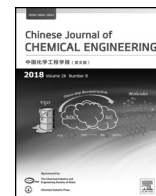




Contents lists available at ScienceDirect

Chinese Journal of Chemical Engineering

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

Review

Process optimization with consideration of uncertainties—An overview

Ying Chen, Zhihong Yuan, Bingzhen Chen *

Department of Chemical Engineering, Tsinghua University, Beijing 10084, China

ARTICLE INFO

Article history:

Received 5 September 2017

Accepted 13 September 2017

Available online 28 September 2017

Keywords:

Optimization under uncertainty

Robust optimization

Stochastic programming

Chance constrained programming

Data-driven optimization

ABSTRACT

Optimization under uncertainty is a challenging topic of practical importance in the Process Systems Engineering. Since the solution of an optimization problem generally exhibits high sensitivity to the parameter variations, the deterministic model which neglects the parametric uncertainties is not suitable for practical applications. This paper provides an overview of the key contributions and recent advances in the field of process optimization under uncertainty over the past ten years and discusses their advantages and limitations thoroughly. The discussion is focused on three specific research areas, namely robust optimization, stochastic programming and chance constrained programming, based on which a systematic analysis of their applications, developments and future directions are presented. It shows that the more recent trend has been to integrate different optimization methods to leverage their respective superiority and compensate for their drawbacks. Moreover, data-driven optimization, which combines mathematical programming methods and machine learning algorithms, has become an emerging and competitive tool to handle optimization problems in the presence of uncertainty based on massive historical data.

© 2017 The Chemical Industry and Engineering Society of China, and Chemical Industry Press. All rights reserved.

1. Introduction

Over the past several decades, people have realized the significance of accounting for uncertainty in the process optimization. Uncertainties are inevitable in the realistic chemical industries, ranging from process design and synthesis, to operational planning, production scheduling and supply chain optimization. An optimal production strategy derived from the deterministic model may become suboptimal or even totally infeasible owing to a slight “drift” around nominal values. When confronted with numerous parameters subjected to uncertainty and fluctuation, decision makers suffer a lot. An over-conservative decision always leads to an unnecessary deterioration of profit expectation, while an aggressive decision may result in severe constraint violations. Process optimization under uncertainty, which helps us to seek a satisfactory compromise between reliability and profitability, has received extensive attention for decades.

Three of the most popular optimization methodologies which incorporate uncertainty into the mathematical models are stochastic programming, robust optimization and chance constrained programming. Stochastic programming can be regarded as a scenario-based mathematical model, which allows recourse to hedge against infeasibility. Robust optimization is a set-induced approach that we do not need to know accurate probability distribution information about uncertainties. Its main feature is to ensure absolute feasibility over a specified uncertainty set. Chance constrained programming is a generation of robust optimization. It allows constraint violation to a certain degree.

Constraints are guaranteed to be satisfied with at least a specified probability at the optimum found.

Each of them has its own advantages, limitations, applicability and main mathematical formulation, which are listed in Table 1.

Great progress has been made in this field, however, its application in Process Systems Engineering still faces many barriers. The key difficulties in process optimization under uncertainty lie in the over-conservativeness of robust optimization, the computational intractability of stochastic programming and chance constrained programming.

The scope of this paper is to provide an overview of recent advances in process optimization under uncertainty over the past ten years. It is worth noting that optimization under uncertainty has a long history of development in Process Systems Engineering and more details and information can be found in some previous excellent reviews [1–5]. The content of this paper is organized as follows. Following these introductory remarks, Sections 2, 3, 4 present the characteristics, developments and applications of robust optimization, stochastic programming, chance constrained programming respectively. Section 5 describes the recent trends of this field and indicates the future directions. Finally, remaining challenges are discussed and conclusions are drawn in Section 6.

2. Robust Optimization

Robust optimization has emerged as a powerful technique to address uncertainty owing to its high reliability to immune against uncertainty and its computational tractability. It ensures feasibility against all possible realizations of uncertain parameters among a predefined uncertainty set. However, traditional robust optimization has always

* Corresponding author.

E-mail address: dcecbz@mail.tsinghua.edu.cn (B. Chen).

Table 1
Comparison among three optimization methods

Optimization method	Robust optimization	Stochastic programming	Chance constrained programming
Probability Distribution Information	Do not need to know	Need to know	Need to know
Computational Expenses	Computationally tractable, especially for linear programming	The most computationally expensive, especially when dealing with both exogenous and endogenous uncertainties. Two-stage and multistage stochastic MINLP problems are difficult to solve	Computationally challenging, joint probabilistic constraints will lead to an intractable NLP problem
Scope of Application Mathematical Formulation	Short-term scheduling where feasibility is a major concern [1] LP, MILP, QP	Long-term production planning and strategic design [1] LP, MILP, MINLP	Design and operation, production planning and optimal control LP, QP, NLP

been criticized by its over-conservativeness since the worst scenario is considered. In fact, the worst case is a rare occurrence, which may lead to an extremely poor economic objective. Therefore, many attempts have been made to reduce the conservativeness of robust optimization by avoiding the worst-case analysis. Here it should be noticed that flexibility analysis is also closely related to the solution of process optimization under uncertainty. Flexibility is concerned with the problem of ensuring feasible steady-state operation over a given range of uncertain operating conditions. Although flexibility analysis and robust optimization are derived from different fields of research, they share some knowledge in common. Both of them use the polyhedral sets to describe uncertainty and ensure feasibility by considering the worst case. Zhang *et al.* [6] made a comprehensive comparison between flexibility analysis and robust optimization and revealed the strong connection between them. In this paper, the discussion is focused on robust optimization rather than flexibility analysis.

Recent advances and methods to reduce the conservatism of robust optimization can be summarized as follows.

2.1. Utilize the probability information to adjust the uncertainty set

The size of the uncertainty set reflects the decision maker's risk tolerance, while the shape of it indicates detailed knowledge the modeler may have toward uncertainties such as correlations among uncertain parameters. The robustness and conservatism of robust optimization models are severely affected by the size of uncertainty set. Since a certain degree of constraint violation is allowed in reality, the uncertainty set is not necessary to cover the whole region of the uncertainty space. To avoid overly conservative solutions, the uncertainty set should be as small as possible on condition that the probability requirement on constraint satisfaction is satisfied. Li *et al.* [7] presented a systematic research on the set-induced robust counterpart optimization based on different types of uncertainty sets (*i.e.*, interval; ellipsoidal; polyhedral; the combination of interval, ellipsoidal, and polyhedral set) and compared their conservatism. Li *et al.* [8,9] and Guzman *et al.* [10–12] demonstrated that the performance of robust optimization can be greatly improved as the probabilistic bounds that determine probabilistic guarantees on constraint violation are tightened by utilizing the probability distribution information. In the real-world optimization problems, some uncertainties have greater effects on the objective than others, so Zhang *et al.* [13] proposed a flexible uncertainty set where uncertainties are considered at different weights to get tighter bounds on every dimension. Yuan *et al.* [14] also illustrated that incorporating more accurate correlation information through the covariance matrix into robust optimization framework will deform the uncertainty set and reduce its conservativeness.

2.2. Introduce a budget parameter to adjust the degree of conservatism

Bertsimas and Sim [15,16] introduced a budget parameter Γ_i (not necessarily integer, that takes the value ranging from 0 to the number

of uncertain parameters) to control the degree of conservatism. The corresponding formulation takes the following form.

$$\begin{aligned} \min & \quad cx \\ \text{s.t.} & \quad \sum_m a_{lm}x_m + \max_{\{S_i \cup \{t_i\} | S_i \subseteq M_i, |S_i| = \lfloor \Gamma_i \rfloor, t_i \in M_i \setminus S_i\}} \left\{ \sum_{m \in S_i} \hat{a}_{lm}|x_m| + (\Gamma_i - \lfloor \Gamma_i \rfloor)\hat{a}_{lt_i}|x_{t_i}| \right\} \leq p_l \end{aligned} \quad (1)$$

It is assumed that some of the parameters a_{lm} are subjected to uncertainty and fluctuate within the range of $[a_{lm} - \hat{a}_{lm}, a_{lm} + \hat{a}_{lm}]$, where a_{lm} represents the nominal value of the parameters and $\hat{a}_{lm}$ denotes the perturbation amplitude. S_i is the subset that contains $\lfloor \Gamma_i \rfloor$ uncertain parameters, and if Γ_i is not an integer, t_i is regarded as an index to describe an additional uncertain parameter. The intention of this formulation is to restrict up to $\lfloor \Gamma_i \rfloor$ parameters that are permitted to reach their worst case simultaneously. Hence, by adjusting Γ_i , we can control the level of conservatism of the solution. Besides, we have the flexibility to assign appropriate budget parameters to the different constraints according to a priori which parameters are allowed to take the worst-case value. Not only does it avoid intractable nonlinear formulation, but it also provides a trade-off between performance and conservatism.

2.3. Incorporate recourse decisions — adaptive robust optimization

The traditional robust optimization (also called static robust optimization, SRO), which requires all decisions to be made once and for all without the consideration of recourse, may lead to an overly conservative solution. Borrowing from the philosophy of multi-stage stochastic programming, adaptive robust optimization is proposed and has received much attention in recent years. Recourse decisions are introduced into static robust optimization so that adjustments can be implemented after the uncertainties are gradually revealed.

General formulation of two-stage adaptive robust optimization is presented in Eq. (2) [17], and it can be easily extended to multi-stage adaptive robust optimization framework.

$$\begin{aligned} \min_x & \quad c^T x + \max_{u \in U} \min_{y \in \Omega(x, u)} b^T y \\ \text{s.t.} & \quad Ax \geq d, \quad x \in S_x \\ & \quad \Omega(x, u) = \{y \in S_y : Wy \geq h - Tx - Mu\} \end{aligned} \quad (2)$$

Two-stage adaptive robust optimization is a sequential decision-making process, which comprises a design stage and an operational stage. x represents the first-stage decision variables that have to be made before the uncertainty u reveals, while y represents the second-stage (recourse) decisions which can be postponed in a “wait and see” mode based on the first-stage decisions and uncertainty realization. The set U represents the uncertainty set. The matrix W is generally called recourse matrix and T stands for technology matrix.

To retain computational tractability, the adjustable variables are generally restricted to be affine functions of the uncertain parameters, which is shown in Eq. (3). It means the recourse variables can tune themselves to the disclosure of uncertainties.

$$\begin{aligned} y &= y(u) \\ y &= q + Qu \end{aligned} \quad (3)$$

By using linear decision rules, adaptive robust optimization not only remains a tractable LP structure, but also accounts for recourse to a certain extent.

Recent advances of robust optimization applications in Process Systems Engineering are shown in Table 2.

3. Stochastic Programming

Stochastic programming is the scenario-based model with fixed recourse that optimizes the expectation of a number of discrete scenarios using either discrete probability distributions or the discretization of continuous probability distributions. There are two types of decision variables in stochastic programming. “Here and now” decisions are those that have to be fixed for the entire horizon before uncertain parameters are revealed. “Wait and see” decisions are recognized as corrective measures that can be made at different stages based on the realization of uncertainties in sequence [2].

Eq. (4) presents the general formulation of two-stage stochastic programming [34] and it can be extended to multi-stage stochastic programming.

$$\begin{aligned} \min z &= c^T x + E_{\xi} \left[\min q(\omega)^T y(\omega) \right] \\ \text{s.t. } Ax &= b, \\ T(\omega)x + Wy(\omega) &= h(\omega), \\ x \geq 0, y(\omega) &\geq 0 \end{aligned} \quad (4)$$

where x denotes the first-stage decision variables that need to be decided prior to the disclosure of uncertain variables ω , and y is regarded as

the second-stage (recourse) decisions which can be postponed until the actual value of ξ becomes known. E_{ξ} represents the mathematical expectation with respect to ξ . T is called the technology matrix while W constitutes the recourse matrix.

The deterministic equivalent program is presented as follows.

$$\begin{aligned} \min z &= c^T x + Q(x) \\ \text{s.t. } Ax &= b, \\ x &\geq 0 \end{aligned} \quad (5)$$

where $Q(x) = E_{\xi} Q(x, \xi(\omega))$ and $Q(x, \xi(\omega)) = \min_y \{ q(\omega)^T y | Wy = h(\omega) - T(\omega)x, y \geq 0 \}$ where $Q(x, \xi(\omega))$ is called recourse function, it exerts recourse on the objective function against constraint violations, and $Q(x)$ represents the expectation of recourse function.

Uncertainties in the stochastic programming can be classified into two types. Exogenous uncertainties generally correspond to market uncertainties, while endogenous uncertainties refer specifically to technical uncertainties. The realization of exogenous uncertainties is independent of process decisions. Since exogenous uncertainties are revealed automatically in each time period (e.g. oil prices), the structure of the exogenous scenario tree is known in advance. On the contrary, process decisions have important influence on the realization of endogenous uncertainties. Given that the timing of realizations relies heavily on the process decisions, the endogenous scenario-tree representation is more complicated and conditional, thus a superstructure is used to describe all potential outcomes. When copying the exogenous scenario tree for each possible combination of realizations of the endogenous uncertainties, we can get a ‘composite’ scenario tree [35]. Taking account of both the exogenous and endogenous uncertainties simultaneously is still an issue with great challenge.

Stochastic programming has played an important role in long-term production planning and strategic design, such as supply chain design and planning [36–42], process networks synthesis and planning [43–45], conceptual design and optimization of chemical processes

Table 2
Recent advances of robust optimization applications in Process Systems Engineering

Authors	Applications	Algorithms
Janak et al. [18], 2007	Short-term batch process scheduling	Uncertainty with known probability distribution
Li et al. [19], 2008	Short-term batch process scheduling	Introduce a budget parameter Γ_I
Ierapetritou et al. [20], 2007	Short-term batch process scheduling	Multi-objective robust optimization solved by normal boundary intersection technique
Verderame et al. [21], 2009	Operational planning of large-scale industrial batch plants	Robust optimization + Operational planning with production disaggregation model (PPDM)
Verderame et al. [22], 2010	Integration of operational planning and medium-term scheduling for large-scale industrial batch plants	Robust operational PPDM + Operational CVAR-PPDM
Verderame et al. [23], 2011	Operational planning of a multisite production and distribution network	Integration of planning and scheduling uncertainty by the rolling horizon framework with a novel feedback loop
Li et al. [24], 2012	Scheduling of crude oil operations	Robust optimization framework + Conditional value-at-risk (CVAR) theory
Wittmann-Hohlbein et al. [25], 2013	Short-term batch process scheduling	Unit-specific event-based continuous-time MINLP formulation + Robust optimization Solved by branch and bound global optimization algorithm
Tong et al. [26], 2014	Design and planning of the advanced hydrocarbon biofuel supply chain	Robust optimization + Multi-parametric programming
Zhang et al. [27], 2015	Scheduling of air separation with cryogenic energy storage	Unit cost objective + Introduce budget parameters. MILFP model solved by parametric algorithm and reformulation-linearization algorithm respectively
Zhang et al. [28], 2016	Scheduling of continuous industrial processes providing interruptible load	Introduce budget parameters
Shi et al. [29], 2016	Production scheduling of multiproduct batch manufacturing processes	Adjustable robust optimization (ARO) + Budgets of uncertainty
Gong et al. [30], 2016	Large-scale bioconversion process and product network design	Two-stage adaptive robust optimization + Budgets of uncertainty
Lappas et al. [31], 2016	Short-term batch process scheduling	Solved by Benders decomposition algorithm and C&CG algorithm respectively
Gong et al. [32], 2017	Large-scale process network design	Two-stage adaptive robust MINLP model + Budgets of uncertainty
Ning et al. [33], 2017	Batch process scheduling and process network planning	Solved by C&CG algorithm + Branch-and-Refine algorithm
		Multi-stage adjustable robust optimization + Decision-dependent uncertainty set
		Two-stage adaptive robust MIFP model + Budgets of uncertainty
		Solved by Parametric algorithm + C&CG algorithm
		Data-driven adaptive robust optimization (Machine-learning model + Adaptive robust optimization). Solved by a novel C&CG algorithm

[46–48], oil-field and gas-field development planning [49–51] and capacity planning for power-intensive continuous processes [52,53]. Stochastic programming is generally assumed to be a risk-neutral, however, modelers have different levels of risk aversion. Therefore, incorporating risk management into stochastic programming will help the decision makers find the most satisfactory solution based on their risk tolerance and preference [40,42].

A challenging aspect of applying stochastic programming to practical use is that the scale of the problem grows exponentially as the number of scenarios increases. Scenario generation is the basis and prerequisite of scenario-based stochastic programming, which can be done by sampling based method, statistic moment matching method and stochastic process simulation method. It is of vital importance to select an acceptable number of scenarios to well represent an approximate discretized representation of the uncertainty. Li and Floudas [54,55] proposed an optimal scenario reduction framework based on distance of uncertainty distribution and output performance, which can select a small number of representative scenarios capable of characterizing the probability density function. Calfa *et al.* [56] incorporated historical data to generate scenario tree by using moment matching along with distribution matching. The aforementioned methods provide a promising way to reduce the number of scenarios and computational complexity while ensuring a good-quality solution of stochastic programming.

To tackle the computational challenge of stochastic programming, integrating tailored solution strategies and decomposition algorithms is an efficient and widely used technique. Decomposition schemes take advantage of the special decomposable structure of stochastic programming (e.g., a block angular structure), largely reducing the scale of the problem and computational expenses. It deals with the problems in which only a relatively small set of complicating variables and constraints connects the scenarios while most of the variables and constraints have no direct interaction with each other. If these complicating variables and constraints are removed by dualizing them, then the full model will be decomposed into many smaller scenario subproblems that can be solved independently. L-shaped method and Lagrangean decomposition are two of the most common decomposition algorithms to cope with large-scale stochastic programming problems and there have been many novel variants of them to improve the efficiency of the algorithms in recent literatures [35,39,42,50,53,57].

Approximating the nonlinear term in the objective is the basic idea of the L-shaped method. It is desirable to avoid numerous function evaluations for the nonlinear objective term (the recourse function) because it involves a solution of all second-stage recourse linear programs. The first-stage variables are generally selected as complicating variables in this method. L-shaped method, which applies Benders decomposition to two-stage linear stochastic programming problems with continuous recourse, is guaranteed to converge to the optimal solution of the original problem in a finite number of iterations.

Lagrangean decomposition tends to be more appropriate for two-stage and multistage stochastic integer programming problems with mixed integer variables in all stages. The scenario tree is split into independent scenarios through the relaxation of non-anticipativity constraints (NACs). The barrier of this method is that it is not always guaranteed to find feasible solutions to the original problem because the solution may violate the relaxed constraints. To eliminate the duality gap between upper bound and lower bound, R.M. Apap *et al.* [35] combined Lagrangean decomposition and a heuristic which helps to find feasible solutions in the branch-and-bound algorithm.

The aforementioned methods are based on the discrete probability distributions of the uncertain parameters. As for the stochastic problems with continuous distributions of the uncertainties, discretizing the probability distributions through approximation or Monte Carlo sampling is the first step we should do.

4. Chance Constrained Optimization

Chance constrained optimization, first proposed by Charnes and Cooper [58,59], is a competitive tool to deal with parameter uncertainty. It allows constraint violation to some extent and does not exert recourse or penalty on the objective function, but the probability of complying with constraints must be satisfied with a specified confidence level. The probability distribution functions of the uncertainties need to be known so that the chance constrained optimization problem can be transformed into an equivalent deterministic formulation. Its main feature is that the decision makers can achieve a satisfactory trade-off between profitability and reliability by adjusting the confidence level.

A typical chance constrained optimization formulation is presented as follows. Eq. (6) represents the individual chance constrained optimization while Eq. (7) is the joint chance constrained optimization.

$$\begin{aligned} \min & f(\mathbf{x}) \\ \text{s.t.} & \Pr\{\mathbf{h}_i(\mathbf{x}, \mathbf{u}, \boldsymbol{\xi}) \leq 0\} \geq \alpha_i, \quad i = 1, \dots, m \end{aligned} \quad (6)$$

$$\begin{aligned} \min & f(\mathbf{x}) \\ \text{s.t.} & \Pr\{\mathbf{h}_i(\mathbf{x}, \mathbf{u}, \boldsymbol{\xi}) \leq 0, i = 1, \dots, m\} \geq \alpha \end{aligned} \quad (7)$$

where f represents the objective function, the vector $\mathbf{h}$ denotes the inequality constraints. $\mathbf{x}, \mathbf{u}, \boldsymbol{\xi}$ are the vectors of state, control and uncertain variables, respectively. $\Pr$ means probability and α is the confidence level specified by the modeler according to profit profiles vs. confidence level.

In the individual chance constraint, different confidence levels are assigned to different constraints based on their respective requirement of reliability. It is generally used when some constraints are more critical than others and should be paid more attention. Joint chance constraint requires all constraints to be satisfied simultaneously with the same predefined probability. When considering the safety and reliability of the process operation, joint chance constraint will be a better choice [60]. When relaxing them into equivalent deterministic problems, the individual chance constraint is reformulated to a linear programming (LP) problem that can be solved with commercial LP solvers. Unfortunately, the relaxation of the joint chance constraint results in a complicated NLP problem. Especially when the uncertain parameters are correlated with each other, it will be converted to a multivariate integration formulation that is generally solved through approximation approach. Undoubtedly, the computational intractability limits its widespread applications in practice.

Major applications of chance constrained programming in recent years are shown as follows. Mitra *et al.* [61] used the chance constrained programming to address the multi-objective mid-term supply chain planning problem. Barz *et al.* [62] proposed a two-layer chance-constrained optimization framework to get the optimal and reliable decisions, which is experimentally verified by a high-pressure distillation pilot plant. Yang *et al.* [63] employed chance constrained programming to achieve the trade-off between economic interest and reliability of decision results for a refinery supply chain operation under yield uncertainty.

To overcome the bottleneck in computational intractability, many efforts have been made in recent years. Klöppel *et al.* [64] used the sparse-grid integration method to solve nonlinear dynamic chance-constrained optimization, which largely decreased the computational expense. Yuan *et al.* [65] and Li *et al.* [66] investigated the computational tractable robust optimization approximation framework for solving the joint chance constrained problem. And data-driven chance constrained programming used kernel-based density estimation to approximate unknown true probability density functions, reformulating chance constraints into algebraic forms [67–69].

5. Future Directions

5.1. Stochastic robust optimization

There is a tendency that a few recent studies have begun to integrate different optimization methods under uncertainties to leverage their respective superiority and compensate for their drawbacks. When coping with multi-scale uncertainties, stochastic robust optimization can achieve a low expected cost while ensuring the system robustness by taking advantage of both of them [70]. Since decision makers have different levels of conservativeness toward strategic and operational uncertainties, stochastic programming and robust optimization are integrated into a nested framework where stochastic programming is used to handle strategic uncertainties while robust optimization guarantees the robustness of operations [71].

Stochastic robust optimization model is presented as follows.

$$\begin{aligned} \min \quad & C^{\text{str}}(x, y) + E_{\omega \in \Omega} [\bar{C}_{\omega}^{\text{opr}}(x, y)] \\ \text{s.t.} \quad & Ax + By \geq d \\ & x \in \{0, 1\}^n, y \in \mathbb{R}^m \\ & \begin{cases} \bar{C}_{\omega}^{\text{opr}}(x, y) = \max_{u_{\omega} \in U_{\omega}} \min_z C_{\omega}^{\text{opr}}(x, y, z, u_{\omega}) \\ \text{s.t.} \quad G_{\omega} z \geq h_{\omega} - E_{\omega} x - F_{\omega} y - K_{\omega} u_{\omega} \\ z_{\omega} \in \mathbb{R}^r \end{cases} \\ & \forall \omega \in \Omega \end{aligned} \quad (8)$$

where x, y represent the binary variables and the continuous variables for strategic decisions respectively, and z denotes the continuous operational decisions. ω means a certain realization of strategic uncertainties, and u_{ω} denotes a certain reveal of operational uncertainties. It should be noted that the operational uncertainty set U_{ω} is indexed by the realization of the strategic uncertainty ω . $C^{\text{str}}(x, y)$ stands for strategic stage cost and $\bar{C}_{\omega}^{\text{opr}}(x, y)$ represents the worst-case operational cost under the realization of strategic uncertainty ω .

Liu *et al.* [72] proposed a stochastic robust model to address two-stage power system optimization problems with uncertainty, and it has a stochastic min-max structure where the inner problem of the second stage reaches the worst cases and the outer problem of the second stage optimizes the expected economic performance under various worst-case scenarios. Yue and You [71] presented a nested stochastic robust optimization model for supply chain design and operation. Amaran *et al.* [73] also applied it to medium-term maintenance turnaround planning for integrated chemical sites to reflect different degrees of robustness toward production levels and maintenance manpower allocation.

5.2. Data-driven optimization under uncertainty

Big data has been an active area of research in recent years, which injects fresh vitality into traditional mathematical programming with consideration of uncertainties. The past ten years witnessed an explosion in the availability of data for Process Systems Engineering applications. Massive amounts of data are collected in chemical industries (e.g., historical demand profiles and historical yield variations). Data-driven optimization are motivated by dramatic advances of machine-learning algorithms and increasing availability of data.

Bertsimas *et al.* [74] utilized data to devise uncertainty sets for robust optimization through statistical hypothesis tests. Campbell *et al.* [75] presented a Bayesian nonparametric, data-driven, nonconvex uncertainty set constructed from a union of simple ellipsoidal sets for robust optimization, which can capture the nature of uncertainties in practical applications and achieve the desired degree of conservatism. Since the outliers caused by anomalous measurements significantly affect the area of the uncertainty set, robust optimization with conventional uncertainty sets could be overly conservative. Ning *et al.* [33] proposed a data-driven adaptive nested robust optimization framework which

has two layers of robustness. A well-designed data-driven uncertainty set is immune against anomalous measurements, and the adaptive robust optimization model is robust to parameter variations.

Moreover, historical data also leverage great power in stochastic programming. The data-driven approach helps to generate the scenario tree through statistical methods and reduce its size while retaining its statistical properties. Calfa *et al.* [56] proposed a data-driven multi-stage scenario tree generation approach via moment matching and distribution matching, which can assign reasonable probabilities and outcomes in scenario trees by taking advantage of historical data. Feng *et al.* [76] proposed a novel heuristic scenario reduction method based on wait-and-see clustering and selected one representative scenario from each cluster. And Xu *et al.* [77] presented a multi-stage scenario tree generation approach by utilizing simulation, clustering, nonlinear time series and moment matching methods simultaneously.

Another challenge for process optimization under uncertainty lies in estimating probability density/distribution of uncertain parameters. Jiang *et al.* [68] proposed a data-driven chance-constrained approach and constructed a density-based confidence set for the uncertainty based on a general φ -divergence measure. Several data-driven nonparametric estimation techniques, such as kernel density estimation (KDE), can be used to estimate unknown probability density/distribution functions of uncertainties so that we can reformulate individual and joint chance constraints into algebraic forms [67].

In short, machine learning sits at the intersection of computer science, mathematical statistics, inference and decision-making under uncertainty [78]. Data-driven optimization, which integrates mathematical programming methods with machine learning algorithms, has become an emerging and competitive tool to cope with optimization problems under uncertainty utilizing massive historical data.

6. Conclusions and Outlook

An overview presented in the preceding sections demonstrates that dramatic progress has been made in the field of process optimization under uncertainties over the past ten years. Different optimization methods focus on specific application areas. Robust optimization is more suitable for short-term scheduling problems. As for long-term production planning and strategic design decisions, such as the development of oilfields, the capacity expansion of process networks and supply chain design and planning, stochastic programming leverages its superiority because it allows recourse decisions to be made at later time points after the gradual realization of the uncertainties. The above detailed discussion indicates that the major challenge of robust optimization is how to reduce the degree of conservatism and to seek a trade-off between robustness and profitability. Several methods, such as adjusting the size and the shape of the uncertainty set by utilizing the probability information, introducing a budget parameter and incorporating recourse decisions, provide greater flexibility for robust optimization and make the solutions of robust optimization less conservative. The key limitation of stochastic programming and chance constrained programming lies in their computational intractability. To solve large-scale stochastic programming problems efficiently, scenario reduction strategies and decomposition algorithms are widely used. Parallel computation and Multi-Agent System will also make a contribution to the large-scale mathematical models in the future. As for chance constrained programming, many scholars have studied the application of the robust counterpart optimization to approximate the joint chance constrained problem in recent years and indicate that robust optimization can provide safe and tractable analytical approximation for solving joint chance constrained problem.

In the era of big data and with the rapid development of advanced machine-learning algorithms, process optimization under uncertainty benefits from valuable information embedded within massive historical data. It is inspiring to see that the machine-learning model is seamlessly integrated with optimization approaches. Since the well-designed

data-driven uncertainty set tends to be tighter than classical ones, it can greatly reduce the conservatism of robust optimization solutions. Also, the obtained data-driven probabilities and outcomes can be used in the scenario tree generation for stochastic programming. Undoubtedly, data-driven machine-learning approaches can help us to capture the characterization of uncertain data and advance the development of process optimization under uncertainty. Despite the recent progress in the data-driven optimization, there are still many problems and challenges left. Not all the uncertainties can be predicted and described based on the historical data (e.g., the oil price suffers dramatic fluctuations due to some sudden political and economic factors, so the historical data do not have the reference value). In addition, the data obtained directly from the industries are not “clean” enough and have to be preprocessed before we use it.

Current trend in this field also demonstrates that different optimization methods have begun to integrate with each other. Since a “one-size-fits-all” approach is not sufficient to reflect the modeler's different levels of conservatism toward multi-scale uncertainties, stochastic robust optimization is proposed. Furthermore, in the presence of uncertainties, decision makers not only pursue the maximization of profits, but they also care about the minimization of risks. Hence, incorporating effective risk management measures into optimization models is another promising direction in the future.

References

- [1] I.E. Grossmann, R.M. Apap, B.A. Calfa, P. García-Herreros, Q. Zhang, Recent advances in mathematical programming techniques for the optimization of process systems under uncertainty, *Comput. Chem. Eng.* 91 (2016) 3–14.
- [2] N.V. Sahinidis, Optimization under uncertainty: state-of-the-art and opportunities, *Comput. Chem. Eng.* 28 (2004) 971–983.
- [3] G.I. Schuëller, H.A. Jensen, Computational methods in optimization considering uncertainties – An overview, *Comput. Methods Appl. Mech. Eng.* 198 (2008) 2–13.
- [4] Z. Li, M. Ierapetritou, Process scheduling under uncertainty: Review and challenges, *Comput. Chem. Eng.* 32 (2008) 715–727.
- [5] P.M. Verderame, J.A. Elia, J. Li, C.A. Floudas, Planning and scheduling under uncertainty: A review across multiple sectors, *Ind. Eng. Chem. Res.* 49 (2010) 3993–4017.
- [6] Q. Zhang, I.E. Grossmann, R.M. Lima, On the relation between flexibility analysis and robust optimization for linear systems, *AIChE J.* 62 (2016) 3109–3123.
- [7] Z. Li, R. Ding, C.A. Floudas, A comparative theoretical and computational study on robust counterpart optimization: I. Robust linear optimization and robust mixed integer linear optimization, *Ind. Eng. Chem. Res.* 50 (2011) 10567–10603.
- [8] Z. Li, C.A. Floudas, A comparative theoretical and computational study on robust counterpart optimization: II. Probabilistic guarantees on constraint satisfaction, *Ind. Eng. Chem. Res.* 51 (2012) 6769–6788.
- [9] Z. Li, C.A. Floudas, A comparative theoretical and computational study on robust counterpart optimization: III. Improving the quality of robust solutions, *Ind. Eng. Chem. Res.* 53 (2014) 13112–13124.
- [10] Y.A. Guzman, L.R. Matthews, C.A. Floudas, New a priori and a posteriori probabilistic bounds for robust counterpart optimization: I. Unknown probability distributions, *Comput. Chem. Eng.* 84 (2016) 568–598.
- [11] Y.A. Guzman, L.R. Matthews, C.A. Floudas, New a priori and a posteriori probabilistic bounds for robust counterpart optimization: II. A priori bounds for known symmetric and asymmetric probability distributions, *Comput. Chem. Eng.* 101 (2017) 279–311.
- [12] Y.A. Guzman, L.R. Matthews, C.A. Floudas, New a priori and a posteriori probabilistic bounds for robust counterpart optimization: III. Exact and near-exact a posteriori expressions for known probability distributions, *Comput. Chem. Eng.* 103 (2017) 116–143.
- [13] Y. Zhang, Y. Feng, G. Rong, New robust optimization approach induced by flexible uncertainty set: Optimization under continuous uncertainty, *Ind. Eng. Chem. Res.* 56 (2017) 270–287.
- [14] Y. Yuan, Z. Li, B. Huang, Robust optimization under correlated uncertainty: Formulations and computational study, *Comput. Chem. Eng.* 85 (2016) 58–71.
- [15] D. Bertsimas, M. Sim, The Price of Robustness, *Oper. Res.* 52 (2004) 35–53.
- [16] D. Bertsimas, M. Sim, Robust discrete optimization and network flows, *Math. Program.* 98 (2003) 49–71.
- [17] B. Zeng, L. Zhao, Solving two-stage robust optimization problems using a column-and-constraint generation method, *Oper. Res. Lett.* 41 (2013) 457–461.
- [18] S.L. Janak, X. Lin, C.A. Floudas, A new robust optimization approach for scheduling under uncertainty, *Comput. Chem. Eng.* 31 (2007) 171–195.
- [19] Z. Li, M.G. Ierapetritou, Robust optimization for process scheduling under uncertainty, *Ind. Eng. Chem. Res.* 47 (2008) 4148–4157.
- [20] M.G. Ierapetritou, Z. Jia, Short-term scheduling of chemical process including uncertainty, *Control. Eng. Pract.* 15 (2007) 1207–1221.
- [21] P.M. Verderame, C.A. Floudas, Operational planning of large-scale industrial batch plants under demand due date and amount uncertainty. I. Robust optimization framework, *Ind. Eng. Chem. Res.* 48 (2009) 7214–7231.
- [22] P.M. Verderame, C.A. Floudas, Integration of operational planning and medium-term scheduling for large-scale industrial batch plants under demand and processing time uncertainty, *Ind. Eng. Chem. Res.* 49 (2010) 4948–4965.
- [23] P.M. Verderame, C.A. Floudas, Multisite planning under demand and transportation time uncertainty: Robust optimization and conditional value-at-risk frameworks, *Ind. Eng. Chem. Res.* 50 (2011) 4959–4982.
- [24] J. Li, R. Misener, C.A. Floudas, Scheduling of crude oil operations under demand uncertainty: A robust optimization framework coupled with global optimization, *AIChE J.* 58 (2012) 2373–2396.
- [25] M. Wittmann-Hohlbein, E.N. Pistikopoulos, Proactive scheduling of batch processes by a combined robust optimization and multiparametric programming approach, *AIChE J.* 59 (2013) 4184–4211.
- [26] K. Tong, F. You, G. Rong, Robust design and operations of hydrocarbon biofuel supply chain integrating with existing petroleum refineries considering unit cost objective, *Comput. Chem. Eng.* 68 (2014) 128–139.
- [27] Q. Zhang, I.E. Grossmann, C.F. Heuberger, A. Sundaramoorthy, J.M. Pinto, Air separation with cryogenic energy storage: Optimal scheduling considering electric energy and reserve markets, *AIChE J.* 61 (2015) 1547–1558.
- [28] Q. Zhang, M.F. Morari, I.E. Grossmann, A. Sundaramoorthy, J.M. Pinto, An adjustable robust optimization approach to scheduling of continuous industrial processes providing interruptible load, *Comput. Chem. Eng.* 86 (2016) 106–119.
- [29] H. Shi, F. You, A computational framework and solution algorithms for two-stage adaptive robust scheduling of batch manufacturing processes under uncertainty, *AIChE J.* 62 (2016) 687–703.
- [30] J. Gong, D.J. Garcia, F. You, Unraveling optimal biomass processing routes from bio-conversion product and process networks under uncertainty: An adaptive robust optimization approach, *ACS Sustain. Chem. Eng.* 4 (2016) 3160–3173.
- [31] N.H. Lappas, C.E. Gounaris, Multi-stage adjustable robust optimization for process scheduling under uncertainty, *AIChE J.* 62 (2016) 1646–1667.
- [32] J. Gong, F. You, Optimal processing network design under uncertainty for producing fuels and value-added bioproducts from microalgae: Two-stage adaptive robust mixed integer fractional programming model and computationally efficient solution algorithm, *AIChE J.* 63 (2017) 582–600.
- [33] C. Ning, F. You, Data-driven adaptive nested robust optimization: General modeling framework and efficient computational algorithm for decision making under uncertainty, *AIChE J.* 63 (2017) 3730–3817.
- [34] J.R. Birge, F. Louveaux, Introduction to stochastic programming, Springer Science & Business Media, 2011.
- [35] R.M. Apap, I.E. Grossmann, Models and computational strategies for multistage stochastic programming under endogenous and exogenous uncertainties, *Comput. Chem. Eng.* 103 (2017) 233–274.
- [36] J. Gao, F. You, Deciphering and handling uncertainty in shale gas supply chain design and optimization: Novel modeling framework and computationally efficient solution algorithm, *AIChE J.* 61 (2015) 3739–3755.
- [37] B.H. Gebreslassie, Y. Yao, F. You, Design under uncertainty of hydrocarbon biorefinery supply chains: Multiobjective stochastic programming models, decomposition algorithm, and a comparison between CVaR and downside risk, *AIChE J.* 58 (2012) 2155–2179.
- [38] K. Tong, Y. Feng, G. Rong, Planning under demand and yield uncertainties in an oil supply chain, *Ind. Eng. Chem. Res.* 51 (2012) 814–834.
- [39] F. You, I.E. Grossmann, Multicut Benders decomposition algorithm for process supply chain planning under uncertainty, *Ann. Oper. Res.* 210 (2011) 191–211.
- [40] F. You, J.M. Wassick, I.E. Grossmann, Risk management for a global supply chain planning under uncertainty: Models and algorithms, *AIChE J.* 55 (2009) 931–946.
- [41] S. Baptista, M. Isabel Gomes, A.P. Barbosa-Povoa, A two-stage stochastic model for the design and planning of a multi-product closed loop supply chain, in: B. Ian David Lockhart, F. Michael (Eds.), Computer Aided Chemical Engineering, Elsevier, Amsterdam 2012, pp. 412–416.
- [42] F. Oliveira, V. Gupta, S. Hamacher, I.E. Grossmann, A Lagrangean decomposition approach for oil supply chain investment planning under uncertainty with risk considerations, *Comput. Chem. Eng.* 50 (2013) 184–195.
- [43] V. Gupta, I.E. Grossmann, Solution strategies for multistage stochastic programming with endogenous uncertainties, *Comput. Chem. Eng.* 35 (2011) 2235–2247.
- [44] B. Tarhan, I.E. Grossmann, A multistage stochastic programming approach with strategies for uncertainty reduction in the synthesis of process networks with uncertain yields, *Comput. Chem. Eng.* 32 (2008) 766–788.
- [45] V. Goel, I.E. Grossmann, A class of stochastic programs with decision dependent uncertainty, *Math. Program.* 108 (2006) 355–394.
- [46] J. Steimel, S. Engell, Conceptual design and optimization of chemical processes under uncertainty by two-stage programming, *Comput. Chem. Eng.* 81 (2015) 200–217.
- [47] J. Steimel, M. Harrmann, G. Schembecker, S. Engell, Model-based conceptual design and optimization tool support for the early stage development of chemical processes under uncertainty, *Comput. Chem. Eng.* 59 (2013) 63–73.
- [48] J. Steimel, S. Engell, Optimization-based support for process design under uncertainty: A case study, *AIChE J.* 62 (2016) 3404–3419 Amsterdam.
- [49] V. Gupta, I.E. Grossmann, Multistage stochastic programming approach for offshore oilfield infrastructure planning under production sharing agreements and endogenous uncertainties, *J. Pet. Sci. Eng.* 124 (2014) 180–197.
- [50] V. Gupta, I.E. Grossmann, A new decomposition algorithm for multistage stochastic programs with endogenous uncertainties, *Comput. Chem. Eng.* 62 (2014) 62–79.
- [51] B. Tarhan, I.E. Grossmann, V. Goel, Stochastic programming approach for the planning of offshore oil or gas field infrastructure under decision-dependent uncertainty, *Ind. Eng. Chem. Res.* 48 (2009) 3078–3097.

- [52] S. Mitra, J.M. Pinto, I.E. Grossmann, Optimal multi-scale capacity planning for power-intensive continuous processes under time-sensitive electricity prices and demand uncertainty. Part I: Modeling, *Comput. Chem. Eng.* 65 (2014) 89–101.
- [53] S. Mitra, J.M. Pinto, I.E. Grossmann, Optimal multi-scale capacity planning for power-intensive continuous processes under time-sensitive electricity prices and demand uncertainty. Part II: Enhanced hybrid bi-level decomposition, *Comput. Chem. Eng.* 65 (2014) 102–111.
- [54] Z. Li, C.A. Floudas, Optimal scenario reduction framework based on distance of uncertainty distribution and output performance: I. Single reduction via mixed integer linear optimization, *Comput. Chem. Eng.* 70 (2014) 50–66.
- [55] Z. Li, C.A. Floudas, Optimal scenario reduction framework based on distance of uncertainty distribution and output performance: II. Sequential reduction, *Comput. Chem. Eng.* 84 (2016) 599–610.
- [56] B.A. Calfa, A. Agarwal, I.E. Grossmann, J.M. Wassick, Data-driven multi-stage scenario tree generation via statistical property and distribution matching, *Comput. Chem. Eng.* 68 (2014) 7–23.
- [57] F. Oliveira, I.E. Grossmann, S. Hamacher, Accelerating Benders stochastic decomposition for the optimization under uncertainty of the petroleum product supply chain, *Comput. Oper. Res.* 49 (2014) 47–58.
- [58] A. Charnes, W.W. Cooper, Chance-constrained programming, *Manag. Sci.* 6 (1959) 73–79.
- [59] Chance-constrained programming, in: S.I. Gass, M.C. Fu (Eds.), *Encyclopedia of Operations Research and Management Science*, Springer US, Boston, MA 2013, p. 160.
- [60] P. Li, H. Arellano-Garcia, G. Wozny, Chance constrained programming approach to process optimization under uncertainty, *Comput. Chem. Eng.* 32 (2008) 25–45.
- [61] K. Mitra, R.D. Gudi, S.C. Patwardhan, G. Sardar, Midterm supply chain planning under uncertainty: A multiobjective chance constrained programming framework, *Ind. Eng. Chem. Res.* 47 (2008) 5501–5511.
- [62] T. Barz, G.N. Wozny, H. Arellano-Garcia, Robust implementation of optimal decisions using a two-layer chance-constrained approach, *Ind. Eng. Chem. Res.* 50 (2011) 5050–5063.
- [63] J. Yang, H. Gu, G. Rong, Supply chain optimization for refinery with considerations of operation mode changeover and yield fluctuations, *Ind. Eng. Chem. Res.* 49 (2010) 276–287.
- [64] M. Klöppel, A. Geletu, A. Hoffmann, P. Li, Using sparse-grid methods to improve computation efficiency in solving dynamic nonlinear chance-constrained optimization problems, *Ind. Eng. Chem. Res.* 50 (2011) 5693–5704.
- [65] Y. Yuan, Z. Li, B. Huang, Robust optimization approximation for joint chance constrained optimization problem, *J. Glob. Optim.* 67 (2016) 805–827.
- [66] Z. Li, Z. Li, Optimal robust optimization approximation for chance constrained optimization problem, *Comput. Chem. Eng.* 74 (2015) 89–99.
- [67] B.A. Calfa, I.E. Grossmann, A. Agarwal, S.J. Bury, J.M. Wassick, Data-driven individual and joint chance-constrained optimization via kernel smoothing, *Comput. Chem. Eng.* 78 (2015) 51–69.
- [68] R. Jiang, Y. Guan, Data-driven chance constrained stochastic program, *Math. Program.* 158 (2015) 291–327.
- [69] Y. Zhang, Y. Feng, G. Rong, Data-driven chance constrained and robust optimization under matrix uncertainty, *Ind. Eng. Chem. Res.* 55 (2016) 6145–6160.
- [70] C. Zhao, Y. Guan, Unified stochastic and robust unit commitment, *IEEE Trans. Power Syst.* 28 (2013) 3353–3361.
- [71] D. Yue, F. You, Optimal supply chain design and operations under multi-scale uncertainties: Nested stochastic robust optimization modeling framework and solution algorithm, *AIChE J.* 62 (2016) 3041–3055.
- [72] C. Liu, C. Lee, H. Chen, S. Mehrotra, Stochastic robust mathematical programming model for power system optimization, *IEEE Trans. Power Syst.* 31 (2016) 821–822.
- [73] Y. Ye, J. Li, Z. Li, Q. Tang, X. Xiao, C.A. Floudas, Robust optimization and stochastic programming approaches for medium-term production scheduling of a large-scale steelmaking continuous casting process under demand uncertainty, *Comput. Chem. Eng.* 66 (2014) 165–185.
- [74] D. Bertsimas, V. Gupta, N. Kallus, Data-driven robust optimization, *Math. Program.* (2017) 1–58.
- [75] T. Campbell, J.P. How, Bayesian nonparametric set construction for robust optimization, *P Amer Contr Conf* 2015, pp. 4216–4221.
- [76] Y. Feng, S.M. Ryan, Scenario construction and reduction applied to stochastic power generation expansion planning, *Comput. Oper. Res.* 40 (2013) 9–23.
- [77] D. Xu, Z. Chen, L. Yang, Scenario tree generation approaches using K-means and LP moment matching methods, *J. Comput. Appl. Math.* 236 (2012) 4561–4579.
- [78] M.I. Jordan, T.M. Mitchell, Machine learning: Trends, perspectives, and prospects, *Science* 349 (2015) 255–260.