

On Multistage Stochastic Mixed 0–1 Optimization with Time-Consistent Stochastic Dominance Risk Averse Management

Laureano F. Escudero and Juan F. Monge

Abstract A new time-consistent risk averse measure is considered, so-called Expected Conditional Stochastic Dominance (ECSD), for multistage stochastic mixed 0–1 optimization, where first- and second-order stochastic dominance risk averse functionals are taken into account. As a result of the ECSD modeling, its problem solving is much more difficult than the Risk Neutral counterpart, so, it is unrealistic to solve the problem up to optimality by plain use of MIP solvers. Instead of it, decomposition algorithms of some type should be used. Computational results are reported for instances of a well-known real-life problem, where a decomposition heuristic algorithm is tested in its efficiency and computing effort, having the plain use of a MIP solver as a benchmark for computational purposes.

1 Introduction

Stochastic optimization is currently one of the most robust tools for decision making. It is broadly used in real-world applications in a wide range of problems from different areas in energy, finance, production, distribution, supply chain management, etc. The continuous optimization problems under uncertainty have been studied in [4] for risk neutral (RN) problems, and [16] for risk averse measures, among others. A survey on exact and inexact decomposition algorithms is performed in [9]. It is well known that an optimization (say, minimization) mixed 0–1 problem under uncertainty with a finite number of possible supporting scenarios has a mixed 0–1 Deterministic Equivalent Model (DEM). Traditionally, special attention has been given to optimizing the DEM by minimizing the objective function expected value

L. F. Escudero (✉)

Area Estadística e Investigación Operativa, Universidad Rey Juan Carlos,
28933 Móstoles (Madrid), Spain
e-mail: laureano.escudero@urjc.es

J. F. Monge

Centro de Investigación Operativa, Universidad Miguel Hernández,
03202 Elche (Alicante), Spain
e-mail: monge@umh.es

© Springer International Publishing AG 2018

E. Gil et al. (eds.), *The Mathematics of the Uncertain*, Studies in Systems, Decision and Control 142, https://doi.org/10.1007/978-3-319-73848-2_12

127

in the scenarios, subject to the satisfaction of all the problem constraints in the defined scenarios, i.e., the so-called risk neutral (RN) approach. Currently, we are able to solve huge DEMs by using different types of decomposition approaches, e.g., see in [1] our last improvement of the Branch-and-Fix Coordination (BFC) methodology. However, the optimization of the RN model has the inconvenience of providing a solution that ignores the potential variability of the objective function value (say, cost) in the scenarios and, so, it does not avoid the potential high negative impact of the proposed solution in the low-probability high-cost scenarios to occur.

However, there are some risk averse approaches that, additionally, deal with risk management, see a good survey in [2], among others. As we know, the first risk averse measure, so-called Chance Constraint (CC) functional, was introduced in the seminal work [5], where the problem's feasible set is restricted to satisfying an upper bound on the probability (in our case, the expected fraction of scenarios) of having shortfall on satisfying each of the modeler-driven constraints. It is also so-called first-order stochastic dominance (FSD) that for a two-stage setting was introduced in [12], where a Lagrangean-based decomposition algorithm was used for problem solving. Another of the first risk averse measures, so-called Integrated Chance Constraint (ICC) functional, was introduced in [14] and expanded in [15]. The ICC type 1 is also so-called second-order stochastic dominance (SSD) that for a two-stage setting was introduced in [11], where a Lagrangean-based decomposition algorithm was used for problem solving. The multistage risk averse time-inconsistent measure based on the SSD functional was introduced in [6] for a set of profiles related to given thresholds on a multifunction setting (including the objective function). As an extension of the two-stage SD mixed-integer linear recourse, FSD and SSD measures were jointly used in a multistage mixed 0–1 setting, so-called Time Stochastic Dominance (TSD), see our works [7, 9], where the decomposition algorithms use scenario clustering in the BFC scheme in the first work and a Lagrangean-based scheme in the other one.

The risk averse functional subject of this work, so-called Expected Conditional Stochastic Dominance (ECSD), is a time-consistent risk averse functional, since it bounds the scenario probability of having shortfall on satisfying different rhs for given constraints, and it also bounds the expected shortfall, both for the scenario groups with a one-to-one correspondence with nodes in the scenario tree related to a chosen set of stages in the time horizon. Roughly, a risk averse measure is time-consistent if the solution at any node in the scenario tree does not depend on the scenarios that cannot occur at that node. Obviously, RN is a time-consistent measure. Some other time-consistent measures, mainly the Expected Conditional Value-at-Risk, have been studied in [3, 13], among others.

The rest of the paper is organized as follows. For completeness and notation presentation, Sect. 2 deals with the main concepts of risk neutral-based mathematical optimization under uncertainty. Section 3 deals with the risk averse measure ECSD. Section 4 reports some results of our computational experiment. Section 5 concludes.

2 Risk Neutral Measure in Multistage Mixed 0–1 Stochastic Problems

The DEM, compact representation, of the multistage mixed 0–1 model for minimizing the objective function expected value in the scenarios (i.e., a RN approach) can be expressed

$$\begin{aligned}
 z_{RN} &= \min \sum_{g \in \mathcal{G}} w_g (a_g x_g + b_g y_g) \\
 \text{s.t.} \quad & \sum_{q \in \mathcal{A}_g} (A_q^g x_q + B_q^g y_q) = h_g \quad \forall g \in \mathcal{G} \\
 & x_g \in \{0, 1\}^{nx(g)}, \quad y_g \in \mathbb{R}^{ny(g)} \quad \forall g \in \mathcal{G},
 \end{aligned} \tag{1}$$

where \mathcal{G} is the set of nodes in the scenario tree, w_g is the weight computed for node $g \in \mathcal{G}$, such that $w_g = \sum_{\omega \in \Omega_g} w_\omega$, where w_ω is the modeler-driven probability of scenario ω , for $\omega \in \Omega$, where Ω is the finite set of scenarios considered in the supporting tree; x_g and y_g are the $nx(g)$ and $ny(g)$ -dimensioned vectors of the 0–1 variables and the continuous variables, resp., for node $g \in \mathcal{G}$; a_g and b_g are the vectors of the objective function coefficients for x_g and y_g , resp.; \mathcal{A}_g is the set included by the same node g and its ancestors in the scenario tree with nonzero coefficients in the constraints of node $g \in \mathcal{G}$; A_q^g and B_q^g are the constraint matrices; and h_g is the rhs for node $g \in \mathcal{G}$. All vectors and matrices are with the adequate dimensions. Notice that the non-anticipativity constraints (NAC) are implicitly satisfied.

Additionally, let the following notation to be used throughout this work. \mathcal{T} is the set of stages $\{1, 2, \dots, T\}$ in the time horizon with $T = |\mathcal{T}|$; $\mathcal{G}_t \subseteq \mathcal{G}$ is the set of nodes in stage $t \in \mathcal{T}$, where $|\mathcal{G}_t| = 1$. Let $\Omega_g \subset \Omega$ be the set of scenarios in group g , where it has a one-to-one correspondence with node g in the scenario tree, for $g \in \mathcal{G}$; for easing notation, let $\omega \equiv g$ for $g \in \mathcal{G}_T$ and, then, $\omega \in \Omega_g$, where $|\Omega_g| = 1$; $t(g)$ is the stage to which node g belongs to, such that $g \in \mathcal{G}_{t(g)}$; $\tilde{\mathcal{A}}_g$ is the set of ancestor nodes in the scenario tree to node g (including itself), for $g \in \mathcal{G}$ (such that $\mathcal{A}_g \subseteq \tilde{\mathcal{A}}_g$); and $\tilde{\mathcal{I}}_g$ is the set included by the same node g and its successors in the scenario tree, for $g \in \mathcal{G} \setminus \mathcal{G}_T$.

3 ECSD Risk Averse Measure in Multistage Mixed 0–1 Stochastic Problems

The RN model (1) aims to minimize the objective function expected value alone subject to the constraint system in the model. As stated above the main criticism that can be made to this very popular strategy is that it ignores the variability of the objective function value in the scenarios and, in particular, the “right” tail of the non-wanted scenarios. The Expected Conditional Stochastic Dominance (ECSD) risk averse measure also minimizes the objective function expected value but, besides the set of RN constraints, a set of profiles is considered for a set of functions (including the objective one) in a given scenario subset. Each profile is included by a threshold

on the function value for any scenario that belong to a given group (to be defined below) up to the end of the time horizon, a bound target on the fraction of those scenarios having excess on reaching it (the so-called first-order ECSD), a bound target on the expected excess (the so-called second-order ECSD) and a bound target on the excess for any of those scenarios. Observe that the type of risk reduction functional that is considered in this work also allows other functions besides the objective one, such as environmental functions in industrial strategic investment, mitigation effects of natural disasters, and preservation of cultural, monumental and strategic assets, among others; all of them can be accommodated in the framework presented below.

So, let the additional notation, \mathcal{F} is the set of indexes of the risk reduction oriented functions to consider, such that a_g^f and b_g^f are the vectors of the coefficients of the variables in x_g and y_g for $g \in \mathcal{G}$ in the function indexed with f , for $f \in \mathcal{F}$. Let us consider that $f = 1$ is the index for the objective function and, so, $a_g^1 \equiv a_g$ and $b_g^1 \equiv b_g$. Let also $\tilde{\mathcal{T}}^f \subseteq \mathcal{T} \setminus \{T\}$ be the modeler-driven stage set for function $f \in \mathcal{F}$, such that the risk reduction measure ECSD is performed for each group of scenarios Ω_g , for $g \in \mathcal{G}$, $t \in \tilde{\mathcal{T}}^f$. Let also \mathcal{P}_g^f denote the set of indexes of the related profiles for the function indexed with $f \in \mathcal{F}$ and the scenarios in set Ω_g , for $g \in \mathcal{G}$, $t \in \tilde{\mathcal{T}}^f$, such that the profile indexed with p , for $p \in \mathcal{P}_g^f$, is included by the 4-tuple $(\phi^p, \tilde{e}^p, \bar{e}^p, \bar{v}^p)$, where,

- ϕ^p , threshold for the value of function f up to the last stage T in the time horizon in scenario ω , for $\omega \in \Omega_g$.
- \tilde{e}^p , bound target on the excess of any of those scenarios on reaching the threshold.
- \bar{e}^p , bound target on the expected excess.
- \bar{v}^p , bound target on the expected fraction of those scenarios with excess.

Let the following additional variables for scenario ω and profile p , for $\omega \in \Omega_g$, $p \in \mathcal{P}_g^f$, $g \in \mathcal{G}$, $t \in \tilde{\mathcal{T}}^f$, $f \in \mathcal{F}$:

- e_ω^p , continuous variable that takes the excess of the value of function f on reaching threshold ϕ^p , for scenario ω .
- v_ω^p , 0–1 variable that takes the value 1 if the value of function f has an excess on reaching threshold ϕ^p and otherwise, 0, for scenario ω .

The ECSD model can be expressed

$$\begin{aligned}
 z_{ECSD} &= \min \sum_{g \in \mathcal{G}} w_g (a_g^1 x_g + b_g^1 y_g) \\
 &+ \sum_{f \in \mathcal{F}} \sum_{t \in \tilde{\mathcal{T}}^f} \sum_{g \in \mathcal{G}_t} \sum_{p \in \mathcal{P}_g^f} (M_{\epsilon_{\tilde{e}^p}} \epsilon_{\tilde{e}^p} + M_{\epsilon_{\bar{e}^p}} \epsilon_{\bar{e}^p} + M_{\epsilon_{\bar{v}^p}} \epsilon_{\bar{v}^p}) \\
 \text{s.t. } &\sum_{q \in \mathcal{A}_g^f} (A_q^g x_q + B_q^g y_q) = h_g \quad \forall g \in \mathcal{G} \\
 &x_g \in \{0, 1\}^{n_x(g)}, \quad y_g \in \mathbb{R}^{n_y(g)} \quad \forall g \in \mathcal{G} \\
 &\sum_{g \in \mathcal{A}_\omega^f} (a_q^f x_q + b_q^f y_q) - e_\omega^p \leq \phi^p \quad \forall \omega \in \Omega_g, p \in \mathcal{P}_g^f, g \in \mathcal{G}_t, t \in \tilde{\mathcal{T}}^f, f \in \mathcal{F} \\
 &0 \leq e_\omega^p \leq \tilde{e}^p v_\omega^p + \epsilon_{\tilde{e}^p},
 \end{aligned}$$

$$\begin{aligned}
 v_\omega^p &\in \{0, 1\}, \quad \epsilon_{\bar{z}^p} \in \mathbb{R}^+ && \forall \omega \in \Omega_g, \quad p \in \mathcal{P}_g^f, \quad g \in \mathcal{G}_t, \quad t \in \tilde{\mathcal{T}}^f, \quad f \in \mathcal{F} \\
 \sum_{\omega \in \Omega_g} w_\omega e_\omega^p &\leq \bar{e}^p + \epsilon_{\bar{e}^p}, \quad \epsilon_{\bar{e}^p} \in \mathbb{R}^+ && \forall p \in \mathcal{P}_g^f, \quad g \in \mathcal{G}_t, \quad t \in \tilde{\mathcal{T}}^f, \quad f \in \mathcal{F} \\
 \sum_{\omega \in \Omega_g} w_\omega v_\omega^p &\leq \bar{v}^p + \epsilon_{\bar{v}^p}, \quad \epsilon_{\bar{v}^p} \in \mathbb{R}^+ && \forall p \in \mathcal{P}_g^f, \quad g \in \mathcal{G}_t, \quad t \in \tilde{\mathcal{T}}^f, \quad f \in \mathcal{F},
 \end{aligned} \tag{2}$$

where the last terms in the objective function prevent the potential infeasibility of the risk averse constraint system, such that the big M -parameters are the related penalization and the ϵ -variables are the slack ones to avoid the infeasibility.

The concept of *expected conditional risk averse measure* (ECRM) is introduced in [13], and its time-consistency is defined and proved. We show elsewhere [10] that ECSD is a member of the family of ECRMs and, therefore, it is time consistent. Let us assume that the decisions in a given problem have been made up to node g , for $g \in \cup_{\{t \in \mathcal{T}: t < T\}} \mathcal{G}_t$, according to the solution obtained in the original model (2) solved at stage $t = 1$. Now, let the submodel solved at node g , such that it is supported by the subtree rooted with node g whose successor nodes are in set $\tilde{\mathcal{S}}_g$. Then, the rationale behind a time-consistent risk averse measure is that the solution value to be obtained in the submodel solved at stage $t(g)$ for node g should have the same value as in the original model solved at stage $t = 1$.

4 Some Computational Experience

Let the tactical supply chain planning (TSCP) problem presented in our work [10] to be the pilot case where to consider the performance of the RN and ECSD risk measures. Its deterministic version is based on a real-life case in the assembly sector. TSCP has a broad applicability, specifically, in sectors such as car, computer and domestic appliances manufacturing, among others. It is the case in which a company with multiple raw material suppliers, plants, products, tiers of production in the bill of material (BoM) and markets needs to satisfy a product demand vector over a given time horizon. The goal is to determine a raw material supplying plan and a master production, inventory and distribution planning that best makes use of the available resources and their capacity extension acquisitions in the whole supply chain for each period of a given time horizon. The resources' best use consists of minimizing the raw material supplying commitment cost, the production and inventory costs in the plants, and the product backlog and demand lost penalization along the time horizon. The raw material supplying commitment cost is frequently modeled by a piecewise linear, concave and nondecreasing function of the total volume to commit for the whole time horizon. Typical types of constraints (some of them related to either-or decisions) are as follows: Balance equations of end-products and components, conditional lower and upper bounds for raw material supplying and product release, resource consumption bounds and capacity extension

Table 1 STSCP Models' dimensions

ID	Scenario tree						RN model			ECSD model		
	Structure	T	t*	$ \mathcal{G}_t^* $	$ \mathcal{G} $	$ \Omega $	nc	n01	m	nc	n01	m
L1	2 ⁴ 2 ⁴	8	5	16	255	128	57064	18126	90364	57064	18126	90424
L2	2 ⁴ 2 ⁵	9	5	16	511	256	113768	36302	180604	113768	36302	180664
L3	2 ⁴ 2 ⁴	8	5	16	255	128	72424	18131	99324	72424	18131	99384
L4	2 ⁴ 2 ⁵	9	5	16	511	256	144488	36307	198524	144488	36307	198584

acquisitions, and balance equations of lost demand and backlogging, among others. There are different types of resources at different levels for groups of consecutive periods (so-called stages) along the time horizon. The cost of the resources' capacity extension acquisition is expressed as a piecewise discrete and nondecreasing function. Another important feature of the problem is that the burden of raw material stocking is frequently transferred to the suppliers.

The experiment was conducted on a PC with a 2.5 GHz dual-core Intel Core i5 processor, 8 Gb of RAM and the operating system was OS X 10.9, where the MIP solver to use is CPLEX v12.5 and its optimality tolerance is set up to 0.001. The decomposition algorithm to use is our matheuristic SDP-ECSD [10].

The problem's dimensions of the testbed under consideration are up to 9 stages, 20 end-products, 30 market centers each, 20 subassemblies, 40 raw materials, and 25 types of resources. Table 1 presents the structure of the scenario tree for each instance as well as the dimensions of the two stochastic formulations. The set of stages \mathcal{T} has been split in two parts for problem solving. The first column of the table is the identifier of the instance, and the second one gives the predefined structure $A_1^{B_1} A_2^{B_2}$ of the scenario tree, where A_i denotes the number of children that each node has in each stage in part i , and B_i is the number of its stages, for $i = 1, 2$. The period subset $\tilde{\mathcal{T}}$ is singleton and $t^* \in \tilde{\mathcal{T}}$, where t^* is the period defining the groups of scenarios for cost risk reduction in the ECSD measure. The headings of the columns for the dimensions of the models are as follows: nc , number of continuous variables; $n01$, number of 0–1 variables; and m , number of constraints.

The results of solving RN model (1) are shown in Table 2. The first column refers again to the identifier of the instance. The following three columns reports the CPLEX

Table 2 RN model (1) solved with CPLEX and the SDP-ECSD matheuristic

ID	CPLEX			SDP-ECSD		
	\bar{z}_{CPX}	t_{CPX}	OG%	\bar{z}_{RN}	t_{RN}	GG%
L1	358305	7200	0.62	362350	1409	1.13
L2	212922	7200	1.14	215236	3884	1.09
L3	218343	7200	0.72	221997	5032	1.67
L4	398303	7200	0.28	403833	8055	1.39

Table 3 ECSD model (2) solved with the SDP-ECSD matheuristic

ID	\bar{z}_{ECSD}	t_{ECSD}	$dev_{ECSD}\%$
L1	375127	21065	3.53
L2	215544	25032	0.14
L3	253152	22103	14.03
L4	431792	19403	6.92

results, where \bar{z}_{CPX} is the RN cost value, t_{CPX} is the elapsed time (in seconds) to obtain it, and $OG\%$ is its optimality gap (in %). The optimization of the instances reaches the allowed time (2h) without proving the 0.1%-optimality of the solution. Another block of columns in Table 2 reports the SDP-ECSD results. The headings are as follows: \bar{z}_{RN} and t_{RN} , RN solution value and related elapsed time (in seconds); and $GG\%$, goodness gap, i.e., the deviation of the solution value obtained by the matheuristic from the value obtained by CPLEX, expressed as $GG\% = (\bar{z}_{RN} - \bar{z}_{CPX})/\bar{z}_{CPX}\%$. We can observe that, generally, the elapsed time that is required by SDP-ECSD is very small. On the other hand, the goodness gap of its RN value versus the one provided by CPLEX is very small as well; notice that it goes from 1.09 to 1.39%.

The violations of the \bar{e}_{RN}^p and \bar{v}_{RN}^p ECSD bounds for $p = 1, 2$ by the RN solution are up to 165 and 700%, respectively. The details are not shown but they are available from the authors under request, see also [10].

The results of solving ECSD model (2) by matheuristic SDP-ECSD are shown in Table 3. Some headings are as follows: \bar{z}_{ECSD} and t_{ECSD} , ECSD incumbent value and related elapsed time (in seconds); and $dev_{ECSD}\%$, deviation of the ECSD cost from the RN one (see Table 2), expressed as $dev_{ECSD}\% = (\bar{z}_{ECSD} - \bar{z}_{RN})/\bar{z}_{RN}\%$.

Notice that the elapsed time required by the matheuristic for solving the ECSD model (2) is much greater than the time required for solving the RN model (1). It confirms the common knowledge, namely, the stochastic dominance strategy are computationally much harder than the RN one (requiring an elapsed time that is one order of magnitude higher than the time required for obtaining the RN solution). It is due to the cross scenario constraints for satisfying the risk reduction measure. Notice that, probably, CPLEX could not even solve the ECSD model, since it could not do it for the RN one. On the other hand, the deviation of the ECSD cost (due to the satisfaction of the cost risk reduction constraint system) could even reach the increment of the 14.04% of the RN cost (where, as notice above, the violations of the two types of risk reduction bounds are up to 165 and 700%).

5 Conclusions

Frequently there are problems with high variability in the functions to consider (beside the one to minimize). So, a risk reduction functional is required for avoiding low-probability high-negative function values incurred by the solutions obtained

from the minimization of the objective function expected (cost) value. In this work we have considered the time-consistent multi-function first- and second-order stochastic dominance functional ECSD for risk management to control. It allows to personalize the type of risk reduction profiles for groups of scenarios. That high risk management could increase the cost function value while satisfying the risk reduction constraint system. Any way, a specialization of decomposition algorithms is required for problem solving in a affordable computing effort. We have used our SDP-ECSD decomposition algorithm for problem solving in a pilot case from the tactical supply chain field; the solution's quality is good enough and the computing effort is very affordable.

References

1. Aldasoro U, Escudero LF, Merino M, Pérez G (2017) A parallel branch-and-fix coordination based matheuristic algorithm for solving large sized multistage stochastic mixed 0–1 problems. *Eur J Oper Res* 258:590–606
2. Alonso-Ayuso A, Carvallo F, Escudero LF, Guignard M, Pi J, Puranmalka R, Weintraub A (2014) On the optimization of copper extraction in mining under uncertainty in copper prices. *Eur J Oper Res* 233:711–726
3. Alonso-Ayuso A, Escudero LF, Guignard M, Weintraub A (2017) Risk management for forestry planning under uncertainty in demand and prices. *Eur J Oper Res*. <http://doi.org/10.1016/j.ejor.2017.12.22>
4. Birge JR, Louveaux FV (2011) *Introduction to stochastic programming*, 2nd edn. Springer, New York
5. Charnes A, Cooper WW, Symonds SH (1958) Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil. *Manag Sci* 4:235–263
6. Dentcheva D, Ruszczyński A (2003) Optimization with stochastic dominance constraints. *SIAM J Optim* 14:548–566
7. Escudero LF, Garín A, Merino M, Pérez G (2016) On time stochastic dominance induced by mixed integer-linear recourse in multistage stochastic programs. *Eur J Oper Res* 249:164–176
8. Escudero LF, Garín A, Monge JF, Unzueta A (2018) On multistage stochastic mixed 0-1 bilinear optimization based on endogenous uncertainty and time consistent stochastic dominance risk management. Submitted
9. Escudero LF, Garín A, Unzueta A (2017) Cluster Lagrangean decomposition for risk averse in multistage stochastic optimization. *Comput Oper Res* 85:154–171
10. Escudero LF, Monge JF, Romero-Morales D (2017) On time-consistent stochastic dominance risk averse measure for tactical supply chain planning under uncertainty. *Comput Oper Res*. <http://doi.org/10.1016/j.ejor.2017.07.11>
11. Gollmer R, Gotzes U, Schultz R (2011) A note on second-order stochastic dominance constraints induced by mixed-integer linear recourse. *Math Program Ser B* 126:179–190
12. Gollmer R, Neise F, Schultz R (2008) Stochastic programs with first-order stochastic dominance constraints induced by mixed-integer linear recourse. *SIAM J Optim* 19:552–571
13. Homem-de-Mello T, Pagnoncelli BK (2016) Risk aversion in multistage stochastic programming: a modeling and algorithmic perspective. *Eur J Oper Res* 249:188–199
14. Klein Haneveld WK (1986) *Duality in stochastic linear and dynamic programming*. Lecture notes in economics and mathematical systems, vol 274. Springer, Berlin
15. Klein Haneveld WK, van der Vlerk MH (2006) Integrated chance constraints: reduced forms and an algorithm. *Comput Manag Sci* 3:245–269
16. Pflug GCh, Römisch W (2007) *Modeling, measuring and managing risk*. World Scientific, Singapore