

A new matheuristic approach for optimizing mineral value chains under uncertainty

A. Lamghari
R. Dimitrakopoulos

G-2019-93

December 2019

La collection *Les Cahiers du GERAD* est constituée des travaux de recherche menés par nos membres. La plupart de ces documents de travail a été soumis à des revues avec comité de révision. Lorsqu'un document est accepté et publié, le pdf original est retiré si c'est nécessaire et un lien vers l'article publié est ajouté.

Citation suggérée : A. Lamghari, R. Dimitrakopoulos (Décembre 2019). A new matheuristic approach for optimizing mineral value chains under uncertainty, Rapport technique, Les Cahiers du GERAD G-2019-93, GERAD, HEC Montréal, Canada.

Avant de citer ce rapport technique, veuillez visiter notre site Web (<https://www.gerad.ca/fr/papers/G-2019-93>) afin de mettre à jour vos données de référence, s'il a été publié dans une revue scientifique.

The series *Les Cahiers du GERAD* consists of working papers carried out by our members. Most of these pre-prints have been submitted to peer-reviewed journals. When accepted and published, if necessary, the original pdf is removed and a link to the published article is added.

Suggested citation: A. Lamghari, R. Dimitrakopoulos (December 2019). A new matheuristic approach for optimizing mineral value chains under uncertainty, Technical report, Les Cahiers du GERAD G-2019-93, GERAD, HEC Montréal, Canada.

Before citing this technical report, please visit our website (<https://www.gerad.ca/en/papers/G-2019-93>) to update your reference data, if it has been published in a scientific journal.

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2019
– Bibliothèque et Archives Canada, 2019

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2019
– Library and Archives Canada, 2019

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Télec. : 514 340-5665
info@gerad.ca
www.gerad.ca

A new matheuristic approach for optimizing mineral value chains under uncertainty

Amina Lamghari ^{a,b,c}

Roussos Dimitrakopoulos ^{a,b,c}

^a GERAD, Montréal (Québec), Canada, H3T 2A7

^b COSMO – Stochastic Mine Planning Laboratory
Montréal (Québec) Canada, H3A 0E8

^c Department of Mining and Materials Engineering,
McGill University, FDA Building, Montréal
(Québec), Canada, H3A 0E8

amina.lamghari@mcgill.ca

roussos.dimitrakopoulos@mcgill.ca

December 2019
Les Cahiers du GERAD
G–2019–93

Copyright © 2019 GERAD, Lamghari, Dimitrakopoulos

Les textes publiés dans la série des rapports de recherche *Les Cahiers du GERAD* n'engagent que la responsabilité de leurs auteurs. Les auteurs conservent leur droit d'auteur et leurs droits moraux sur leurs publications et les utilisateurs s'engagent à reconnaître et respecter les exigences légales associées à ces droits. Ainsi, les utilisateurs:

- Peuvent télécharger et imprimer une copie de toute publication du portail public aux fins d'étude ou de recherche privée;
- Ne peuvent pas distribuer le matériel ou l'utiliser pour une activité à but lucratif ou pour un gain commercial;
- Peuvent distribuer gratuitement l'URL identifiant la publication.

Si vous pensez que ce document enfreint le droit d'auteur, contactez-nous en fournissant des détails. Nous supprimerons immédiatement l'accès au travail et enquêterons sur votre demande.

The authors are exclusively responsible for the content of their research papers published in the series *Les Cahiers du GERAD*. Copyright and moral rights for the publications are retained by the authors and the users must commit themselves to recognize and abide the legal requirements associated with these rights. Thus, users:

- May download and print one copy of any publication from the public portal for the purpose of private study or research;
- May not further distribute the material or use it for any profit-making activity or commercial gain;
- May freely distribute the URL identifying the publication.

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Abstract: Mineral value chains or mining complexes involve mining, processing, stockpiling, waste management, and transportation activities. An integrated stochastic optimization of all related aspects and activities has been shown to increase the net present value of related mining projects and operations, reduce risk in meeting production targets, and lead to more robust and coordinated production schedules. However, it entails solving a substantially large and more complex stochastic optimization problem than separately optimizing individual components of a mineral value chain. To tackle this complex optimization problem, a new matheuristic that integrates components from exact algorithms (relaxation and decomposition), machine learning techniques (reinforcement learning and artificial neural networks), and heuristics (local improvement and randomized search) is proposed. A general mathematical formulation that serves as the basis for the proposed methodology is also introduced.

Keywords: Mineral value chains/mining complexes, stochastic simultaneous optimization, large-scale optimization, matheuristic, decomposition

1 Introduction

Production scheduling and downstream optimization are two classical problems in strategic mine planning. Production scheduling concerns strategic aspects such as designing a mining sequence of extraction over the life-of-the-mine, while downstream optimization deals with the material flow leading to metal products aspects and assumes that major strategic decisions have already been made. These two problems have been intensively studied in the mining literature. However, most related studies focus on a single problem although the two are closely related, leaving behind the benefits of coordinating and integrating mine planning and downstream production decisions, one main reason being computational complexity. The joint optimization of production scheduling and downstream optimization while explicitly accounting for uncertainty (geological and/or financial), henceforth referred to as the simultaneous stochastic optimization of mining complexes (SSOMC), has received attention over the last few years, however, despite new developments computational complexity remains an aspect to be further addressed.

A mining complex or mineral value chain can be seen as a system where raw materials are extracted from one or several mineral deposits, transformed into sellable products using different processing streams, and delivered to the spot-market or customers (Goodfellow and Dimitrakopoulos, 2017; Montiel and Dimitrakopoulos 2015). Goodfellow and Dimitrakopoulos (2016) were the first to introduce a mathematical model that specifically addresses mining complexes optimization under geological uncertainty. The authors employed a multi-neighborhood simulated annealing algorithm to solve the problem. The algorithm uses three neighborhoods. The first neighborhood changes the period in which a block is extracted, the second changes the destination of a cluster of blocks, and the third one modifies the amount of material sent from one processor to another. The authors compared the proposed algorithm to two other methods that combine multi-neighborhood simulated annealing with particle swarm and differential evolution, respectively. The latter proved to be more efficient, but also more computationally expensive. Multi-neighborhood simulated annealing was further studied in Montiel and Dimitrakopoulos (2015), Montiel et al. (2016), and Montiel and Dimitrakopoulos (2017, 2018). In the first paper, the authors simultaneously optimize mining, processing, and transportation decisions. The initial solution is improved by shifting a single block to another period, changing the operating mode in a processor, and changing the transportation arrangement in a processor. In the other two papers, the model and the algorithms are designed to handle mining complexes with multiple mines. Del Castillo and Dimitrakopoulos (2019) investigated a similar problem but in which capital expenditure decisions are included. There are some studies that considered exact methods. Zhang et al. (2019) studied mining complexes under market uncertainty. They combined classical Benders decomposition with aggregation techniques to reduce the size of the problem so that the resulting model is of tractable size and can be solved by the proposed method.

Nearly all papers on SSOMC, except the one above by Zhang et al. (2019), employs metaheuristics, mainly multi-neighborhood simulated annealing, to meet the challenges of scale, complexity, and uncertainty in optimizing mining complexes. In this paper, a different approach, namely a new metaheuristic, is introduced. The proposed methodology capitalizes on the synergies between artificial intelligence and optimization techniques and integrates components from exact algorithms (relaxation and decomposition), machine learning techniques (reinforcement learning and artificial neural networks), and heuristics (local improvement and randomized search).

The remainder of the paper is divided into four sections. Section 2 gives a formal mathematical description of the problem studied. The proposed solution procedure is described in Section 3. Preliminary numerical results are reported in Section 4. Section 5 provides conclusions and directions for future research.

2 Problem description and formulation

2.1 Problem description

A typical mining complex (mineral value chain) consists of multiple mines that provide material to multiple downstream facilities where the supply is blended and transformed into final sellable products. A downstream facility could be a stockpile, a crusher, a mill, a concentrator, a leach pad, a refinery, a waste dump, etc. and the output from one facility may be used as an input for another facility. In this general context, three kinds of decisions must be made in each period of the planning horizon: whether or not to mine a block (extraction decisions), where to send extracted blocks (first-stage destination decisions), and how to route the flow from the facilities fed by the mines to the facilities where final sellable products are produced (routing decisions). The constraints to be taken into account can be grouped into four main classes: precedence constraints, resource constraints, blending constraints, and flow conservation constraints. The information about blocks' mineral properties required for decision making is not known with certainty a priori, but only approximately in the form of equiprobable geological realizations (scenarios) of the orebodies. Such realizations are used to derive a robust solution that is well-hedged against uncertainty and where the risk of not meeting production targets is minimized.

In this paper, while extraction decisions are made at the block level, first-stage destination decisions are made at an aggregate level based on a clustering method similar to that proposed in Goodfellow and Dimitrakopoulos (2016). More specifically, in a pre-processing step, blocks' mineral properties are used to measure blocks' similarities and partition them into $\mathcal{G}$ subsets, henceforth called *groups*. We then refer to block i being in group g under geological scenario s when its mineral properties under that scenario fall into that group. Grouping blocks to define first-stage destinations is interesting from a practical point of view as it provides the mine planners with guidelines for what cut-off grade they should operate with for each year. It is also useful computationally as it significantly reduces the size of the optimization problem and consequently, the computational burden of solving it, but at the cost of a loss in selectivity. The selectivity drawback can be overcome by setting $\mathcal{G}$ to high values close to the number of blocks.

A mathematical formulation of the optimization problem described above is presented in the next section.

2.2 Problem formulation

The following notation is used to formulate the problem.

Sets

- $\mathcal{T}$: Set of time periods, indexed by t .
- $\mathcal{B}$: Set of blocks, considering all mines in the mining complex (mineral value chain), indexed by b .
- $\mathcal{P}_b \subset \mathcal{B}$: Set of immediate predecessors of block b , indexed by b' .
- $\mathcal{S}$: Set of scenarios modelling geological uncertainty, indexed by s .
- $\mathcal{G}$: Set of groups, indexed by g . Recall that the groups are defined based on the mineral properties of all blocks in the different mines, considering all geological scenarios.
- $\mathcal{A}$: Set of attributes or specificities to be tracked across the mineral value chain, indexed by a .
- $\mathcal{M} \subset \mathcal{A}$: Set of final sellable products (minerals), indexed by m .
- $\mathcal{R}_a \subset \mathcal{A}$: Set of attributes whose proportion relatively to the proportion of attribute a has to be controlled, indexed by a' .

To characterize the downstream part of the mining complex that goes from the first-stage facilities fed from the mines where blocks are first mixed, to intermediate facilities where material is further mixed and processed, to last-stage facilities where final sellable products are recovered, an acyclic direct graph is used. Nodes represent any downstream facility, and arcs between two nodes define an existing flow. The following notation is used to define the graph.

- $\mathcal{I}$: Set of nodes associated with first-stage facilities, indexed by i .
- $\mathcal{L}$: Set of nodes associated with last-stage facilities, indexed by l .
- $\mathcal{N} = \mathcal{I} \cup \mathcal{J} \cup \mathcal{L}$: Set of all nodes, indexed by n . Therefore, a node $j \in \mathcal{J}$ is neither associated with a first-stage facility nor associated with a last-stage facility. It is associated with an intermediate facility.
- A : Set of arcs. An arc (i, j) exists only if facility i can potentially send material to facility j . Without loss of generality, we assume that $A \subseteq (\mathcal{I} \times \mathcal{J}) \cup (\mathcal{J} \times \mathcal{J}) \cup (\mathcal{J} \times \mathcal{L}) \cup (\mathcal{I} \times \mathcal{L})$

Parameters

- w_b : Weight of block b (tonnage).
- α_{ab}^s : Grade of attribute a in block b under scenario s . The grade is defined to be the proportion of attribute to rock.
- $\beta_{bg}^s = \begin{cases} 1 & \text{if block } b \text{ belongs to group } g \text{ under scenarios} \\ 0 & \text{otherwise} \end{cases}$

Recall that groups $g \in \mathcal{G}$ are disjoint and are defined based on the blocks' mineral properties. Therefore, for a given scenario, a block cannot appear in different groups; that is, $\sum_{g \in \mathcal{G}} \beta_{bg}^s = 1 \forall b \in \mathcal{B}, s \in S$. However, if we consider all scenarios, a block might appear in multiple groups; that is,

$$\sum_{s \in S} \sum_{g \in \mathcal{G}} \beta_{bg}^s \geq 1 \quad \forall b \in \mathcal{B}.$$

- $\epsilon_{gi} = \begin{cases} 1 & \text{if group } g \text{ is eligible for node } i \in \mathcal{I} \\ 0 & \text{otherwise.} \end{cases}$

Recall that each $i \in \mathcal{I}$ is associated with one first-stage facility. An available (extracted) block cannot be sent to just any first-stage facility. It can go only to a predetermined subset of facilities depending on the block's material type and also on the mine from which the block has been extracted. The parameter ϵ_{gi} is used to identify such subset.

- $\overline{W}^t$: Maximum amount of material that can be mined in period t .
- $\overline{C}_n^t$: Maximum amount of material that can be processed at node $n \in \mathcal{N} = \mathcal{I} \cup \mathcal{J} \cup \mathcal{L}$ in period t (capacity of node n).
- $\overline{P}_{al}^t$: Upper limit (bound requirement) on the proportion (level) of attribute a at node $l \in \mathcal{L}$ in period t .
- $\overline{R}_{aa'l}^t$: Upper limit on the proportion of attribute a relatively to attribute a' at node $l \in \mathcal{L}$ in period t .
- r_{ml}^t : Recovery of final sellable product $m \in \mathcal{M}$ at node $l \in \mathcal{L}$ during period t ($r_{ml}^t \in [0, 1]$). Recall that the set of final sellable products $\mathcal{M}$ is a subset of the set of attributes $\mathcal{A}$. The recovery represents the amount of output of attribute m obtained for each ton of input of m .
- c_b^t : Per-unit cost of mining block b in period t .
- c_{ij}^t : Per-unit cost of processing material at node j and/or transporting material on arc $(i, j) \in A$.
- σ_n^{ts+} : Unit surplus cost incurred if the amount of material received at node $n \in \mathcal{N} = \mathcal{I} \cup \mathcal{J} \cup \mathcal{L}$ during period t exceeds the capacity $\overline{C}_n^t$ of that node during that period.

- c_{al}^{t+} : Unit surplus cost incurred if the proportion of attribute a in node $l \in \mathcal{L}$ exceeds $\overline{P_{al}^t}$ during period t .
- $c_{aa'l}^{t+}$: Unit surplus cost incurred if the proportion of attribute a relatively to that of a' in node $l \in \mathcal{L}$ exceeds $\overline{P_{aa'l}^t}$ during period t .
- π_m^t : Per-unit profit from selling final sellable product (mineral) m in period t , defined as being the revenue minus the selling cost.

Decision variables

- For each block $b \in \mathcal{B}$ and each period $t \in \mathcal{T}$, we define the following binary variables:

$$x_b^t = \begin{cases} 1 & \text{if block } b \text{ is mined in period } t \\ 0 & \text{otherwise.} \end{cases}$$

- For each group $g \in \mathcal{G}$, each node $i \in \mathcal{I}$ (associated with a first-stage facility) and each period $t \in \mathcal{T}$, we define the following binary variables:

$$\kappa_{gi}^t = \begin{cases} 1 & \text{if blocks in group } g \text{ are sent to destination } i \text{ in period } t \\ 0 & \text{otherwise.} \end{cases}$$

As discussed in Section 2.1, variables κ_{gi}^t control the cut-off grade applied at i during period t . Other decision variables in the formulation are the following:

- y_{ij}^{ts} : Flow of material on arc (i, j) in period t under scenario s .
- Δ_n^{ts+} : Surplus in the amount of material received at node $n \in \mathcal{N}$ in period t under scenario s . Recall that the set of nodes $\mathcal{N}$ comprises nodes associated with first-stage facilities $i \in \mathcal{I}$, with intermediate facilities node $j \in \mathcal{J}$, and with last-stage facilities $l \in \mathcal{L}$.
- p_{aj}^{ts} : Proportion value (grade) of attribute $a \in \mathcal{A}$ in material leaving node $j \in \mathcal{I} \cup \mathcal{J}$ at period t under scenario s .
- δ_{al}^{ts+} : Surplus in quantity of attribute a in node $l \in \mathcal{L}$ in period t under scenario s (before considering the arriving material).
- $d_{aa'l}^{ts+}$: Surplus in the proportion of quantity of attribute $a \in \mathcal{A}$ compared to the quantity of attribute $a' \in \mathcal{R}_a$ in node $l \in \mathcal{L}$ in period t under scenario s (before considering the arriving material).

To simplify the presentation, we introduce the following three variables, although not necessary:

- μ_i^{ts} : Quantity of material (rock) supplied from the mines to node $i \in \mathcal{I}$ at period t under scenario s .
- ϑ_{ai}^{ts} : Proportion value (grade) of attribute $a \in \mathcal{A}$ in material supplied from the mines to node $i \in \mathcal{I}$ during period t under scenario s .
- o_m^{ts} : Quantity produced of sellable product (mineral) $m \in \mathcal{M}$ in period t under scenario s .

Denote by ζ the random data vector (uncertain parameters) and by $\zeta(s)$ one of its particular realizations. Assuming that all the scenarios are equiprobable, the two-stage stochastic formulation of the problem can be written as follows:

$$\max f(x, \kappa) = - \sum_{t \in \mathcal{T}} \sum_{b \in \mathcal{B}} c_b^t x_b^t + \frac{1}{S} \sum_{s \in \mathcal{S}} Q(x, \kappa, \zeta(s)) \quad (1)$$

Subject to

$$\sum_{t \in \mathcal{T}} x_b^t \leq 1 \quad \forall b \in \mathcal{B} \quad (2)$$

$$x_b^t \leq \sum_{\tau=1}^t x_{b'}^{\tau} \quad \forall b \in \mathcal{B}, b' \in \mathcal{P}_b, t \in \mathcal{T} \quad (3)$$

$$\sum_{b \in \mathcal{B}} w_b x_b^t \leq \overline{W}^t \quad \forall t \in \mathcal{T} \quad (4)$$

$$\sum_{i \in \mathcal{I}} \kappa_{gi}^t = 1 \quad \forall g \in \mathcal{G}, t \in \mathcal{T} \quad (5)$$

$$\kappa_{gi}^t \leq \epsilon_{gi} \quad \forall g \in \mathcal{G}, i \in \mathcal{I}, t \in \mathcal{T} \quad (6)$$

$$x_b^t \in \{0, 1\} \quad \forall b \in \mathcal{B}, t \in \mathcal{T} \quad (7)$$

$$\kappa_{gi}^t \in \{0, 1\} \quad \forall g \in \mathcal{G}, i \in \mathcal{I}, t \in \mathcal{T} \quad (8)$$

where $Q(x, \kappa, \zeta(s))$ is the optimal value of the following problem (second-stage problem):

$$\begin{aligned} Q(x, \kappa, \zeta(s)) = & \max \sum_{t \in \mathcal{T}} \left\{ - \sum_{(i,j) \in A} c_{ij}^t y_{ij}^{ts} \right. \\ & - \sum_{n \in \mathcal{N}} \sigma_n^{t+} \Delta_n^{ts+} - \sum_{a \in \mathcal{A}} \sum_{l \in \mathcal{L}} c_{al}^{t+} \delta_{al}^{ts+} - \sum_{a \in \mathcal{A}} \sum_{a' \in \mathcal{R}_a} \sum_{l \in \mathcal{L}} c_{aa'l}^{t+} d_{aa'l}^{ts+} \\ & \left. + \sum_{m \in \mathcal{M}} \pi_m^t o_m^{ts} \right\} \end{aligned} \quad (9)$$

Subject to

$$\sum_{g \in \mathcal{G}} \sum_{b \in \mathcal{B}} \beta_{bg}^s w_b x_b^t \kappa_{gi}^t = \mu_i^{ts} \quad \forall i \in \mathcal{I}, t \in \mathcal{T} \quad (10)$$

$$\mu_i^{ts} = \sum_{j \in \mathcal{J} \cup \mathcal{L}} y_{ij}^{ts} \quad \forall i \in \mathcal{I}, t \in \mathcal{T} \quad (11)$$

$$\sum_{i \in \mathcal{I} \cup \mathcal{J}} y_{ij}^{ts} = \sum_{l \in \mathcal{J} \cup \mathcal{L}} y_{jl}^{ts} \quad \forall j \in \mathcal{J}, t \in \mathcal{T} \quad (12)$$

$$\mu_i^{ts} - \Delta_i^{ts+} \leq \overline{C}_i^t \quad \forall i \in \mathcal{I}, s \in \mathcal{S} \quad (13)$$

$$\sum_{i \in \mathcal{I} \cup \mathcal{J}} y_{ij}^{ts} \leq \overline{C}_j^t \quad \forall j \in \mathcal{J} \cup \mathcal{L}, t \in \mathcal{T} \quad (14)$$

$$\sum_{g \in \mathcal{G}} \sum_{b \in \mathcal{B}} \beta_{bg}^s \alpha_{ab}^s w_b x_b^t \kappa_{gi}^t = \vartheta_{ai}^{ts} \mu_i^{ts} \quad \forall a \in \mathcal{A}, i \in \mathcal{I}, t \in \mathcal{T} \quad (15)$$

$$\sum_{i \in \mathcal{I}} \vartheta_{ai}^{ts} y_{ij}^{ts} + \sum_{j' \in \mathcal{J}} p_{aj'}^{ts} y_{j'j}^{ts} = p_{aj}^{ts} \sum_{i \in \mathcal{I} \cup \mathcal{J}} y_{ij}^{ts} \quad \forall a \in \mathcal{A}, j \in \mathcal{J} \cup \mathcal{L}, t \in \mathcal{T} \quad (16)$$

$$p_{al}^{ts} - \delta_{al}^{ts+} \leq \overline{P}_{al}^t \quad \forall a \in \mathcal{A}, l \in \mathcal{L}, t \in \mathcal{T} \quad (17)$$

$$p_{al}^{ts} - d_{aa'l}^{ts+} \leq \overline{R}_{aa'l}^t p_{a'l}^{ts} \quad \forall a \in \mathcal{A}, a' \in \mathcal{R}_a, l \in \mathcal{L}, t \in \mathcal{T} \quad (18)$$

$$o_m^{ts} = \sum_{l \in \mathcal{L}} p_{ml}^{ts} \sum_{i \in \mathcal{I} \cup \mathcal{J}} y_{il}^{ts} \quad \forall m \in \mathcal{M}, l \in \mathcal{L}, t \in \mathcal{T} \quad (19)$$

$$y, p, \mu, \vartheta, o, \Delta^+, \delta^+, d^+ \geq 0 \quad (20)$$

The objective function (1) minimizes the first-stage mining costs and maximizes the expected second-stage profit. The objective function of the second-stage (9) consists of five parts: the processing and transportation costs (first term), the penalty costs (second, third, and fourth terms), and the revenues from the minerals (final products) produced (fifth term). It is assumed that all minerals

produced in period t can be sold in the same period. Penalties are incurred for failing to meet downstream facilities' requirements. Indeed, given that the blocks are sent as groups to first-stage facilities, it might be impossible to respect the facilities' capacity for certain realizations. Therefore, an additional cost (second term) is included to penalize surplus. Similarly, given that it might be impossible to meet upper bounds on the grades and proportions for certain realizations of the uncertain parameters, additional cost terms are included to penalize surplus in the attributes that need to be controlled (third and fourth term).

Constraints (2–8) define the feasible set of the first-stage variables. More specifically, constraints (2) ensure that a block is mined at most once while constraints (3) prevent a block from being mined before its predecessors. Constraints (4) impose maximal mining level at each period. Constraints (5) allow a group of blocks to be sent to one first-stage facility and only if it is *eligible* for this facility (constraints (6)).

Constraints (10) link the first-stage variables to the second-stage variables. More specifically, they define the total tonnage available at each first-stage facility under each scenario. Constraints (11) ensure that the supply from the mines is transported to a downstream facility. Constraints (12) are the flow balance at each intermediate facility and prevent stocks at these facilities. The total quantity in a facility at each period can be at most the facility capacity as specified by constraints (13) and (14). Surplus is allowed but at the expense of an additional cost.

Constraints (15) define the grade of each attribute at each first-stage facility; that is, in the material supplied from the mines. Constraints (16) are similar and apply to the intermediate and last-stage facilities. Indeed, recall that the grade is the proportion of the attribute to rock. The grade p_{aj}^{ts} of attribute a in node j under scenario s at the end of period t is thus a variable that depends on the amount of this attribute and the amount of rock sent to this node from the nodes of the previous stage. Also, recall that for each node associated with a first-stage or intermediate facility, the amount of input attribute and the output attribute are equal. Given these considerations, p_{aj}^{ts} of the material leaving node j can be calculated as follows:

$$p_{aj}^{ts} = \frac{\sum_{i \in \mathcal{I}} v_{ai}^{ts} y_{ij}^{ts} + \sum_{j' \in \mathcal{J}} p_{aj'}^{ts} y_{j'j}^{ts}}{\sum_{i \in \mathcal{I} \cup \mathcal{J}} y_{ij}^{ts}}. \quad (21)$$

These equations simplify to the bilinear constraints (16).

Constraints (17) ensure that the proportion of each attribute in the combined material in each last-stage facility lies below a specific level, and if not, a penalty cost is incurred. Note that an attribute can be a pollutant and thus these constraints handle sustainability constraints at waste dumps, tailings and slags for example. Constraints (18) are similar and take care of the proportion of an attribute a relative to another attribute $a' \in \mathcal{R}_a$. Constraints (19) define the amount of each mineral produced at each period under each scenario at the final nodes. Finally, constraints (20) are the non-negativity constraints. Indices are omitted. Although in this paper we assume that the arcs have unlimited capacities, limits on the flows can be easily accommodated without any added conceptual difficulties.

In the next section, the algorithmic strategy for addressing the intrinsic difficulties of problem (1)–(20); i.e., non-linear constraints, complex objective function, large size, is detailed.

3 Matheuristic

3.1 General structure of the algorithm

As mentioned in Section 1, to tackle the SSOMC, a matheuristic that exploits mathematical programming techniques and machine learning techniques in a heuristic framework is used. Within this algorithmic framework, the problem is separated into two sub-problems. The first one, referred to

as the upstream sub-problem, determines a mining sequence; that is, which blocks to extract at each period of the life-of-the-mine. Given a mining sequence, the second sub-problem, referred to as the downstream sub-problem, determines where to send the extracted blocks and the most profitable way to mix them so as to meet the different processing facilities' capacities and blending requirements, as well as production targets. The two sub-problems are iteratively solved until a stopping criterion is met.

Solving the upstream sub-problem involves modifying the mining sequence to explore different parts of the solution space. This is done using the hyper-heuristic proposed in Lamghari and Dimitrakopoulos (2018) that uses a set of 27 simple perturbative low-level heuristics and a score-based learning mechanism along with a tabu list to select the low-level heuristic to apply at each step of the solution process. To evaluate the mining sequences generated by the selected low-level heuristic, the downstream sub-problem, whose formulation involves binary variables and bilinear terms, must be solved. This is done by using an exact algorithm inspired by Benders decomposition (Benders, 1962). The original downstream sub-problem is reformulated into a master problem and a series of blending sub-problems. The master is a relaxation of the original problem in which McCormick envelopes are used to relax the bilinear terms and valid inequalities are used to strengthen the resulting linear formulation. At each iteration of this decomposition algorithm, the master problem is solved to obtain an upper bound on the optimal objective function value; then the blending sub-problems are solved by holding the master binary variables fixed to compute a lower bound and generate an optimality cut. This cut is added to the master and the process is repeated until a threshold optimality gap is reached. As for the upstream sub-problem, machine learning techniques are also exploited when solving the downstream sub-problem. More specifically, in order to speed up the computational times, an artificial neural network surrogate is used instead of the decomposition algorithm to approximate the objective function values and identify the best neighbor solutions generated by the upstream low-level heuristic. The decomposition algorithm is then employed only for these solutions, and its output is used to update the scores of the low-level heuristics and the training set of the neural network for the next iteration.

A more detailed description of the procedures proposed to solve the downstream sub-problem is provided in the following Section. Details about the algorithm for solving the upstream sub-problem can be found in Lamghari and Dimitrakopoulos (2018).

3.2 Solving the downstream sub-problem

Let $\bar{x} = (\bar{x}_b^t)$ be a mining sequence generated by the hyper-heuristic. Denote by:

- $w_g^{ts} = \sum_{b \in \mathcal{B}} \beta_{bg}^s w_b \bar{x}_b^t$ the total tonnage of material that belongs to group g under scenario s , available at period t
- $\lambda_{ag}^{ts} = \frac{\sum_{b \in \mathcal{B}} \beta_{bg}^s w_b \bar{x}_b^t}{w_g^{ts}}$, the grade of attribute a in the material that belongs to group g under scenario s , available at period t .

Evaluating the objective function $f(\bar{x})$ involves identifying the flow of the extracted material through the downstream part of the mining complex; i.e., solving the downstream sub-problem. Assuming that there are no stockpiles that link the periods to each other, the sub-problems associated with the periods can be solved independently. Therefore, for notational simplicity, we drop the subscript designating the period in the rest of this section. All variables and parameters should be understood as related to a particular period t .

The sub-problem associated with a period t , which we will refer to as DP, can informally be defined as follows: Given a set of groups (raw material with different properties) characterized by scenario-dependent tonnages and attribute levels (w_g^s and λ_{ag}^s , respectively) that have to be first mixed in first-stage facilities ($i \in \mathcal{I}$), then in intermediate facilities ($j \in \mathcal{J}$), and finally sent forth from the intermediate facilities to be mixed again at last-stage facilities ($l \in \mathcal{L}$) to recover final sellable

products, what is the most profitable way to mix these groups so as to meet to the best extent possible the different facilities' capacities as well as the attribute requirements at the last-stage facilities?

The DP can be viewed as a stochastic variant of the well-known pooling problem, which is a generalization of the blending problem (Gupte et al., 2017). Different formulations have been proposed in the literature for the pooling problem, one of which is extended below to account for the specificities and handle the features of the problem addressed in this paper. It is based on that proposed by Haverly (1978) and uses the so-called *quality* variables that control the concentration of each attribute in the material leaving each facility, which lead to a non-linear formulation. In what follows, we first provide the mathematical model and then its linear relaxation. In addition to the pooling problem requirements and restrictions, binary variables are required to model the flow of materials in first-stage facilities and scenario-dependent variables are required to account for the stochastic nature (uncertainty) of supply.

3.2.1 Mathematical formulation

We seek to maximize the expected profit; that is, the expected revenue from the final sellable products minus the total expected cost, which is the sum of the expected transportation and processing costs at the first-stage, intermediate and last-stage facilities and of the expected penalty costs incurred whenever the capacity, blending, and ratio constraints are violated. Using the same notation as in the previous section (Section 2) omitting the t subscript, the objective function can be expressed as follows:

$$\max \frac{1}{S} \sum_{s \in S} \left\{ \sum_{m \in \mathcal{M}} \pi_m o_m^s - \sum_{(i,j) \in \mathcal{A}} c_{ij} y_{ij}^s - \sum_{j \in \mathcal{N}} \sigma_j^+ \Delta_j^{s+} - \sum_{a \in \mathcal{A}} \sum_{l \in \mathcal{L}} c_{al}^+ \delta_{al}^{s+} - \sum_{a \in \mathcal{A}} \sum_{a' \in \mathcal{R}_a} \sum_{l \in \mathcal{L}} c_{aa'l}^+ d_{aa'l}^{s+} \right\} \quad (22)$$

The constraints are as follows:

1. *First-stage facilities constraints:* This set of constraints ensure that each group $g \in \mathcal{G}$ is sent to exactly one single first-facility for which it is eligible:

$$\sum_{i \in \mathcal{I}} \kappa_{gi} = 1 \quad \forall g \in \mathcal{G} \quad (23)$$

$$\kappa_{gi} \leq \epsilon_{gi} \quad \forall g \in \mathcal{G}, i \in \mathcal{I} \quad (24)$$

2. *Flow conservation constraints:* These constraints ensure that the amount entering and leaving each first stage facility $i \in \mathcal{I}$ and each intermediate facility $j \in \mathcal{J}$ under each scenario $s \in S$ are the same:

$$\sum_{g \in \mathcal{G}} w_g^s \kappa_{gi} = \sum_{j \in \mathcal{J} \cup \mathcal{L}} y_{ij}^s \quad \forall i \in \mathcal{I}, s \in S \quad (25)$$

$$\sum_{i \in \mathcal{I} \cup \mathcal{J}} y_{ij}^s = \sum_{l \in \mathcal{J} \cup \mathcal{L}} y_{jl}^s \quad \forall j \in \mathcal{J}, s \in S \quad (26)$$

3. *Capacity constraints:* These constraints limit the quantity received at each facility $n \in \mathcal{N}$ under each scenario $s \in S$ to the facility's capacity; otherwise, an extra cost $\sigma_n^+ \Delta_n^{s+}$ is incurred:

$$\sum_{j \in \mathcal{J} \cup \mathcal{L}} y_{ij}^s - \Delta_i^{s+} \leq \overline{C}_i \quad \forall i \in \mathcal{I}, s \in S \quad (27)$$

$$\sum_{l \in \mathcal{J} \cup \mathcal{L}} y_{jl}^s - \Delta_j^{s+} \leq \overline{C}_j \quad \forall j \in \mathcal{J}, s \in S \quad (28)$$

$$\sum_{i \in \mathcal{I} \cup \mathcal{J}} y_{il}^s - \Delta_l^{s+} \leq \overline{C}_l \quad \forall l \in \mathcal{L}, s \in S \quad (29)$$

4. *Grade definition at the first-stage and intermediate facilities:* Assuming that at least one group whose tonnage is non-null has been sent to first-stage facility $i \in \mathcal{I}$ (i.e., that $\sum_{g \in \mathcal{G}} w_g^s \kappa_{gi} \neq 0$) and that the mixing process follows a linear blending, the grade of attribute $a \in \mathcal{A}$ in the material leaving node i under scenario s is given by:¹

$$p_{ai}^s = \frac{\sum_{g \in \mathcal{G}} \lambda_{ag}^s w_g^s \kappa_{gi}}{\sum_{g \in \mathcal{G}} w_g^s \kappa_{gi}}.$$

This expression can be equivalently rewritten in a bilinear form as:

$$\sum_{g \in \mathcal{G}} \lambda_{ag}^s w_g^s \kappa_{gi} = p_{ai}^s \sum_{g \in \mathcal{G}} w_g^s \kappa_{gi} \quad \forall a \in \mathcal{A}, i \in \mathcal{I}, s \in S \quad (30a)$$

which can also be rewritten, given constraints (25), as follows:

$$\sum_{g \in \mathcal{G}} \lambda_{ag}^s w_g^s \kappa_{gi} = p_{ai}^s \sum_{j \in \mathcal{I} \cup \mathcal{L}} y_{ij}^s \quad \forall a \in \mathcal{A}, i \in \mathcal{I}, s \in S \quad (30b)$$

Similarly, at the intermediate facilities $j \in \mathcal{J}$, we have:

$$\sum_{i \in \mathcal{I}} p_{ai}^s y_{ij}^s + \sum_{j' \in \mathcal{J}} p_{aj'}^s y_{j'j}^s = p_{aj}^s \sum_{i \in \mathcal{I} \cup \mathcal{J}} y_{ij}^s \quad \forall a \in \mathcal{A}, j \in \mathcal{J}, s \in S \quad (31a)$$

which can also be rewritten as follows, considering constraints (26):

$$\sum_{i \in \mathcal{I}} p_{ai}^s y_{ij}^s + \sum_{j' \in \mathcal{J}} p_{aj'}^s y_{j'j}^s = p_{aj}^s \sum_{l \in \mathcal{J} \cup \mathcal{L}} y_{jl}^s \quad \forall a \in \mathcal{A}, j \in \mathcal{J}, s \in S \quad (31b)$$

5. *Blending constraints:* These constraints model the mixing process requirements at each last-stage facility $l \in \mathcal{L}$. Constraints (32) limit the grade of attribute $a \in \mathcal{A}$ in the material received at l to the facility's upper bound, $\overline{P}_{al}$, making it impossible to have more than what would exceed this upper bound unless a unit cost of c_{al}^+ is paid. Constraints (33) model the ratio requirements, ensuring that in the material processed at l , the ratio between attribute $a \in \mathcal{A}$ to related attribute $a' \in \mathcal{R}_a$ does not exceed the upper bound $\overline{R}_{aa'l}$; otherwise, a cost $c_{aa'l}^+$ has to be paid for each unit in surplus:

$$\sum_{i \in \mathcal{I}} p_{ai}^s y_{il}^s + \sum_{j \in \mathcal{J}} p_{aj}^s y_{jl}^s - \delta_{al}^{s+} \leq \overline{P}_{al} \sum_{i \in \mathcal{I} \cup \mathcal{J}} y_{il}^s \quad \forall a \in \mathcal{A}, l \in \mathcal{L}, s \in S \quad (32)$$

$$\sum_{i \in \mathcal{I}} p_{ai}^s y_{il}^s + \sum_{j \in \mathcal{J}} p_{aj}^s y_{jl}^s - d_{aa'l}^{s+} \leq \overline{R}_{aa'l} \left(\sum_{i \in \mathcal{I}} p_{a'i}^s y_{il}^s + \sum_{j \in \mathcal{J}} p_{a'j}^s y_{jl}^s \right) \quad \forall a \in \mathcal{A}, a' \in \mathcal{R}_a, \forall l \in \mathcal{L}, s \in S \quad (33)$$

6. *Quantities produced:* The amount produced of each sellable final product $m \in \mathcal{M} \subset \mathcal{A}$ under each scenario $s \in S$ is given by the total amount recovered in the different last-stage facilities:

$$o_m^s = \sum_{l \in \mathcal{L}} r_{ml} \left(\sum_{i \in \mathcal{I}} p_{mi}^s y_{il}^s + \sum_{j \in \mathcal{J}} p_{mj}^s y_{jl}^s \right) \quad \forall m \in \mathcal{M}, s \in S \quad (34)$$

7. *Integrality and non-negativity constraints:*

$$\kappa_{gi} \in \{0, 1\} \text{ and } y, p, o, \Delta^+, \delta^+, d^+ \geq 0. \quad (35)$$

¹Note that, in this paper, it is assumed that the recovery factor at the first-stage and intermediate facilities is equal to 1, while it is smaller than 1 in the last-stage facilities. Different values for the recovery factor can be easily accommodated without any added conceptual difficulties.

3.2.2 Linear relaxation of the bilinear terms

As mentioned earlier, the problem described above is a stochastic variant of the pooling problem, which is a nonconvex, strongly NP-hard bilinear problem (Gupte et al., 2017). It can be relaxed by linearizing the bilinear terms in constraints (30)–(34). Recall that the bilinear constraints defining the grade at first-stage and intermediate facilities can be written in two different forms. For first-stage facilities i , constraints (30a) will be used instead of constraints (30b) as the former allow an exact linearization as explained next.² For intermediate facilities j , constraints (31b) will be used instead of (31a) in order to have all the non-linearities in the model in the form $p_{ai}^s y_{ij}^s$ (using constraints (31a) would introduce bilinear terms of the form $p_{aj}^s y_{ij}^s$). Before describing the proposed linear relaxation techniques, some extra notation and definitions that will be used through this section to facilitate further discussions must first be introduced.

For every facility $j \in \mathcal{I} \cup \mathcal{J}$, let $\mathcal{G}_j$ be the subset of groups that might end up being processed at j . This means that for intermediate facilities $j \in \mathcal{J}$, $\mathcal{G}_j = \{g \in \mathcal{G} : \text{there exists at least a first-stage facility } i \in \mathcal{I} \text{ to which } g \text{ is eligible and from which there exists a path in the graph to } j\}$, while for first-stage facilities $i \in \mathcal{I}$, $\mathcal{G}_i$ reduces to the subset of groups eligible to i ($\mathcal{G}_i = \{g \in \mathcal{G} : \epsilon_{gi} = 1\}$). Also, for every attribute $a \in \mathcal{A}$, every facility $j \in \mathcal{I} \cup \mathcal{J}$, and every scenario $s \in S$, let $\underline{\Lambda}_{aj}^s = \min_{g \in \mathcal{G}_j} \lambda_{ag}^s$ be a lower bound on the grade of attribute a in the material leaving j under scenario s . Note that this lower bound is valid since the flow leaving j originates at the groups $g \in \mathcal{G}_j$. The upper bound is denoted by $\overline{\Lambda}_{aj}^s = \max_{g \in \mathcal{G}_j} \lambda_{ag}^s$. Finally, for each arc $(i, j) \in A$ and each scenario $s \in S$, define $\overline{Y}_{ij}^s = \sum_{g \in \mathcal{G}_i} w_g^s$ to be an upper bound on the flow along arc (i, j) . We assume the lower bound on flow on arc (i, j) to be zero (i.e., $\underline{Y}_{ij}^s = 0 \forall (i, j) \in A, s \in S$).

Using the notation above, constraints (30a) can be linearized by substituting every bilinear term $p_{ai}^s \kappa_{gi}$ with a new variable u_{agi}^s and adding the following constraints:

$$u_{agi}^s \leq \overline{\Lambda}_{ai}^s \kappa_{gi} \quad (36a)$$

$$u_{agi}^s \leq p_{ai}^s \quad (36b)$$

$$u_{agi}^s \geq p_{ai}^s - (1 - \kappa_{gi}) \overline{\Lambda}_{ai}^s \quad (36c)$$

$$u_{agi}^s \geq 0 \quad (36d)$$

This way, if κ_{gi} is zero, then inequalities (36a) and (36d) ensure that u_{agi}^s will be zero as well, and the other inequalities are redundant. On the other hand, if κ_{gi} is equal to 1, inequalities (36b) and (36c) imply that u_{agi}^s will be p_{ai}^s , exactly as required. Note that in this case, inequality (36a) is redundant since it ensures that $u_{agi}^s = p_{ai}^s$ is less than $\overline{\Lambda}_{ai}^s = \max_{g \in \mathcal{G}_i} \lambda_{ag}^s$, which is true by definition.

The standard McCormick approximation (McCormick, 1976) can be used to relax the bilinear terms in constraints (31b) and (32)–(34), therefore obtaining a Mixed Integer Linear Programming Problem. This approach consists of introducing new variables and adding linear constraints that tie these variables close to the bilinear terms that they replace. More precisely, any bilinear term xy , where $x \in [\underline{x}, \overline{x}]$ and $y \in [\underline{y}, \overline{y}]$, is replaced by a new variable v . This new variable is linked to the variables x and y via the following set of constraints:

$$v \geq \underline{x}y + \underline{y}x - \underline{x}\underline{y} \quad (37a)$$

$$v \geq \overline{x}y + \overline{y}x - \overline{x}\overline{y} \quad (37b)$$

$$v \leq \overline{x}y + \underline{y}x - \overline{x}\underline{y} \quad (37c)$$

$$v \leq \underline{x}y + \overline{y}x - \underline{x}\overline{y} \quad (37d)$$

It is worth noting that the tightness of the relaxation depends on the tightness of the variables bounds. It has been shown in Androulakis et al. (1995) that the maximum difference between the

²Recall that the bilinear terms in (30a) are the product of a binary and a continuous variable as opposed to the product of two continuous variables in (30b).

variable v and the bilinear term xy is $d_{\max} = \frac{1}{4}(\bar{x} - \underline{x})(\bar{y} - \underline{y})$. This explains why constraints (30a) were chosen over constraints (30b) to model the grade definition constraints at first-stage facilities.

Without loss of generality, consider the bilinear terms $p_{ai}^s y_{ij}^s$ in constraints (31b). Using the notation introduced at the beginning of this section, the McCormick inequalities reduce to the following:

$$v_{aij}^s \geq \underline{\Lambda}_{ai}^s y_{ij}^s \quad (38a)$$

$$v_{aij}^s \geq \overline{\Lambda}_{ai}^s y_{ij}^s + \overline{Y}_{ij}^s p_{ai}^s - \overline{\Lambda}_{ai}^s \overline{Y}_{ij}^s \quad (38b)$$

$$v_{aij}^s \leq \underline{\Lambda}_{ai}^s y_{ij}^s \quad (38c)$$

$$v_{aij}^s \leq \underline{\Lambda}_{ai}^s y_{ij}^s + \overline{Y}_{ij}^s p_{ai}^s - \underline{\Lambda}_{ai}^s \overline{Y}_{ij}^s \quad (38d)$$

where v_{aij}^s is defined as:

$$v_{aij}^s = p_{ai}^s y_{ij}^s. \quad (38e)$$

Note that the variable v_{aij}^s can be interpreted as the total amount of attribute a in the flow along arc (i, j) under scenario s . With this definition, constraints (31b), defining the grade at intermediate facilities $j \in \mathcal{J}$, simply state that the amounts of attribute a entering and leaving node (facility) j are equal.

3.2.3 Decomposition algorithm

As mentioned earlier, to solve DP (the downstream sub-problem), an exact algorithm, inspired by Benders decomposition (Benders, 1962), is proposed. Benders decomposition can be summarized as follows: The original problem is reformulated into a master and a number of sub-problems. At each iteration, the values of the master problem variables are first determined, and the sub-problems are solved by holding these variables fixed. If all the sub-problems are feasible and bounded, an optimal cut is added to the master problem; otherwise, a feasibility cut is added. A lower bound can be computed from the sub-problems and an upper bound is obtained if the master problem is solved to optimality. The process continues until an optimal solution is found or until the optimality gap is smaller than a given threshold value. In the context of stochastic integer programming literature, this algorithmic scheme is also known as the L-shaped method (Birge and Louveaux, 2011). In general, there are two decomposition strategies: one that projects out all variables of the subproblems from the master, and one that does not project out any variables. In this paper, the second strategy is used. Namely, the master formulation includes all variables. It also includes linear approximations of the non-linear constraints (all bilinear terms are replaced with the corresponding McCormick envelopes or exact linearization, described in the previous section). The master is thus a mixed-integer linear programming relaxation of the original problem. During the course of the algorithm, additional constraints (Benders cuts) are also included in the formulation of the master problem to obtain a better approximation of the objective function value as discussed in detail below.

Define:

- θ : an approximation of the objective function value of the original problem DP (the non-linear stochastic mixed integer program).
- U : an upper bound on the value of the objective function value of the original problem DP described in Section 3.3.1. Since DP is bounded, U is finite.
- $\mathcal{R}$: the set of first-stage feasible solutions, indexed by r . Note that since first-stage decision variables (κ_{gi}) are binary, the set $\mathcal{R}$ is finite.

Moreover, for every r , let:

- θ_r : a constant equal to the expected second stage value of the r^{th} solution. To compute θ_r , the κ_{gi} are fixed as in the r^{th} solution and the so-obtained S blending problems are solved to get

the values of the other variables. Let $(y_{ij}^{s*}, o_m^{s*}, \Delta_j^{s*+}, \delta_{al}^{s*+}, d_{aa'l}^{s*+})$ be the optimal solution of the blending problem associated with scenario s . Then θ_r is calculated as follows:

$$\begin{aligned} \theta_r = \frac{1}{S} \sum_{s \in S} \left\{ \sum_{m \in \mathcal{M}} \pi_m o_m^{s*} - \sum_{(i,j) \in A} c_{ij} y_{ij}^{s*} \right. \\ \left. - \sum_{j \in \mathcal{N}} \sigma_j^+ \Delta_j^{s*+} - \sum_{a \in \mathcal{A}} \sum_{l \in \mathcal{L}} c_{al}^+ \delta_{al}^{s*+} - \sum_{a \in \mathcal{A}} \sum_{a' \in \mathcal{R}_a} \sum_{l \in \mathcal{L}} c_{aa'l}^+ d_{aa'l}^{s*+} \right\} \end{aligned} \quad (39)$$

S_r : the subset of indices (g, i) such as $\kappa_{gi} = 1$ in the r^{th} feasible solution, in other words, $S_r = \{(g, i) : g \in \mathcal{G}, i \in \mathcal{I} \text{ and } g \text{ has been sent to first-stage facility } i \text{ in the } r^{th} \text{ feasible solution}\}$.

Using this notation, the master problem can be formally written as:

$$\max \theta \quad (40)$$

$$\begin{aligned} \theta \leq \frac{1}{S} \sum_{s \in S} \left\{ \sum_{m \in \mathcal{M}} \pi_m o_m^s - \sum_{(i,j) \in A} c_{ij} y_{ij}^s - \sum_{j \in \mathcal{N}} \sigma_j^+ \Delta_j^{s+} \right. \\ \left. - \sum_{a \in \mathcal{A}} \sum_{l \in \mathcal{L}} c_{al}^+ \delta_{al}^{s+} - \sum_{a \in \mathcal{A}} \sum_{a' \in \mathcal{R}_a} \sum_{l \in \mathcal{L}} c_{aa'l}^+ d_{aa'l}^{s+} \right\} \end{aligned} \quad (41)$$

$$\begin{aligned} \theta \leq - (U - \theta_r) \left(\sum_{(g,i) \in S_r} \kappa_{gi} - \sum_{(g,i) \notin S_r} \kappa_{gi} \right) \\ + (U - \theta_r) (|S_r| - 1) + U \quad \forall r \in \mathcal{R} \end{aligned} \quad (42)$$

$$\sum_{i \in \mathcal{I}} \kappa_{gi} = 1 \quad \forall g \in \mathcal{G} \quad (43)$$

$$\kappa_{gi} \leq \epsilon_{gi} \quad \forall g \in \mathcal{G}, i \in \mathcal{I} \quad (44)$$

$$\sum_{g \in \mathcal{G}} w_g \kappa_{gi} = \sum_{l \in \mathcal{L}} y_{il}^s \quad \forall i \in \mathcal{I}, s \in S \quad (45)$$

$$\sum_{l \in \mathcal{L}} y_{il}^s - \Delta_i^{s+} \leq \overline{C_i} \quad \forall i \in \mathcal{I}, s \in S \quad (46)$$

$$\sum_{i \in \mathcal{I}} y_{il}^s - \Delta_l^{s+} \leq \overline{C_l} \quad \forall l \in \mathcal{L}, s \in S \quad (47)$$

$$\sum_{g \in \mathcal{G}} \lambda_{ag}^s w_g \kappa_{gi} = \sum_{g \in \mathcal{G}} w_g u_{agi} \quad \forall a \in \mathcal{A}, i \in \mathcal{I}, s \in S \quad (48)$$

$$\sum_{i \in \mathcal{I}} v_{ail}^s - \delta_{al}^{s+} \leq \overline{P_{al}} \sum_{i \in \mathcal{I}} y_{il}^s \quad \forall a \in \mathcal{A}, l \in \mathcal{L}, s \in S \quad (49)$$

$$\sum_{i \in \mathcal{I}} v_{ail}^s - d_{aa'l}^{s+} \leq \overline{R_{aa'l}} \sum_{i \in \mathcal{I}} v_{a'il}^s \quad \forall a \in \mathcal{A}, a' \in \mathcal{R}_a, l \in \mathcal{L}, s \in S \quad (50)$$

$$o_m^s = \sum_{l \in \mathcal{L}} r_{ml} \sum_{i \in \mathcal{I}} v_{mil}^s \quad \forall m \in \mathcal{M}, s \in S \quad (51)$$

$$u_{agi}^s \leq \overline{\Lambda_{ai}^s} \kappa_{gi} \quad \forall a \in \mathcal{A}, g \in \mathcal{G}, i \in \mathcal{I}, s \in S \quad (52)$$

$$u_{agi}^s \leq p_{ai}^s \quad \forall a \in \mathcal{A}, g \in \mathcal{G}, i \in \mathcal{I}, s \in S \quad (53)$$

$$u_{agi}^s \geq p_{ai}^s - (1 - \kappa_{gi}) \overline{\Lambda_{ai}^s} \quad \forall a \in \mathcal{A}, g \in \mathcal{G}, i \in \mathcal{I}, s \in S \quad (54)$$

$$v_{ail}^s \geq \underline{\Lambda}_{ai}^s y_{il}^s \quad \forall a \in \mathcal{A}, (i, l) \in A, s \in S \quad (55)$$

$$v_{ail}^s \geq \underline{\Lambda}_{ai}^s y_{il}^s + \overline{Y}_{il}^s p_{ai}^s - \overline{\Lambda}_{ai}^s \overline{Y}_{il}^s \quad \forall a \in \mathcal{A}, (i, l) \in A, s \in S \quad (56)$$

$$v_{ail}^s \leq \overline{\Lambda}_{ai}^s y_{il}^s \quad \forall a \in \mathcal{A}, (i, l) \in A, s \in S \quad (57)$$

$$v_{ail}^s \leq \underline{\Lambda}_{ai}^s y_{il}^s + \overline{Y}_{il}^s p_{ai}^s - \underline{\Lambda}_{ai}^s \overline{Y}_{il}^s \quad \forall a \in \mathcal{A}, (i, l) \in A, s \in S \quad (58)$$

$$\kappa_{gi} \in \{0, 1\} \text{ and } y, p, o, \Delta^+, \delta^+, d^+, u, v, \theta \geq 0. \quad (59)$$

The constraints of the master can be divided into five categories: upper bounds on the objective function, capacity constraints, upper bounds on the grade and ratio, approximation of the bilinear terms, and integrality and non-negativity constraints. While the last four categories are self-explanatory given the discussion in the previous sections, the constraints in the first category (constraints (41) and (42)) deserve some discussion.

Recall that variable θ gives an approximation of the objective function value of DP (the non-linear stochastic mixed integer program). Constraints (41) simply state that θ can be at most the objective function value; that is, it has to be chosen in a way that is coherent with the values of the other variables of the master. Constraints (42) are similar to those introduced by Laporte and Louveaux (1993) in the context of stochastic integer programs with complete recourse, and they are gradually added to the master during the solution procedure to ensure that, if an already visited first-stage solution r is selected, then a reward equal to θ_r is incurred. This reward is nothing else but the exact value of this solution, which is known since it has been already calculated during the previous iterations of the algorithm. Otherwise, there are no restrictions on θ except (41) and non-negativity. Note that θ is non-negative because it is always possible to send all groups to the waste dump and get a solution of value 0 (no costs, no revenues, no constraints to be violated since it is assumed that the waste dump has an unlimited capacity).

As discussed above, the master provides an upper bound. Moreover, once solved, a feasible solution to the original non-linear formulation can be determined. This is done as follows. Variables κ_{gi} are fixed, which allows setting and fixing variables p_{ai}^s as well (recall that $p_{ai}^s = \frac{\sum_{g \in \mathcal{G}} \lambda_{ag}^s w_g^s \kappa_{gi}}{\sum_{g \in \mathcal{G}} w_g^s \kappa_{gi}}$), and in turn eliminating all bilinearities from the model. The resulting Benders sub-problem can be separated by scenario and reduces to independent problems that have the structure of a blending problem and can be solved as linear programs (LPs). Due to the presence of the variables Δ_i^{s+} , Δ_l^{s+} , δ_{al}^{s+} and $d_{aa'l}^{s+}$ the sub-problems are always feasible. Furthermore, since the objective function coefficients (costs and profit parameters) are finite, any feasible solution must be bounded.

A formal statement of the proposed decomposition algorithm is given below.

Step 0 (Initialization) Set $r := 0$, $R := \emptyset$, $\overline{Z} = UB := \infty$ or an appropriate upper bound, $\underline{Z} = LB := -\infty$ or the value of the best-known feasible solution. Specify an optimality gap or tolerance ϵ .

Step 1 (Solving the master problem) Set $r := r + 1$ and solve the current master problem. Let $(\kappa^r, y^r, p^r, o^r, \Delta^{+r}, \delta^{+r}, d^{+r}, u^r, v^r, \theta^r)$ be the so-obtained optimal solution and U^r be the value of this solution. Set $\overline{Z} := U^r$.

Step 2 (Solving the Benders sub-problem) Fix the first-stage variables κ to κ^r , solve the resulting S blending sub-problems, and compute the expected true value θ^r of this r^{th} feasible solution.

Step 3 (Updating the bounds and the incumbent) If $\theta^r > \underline{Z}$, then set $\underline{Z} := \theta^r$ and update the best-known feasible solution (the incumbent).

Step 4 (Verifying the stopping criterion and updating the master) If $\frac{\overline{Z} - \underline{Z}}{\underline{Z}} \leq \epsilon$, stop and return the best-known feasible solution. Otherwise, include r in $\mathcal{R}$, add to the master the optimality cut (42) and return to Step 1.

The algorithm summarized above yields a global optimal solution, within ϵ optimality gap, in a finite number of iterations. The convergence of the algorithm is guaranteed by the finite number of feasible first-stage solutions.

3.2.4 Using a surrogate

As can be seen from the results in the next section (Section 4), the proposed decomposition algorithm converges rapidly. However, using it to evaluate each mining sequence (neighbor solution) generated using the hyper-heuristic might not be the most suitable approach as it would lead to prohibitive computational times (recall that computing the objective function of any neighbor solution entails solving the associated downstream sub-problem). Instead, a surrogate model can be used to provide fast approximations of the objective function and thus speed up the search. We propose using machine learning techniques, namely an artificial neural network (ANN) as a surrogate model (surrogate evaluation function). The most promising neighbor solutions; that is, those achieving the best value for the surrogate, are afterward evaluated using the decomposition algorithm.

The ANN learns how to approximate the objective function value based on the inputs, which are the characteristics of a mining sequence such as group tonnages, attribute grades, etc. It requires a training set. At the beginning of the algorithm, all neighbor solutions are evaluated using the exact objective function (i.e., by solving the downstream sub-problem using the decomposition algorithm). The training set is initialized with information gathered from such solutions. After a few iterations, the surrogate function is used instead of the exact objective function, which, as aforementioned, allows us to sample a subset of the neighborhood for exact evaluation. Not only does the exact evaluation of this subset allow identification of the new current solution, but it also provides new data for the training set.

4 Preliminary computational results

In this section, the efficiency and the robustness of the decomposition algorithm described in Section 3.2.3 are assessed. Numerical experiments have been performed on the Escondida-Norte mining complex, which is briefly described below. The decomposition algorithm was coded in C++ using IBM ILOG CPLEX 12.6.1. The experiments were run on an Intel(R) Xeon(R) CPU X5675 computer (3.07 GHz) with 96 Go of RAM running under Linux.

4.1 Dataset: The Escondida-Norte mining complex

The Escondida-Norte mining complex comprises two mines, five crushers, three mills, one leach pad, one bioleach, and one waste dump. As can be seen from Figure 1, material comes out of the mines and goes to one of the five crushers or to the bioleach or the waste dump. It gets further processed in one of the three mills or in a leach pad and ends up at the port or as copper cathodes. Note that the main sellable final product is copper, but the two mines also produce gold, iron, and molybdenum.

There are six main material types (sulphide high grade, sulphide low grade, oxide high grade, oxide low grade, mixed, and waste) and seven attributes of interest (copper total, copper soluble, gold, silver, arsenic, iron, and molybdenum). A set of 15 equiprobable scenarios is used to model geological uncertainty (grades and material types).

Of the Escondida-mining complex's two mines, the first one (mine 1) has 68,450 blocks, while the second one (mine 2) has 103,390 blocks. In total, there are 171,840 blocks to be scheduled over 8 years. In addition to the slope, eligibility, mining and processing capacity constraints, there are also blending and ratio constraints that must be taken into account. Namely, at the mills, the level of arsenic, iron and the ratio of copper total to iron need to be controlled, while at the bioleach, the ratio of copper soluble to copper total must be controlled.

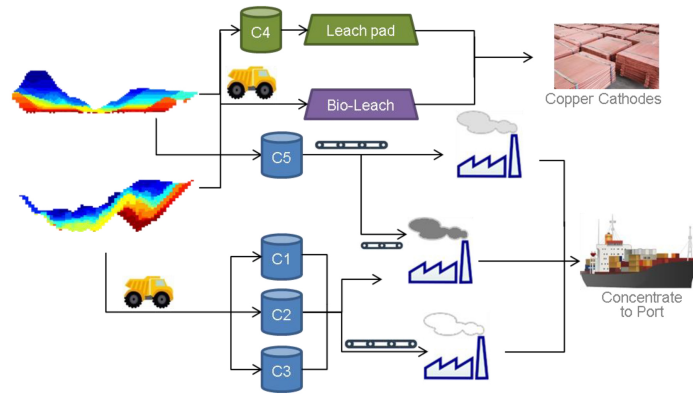


Figure 1: Schematic representation of the Escondida-Norte mining complex

4.2 Evaluation of the decomposition algorithm

In this section, results obtained with the decomposition algorithm are reported. For each mine, a mining sequence was generated using the random heuristic proposed in Lamghari and Dimitrakopoulos (2012). The decomposition algorithm was applied to determine the flow of the extracted material.

Table 1: Summary of the computational times in seconds obtained when considering each of the two mines individually and simultaneously

Period	CPU time (in seconds)		
	Mine 1	Mine 2	Mine 1 and 2
1	1.91	2.48	5.4
2	1.73	2.96	4.43
3	1.26	3.66	7.99
4	1.73	2.59	3.94
5	1.68	3.18	8.47
6	1.21	3.29	4.77
7	0.37	3.84	8.21
8	1.02	10.09	5.71

As can be seen from the results in Table 1, the decomposition algorithm converges quite rapidly. It obtains an optimal solution of the downstream sub-problem in few seconds. A comparison of the results in columns 2, 3, and 4 of this table show that the algorithm scales well with size. The difference in solution time can be explained by the number of major iterations of the decomposition algorithm, or equivalently, the number of cuts added to the master problem during the course of the algorithm. The number of cuts generated when solving the instance with two mines is given in Table 2. Note that in our implementation, instead of solving one master at each iteration, a *branch-and-Benders-cut method*, also known as *Benders-based branch-and-cut method*, was used. In this method, a single branch-and-cut tree is constructed, and the cuts are added to this tree during its exploration.

Table 2: Number of cuts generated when considering blocks extracted from both mines

Period	Number of cuts
1	36
2	27
3	59
4	21
5	58
6	25
7	46
8	44

5 Conclusions

This paper deals with the development of new models and solution approaches to address the large and complex problems faced by the mining industry when optimizing mineral value chains (mining complexes) under uncertainty. A general mathematical model that integrates geological uncertainty and multiple mines sharing the same downstream infrastructure has been proposed to simultaneously optimize mining decisions and the flow of the extracted material. A new methodology that integrates components from exact algorithms (relaxation and decomposition), machine learning techniques (reinforcement learning and artificial neural networks), and heuristics (local improvement and randomized search) has been developed to efficiently solve the proposed formulation.

The proposed approach opens doors for new methodologies that capitalize on the synergies between machine learning and optimization techniques. Future work will follow this direction.

6 References

- Androulakis, I., Maranas, C., Floudas, C. (1995) α bb: a global optimization method for general constrained nonconvex problems. *Journal of Global Optimization*, 7: 337–363.
- Benders, J. F. (1962) Partitioning procedures for solving mixed variables programming problems. *Numerische Mathematik*, 4(1): 238–252.
- Birge, J., Louveaux, F. (2011) *Introduction to stochastic programming*, Second Edition. Springer.
- Del Castillo, M.F., Dimitrakopoulos, R. (2019) Dynamically optimizing the strategic plan of mining complexes under supply uncertainty. *Resources Policy*, 60: 83–93.
- Godoy, M., Dimitrakopoulos, R. (2004) Managing risk and waste mining in long-term production scheduling. *SME Transactions* 316: 43–50.
- Goodfellow, R., Dimitrakopoulos, R. (2016) Global optimization of open pit mining complexes with uncertainty. *Applied Soft Computing* 40: 292–304.
- Goodfellow, R., Dimitrakopoulos, R. (2017) Simultaneous stochastic optimization of mining complexes and mineral value chains. *Mathematical Geosciences* 49(3): 341–360.
- Gupte, A., Ahmed, S., Dey, S., Cheon, M. (2017) Relaxations and discretizations for the pooling problem. *Journal of Global Optimization*, 7: 337–363.
- Haverly, C. (1978) Studies of the behavior of recursion for the pooling problem. *ACM SIGMAP Bull.*, 25: 19–28.
- Laporte, G., Louveaux, F.V. (1993) The integer L-shaped method for stochastic integer programs with complete recourse. *Operations Research Letters*, 13: 133–142.
- Lamghari, A., Dimitrakopoulos, R. (2012) A diversified tabu search approach for the open-pit mine production scheduling problem with metal uncertainty. *European Journal of Operational Research*, 222: 642–652.
- Lamghari, A., Dimitrakopoulos, R. (2018) Hyper-heuristic approaches for strategic mine planning under uncertainty. *Computers & Operations Research*, doi.org/10.1016/j.cor.2018.11.010.
- McCormick, G. (1976) Computability of global solutions to factorable nonconvex programs: Part I – convex underestimating problems. *Mathematical Programming*, 10: 147–175.
- Montiel, L., Dimitrakopoulos, R. (2015) Optimizing mining complexes with multiple processing and transportation alternatives: An uncertainty-based approach. *European Journal of Operational Research*, 247(1): 166–178.
- Montiel, L., Dimitrakopoulos, R., Kawahata, K. (2016) Globally optimising open-pit and underground mining operations under geological uncertainty. *Mining Technology*, 125(1): 2–14.
- Montiel, L., Dimitrakopoulos, R. (2018) Simultaneous stochastic optimization of production scheduling at Twin Creeks mining complex, Nevada. *Mining Engineering*, 70(12): 48–56.

Zhang, J., Nault, B.R., Dimitrakopoulos, R. (2019) Optimizing a mineral value chain with market uncertainty using benders decomposition. *European Journal of Operational Research*, 274(1): 227-239.