Medium range optimization of copper extraction planning under uncertainty in future copper prices *

Antonio Alonso-Ayuso
Dept. de Estadística e Investigación Operativa,
Universidad Rey Juan Carlos, Madrid, Spain

Felipe Carvallo
Dept. de Ingeniería Industrial, Universidad de Chile, Santiago, Chile

Laureano F. Escudero
Dept. de Estadística e Investigación Operativa,
Universidad Rey Juan Carlos, Madrid, Spain

Monique Guignard
Dept. of Operations and Information Management, The Wharton School
University of Pennsylvania, Philadelphia, PA, USA

Jiaxing Pi
Dept. of Electrical and Systems Engineering,
University of Pennsylvania, Philadelphia, PA, USA

Raghav Puranmalka
Dept. of Bioengineering
University of Pennsylvania, Philadelphia, PA, USA

Andres Weintraub
Dept. de Ingeniería Industrial, Universidad de Chile, Santiago, Chile

November 7, 2011

Abstract

Deterministic mine planning models along a time horizon have proved to be very effective in supporting decisions on sequencing the extraction of material in copper mines. Some of these models have been developed for, and used successfully by CODELCO, the Chilean state copper company. These models are extremely large. In this paper, we wish to consider the uncertainty in a very volatile parameter of the problem, namely, the copper price along a

*This research has been partially supported by the projects OPTIMOS2 MTM2009-14039-C06-03 (A. Alonso-Ayuso) and PLANIN MTM2009-14017-C04-01 (L.F. Escudero) from the Spanish Ministry of Science and Innovation, RIESGOS-CM from Comunidad de Madrid (A. Alonso-Ayuso and L.F. Escudero) from Spain, the National Science Foundation grant DMI0400155 (M. Guignard-Spielberg) from USA, and the Institute for Complex Enginerring Systems and Fondecyt, grant 100265 (A. Weintraub) from Chile.

given time horizon. We represent the uncertainty by a multistage scenario tree. The resulting stochastic model is then converted into a mixed 0-1 Deterministic Equivalent Model using a compact representation. We first introduce the stochastic model that maximizes the expected profit along the time horizon over all scenarios (i.e., as in a risk neutral environment). We then present several approaches for risk management, in a risk averse environment. Specifically, we consider the maximization of the Value-at-Risk and several variants of the Conditional Value-at-Risk, the maximization of the expected profit minus the weighted probability of having a "bad" scenario in the solution provided by the model, and the maximization of the expected profit subject to stochastic dominance constraints for a set of profiles given by the pairs of target profits and bounds on either the probability of not reaching them or the expected profit deficit over the targets. We present an extensive computational experience on the actual problem, by comparing the risk neutral approach, the tested risk averse strategies and the performance of the traditional deterministic approach that uses the expected value of the uncertain parameters. The results clearly show the advantage of using any risk neutral strategy over the traditional deterministic approach, as well as the advantage of using the risk averse strategy over the risk neutral one, although the plain use of the MIP solvers should be replaced by decomposition algorithms.
Keywords: Mining, planning models, copper price uncertainty, multistage stochastic mixed 0-1 optimization, scenario analysis, mixed 0-1 Deterministic Equivalent Model, risk measures, risk aversion, VaR, CVAR, mean-risk, stochastic dominance constraints.

1 Introduction

We consider a planning problem for a large underground mine, El Teniente. The mine has used a large scale deterministic MIP, with heuristics, to support long range planning, see [9, 15, 16]. The planning involves decisions which blocks to extract each year, which machinery to use, how to transport the material for grinding and processing into commercial copper, etc. The use of the deterministic model has been very successful, yielding an improvement in net present value (NPV) over a 25 year horizon of about 5%, over 100 million US $. In this paper we introduce explicitly the issue of uncertainty, which is of concern to the firm. There are two major sources of uncertainty in the firm's planning. One is derived from the market. Copper prices are highly volatile, especially lately, and hard to predict. This uncertainty is represented in the paper by a multistage scenario tree, and is treated by a scenario analysis scheme. The other main source of uncertainty is the grade (% of copper) in different parts of the mine. Prospecting can improve knowledge about the grade. In addition, as the extraction of the mine progresses, more information on grades becomes available, so we do not deal with that uncertainty in this paper.

In this work we consider a time horizon of 5 years, which is appropriate for making commercial decisions related to copper sales. The uncertainty is reflected through price scenarios, which define a possible price for each period (typically, a year) throughout the horizon. The decisions to be made concern the quantity of copper to extract in each period and the blocks from which it will be extracted. To solve the problem, we need to satisfy the nonanticipativity principle. The satisfaction of the related constraints is implicit in the Deterministic Equivalent Model (DEM) presented below in compact form, which is equivalent to the stochastic model.

The deterministic problem is already very large, as all 30 by 30 by 30 meter blocks are defined explicitly. The detailed mixed 0-1 DEM is therefore also very large, even considering only a 5 year horizon. Since on the one hand, in stochastic modeling the basic problem needs to be solved repeatedly and, on the other hand, decision making at this level does not require a high level of detail, we first reduce the size of the problem through an aggregation procedure. The aggregated model will then be solved for a set of scenarios that represent the uncertainty in copper prices. We will present a tight mixed 0-1 DEM, such that a state-of-the-art commercial solver can solve the instances for a moderate number of scenarios.

We introduce the stochastic model by maximizing the expected profit along the time horizon over all scenarios (i.e., as in in a risk neutral environment), subject to the satisfaction of all the problem constraints in all defined scenarios. We are able to solving very large tight DEMs (up to 800,000+ constraints, 274,000 0-1 variables and 1,900+ continuous variables) using a state-of-the-art MIP solver for cases with 77-scenarios in a reasonable amount of elapsed time; see Section 6. The expected profit is maximized but the DEM does not consider the risk of scenarios with bad consequences. Thus, in addition, we present several approaches for risk management in a risk averse environment, namely, (1) the maximization of the well known Value-at-Risk (VaR), (2) the maximization of several variants of the Conditional Value-at-Risk (CVaR), (3) the maximization of the mean-risk, i.e., the expected profit minus the weighted probability of having a "bad" scenario occurring for the given solution provided by the model, (4) the maximization of the expected profit subject to first-order stochastic dominance constraints (sdc) for a set of profiles given by the pairs of target profits and bounds on the probability of not reaching them, and (5) the maximization of the expected profit subject to second-order sdc whose set of profiles are given by the pairs of target profits and bounds on the expected profit deficit over the targets. An extensive computational experience on the real-life problem is presented by comparing the risk neutral approach, the above risk averse strategies and the performance of the traditional deterministic approach by considering the expected value of the uncertain parameters.

The remainder of the paper is organized as follows. In Section 2 we describe the mining problem, including the aggregation process. In Section 3, a mixed 0-1 optimization model is presented for the deterministic version of the problem, in which all parameters are assumed to be known in advance. In Section 4, the copper price uncertainties are described, the scenario analysis methodology that is used to deal with the uncertainty is presented and the model for the risk neutral environment is given. Section 5 presents the risk aversion that we choose to use, namely, VaR, several variants of CVaR, mean-risk and the sdc strategies. Section 6 reports on the computational experiment using a state-of-the-art MIP solver, with three illustrative large-scale 27-, 45- and 75-scenarios cases, and then compares the risk management strategies considered in the paper with the traditional deterministic strategy. Section 7 concludes, outlining our future work that, basically, will consist of using decomposition algorithms on the splitting variable representation of the problem, to permit the solution of larger instances, given the excessive computational effort required by the MIP engine of choice for solving large-scale instances of the stochastic MIP model even when using the compact representation.

2 The mining problem

We consider an underground mine located in Chile. Mining is carried out in several sectors of the mine El Teniente. The overall number of sectors is 18, but given our horizon of five years, only three of them are considered as active in the horizon. In the three active sectors we consider that extraction can only be carried out either on the two smaller sectors, or on the largest one. There is a cost associated with increasing or decreasing production in a sector from one period to the next.

In each sector there are a number of vertical columns, composed of blocks of 30 meters by 30 meters by 30 meters. The columns are neighboring, and are extracted in sequence. The columns numbers go from 70 to 2500, each with a basal area between 250m^2 and 400m^2 and a height between 549m and 959m. Thus, a column may consists of between 18 and 32 vertical blocks. The extraction method used is called block caving: At each drawpoint of a column, a void is created so that the rock breaks and falls due to gravity. The following specific rules must be respected in the mining process:

- The columns enter production in a specified sequence,
- The extraction of columns must have properties of neighborhood smoothness, implying that after extraction, neighboring columns cannot have much difference in height in what remains, and
- at each drawpoint there is a maximum extraction rate to prevent the roof from collapsing,

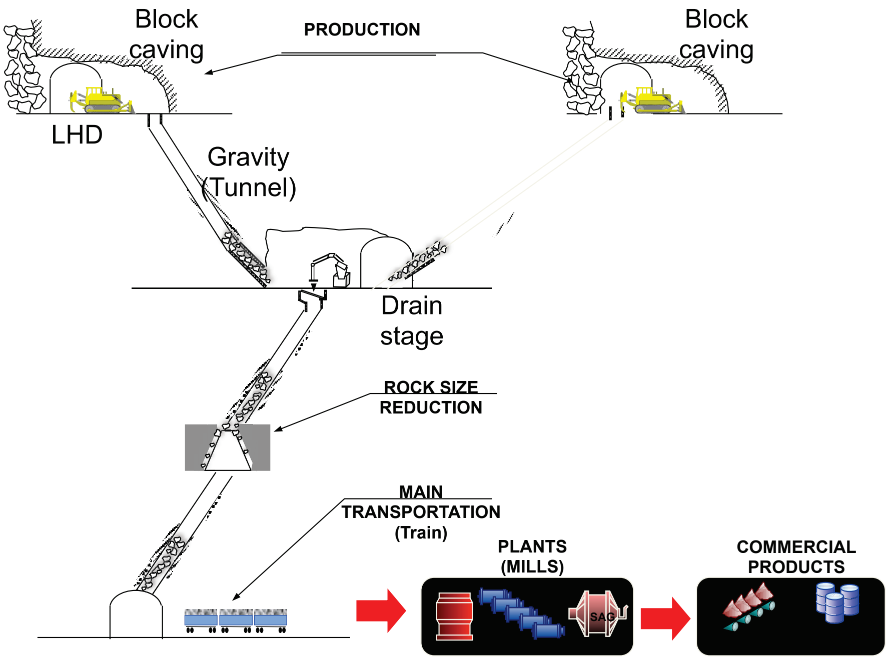


Figure 1: Underground mine with processing streams B and C in the Plants

as well as a minimum number of blocks to be extracted from each column to ensure a proper structure of the remaining mine.

Fig. 1 depicts a typical flow process, see [15]. In the block caving process, the broken rock is removed from the bottom of the columns and hauled using specialized machinery to a dumping point. There, through gravity rock is driven through a draining process and then the ore reaches a crusher where a rock reduction process is carried out. Finally, the crushed rock reaches a lower level, where it is sent via train to downstream processes. There are two processing streams in the plants, called B and C.

Decisions need to be made on when to cave in each column, when to move to the next column, and how far up in the column to extract. The rate of copper content tends to drop as we go up the column, so, depending on copper prices, it may be preferable at some point to drop the present column and move to the next one. In any case, once a column is dropped from production, it cannot be re-entered due to mechanical and stability issues. Another important decision consists of selecting which of the available sectors should be worked on. Then, downstream operations need to be integrated, including transportation, rock reductions, operations at mills and concentration plants until commercial copper is produced. The mine process can be represented in a network form, where the nodes represent specific activities and the arcs represent transportation, as shown in Fig. 2.

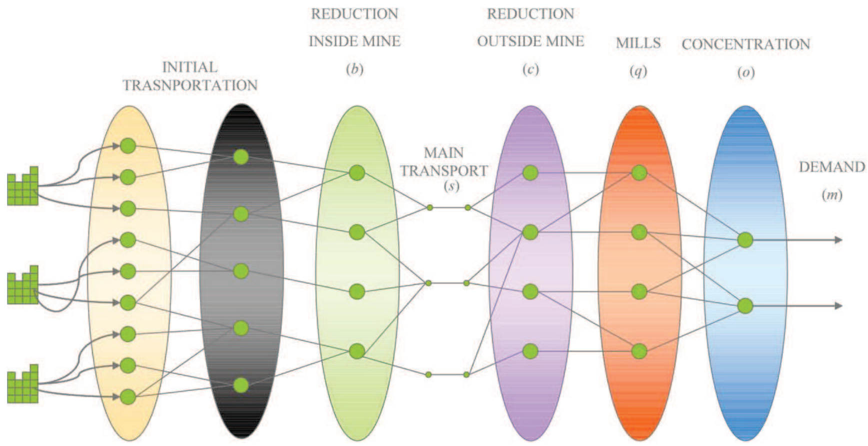


Figure 2: Network flow representation

An aggregation procedure was presented in [37] to reduce the size of the deterministic version of the problem. The aggregation was based on a cluster analysis [40], where the blocks of the original problem were aggregated based on spatial neighborhoods and similarities on the grade contents in copper and molibdenum, i.e., tons produced and extraction speed. The a priori aggregation process is carried out along the ideas presented in [38] for a forestry aggregation process. The aggregation process involves defining components which are of significance to obtain an aggregate value. In this form, weights are assigned to each component according to their significance for the final value. In the mining problem, the most important component was speed of extraction. The clustering method was a modification of the approach proposed in [25]. This aggregation process insures feasibility when the solution is again disaggregated. Fig. 3 show and example with 50 blocks disposed in 10 columns. The figure depicts an example of how the 50 blocks have been aggregated into 20 clusters. Note that a cluster can include blocks from different columns; for instance, cluster 16 consists of 4 blocks that form a J. It is worth pointing out that, due to the extraction method, precedence relations do exist. For example, clusters 20 and 16 need to be extracted before cluster 15, since extraction is performed by gravity. In special cases, clusters are linked so that they have to be extracted simultaneously. In the aggregated form, the smoothness of extracted columns is not preserved directly. When dis-aggregating the cluster solution smoothness is imposed again. We use the aggregated model in the stochastic version of the problem presented in this paper.

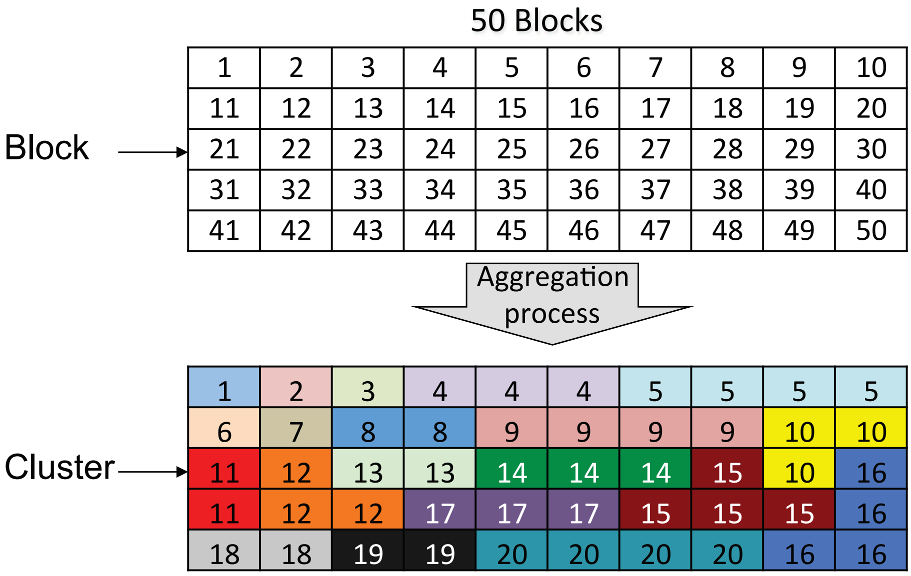


Figure 3: Clustering process

3 Model representation

We now present the deterministic model for the aggregated version of the problem, where some blocks have been aggregated into clusters. The notation to be used through the paper is as follows:

Sets

- $\mathcal{T}$, set of periods in the time horizon.
- $\mathcal{S}$, set of sectors, where $\mathcal{S} = \{ES, FW, NN\}$ and ES, FW, NN are the three given sectors.
- $\mathcal{K}^s$, set of clusters in sector s, for $s \in \mathcal{S}$.
- $\mathcal{T}^k$, set of periods when the cluster k could have been reached in the extraction process, for $k \in \mathcal{K}^s, s \in \mathcal{S}$.
- $\mathcal{P}^s$, set of subsets of clusters in sector s, such that each element $\mathcal{P}_i$ in $\mathcal{P}^s$ is a set of clusters that must be extracted simultaneously, for $s \in \mathcal{S}$. Note 1: $\mathcal{T}^k = \mathcal{T}^{k'}$, for $k, k' \in \mathcal{P}_i, \mathcal{P}_i \in \mathcal{P}^s, s \in \mathcal{S}$. Note 2: $\mathcal{P}^s$ defines a partition of $\mathcal{K}^s$, for $s \in \mathcal{S}$.
- $Pred_k$, set of predecessor clusters of cluster k, such that all clusters in $Pred_k$ must be extracted by the time cluster k is extracted, for $k \in \mathcal{K}^s, s \in \mathcal{S}$. Note 1: $Pred_k \subset \mathcal{K}^s$, for $k \in \mathcal{K}^s, s \in \mathcal{S}$. Note 2: $\mathcal{T}^k$ should be defined taking into account the precedence relationships defined by set $Pred_k$.

Parameters

- $percent^{\text{cu}}_k$ and $percent^{\text{mo}}_k$, percentage of copper and molybdenum in cluster k, respectively, for $k \in \mathcal{K}^s, s \in \mathcal{S}$.
- $area_k$, maximum area (in m^2) that can be mined in cluster k, for $k \in \mathcal{K}^s, s \in \mathcal{S}$.
- TON_k, number of tons of rock that can be processed in cluster k at any period (i.e., a year), for $k \in \mathcal{K}^s, s \in \mathcal{S}$.
- TON^{ini}_s, number of tons of rock that have been processed in sector s at the pre-initial period (i.e., period 0), for $s \in \mathcal{S}$.
- $\overline{TON}$, maximum quantity of tons of rock that can be processed per period.
- $\underline{TON}_{st}$, minimum number of tons of rock that can be processed, if any, in sector s at period t, for $s \in \mathcal{S}, t \in \mathcal{T}$.
- $\overline{TON}^+_{st}$ and $\overline{TON}^-_{st}$, maximum increase and decrease of tons of rock processed in sector s from period t-1 to period t, respectively, for $s \in \mathcal{S}, t \in \mathcal{T}$.
- $\overline{area}_s$ and $\underline{area}_s$, maximum and minimum active area (in m^2) in sector s, respectively, for $s \in \mathcal{S}$.
- $\overline{TON}^{\text{B}}_t$, maximum number of tons of rock that can be processed at stream B at period t, for $t \in \mathcal{T}$.
- $discount_t$, discount factor in period t for prices and costs, for $t \in \mathcal{T}$.
- $price^{\text{cu}}_t$ and $price^{\text{mo}}_t$, copper price and molybdenum price per ton at period t, respectively, for $t \in \mathcal{T}$.
- $cost^{\text{m}}_{st}$, cost per unit of mining in sector s at period t, for $s \in \mathcal{S}, t \in \mathcal{T}$.
- $cost^{\text{a}}_{st}$, cost per unit of area in sector s at period t, for $s \in \mathcal{S}, t \in \mathcal{T}$.
- $cost^+_{st}$ and $cost^-_{st}$, cost per unit of production increase and decrease from period t-1 to period t, respectively, for $s \in \mathcal{S}, t \in \mathcal{T}$.
- $cost^{\text{B}}_t$ and $cost^{\text{C}}_t$, unit ton cost for processing streams B and C at period t, respectively, for $t \in \mathcal{T}$.

0-1 variables

z_{kt} takes the value 1 if cluster k is extracted in period t and 0 otherwise, for $t \in \mathcal{T}^k, k \in \mathcal{K}^s, s \in \mathcal{S}$.

x_s takes the value 1 if sector s is extracted and 0 otherwise, for $s \in \mathcal{S}$.

Continuous variables

ton_{st}, number of tons of rock extracted in sector s at period t, for $s \in \mathcal{S}, t \in \mathcal{T}$.

ton^+_{st} **and** ton^-_{st}, increase and decrease in the number of tons extracted in sector s at period t, for $s \in \mathcal{S}, t \in \mathcal{T}$.

ton^{B}_t **and** ton^{C}_t**,** number of tons sent to process in period t in processing streams B and C, respectively, for $t \in \mathcal{T}$.

Objective function

$$\max \sum_{t \in \mathcal{T}} discount_t \Bigg[\sum_{s \in \mathcal{S}} \sum_{k \in \mathcal{K}^s} TON_k \Big(price^{\mathrm{cu}}_t \, percent^{\mathrm{cu}}_k + price^{\mathrm{mo}}_t \, percent^{\mathrm{mo}}_k \Big) z_{kt} - \sum_{s \in \mathcal{S}} \Big(cost^{\mathrm{m}}_{st} \, ton_{st} + cost^{\mathrm{a}}_{st} \sum_{k \in \mathcal{K}^s} area_k \, z_{kt} + cost^+_{st} \, ton^+_{st} + cost^-_{st} \, ton^-_{st} \Big) - cost^{\mathrm{B}}_t \, ton^{\mathrm{B}}_t - cost^{\mathrm{C}}_t \, ton^{\mathrm{C}}_t \Bigg]. \tag{1}$$

(1) maximizes the NPV (net present value) of the total profit along the time horizon. The profit for each period includes the income from selling the extracted copper and molibdenum, reduced by the mining and sector costs, the cost of production increase and decrease from one period to the next one, and processing costs.

Constraints

$$x_{ES} + x_{FW} \leq 1 \tag{2}$$

$$x_{ES} + x_{NN} \leq 1 \tag{3}$$

$$\sum_{t \in \mathcal{T}^k} z_{kt} \leq x_s, \quad \forall k \in \mathcal{K}^s, s \in \mathcal{S} \tag{4}$$

$$\sum_{t' \in \mathcal{T}^k : t' \leq t} z_{kt'} \leq \sum_{t' \in \mathcal{T}^k : t' \leq t} z_{jt'} \quad \forall j \in Pred_k, t \in \mathcal{T}^k, k \in \mathcal{K}^s, s \in \mathcal{S} \tag{5}$$

$$z_{kt} = z_{k't} \quad \forall t \in \mathcal{T}^k, k, k' \in \mathcal{P}_i, \mathcal{P}_i \in \mathcal{P}^s, s \in \mathcal{S} \tag{6}$$

$$ton_{st} = \sum_{k \in \mathcal{K}^s : t \in \mathcal{T}^k} TON_k \, z_{kt} \quad \forall s \in \mathcal{S}, t \in \mathcal{T} \tag{7}$$

$$ton_t^{\mathrm{B}} + ton_t^{\mathrm{C}} = \sum_{s \in \mathcal{S}} ton_{st} \quad \forall t \in \mathcal{T} \tag{8}$$

$$ton_{st}^{+} - ton_{st}^{-} = \begin{cases} ton_{st} - TON_s^{\mathrm{ini}}, & \text{if } t = 1 \\ ton_{st} - ton_{s,t-1}, & \text{if } t > 1 \end{cases} \quad \forall s \in \mathcal{S}, t \in \mathcal{T} \tag{9}$$

$$\underline{area}_s x_s \leq \sum_{k \in \mathcal{K}^s} area_k \sum_{t \in \mathcal{T}^k} z_{kt} \leq \overline{area}_s x_s \quad \forall s \in \mathcal{S} \tag{10}$$

$$\sum_{s \in \mathcal{S}} ton_{st} \leq \overline{TON} \quad \forall t \in \mathcal{T} \tag{11}$$

$$\underline{TON}_{st} x_s \leq ton_{st} \quad \forall s \in \mathcal{S}, t \in \mathcal{T} \tag{12}$$

$$0 \leq ton_t^{\mathrm{B}} \leq \overline{TON}_t^{\mathrm{B}} \quad \forall t \in \mathcal{T} \tag{13}$$

$$0 \leq ton_{st}^{+} \leq \overline{TON}_{st}^{+} x_s \quad \forall s \in \mathcal{S}, t \in \mathcal{T} \tag{14}$$

$$0 \leq ton_{st}^{-} \leq \overline{TON}_{st}^{-} x_s \quad \forall s \in \mathcal{S}, t \in \mathcal{T} - \{1\} \tag{15}$$

$$0 \leq ton_t^{\mathrm{C}} \quad \forall t \in \mathcal{T} \tag{16}$$

$$z_{kt} \in \{0, 1\} \quad \forall t \in \mathcal{T}^k, k \in \mathcal{K}^s, s \in \mathcal{S} \tag{17}$$

$$x_s \in \{0, 1\} \quad \forall s \in \mathcal{S}. \tag{18}$$

Constraints (2) and (3) ensure that if sector EN is selected, none of the sectors FW and NN are selected, and vice-versa. Constraints (4) guarantee that no cluster is processed in an unselected sector and, additionally, they ensure that each cluster is processed at most once. There are different alternatives for modeling precedence relationships, but constraints (5), that guarantee that if a cluster is processed at a given period then all predecessor clusters are also processed *by* that period, have given very good results in other contexts, see [1, 2, 4, 7], since it results in a stronger model. Our results confirm those other results. Notice that this type of constraints does not force a cluster to be processed if any of its predecessors is processed. Additionally, if a predecessor cluster has not been processed, then the given cluster cannot be processed yet. Constraints (6) force the clusters in set $\mathcal{P}_i$ to be extracted simultaneously in each sector. Constraints (7) evaluate the number of tons processed in each sector at each period. Constraints (8) are the flow conservation constraint for the processing stream. Constraint (9) calculates the increase and decrease in the number of tons processed in each period, respectively. Constraints (10) impose upper and lower bounds for the total area processed in each sector. Constraints (11)-(12) impose bounds on the number of tons processed in each period. Constraints (13) impose an upper bound due to the capacity of processing stream B. Constraints (14) and (15) bound the maximum increase and decrease of tons in each sector in each period.

Figure 4: Historical data for copper prices

4 Uncertainty in the (future) prices of copper

The deterministic model assumes that prices are known in advance to the planning decision. However, as we can see in Fig. 4, copper prices can vary along the planning horizon. Notice the volatility of the uncertain parameters and, so, they are very difficult to predict.

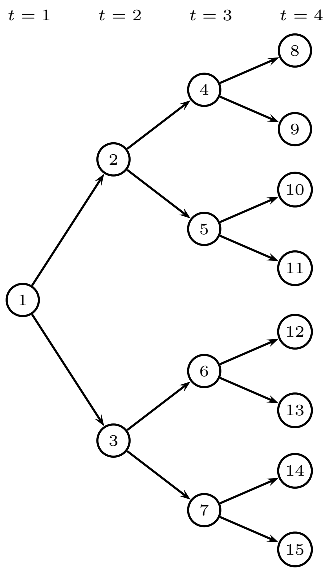
There are several ways in which to express future uncertainties in copper prices. One the most used consists of representing it by considering scenarios with known or estimated probabilities. A well known approach is the mean returning random walks, where prices are assumed to vary randomly from one period to the next, eventually returning to a given mean future value. In this case, the number of scenarios to express a proper range of prices may grow quite rapidly. Real options is another form, see e.g., [36], which has the weakness of not being able to integrate the approach within a reasonably detailed description of the mining problem.

Let us consider the following deterministic problem

$$\begin{aligned} \max \quad & ax + cy \\ \text{s.t.} \quad & Ax + By = b \\ & x \in \{0,1\}^n, y \geq 0, \end{aligned} \tag{19}$$

where m, n and n_c are the number of constraints, and 0-1 and continuous variables, respectively, a and c are n- and n_c-dimensional objective function coefficient vectors, respectively; b is the m-dimensional right-hand-side (rhs) of the constraint system; A and B are $m \times n$ and $m \times n_c$ constraint matrices, respectively; and x and y are the n-vector of 0-1 variables and the n_c-vector of continuous variables along the set $\mathcal{T}$ of periods, respectively.

For representing the uncertainty, we use a scenario tree approach in which uncertainty is modeled in terms of a set of scenarios. See e.g., [3, 5, 20] for symmetric scenario trees, among



$$\Omega = \Omega^1 = \{8, \ldots, 15\};\ \Omega^2 = \{8, \ldots, 11\}$$
$$\mathcal{G}^2 = \{2, 3\};\ \gamma(4) = 2;\ \mathcal{N}_7 = \{1, 3\}$$

Figure 5: A multistage scenario tree

many others, and [21] for nonsymmetric ones. For this purpose we need the following definitions. A *stage* of a time horizon is a set of one or various periods (in our case, years) in which the random parameters are realized; a *scenario* is the realization of uncertain parameters during the stages of the time horizon; and a *scenario group* for a given stage is the group of scenarios with the same uncertain parameter realizations up to that stage. (That is, a scenario group defines a group of partial scenarios).

To illustrate the multistage scenario tree concept, let Fig. 5 depict a symmetric scenario tree in which each node represents a stage, where a decision can be taken. Once a decision has been taken, various possible situations may occur. In our example there are two situations in stage $t = 2$. This information is generally presented in the form of a tree in which each path from the root to a leaf represents a scenario and corresponds to the realization of the entire set of uncertain parameters. For example, path $\{1, 3, 6, 12\}$ represents one scenario, and it is customary to call it scenario 12. In what follows, we do not distinguish between a scenario (or a group) and the corresponding node on the tree (with the same number). Each node in the tree must be associated with a scenario group in such a manner that any two scenarios belong to the same group (i.e., they have the same partial scenario) in a given stage if they include the same occurrences of uncertain parameters up to that stage. In this case, the well known nonanticipativity principle applies. It was stated in [39] and restated in [33]; see also [8], among others. This principle ensures that the solution at each t does not depend on information that is yet unavailable and requires that the decisions pertaining to scenarios in the same group (i.e., partial scenarios with the same value in the parameters) be the same. For example, for stage 3, scenarios 12 and 13 belong to the same group associated with path $\{1, 3, 6\}$, i.e., with group $g = 6$. Notice the difference between a scenario (a path from the root node to a leaf node) and a partial scenario (a path from the root to an intermediate node).

The notation for the scenario tree to be used in the paper is as follows:

$\mathcal{T}$, set of stages in the time horizon $1, 2, ..., T$.

$\mathcal{T}^-$, set of all stages except the last one.

Ω, set of scenarios.

$\mathcal{G}$, set of scenario groups.

$\mathcal{G}^t$, set of scenario groups in stage t ($\mathcal{G}^t \subseteq \mathcal{G}$), for $t \in \mathcal{T}$.

Ω^g, set of scenarios in group g ($\Omega^g \subseteq \Omega$), for $g \in \mathcal{G}$.

$\sigma(g)$, immediate ancestor node of node g, for $g \in \mathcal{G}$.

$\mathcal{N}_g$, set of ancestor groups (i.e, nodes) to group g, including itself.

If the parameters of vector c (i.e, copper prices) in problem (19) are random parameters with a set of discrete occurrences, say, c^ω over the set Ω of scenarios $\omega \in \Omega$, we will model our problem for maximizing the expected profit over the scenarios as follows,

$$\begin{array}{lll} \max & \sum_{\omega \in \Omega} w^\omega (a x^\omega + c^\omega y^\omega) & \\ \text{s.t.} & A x^\omega + B y^\omega = b & \forall \omega \in \Omega \\ & (x, y) \in \mathcal{N} & \\ & x^\omega \in \{0,1\}^n, y^\omega \geq 0 & \forall \omega \in \Omega, \end{array} \tag{20}$$

where w^ω is a positive weight assigned to scenario ω, for instance its probability ($\sum_{\omega \in \Omega} w^\omega = 1$); x^ω and y^ω represent the x and y variables for scenario ω, respectively. $\mathcal{N}$ The nonanticipativity set is defined by

$$\mathcal{N} = \{v | v_t^\omega = v_t^{\omega'}, \forall \omega, \omega' \in \Omega^g, g \in \mathcal{G}^t, t \in \mathcal{T}^-\}, \tag{21}$$

where $v = (x, y)$ and v_t^ω is such that $v^\omega = (v_t^\omega, \forall t \in \mathcal{T})$. Upon incorporating the set (21) in model (20), we can obtain the related multistage Deterministic Equivalent Model (*DEM*). The nonanticipativity set can be represented implicitely, through the variable definition (compact formulation) or explicitely, including new constraints in the model (sppliting variable representation), see [18], among others. For the purpose of the paper, we will only consider the compact representation since it is to be solved by a commercial MIP solver, and sppliting variable representation is more focussed on decomposition methods.

For simplifying the presentation of the model, let us assume that variables may have nonzero coefficients only in the constraints related to two consecutive stages (in our case, yearly periods). The aim consists of maximizing the expected value (i.e., mean) of the objective function (risk neutral). So, the model is as follows,

$$
\begin{aligned}
Q_E = \max \quad & \sum_{g \in \mathcal{G}} w_g (a_g x_g + c_g y_g) & \\
\text{s.t.} \quad & A' x_{\sigma(g)} + A x_g + B' y_{\sigma(g)} + B y_g = b & \forall g \in \mathcal{G} \\
& x_g \in \{0,1\}^{n_t}, y_g \geq 0 & \forall g \in \mathcal{G},
\end{aligned} \tag{22}
$$

where $w_g = \sum_{\omega \in \Omega^g} w^\omega$ gives the weight assigned to scenario group g, n_t is the number of x and y variables at stage t, and a_g and c_g are the counterparts of parameters a and c^ω related to scenario group g, for $g \in \mathcal{G}$, such that the values of the parameters for each scenario in the group are identical. Additionally, x_g and y_g represent the x and y variables for scenario group g, respectively, and A' and B' are the constraint matrices in scenario group g for the x and y variables related to the immediate ancestor of group g. Notice that in our case, the values of the vectors a_g are the same for the scenario groups $\{g\}$ that belong to the same stage t, i.e, the groups in $\mathcal{G}^t$.

It is out of the scope of this work to present a methodology for generating multistage scenario trees generation and reduction; see e.g., [13, 26, 27] and references therein. However, as an illustrative instance, Fig. 6 depicts an ad-hoc generated multistage 27-scenario tree for the copper prices during the years 2006 to 2010. All scenarios clearly have the same price, 2576 US$/ton for stage 1 (i.e., year 2006), then prices go up or down. For stage 5 (i.e., year 2010) the range is from 705 to 6316 US$/ ton, see the statistical historical information in [11]. Results of the computational experience for cases with 27-, 45- and 75 scenario trees are reported in Section 6.

5 Risk aversion management

The model that we have considered in the previous section aims to maximize the objective function expected value alone and, then, a so called risk neutral strategy is considered. The main criticism that can be made about this very popular mean strategy is that it ignores the variance on the objective function value over the scenarios and, in particular, the "left" tail of the non-wanted scenarios. However, there are some other approaches that additionally deal with risk management in an averse approach by considering, e.g., scenario immunization, see [12] and its treatment in

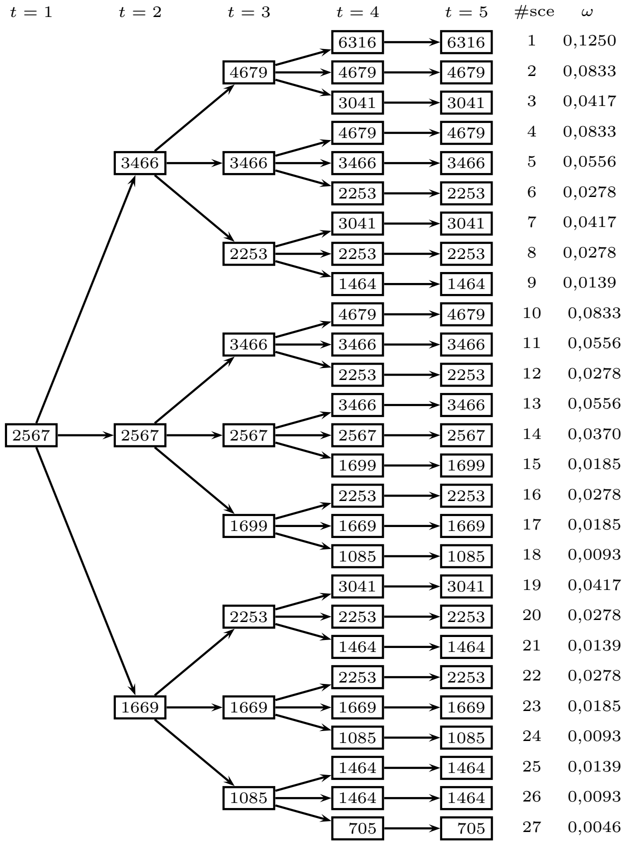


Figure 6: Copper prices (US$ per Ton) in the 27-scenario tree instance

[17], semi-deviations [31], value-and-risk and conditional value-at-risk [32, 35], excess probabilities [34], and first- and second-order stochastic dominance constraints strategies [23, 24] and therein references, among others. In this section we present a modification of model (22) to allow the risk aversion environment.

So, we consider the following risk averse measures, that take into account the bad tail of the profit distribution over the scenarios:

- Value-at-Risk: Well known theoretical research suggests that the measures based on quantiles are good functions for risk management. Among them, the Valued-at-Risk (VaR) has turned into a reference to many applications in the financial, transportation and productions planning sectors, among others. That approach is very attractive since it is easy to interpret. By definition, the β-VaR of an objective function over a set of scenarios is its lowest value, say α, such that the objective function value of the scenario to occur is over α with β probability (that is provided by the modeler).

- Conditional Value-at-Risk, [32, 35]: The advantage of the VaR strategy over the traditional maxmin strategy is obvious, since it takes into account a bound $1-\beta$ on the probability of the occurrence of a scenario whose profit is below α. However, it does not consider how bad the scenarios with a profit below VaR can be. The β-Conditional Value-at-Risk (β-CVaR) strategy takes into account the profit of the bad scenarios for a $\frac{1}{\beta}$ weighting parameter, where CVaR is the conditional expectation of profit below that amount α.

- Deficit Probability: As an alternative to the VaR and CVaR strategies, a risk measure is proposed in [34] for weighting the probability that a non-desired scenario will occur, that is, a scenario where the profit for the given solution of the model is below a given threshold, say ϕ, that is provided by the modeler.

- Stochastic dominance constraints strategies (sdc), see [23, 24] and the references therein, where the expected objective function value is optimized (in our case, maximized), such that a set of thresholds of the objective function value for each scenario is intended to be satisfied, with either a bound on the failure's probability for each threshold (so called first-order sdc) or a bound on the expected deficit on reaching each threshold (so called second-order sdc). Note: A profile is said to be included by a threshold and a bound of either its failure probability or its expected deficit and, in any case, it is provided by the modeler.

Some of these risk measures, and many other approaches in the literature, try to reduce the probability of the occurrence of non-wanted scenarios or the maximization of the solution value for the worst scenario with a given failure's probability, but they do not pay attention to the best scenarios (except the last strategy depending on the number of profiles). On the contrary, decision makers usually look for a trade-off between the risk minimization and the profit maximization. For this reason, the risk measures are usually combined with the optimization of the objective function value, leading to strategies as the combination of the Expected Value and Deficit Probability [34] and the combination of Expected Value and CVaR [35], among others. In the following, we will present how these approaches are modeled in a multistage stochastic mixed 0-1 program by including some new 0-1 variables and/or constraints. See [6] in a different context for an application in finance.

5.1 $\mathrm{Q_e}$ & VaR: Value-at-Risk strategy

The model that maximizes a combination of the Expected objective function Value and the β-VaR can be represented as follows,

$$
\begin{array}{lll}
\max & \gamma \sum_{g \in \mathcal{G}} w_g(a_g x_g + c_g y_g) + \rho\alpha & \\
\text{s.t.} & A' x_{\sigma(g)} + A x_g + B' y_{\sigma(g)} + B y_g = b & \forall g \in \mathcal{G} \\
& \sum_{g \in \mathcal{N}_d} (a_g x_g + c_g y_g) + M^\omega \nu^\omega \geq \alpha & \forall \omega \in \Omega \\
& \sum_{\omega \in \Omega} w^\omega \nu^\omega \leq 1 - \beta & \\
& x_g \in \{0,1\}^{n_t}, y_g \geq 0 & \forall g \in \mathcal{G} \\
& \nu^\omega \in \{0,1\} & \forall \omega \in \Omega \\
& \alpha \in \mathbb{R}, &
\end{array}
\tag{23}
$$

such that ω is the unique scenario in set Ω^d where $d \in \mathcal{G}^T$, α is a rational variable, ν^ω is a 0-1 variable, such that its value is 1 if the objective function value (i.e., the profit) for scenario ω is smaller than α and, 0 otherwise, M^ω is the "big M" parameter, preferably to be the smallest one, which does not eliminate any feasible solution of the stochastic program under scenario ω, ρ is a weighting factor and $\gamma \in \{0,1\}$ is another parameter such that for $\gamma = 0$ and $\rho = 1$ it results the classical VaR objective function. Remember that $\mathcal{N}_d$ gives the set of ancestor nodes (i.e., scenario groups) in the back path from leaf node d to root node 1. Note: The $1-\beta$ failure's probability of not satisfying a given constraint may have its roots in the concept of Chance Constraints introduced in [10].

5.2 VaR & overCVaR: Conditional expectation above VaR strategy

As it is state at the beginning of Section 5, the advantage of the VaR strategy over the traditional maxmin strategy is obvious since it takes into account a bound on the probability of the occurrence of scenarios whose profit is below VaR. However, it does not consider how bad the scenarios with a profit above VaR can be. On the contrary, the so named VaR & overCVaR strategy maximizes a combination of the the VaR and the weighted Conditional expectation *above* VaR, such that the model is as follows,

$$
\begin{array}{lll}
\max & \alpha + \rho \sum_{d \in \mathcal{G}^T} w_d \Big(\sum_{g \in \mathcal{G}_d} w_g (a_g x_g + c_g y_g) - \alpha \Big)_+ & \\
\text{s.t.} & A' x_{\sigma(g)} + A x_g + B' y_{\sigma(g)} + B y_g = b & \forall g \in \mathcal{G} \\
& \sum_{g \in \mathcal{N}_d} (a_g x_g + c_g y_g) + M^\omega \nu^\omega \geq \alpha & \forall \omega \in \Omega \\
& \sum_{\omega \in \Omega} w^\omega \nu^\omega \leq 1 - \beta & \\
& x_g \in \{0,1\}^{n_t}, y_g \geq 0 & \forall g \in \mathcal{G} \\
& \nu^\omega \in \{0,1\} & \forall \omega \in \Omega.
\end{array}
\tag{24}
$$

5.3 $\mathrm{Q_e}$ & CVaR: Conditional expectation below VaR strategy

As an alternative to the Conditional expectation above VaR strategy, see model (24), the so named $\mathrm{Q_e}$ & CVaR strategy takes into account the conditional expectation of the profit *below* VaR. The model that maximizes a combination of the Expected objective function Value and the β-CVaR can be expressed as follows,

$$
\begin{array}{lll}
\max & \gamma \sum_{g \in \mathcal{G}} w_g (a_g x_g + c_g y_g) + \rho \Big(\alpha - \frac{1}{\beta} \sum_{d \in \mathcal{G}^T} w_d \, sup \Big(\alpha - \sum_{g \in \mathcal{N}_d} (a_g x_g + c_g y_g) \Big)_+ \Big) & \\
\text{s.t.} & A' x_{\sigma(g)} + A x_g + B' y_{\sigma(g)} + B y_g = b & \forall g \in \mathcal{G} \\
& x_g \in \{0,1\}^{n_t}, y_g \geq 0 & \forall g \in \mathcal{G} \\
& \alpha \in \mathbb{R} &
\end{array}
\tag{25}
$$

where $z_+ = \max\{0, z\}$. Note: For $\gamma = 0$ and $\rho = 1$ it results the CVaR strategy introduced in [32].

A more amenable representation of model (25) is as follows, see [35],

$$
\begin{array}{lll}
\max & \gamma \sum_{g \in \mathcal{G}} w_g (a_g x_g + c_g y_g) + \rho \Big(\alpha - \frac{1}{\beta} \sum_{\omega \in \Omega} w^\omega v^\omega \Big) & \\
\text{s.t.} & A' x_{\sigma(g)} + A x_g + B' y_{\sigma(g)} + B y_g = b & \forall g \in \mathcal{G} \\
& \alpha - \sum_{g \in \mathcal{N}_d} (a_g x_g + c_g y_g) \leq v^\omega & \forall \omega \in \Omega \\
& x_g \in \{0,1\}^{n_t}, y_g \geq 0 & \forall g \in \mathcal{G} \\
& v^\omega \geq 0 & \forall \omega \in \Omega \\
& \alpha \in \mathbb{R}, &
\end{array}
\tag{26}
$$

such that v^ω is a non-negative variable equal to the difference (if its positive) between α and the profit for scenario ω, so named negative profit over VaR α.

5.4 $\mathrm{Q_e}$ & DP: Deficit probability

Let the model (27) that maximizes the Expected objective function Value minus the weighted probability of the scenario to occur whose objective function value is below a given threshold, see [34]. As in the VaR strategy, a new 0-1 variable per scenario, say ν^ω is needed, such that its value is 1 if the objective function value for scenario ω is smaller than a threshold, say ϕ and, 0 otherwise.

$$\begin{aligned}
&\max \textstyle\sum_{g\in\mathcal{G}} w_g(a_g x_g + c_g y_g) - \ \rho \sum_{\omega\in\Omega} w^\omega \nu^\omega \\
&\text{s.t.} A' x_{\sigma(g)} + A x_g + B' y_{\sigma(g)} + B y_g = b \qquad \forall g \in \mathcal{G} \\
&\quad \sum_{g\in\mathcal{N}_d} (a_g x_g + c_g y_g) + M^\omega \nu^\omega \geq \phi \qquad \forall \omega \in \Omega \\
&\quad x_g \in \{0,1\}^{n_t}, y_g \geq 0 \qquad \forall g \in \mathcal{G} \\
&\quad \nu^\omega \in \{0,1\} \qquad \forall \omega \in \Omega.
\end{aligned} \tag{27}$$

Notice that the threshold satisfaction constraints allow to impose a lower limit A^ω in the objective function value per scenario, just by fixing $M^\omega = \phi - A^\omega$. For example, $M^\omega = \phi$ means that scenarios with negative objective function values are not allowed.

5.5 Stochastic dominance constraints strategies

As an alternative to the previous strategy, let the recent approaches based on the first-order and second-order stochastic dominance constraints (sdc), see [23] and [24], respectively. The concept of sdc aims to identify acceptable feasible solutions so that the strategy optimizes over them. "A random variable X is said to be stochastically greater in first order, respectively second order, than a random variable Y, i.e., $X \succeq_1 Y$, respectively $X \succeq_2 Y$, iff $Eh(X) \geq Eh(Y)$ for all nondecreasing, respectively nondecreasing convex, functions h for which both expectations exists", see [23], respectively [24].

Let us implement the first-order stochastic dominance constraints strategy (so named sdc-1) by proposing the following model (28), such that the function value $\sum_{g\in\mathcal{N}_d}(a_g x_g + c_g y_g)$ (i.e., the profit of the unique scenario ω in set Ω^d where $d \in \mathcal{G}^T$) is not below the threshold ϕ^p with probability, say β^p, for $p \in \mathcal{P}$, where $\mathcal{P}$ is the set of profiles under consideration.

$$
\begin{array}{lll}
\max & \sum_{g \in \mathcal{G}} w_g(a_g x_g + c_g y_g) & \\
\text{s.t.} & A' x_{\sigma(g)} + A x_g + B' y_{\sigma(g)} + B y_g = b & \forall g \in \mathcal{G} \\
& \sum_{g \in \mathcal{N}_d} (a_g x_g + c_g y_g) + M^\omega \nu^{\omega p} \geq \phi^p & \forall \omega \in \Omega, p \in \mathcal{P} \\
& \sum_{\omega \in \Omega} w^\omega \nu^{\omega p} \leq 1 - \beta^p & \forall p \in \mathcal{P} \\
& x_g \in \{0,1\}^{n_t}, y_g \geq 0 & \forall g \in \mathcal{G} \\
& \nu^{\omega p} \in \{0,1\} & \forall \omega \in \Omega, p \in \mathcal{P},
\end{array}
\tag{28}
$$

where $\nu^{\omega p}$ is a 0-1 variable such that its value is 1 if the objective function value for scenario ω is smaller than threshold ϕ^p and 0, otherwise.

The second-order stochastic dominance constraints strategy (so named sdc-2) requires a set of profiles given by the pairs $(\phi^p, e^p) \, \forall p \in \mathcal{P}$, where e^p is the upper bound of the expected deficit of the profit over the scenarios on reaching the threshold ϕ^p. It is implemented by the model,

$$
\begin{array}{lll}
\max & \sum_{g \in \mathcal{G}} w_g(a_g x_g + c_g y_g) & \\
\text{s.t.} & A' x_{\sigma(g)} + A x_g + B' y_{\sigma(g)} + B y_g = b & \forall g \in \mathcal{G} \\
& \phi^p - \sum_{g \in \mathcal{N}_d} (a_g x_g + c_g y_g) \leq v^{\omega p} & \forall \omega \in \Omega, p \in \mathcal{P} \\
& \sum_{\omega \in \Omega} w^\omega v^{\omega p} \leq e^p & \forall p \in \mathcal{P} \\
& x_g \in \{0,1\}^{n_t}, y_g \geq 0 & \forall g \in \mathcal{G} \\
& v^{\omega p} \geq 0 & \forall \omega \in \Omega, p \in \mathcal{P},
\end{array}
\tag{29}
$$

such that $v^{\omega p}$ is a non-negative variable equal to the difference (if it is positive) between the threshold ϕ^p and the profit for scenario ω, so named negative profit over threshold ϕ^p. Notice that this strategy does not require additional 0-1 variables. The concept of the expected deficit of the profit on reaching a given threshold may have its roots in the Integrated Chance Constraints concept introduced in [29], see also [14, 30].

Solution considerations: We must point out that the models (23), (24), (28) and (29) have a computational disadvantage when comparing them with the models (26) and (27), since they have constraints linking 0-1 variables from different scenarios. Notice that the disadvantage is stronger for the models (28) and (29) with $|\mathcal{P}| > 1$ than for the models (23) and (24). In any case, a decomposition approach must be used for problem solving of very large instances. A Lagrange relaxation can be proposed for dualizing those linking constraints as done in the strategy presented

Table 1: Model dimensions

	m	$n01$	nc	nel	den
Deterministic	10708	6613	55	47699	0.067
27-scen tree risk neutral DEM	288272	100974	737	1208798	0.004
45-scen tree risk neutral DEM	480490	167951	823	2651860	< 0.004
75-scen tree risk neutral DEM	800694	274813	1992	3628801	< 0.002

in [24] for the second-order sdc. Moreover, see in Section 7 the outline of our future research work on this subject.

6 Computational experience

The models presented in this work have been optimized by using the following HW/SW platform: 2 quad-core Xeon E5450 3.0Ghz 64-bit processors with 6Mb of cache each and 64Gb (8 x 8Gb) of 667MHz fully buffered RAM memory, GAMS 23.6 [22] as a modeler system and CPLEX v12.2 [28] as the optimizer. It uses realistic data from a copper extraction real-life planning.

The instances that are considered have three sectors for copper extraction with 2100, 664 and 2640 clusters, respectively, and a time horizon of 5 years (from 2006 to 2010). We have performed two different studies: First, we have compared the solution obtained by the deterministic model versus the solution obtained by the stochastic one, where the risk neutral strategy is used. In a second step, we have analyzed the impact of considering risk adverse measures in the model . We use three illustrative cases related to 27-, 45- and 75-scenario trees, respectively. Note that we use a discrete distribution for representing the uncertainty, while VaR and CVaR were initially proposed for continuous distributions. Therefore, the weight of each scenario for small trees can be so big that only a small number of scenarios can represent risky situations and, then, the VaR and CVaR values can not be representative. Notice that changes in only one scenario can cause great changes in the value of these risk measures. So, for this reason, among other, a higher number of scenarios is more interesting than a smaller one.

Table 1 shows the dimensions of the deterministic model based on the expected value of the uncertain parameters and the risk neutral DEM (22). The headings are as follows: m, number of constraints; $n01$, number of 0-1 variables; nc, number of continuous variables; nel, number of nonzero elements in the constraint matrix; and den, matrix density (in %). The dimensions are quite large for the DEM instances.

Table 2: 27-scen tree. Stochastic solution values. Risk neutral model (22)

	Max GAP	
	0.01%	0.50%
LP relaxation solution value	354.72	354.72
Solution value	354.67	354.18
Optimality GAP (%)	0,01	0,15
Greatest scen. solution value	990.53	991.02
Median	308.70	307.81
0.90-VaR	-10.53	-11.85
0.95-VaR	-38.90	-39.10
0.90-CVaR	-39.57	-39.77
Smallest scen. solution value	-52.19	-52.34
Weight of scenarios with negative profit	0,157	0,157
Conditional expected negative profit	-25.50	-25.89
CPU time (secs.)	70973	241

6.1 Risk neutral approach: Deterministic versus stochastic models

Let us start with the stochastic model (22) applied to the 27-scenario tree case depicted in Fig. 6 for the copper prices uncertainty. This model has been solved by using two different values for the termination criteria, namely, quasi-optimality gap bounds GAP= 0.01% and 0.5%. Table 2 shows the main statistics of the solution obtained in both cases. Notice that a slightly better solution has been obtained for the 0.01% maximum gap, but 20 hours of computation were required, being the gap 0.02% after 2000 secs. of elapsed time. Given the solution's quality and the required elapsed time, we have considered reasonable to set up the maximum gap to 0.5%. Notice that this gap only refers to the profit (here, the objective function to maximize) so named the solution value in the table.

Let EV (Expected Value) denote the traditional deterministic model where the uncertain parameters have been replaced by their expected values, EEV be the Expected result of the Expected Value, obtained by applying over the scenarios the EV solution, and WS (Wait-and-See) be the average over all scenarios of the optimal solution value for each independent scenario (which is an upper bound of the solution value for the original model), such that the WS solution usually does not satisfy the nonanticipativity constraints since they have been relaxed. The methodology for obtaining the EEV is very well established in the two-stage environment, see [8], but it is a difficult one for multistage problems, see [19]. Alternatively, we propose the following

Table 3: 27-scen tree. Deterministic solutions. WS and EEV

	WS	EEV
Solution value	404,19	352,18
Optimality GAP (%)	0,02	0,02
Greatest scen solution value	1031,72	986,39
Median	314,72	310,38
0.90-VaR	71,74	-19,20
0.95-VaR	39,73	-51,74
0.90-CVaR	44,33	-53,53
Smallest scen solution value	28,88	-66,67
Weight of scenarios with negative profit	0,000	0,157
Conditional expected negative profit	0,00	-37,41
CPU time	207	80

methodology for obtaining EEV in a rolling horizon type of calculation (see [1] for more details): (1) The solution for the first stage is obtained from the EV solution, (2) Once the solution up to a stage t is fixed, $|\mathcal{G}^t|$ independent scenario subtrees remain, (3) The EV solution is independently obtained for the scenario subtrees, whose root nodes are the scenario groups in $\mathcal{G}^t$, so that the solution for each scenario group is fixed to its EV solution, (4) The procedure continues until the last stage, where the mixed 0-1 two-stage problem for each related scenario group is solved. So, at the end of the process there is a solution for each scenario and EEV is obtained by weighting the solution values for the scenarios as calculated by the procedure. Table 3 shows the main results related to the WS and EEV solutions. We can observe that the EEV solution value (in our case, the expected profit) is only a 0.6% smaller than the one provided by the risk neutral model (22), but it has worse VaR, CVaR, and conditional expected negative profit than most of the alternative approaches presented above, even the risk neutral maximization of the expected profit over the scenarios along the time horizon.

Therefore, the stochastic model, even without including risk averse measures, slightly allows to obtain a better solution in terms of the expected profit, but with much smaller risk of bad scenarios; notice that the VaR and CVaR are much better.

6.2 Risk averse measures

In this section the results for the 27-scenarios tree instance are reported by solving the different risk averse models presented in Section 5. A computational comparison with the risk neutral

Table 4: 27-scen tree. Stochastic solution values. Q_e & VaR model (23)

ρ	Z_{LP}	Z_{MIP}	$GAP(\%)$	β	$\%\ Q_E$	VaR	CVaR	$P(<0)$	$E(<0)$	t (secs)
0	354.72	354.18	0.15	1.00	100.00	-11.85	-39.77	0.157	-25.88	241
0.5	382.78	343.26	0.05	0.95	97.76	23.11	-0.06	0.069	-8.60	3122
1	424.99	351.26	0.05	0.95	91.21	49.07	28.45	0.000	0.00	2448
5	855.81	478.30	0.06	0.95	87.89	54.70	33.91	0.000	0.00	2342
10	1414.96	646.76	0.04	0.95	87.20	55.06	34.27	0.000	0.00	3190
100	11540.70	4257.68	0.12	0.95	65.54	61.27	37.70	0.000	0.00	3174
10000	1125725.59	402648.97	0.16	0.95	65.60	61.28	37.14	0.000	0.00	5570
0.5	406.36	361.25	0.45	0.90	96.14	41.44	3.37	0.056	-8.28	4256
1	471.34	385.44	0.03	0.90	94.44	50.93	5.87	0.056	-7.21	3405
5	1076.65	649.52	0.50	0.90	85.98	69.00	31.29	0.000	0.00	2842
10	1848.47	995.65	0.40	0.90	85.40	69.32	31.32	0.000	0.00	4946
100	15893.13	7230.05	0.34	0.90	85.21	69.28	31.46	0.000	0.00	1968
10000	1561747.00	693547.64	0.25	0.90	85.17	69.32	31.57	0.000	0.00	3807

model (22) is presented in order to analyze the impact of these measures in the solution. Tables 4 and 5 present the results of the VaR and CVaR models (23) and (26), respectively, for $\gamma = 1$ and $\rho \geq 0$. The headings are as follow: Z_{LP}, solution value of the LP relaxation; Z_{MIP}, solution value of the incumbent solution; GAP, related optimality gap (remind that the bound for the optimality GAP has been set up to 0.5%); β, lower bound on the success probability given by the user for the scenario to occur; VaR is the 0.90-VaR; CVaR is the 0.90-CVaR; $\%Q_E$, normalized weighted expected profit over all scenarios, where 100 is the expected profit for the risk neutral model (22); $P(<0)$, sum of the weights of the scenarios with a negative profit; $E(<0)$, conditional expectation of the scenarios with a negative profit; t, elapsed time (in secs.) for obtaining the incumbent solution.

We can observe in Tables 4 and 5 that the weight parameter $\rho > 1$ provides good VaR and CVaR values in both models. However, it does no provide good expected profit, since there is a 12% reduction (approx.) with respect to the risk neutral model (22) for $\rho = 5$ and $\beta = 0.95$, and an even larger for very large ρ values. Notice that the VaR based model (23) requires much more elapsed time that the CVaR model (26) (on average, 1 hour and 3-4 minutes, respectively).

Table 6 presents the results of the VaR+Over-CVaR model (24). The headings are the same as in Tables 4 and 5. We can observe that, in general, this strategy provides good results, but it requires much more elapsed time than the strategy VaR in some cases and the strategy CVaR in all cases.

Table 7 presents the results of the $Q_e\&DP$ model (27). The additional headings are as follows:

Table 5: 27-scen tree. Stochastic solution values. Q_e & CVaR model (26)

ρ	Z_{LP}	Z_{MIP}	$GAP(\%)$	β	% Q_E	VaR	CVaR	$P(<0)$	$E(<0)$	t (secs)
0	354.72	354.18	0.15	1.00	100.00	−11.85	−39.77	0.157	−25.88	241
0.01	354.25	354.15	0.15	0.95	100.00	−10.71	−39.94	0.157	−25.89	301
0.5	341.05	339.74	0.39	0.95	97.50	22.25	−5.78	0.069	−8.97	193
0.75	342.57	341.21	0.40	0.95	92.71	42.60	24.51	0.000	0.00	335
1	347.84	346.90	0.27	0.95	91.03	48.62	28.53	0.000	0.00	206
5	465.12	459.16	0.33	0.95	87.78	53.99	33.86	0.000	0.00	280
10	631.78	608.38	0.48	0.95	87.08	54.06	34.03	0.000	0.00	239
100	3663.48	3326.77	0.47	0.95	86.47	54.78	34.38	0.000	0.00	298
10000	337426.10	303513.62	0.12	0.95	86.39	54.26	34.33	0.000	0.00	335
0.01	354.34	353.36	0.37	0.90	99.78	−10.75	−39.23	0.157	−25.29	315
0.5	347.07	346.07	0.29	0.90	97.66	21.98	−5.52	0.069	−9.17	279
0.75	349.43	348.09	0.38	0.90	93.27	42.15	23.15	0.000	0.00	241
1	356.62	355.48	0.32	0.90	91.01	53.19	32.57	0.000	0.00	224
5	511.27	503.47	0.48	0.90	87.59	62.57	37.96	0.000	0.00	318
10	726.03	700.20	0.40	0.90	86.60	63.62	38.65	0.000	0.00	313
100	4630.34	4347.86	0.38	0.90	64.81	60.31	36.01	0.000	0.00	373
10000	434340.60	412059.55	0.32	0.90	64.82	60.31	36.01	0.000	0.00	289

Table 6: 27-scen tree. Stochastic solution values. VaR & overCVaR model (24)

ρ	Z_{LP}	Z_{MIP}	$GAP(\%)$	β	% Q_E	VaR	CVaR	$P(<0)$	$E(<0)$	t (secs)
0.5	774.54	769.97	0.50	0.95	57.53	59.44	33.45	0.000	0.00	353
1	359.72	355.15	0.50	0.95	100.11	−11.59	−42.01	0.130	−32.85	25899
5	642.28	444.86	0.50	0.95	88.13	53.97	28.89	0.009	−11.16	4476
10	1068.16	613.19	0.31	0.95	87.22	55.07	38.13	0.000	0.00	2672
100	8793.76	4218.26	0.20	0.95	65.22	61.21	35.75	0.000	0.00	3550
10000	858951.30	402870.22	0.09	0.95	65.60	61.19	37.23	0.000	0.00	2135
0.5	774.71	769.93	0.44	0.90	57.52	59.14	33.46	0.000	0.00	468
1	364.72	359.17	1.49	0.90	99.76	−3.22	−63.39	0.157	−40.37	72000[1]
5	756.29	584.88	0.44	0.90	86.18	68.69	18.16	0.000	0.00	2305
10	1312.81	930.63	0.10	0.90	85.13	69.37	20.47	0.000	0.00	2129
100	11438.15	7176.66	0.22	0.90	85.11	69.40	26.51	0.000	0.00	4688
10000	1125623.00	693796.13	0.21	0.90	84.95	69.36	22.60	0.000	0.00	4921

[1] Time limit reached

Table 7: 27-scen tree. Stochastic solution values. Q_e & DP model (27)

ρ	Z_{LP}	Z_{MIP}	$GAP(\%)$	ϕ	$P(<\phi)$	$\%\ Q_E$	VaR	CVaR	$P(<0)$	$E(<0))$	t (secs)
0	354.72	354.18	0.15		1.00	100.00	−11.85	−39.77	0.157	−25.88	241
25	353.57	349.17	0.50	30.00	0.21	100.09	−7.89	−37.81	0.111	−34.07	2830
100	350.44	338.63	0.35	30.00	0.13	99.14	11.89	−19.56	0.079	−24.65	6819
200	346.82	327.74	0.26	30.00	0.07	96.45	30.44	4.27	0.056	−7.84	14599
400	341.56	314.45	0.45	30.00	0.07	96.62	30.41	4.38	0.056	−7.64	5083
800	334.79	290.49	0.23	30.00	0.04	90.38	50.87	30.54	0.000	0.00	4800
1200	329.34	275.44	0.33	30.00	0.00	90.32	50.86	30.35	0.000	0.00	6937
25	351.84	348.25	0.46	90.00	0.24	100.02	−7.28	−39.06	0.157	−25.03	478
100	343.59	330.51	0.48	90.00	0.24	100.11	−9.84	−39.99	0.157	−25.68	1256
200	333.68	309.18	0.39	90.00	0.21	99.32	2.31	−34.77	0.097	−34.77	5537
400	319.10	267.05	0.49	90.00	0.20	97.88	13.47	−15.12	0.069	−23.29	4023
800	294.59	194.78	0.18	90.00	0.15	89.50	55.49	28.36	0.000	0.00	4586
1200	277.26	133.67	0.26	90.00	0.00	89.50	55.52	28.33	0.000	0.00	5899

ϕ, threshold profit; and $P(<\phi)$, sum of the weights of the scenarios with negative profit over the threshold. The results shown in the table clearly have a risk reduction versus the risk neutral model (22) for big ρ values. The larger the threshold, the worse the solution value is in terms of the CVaR and the probability of having a scenario with negative profit. Therefore, the best combination for this situation seems to be small threshold (say, ϕ=30) and large parameter ρ (say, 400); however, notice that at the end the combination is chosen by the user to satisfy its own requirements. This combination provides very similar solutions to the ones provided by the CVaR model (26), see Table 5. However, the former requires more than 5000 secs. and the latter requires 200 secs. (approx.) of elapsed time.

Finally, Tables 8 and 9 present the results of the Stochastic Dominance Constraint models (28) and (29), respectively. We can observe that we have not reported results for big sets $\mathcal{P}$ of profiles due to unaffordable elapsed time that is required by plain use of the MIP solver. However, we have experimented with $|\mathcal{P}|$=2 and $|\mathcal{P}|$=1 profiles for the sdc-1 and sdc-2 strategies, respectively. We can observe that, in general, the results are very good but, in some cases, the elapsed time is very high. Although more computational experience is needed it seems that these strategies are worth to be considered in the future by using more profiles, provided that a quick exact decomposition algorithm is used instead of plain MIP solver, see Section 7.

For a better comparison of different solutions proposed by the risk averse approaches presented above, Fig. 7 depict the Expected Value (in abscissas) and the 0.90-CVaR value (in ordinates) presented in Tables 4 to 9 for the 27-scenario tree case. We can observe how the strategies approach

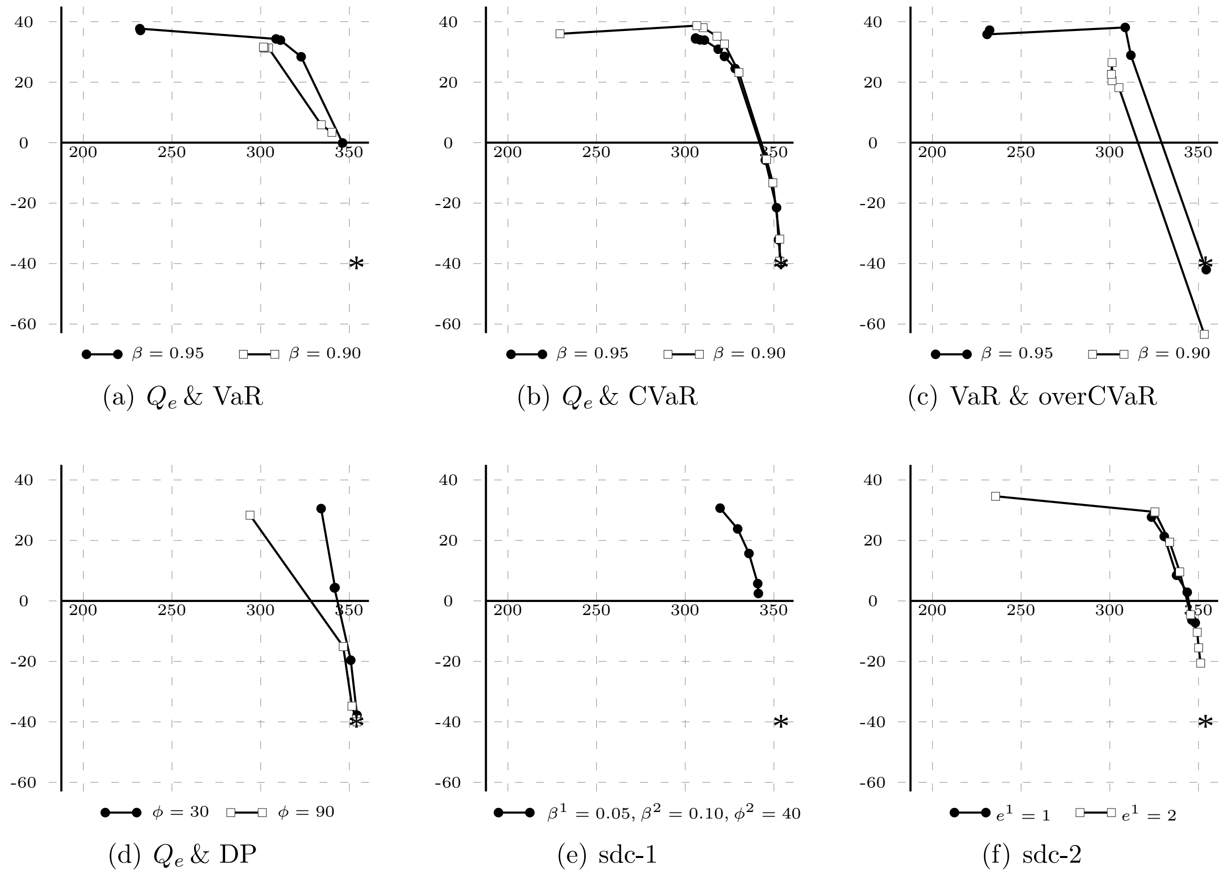
Table 8: 27-scen tree. Stochastic solution values. sdc-1 model (28)

ϕ^1	β^1	ϕ^2	β^2	Z_{LP}	Z_{MIP}	$GAP(\%)$	$\% Q_E$	VaR	CVaR	$P(<0)$	$E(<0)$	t (secs)
-10	0.05	40	0.10	352.69	341.28	0.51	96.36	40.00	2.462	0.056	-9.08	25410
0	0.05	40	0.10	352.15	341.05	1.34	96.29	40.00	5.716	0.037	-7.34	42583[1]
10	0.05	40	0.10	351.59	336.19	0.18	94.92	40.05	15.680	0.000	0.00	4778
20	0.05	40	0.10	351.02	329.79	0.13	93.11	42.57	23.813	0.000	0.00	2641
30	0.05	40	0.10	350.34	319.66	0.36	90.25	50.21	30.676	0.000	0.00	7881
40	0.05	40	0.10	infeasible								

[1] Cplex stop, out of memory

Table 9: 27-scen tree. Stochastic solution values. sdc-2 model (29)

ϕ^1	e^1	Z_{LP}	Z_{MIP}	$GAP(\%)$	$\% Q_E$	VaR	CVaR	$P(<0)$	$E(<0)$	t (secs)
0	1	349.07	348.38	0.20	98.36	14.92	-7.24	0.069	-14.31	866
10	1	344.67	343.84	0.24	97.08	21.32	2.86	0.056	-7.33	876
20	1	338.66	337.82	0.24	95.38	29.20	8.48	0.000	0.00	639
30	1	332.35	330.90	0.41	93.42	37.43	21.32	0.000	0.00	3216
40	1	324.38	323.68	0.22	91.39	45.00	27.72	0.000	0.00	567
50	1	infeasible								
60	1	infeasible								
0	2	352.66	351.43	0.35	99.22	8.93	-20.56	0.097	-20.56	2744
10	2	350.23	349.48	0.22	98.67	13.60	-10.36	0.069	-18.04	378
20	2	346.90	345.76	0.33	97.62	20.46	-4.57	0.056	-10.60	4740
30	2	341.01	339.64	0.40	95.89	30.34	9.59	0.037	-2.37	4455
40	2	334.33	333.92	0.12	94.28	40.57	19.44	0.000	0.00	5442
50	2	326.48	325.60	0.27	91.93	50.13	29.43	0.000	0.00	269
60	2	308.41	235.76	0.10	66.56	59.95	34.61	0.000	0.00	1726



(a) Q_e & VaR (b) Q_e & CVaR (c) VaR & overCVaR

(d) Q_e & DP (e) sdc-1 (f) sdc-2

Figure 7: 27-scen tree. Expected Value versus CVaR

a trade-off between maximizing the profit (expected value) and minimize the risk (CVaR). Figs. 7(a), 7(b) and 7(c) represent the results for β equal to 0.90 and 0.95, and the different values of ρ in Tables 4, 5, and 6, respectively. Fig. 7(d) represents the results for ϕ equal to 30 and 90 and the different values of ρ in Table 7. Fig. 7(e) represents the results for the different profiles in Table 8, where ϕ^1 vary from -10 to 30, $\beta^1 = 0.05$, $\phi^2 = 40$ and $\beta^2 = 0.10$. Finally, Fig. 7(f) represents the results for the different profiles in Table 9, where ϕ vary from 0 to 60 and e^1 is equal to 1 and 2. We can observe that the strategies $\mathrm{Q_e}$ & CVaR and sdc-2 provide the best results in our testbed. These figures represent for cases (a) to (d) the Pareto frontier of the bi-objective problem that consists of maximizing the net profit and minimizing the risk. The asterisk represents the value for the risk-neutral approach. It can be observed that a great reduction in risk can be obtained with some reduction in net profit for the adequate weighting parameters of the risk averse measures.

Based on the elapsed time reported in Tables 4 to 9 for the 27-scenario tree case, we only experimented for the 45- and 75-scenario tree cases with the following strategies:

- risk neutral model (22),

Table 10: 45-scen tree. Stochastic solution values

Model	ρ	Z_{LP}	Z_{MIP}	$GAP(\%)$	$\%\ Q_E$	VaR	CVaR	$P(<0)$	$E(<0)$	t (secs)
Q_E	0	350.15	349.82	0.09	100.00	-0.23	-33.47	0.102	-29.84	915
Q_E & DP	400	340.36	316.98	0.50	93.79	53.70	29.88	0.000	0.00	29892
Q_E & VaR	1	482.78	388.56	0.10	97.08	48.95	14.61	0.028	-6.28	29731
Q_E & CVaR	0.50	336.04	334.61	0.42	97.39	36.71	8.42	0.028	-10.39	529
	0.75	337.40	336.77	0.19	92.86	52.34	31.99	0.000	0.00	398
sdc-2		343.40	342.71	0.20	97.97	33.50	8.60	0.028	-10.90	14221

Table 11: 75-scen tree. Stochastic solution values.

Model	ρ	Z_{LP}	Z_{MIP}	$GAP(\%)$	$\%\ Q_E$	VaR	CVaR	$P(<0)$	$E(<0)$	t (secs)
Q_E	0	345.61	344.99	0.18	100.00	2.20	-26.85	0.091	-28.82	2115
Q_E & DP	400	332.29	302.12	3.39	93.76	56.99	34.09	0.000	0.00	72000[1]
Q_E & VaR	1	485.01	392.49	0.07	92.35	73.91	38.35	0.000	0.00	67387
Q_E & CVaR	0.50	347.06	346.06	0.29	97.66	46.45	16.78	0.023	-8.91	1780
	0.75	352.35	351.40	0.27	96.40	49.75	23.75	0.016	-3.71	2276
sdc-2	1	340.26								19126[2]

[1] Time limit reached. [2] Cplex stop, out of memory

- $\mathrm{Q_e}$ & VaR model (23) with $\gamma = \rho = 1$ and $\beta = 0.90$,
- $\mathrm{Q_e}$ & CVaR model (26) with $\gamma = 1, \rho \in \{0.5, 0.75\}$ and $\beta = 0.90$,
- $\mathrm{Q_e}$ & DP model (27), with $\rho = 400$ and $\phi = 30$,
- sdc-2 model (29) with $\phi^1 = 30$ and $e^1 = 2$,

whose main results are shown in Tables 10 and 11, respectively. Similarly to the smaller instances with 27 scenarios, we can observe that the tested risk averse strategies provide better solutions than the risk neutral approach, in the meaning that, with a moderate reduction in the expected profit (from 2.03% to 7.65%), the different risk measures give an interesting risk reduction as measured by the non-negative 0.90-VaR and 0.90-CVaR values. In any case, observe in Table 1 the very large dimensions of the risk neutral model (22).

It is worth to point out that the risk averse strategies that have been experimented with in this work offer much better results in the VaR, CVaR and conditional expected negative profit than the risk neutral strategy without, on the other hand, reducing too-much the expected profit.

7 Conclusions and future work

In his work we have presented the stochastic version of the copper extraction planning along a time horizon (i.e, years) under uncertainty in the (volatile) copper prices. The problem is a very large large-scale MIP model, even for a reduced version of the original one, since it requires to represent the uncertainty in the volatile copper prices by a multistage scenario tree. We have presented the Deterministic Equivalent Model (DEM) to the stochastic problem in a compact representation. The first conclusion that can be drawn from this work is that even the risk neutral approach, see model (22), provides a better solution than the traditional (and myopic) deterministic solution by considering the expected value of the uncertain parameters.

A second result from the 27-, 45- and 75-scenario tree cases that we have experimented with is that the risk adverse Expected value & CVaR strategy, see model (26), and the second-order stochastic dominance constraints (sdc) strategy, see model (29), seem to provide better results in the solution's quality (since they reduce the risk of bad scenarios without reducing to-much the expected profit) and they require less elapsed time than any other ones. In any case, these results have to be validated by an extensive computational experience with larger cases. Notice that we would not fully experimented with bigger profiles in the sdc strategies, nor with the Conditional expectation above CVaR strategy (24) and the other ones, due to the excessive computer requirements (i.e., memory and elapsed time) of the MIP engine to solve the very large scale instances of the real-life copper extraction problem under uncertainty in the copper prices.

The third definitive conclusion that can be drawn from the analysis of the three scenario tree illustrative cases is that solving the DEM by using even a state-of-the-art optimization engine may require, even in compact representation, so much computing effort for large-scale instances that a decomposition algorithmic is required. This scheme would allow to use scenario bundles based on a so named *break stage*, see [21]. So, we need to develop a decomposition approach for solving large scale real-life instances, such that smaller MIP submodels can be solved in parallel. For that purpose we can represent the DEM by a mixture of the splitting variable and compact representations. The first one will allow (based on the break stage) to decompose the model by scenario bundles. A Branch-and-Fix Coordination (BFC) [3] type of scheme can be used for handling the splitting variable representation to obtain the solution value of the original stochastic MIP problem, where the compact representation of the independent submodels related to the scenario bundles for the stages after the break stage will be optimized by plain use of the MIP optimization engine of choice; see [20, 21]. Alternatively to the BFC approach, the splitting variable representation will allow to use a Lagreangean Decomposition approach to obtain strong lower bounds of the solution value of the original stochastic problem, by dualizing the nonanticipativity constraints related to the stages up to the break stage; so, several iterative Lagrange multipliers updating schemes will be computationally analyzed and feasible solutions from the Lagrangean dual one can be obtained. In any case, the strategies Value-at-Risk (23), Conditional expectation above VaR (24) and stochastic dominance constraints (28) and (29) require constraints linking variables from different scenarios. Although those constraints are very few, their dualization via Lagrange Relaxation is an additional challenge for the decomposition approaches.

Acknowledgements

The authors would like to thank to Rüdiger Schultz for the time devoted to the paper. The fruitful discussions that some authors have had with him have contributed to a better insight on the meaning and applicability of the different risk measures discussed in the paper, as well as on the understanding of the frequent convergency between the deterministic and risk neutral solution values.

References

[1] A. Agustín, A. Alonso-Ayuso, L.F. Escudero and C. Pizarro. On Air Traffic Flow Management with Rerouting. Part I: Deterministic case. *Submitted for publication*, 2011.

[2] M. Albareda-Sambola, A. Alonso-Ayuso, L.F. Escudero, E. Fernández, Y. Hinojosa and C. Pizarro. A computational comparison of several formulations for the multi-period location-assignment problem. *TOP*, 18:62-80, 2010.

[3] 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.

[4] A. Alonso-Ayuso, L.F. Escudero, M.T. Ortuño and C. Pizarro. On a Stochastic Sequencing and Scheduling Problem. *Computers & Operations Research*, 34:2604-2624, 2007.

[5] A. Alonso-Ayuso, L.F. Escudero, A. Garín, M.T. Ortuño and G. Pérez. A Stochastic 0-1 Program based approach for Strategic Supply Chain Planning under Uncertainty. *Journal of Global Optimization*, 26:97-124, 2003.

[6] L. Aranburu, L.F. Escudero, A Garín and G. Pérez. Stochastic models for optimizing immunization strategies in fixed-income security portfolios under some sources of uncertainty. *Submitted for publication*, 2011.

[7] D. Bertsimas and S. Stock. The air traffic flow management problem with enroute capacities. *Operations Research*, 46:406-422, 1998.

[8] J.R. Birge and F.V. Louveaux. *Introduction to Stochastic Programming.* Springer, 1997.

[9] R. Caro, R. Epstein, P. Santibañez and A.Weintraub. Handbook of Operation Research in Natural Resources, chapter 32. Springer Science, 2007.

[10] A. Charnes and W.W. Cooper. Chance-constrained programming. *Management Science*, 5:73-79, 1959.

[11] COCHILCO. Anuario de estadísticas del cobre y otros minerales, 1987-2006, `http://www.cochilco.cl/productos/grafico_ingreso.asp?tipo_metal=1`, Chile.

[12] R. Dembo. Scenario immunization. *Annals of Operations Research*, 30:63-90, 1991.

[13] J. Dupacova, G. Consigli and S.W. Wallace. Scenarios for multistage stochastic programs. *Annals of Operations Research*, 100:25-53, 2000.

[14] G.D. Eppen, R.K. Martin and L. Schrage. Scenario approach to capacity planning. *Operations Research*, 34:517-527, 1989.

[15] R. Epstein, S. Gaete, F. Caro, A. Weintraub, P. Santibañez and J. Catalán. Optimizing long term planning for underground copper mines. International Copper Conference, Santiago, Chile, 2003.

[16] Epstein R, Goig M, Weintraub A, Catalán J, Santibañez P, Urrutia R, Cancino R, Gaete S, Aguayo A. Optimizing Long - Term Production Plans in Underground and Open Pit Copper Mines. *To appear in Operations Research*, 2011.

[17] L.F. Escudero. Robust Portfolios for Mortgage-Backed Securities. In S.A. Zenios (ed.), *Quantitative Methods. AI and Supercomputers in Finance.* Unicom, London, 201-228, 1995.

[18] L.F. Escudero. On a mixture of the Fix-and-Relax Coordination and Lagrangean Substitution schemes for multistage stochastic mixed integer programming. *TOP*, 17:5-29, 2009.

[19] L.F. Escudero, A. Garín and G. Pérez. The value of the stochastic solution in multistage problems. *TOP*, 15:48-64, 2007.

[20] L.F. Escudero, A. Garín, M. Merino and G. Pérez. On BFC-MSMIP strategies for scenario cluster parttioning and Twin Nodes Families branching slection and bounding for multistage stochastic mixed integer programming. *Computers & Operations Research*, 37:738-753, 2010.

[21] L.F. Escudero, A. Garín, M. Merino and G. Pérez. An algorithmic framework for solving large scale multistage stochastic mixed 0-1 problems with nonsymmetric scenario trees. *Computers & Operations Research*, doi: 10.1016/j.cor.2011.06.021, 2011.

[22] GAMS CO. GAMS user manual, 2010.

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

[24] R. Gollmer, U. Gotzes and R. Schultz. A note on second-order stochastic dominance constraints induced by mixed-integer linear recourse. *Mathematical Programming, Ser. A*, 126:179-190, 2011.

[25] J. Hartigan. *Clustering Algorithms.* J. Wiley, p. 686, 1975.

[26] H. Heitsch and W. Römisch. Scenario tree reduction for multistage stochastic programs. *Computational Management Science*, 6:117-133, 2009.

[27] K. Hoyland, M. Kaut and S.W. Wallace. A heuristic for moment-matching scenario generation. *Computational Optimization and Applications*, 24.169-185, 2003.

[28] IBM ILOG. CPLEX user manual, 2009.

[29] W.K. Klein. *Duality in stochastic linear and dynamic porogramming.* Spirnger-Varlag, 1986.

[30] W.K. Klein and M.H. van der Vlerk. Integrated chance constraints: reduced forms an algorithm. *Computational Management Science*, 3:245-269, 2006.

[31] W. Ogryczak and A. Ruszczynski. From stochastic dominance to mean-risk models: semideviations as risk measures. *European Journal of Operational Research*, 116: 33-50, 1999.

[32] R.T. Rockafellar and S. Uryasev. Optimization on conditional value-at-risk. *Journal of Risk*, 2:21-41, 2000.

[33] R.T. Rockafellar and R.J-B Wets. Scenario and policy aggregation in optimisation under uncertainty. *Mathematics of Operations Research*, 52:359-375, 1991.

[34] R. Schultz and S. Tiedemann. Risk aversion via excess probabilities in stochastic programs with mixed-integer recourse. *SIAM Journal on Optimization*, 14:115-138, 2003.

[35] R. Schultz and S. Tiedemann. Conditional value-at-risk in stochastic programs with mixed integer recourse. *Mathematical Programming, Ser. B*, 105:365-386, 2006.

[36] L. Trigeorgis *Real options.* MIT Press, 1996.

[37] A. Weintraub, M. Pereira and X. Schultz. A priori and a posteriori aggregation procedures to reduce model size in MIP mine planning models. Working paper, Department of Industrial Engineering, University of Chile, Santiago, Chile, 2007.

[38] Weintraub A, Saez G, Yadlin M. Aggregation Procedures in Forest Management Planning Using Cluster Analysis. *Forest Science*, 43(2)274-284, 1997.

[39] R.J-B Wets. Stochastic programs with fixed recourse: The equivalent deterministic program. *SIAM Review*, 16:309-339, 1974.

[40] P. Zipkin. Bounds on the effect of aggregating variables in linear programs. *Operations Research*, 28:1299-1309, 1980.