that could be used by the airline in the absence of explicitly modeling robustness [73]. In particular, when comparing the performance of the solutions over multiple scenarios over the course of one month of various airline disruptions, the authors note that the performance of the solutions is very similar on low-delay scenarios such as scenario 10 (horizontal-axis), but varies considerably in high-delay scenarios such as scenario 5.

Uncertainty models. There may be 'simple' ways of capturing uncertainty. For example, suppose we add slack, or virtual gaps, to each flight to absorb delay. This might be a rule-of-thumb approach adopted by the airline. However, this poses the following problems – it is difficult to find out where to add the slack, and by how much. Ad-hoc placement of slack, especially in highly resource-constrained problems like airline schedules would often require more resources (planes) than are available.

Robust Models and challenges: The challenges in building robust models using the approaches we consider in this paper are two-fold: First, translating the business indicator (flight or passenger delays) into

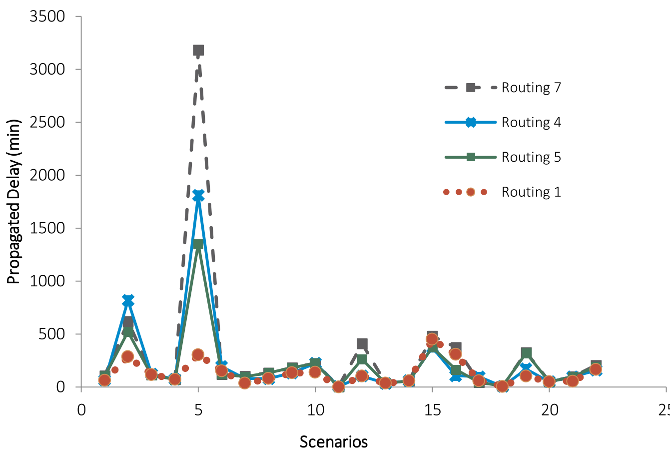

Fig. 8. Variability in performance of multiple nominal optimal solutions.

a standard risk measure (such as variance, value-at-risk or CVaR) is non-intuitive. Hence, it is not straightforward to decide which methods are better applicable. Second, to model this problem, we first have to understand the phenomenon of delays and introduce new parameters into the formulation that reflect delays.

Robust modeling: Marla, Vaze and Barnhart [73] first use the phenomenon of *propagated* delays to characterize uncertainty. The key understanding behind this model is that delay experienced by a flight is a function of the entire aircraft route (captured as a composite variable in the formulation) and not of a single flight, because upstream flights can cause propagation of delay to downstream flights. They model uncertainty as the amount of delay experienced by a flight when operated as part of specific aircraft route (note that each flight can be a part of several routes due to the network structure, but only one has to be chosen in the optimization). The uncertain parameters, then are the worst-case or probabilistic delay levels experienced by a flight i when operated as part of route r. This uncertainty can be physically interpreted through the following parameter(s) in the set-partitioning constraints: if a flight i incurs excessive delay (delay above a certain threshold) because of being operated as part of route r, we denote the chances of such excessive delay probabilistically (in the CCP paradigm) or the chances of such excessive delay in the worst-case (in the RO paradigm). For details, we refer the reader to Marla [72] and Marla, Vaze and Barnhart [73]. Empirically, route delay distributions are found to be approximately bi-modal, that is, they are small for the most part, but can be very large about 5–10% of the time.

5.1. The RO approach

The set-partitioning model parameters that capture if flight i is operated as part of route r are modified to also capture if flight i experiences a worst-case delay greater than a threshold if it is operated as part of route r. The nominal value of the parameter is 1 if flight i is part of route r. The uncertainty is asymmetrically distributed about the nominal value of 1 for the set-partitioning constraint parameters. The protection level Γ defined by the RO approach is associated with the 'number of aircraft routes to protect against realizing their worst-case propagated delay'. This means the formulation chooses those routes that minimize the level of propagation by protecting against the case when Γ strings attain their worst-case delay propagation levels.

Finding: Non-monotonicity of solutions from Bertsimas and Sim's RO approach. As the value of Γ increases, we find that solutions do not necessarily become more robust. For example, Fig. 9 shows the simulated performance of solutions with respect to propagated delays, for different values of robustness parameter Γ. $\Gamma = 2$ shows improved (decreased) delay with respect to $\Gamma = 1$, however, $\Gamma = 3$ is a worse solution than $\Gamma = 2$, with increased delay. Thus, a non-monotonicity in solution

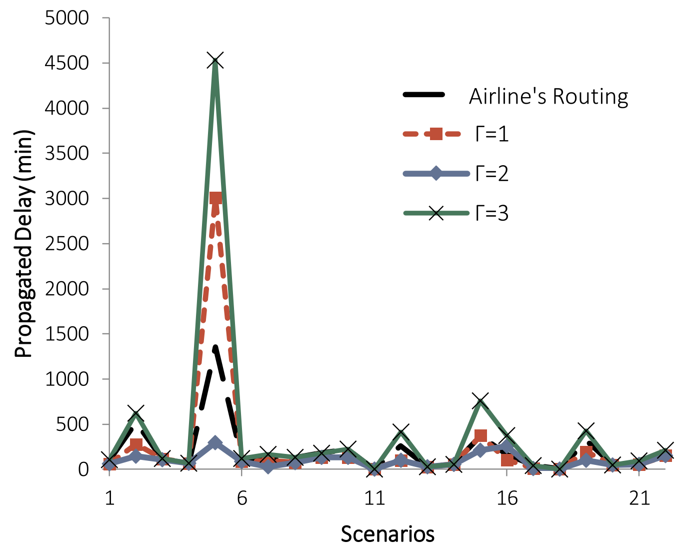

Fig. 9. Non-monotonicity in solutions with increase in robustness parameter values (Source: [33]).

behavior is observed. This is observed for several metrics (KPIs) of interest, both for propagated delays, as well as total flight delays that we optimize in the RO model as robustness metrics. Similar behavior is observed in network-based routing and scheduling problems for applications involving UAV routing and task scheduling by Sakamoto [85] and Bryant [86]; who report that as the value of Γ increases, the true robustness metric as measured through simulation can deteriorate, rendering the model counter-intuitive.

Marla [72], Marla, Vaze and Barnhart [73], Sakamoto [85] and Bryant [86] explain this phenomenon as follows. The RO model, when solved for a particular value of Γ, ensures that the constraint *at least* satisfies the chosen value of Γ while minimizing the objective (cost). In the case of large-scale network optimization problems with many constraints (and therefore many Γ values to set), discrete values of variables and multiple optima with same cost, it is possible that multiple solutions satisfy the *minimum* value of Γ but also can satisfy other (higher) values of Γ at the same cost. In other words, multiple optimal solutions can exist for a chosen value of Γ, each satisfying the *minimum* probability of constraint protection related to Γ, but also satisfy a higher value of Γ – indicating a *lack of pareto-optimality* (finding the maximum Γ at a given cost). Therefore, the solver can choose among multiple optima, a higher protection (Γ) solution can be chosen for $\Gamma = 2$ rather than for $\Gamma = 3$, resulting in a *non-monotonicity* in solution performance. This means that the placement of slack in the network is not better than for $\Gamma = 2$. This is also confirmed by observations by Iancu and Trichakis [94].

The implications of the non-monotonicity resulting from a lack of pareto-optimality are as follows: (i) it becomes more difficult to choose a value of Γ a priori (in addition to the weak bounds discussed in the previous examples), resulting in a trial-and-error testing for the 'right' value of Γ, and (ii) if multiple constraints (for example, for each individual flight) with uncertainty exist, this problem is exacerbated, because of a combinatorial manner of choosing a vector of Γ values. This strengthens our remarks from the previous case studies about the difficulty of Γ value specification.

5.2. The $\Delta - EV$ approach

The alternative approach presented in Marla [72] and Marla, Vaze and Barnhart [73] presents a way to prevent solution non-monotonicity due to lack of pareto-optimality. Because the solution search is not dependent on the input value of Γ but on a budget on the objective (in this case, a budget on propagated delay) set by the user, solutions that

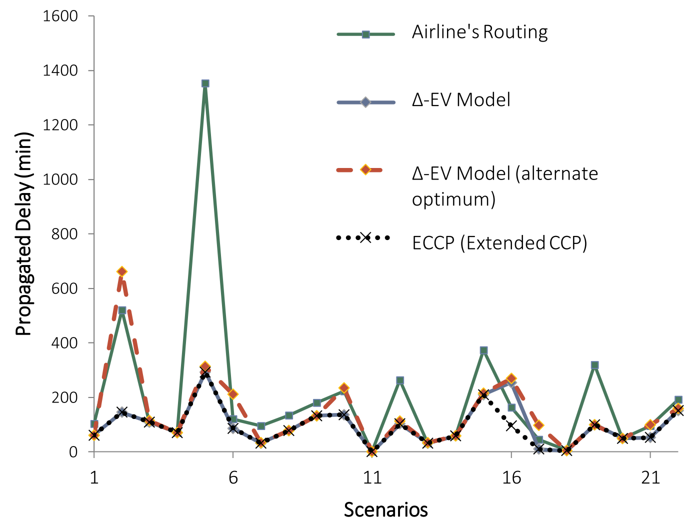

Fig. 10. Comparison of solutions from nominal, ECCP and Δ-EV approaches (Source: [19]).

satisfy the highest value of Γ possible within the budget can be found. Though specifying an acceptable level of propagated delay is also difficult to set a priori (and this approach is iterative with respect to the delay budget), it is more intuitive for manager to specify this bound from historical data. This arrangement allows solutions with better slack placement (and thus, lower propagated delay) to dominate. Among the solutions in Fig. 10, the $\Delta - EV$ model sets the cost as a budget and finds the maximum Γ solution to be chosen (Γ = 2 in this case) because it dominates other solutions. Additionally, solving the $\Delta - EV$ model requires fewer iterations and less time than the original RO approach [73].

5.3. The CCP model

To capture uncertainty using CCP, Marla and Barnhart [73] take advantage of the set-partitioning nature of the nominal formulation. Because each flight is contained in only one route in the solution, the probabilistic CCP constraint can be made deterministic by formulating it as a binary integer constraint, with each coefficient indicating the probability of delay a flight incurs on the route. Through uncertainty is in the left-hand-side of the constraint, due to the set-partitioning nature, they can avoid the need for a cumbersome formulation involving joint probability distributions [73].

The CCP model solutions exhibit the same type of non-monotonicity as RO approach solutions. As seen in Table 2, non-monotonicity in the CCP model solutions occur as well: the solution with protection percent 94 is not as robust as the solution with protection percent 92.

For each input value of protection level α, the solution is seen to satisfy the chance-constraint with *at least* probability α. However, non-monotonicity arises because the constraints can be satisfied with higher than probability α, and we do not bound from *above* the probability of violation. The non-monotonicity in performance gives rise to similar questions as in the RO approach, namely - how can we set a priori a protection level for multiple constraints if we are unable to predict the realized level of protection?

5.4. The ECCP approach

The alternative approach presented by Marla and Barnhart [73] allows the model to pick solutions with the highest constraint protection (α) possible within an allowed budget of delay. Because we now force the model to choose the solution with the higher protection level α, non-monotonicity is avoided and solutions with higher α and better slack placement dominate. This method is iterative with respect to the budget of delay, but nevertheless, requires fewer iterations than the CCP model.

5.5. CVaR approach

It is difficult to obtain a closed-form expression for CVaR for the problem. To model CVaR using a scenario-based approach, millions of scenarios should be generated to capture the uncertainty in the hundreds of thousands of parameters representing aircraft routes. However, we do not have enough information about the distribution, in particular, about joint probability distributions, to generate scenarios. In the event we were able to generate the required number of scenarios, the size of the large-scale binary program would have led to intractability. The CVaR model is thus unsuitable for this problem.

5.6. Solution comparisons and modeling insights

In Fig. 10, we compare the solution quality of the $\Delta - EV$ and ECCP model solutions against the airline's current routing. We do not include the RO approach and the CCP approach solutions because they are dominated by solutions from the $\Delta - EV$ and ECCP models. First, note that these solutions all improve upon the airline's routing [73], and generate solutions with lower delay. Second, the solutions to the $\Delta - EV$ and the ECCP models behave differently. Because of the focus on extreme-delay scenarios, the $\Delta - EV$ models (as also the RO model) produce solutions that are good for extreme scenarios but not for average scenarios. In fact, there exist multiple optimal solutions to the

Table 2
Non-monotonicity in the CCP model solutions.

	% Flight delays ≤15 min	≤60 min	≤90 min	≤120 min	Passenger disruptions Num disrupted	% disruptions reduced
α = 90 per flight	78.54	93.10	95.63	97.82	1025	6.77
α = 92 per flight	77.54	92.54	95.00	97.36	1209	−9.90
α = 94 per flight	79.54	93.73	96.00	98.18	987	10.20
Airline's routing	77.72	92.82	95.30	97.73	1100	0.00

$\Delta - EV$ model which 'protect' route subsets of the same size, but such solutions perform similarly on high delay scenarios (scenario 5) but very differently on average delay scenarios (scenarios 2, 17). On the other hand, the ECCP model because it captures delays probabilistically, focuses on high probability scenarios, and dominates over the Δ-EV model over the multiple scenarios of interest.

It is not straightforward to apply general approaches that protect using risk metrics such as 'number of uncertain parameters at worst-case' or 'probability of constraint violation' to the aircraft routing formulation; because they do not translate directly to delay metrics, and in particular, capturing how delays occur in networks. Because delays in networks are not static, in order to apply these methods, it is necessary to use domain knowledge that allows us to capture dynamic network properties such as delay propagation as parameters of the formulation. Subsequently, generic robust optimization models capturing parameter uncertainty may be applied.

Moreover, it is not straightforward to decide which methods of modeling uncertainty are most suitable. This is because the robust measures, which tend to protect constraint violation (number of coefficients at worst-case Γ and probability of violation α) are do not translate easily into delay-based metrics. The RO approach and CCP approaches are the most applicable to this problem (after modeling problem-specific characteristics in the form of delay propagation) because they can be used for large-scale problems. CVaR and other scenario-based approaches, on the other hand, are intractable for typical problem sizes of this application.

The RO and CCP approaches require multiple iterations in identifying good solutions because they are not geared to choosing the lowest *delay* solution among multiple optimal solutions with the same values of Γ and α. This is because these models do not pick the highest possible values of Γ and α for a given constraint; and moreover, these methods capture proxies for delays through Γ and α.

The $\Delta - EV$ and *ECCP* models can search the solution space to find, among multiple optimal solutions, those with the highest protection defined in terms of Γ or α. Therefore, these models also decrease the need to re-iterate to find good solutions.

While the solutions from the two modeling approaches perform similarly in many scenarios, they also differ significantly in specific scenarios. The $\Delta - EV$ model does not differentiate among solutions that perform similarly in the worst-case but very differently in the average-case; whereas the *ECCP* model, because of its probabilistic modeling, can find solutions with good performance in the more frequently occurring cases, thus limiting the total delay minutes experienced. For the real-world data of interest, the *ECCP* method solutions dominate in terms of robustness (delay) over other models. This indicates that worst-case-uncertainty based models should be used with more caution when worst-case delays are rare.

6. Lessons from the case studies: Discussion

Process of building robust solutions. Unlike for deterministic problems, the approach to generating robust solutions requires a series of steps. For completeness, we describe these here as follows: (i) solving the nominal problem or an existing implemented solution, (ii) simulating uncertainly and assess impact on feasibility/quality of the nominal solution, (iii) determining if the nominal solution needs to be improved by building in more robustness, (iv) choosing a method among multiple classes of methods such as RO, CVaR and CCP, based on the data available, (v) computing solutions with multiple robustness parameters for each method applied, (vi) evaluating solution quality of robust solutions and iterating if needed. Note that in this process, the construction of a reliable simulation tool, and updating of data used in the simulation tool, goes hand-in-hand with the development of the robust modeling frameworks. It is also important to note that generating robust solutions is an iterative process, combining the robust optimization tools and simulation.

Role of Simulation. It is key to note that simulation is a tool for 'what-if' analysis, and is *complementary* to the application of robust models, in the process of generating robust solutions. For the simulation framework to be effective, a reliable model of the underlying uncertainty is required. The model of uncertainty for the simulation can be from historical data, using a data-driven approach (as in the case of the aircraft routing case); using exact (or approximate) distributions from the underlying data if available (as in the pharmaceutical supply chain case); and/or testing using hypothetical distributions (as in the corporate portfolio case). Simulation can be used to evaluate the performance of both nominal solutions and `robust' solutions, to understand the tradeoff each methods provides in solution performance defined by robustness or mean objective value. While simulation is a common tool used in many industry settings, it cannot, for example, point to the existence (or not) of more robust solutions. Therefore, it cannot be a substitute for optimization-under-uncertainty approaches.

Different classes of robust modeling approaches can be equivalent, as described by theory. However, realizing it in practice requires careful modeling and tuning of parameters. Different classes of methods capture robustness or protection according to different definitions. As is discussed in many of the theoretical papers, however, despite protection levels being defined very differently in each method, there is equivalence among these methods. We show that, while it is difficult to realize in practice the equivalence between the methods exactly as described in theory, this can be achieved by appropriately tuning parameters; sometimes in non-intuitive ways. If fine-tuned appropriately and tested extensively using a framework in which the optimization models are accompanied by simulation, nearly all methods can perform quite well. Each robust approach trades off objective function and robustness/reliability in its own way, which is finally evident after evaluation through simulation.

For all the cases in this paper, applying approaches for optimization under uncertainty (with carefully calibrated parameters) has generated solutions that are more robust compared to the nominal solution; and have added value to the customer. In particular, these solutions protect the customer's investment in many realizations of uncertainty, while maintaining the expected performance in line with the customer's objectives and key performance metrics.

Fine tuning the various methods of capturing uncertainty can be performed through multiple techniques.

(a) *Modeling uncertainty.* Sources of uncertainty can vary significantly – for example, in the corporate portfolio case, parameters describing the return estimator are subject to uncertainty. In the pharmaceutical supply chain case, the uncertainty is in the 'risk' parameters themselves. In the aircraft routing case, we have to first model the phenomenon of delays to create uncertain parameters for the problem. This is not a case of uncertainty in 'problem parameters', but of creating parameters as a proxy for delay. While in the first two case studies, uncertainty is two-sided, in the third, uncertainty is one-sided (as it models delays). Modeling or defining uncertainty parameters differently depending on the source of uncertainty, and alternative, more efficient, formulations can be explored to represent the uncertainty in different ways.

(b) Often, the RO Approach and the CCP Approach will not give solutions that satisfy pareto-optimality in terms of the robustness parameter values; alternatively the Δ-EV (Extended RO) or Extended CCP (ECCP) can be employed to find pareto-optimal solutions

(c) Intelligently tuning parameters such as uncertainty ranges (RO and Δ-EV) and quantiles of uncertainty (ECCP), and recognizing that more parameters are tunable beyond the robustness parameters. Because some values of uncertainty ranges can lead to over-conservatism, it is important to recognize that those inputs can be tuned as well, for a more appropriate/less conservative tradeoff. Moreover, instead of using the robustness parameters to tune models, robustness can be defined in terms of the robustness budget available to the practitioner, through the Δ-EV and ECCP methods. We discuss these in more detail in the next lesson.

Finding robust solutions is an iterative and interactive process.

Robust approaches aim to translate a KPI of robustness desired by the analyst or customer into a risk measure of the approach. If the robustness metric is the same as a standard risk measure (such as variance, CVaR, VaR), standard methods might be easily chosen that optimize based on these metrics. For non-standard KPIs (such as delays in the aircraft routing case) or non-coherent risk measures (such as VaR) it is hard to identify the best approach or the best way to define uncertainty sets (as mentioned in Bertsimas and Thiele [95]).

Generating uncertainty sets and parameters for models from historical data is a challenge in practice. For all the different robust modeling approaches, using data-driven ways to describe uncertainty sets or parameters using qualitative (is the distribution heavy-tailed or light-tailed?) or quantitative information (90th quantile) is necessary. Iterating among parameter values, to strike the most acceptable tradeoff between the multiple performance metrics, is inevitable.

Bertsimas and Sim's RO approach: From empirical experience, it is difficult to specify a value for the risk parameter Γ, as it is not described by tight bounds for the probability of violation. In fact, the approach is seen to be highly conservative, especially so in the case of integer programs. Therefore it is suitable for highly risk-averse users. However, the bound also provides a helpful starting point to choose a better value of Γ.

We can find the 'most suitable' cost-robustness tradeoff for the manager by repeatedly solving the model for different values of Γ. This is possible for small-to-medium sizes like the corporate portfolio problem and the pharmaceutical supply chain problem. For larger-size problems, such as aircraft routing, repeated re-solving is cumbersome and possibly even intractable. For such problems we recommend the use of the $\Delta - EV$ model. Additionally, controlling the level of conservatism can be done in a data-driven way, to guide the choice of uncertainty sets (corporate portfolio and pharmaceutical supply chain case studies).

$\Delta - EV$ *model:* When specifying Γ becomes cumbersome in the RO approach, such as for large-scale problems, we recommend the use of the $\Delta - EV$ model that drives the trade-off between cost and robustness through the budget constraint [73]. The $\Delta - EV$ model solutions dominate those of the RO model solutions for fixed values of 'robustness budget' and uncertainty sets.

CCP is applicable only when some knowledge of the underlying uncertainty distribution is available, either in partial or complete form. Quantile information from the manager and system experts may also be used. CCP is more intuitive to interpret by the manager but more difficult to find an equivalent deterministic form (unless scenarios from large amount of data are generated [41]). In the corporate portfolio case, CCP resulted in a quadratic deterministic formulation; whereas in the pharmaceutical supply chain case, an additional proxy constraint is added, to capture left-hand-side uncertainty. Finally, in the aircraft routing case, set-partitioning structure of the constraint containing uncertain parameters allowed a simpler deterministic formulation [73].

ECCP model: We recommend the use of the ECCP model when it is difficult to specify the probability of constraint violation for multiple constraints, or when one among multiple optima of the CCP model needs to be chosen. ECCP solutions dominate in terms of robustness, over solutions of the CCP.

CVaR is a good approach to use for small to medium-scale problems where a large quantity of data is available to help define the uncertainty through possible scenarios. However, translating the business indicator into a numerical value for the critical value is not straightforward, as the value of CVaR is not intuitive. The solution process can be iterative because VaR and CVaR are simultaneously optimized for. For a small number of uncertain parameters (corporate portfolio case), the CVaR formulation may be possible to solve but quickly becomes intractable for medium (pharmaceutical supply chain) or larger (aircraft routing) problems.

Recommendation: We suggest that when it becomes difficult to set the robustness parameters for multiple constraints, the manager can be more easily able to make decisions by specifying a cost budget of uncertainty, described in terms of the objective function (return, or cost) of the deterministic formulation. We have provided pointers to alternative models proposed by Marla [72] and Marla, Vaze and Barnhart [73] that define robustness in terms of 'loss in nominal objective function to gain robustness'. Particularly in the case of problems with *multiple optima* (such as network-based settings) we recommend thorough investigation of multiple optimal solutions satisfying the same values of Γ or α, but with different values of robustness metrics, or multi-objective optimization. In these scenarios, the $\Delta - EV$ and ECCP models become more relevant for application.

Criteria in choosing robust modeling approaches are ease of implementation, ease of determining sources of uncertainty and ease of characterizing them, intuitiveness of approach in problem setting, applicability of approach to the setting, level of conservatism showed, and tractability. In practice, also, given a limited budget of investment for software packages that implement such methods, investment for practical implementation should also be considered. Therefore, we recommend that practitioners certainly set up a reliable simulation system to accompany the evaluation of robust solutions, as well as implement method(s) in the order of ease of implementation in the organization.

We recommend that approaches be used in the order of ease of implementation. For example, for the case of the three classes of approaches discussed here, we recommend first the application of the RO approach, as the framework of translating the nominal formulation to robust formulation is simplest for all linear and integer programs. Second, we suggest an approach like CCP which captures more details about data (even partial information in the form of quantiles) but requires more effort in formulating. However, if RO or CCP face tractability or pareto-optimality issues, we recommend $\Delta - EV$ and ECCP as alternatives. Finally, CVaR requires more modeling effort. We recommend this when large data sets are available to help describe the uncertainty. Using this sequence of approaches helps us to first judge the benefits of capturing uncertainty. It also indicates the possible necessity of collecting further data to explore the nature of uncertainty, in order to investigate the use of more complex and sophisticated models requiring more data to capture uncertainty.

It should be of reassurance to a practitioner that they can choose a method that can be easiest implemented in their systems, or easily interpretable to the manager, or more natural or intuitive to developers. By appropriately tuning multiple parameters used as inputs to these methods and testing them via a reliable simulation framework, tradeoff curves between robustness parameters and mean objective values can be generated, allowing the manager to choose a suitable operating point.

Author credit

Lavanya Marla: conceptualization, formulation (case studies 2 and 3), data curation and lead writing; **Alex Rikun:** formulation, solving, case study 1; **Gautier Stauffer:** validation, investigation, resources, data sources, second lead in writing; **Eleni Pratisini:** data sources, supervision, project administration, funding acquisition

Declaration of competing interest

None.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.orp.2020.100150.

Appendix A. Formulations for the Corporate Portfolio Case

A.1 RO Approach Formulation and Solutions

$$\max_{x}\left(\min_{(\tilde{a},\tilde{b})}\sum_{i\in I}(\tilde{a}_i+\tilde{b}_i x_i)\right) \tag{17}$$

$$\begin{pmatrix}\tilde{a}_i\\ \tilde{b}_i\end{pmatrix}=\begin{pmatrix}\overline{a_i}\\ \overline{b_i}\end{pmatrix}+G\begin{pmatrix}y_1^i\\ y_2^i\end{pmatrix}\ \forall\ i\in I \tag{18}$$

$$l_i\le x_i\le u_i\ \forall\ i\in I \tag{19}$$

$$0\le\left|y_j^i\right|\le z^i\le 1\quad\forall\ j\in\{1,2\},\ \forall\ i\in I \tag{20}$$

$$\sum_i z^i\le\Gamma \tag{21}$$

$$x_i\ge 0\quad\forall\ i\in I \tag{22}$$

$$y_j^i\ge 0\ \ \forall\ j\in\{1,2\},\ \forall\ i\in I \tag{23}$$

In the above formulation, G is the standard deviation matrix and $GG'=C$, the covariance matrix as defined in Section 3.1.

To complement Fig. 3, we study the mean-variance tradeoffs when the same values of variance and covariance are input into the RO Approach, but the underlying data actually arises from various distributions – specifically, uniform, normal and discrete distributions. As discussed in Section 3.1, we again find (see Fig. 11) that using a convex subset of the uncertainty set results in a less conservative tradeoff between the mean and standard deviation.

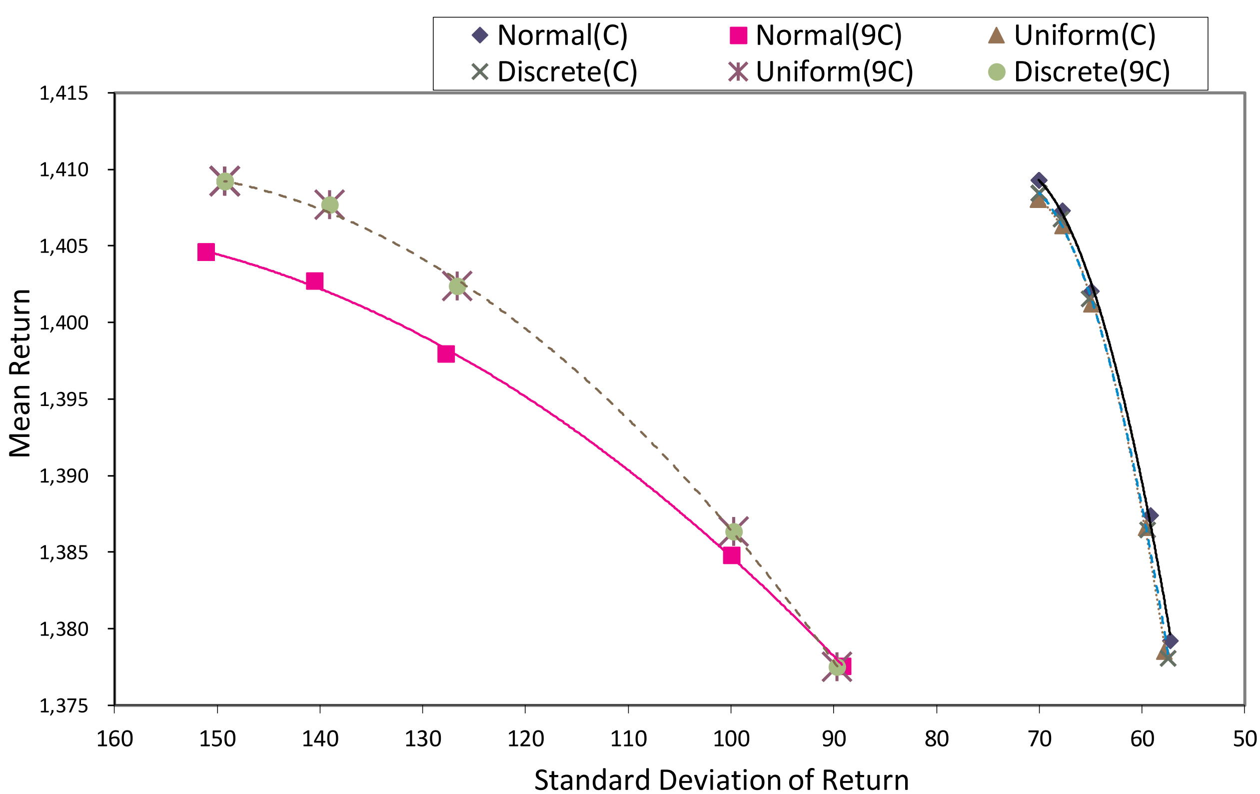

Fig. 11. Mean-standard deviation tradeoff of Return.

A.2 CCP Approach Formulation

$$\max_{x} \sum_{i\in I} (\overline{a_i} + \overline{b_i}x_i) \tag{24}$$

$$\sum_{i\in I} (\overline{a_i} + \overline{b_i}x_i) + \Phi^{-1}(1-\alpha)\sqrt{\sum_{i\in I} 1, x^T C_i 1, x} \geq c.\ v. \tag{25}$$

$$\sum_{i\in I} x_i \leq B \tag{26}$$

$$l_i \leq x_i \leq u_i \quad \forall \quad i \in I \tag{27}$$

$$x_i \geq 0 \ for\ all\ i \in I \tag{28}$$

Here $c.v.(\alpha)$ is the critical value (obtained from statistical data) that the return from the portfolio should exceed, with probability α. Φ^{-1} indicates the inverse cumulative standard normal function. Note that $\alpha = 0.5$ is equivalent to the nominal problem, and for $\alpha \geq 0.5$, we obtain a second-order convex formulation.

A.3 CVaR Approach Formulation and Solutions

$$\max_{x, z_j, \beta} \sum_{i\in I} (\overline{a_i} + \overline{b_i}x_i) \tag{29}$$

$$\sum_{i\in I} (\widetilde{a_{ij}} + \widetilde{b_{ij}}x_i) + z_j \geq \beta \ \ \forall j \ \ 1, \ldots, M \tag{30}$$

$$\beta + \frac{1}{(1-\alpha)M} \sum_{j=1}^{M} z_j \geq c.\ v.\ (\alpha) \tag{31}$$

$$\sum_{i\in I} x_i \leq B \tag{32}$$

$$l_i \leq x_i \leq u_i \quad \forall \quad i \in I \tag{33}$$

$$x_i \geq 0 \ for\ all\ i \in I \tag{34}$$

$$z_j \geq 0 \quad \forall\, j = 1, \ldots, M \tag{35}$$

(29)–(35) is the CVaR formulation of the nominal formulation (1)–(4). Here, $\widetilde{a_{ij}}$ and $\widetilde{b_{ij}}$ represent the realizations of the uncertain parameters as observed in scenario j of the M sampled scenarios. (30) and (31) together approximate the true CVaR equation. z_j is a dummy variable that helps in the approximation. β is the approximation of the VaR value when CVaR is constrained as shown in this formulation. The remaining constraints are from the nominal formulation.

Appendix B. Formulations for the Pharmaceutical Supply Chain Case

B.1 RO Approach Formulation

According to this model, the parameters $H(t, s, e)$ are assumed to realize values in a range of uncertainty $\hat{H}(t, s, e)$ around the mean hazard rates $\bar{H}(t, s, e)$. That is, the hazard rates take values in the interval $[\bar{H}(t, s, e) - \hat{H}(t, s, e), \bar{H}(t, s, e) + \hat{H}(t, s, e)]$. The uncertainty set in this model is defined as the case when Γ of the hazard rate parameters realize values at the worst-case bounds of their uncertainty ranges. The linearized formulation, as described in Bertsimas and Sim [68], is as follows.

$$\max_{x\in X}\left(\sum_{p,t,s,e} R(p, t, s, e)x(p, t, s, e) - \sum C(p, t, s, e)x(p, t, s, e)\right) \tag{36}$$

$$s.\ t. \sum_{p,t,s,e} \bar{H}(t, s, e)R(p, e)x(p, t, s, e) + z\Gamma + v(t, s, e) \leq CV \tag{37}$$

$$z + v(t, s, e) \geq y(t, s, e) \quad \forall \quad t, s, e \tag{38}$$

$$-y(t, s, e) \leq \left(\sum_{p} x(p, t, s, e)R(p, e)\right)\hat{H}(t, s, e) \quad \forall\, t, s, e \tag{39}$$

$$\left(\sum_{p} x(p, t, s, e)R(p, e)\right)\hat{H}(t, s, e) \leq y(t, s, e) \quad \forall\, t, s, e \tag{40}$$

$$\sum_{t,s} x(p, t, s, e) \leq 1 \quad \forall\, p, \ e \tag{41}$$

$$x(p,\ t,\ s,\ e) \in \{0,\ 1\} \quad \forall\ p,\ \ t,\ \ s,\ \ e \tag{42}$$

$$v(t,\ s,\ e) \geq 0 \quad \forall\ t,\ s,\ e \tag{43}$$

$$y(t,\ s,\ e) \geq 0 \quad \forall\ t,\ s,\ e \tag{44}$$

B.2 $\Delta - EV$ Formulation

Note that in the $\Delta - EV$ model, we aim to maximize the number of uncertain coefficients that can take on their worst-case values, ensuring costs within a robustness budget δ. Therefore, for this model, we order the ranges of the uncertain coefficients $\hat{H}(t,\ s,\ e)R(p,\ e)$ in increasing order. After ordering, the rank of the (p, t, s, e)-th coefficient is denoted by (k, p, t, s, e). Also, the original index (p, t, s, e) of the variable that takes the kth position in the sorted $\hat{H}(t,\ s,\ e)R(p,\ e)$ values is denoted by (p, t, s, e, k). Thus, the value K of the index in the last position in the sorted list is described by $K = |P| + |T| + |S| + |E|$. We define Δ equal to the maximum number of variables $x(p, t, s, e)$ in the solution with $x = 1$ whose coefficient values must assume their nominal values for the solution to remain feasible.

Let P^* be the optimal profit of the nominal problem (9)–(12). Let δ *be* the user-specified incremental cost that is acceptable for increased robustness, that is, the profit of a robust solution from the Delta formulation is at least $P^* - \delta$. Let variables $v(p, t, s, e)$ equal 1 if the uncertain coefficient $H(t, s, e)R(p, e)$ is not allowed to take on its extreme value, and takes on its nominal value in the solution of the Delta model. Variables $w(k)$ equal 1 for all k for which there exists a $l \geq k$ with $v(p,\ t,\ s,\ e,\ l) = 1$. $w(k)$ variables in the sorted order of $\hat{H}(t,\ s,\ e)R(p,\ e)$ values follow a step function. This leads to the $\Delta - EV$ formulation, which is as follows.

$$\min_{x \in X} \Delta \tag{45}$$

$$s.\ t.\ \sum_{p,t,s,e} R(p,\ t,\ s,\ e)x(p,\ t,\ s,\ e) - \sum C(p,\ t,\ s,\ e)x(p,\ t,\ s,\ e) \geq P^* - \delta \tag{46}$$

$$\Delta \geq \sum_{p,t,s,e} v(p,\ t,\ s,\ e) \tag{47}$$

$$v(p,\ t,\ s,\ e) \leq x(p,\ t,\ s,\ e) \ \ \forall\ p,\ t,\ s,\ e \tag{48}$$

$$v(p,\ t,\ s,\ e;\ k) \leq w(k)\ \forall\ k = 1,\ \ldots,K \tag{49}$$

$$v(p,\ t,\ s,\ e) \geq x(p,\ t,\ s,\ e) + w(k;\ p,\ t,\ s,\ e) - 1\ \ \forall\ p,\ t,\ s,\ e \tag{50}$$

$$w(k+1) \leq w(k)\ \forall\ k = 1,\ \ldots,K \tag{51}$$

$$w(0) = 1 \tag{52}$$

$$w(K+1) = 0 \tag{53}$$

$$\sum_{t,s} x(p,\ t,\ s,\ e) \leq 1 \quad \forall\ p,\ \ e \tag{54}$$

$$x(p,\ t,\ s,\ e) \in \{0,\ 1\} \quad \forall\ p,\ t,\ s,\ e \tag{55}$$

$$v(p,\ t,\ s,\ e) \in [0,\ 1] \quad \forall\ p,\ t,\ s,\ e \tag{56}$$

$$w(k) \in \{0,\ 1\}\ \forall\ k = 0,\ \ldots,K \tag{57}$$

B.3 ECCP Formulation

$$\max \alpha \tag{58}$$

$$s.\ t.\ \sum_{p,t,s,e} R(p,\ t,\ s,\ e)x(p,\ t,\ s,\ e) - \sum C(p,\ t,\ s,\ e)x(p,\ t,\ s,\ e) \geq P^* - \delta \tag{59}$$

$$\sum_{p,t,s,e} \bar{H}(t,\ s,\ e)R(p,\ e)x(p,\ t,\ s,\ e) \leq CV \tag{60}$$

$$\sum_{p,t,s,e} F^{-1}_{H(t,s,e)}(\beta)R(p,\ e)x(p,\ t,\ s,\ e) \leq \sum_{k=1}^{K} F^{-1}_{CV,CCP}(\alpha_k)(y_k - y_{k-1}) \tag{61}$$

$$\sum_{t,s} x(p,\ t,\ s,\ e) \leq 1 \quad \forall\ p,\ \ e \tag{62}$$

$$y_k \geq y_{k-1} \tag{63}$$

$$y_0 = 0 \tag{64}$$

$$y_K = 1 \tag{65}$$

$$\alpha \le \sum_{k=1}^{K} \alpha_k (y_k - y_{k-1}) \quad (66)$$

$$x(p, t, s, e) \in \{0, 1\} \quad \forall\, p, t, s, e \quad (67)$$

$$y_k \in \{0, 1\} \quad \forall\, k = 1, \ldots, K \quad (68)$$

The extended chance-constrained model builds on the chance-constraints in (14)-(15). We assume that some quantiles α_k, $k = 1, \ldots, K$ of the critical value CV_{CCP} are known, from analysis of historical data. Instead of choosing one particular value of α_k, we try to attain the highest protection level possible, within a budget δ on the profit. y_k are binary variables that equal 1 if the protection level α_k is attained by the solution. The objective (58) maximizes the protection level realized by the solution, which is described by (61) and (65). The protection level variables y_k take the form of a step function.

B.4 Supplementary Results under various distributions of uncertainty

Figs. 12, 13, 14

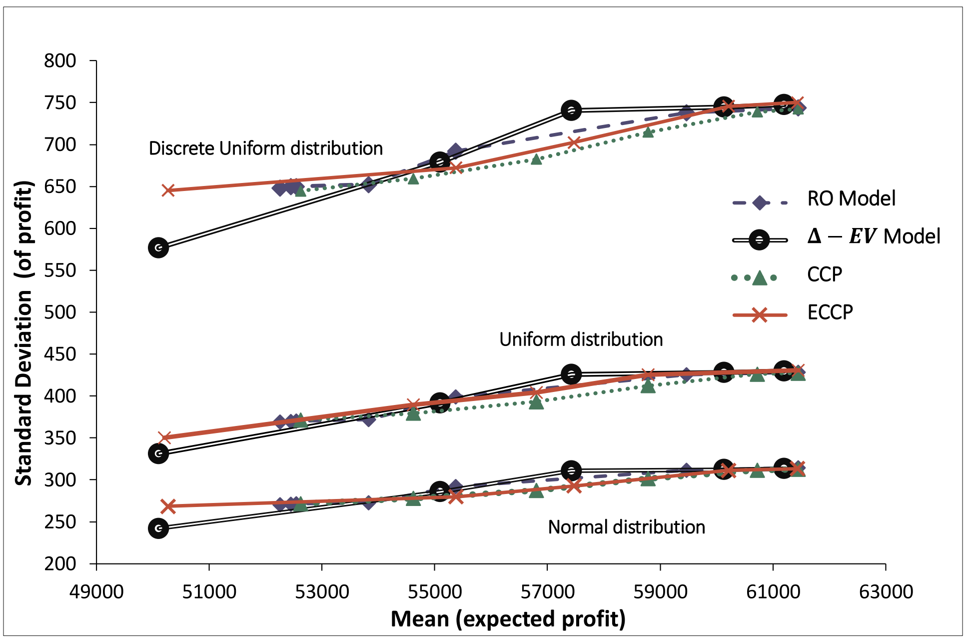

Fig. 12. Mean-standard deviation tradeoffs using the four approaches, for normal, uniform and discrete uniform distributions. The input to optimization uses a range of $\hat{H}=0.04$, and the simulation to test the robustness also uses 0.04 as the standard deviation of uncertainty. The robust parameter values are the same as in Fig. 6; and the same budgets are used for the Δ-EV and ECCP models.

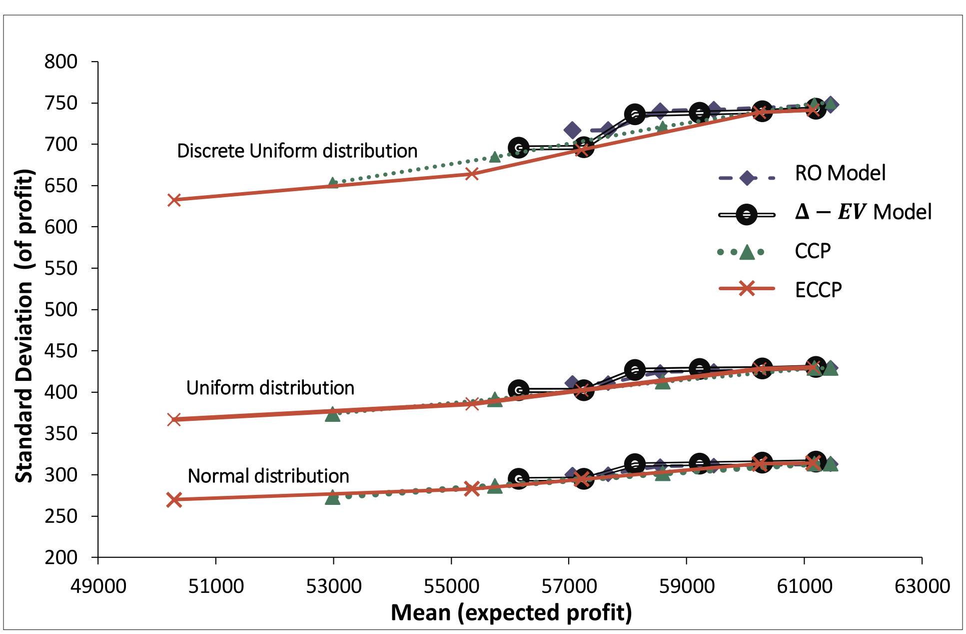

Fig. 13. Mean-standard deviation tradeoffs using the four approaches, for normal, uniform and discrete uniform distributions. The optimization input has a range of $\hat{H} = 0.02$, and simulation uses 0.04 as the std. deviation. All parameters same as Fig. 12.

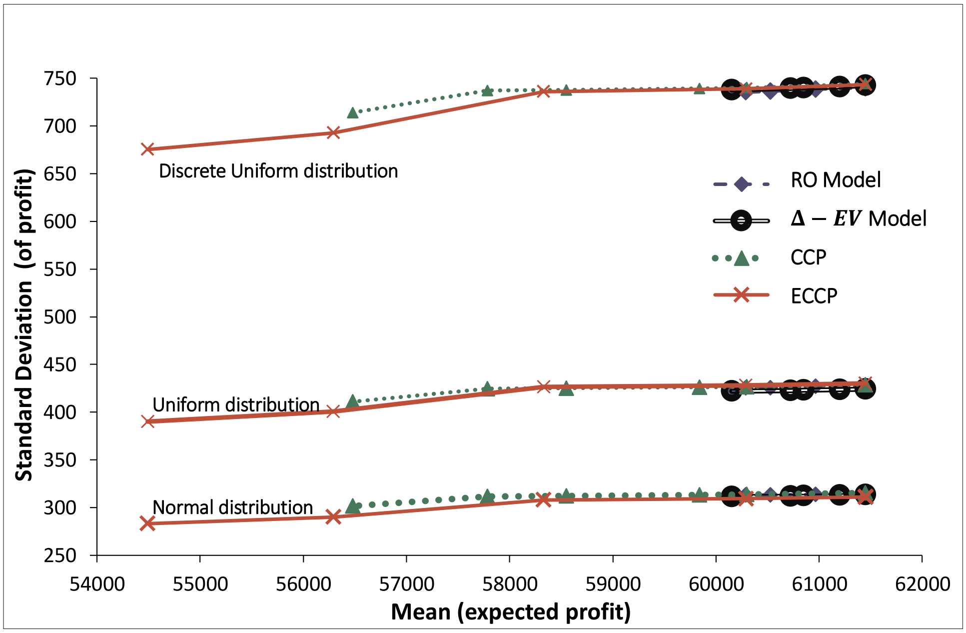

Fig. 14. Mean-standard deviation tradeoffs using the four approaches, for normal, uniform and discrete uniform distributions. The optimization input has a range of $\hat{H} = 0.02$, and simulation uses 0.04 as the std. deviation. All parameters same as Fig. 12.

References

[1] Brown GG, Rosenthal RE. Optimization tradecraft: hard-won insights from real-world decision support* Boston, MA: Springer; 2010. p. 99–114.
[2] Mulvey JM, Vanderbei RJ, Zenios SA. Robust optimization of large-scale systems. Oper Res 1995;43(2):264–81.
[3] Greenberg HJ. Representing uncertainty in decision support. OR/MS Today. INFORMS; 2007.
[4] Lejeune MA, Shen S. Multi-objective probabilistically constrained programs with variable risk: models for multi-portfolio financial optimization. Eur J Oper Res 2016;252(2):522–39.
[5] Mulvey JM, Erkan HG. Applying cvar for decentralized risk management of financial companies. J Bank Financ 2006;30(2):627–44.
[6] Trindade AA, Uryasev S, Shapiro A, Zrazhevsky G. Financial prediction with constrained tail risk. J Bank Financ 2007;31(11):3524–38.
[7] Bogentoft E, Romeijn EH, Uryasev S. Asset/Liability management for pension funds using CVaR constraints. J Risk Finance 2001;3(1):57–71.
[8] Goldfarb D, Iyengar G. Robust portfolio selection problems. Math Oper Res 2003;28(1):1–38.
[9] (Frank) Chen Y, Xu M, Zhang ZG. Technical note—a risk-averse newsvendor model under the CVaRcriterion. Oper Res 2009;57(4):1040–4.
[10] Gönsch J, Hassler M. Optimizing the conditional value-at-risk in revenue management. Rev Manag Sci 2014;8(4):495–521.
[11] Gönsch J. A survey on risk-averse and robust revenue management. Eur J Oper Res 2017;263(2):337–48.
[12] Perakis G, Roels G. Robust controls for network revenue management. Manuf Serv Oper Manag 2010;12(1):56–76.
[13] Rusmevichientong P, Topaloglu H. Robust assortment optimization in revenue management under the multinomial logit choice model. Oper Res 2012;60(4):865–82.
[14] Lan Y, Gao H, Ball MO, Karaesmen I. Revenue management with limited demand information. Manag Sci 2008;54(9):1594–609.
[15] Serraino G, Uryasev S. Conditional value-at-risk (CVaR). Encyclopedia of operations research and management science. Boston, MA: Springer US; 2013. p. 258–66.
[16] Bandi C, Bertsimas D. Tractable stochastic analysis in high dimensions via robust optimization. Math Program 2012;134(1):23–70.
[17] Bandi C, Bertsimas D, Youssef N. Robust queueing theory. Oper Res 2015;63(3):676–700.
[18] Lopatatzidis S, De Bock J, de Cooman G, De Vuyst S, Walraevens J. Robust queueing theory: an initial study using imprecise probabilities. Queueing Syst 2016;82(1–2):75–101.
[19] Marla L, Vaze V, Barnhart C. Robust optimization: lessons learned from aircraft routing. Comput Oper Res 2018;98:165–84.
[20] Marla L. Robust optimization for network-based resource allocation problems under uncertainty Master's Thesis 2007
[21] Erera AL, Morales JC, Savelsbergh M. Robust optimization for empty repositioning problems. Oper Res 2009;57(2):468–83.
[22] Sohoni M, Lee Y-C, Klabjan D. Robust airline scheduling under block-time uncertainty. Transp Sci 2011;45(4):451–64.
[23] Yan C, Kung J. Robust aircraft routing. Transp Sci 2018;52(1):118–33.
[24] Shen S, Smith JC, Ahmed S. Expectation and chance-constrained models and algorithms for insuring critical paths. Manag Sci 2010;56(10):1794–814.
[25] Deng Y, Shen S. Decomposition algorithms for optimizing multi-server appointment scheduling with chance constraints. Math Program 2016;157(1):245–76.
[26] Shen S, Kurt M, Wang J. Chance-Constrained programming models and approximations for general stochastic bottleneck spanning tree problems. INFORMS J Comput 2015;27(2):301–16.
[27] Bertsimas D, Sim M. Robust discrete optimization and network flows. Math Program 2003;98(1–3):49–71.
[28] Panda D, Kar S, Maity K, Maiti M. A single period inventory model with imperfect production and stochastic demand under chance and imprecise constraints. Eur J Oper Res 2008;188(1):121–39.
[29] Bertsimas D, Thiele A. A robust optimization approach to inventory theory. Oper Res 2006;54(1):150–68.
[30] Zhang Y, Shen S, Mathieu J. Distributionally robust chance-constrained optimal power flow with uncertain renewables and uncertain reserves provided by loads. IEEE Trans Power Syst 2016:1.
[31] Tavakoli M, Shokridehaki F, Funsho Akorede M, Marzband M, Vechiu I, Pouresmaeil E. CVaR-based energy management scheme for optimal resilience and operational cost in commercial building microgrids. Int J Electr Power Energy Syst 2018;100:1–9.
[32] Bertsimas D, Litvinov E, Sun XA, Zhao J, Zheng T. Closure to the discussion of 'Adaptive robust optimization for the security constrained unit commitment problem,' IEEE Trans Power Syst 2014;29(2):996–7.
[33] Lorca A, Sun XA. The adaptive robust multi-period alternating current optimal power flow problem. IEEE Trans Power Syst 2018;33(2):1993–2003.
[34] Dantzig GB. Linear programming under uncertainty. Manage Sci 1955;1(3–4):197–206.
[35] Beale EML. On minimizing a convex function subject to linear inequalities. Royal Stat Soc 1955;17(2):173–84.
[36] Sahinidis NV. Optimization under uncertainty: state-of-the-art and opportunities. Comput Chem Eng 2004;28(6–7):971–83.
[37] Birge JR, Louveaux F. Introduction to stochastic programming. New York, NYNew York: Springer; 2011.
[38] Kall P, Mayer J. Stochastic linear programming: models, theory, and computation. Springer Science; 2005.
[39] Charnes A, Cooper WW, Symonds GH. Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil. Manag Sci 1958;4(3):235–63.
[40] Charnes A, Cooper WW. Chance-Constrained programming. Manag Sci 1959;6(1):73–9.
[41] A. Nemirovski and A. Shapiro, "Scenario approximations of chance constraints", in *Probabilistic and randomized methods for design under uncertainty*, London: Springer-Verlag, pp. 3–47.
[42] Nemirovski A, Shapiro A. Convex approximations of chance constrained programs. SIAM J Optim 2007;17(4):969–96.
[43] Miller BL, Wagner HM. Chance constrained programming with joint constraints. Oper Res 1965;13(6):930–45.
[44] Küçükyavuz S. On mixing sets arising in chance-constrained programming. Math Program 2012;132(1–2):31–56.
[45] Ben-Israel A. On some problems of mathematical programming. Northwestern University; 1962.
[46] Bawa VS. On chance constrained programming problems with joint constraints. Manage Sci 1973;19(11):1326–31.
[47] Jagannathan R. Chance-Constrained programming with joint constraints. Oper Res 1974;22(2):358–72.

[48] Prékopa A. Contributions to the theory of stochastic programming. Math Program 1973;4(1):202–21.
[49] Song Y, Luedtke JR. Branch-and-cut approaches for chance-constrained formulations of reliable network design problems. Math ProgrammComput 2013;5(4):397–432.
[50] Song Y, Luedtke JR, Küçükyavuz S. Chance-Constrained binary packing problems. INFORMS J Comput 2014;26(4):735–47.
[51] Calafiore G, Campi MC. Uncertain convex programs: randomized solutions and confidence levels. Math Program 2005;102(1):25–46.
[52] Calafiore GC, Campi MC. The scenario approach to robust control design. IEEE Trans Autom Control 2006;51(5):742–53.
[53] Campi MC, Garatti S. A sampling-and-discarding approach to chance-constrained optimization: feasibility and optimality. J Optim Theory Appl 2011;148(2):257–80.
[54] Luedtke J, Ahmed S, Nemhauser GL. An integer programming approach for linear programs with probabilistic constraints. Math Program 2010;122(2):247–72.
[55] Sen S. Relaxations for probabilistically constrained programs with discrete random variables. Oper Res Lett 1992;11(2):81–6.
[56] Dentcheva D, Prékopa A, Ruszczynski A. Concavity and efficient points of discrete distributions in probabilistic programming. Math Program 2000;89(1):55–77.
[57] Beraldi P, Ruszczyński A. A branch and bound method for stochastic integer problems under probabilistic constraints. Optim Methods Softw 2002;17(3):359–82.
[58] Ruszczyski A. Probabilistic programming with discrete distributions and precedence constrained knapsack polyhedra. Math Program 2002;93(2):195–215.
[59] Artzner P, Delbaen F, Eber J-M, Heath D. Coherent measures of risk. Math Finance 1999;9(3):203–28.
[60] Krokhmal P, Uryasev S, Palmquist J. Portfolio optimization with conditional value-at-risk objective and constraints. J Risk 2001;4(2):43–68.
[61] Rockafellar RT, Uryasev S. Optimization of conditional value-at-risk. J Risk 2000;2:21–41.
[62] Koenig M, Meissner J. List pricing versus dynamic pricing: impact on the revenue risk. Eur J Oper Res 2010;204(3):505–12.
[63] Koenig M, Meissner J. Risk management policies for dynamic capacity control. Comput Oper Res 2015;59:104–18.
[64] Pflug GC. Some remarks on the value-at-risk and the conditional value-at-risk. In: Uryasev S, editor. Probabilistic constrained optimization: methodology and applications. Kluwer Academic Publishers; 2000. p. 272–81.
[65] Soyster AL. Technical note—convex programming with set-inclusive constraints and applications to inexact linear programming. Oper Res Oct. 1973;21(5):1154–7.
[66] Ben-Tal A, Nemirovski A. Robust solutions of uncertain linear programs. Oper Res Lett 1999;25:1–13.
[67] Ben-Tal A, Nemirovski A. Robust convex optimization. Math Oper Res 1998;23(4):769–805.
[68] Bertsimas D, Sim M. The price of robustness. Oper Res 2004;52(1):35–53.
[69] Ben-Tal A, Nemirovski A. Robust solutions of linear programming problems contaminated with uncertain data. Math Program 2000;88(3):411–24.
[70] Ben-Tal A, El Ghaoui L, Nemirovski AS. Robust optimization. Princeton University Press; 2009.
[71] Bertsimas D, Brown DB. Constructing uncertainty sets for robust linear optimization. Oper Res 2009;57(6):1483–95.
[72] Marla L. Airline schedule planning and operations: optimization-based approaches for delay mitigation. Massachusetts Institute of Technology; 2010.
[73] Marla L, Vaze V, Barnhart C. Robust optimization: lessons learned from aircraft routing. Comput Oper Res 2018;98:165–84.
[74] Bertsimas D, Gupta V, Kallus N. Data-driven robust optimization. Math Program 2018;167(2):235–92.
[75] Gabrel V, Murat C, Thiele A. Recent advances in robust optimization: an overview. Eur J Oper Res 2014;235(3):471–83.
[76] Delage E, Ye Y. Distributionally robust optimization under moment uncertainty with application to data-driven problems. Oper Res 2010;58(3):595–612.
[77] Wiesemann W, Kuhn D, Sim M. Distributionally robust convex optimization. Oper Res 2014;62(6):1358–76.
[78] Hanasusanto GA, Roitch V, Kuhn D, Wiesemann W. A distributionally robust perspective on uncertainty quantification and chance constrained programming. Math Program 2015;151(1):35–62.
[79] Esfahani PM, Kuhn D. Data-driven distributionally robust optimization using the Wasserstein metric: performance guarantees and tractable reformulations. Math Program 2018;171(1–2):115–66.
[80] Goh J, Sim M. Distributionally robust optimization and its tractable approximations. Oper Res 2010;58(no. 4–part–1):902–17.
[81] Chen W, Sim M, Sun J, Teo C-P. From cvar to uncertainty set: implications in joint chance-constrained optimization. Oper Res 2010;58(2):470–85.
[82] Lilien GL, Rangaswamy A. Marketing engineering : computer-assisted marketing analysis and planning. 2nd ed. Upper Saddle River, NJ: Prentice-Hall; 2004.
[83] Lu JW, Beamish PW. International diversification and firm performance: the S-curve hypothesis. Acad Manag J 2004;47(4):598–609.
[84] Epiney L. Corporate portfolio management: A robust optimization perspective. Masters Thesis IBM Research ZRL and Ecole Polytechnique Federale de Lausanne; 2007.
[85] Sakamoto P. UAV mission planning under uncertainty Masters Thesis Massachusetts Institute of Technology; 2006
[86] Bryant CH. Robust Planning for Effects-Based Operations Masters Thesis Massachusetts Institute of Technology; 2006
[87] Rockafellar RT, Uryasev S. Conditional value-at-risk for general loss distributions. J Bank Financ 2002;26(7):1443–71.
[88] "Top 25 pharmaceutical company pipeline analysis and sales projections to 2025 - KaloramaInformation." [Online]. Available:https://kaloramainformation.com/product/top-25-pharmaceutical-company-pipeline-analysis-and-sales-projections-to-2025/. [Accessed: 24 July 2019].
[89] Gallay O. Robustness and regulatory risk in supply chains. Masters Thesis IBM Zurich Research Laboratory and Ecole Polytechnique Federale de Lausanne; 2005
[90] Barnhart C, Belobaba P, Odoni AR. Applications of operations research in the air transport industry. Transp Sci 2003;37(4):368–91.
[91] Clarke L, Johnson E, Nemhauser G, Zhu Z. The aircraft rotation problem. Ann Oper Res 1997;69(0):33–46.
[92] C. Barnhart, "Airline schedule optimization", in *The global airline industry*, Chichester, UK: John Wiley & Sons, Ltd, pp. 183–211.
[93] Lan S, Clarke J-P, Barnhart C. Planning for robust airline operations: optimizing aircraft routings and flight departure times to minimize passenger disruptions. Transp Sci 2006;40(1):15–28.
[94] Iancu DA, Trichakis N. Pareto efficiency in robust optimization. Manag Sci 2014;60(1):130–47.
[95] Bertsimas D, Thiele A. Robust and data-driven optimization: modern decision making under uncertainty. Models, methods, and applications for innovative decision making. INFORMS; 2006. p. 95–122.