Two-Stage Stochastic Mixed-Integer Linear Programming: The Conditional Scenario Approach

C. Beltran-Royo*

30/08/2016

Abstract

In this paper we consider the two-stage stochastic mixed-integer linear programming problem with recourse, which we call the RP problem. A common way to approximate the RP problem, which is usually formulated in terms of *scenarios*, is to formulate the so-called Expected Value (EV) problem, which only considers the expectation of the random parameters of the RP problem. In this paper we introduce the *Conditional Scenario* (CS) problem which represents a midpoint between the RP and the EV problems regarding computational tractability and ability to deal with uncertainty. In the theoretical section we have analyzed some useful bounds related to the RP, EV and CS problems. In the numerical example here presented, the CS problem has outperformed both the EV problem in terms of solution quality, and the RP problem with the same number of scenarios as in the CS problem, in terms of solution time.

Keywords: Stochastic mixed-integer linear programming, conditional expectation, scenario, conditional scenario.

1 Introduction

The relevance for managerial purposes, properties, solution methods and applications of two-stage stochastic mixed-integer linear programming can be found in surveys such as [35] and in books such as [10]. Among the applications one finds: the facility location problem with Bernoulli demands [2], scheduling of a multiproduct batch plant with uncertain demand [15], strategic production planning under uncertainty [4], thermal power system expansion [3] and employee scheduling in retail outlets with uncertain demand [30].

To address stochastic optimization problems different approaches can be used: robust optimization [7], chance constraint optimization [13, 32], sampling based methods [21] and scenario based optimization [10, 25], among others. In this paper we focus on the last approach. The RP problem here considered corresponds to a *two-stage stochastic mixed-integer linear* optimization problem with *recourse* and *risk neutral*. Several variants of this RP problem have been proposed in the literature: multi-stage versions [8, 28], risk aversion versions [37], non-linear versions [1], etc.

The two-stage stochastic mixed-integer linear programming problem formulated in terms of a continuous random vector which accounts for all the uncertain parameters of the problem is, in general, numerically intractable. To address this difficulty one can approximate the original random vector by a random vector with a finite number of realizations (the *scenario tree*). Thus, in the first step of scenario

*cesar.beltran@urjc.es, Computer Science and Statistics, Rey Juan Carlos University, Madrid, Spain.

based optimization one calculates a representative and tractable scenario tree. See [14, 20, 22, 27, 29], among others. In the second step, one formulates and solves the so-called deterministic equivalent, which we call the RP problem. A general purpose optimization software, such as CPLEX [23], can be used to solve RP instances of moderate size. However, given the high complexity of the RP problem, in many cases one needs to use specialized approaches: branch-and-fix coordination [5], Benders decomposition [36], Lagrangian relaxation [41], decomposition with branch-and-cut [38], etc. In this paper we will assume that the RP problem has been formulated in terms of a random vector with discrete support consisting of S scenarios which corresponds to a structured Mixed-Integer Linear Programming (MILP) problem.

As an alternative to the RP problem one can use the Expected Value (EV) problem [10], where the random parameters of the RP problem are approximated by their expectations. That is, the EV problem approximates the RP problem by ignoring the parameter uncertainty. It is normally recognized that the EV problem requires a low computational effort but, in general, it is not an adequate approximation to the stochastic programming problems [10, 25]. Therefore, it seems that scenario based optimization (the RP problem) and deterministic optimization (the EV problem) represent two extreme choices regarding computational effort and ability to deal with uncertainty. A natural question can be raised: would it be possible to propose a midpoint between these two approaches? That is, would it be possible to suggest an approach with a moderate computational effort and with a reasonable ability to deal with uncertainty? In this paper we give a positive answer to this question by introducing the Conditional Scenario (CS) problem. As we will see, in the CS problem the random parameters of the RP problem are approximated by their *conditional expectations*. Thus, the CS problem improves the ability of the EV problem to deal with uncertainty by considering conditional expectations instead of expectations. On the other hand, the CS problem reduces the computational burden of the RP problem by considering *conditional scenarios* instead of *scenarios* (the conditional scenario concept will be introduced in Section 3). Of course, there is no such thing as a free lunch and the optimal CS solution is, in general, suboptimal for the RP problem but hopefully better than the optimal EV solution.

The CS approach can be seen as an aggregation method. The idea of aggregate constraints and/or variables for solving large-scale problems is well known. A survey presented in [34] demonstrates aggregation and disaggregation techniques for optimization problems. Different aggregation methods have been analyzed for several optimization problems including linear programming [43, 44], mixed-integer linear programming [18] and non-linear programming [16]. In the context of aggregation methods applied to stochastic programming, [9] obtains bounds on the optimal value of a multistage stochastic linear program with random right-hand sides. These results are extended in [42] to the general case which allows for randomness not only in the right-hand sides but also in the constraint matrices and costs. In [40] the aggregation approach is used not only to derive optimality bounds but also to solve the original two-stage stochastic linear program (with fixed recourse and fixed costs). In the CS approach, seen as an aggregation method, the aggregation weights are given by conditional probability functions (see Remark 3 in the Appendix). The main theoretical contribution of the CS approach is to provide stronger bounds for the RP problem, compared to the EV ones (see Section 5). Furthermore, in the numerical example in Section 6 the CS approach clearly outperforms the EV approach and the RP approach (with the same number of scenarios as in the CS approach) in order to efficiently compute good quality solutions.

In summary, the objectives of this paper are to introduce the CS approach, to study some of its theoretical properties and to show by example how one can efficiently obtain good solutions in two-stage stochastic mixed-integer linear programming. With these objectives in mind, in Section 2 we formulate the RP and EV problems. In Section 3 we introduce the conditional scenario concept to be used to formulate the CS problem in Section 4. In Section 5 some useful bounds related to the RP, CS and EV problems are stated. In Section 6 a numerical example is used to compare the EV, CS and RP problems. Section 7 concludes the paper and outlines future research. Finally, in the Appendix we prove the theoretical results of Section 5.

2 The recourse problem

2.1 Notation

Indexes:

t	Decision stages	$t \in \mathcal{T} = \{1, 2\}$
j	Integer components of decision vectors at stage t	$j \in \mathcal{J}_t = \{1, \dots, J_t\},\ t \in \mathcal{T}$
e	Realizations of random variables	$e \in \mathcal{E} = \{1, \dots, E\}$
r	Components of random vectors	$r \in \mathcal{R} = \{1, \dots, R\}$
s	Scenarios	$s \in \mathcal{S} = \{1, \dots, S\}$
re	Index pair for conditional scenarios where $\mathcal{E}_r = \{1, \dots, E_r\}$	$re \in \mathcal{RE}_r = \mathcal{R} \times \mathcal{E}_r$

Random vectors:

$\tilde{\boldsymbol{\xi}}$	Random vector with finite support which accounts for all the random parameters of the RP problem	
$\tilde{\xi}^s$	Scenario or realization of the random vector $\tilde{\boldsymbol{\xi}}$	$s \in \mathcal{S}$
$\tilde{p}^s$	Probability of $\tilde{\xi}^s$, that is, $\tilde{p}^s = P\big(\tilde{\boldsymbol{\xi}} = \tilde{\xi}^s\big)$	$s \in \mathcal{S}$
$\tilde{\boldsymbol{\xi}}_r$	Component r of the random vector $\tilde{\boldsymbol{\xi}}$	$r \in \mathcal{R}$
$\tilde{\xi}_{re}$	Realization of the random variable $\tilde{\boldsymbol{\xi}}_r$	$r \in \mathcal{R}, e \in \mathcal{E}_r$
$\hat{\boldsymbol{\xi}}^r$	Random vector which approximates $\tilde{\boldsymbol{\xi}}$ by conditional expectation, that is, $\hat{\boldsymbol{\xi}}^r = \mathbb{E}[\,\tilde{\boldsymbol{\xi}} \mid \tilde{\boldsymbol{\xi}}_r\,]$	$r \in \mathcal{R}$
$\hat{\xi}^{re}$	Conditional scenario or realization of the random vector $\hat{\boldsymbol{\xi}}^r$ $\hat{\xi}^{re} = \mathbb{E}[\,\tilde{\boldsymbol{\xi}} \mid \tilde{\xi}_{re}\,]$	$r \in \mathcal{R}, e \in \mathcal{E}_r$
$\hat{p}^{re}$	Probability of $\hat{\xi}^{re}$, that is, $\hat{p}^{re} = P\big(\hat{\boldsymbol{\xi}}^r = \hat{\xi}^{re}\big)$	$r \in \mathcal{R}, e \in \mathcal{E}_r$
$\bar{\xi}$	Expectation of $\tilde{\boldsymbol{\xi}}$, that is, $\bar{\xi} = \mathbb{E}[\,\tilde{\boldsymbol{\xi}}\,]$	

2.2 Problem formulation

The RP problem considered in this paper can be written as follows (stochastic version):

$$\min_{\tilde{x}} \quad z_{RP} = c_1^\intercal \tilde{x}_1 + \mathbb{E}[\,\tilde{\mathbf{c}}_2^\intercal \tilde{x}_2(\tilde{\boldsymbol{\xi}})\,] \tag{1}$$

$$\text{s.t.} \quad A_1 \tilde{x}_1 = b_1 \tag{2}$$

$$\tilde{\mathbf{A}}_2 \tilde{x}_1 + \tilde{\mathbf{B}}_2 \tilde{x}_2(\tilde{\boldsymbol{\xi}}) = \tilde{\mathbf{b}}_2 \qquad \text{w.p.1} \tag{3}$$

$$\tilde{x}_1 \geq 0 \tag{4}$$

$$\tilde{x}_2(\tilde{\boldsymbol{\xi}}) \geq 0 \qquad \text{w.p.1} \tag{5}$$

$$\tilde{x}_{1j} \text{ integer} \qquad j \in \mathcal{J}_1 \tag{6}$$

$$\tilde{x}_{2j}(\tilde{\boldsymbol{\xi}}) \text{ integer} \qquad \text{w.p.1},\ j \in \mathcal{J}_2, \tag{7}$$

where 'w.p.1' stands for 'with probability one' and $\mathcal{J}_t$ is the index set for the integer variables at stage t for all $t \in \mathcal{T}$. It is assumed that some or all the entries of $\tilde{\mathbf{c}}_2, \tilde{\mathbf{A}}_2, \tilde{\mathbf{B}}_2$ and $\tilde{\mathbf{b}}_2$ are random variables with finite support. In this context, it is common to define a random vector with all the random parameters of the problem, say $\tilde{\boldsymbol{\xi}} = \text{vec}\big(\tilde{\mathbf{c}}_2, \tilde{\mathbf{A}}_2, \tilde{\mathbf{B}}_2, \tilde{\mathbf{b}}_2\big)$, where 'vec' is the operator that stacks vectors and matrix columns into a single vector. The support of $\tilde{\boldsymbol{\xi}}$ is the set of scenarios or realizations $\tilde{\xi}^s$ and the corresponding probabilities are $\tilde{p}^s = P\big(\tilde{\boldsymbol{\xi}} = \tilde{\xi}^s\big)$ for all $s \in \mathcal{S}$. The random parameters are written in boldface to distinguish them from the deterministic ones, namely, c_1, A_1 and b_1.

In the RP problem the first stage decision vector $\tilde{x}_1$ does not depend on the uncertain parameters in contrast with the second stage decision vector x_2, which depends on $\tilde{\boldsymbol{\xi}}$ and for this reason one usually writes $\tilde{x}_2(\tilde{\boldsymbol{\xi}})$ and calls it a *policy* [39]. The objective of the RP problem here formulated is to minimize the first stage cost plus the expected second stage cost.

For numerical purposes the RP problem is written in the so-called deterministic equivalent version:

$$\begin{aligned}
\min_{\tilde{x}} \quad & z_{RP} = c_1^\intercal \tilde{x}_1 + \sum_{s\in\mathcal{S}} \tilde{p}^s \, \tilde{c}_2^{s\intercal} \tilde{x}_2^s && && (8)\\
\text{s.t.} \quad & A_1 \tilde{x}_1 = b_1 && && (9)\\
& \tilde{A}_2^s \tilde{x}_1 + \tilde{B}_2^s \tilde{x}_2^s = \tilde{b}_2^s && s \in \mathcal{S} && (10)\\
& \tilde{x}_1 \geq 0 && && (11)\\
& \tilde{x}_2^s \geq 0 && s \in \mathcal{S} && (12)\\
& \tilde{x}_{1j} \text{ integer} && j \in \mathcal{J}_1 && (13)\\
& \tilde{x}_{2j}^s \text{ integer} && s \in \mathcal{S},\ j \in \mathcal{J}_2, && (14)
\end{aligned}$$

such that $\tilde{\xi}^s = \operatorname{vec}\big(\, \tilde{c}_2^s, \tilde{A}_2^s, \tilde{B}_2^s, \tilde{b}_2^s \,\big)$ for all $s \in \mathcal{S}$.

The RP problem can be approximated by the so-called expected value problem. More precisely, the EV problem corresponds to approximate $\tilde{\boldsymbol{\xi}}$ by $\bar{\xi} = \mathbb{E}[\, \tilde{\boldsymbol{\xi}} \,] = \operatorname{vec}\big(\, \bar{c}_2, \bar{A}_2, \bar{B}_2, \bar{b}_2 \,\big)$ and can be stated as

$$\begin{aligned}
\min_{\bar{x}} \quad & z_{EV} = c_1^\intercal \bar{x}_1 + \bar{c}_2^\intercal \bar{x}_2 && && (15)\\
\text{s.t.} \quad & A_1 \bar{x}_1 = b_1 && && (16)\\
& \bar{A}_2 \bar{x}_1 + \bar{B}_2 \bar{x}_2 = \bar{b}_2 && && (17)\\
& \bar{x}_1 \geq 0 && && (18)\\
& \bar{x}_2 \geq 0 && && (19)\\
& \bar{x}_{tj} \text{ integer} && t \in \mathcal{T}, j \in \mathcal{J}_t. && (20)
\end{aligned}$$

Notice that notation $\bar{x}$ and $\tilde{x}$ is used to distinguish the decision vectors of the EV and RP problems. To measure the quality of an optimal EV solution $\bar{x}^*$, one usually considers the so-called 'Expected result of using the EV solution' (E-EV) [10]. The E-EV can be computed by solving the RP problem with the additional constraint $\tilde{x}_1 = \bar{x}_1^*$ (here we are assuming that $\bar{x}_1^*$ is feasible for the RP problem). Further details can be found in Section 6.4.

3 Conditional scenarios

In this section we introduce the conditional scenario concept. As already pointed out, a two-stage stochastic MILP problem formulated in terms of a continuous random vector, say $\boldsymbol{\xi} = \big(\, \boldsymbol{\xi}_1 \ldots \boldsymbol{\xi}_R \,\big)^\intercal$, is in general numerically intractable. To address this difficulty one can approximate $\boldsymbol{\xi}$ by a random vector, say $\tilde{\boldsymbol{\xi}} = \big(\, \tilde{\boldsymbol{\xi}}_1 \ldots \tilde{\boldsymbol{\xi}}_R \,\big)^\intercal$, with a discrete support consisting of S scenarios, and solve the corresponding RP problem. Since the computational complexity to solve the RP problem increases with the number of scenarios, some techniques to compute a reduced number of representative scenarios are normally used: moment matching methods [22, 26], the Sample Average Approximation (SAA) method [21, 27], approaches based on probability metrics [14, 31], among others. If the computational effort of using scenarios is high, in this paper we propose to approximate $\tilde{\boldsymbol{\xi}}$ by the rth conditional expectation $\hat{\boldsymbol{\xi}}^r = \mathbb{E}[\, \tilde{\boldsymbol{\xi}} \mid \tilde{\boldsymbol{\xi}}_r \,]$ for all $r \in \mathcal{R}$, and then formulate the CS problem in terms of these conditional expectations (this problem will be defined in the next section). The conditional expectation $\hat{\boldsymbol{\xi}}^r = \big(\, \hat{\boldsymbol{\xi}}_1^r \ldots \hat{\boldsymbol{\xi}}_R^r \,\big)^\intercal$ is a random vector with a finite number of realizations $\hat{\xi}^{re}$ and the corresponding probability values $\hat{p}^{re} = P\big(\, \hat{\boldsymbol{\xi}}^r = \hat{\xi}^{re} \,\big)$ for all $e \in \mathcal{E}_r$. Each realization $\hat{\xi}^{re}$ is called a conditional expectation scenario or, for short, *conditional scenario*. Notice that we use symbol $\tilde{\xi}^s$ with $s \in \mathcal{S}$ for scenarios and symbol $\hat{\xi}^{re}$ with $re \in \mathcal{RE}_r$ for conditional scenarios. Therefore, above a *'discretization + conditional expectation'* scheme has been outlined, which can be summarized by the chain $\boldsymbol{\xi} \rightarrow \tilde{\boldsymbol{\xi}} \rightarrow \{\hat{\boldsymbol{\xi}}^r\}_{r\in\mathcal{R}}$. That is, the conditional scenarios are computed from a given set of scenarios.

As we will see in Example 1, it is also possible to compute conditional scenarios directly from a given continuous random vector by applying a *'conditional expectation + discretization'* scheme which can be summarized by the chain $\boldsymbol{\xi} \to \{\boldsymbol{\xi}^r\}_{r\in\mathcal{R}} \to \{\hat{\boldsymbol{\xi}}^r\}_{r\in\mathcal{R}}$. That is, one approximates the continuous random vector by a set of (continuous) conditional expectation vectors and then discretizes them into a set of (discrete) conditional expectation vectors. Therefore, the two outlined approaches can be used to generate a representative set of conditional scenarios. The equivalence and effectiveness of these two approaches is a matter of further research.

Next, we present two discretization methods: Method 1 is intended to discretize a normal random variable. Method 2 is intended to discretize a multinormal random vector into conditional scenarios (it uses Method 1).

Method 1. *(Discretization of a normal random variable)*

- *Objective:* To approximate a normal random variable $\mathbf{x}$ by a finite support one $\tilde{\mathbf{x}}$.
- *Input:* $\mathbf{x}$, random variable $N(\mu, \sigma)$. E, number of discretization points. $\mathcal{I}_0 = [a,\ b[$, bounded interval such that $P\big(\mathbf{x} \in \mathcal{I}_0\big) \approx 1$.
- *Output:* $\tilde{\mathbf{x}}$, discrete random variable with support $\{\tilde{x}_e\}_{e\in\mathcal{E}}$ and the corresponding probabilities $\{\tilde{p}_e\}_{e\in\mathcal{E}}$, such that $\tilde{p}_e = P\big(\tilde{\mathbf{x}} = \tilde{x}_e\big)$ for all $e \in \mathcal{E} = \{1\ldots, E\}$.
- *Steps:*
  1) Divide $\mathcal{I}_0$ into E intervals of equal length $\mathcal{I}_e = [a_e, b_e[$ for all $e \in \mathcal{E}$, such that the union is $\mathcal{I}_0$ and the intersection is empty for any pair of them (i.e., $\{\mathcal{I}_e\}_{e\in\mathcal{E}}$ is a partition of $\mathcal{I}_0$).
  2) Compute the discretization points
  $$\tilde{x}_e = \mathbb{E}[\,\mathbf{x} \mid \mathbf{x} \in \mathcal{I}_e\,] = \mu + \sigma\,\frac{\phi(\alpha_e) - \phi(\beta_e)}{\Phi(\beta_e) - \Phi(\alpha_e)} \qquad e \in \mathcal{E}, \tag{21}$$
  where $\alpha_e = (a_e - \mu)/\sigma$ and $\beta_e = (b_e - \mu)/\sigma$. As usual, ϕ and Φ are the probability density function and the cumulative distribution function, respectivelly, of a standard normal random variable.
  3) Compute the corresponding probabilities
  $$\tilde{p}_e = P\big(\mathbf{x} \in \mathcal{I}_e\big) = \Phi(\beta_e) - \Phi(\alpha_e) \qquad e \in \mathcal{E}.$$

■

Notice that Equation (21) corresponds to the expectation of a truncated normal distribution ([24], Section 10.1). In this context, the 'best' discretization of a continuous random variable or vector according to a given criterion is known in the literature as the *optimal quantization problem* [31]. Method 1 is a quantization heuristic used in this paper to illustrate the conditional scenario approach and, of course, other discretization methods could be used. The impact of the quantization method that one uses in the conditional scenario approach is a matter that deserves further research.

As already pointed out, the following method can be used to discretize a multinormal random vector into conditional scenarios (it follows the 'conditional expectation + discretization' scheme).

Method 2. *(Conditional scenarios of a multinormal random vector)*

For all $r \in \mathcal{R}$:

1) Given $\boldsymbol{\xi} = \big(\boldsymbol{\xi}_1 \ldots \boldsymbol{\xi}_R\big)^\intercal$, a multinormal random vector such that $\boldsymbol{\xi} \sim N_R(\mu, \Sigma)$, compute the rth conditional expectation (see [19], Section 5.1):
$$\boldsymbol{\xi}^r = \mathbb{E}[\,\boldsymbol{\xi} \mid \boldsymbol{\xi}_r\,] = \mu + \frac{\boldsymbol{\xi}_r - \mu_r}{\sigma_r^2}\,\Sigma_{*r}, \tag{22}$$
where Σ_{*r} is the r-th column of the covariance matrix Σ.

2) Discretize the normal random variable $\boldsymbol{\xi}_r$ into the random variable $\tilde{\boldsymbol{\xi}}_r$ which has finite support $\{\tilde{\xi}_{re}\}_{e\in\mathcal{E}_r}$ and the corresponding probabilities $\{\tilde{p}_{re}\}_{e\in\mathcal{E}_r}$ (see Method 1).

3) By using the discrete random variable $\tilde{\boldsymbol{\xi}}_r$, discretize the continuous random vector $\boldsymbol{\xi}^r$ into the random vector $\hat{\boldsymbol{\xi}}^r$ which has finite support $\{\hat{\xi}^{re}\}_{e\in\mathcal{E}_r}$ and the corresponding probabilities $\{\hat{p}^{re}\}_{e\in\mathcal{E}_r}$. More precisely:

$$\hat{\xi}^{re} = \mathbb{E}[\,\boldsymbol{\xi} \mid \tilde{\xi}_{re}\,] = \mu + \frac{\tilde{\xi}_{re} - \mu_r}{\sigma_r^2}\,\Sigma_{*r} \qquad e \in \mathcal{E}_r$$
$$\hat{p}^{re} = \tilde{p}_{re} \qquad e \in \mathcal{E}_r.$$

■

Considering that by definition the random vector $\boldsymbol{\xi}^r$ is a transformation of the random variable $\boldsymbol{\xi}_r$, the number of the corresponding conditional scenarios is equal to the number of points used to discretize $\boldsymbol{\xi}_r$ into $\tilde{\boldsymbol{\xi}}_r$, that is, E_r. Otherwise said, we have a conditional scenario $\hat{\xi}^{re}$ per each discretization point $\tilde{\xi}_{re}$. Taking into account that the vectors to be discretized are $\boldsymbol{\xi}^1, \ldots, \boldsymbol{\xi}^R$, the total number of conditional scenarios is $\sum_{r=1}^{R} E_r$. Of course, the CS problem, which is based on conditional scenarios, is an approximation to the RP problem and for this reason, the optimal CS solution is, in general, suboptimal for the RP problem but hopefully better than the EV counterpart.

Example 1. *(Conditional scenarios of a multinormal random vector)* Let us consider the multinormal random vector $\boldsymbol{\xi} = \begin{pmatrix}\boldsymbol{\xi}_1 & \boldsymbol{\xi}_2\end{pmatrix}^\intercal \sim N_2(\mu, \Sigma)$ such that

$$\mu = \begin{pmatrix}100 & 200\end{pmatrix}^\intercal \qquad \Sigma = \begin{pmatrix}400 & 480\\ 480 & 1600\end{pmatrix}.$$

In order to illustrate the previous methods, let us approximate $\boldsymbol{\xi}$ by six conditional scenarios derived from the 1st conditional expectation. In the first step of Method 2 one computes the 1st conditional expectation by using (22):

$$\boldsymbol{\xi}^1 = \mathbb{E}[\,\boldsymbol{\xi} \mid \boldsymbol{\xi}_1\,] = \begin{pmatrix}\boldsymbol{\xi}_1\\ 80 + 1.2\boldsymbol{\xi}_1\end{pmatrix}, \tag{23}$$

where $\boldsymbol{\xi}_1 \sim N(\mu_1,\ \sigma_1)$ with $\mu_1 = 100$ and $\sigma_1 = 20$. In the second step of Method 2 one discretizes the random variable $\boldsymbol{\xi}_1$ into six representative points. By using Method 1 with $E = 6$ points and $\mathcal{I}_0 = [\mu_1 - 3\sigma_1, \mu_1 + 3\sigma_1[$, one obtains:

$$\begin{aligned}
\tilde{\xi}_{1,1} &= 53.7 & \tilde{p}_{1,1} &= 0.0214\\
\tilde{\xi}_{1,2} &= 72.3 & \tilde{p}_{1,2} &= 0.1363\\
\tilde{\xi}_{1,3} &= 90.8 & \tilde{p}_{1,3} &= 0.3423\\
\tilde{\xi}_{1,4} &= 109.2 & \tilde{p}_{1,4} &= 0.3423\\
\tilde{\xi}_{1,5} &= 127.7 & \tilde{p}_{1,5} &= 0.1363\\
\tilde{\xi}_{1,6} &= 146.3 & \tilde{p}_{1,6} &= 0.0214.
\end{aligned}$$

In the third step of Method 2 one computes the conditional scenarios combining the previous discretization of $\boldsymbol{\xi}_1$ with (23). For example the first conditional scenario $\hat{\xi}^{1,1}$ can be computed as follows:

$$\hat{\xi}^{1,1} = \mathbb{E}[\,\boldsymbol{\xi} \mid \tilde{\xi}_{1,1}\,] = \begin{pmatrix}\tilde{\xi}_{1,1}\\ 80 + 1.2\tilde{\xi}_{1,1}\end{pmatrix} = \begin{pmatrix}53.7\\ 80 + 1.2\cdot 53.7\end{pmatrix} = \begin{pmatrix}53.7\\ 144.4\end{pmatrix}$$
$$\hat{p}^{1,1} = \tilde{p}_{1,1} = 0.0215.$$

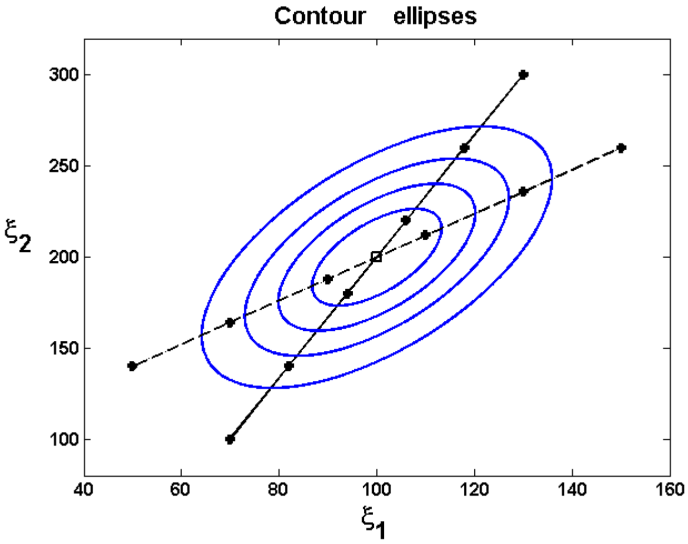

Figure 1: Equiprobability contour ellipses of the multinormal random vector $\boldsymbol{\xi}$ of Example 1. The dashed and the solid lines correspond to the first ($\boldsymbol{\xi}^1$) and the second ($\boldsymbol{\xi}^2$) conditional expectations, respectively. The dots represent the twelve conditional scenarios and the square represents the 'expected scenario' μ.

In summary we obtain:

$$\begin{array}{ll} \hat{\xi}^{1,1} = \begin{pmatrix} 53.7 & 144.4 \end{pmatrix}^\intercal & \hat{p}^{1,1} = \tilde{p}_{1,1} \\ \hat{\xi}^{1,2} = \begin{pmatrix} 72.3 & 166.8 \end{pmatrix}^\intercal & \hat{p}^{1,2} = \tilde{p}_{1,2} \\ \hat{\xi}^{1,3} = \begin{pmatrix} 90.8 & 189.0 \end{pmatrix}^\intercal & \hat{p}^{1,3} = \tilde{p}_{1,3} \\ \hat{\xi}^{1,4} = \begin{pmatrix} 109.2 & 211.0 \end{pmatrix}^\intercal & \hat{p}^{1,4} = \tilde{p}_{1,4} \\ \hat{\xi}^{1,5} = \begin{pmatrix} 127.7 & 233.2 \end{pmatrix}^\intercal & \hat{p}^{1,5} = \tilde{p}_{1,5} \\ \hat{\xi}^{1,6} = \begin{pmatrix} 146.3 & 255.6 \end{pmatrix}^\intercal & \hat{p}^{1,6} = \tilde{p}_{1,6}. \end{array}$$

The approximation of $\boldsymbol{\xi}$ by six conditional scenarios derived from the 2nd conditional expectation $\hat{\boldsymbol{\xi}}^2$ could be done analogously. In total we would obtain twelve conditional scenarios which are represented in Figure 1. As we will see in the next section, the CS problem is formulated in terms of conditional scenarios in contrast with the EV problem, which is formulated in terms of the 'expected scenario' μ (the square in Figure 1). ■

4 The conditional scenario problem

Now let us see that, given a random vector $\tilde{\boldsymbol{\xi}}$, its conditional expectation $\hat{\boldsymbol{\xi}}^r$ is an 'optimal approximation' of $\tilde{\boldsymbol{\xi}}$ for any $r \in \mathcal{R}$. The *conditional expectation* $\hat{\boldsymbol{\xi}}^r = \mathbb{E}[\, \tilde{\boldsymbol{\xi}} \mid \tilde{\boldsymbol{\xi}}_r \,]$, viewed as a function $h_r^*(\tilde{\boldsymbol{\xi}}_r)$, can be interpreted as an approximation of the random vector $\tilde{\boldsymbol{\xi}}$ by a function of the random variable $\tilde{\boldsymbol{\xi}}_r$. The error of this approximation is then given by

$$U = \tilde{\boldsymbol{\xi}} - h_r^*(\tilde{\boldsymbol{\xi}}_r).$$

The approximation h_r^* has appealing properties (Theorem 4.3 in [19]): a) It has a null expected error, that is, $\mathbb{E}[\, U \,] = 0$. b) It is an optimal approximation of $\tilde{\boldsymbol{\xi}}$ in the sense that it minimizes the mean squared error (MSE), where

$$MSE(h_r) = \mathbb{E}[\, \big(\tilde{\boldsymbol{\xi}} - h_r(\tilde{\boldsymbol{\xi}}_r) \big)^\intercal \big(\tilde{\boldsymbol{\xi}} - h_r(\tilde{\boldsymbol{\xi}}_r) \big) \,] = \mathbb{E}[\, \|\tilde{\boldsymbol{\xi}} - h_r(\tilde{\boldsymbol{\xi}}_r)\|^2 \,].$$

Now that $\mathbb{E}[\,\tilde{\boldsymbol{\xi}} \mid \tilde{\boldsymbol{\xi}}_r\,]$, known as the regression function of $\tilde{\boldsymbol{\xi}}$ on $\tilde{\boldsymbol{\xi}}_r$, is an optimal approximation to $\tilde{\boldsymbol{\xi}}$ as a function of $\tilde{\boldsymbol{\xi}}_r$, the question is: Which approximation $\hat{\boldsymbol{\xi}}^r$ should one use? A possible answer is to use all of them: $\hat{\boldsymbol{\xi}}^1, \ldots, \hat{\boldsymbol{\xi}}^R$. Since we do not wish simultaneous approximations, we consider a random index $\mathbf{r}$ with uniform distribution in $\mathcal{R}$, that is, $P\big(\mathbf{r} = r\big) = 1/R$ for all $r \in \mathcal{R}$. Then we consider the *randomized conditional expectation* $\hat{\boldsymbol{\xi}}^{\mathbf{r}}$ (notice the random superscript written in boldface) which is equivalent to consider each $\hat{\boldsymbol{\xi}}^r$ with probability $1/R$ for all $r \in \mathcal{R}$.

Next we define the Conditional Scenario (CS) problem as the approximation to the RP problem based on the approximation of the random vector $\tilde{\boldsymbol{\xi}}$ by the set of conditional expectations $\{\hat{\boldsymbol{\xi}}^r\}_{r\in\mathcal{R}}$, each one taken with probability $1/R$. The CS problem can be stated as follows (stochastic version):

$$
\begin{aligned}
\min_{\hat{x}} \quad & z_{CS} = c_1^\intercal \hat{x}_1 + \frac{1}{R}\sum_{r\in\mathcal{R}} \mathbb{E}[\hat{\mathbf{c}}_2^{r\,\intercal}\, \hat{x}_2(\hat{\boldsymbol{\xi}}^r)] && && (24)\\
\text{s.t.} \quad & A_1\hat{x}_1 = b_1 && && (25)\\
& \hat{\mathbf{A}}_2^r \hat{x}_1 + \hat{\mathbf{B}}_2^r \hat{x}_2(\hat{\boldsymbol{\xi}}^r) = \hat{\mathbf{b}}_2^r && \text{w.p.1},\ r\in\mathcal{R} && (26)\\
& \hat{x}_1 \geq 0 && && (27)\\
& \hat{x}_2(\hat{\boldsymbol{\xi}}^r) \geq 0 && \text{w.p.1},\ r\in\mathcal{R} && (28)\\
& \hat{x}_{1j} \text{ integer} && j\in\mathcal{J}_1 && (29)\\
& \hat{x}_{2j}(\hat{\boldsymbol{\xi}}^r) \text{ integer} && \text{w.p.1},\ r\in\mathcal{R},\ j\in\mathcal{J}_2, && (30)
\end{aligned}
$$

where

$$
\begin{aligned}
\hat{\boldsymbol{\xi}}^r &= \mathbb{E}[\,\tilde{\boldsymbol{\xi}} \mid \tilde{\boldsymbol{\xi}}_r\,] = \text{vec}\big(\hat{\mathbf{c}}_2^r, \hat{\mathbf{A}}_2^r, \hat{\mathbf{B}}_2^r, \hat{\mathbf{b}}_2^r\big) && r\in\mathcal{R}\\
\hat{\mathbf{c}}_2^r &= \mathbb{E}[\,\tilde{\mathbf{c}}_2 \mid \tilde{\boldsymbol{\xi}}_r\,], \quad \hat{\mathbf{A}}_2^r = \mathbb{E}[\,\tilde{\mathbf{A}}_2 \mid \tilde{\boldsymbol{\xi}}_r\,] && r\in\mathcal{R}\\
\hat{\mathbf{B}}_2^r &= \mathbb{E}[\,\tilde{\mathbf{B}}_2 \mid \tilde{\boldsymbol{\xi}}_r\,], \quad \hat{\mathbf{b}}_2^r = \mathbb{E}[\,\tilde{\mathbf{b}}_2 \mid \tilde{\boldsymbol{\xi}}_r\,] && r\in\mathcal{R}.
\end{aligned}
$$

Finally, we state the CS problem in the deterministic equivalent version:

$$
\begin{aligned}
\min_{\hat{x}} \quad & z_{CS} = c_1^\intercal \hat{x}_1 + \frac{1}{R}\sum_{re\in\mathcal{RE}_r} \hat{p}^{re}\, \hat{c}_2^{re\intercal} \hat{x}_2^{re} && && (31)\\
\text{s.t.} \quad & A_1\hat{x}_1 = b_1 && && (32)\\
& \hat{A}_2^{re}\hat{x}_1 + \hat{B}_2^{re}\hat{x}_2^{re} = \hat{b}_2^{re} && re\in\mathcal{RE}_r && (33)\\
& \hat{x}_1 \geq 0 && && (34)\\
& \hat{x}_2^{re} \geq 0 && re\in\mathcal{RE}_r && (35)\\
& \hat{x}_{1j} \text{ integer} && j\in\mathcal{J}_1 && (36)\\
& \hat{x}_{2j}^{re} \text{ integer} && re\in\mathcal{RE}_r,\ j\in\mathcal{J}_2. && (37)
\end{aligned}
$$

The deterministic equivalent version is used for numerical purposes (see Section 6) and the stochastic one for theoretical analysis (see the Appendix).

5 Useful bounds

In this section we give some bounds which can be used to assess the quality of the optimal EV and CS solutions. Normally, the EV and CS problems are solved to approximate difficult RP instances. In this case, one cannot assess the quality of the optimal EV and CS solutions by using the optimal RP cost, that is, one cannot compute the EV and CS optimality gaps. If one has a lower bound for the RP optimal cost, a worst-case optimality gap can be computed for the optimal EV and CS solutions (see Section 6.4 for details). In what follows the values z^*_{EV}, z^*_{CS} and z^*_{RP} stand for the optimal cost of the EV, CS and RP problems, respectively. The proofs of the following theoretical results are in the Appendix.

Proposition 1. *Let us consider the RP problem (1)–(7) where some or all of the components of A_2 and/or b_2 are stochastic (the other parameters being deterministic). Let us assume that $\mathcal{J}_2 = \emptyset$, that is, all the second stage variables are continuous in the EV, CS and RP problems. Then*

$$z^*_{EV} \leq z^*_{CS} \leq z^*_{RP}.$$

Remark 1. This result also applies if the recourse problem has integer second stage variables. To explain this remark, let us consider a recourse problem, say RP_1, with stochastic A_2 and/or b_2 such that $\mathcal{J}_2(RP_1) \neq \emptyset$. One can consider: problem RP, obtained from problem RP_1 by dropping the integrality condition ($\mathcal{J}_2 = \emptyset$) and the corresponding problems CS and EV (they have the same $\mathcal{J}_2$). Then, by Proposition 1, it is clear that

$$z^*_{EV} \leq z^*_{CS} \leq z^*_{RP} \leq z^*_{RP_1}.$$

Proposition 2. *Let us consider the RP problem (1)–(7) where some or all the components of c_2 are stochastic (the other parameters being deterministic). Then*

$$z^*_{RP} \leq z^*_{CS} \leq z^*_{EV}.$$

Approximating the RP problem by the EV problem has three main drawbacks. First, as one would expect, any optimal EV decision, say $\bar{x}^*_1$, is in general suboptimal for the RP problem (assuming that $\bar{x}^*_1$ is feasible for the RP problem). Second, the optimal EV cost z^*_{EV} is usually misleading, since it does not correspond to the true expected cost associated with $\bar{x}^*_1$, the so-called *expected result of using the EV solution (E-EV)* [10]. It can be computed by solving the E-EV problem which corresponds to the RP problem with the additional constraint $\tilde{x}_1 = \bar{x}^*_1$ (see Section 6.4). Third, the EV problem does not take parameter uncertainty, if any, into account. On the other hand, approximating the RP problem by the CS problem, only avoids the third drawback. For this reason, we define the *expected result of using the CS solution (E-CS)* as the counterpart of the E-EV, which can be computed analogously.

Proposition 3. *Let us consider the RP problem (1)–(7) where some or all the second stage parameters are stochastic (c_1, A_1 and b_1 are always deterministic). Let us assume that $\bar{x}^*_1$ and $\hat{x}^*_1$, optimal first stage solutions for the EV and CS problems, respectively, are feasible for the RP problem. Then*

$$z^*_{RP} \leq z^*_{E\text{-}EV} \quad \textit{and} \quad z^*_{RP} \leq z^*_{E\text{-}CS}.$$

Proposition 4. *Let us consider the RP problem (1)–(7) where some or all the components of c_2 are stochastic (the other parameters being deterministic). Then*

$$z^*_{RP} \leq z^*_{E\text{-}EV} \leq z^*_{EV} \quad \textit{and} \quad z^*_{RP} \leq z^*_{E\text{-}CS} \leq z^*_{CS}.$$

Remark 2. A particularly useful case is the RP problem such that $\mathcal{J}_2 = \emptyset$ and only A_2 and/or b_2 are stochastic (the RP problem has fixed recourse and deterministic c_2). In this case we have that

$$z^*_{CS} \leq \; z^*_{RP} \leq z^*_{E\text{-}CS} \quad \text{and that} \quad z^*_{CS} \leq \; z^*_{RP} \leq z^*_{E\text{-}EV},$$

which allows to compute worst-case optimality gaps for the optimal EV and CS solutions (notice that we use z^*_{CS} in both cases since by Proposition 1 it is a tighter bound for z^*_{RP} than z^*_{EV}). In the next section we will illustrate this case.

6 Numerical example

In order to illustrate the CS approach, in this section we use a multi-farm feed manufacturer problem which can be modelled as an RP problem with uncertain left-hand sides. This problem has been inspired by the farmer's problem in [10]. Our objective is to compare the EV, CS and SAA approaches as approximations to the RP problem intended to reduce the computational burden. Computations have been conducted on a PC using Windows 7 (64 bits), with a Intel Core i5 processor, 2.67GHz and 8 GB of RAM. The corresponding MILP problems have been solved by CPLEX 12.6 with default parameters.

Table 1: Parameters of the EV multi-farm feed manufacturer problem.

Parameter	Value	Unit	Description	
I	10	-	Number of crops	
$\mathcal{I}$	$\{1,\ldots,I\}$	-	Index set for crops	
i	-	-	Index for crops	$i \in \mathcal{I}$
J	5	-	Number of farms	
$\mathcal{J}$	$\{1,\ldots,J\}$	-	Index set for farms	
j	-	-	Index for farms	$j \in \mathcal{J}$
c_{1ij}	$100 + 3\big(I(j-1)+i\big)$	euros/acre	Planting cost of crop i in farm j	$i \in \mathcal{I}, j \in \mathcal{J}$
$\bar{q}_{2ij}$	$2 + 0.02\big(I(j-1)+i\big)$	tonnes/acre	Expected yield rate of crop i in farm j	$i \in \mathcal{I}, j \in \mathcal{J}$
b_{1j}	$150 + 50j$	acres	Available land in farm j	$j \in \mathcal{J}$
b_{2i}	$400 + 12i$	tonnes	Amount of crop i needed	$i \in \mathcal{I}$
f_{2i}	$2.3\big(\sum_{j\in\mathcal{J}}(c_{1ij}/\bar{q}_{2ij})\big)/J$	euros/tonne	Buying price of crop i	$i \in \mathcal{I}$
g_{2i}	$0.5 f_{2i}$	euros/tonne	Selling price of crop i	$i \in \mathcal{I}$

6.1 The EV multi-farm feed manufacturer problem

Example 2. *(The EV multi-farm feed manufacturer problem: expected yields)*

Consider a feed manufacturer who specializes in raising and manufacturing several types of feed ingredients as for example corn, soybeans, sorghum, oats and barley. In total he plants I types of crops on his J farms. Each farm has b_{1j} acres of land. Based on experience, the manufacturer knows for each crop $i \in \mathcal{I}$ and for each farm $j \in \mathcal{J}$: a) That the mean yield is $\bar{q}_{2ij}$ tonnes/acre. b) That the planting cost is c_{1ij} euros/acre. c) That b_{2i} tonnes are needed for the next feed manufacturing season. These amounts can be raised on the manufacturer farms or purchased from the market and any production in excess is sold. For each crop i, the purchasing and selling prices for the manufacturer are f_{2i} and g_{2i}, respectively. Furthermore, for economic reasons each crop can be raised, at most, on two farms. Table 1 summarizes the parameters of this problem. It is clear that this is a synthetic example which will be used to illustrate the use of the conditional scenario (CS) approach.

The feed manufacturer wants to decide how much land to devote to each crop in order to obtain the feed ingredients for the next season at the minimum cost. Furthermore, he also wishes to decide the farms on which to raise each crop (at most two farms per crop).

The first step to solve this problem is to define the decision variables (Table 2). The second step is to

Table 2: Decision variables of the EV multi-farm feed manufacturer problem.

Decision	Unit	Description	
$\bar{u}_{1ij}$	-	$\bar{u}_{1ij} = 1$, if crop i is raised in farm j	$i \in \mathcal{I}, j \in \mathcal{J}$
		$\bar{u}_{1ij} = 0$, otherwise	
$\bar{x}_{1ij}$	acres	Land devoted to crop i in farm j	$i \in \mathcal{I}, j \in \mathcal{J}$
$\bar{y}_{2i}$	tonnes	Amount of crop i purchased	$i \in \mathcal{I}$
		(under production shortage)	
$\bar{z}_{2i}$	tonnes	Amount of crop i sold	$i \in \mathcal{I}$
		(under production excess)	

formulate the EV problem:

$$\min_{\bar{u},\bar{x},\bar{y},\bar{z}} \quad z_{EV} = \sum_{i\in\mathcal{I},j\in\mathcal{J}} c_{1ij}\,\bar{x}_{1ij} + \sum_{i\in\mathcal{I}} \big(f_{2i}\,\bar{y}_{2i} - g_{2i}\bar{z}_{2i} \big) \tag{38}$$

$$\text{s.t.} \quad \sum_{i\in\mathcal{I}} \bar{x}_{1ij} \le b_{1j} \qquad j \in \mathcal{J} \tag{39}$$

$$\sum_{j\in\mathcal{J}} \bar{q}_{2ij}\,\bar{x}_{1ij} + \bar{y}_{2i} - \bar{z}_{2i} = b_{2i} \qquad i \in \mathcal{I} \tag{40}$$

$$\sum_{j\in\mathcal{J}} \bar{u}_{1ij} \le 2 \qquad i \in \mathcal{I} \tag{41}$$

$$\bar{x}_{1ij} \le \bar{u}_{1ij}\, b_{1j} \qquad i \in \mathcal{I}, j \in \mathcal{J} \tag{42}$$

$$\bar{u}_{1ij} \in \{1, 0\} \qquad i \in \mathcal{I}, j \in \mathcal{J} \tag{43}$$

$$\bar{x}_1 \ge 0 \tag{44}$$

$$\bar{y}_2 \ge 0, \quad \bar{z}_2 \ge 0. \tag{45}$$

- In (38) the total cost is computed as the planting cost $c_1^\intercal \bar{x}_1$ plus the crop purchasing cost $f_2^\intercal \bar{y}_2$ (under production shortage) minus the crop selling revenue $g_2^\intercal \bar{z}_2$ (under production excess).
- Constraint (39) states that, for each farm j, the planted land must be no more than the available land b_{1j}.
- In (40) there is a balance equation such that, for each crop i, the corresponding production must be equal to the crop demand b_{2i} with the help of purchasing or selling some crops, $\bar{y}_{2i}$ or $\bar{z}_{2i}$, respectively, if necessary.
- Constraints (41) and (42) ensure that each crop can be raised, at most, on two farms.
- We remark this MILP problem has 50 binary variables, 70 continuous variables and 75 constraints.

After solving this EV problem by using the MILP solver (CPLEX) one obtains:

$$z_{EV}^* = 401,274$$

$$\bar{x}_1^* = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 27 \\ 0 & 0 & 0 & 0 & 152 \\ 0 & 0 & 0 & 0 & 156 \\ 0 & 0 & 0 & 101 & 65 \\ 0 & 0 & 0 & 174 & 0 \\ 0 & 0 & 109 & 75 & 0 \\ 0 & 3 & 191 & 0 & 0 \\ 0 & 214 & 0 & 0 & 0 \\ 200 & 33 & 0 & 0 & 0 \end{pmatrix} \quad \bar{y}_2^* = \begin{pmatrix} 412 \\ 346 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \bar{z}_2^* = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

That is, for each $i \in \mathcal{I}, j \in \mathcal{J}$, devote $\bar{x}_{1ij}^*$ acres to crop i in farm j and purchase $\bar{y}_{2i}^*$ tonnes of crop i to obtain the optimal EV cost (401,274 euros). Notice that there is no production excess ($\bar{z}_2^* = 0$). ■

6.2 The CS multi-farm feed manufacturer problem

Example 3. *(The CS multi-farm feed manufacturer problem: uncertain yields)*

In the previous example, although the yield rates are uncertain prior to raising the crops, the EV problem has been formulated in terms of the corresponding expected values, which are deterministic. In the current example we show how the CS problem takes into account the uncertainty of the yield rates. Let us assume that the yield rates can be modeled as a multinormal random vector $\mathbf{q}_2 = \text{vec}\big((\mathbf{q}_{2ij})_{i \in \mathcal{I}, j \in \mathcal{J}} \big)$. Now, the CS problem is a two-stage one: In the first stage, the manufacturer decides on the amount of land and the farms to devote to each crop. In the second stage, the manufacturer will balance the need of each crop by purchasing or selling according to the yield observed. The description of the random yield rates and the related parameters can be found in Table 3.

The objective of the CS problem is to decide how much land to devote to each crop on each farm in order to obtain the feed ingredients for the next season at the minimum *expected* cost.

In Table 4 we have the CS decision variables. To formulate the CS problem we calculate the conditional scenarios that approximate the multinormal random vector $\mathbf{q}_2$, by using Method 2 with $R = 50$ random parameters and Method 1 with $E = 33$ discretization points and $\mathcal{I}_0 = [\mu - 4\sigma, \mu + 4\sigma[$. The choice of $E = 33$ was as follows. In order to tune E, we considered the news vendor problem [10] with a normal random demand and solved it analytically. Then we considered its scenario version by discretizing the continuous random demand (Method 1 with $E = 10, 20, \dots, 100$). We observed that the scenario optimal solution was of good quality compared to the analytical one for the discretizations with $E \geq 30$. For this reason, in the CS multi-farm feed manufacturer problem, we decided to use $E \approx 30$ in order to balance quality of the discretization and computational burden. Finally, we took $E = 33$ in order to have a discrete distribution with a single mode at μ, and in order to divide $\mathcal{I}_0$ into 33 subintervals of length $\sigma/4$ (approximately) which we found reasonable. Of course this was an heuristic approach and, as already pointed in Section 3, the impact of the quantization method that one uses in the conditional scenario approach is a matter that deserves further research.

Given that we set $E_r = E = 33$ for all $r \in \mathcal{R}$, the CS problem considers a total of $R \cdot E = 1650$

Table 3: Parameters of the CS multi-farm feed manufacturer problem.

Parameter	Value	Unit	Description	
$\mathbf{q}_2$	-	tonnes/acre	Random yield rates $\mathbf{q}_2 \sim N_R(\mu, \Sigma)$	
μ_{ij}	$\bar{q}_{2ij}$	tonnes/acre	Expected yield rate of crop i in farm j	$i \in \mathcal{I}, j \in \mathcal{J}$
σ_{ij}	$\bar{q}_{2ij}/4$	tonnes/acre	Standard deviation of $\mathbf{q}_{2ij}$	$i \in \mathcal{I}, j \in \mathcal{J}$
$\rho_{i_1 j_1, i_2 j_2}$	0.7	-	Correlation between $\mathbf{q}_{2i_1j_1}$ and $\mathbf{q}_{2i_2j_2}$, $\quad i_1j_1 \neq i_2j_2$	$i_1 \in \mathcal{I},\ j_1 \in \mathcal{J}$ $i_2 \in \mathcal{I},\ j_2 \in \mathcal{J}$
$\sigma_{i_1 j_1, i_2 j_2}$	$\rho_{i_1 j_1, i_2 j_2}\, \sigma_{i_1 j_1} \sigma_{i_2 j_2}$	(tonnes/acre)^2	Covariance between $\mathbf{q}_{2i_1j_1}$ and $\mathbf{q}_{2i_2j_2}$, $\quad i_1j_1 \neq i_2j_2$	$i_1 \in \mathcal{I},\ j_1 \in \mathcal{J}$ $i_2 \in \mathcal{I},\ j_2 \in \mathcal{J}$
R	50	-	Number of random parameters	$R = I \cdot J$
E_r	33	-	Number of conditional scenarios derived from the rth conditional expectation	$r \in \mathcal{R}$

Table 4: Decision variables of the CS multi-farm feed manufacturer problem.

Decision	Unit	Description	
$\hat{u}_{1ij}$	-	$\hat{u}_{1ij} = 1$, if crop i is raised in farm j $\hat{u}_{1ij} = 0$, otherwise	$i \in \mathcal{I}, j \in \mathcal{J}$
$\hat{x}_{1ij}$	acres	Land devoted to crop i in farm j	$i \in \mathcal{I}, j \in \mathcal{J}$
$\hat{y}_{2i}^{re}$	tonnes	Amount of crop i purchased assuming the conditional scenario re (under production shortage)	$re \in \mathcal{R}\mathcal{E}_r,\ i \in \mathcal{I}$
$\hat{z}_{2i}^{re}$	tonnes	Amount of crop i sold assuming the conditional scenario re (under production excess)	$re \in \mathcal{R}\mathcal{E}_r,\ i \in \mathcal{I}$

conditional scenarios. Then the CS problem (31)–(37), in this case, can be written as follows:

$$\min_{\hat{u},\hat{x},\hat{y},\hat{z}} \quad z_{CS} = \sum_{i\in\mathcal{I},j\in\mathcal{J}} c_{1ij}\,\hat{x}_{1ij} + \frac{1}{R}\sum_{re\in\mathcal{RE}_r}\sum_{i\in\mathcal{I}} \hat{p}^{re}\big(f_{2i}\,\hat{y}_{2i}^{re} - g_{2i}\,\hat{z}_{2i}^{re}\big) \tag{46}$$

$$\text{s.t.} \quad \sum_{i\in\mathcal{I}} \hat{x}_{1ij} \le b_{1j} \qquad j\in\mathcal{J} \tag{47}$$

$$\sum_{j\in\mathcal{J}} \hat{q}_{2ij}^{re}\,\hat{x}_{1ij} + \hat{y}_{2i}^{re} - \hat{z}_{2i}^{re} = b_{2i} \qquad re\in\mathcal{RE}_r, i\in\mathcal{I} \tag{48}$$

$$\sum_{j\in\mathcal{J}} \hat{u}_{1ij} \le 2 \qquad i\in\mathcal{I} \tag{49}$$

$$\hat{x}_{1ij} \le \hat{u}_{1ij}\,b_{1j} \qquad i\in\mathcal{I}, j\in\mathcal{J} \tag{50}$$

$$\hat{u}_{1ij} \in \{1,0\} \qquad i\in\mathcal{I}, j\in\mathcal{J} \tag{51}$$

$$\hat{x}_1 \ge 0 \tag{52}$$

$$\hat{y}_2^{re} \ge 0, \quad \hat{z}_2^{re} \ge 0 \qquad re\in\mathcal{RE}_r. \tag{53}$$

The main difference with the EV problem is that now the uncertainty is taken into account (by means of 1650 conditional scenarios $\hat{q}_2^{re}$). The resulting MILP problem has 50 binary variables, 33,050 continuous variables and 16,565 constraints. After solving the CS problem by using the MILP solver (CPLEX), one obtains:

$$z_{CS}^* = 409,595 \text{ euros}$$

$$\hat{x}_1^* = \begin{pmatrix} 48 & 101 & 0 & 0 & 0 \\ 77 & 80 & 0 & 0 & 0 \\ 75 & 0 & 79 & 0 & 0 \\ 0 & 69 & 86 & 0 & 0 \\ 0 & 0 & 71 & 75 & 0 \\ 0 & 0 & 64 & 84 & 0 \\ 0 & 0 & 0 & 87 & 60 \\ 0 & 0 & 0 & 34 & 112 \\ 0 & 0 & 0 & 35 & 113 \\ 0 & 0 & 0 & 35 & 115 \end{pmatrix} \text{acres.}$$

Unlike in Example 2, here we do not report the optimal vectors $\hat{y}_2^*$ and $\hat{z}_2^*$ given their high dimensions. ■

6.3 The SAA multi-farm feed manufacturer problem

In this context a natural question which arises is: what would be the performance of the RP problem using the same number of scenarios as in the CS problem? We study this question in the following example where the scenarios are randomly sampled. Solving the RP problem formulated in terms of a set of randomly sampled scenarios can be considered as a simple version of the SAA method [27].

Example 4. *(The SAA multi-farm feed manufacturer problem: uncertain yields)*

In Table 5 we have the SAA decision variables: To formulate the SAA problem we consider $S = 1,650$ scenarios randomly sampled from the multinormal random vector $\mathbf{q}_2$ defined in the previous example.

Table 5: Decision variables of the SAA multi-farm feed manufacturer problem.

Decision	Unit	Description	
$\tilde{u}_{1ij}$	-	$\tilde{u}_{1ij} = 1$, if crop i is raised in farm j $\tilde{u}_{1ij} = 0$, otherwise	$i \in \mathcal{I}, j \in \mathcal{J}$
$\tilde{x}_{1ij}$	acres	Land devoted to crop i in farm j	$i \in \mathcal{I}, j \in \mathcal{J}$
$\tilde{y}^s_{2i}$	tonnes	Amount of crop i purchased (under production shortage)	$s \in \mathcal{S},\ i \in \mathcal{I}$
$\tilde{z}^s_{2i}$	tonnes	Amount of crop i sold (under production excess)	$s \in \mathcal{S},\ i \in \mathcal{I}$

Then the SAA feed manufacturer problem can be written as follows:

$$\min_{\tilde{u},\tilde{x},\tilde{y},\tilde{z}} \quad z_{SAA} = \sum_{i\in\mathcal{I},j\in\mathcal{J}} c_{1ij}\, \tilde{x}_{1ij} + \frac{1}{S}\sum_{s\in\mathcal{S}}\sum_{i\in\mathcal{I}} f_{2i}\, \tilde{y}^s_{2i} - g_{2i}\, \tilde{z}^s_{2i} \tag{54}$$

$$\text{s.t.} \quad \sum_{i\in\mathcal{I}} \tilde{x}_{1ij} \le b_{1j} \qquad j \in \mathcal{J} \tag{55}$$

$$\sum_{j\in\mathcal{J}} \tilde{q}^s_{2ij}\, \tilde{x}_{1ij} + \tilde{y}^s_{2i} - \tilde{z}^s_{2i} = b_{2i} \qquad s \in \mathcal{S}, i \in \mathcal{I} \tag{56}$$

$$\sum_{j\in\mathcal{J}} \tilde{u}_{1ij} \le 2 \qquad i \in \mathcal{I} \tag{57}$$

$$\tilde{x}_{1ij} \le \tilde{u}_{1ij}\, b_{1j} \qquad i \in \mathcal{I}, j \in \mathcal{J} \tag{58}$$

$$\tilde{u}_{1ij} \in \{1, 0\} \qquad i \in \mathcal{I}, j \in \mathcal{J} \tag{59}$$

$$\tilde{x}_1 \ge 0 \tag{60}$$

$$\tilde{y}^s_2 \ge 0, \quad \tilde{z}^s_2 \ge 0 \qquad s \in \mathcal{S}. \tag{61}$$

Notice that this problem corresponds to the RP problem (8)–(14) formulated in terms of $1,650$ randomly sampled scenarios. The only difference with the CS multi-farm feed manufacturer problem is the way it is used to approximate the random vector $\mathbf{q}_2$ (scenarios $\tilde{q}^s_2$ versus conditional scenarios $\hat{q}^{re}_2$). In this case, the resulting MILP problems have the same dimensions for the two approaches. After solving the SAA problem by using the MILP solver (CPLEX) one obtains:

$$z^*_{SAA} = 410,654 \text{ euros}$$

$$\tilde{x}^*_1 = \begin{pmatrix} 0 & 58 & 0 & 75 & 0 \\ 0 & 55 & 0 & 0 & 77 \\ 0 & 46 & 0 & 0 & 87 \\ 0 & 0 & 83 & 0 & 56 \\ 0 & 0 & 0 & 47 & 87 \\ 0 & 0 & 102 & 51 & 0 \\ 91 & 91 & 0 & 0 & 0 \\ 0 & 0 & 50 & 106 & 0 \\ 109 & 0 & 0 & 71 & 0 \\ 0 & 0 & 65 & 0 & 93 \end{pmatrix} \text{acres.}$$

■

Table 6: Comparing the EV, CS and SAA solutions.

	EV	CS	SAA	E-SAA	E-CS	E-EV
Cost (euros)	401,274	409,595	410,654	413,724	413,758	431,096
Variation (%)	–2.03	-	-	+1.01	+1.02	+5.25
Time (seconds)	0.006	0.800	93.800	190	190	190

6.4 Comparing the EV, CS and SAA solutions

Example 5. *(The E-EV, E-CS and E-SAA multi-farm feed manufacturer problems: uncertain yields)*

With the data of Examples 2, 3 and 4, let us compare the values EV, CE, SAA, E-EV, E-CS and E-SAA which are the optimal values of the corresponding problems (E-EV stands for the Expected result of using the EV solution, E-CS and E-SAA, analogously). To compute the E-EV we formulate the E-EV problem:

$$\min_{\tilde{u},\tilde{x},\tilde{y},\tilde{z}} \quad z_{E\text{-}EV} = \sum_{i\in\mathcal{I},j\in\mathcal{J}} c_{1ij}\,\tilde{x}_{1ij} + \sum_{s\in\mathcal{S}}\sum_{i\in\mathcal{I}} \tilde{p}^s\big(\,f_{2i}\,\tilde{y}^s_{2i} - g_{2i}\,\tilde{z}^s_{2i}\,\big) \tag{62}$$

$$\text{s.t.} \quad \sum_{i\in\mathcal{I}} \tilde{x}_{1ij} \le b_{1j} \qquad j\in\mathcal{J} \tag{63}$$

$$\sum_{j\in\mathcal{J}} \tilde{q}^s_{2ij}\,\tilde{x}_{1ij} + \tilde{y}^s_{2i} - \tilde{z}^s_{2i} = b_{2i} \qquad s\in\mathcal{S}, i\in\mathcal{I} \tag{64}$$

$$\sum_{j\in\mathcal{J}} \tilde{u}_{1ij} \le 2 \qquad i\in\mathcal{I} \tag{65}$$

$$\tilde{x}_{1ij} \le \tilde{u}_{1ij}\,b_{1j} \qquad i\in\mathcal{I}, j\in\mathcal{J} \tag{66}$$

$$\tilde{u}_{1ij} \in \{1,0\} \qquad i\in\mathcal{I}, j\in\mathcal{J} \tag{67}$$

$$(\tilde{x}_1,\tilde{u}_1) = (x_{ref}, u_{ref}) \tag{68}$$

$$\tilde{x}_1 \ge 0 \tag{69}$$

$$\tilde{y}^s_2 \ge 0, \quad \tilde{z}^s_2 \ge 0 \qquad s\in\mathcal{S}, \tag{70}$$

where we set the reference vector as $(x_{ref}, u_{ref}) = (\bar{x}^*_1, \bar{u}^*_1)$ obtained by solving the EV problem. Notice that this is nothing but the RP problem with the additional constraint $(\tilde{x}_1,\tilde{u}_1) = (\bar{x}^*_1, \bar{u}^*_1)$. Also notice that in this example the RP problem has *relatively complete recourse* [10]. Therefore, any reference solution (x_{ref}, u_{ref}) which is feasible for the EV or CS problems will also be feasible for the RP problem. It is well known that in this case $z^*_{E\text{-}EV}$ can be computed as follows [11]:

$$z^*_{E\text{-}EV} = \sum_{i\in\mathcal{I},j\in\mathcal{J}} c_{1ij}\,\bar{x}^*_{1ij} + \sum_{s\in\mathcal{S}}\sum_{i\in\mathcal{I}} \tilde{p}^s\Big(\,f_{2i}\,\Big[b_{2i} - \sum_{j\in\mathcal{J}} \tilde{q}^s_{2ij}\,\bar{x}^*_{1ij}\Big]_+ - g_{2i}\,\Big[\sum_{j\in\mathcal{J}} \tilde{q}^s_{2ij}\,\bar{x}^*_{1ij} - b_{2i}\Big]_+\Big),$$

where $[\;]_+$ is the positive part, that is, given a scalar, say v, then $[v]_+ = \max\{0, v\}$. To set the E-EV problem we have used $S = 10^6$ scenarios $\{\tilde{q}^s_2\}_{s\in\mathcal{S}}$ randomly sampled from the multinormal random vector $\mathbf{q}_2 \sim N_{50}(\mu,\Sigma)$. In this case one has $\tilde{p}^s = 1/S$ for all $s\in\mathcal{S}$. On the other hand, the E-CS can be obtained by solving the previous problem but setting the reference vector as $(x_{ref}, u_{ref}) = (\hat{x}^*_1, \hat{u}^*_1)$ obtained by solving the CS problem. The E-SAA can be obtained analogously. The results have been summarized in Table 6. In the first row we observe EV $\le$ CS $\le$ (E-SAA, E-CS and E-EV) as stated in Propositions 1 and 3 and by the definition of the E-SAA. Notice that the EV problem 'promises' an optimal cost which is misleading, since the true expected costs E-EV (431,096 euros) is worse than the 'promised' one (401,274 euros). The same applies for the optimal SAA and CS costs and the corresponding values E-SAA and E-CS. In the second row we have the variation relative to the CS value, the tightest bound. Notice that although in this case SAA is a lower bound for E-SAA, E-CS and E-EV, there is not a theoretical guarantee for it. Considering that in this problem CS $\leq$ RP $\leq$ (E-SAA, E-CS and E-EV), this variation can be interpreted as a worst-case optimality gap, which is (P $-$ CS) / CS = 1.01%, 1.02% and 5.25% for P = SAA, CS and EV, respectively. The EV lower bound is 2.03% worse than the CS counterpart. Finally, in the third row we have the CPU time (in seconds) for solving the MILP problems corresponding to the EV, CS and SAA approaches. To compute the E-SAA, E-CS and E-EV, we have used one million scenarios randomly sampled from $\mathbf{q}_2 \sim N_{50}(\mu, \Sigma)$ and the CPU time has been 190 seconds in all the cases. ■

6.5 Comparing the CS and the SAA performances

In the previous numerical example, it is clear that, the CS approach outperforms the EV approach in terms of obtaining a stronger lower bound and a good expected cost in a reasonable time (see Table 6). On the other hand, the CS approach, compared to the SAA approach with the same number of scenarios, has the advantage of providing a theoretical lower bound for the optimal RP value but, regarding the expected costs E-SAA and E-CS, the two methods obtain similar results (E-SAA = 413,724 and E-CS = 413,758, respectively).

The short CPU time of the EV problem is not surprising since it is a much smaller problem than the corresponding CS and SAA problems. However, the difference is surprising between the CS and SAA CPU times (0.8 and 93.8 seconds, respectively). In this section we further investigate this difference by solving the multi-farm feed manufactured problem for different numbers of crops I and farms $J = \lfloor I/2 \rfloor$ (largest integer that does not exceed $I/2$) reported in Table 7. The number of random parameters R and the number of the rows and columns of the constraint matrix of the corresponding MILP problems are also reported. Notice that instance 5 is the case analyzed in the previous sections.

The CS and SAA solutions are compared in Table 8, where the optimal value of problems CS, E-CS, SAA and E-SAA is reported. The worst case optimality gap of the E-CS and E-SAA values is computed by using the optimal CS value reported in the second column. The worst case optimality gap is similar for the two approaches (slightly better for the SAA approach).

The CS and SAA solution times are compared in Table 9 where, for each approach, the LP bound (obtained by solving LP relaxation of the corresponding MILP problem), the LP gap and the CPU time to solve the MILP problem are reported. In the last column the ratio of the two LP gaps: "SAA LP gap" / "CS LP gap" is reported. All the CS instances could be solved in less than three seconds. However, in general, the SAA instances required (much) more time to be solved to the point that the last four instances could not be solved within 3,600 seconds. In this case we only solved the corresponding LP relaxation, as reported in Table 9, which took a few seconds for each instance.

The SAA results in Tables 8 and 9 were obtained by drawing a random sample of scenarios for each instance. Therefore, the optimal value of the SAA problem is itself random. In order to analyze the variability of the SAA results, we solved the SAA instances described in Table 7 ten times each, by drawing a different random sample of scenarios. In Table 10 we report the resulting average and standard deviation for the optimal value and the CPU time. These results are similar to the ones reported in Tables 8 and 9 but with the additional variability information. Instances 7 to 10 could not be solved within 3600 seconds for any of the scenario samples.

A possible reason for the faster performance of the CS approach is its smaller LP gap. Although for the two approaches the LP gap is small, it happens that the SAA LP gap is several times the CS LP gap (see the last column in Table 9). Of course, the results here reported correspond to a very small test and therefore do not prove anything. They only show that the scenario selection may have some relevant impact on the corresponding MILP problem (LP gap and solution time). This question deserves further research.

Table 7: Size of the CS and SAA instances.

Instance	I	J	R	Scenarios	Rows	Columns
1	6	3	18	594	3,591	7,164
2	7	3	21	693	4,882	9,744
3	8	4	32	1,056	8,492	16,960
4	9	4	36	1,188	10,741	21,456
5	10	5	50	1,650	16,565	33,100
6	11	5	55	1,815	20,036	40,040
7	12	6	72	2,376	28,602	57,168
8	13	6	78	2,574	33,559	67,080
9	14	7	98	3,234	45,395	90,748
10	15	7	105	3,465	52,102	104,160

Table 8: Comparing the CS and SAA optimal values.

Instance	CS (euros)	E-CS (euros)	Worst case optimality gap (%)	SAA (euros)	E-SAA (euros)	Worst case optimality gap (%)
1	231,118	231,832	0.31	230,115	231,899	0.34
2	299,189	299,957	0.26	298,310	299,881	0.23
3	320,449	322,234	0.56	320,199	321,685	0.39
4	397,609	399,361	0.44	396,948	398,997	0.35
5	409,595	413,758	1.02	410,654	413,724	1.01
6	494,636	497,948	0.67	492,815	497,372	0.55
7	500,122	513,377	2.65	-	-	-
8	583,876	595,857	2.05	-	-	-
9	607,139	624,573	2.87	-	-	-
10	684,506	704,643	2.94	-	-	-

Table 9: Comparing the CS and SAA solution times.

Instance	CS problem (MILP)			SAA problem (MILP)			Ratio
	LP bound (euros)	LP gap (%)	CPU time (seconds)	LP bound (euros)	LP gap (%)	CPU time (seconds)	LP gaps
1	231,118	0.000081	0.1	230,103	0.005229	0.3	65
2	299,187	0.000713	0.2	298,310	0.000000	0.3	0
3	320,445	0.001387	0.4	320,178	0.006368	1.5	5
4	397,609	0.000298	0.4	396,928	0.005067	0.9	17
5	409,594	0.000279	0.8	410,315	0.082779	93.8	297
6	494,634	0.000456	0.9	492,739	0.015500	12.6	34
7	500,122	0.000000	1.3	505,185	-	$> 3{,}600$	-
8	583,875	0.000212	1.5	588,285	-	$> 3{,}600$	-
9	607,139	0.000000	2.2	615,554	-	$> 3{,}600$	-
10	684,506	0.000000	2.7	693,843	-	$> 3{,}600$	-

Table 10: **Average results for the SAA approach.**

Instance	Optimal value		CPU time	
	Average (euros)	Std. deviation (euros)	Average (seconds)	Std. deviation (seconds)
1	231,997	1,997	0.3	0.1
2	299,882	1,878	0.4	0.1
3	321,476	2,600	1.2	0.5
4	398,634	2,304	1.6	0.8
5	412,764	2,649	141.8	126.3
6	496,075	2,504	24.0	10.8
7	-	-	$> 3{,}600$	-
8	-	-	$> 3{,}600$	-
9	-	-	$> 3{,}600$	-
10	-	-	$> 3{,}600$	-

7 Conclusions

In this paper we have considered the two-stage stochastic mixed-integer linear programming problem with recourse which we have called the RP problem. We have also considered the EV problem as an approximation to the RP problem. The contributions of this paper have been: to introduce the CS problem, a new approximation to the RP problem based on conditional scenarios (a new concept here introduced) and to propose and analyze some useful bounds related to the RP, EV and CS problems. The RP problem is, in general, a good choice to deal with parameter uncertainty. However, if the RP problem results intractable, our suggestion is to use the CS problem, which requires a moderate computational effort and favorably compares to the EV problem in order to deal with parameter uncertainty.

From a practical point of view, the CS approach is appealing since the computation of the conditional scenarios is straightforward. It only requires the computation and discretization of a set of conditional expectations of the random vector that models the uncertain problem parameters. The number of conditional scenarios thus constructed, say N, grows linearly with the number of random parameters of the RP problem. In this way, the resulting CS problem has a moderate size (N times the size of the EV problem). Of course, there is a price to be paid for this: the optimal CS solution is, in general, suboptimal for the RP problem but hopefully better than the optimal EV solution.

In Section 6, we have also analyzed the performance of the RP problem with the same number of (randomly generated) scenarios as in the CS problem (we have called it the SAA problem). We have observed that the solution quality is similar for the CS and SAA approaches. However, we have also observed that the solution time is much shorter for the CS method as reported in Tables 9, where the CS instances show a smaller LP gap than the corresponding SAA instances (this could explain the shorter CS solution time). Of course, with the results reported in this paper we can only formulate the following conjecture which, in our opinion, deserves further research: *Conjecture 1: "The CS problem usually has a smaller LP gap than the corresponding RP problem with the same number of (randomly generated) scenarios".*

From a theoretical point of view, we have shown that in some cases the optimal RP cost can be bounded by the optimal CS and EV costs. In these cases, the CS bound dominates the EV bound. The reason is that the CS problem is a better approximation to the RP problem than the EV counterpart. Therefore, one would expect that the E-CS would dominate the E-EV *(Conjecture 2)*.

As a matter of further research, apart from trying to prove the previous two conjectures, we are planning to analyze the use of conditional scenarios in non-linear programming. Furthermore, as a first step, we have applied the conditional scenario concept in a risk-neutral model. That is, it is based on the expected cost and it does not incorporate any risk measure such as conditional value-at-risk [33] and stochastic dominance [17]. As a second step, we are planning to incorporate some risk measure into models based on conditional scenarios.

Acknowledgments: We are grateful to the two referees for their constructive and stimulating suggestions and to professors Jean-Philippe Vial and Alain Haurie for their comments and support at Logilab, University of Geneva, Switzerland. We are also grateful to the 'Comunidad de Madrid' (Spain) for financial support under grant S2009/esp-1594 (Riesgos CM) and to the Spanish Ministry of Economy and Competitiveness under grants MTM2012-36163-C06-06 and MTM2015-63710-P.

8 Appendix: Proofs of the theoretical results of Section 5

8.1 Proof of Proposition 1

To prove the inequality $z^*_{CS} \leq z^*_{RP}$ we proceed as follows. Let us consider $\tilde{x}^* = (\tilde{x}^*_1, \tilde{x}^*_2(\tilde{\boldsymbol{\xi}}))$, an optimal solution for the RP problem (1)–(7) with stochastic A_2 and/or b_2. Let us also assume that

z^*_{RP} is the corresponding optimal value. Now, based on this solution, we define a new point $\hat{x}^+ = (\hat{x}^+_1, \hat{x}^+_2(\hat{\boldsymbol{\xi}}^1), \dots, \hat{x}^+_2(\hat{\boldsymbol{\xi}}^R))$ such that $\hat{x}^+_1 = \tilde{x}^*_1$, $\hat{x}^+_2(\hat{\boldsymbol{\xi}}^r) = \mathbb{E}[\, \tilde{x}^*_2(\tilde{\boldsymbol{\xi}}) \mid \tilde{\boldsymbol{\xi}}_r \,]$, for all $r \in \mathcal{R}$ and with objective value $z_{CS}(\hat{x}^+)$. Notice that by definition $\hat{x}^+_2$ is a function of $\tilde{\boldsymbol{\xi}}_r$ such that we could write $\hat{x}^+_2(\tilde{\boldsymbol{\xi}}_r)$. However, we prefer to write $\hat{x}^+_2(\hat{\boldsymbol{\xi}}^r)$ to indicate that $\hat{x}^+_2$ is the policy associated with the conditional scenarios $\{\hat{\xi}^{re}\}_{e \in \mathcal{E}_r}$, the support of $\hat{\boldsymbol{\xi}}^r$. There is no ambiguity in using this notation since there is a one-to-one correspondence between the set of random components of $\tilde{\boldsymbol{\xi}}$, that is, $\{\tilde{\boldsymbol{\xi}}_r\}_{r\in\mathcal{R}}$ and the corresponding set of conditional expectations $\{\hat{\boldsymbol{\xi}}^r\}_{r\in\mathcal{R}}$. Therefore in this paper $\hat{x}_2(\hat{\boldsymbol{\xi}}^r)$ and $\hat{x}_2(\tilde{\boldsymbol{\xi}}_r)$ denote the same policy.

Let us see that $\hat{x}^+$ thus defined is feasible for the CS problem (24)–(30). On the one hand, $\hat{x}^+_1$ satisfies equation (25) by definition. On the other hand, we can see that $\hat{x}^+_2$ satisfies equations (26) and (28) as follows. We know that $\tilde{x}^*_2(\tilde{\boldsymbol{\xi}})$ fulfills constraints (3) and (5). Then, by applying the conditional expectation operator $\mathbb{E}[\, \cdot \mid \tilde{\boldsymbol{\xi}}_r \,]$ to these constraints one obtains the following aggregated constraints (further details on constraint aggregation by the conditional expectation operator can be found in [6]):

$$\begin{aligned} &\mathbb{E}[\, \tilde{\mathbf{A}}_2 \mid \tilde{\boldsymbol{\xi}}_r \,]\, \tilde{x}^*_1 + B_2\, \mathbb{E}[\, \tilde{x}^*_2(\tilde{\boldsymbol{\xi}}) \mid \tilde{\boldsymbol{\xi}}_r \,] = \mathbb{E}[\, \tilde{\mathbf{b}}_2 \mid \tilde{\boldsymbol{\xi}}_r \,] && \text{w.p.1},\ r \in \mathcal{R} \\ &\mathbb{E}[\, \tilde{x}^*_2(\tilde{\boldsymbol{\xi}}) \mid \tilde{\boldsymbol{\xi}}_r \,] \geq 0 && \text{w.p.1},\ r \in \mathcal{R}, \end{aligned}$$

or equivalently

$$\begin{aligned} &\hat{\mathbf{A}}^r_2 \hat{x}^+_1 + B_2 \hat{x}^+_2(\hat{\boldsymbol{\xi}}^r) = \hat{\mathbf{b}}^r_2 && \text{w.p.1},\ r \in \mathcal{R} \\ &\hat{x}^+_2(\hat{\boldsymbol{\xi}}^r) \geq 0 && \text{w.p.1},\ r \in \mathcal{R}. \end{aligned}$$

Thus, $\hat{x}^+$ fulfils constraints (26) and (28) and therefore $\hat{x}^+$ is feasible for the CS problem. Now, let us see that $z_{CS}(\hat{x}^+) = z^*_{RP}$.

$$\begin{aligned} z_{CS}(\hat{x}^+) &= c_1^\intercal \hat{x}^+_1 + \frac{1}{R} \sum_{r\in\mathcal{R}} c_2^\intercal\, \mathbb{E}[\, \hat{x}^+_2(\hat{\boldsymbol{\xi}}^r) \,] \\ &= c_1^\intercal \tilde{x}^*_1 + \frac{1}{R} \sum_{r\in\mathcal{R}} c_2^\intercal\, \mathbb{E}[\mathbb{E}[\, \tilde{x}^*_2(\tilde{\boldsymbol{\xi}}) \mid \tilde{\boldsymbol{\xi}}_r \,]] \qquad (71) \\ &= c_1^\intercal \tilde{x}^*_1 + \frac{1}{R} \sum_{r\in\mathcal{R}} c_2^\intercal\, \mathbb{E}[\, \tilde{x}^*_2(\tilde{\boldsymbol{\xi}}) \,] \\ &= c_1^\intercal \tilde{x}^*_1 + \mathbb{E}[\, c_2^\intercal\, \tilde{x}^*_2(\tilde{\boldsymbol{\xi}}) \,] \\ &= z^*_{RP}, \end{aligned}$$

where, in (71) we have used the law of total expectation [12, 19]. All in all, we have proved that given an optimal RP solution $\tilde{x}^*$, there exists a feasible CS solution $\hat{x}^+$ with objective value z^*_{RP} and this proves $z^*_{CS} \leq z^*_{RP}$.

The other inequality, $z^*_{EV} \leq z^*_{CS}$, can be proved as follows. First, we take

$$\hat{x}^* = (\hat{x}^*_1, \hat{x}^*_2(\hat{\boldsymbol{\xi}}^1), \dots, \hat{x}^*_2(\hat{\boldsymbol{\xi}}^R)),$$

an optimal CS solution with objective value z^*_{CS}. Then, we define R new points $\bar{x}^r = (\bar{x}^r_1, \bar{x}^r_2)$ such that $\bar{x}^r_1 = \hat{x}^*_1$ and $\bar{x}^r_2 = \mathbb{E}[\, \hat{x}^*_2(\hat{\boldsymbol{\xi}}^r) \,]$ for all $r \in \mathcal{R}$. It is easy to see that $\bar{x}^r$ is feasible for the EV problem, for all $r \in \mathcal{R}$ (it is enough to aggregate the CS constraints (26) and (28) by the expectation operator $\mathbb{E}[\,\cdot\,]$). Now, given that each $\bar{x}^r$ is feasible for the EV problem, one has:

$$z^*_{EV} \leq c_1^\intercal \bar{x}^r_1 + c_2^\intercal \bar{x}^r_2 = c_1^\intercal \bar{x}^r_1 + c_2^\intercal \mathbb{E}[\, \hat{x}^*_2(\hat{\boldsymbol{\xi}}^r) \,] \qquad r \in \mathcal{R}.$$

By adding these R inequalities one obtains:

$$R\, z^*_{EV} \leq R\, c_1^\intercal \bar{x}^r_1 + \sum_{r\in\mathcal{R}} c_2^\intercal \mathbb{E}[\, \hat{x}^*_2(\hat{\boldsymbol{\xi}}^r) \,],$$

which implies

$$z_{EV}^* \leq c_1^\intercal \bar{x}_1^r + \frac{1}{R} \sum_{r \in \mathcal{R}} \mathbb{E}[\, c_2^\intercal \hat{x}_2^*(\hat{\boldsymbol{\xi}}^r)\,] = z_{CS}^*,$$

as we wanted to prove. ■

Remark 3. As pointed out in Section 1, the CS approach can be seen as an aggregation method. More specifically, under the hypothesis of Proposition 1 (A_2 and/or b_2 stochastic and $\mathcal{J}_2 = \emptyset$), it can be seen that the CS problem is equivalent to

$$\begin{aligned}
\min_{\tilde{x}} \quad & z = c_1^\intercal \tilde{x}_1 + \mathbb{E}[\, c_2^\intercal \tilde{x}_2(\tilde{\boldsymbol{\xi}})\,] \\
\text{s.t.} \quad & A_1 \tilde{x}_1 = b_1 \\
& \mathbb{E}[\, \tilde{\mathbf{A}}_2 \mid \tilde{\boldsymbol{\xi}}_r\,]\tilde{x}_1 + B_2 \mathbb{E}[\, \tilde{x}_2(\tilde{\boldsymbol{\xi}}) \mid \tilde{\boldsymbol{\xi}}_r\,] = \mathbb{E}[\, \tilde{\mathbf{b}}_2 \mid \tilde{\boldsymbol{\xi}}_r\,] && \text{w.p.1}, \ r \in \mathcal{R} \qquad (72) \\
& \mathbb{E}[\, \tilde{x}_2(\tilde{\boldsymbol{\xi}}) \mid \tilde{\boldsymbol{\xi}}_r\,] \geq 0 && \text{w.p.1}, \ r \in \mathcal{R} \qquad (73) \\
& \tilde{x}_1 \geq 0 \\
& \tilde{x}_{1j} \text{ integer} && j \in \mathcal{J}_1,
\end{aligned}$$

where, as we already said, constraints (72) and (73) are obtained by aggregating constraints (3) and (5) by the conditional expectation operator $\mathbb{E}[\, \cdot \mid \tilde{\boldsymbol{\xi}}_r\,]$. Thus, for example in (72), given any realization $\tilde{\xi}_r$, one computes

$$\mathbb{E}[\, \tilde{\mathbf{b}}_2 \mid \tilde{\xi}_r\,] = \sum_{\tilde{b}_2 \in S_{\tilde{\mathbf{b}}_2}} \tilde{b}_2 \, g(\tilde{b}_2 \mid \tilde{\xi}_r),$$

where $S_{\tilde{\mathbf{b}}_2}$ is the support of $\tilde{\mathbf{b}}_2$ and the aggregation weights are given by $g(\cdot \mid \tilde{\xi}_r)$, the conditional probability function of $\tilde{\mathbf{b}}_2$ given $\tilde{\boldsymbol{\xi}}_r = \tilde{\xi}_r$.

8.2 Proof of Proposition 2

To prove the inequality $z_{CS}^* \leq z_{EV}^*$ we proceed as follows. Let us consider $\bar{x}^* = (\bar{x}_1^*, \bar{x}_2^*)$, an optimal solution for the EV problem (15)–(20) with objective value z_{EV}^*. Now, based on this solution, we define a new point $\hat{x}^+ = (\hat{x}_1^+, \hat{x}_2^+(\hat{\boldsymbol{\xi}}^1), \ldots, \hat{x}_2^+(\hat{\boldsymbol{\xi}}^R))$ such that $\hat{x}_1^+ = \bar{x}_1^*, \ \hat{x}_2^+(\hat{\boldsymbol{\xi}}^r) = \bar{x}_2^*$ for all $r \in \mathcal{R}$. Since A_2, B_2 and b_2 are deterministic, it is clear that $\hat{x}^+$ is feasible for the CS problem. Now let us see that $z_{CS}(\hat{x}^+) = z_{EV}^*$.

$$\begin{aligned}
z_{CS}(\hat{x}^+) &= c_1^\intercal \hat{x}_1^+ + \frac{1}{R} \sum_{r \in \mathcal{R}} \mathbb{E}[\, \hat{\mathbf{c}}_2^{r\,\intercal} \, \hat{x}_2^+(\hat{\boldsymbol{\xi}}^r)\,] \\
&= c_1^\intercal \bar{x}_1^* + \frac{1}{R} \sum_{r \in \mathcal{R}} \mathbb{E}[\, \hat{\mathbf{c}}_2^{r\,\intercal} \, \bar{x}_2^*\,] \\
&= c_1^\intercal \bar{x}_1^* + \frac{1}{R} \sum_{r \in \mathcal{R}} \bar{c}_2^\intercal \bar{x}_2^* && (74) \\
&= z_{EV}^*, && (75)
\end{aligned}$$

where, in (74) we have used that $\mathbb{E}[\, \hat{\mathbf{c}}_2^r\,] = \mathbb{E}[\mathbb{E}[\, \tilde{\mathbf{c}}_2 \mid \tilde{\boldsymbol{\xi}}_r\,]] = \mathbb{E}[\, \tilde{\mathbf{c}}_2\,] = \bar{c}_2$ (the second equality is justified by the law of total expectation). We have shown that given an optimal EV solution $\bar{x}^*$ there exists a feasible CS solution $\hat{x}^+$ with objective value z_{EV}^* and this proves $z_{CS}^* \leq z_{EV}^*$.

The other inequality, $z_{RP}^* \leq z_{CS}^*$, can be proved as follows. Take $\hat{x}^* = (\hat{x}_1^*, \hat{x}_2^*(\hat{\boldsymbol{\xi}}^1), \ldots, \hat{x}_2^*(\hat{\boldsymbol{\xi}}^R))$, an optimal CS solution with objective value z_{CS}^*. Then, define R new points $\tilde{x}^r = (\tilde{x}_1^r, \tilde{x}_2^r)$ such that

$\tilde{x}_1^r = \hat{x}_1^*$ and $\tilde{x}_2^r = \hat{x}_2^*(\hat{\boldsymbol{\xi}}^r)$, for all $r \in \mathcal{R}$. We also need the following result (based on the law of total expectation):

$$\begin{aligned}\mathbb{E}[\,\tilde{\mathbf{c}}_2^\intercal \hat{x}_2^*(\hat{\boldsymbol{\xi}}^r)\,] &= \mathbb{E}[\,\tilde{\mathbf{c}}_2^\intercal \hat{x}_2^*(\tilde{\boldsymbol{\xi}}_r)\,] = \mathbb{E}[\,\mathbb{E}[\,\tilde{\mathbf{c}}_2^\intercal \hat{x}_2^*(\tilde{\boldsymbol{\xi}}_r) \mid \tilde{\boldsymbol{\xi}}_r\,]\,] \\ &= \mathbb{E}[\,\mathbb{E}[\,\tilde{\mathbf{c}}_2^\intercal \mid \tilde{\boldsymbol{\xi}}_r\,]\,\hat{x}_2^*(\tilde{\boldsymbol{\xi}}_r)\,] = \mathbb{E}[\,\hat{\mathbf{c}}_2^{r\,\intercal}\,\hat{x}_2^*(\tilde{\boldsymbol{\xi}}_r)\,] = \mathbb{E}[\,\hat{\mathbf{c}}_2^{r\,\intercal}\,\hat{x}_2^*(\hat{\boldsymbol{\xi}}^r)\,], \end{aligned} \tag{76}$$

where we have used that $\hat{x}_2^*(\hat{\boldsymbol{\xi}}^r)$ can also be written as $\hat{x}_2^*(\tilde{\boldsymbol{\xi}}_r)$. It is easy to see that $\tilde{x}^r$ is feasible for the RP problem, for all $r \in \mathcal{R}$ (by hypothesis, only c_2 is stochastic). Therefore one can write:

$$\begin{aligned} z_{RP}^* &\le c_1^\intercal \tilde{x}_1^r + \mathbb{E}[\,\tilde{\mathbf{c}}_2^\intercal \tilde{x}_2^r\,] \\ &= c_1^\intercal \hat{x}_1^* + \mathbb{E}[\,\tilde{\mathbf{c}}_2^\intercal \hat{x}_2^*(\hat{\boldsymbol{\xi}}^r)\,] \\ &= c_1^\intercal \hat{x}_1^* + \mathbb{E}[\,\hat{\mathbf{c}}_2^{r\,\intercal} \hat{x}_2^*(\hat{\boldsymbol{\xi}}^r)\,] \qquad r \in \mathcal{R}, \end{aligned}$$

where we have used (76). Now, by adding these R inequalities one obtains:

$$R\, z_{RP}^* \le R\, c_1^\intercal \hat{x}_1^* + \sum_{r\in\mathcal{R}} \mathbb{E}[\,\hat{\mathbf{c}}_2^{r\,\intercal} \hat{x}_2^*(\hat{\boldsymbol{\xi}}^r)\,],$$

which implies

$$z_{RP}^* \le c_1^\intercal \hat{x}_1^* + \frac{1}{R}\sum_{r\in\mathcal{R}} \mathbb{E}[\,\hat{\mathbf{c}}_2^{r\,\intercal} \hat{x}_2^*(\hat{\boldsymbol{\xi}}^r)\,] = z_{CS}^*, \tag{77}$$

as we wanted to prove. ■

8.3 Proof of Proposition 3

Let us prove the left-hand side inequality first. The reason is that to obtain $z_{E\text{-}EV}^*$ we have to solve the RP problem with the additional constraint $\tilde{x}_1 = \bar{x}_1^*$, which tightens it. The analogous reason applies for the E-CS case. ■

8.4 Proof of Proposition 4

On the one hand, by Proposition 3 one has $z_{RP}^* \le z_{E\text{-}EV}^*$. Thus, to prove the left-hand side chain of inequalities it is enough to show that $z_{E\text{-}EV}^* \le z_{EV}^*$. To prove it, let us consider $\bar{x}^* = (\bar{x}_1^*, \bar{x}_2^*)$ an optimal solution for the EV problem (15)–(20). Let us also assume that z_{EV}^* is the corresponding optimal value. Now, based on this solution, we define a new point $\tilde{x}^+ = (\tilde{x}_1^+, \tilde{x}_2^+(\tilde{\boldsymbol{\xi}}))$ such that $\tilde{x}_1^+ = \bar{x}_1^*$, $\tilde{x}_2^+(\tilde{\boldsymbol{\xi}}) = \bar{x}_2^*$. It is easy to see that the point $\tilde{x}^+$ is feasible for the E-EV problem (the RP problem plus the additional constraint $\tilde{x}_1 = \bar{x}_1^*$). Now, let us see that $z_{E\text{-}EV}(\tilde{x}^+) = z_{EV}^*$.

$$\begin{aligned} z_{E\text{-}EV}(\tilde{x}^+) &= c_1^\intercal \tilde{x}_1^+ + \mathbb{E}[\,\tilde{\mathbf{c}}_2^\intercal\, \tilde{x}_2^+(\tilde{\boldsymbol{\xi}})\,] \\ &= c_1^\intercal \bar{x}_1^* + \mathbb{E}[\,\tilde{\mathbf{c}}_2^\intercal\, \bar{x}_2^*\,] \\ &= c_1^\intercal \bar{x}_1^* + \bar{c}_2^\intercal \bar{x}_2^* \\ &= z_{EV}^*. \end{aligned}$$

We have shown that given an optimal EV solution $\bar{x}^*$ there exists a feasible E-EV solution $\hat{x}^+$ with objective value z_{EV}^* and this proves $z_{E\text{-}EV}^* \le z_{EV}^*$.

On the other hand, also by Proposition 3, one has $z_{RP}^* \le z_{E\text{-}CS}^*$. Thus, to prove the right-hand side chain of inequalities it is enough to show that $z_{E\text{-}CS}^* \le z_{CS}^*$. This inequality can be proved in the way as the inequality $z_{RP}^* \le z_{CS}^*$ of Proposition 2. The only difference is that the E-CS problem corresponds to the RP problem plus the additional constraint $\tilde{x}_1 = \hat{x}_1^*$, and this constraint is also fulfilled by the new points $\tilde{x}^r = (\tilde{x}_1^r, \tilde{x}_2^r) = (\hat{x}_1^*, \hat{x}_2^*(\hat{\boldsymbol{\xi}}^r))$ for all $r \in \mathcal{R}$, defined in the proof of Proposition 2. ■

References

[1] W. van Ackooij. A comparison of four approaches from stochastic programming for large-scale unit-commitment. *EURO Journal on Computational Optimization*, pages 1–29, DOI 10.1007/s13675-015-0051-x, 2015.

[2] M. Albareda-Sambola, E. Fernandez, and F. Saldanha-da-Gama. The facility location problem with Bernoulli demands. *Omega*, 39(3):335–345, 2011.

[3] V. M. Albornoz, P. Benario, and M. E. Rojas. A two-stage stochastic integer programming model for a thermal power system expansion. *International Transactions in Operational Research*, 11(3):243–257, 2004.

[4] A. Alonso-Ayuso, L. F. Escudero, A. Garin, M. T. Ortuño, and G. Perez. On the product selection and plant dimensioning problem under uncertainty. *Omega*, 33(4):307–318, 2005.

[5] A. Alonso-Ayuso, L.F. Escudero, and M.T. Ortuño. BFC, a branch-and-fix coordination algorithmic framework for solving some types of stochastic pure and mixed 0-1 programs. *European Journal of Operational Research*, 151:503–519, 2003.

[6] C. Beltran-Royo, L. F. Escudero, J. F. Monge, and R. E. Rodriguez-Ravines. An effective heuristic for multistage linear programming with a stochastic right-hand side. *Computers and Operations Research*, 51:237–250, 2014.

[7] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust Optimization*. Princeton University Press, 2009.

[8] L. Bertazzi, A. Bosco, and D. Lagana. Managing stochastic demand in an inventory routing problem with transportation procurement. *Omega*, 56:112–121, 2015.

[9] J. R. Birge. Aggregation bounds in stochastic linear programming. *Mathematical Programming*, 31(1):25–41, 1985.

[10] J. R. Birge and F. Louveaux. *Introduction to Stochastic Programming*. Springer (New York), Second Edition, 2011.

[11] G. B. Dantzig. *Linear Programming and Extensions*. Princeton University Press, 1998.

[12] M. H. DeGroot and M. J. Schervish. *Probability and Statistics*. Addison Wesley (Boston, USA), Fourth Edition, 2012.

[13] D. Dentcheva. Optimization models with probabilistic constraints. In A. Shapiro, D. Dentcheva and A. Ruszczynski, editors, *Lectures on Stochastic Programming. Modeling and Theory*, volume 9 of *MPS-SIAM series on optimization*, pages 87–154. SIAM and MPS, Philadelphia, 2009.

[14] J. Dupacova, N. Growe-Kuska, and W. Romisch. Scenario reduction in stochastic programming: An approach using probability metrics. *Mathematical Programming*, 95(3):493–511, 2003.

[15] S. Engell, A. Markert, G. Sand, and Rudiger Schultz. Aggregated scheduling of a multiproduct batch plant by two-stage stochastic integer programming. *Optimization and Engineering*, 5(3):335–359, 2004.

[16] Y. M. Ermoliev, A. V Kryazhimskii, and A. Ruszczynski. Constraint aggregation principle in convex optimization. *Mathematical Programming*, 76(3):353–372, 1997.

[17] R. Gollmer, F. Neise, and R. Schultz. Stochastic programs with first-order dominance constraints induced by mixed-integer linear recourse. *SIAM Journal on Optimization*, 19(2):552–571, 2008.

[18] A. Hallefjord and S. Storoy. Aggregation and disaggregation in integer programming problems. *Operations Research*, 38(4):619–623, 1990.

[19] W. Hardle and L. Simar. *Applied Multivariate Statistical Analysis*. Springer, Third Edition, 2012.

[20] H. Heitsch and W. Romisch. A note on scenario reduction for two-stage stochastic programs. *Operations Research Letters*, 35(6):731–738, 2007.

[21] T. Homem-de-Mello and G. Bayraksan. Monte carlo sampling-based methods for stochastic optimization. *Surveys in Operations Research and Management Science*, 19(1):56–85, 2014.

[22] K. Hoyland, M. Kaut, and S. W. Wallace. A heuristic for moment-matching scenario generation. *Computational Optimization and Applications*, 24(2-3):169–185, 2003.

[23] IBM ILOG CPLEX Optimizer,
www-01.ibm.com/software/commerce/optimization/cplex-optimizer/.

[24] N. L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions, Vol. 1*. New York: John Wiley and Sons, Second Edition, 1994.

[25] P. Kall and J. Mayer. *Stochastic Linear Programming*. Kluwer academic publishers (New York, USA), 2005.

[26] A. J. King and S. W. Wallace. *Modeling with Stochastic Programming*. Springer Science and Business Media, 2012.

[27] A. J. Kleywegt, A. Shapiro, and T. Homem-de Mello. The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, 12(2):479–502, 2002.

[28] S. Nickel, F. Saldanha-da-Gama, and H.-P. Ziegler. A multi-stage stochastic supply network design problem with financial decisions and risk management. *Omega*, 40(5):511–524, 2012.

[29] G. Pages and J. Printems. Optimal quadratic quantization for numerics: the Gaussian case. *Monte Carlo Methods and Applications*, 9(2):135–165, 2003.

[30] A. Parisio and C. N. Jones. A two-stage stochastic programming approach to employee scheduling in retail outlets with uncertain demand. *Omega*, 53:97–103, 2015.

[31] G. Ch. Pflug and A. Pichler. Approximations for probability distributions and stochastic optimization problems. In *Stochastic Optimization Methods in Finance and Energy*, pages 343–387. Springer, 2011.

[32] A. Prekopa. *Stochastic Programming*. Springer (Dordrecht, The Netherlands), 1995.

[33] R. T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–42, 2000.

[34] D. F. Rogers, R. D. Plante, R. T. Wong, and J. R. Evans. Aggregation and disaggregation techniques and methodology in optimization. *Operations Research*, 39(4):553–582, 1991.

[35] N. V. Sahinidis. Optimization under uncertainty: state-of-the-art and opportunities. *Computers and Chemical Engineering*, 28(6):971–983, 2004.

[36] T. Santoso, S. Ahmed, M. Goetschalckx, and A. Shapiro. A stochastic programming approach for supply chain network design under uncertainty. *European Journal of Operational Research*, 167(1):96–115, 2005.

[37] R. Schultz and S. Tiedemann. Conditional value-at-risk in stochastic programs with mixed-integer recourse. *Mathematical Programming*, 105(2-3):365–386, 2006.

[38] S. Sen and H. D. Sherali. Decomposition with branch-and-cut approaches for two-stage stochastic mixed-integer programming. *Mathematical Programming*, 106(2):203–223, 2006.

[39] A. Shapiro, D. Dentcheva, and A. Ruszczynski. *Lectures on Stochastic Programming: Modeling and Theory*, volume 9. Society for Industrial Mathematics, 2009.

[40] Y. Song and J. Luedtke. An adaptive partition-based approach for solving two-stage stochastic programs with fixed recourse. *SIAM Journal on Optimization*, 25(3):1344–1367, 2015.

[41] S. Takriti and J. R. Birge. Lagrangian solution techniques and bounds for loosely coupled mixed-integer stochastic programs. *Operations Research*, 48(1):91–98, 2000.

[42] S. E. Wright. Primal-dual aggregation and disaggregation for stochastic linear programs. *Mathematics of Operations Research*, 19(4):893–908, 1994.

[43] P. H. Zipkin. Bounds for row-aggregation in linear programming. *Operations Research*, 28(4):903–916, 1980.

[44] P. H. Zipkin. Bounds on the effect of aggregating variables in linear programs. *Operations Research*, 28(2):403–418, 1980.