The risk-averse ultimate pit problem*

Gianpiero Canessa† Eduardo Moreno‡
Bernardo K. Pagnoncelli§

Abstract

In this work, we consider a risk-averse ultimate pit problem where the grade of the mineral is uncertain. We propose a two-stage formulation of the problem and discuss which properties are desirable for a risk measure in this context. We show that the only risk measure that satisfies these properties is the entropic. We propose an efficient approximation scheme to solve the risk-averse version of the problem and show its viability in large-scale mines.
Keywords: Ultimate Pit; Mining; Risk-Averse Optimization; Integer programming.

1 Introduction

A fundamental problem in open-pit mining is the determination of the ultimate pit (UP), which consists in finding the contour of the mine that maximizes the difference between profits obtained from minerals minus extraction costs. The problem is formulated on a representation of the mine into blocks of a certain size, taking into account engineering requirements such as slope and precedence constraints. The formulation of the UP problem can traced back to 1965: the seminal paper of [8], which was one of the first to propose an algorithm to solve this problem, and their methodology has been widely used in the mining industry.

The UP problem is relevant in practice mainly for two reasons: first, it gives an estimate of the total mineral that can be potentially extracted from an economical point of view. Second, the solution allows one to generate the so-called *nested pits*: by solving the problem for different prices (or "revenue factors") of the mineral it is possible to obtain a sequence of pits such that higher prices generate larger pits that contain the previous ones. Such nested pits are the input for heuristics capable of generating a long term order of extraction

*Pagnoncelli, Moreno and Canessa thank the support of Anillo grant ACT 1407 and Codelco Tech Grant Piensa Minería 2018
†Diagonal las Torres 2640, Santiago, Chile 7941169. Contact: gianpiero.canessa@edu.uai.cl
‡Diagonal las Torres 2640 oficina 304a-D, Santiago, Chile 7941169.
§Diagonal las Torres 2640 oficina 533-C, Santiago, Chile 7941169.

of different sectors of the mine. Moreover, they provide an estimation of the different sizes of the pit under less favorable price conditions.

The vast majority of work in the UP problem assumes the parameters of the problem are known. In particular, the ore distribution is inferred by drilling, using techniques from geostatistics, such as *kriging* [4]. Such estimates are usually an approximation, and determining the ultimate pit by replacing the ore distribution by a single number per block, such as the average grade, can have negative consequences. There are few works that incorporate the distribution of ore grades into the UP problem. In [9] the authors use expected profits in the objective function, and conclude that larger profits can be obtained by incorporating uncertainty into the UP formulation with respect to the deterministic case, and that the relative gains of the stochastic approach increase with the treatment costs. In [6] the authors move one step further and study the risk-averse UP problem, replacing the expected value by risk measures such as the Conditional Value-at-Risk (CVaR) [11] which has been widely used in applications such as finance and energy.

We propose a detailed study of the risk-averse UP problem, assuming that the grade of each block follows a known distribution. First, we discuss desirable properties—risk-nestedness and additive consistency—that a risk measure should have in defining the UP, and show that the only risk measure that satisfies those properties is the entropic risk measure. In particular, the CVaR does not satisfy any of these properties and is probably not the best choice to capture risk-aversion in the mining context. Second, we derive conditions under which the entropic risk measure generates nested pits by varying the *risk-aversion level* of the decision maker.

We apply our methodology to a small case study of a self-constructed instance to illustrate the gains of our approach. We also validate methodology on a real-world mine. Furthermore, we show how ore grade uncertainty can drastically change the solutions obtained, while the current methodology remains stoic to the stochastic nature of the ore grade.

The structure of this paper is as follows: Section 2 is an overall view of the UP problem, and we present a two-stage stochastic formulation as well as summarize the concept of nested pits. In Section 3 we define risk in the UP problem and propose desirable properties risk measures should have. In Section 4 we show the results of our methodology using a small instance to visualize the differences between the our and the classic methodology, and using a real mine to show the potential of its use in real life instances. Section 5 concludes the work and points out futures avenues of research.

2 Ultimate Pit problem and nested pits

2.1 UP problem

2.1.1 Deterministic version

The UP problem consists in finding the last set of blocks to be extracted in order to maximize the value of the mine, respecting precedence constraints. It is assumed that capacity is unlimited and that the ore can be mined immediately, in the sense that time considerations are left out of the formulation. In other words, the UP solution gives the contour of the mine without specifying *when* the blocks will be extracted.

We will define the parameters which will be used in the paper.

- B: set of blocks of the mine.
- $P \subseteq B \times B$: set of precedences for every block in the mine, that is, if $(b, b') \in P$ then in order to extract block b we must extract block b'.
- c_b^e: cost of extracting block b.
- c_b^p: cost of processing block b.
- r_b: profit for processing block b in the case that its grade is 1.0.
- g_b: ore grade of block b.

Parameter r_b includes, among other factors, the tonnage of the block, the recovery of the metal after processing the block, and the price of the metal in the market. As a convention, $c^e := \{c_b^e\}_{b \in B}$ and $c^p := \{c_b^p\}_{b \in B}$ will be defined as the vector of all extracting costs and processing costs for every block $b \in B$, respectively.

We can formulate the UP problem as the following mixed-integer optimization problem:

$$
\begin{aligned}
\mathrm{UP}(g, r) = \min_{x^e, x^p} \quad & \sum_{b \in B} (c_b^p - r_b g_b) x_b^p + c_b^e x_b^e \\
s.t. \quad & x_b^e \leq x_{b'}^e \ \forall (b, b') \in P, \\
& x_b^p \leq x_b^e \ \forall b \in B, \\
& x_b^e, x_b^p \in \{0, 1\} \ \forall b \in B,
\end{aligned}
\tag{2.1}
$$

where the decision of extracting and processing each block $b \in B$ is represented by the binary variables x_b^e and x_b^p respectively. Since (2.1) is a minimization problem we will assume costs as positive and benefits as negative values, a convention we will use throughout the paper. The first set of constraints represents the extraction precedents for every block and the second set of constraints condition the processing of a block to its extraction.

2.1.2 UP with uncertainty on ore grades

The first step to develop a risk-averse model is to define a way to handle ore grade uncertainty for each block. In practice, formulation (2.1) is commonly solved by assuming g_b is deterministic, usually the mean value of different scenarios given by a geostatistical model.

For each $b \in B$, let $\tilde{g}_b$ be the random variable the represents the ore grade of block b. With this modification, the UP problem under uncertainty can be formulated as a two-stage stochastic optimization problem in which first stage variables represent block extraction, and in the second stage we have processing decisions, given extraction decisions taken in the first stage problem and the observed ore grade of extracted blocks. The two-stage stochastic UP problem can be formulated as follows:

$$\begin{aligned} \mathrm{UP}_U := \min_{x^e} \quad & \sum_{b\in B} c_b^e x_b^e + \rho_\alpha[Q(x^e,\tilde{g},r)] \\ \text{s.t.} \quad & x_b^e \le x_{b'}^e \ \forall (b,b') \in P, \\ & x_b^e \in \{0,1\} \ \forall b \in B, \end{aligned} \tag{2.2}$$

where

$$\begin{aligned} Q(x^e,\tilde{g},r) := \min_{x^p} \quad & \sum_{b\in B} (c_b^p - r_b \tilde{g}_b) x_b^p \\ \text{s.t.} \quad & x_b^p \le x_b^e \ \forall b \in B, \\ & x_b^p \in \{0,1\} \ \forall b \in B. \end{aligned} \tag{2.3}$$

Function $\rho_\alpha : L^1 \to \mathbb{R}$ will be defined as a *deviation or risk measure*, which deals with the uncertainty of the ore grade in the second stage. The parameter $\alpha \in \mathbb{R}$, whose range depends on the risk measure ρ_α under consideration, represents the risk-aversion of the decision maker. The work [9] did an extensive investigation when $\rho_\alpha = \mathbb{E}$, showing the importance of using the stochastic UP problem versus the deterministic version, and how it affects the pits obtained.

2.2 Nested pits

Usually, after the UP is established, mine planners run a complete scheduling model that defines in which period each block, or clusters of blocks, will be extracted. To this end, they solve $\mathrm{UP}(g,(1-\beta)r)$ as in (2.1) where $\beta \in [0,1[$ is a multiplier of the benefit r_b. Figure 1 shows the pits generated by solving a small instance of problem (2.1) with $\beta_1 > \beta_2 > \beta_3 > \beta_4$.

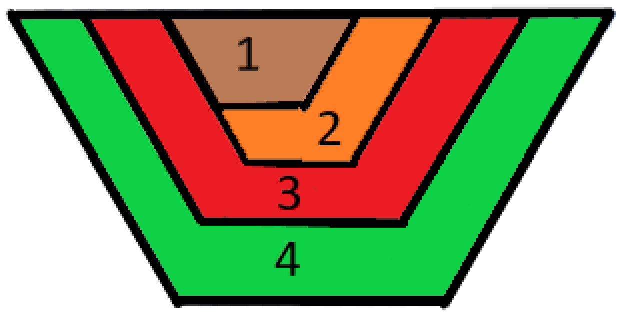


Figure 1: Example of nested pits.

As seen in the figure, the solution of $\mathrm{UP}(g,(1-\beta_i)r)$ will be a subset of blocks for any solution of $\mathrm{UP}(g,(1-\beta_{i+1})r)$ where $\beta_i > \beta_{i+1}$. This property is known as *nested pits*, and it is widely used in the mining context (for more information, please refer to [8], [2] and [7]). These pits are an important input for scheduling heuristics to determine the future workload and planning effort in mining.

Such methodology can be seen as a simple heuristic to deal with ore and price uncertainty. Using larger values of β translates into smaller pits where it is *safe* to start working as the mineral found in that pit justifies the extracting and processing costs even in low price scenarios. However, since ore grade is assumed to be deterministic, this approach does not capture the effects of ore grade variability directly.

An important aspect in our work is to propose an adequate risk measure ρ_α for mining problems, and to be able to solve problem (2.2)-(2.3) efficiently. Uncertainty of the ore grade is taken into account, and we want to derive conditions under which nested pits are obtained by varying the parameter α.

3 Risk measurement

3.1 Defining risk under ore grade uncertainty

In order to select the appropriate risk measure for the UP problem, we need to define which are the desirable characteristics in the pits generated by a given risk measure ρ_α. First, it needs to yield nested pits by varying α: *conservative* pits must be contained in *riskier* pits. The ability to generate nested pits based on parameter variation should be considered as a mandatory property for a risk measure since mine exploitation scheduling uses nested pits as a main input. The precise definition of *risk nestedness* is as follows:

Definition 1. *Assuming as the level of risk-aversion of the decision maker rises then the value of $\alpha \in \mathbb{R}$ increases and the ore grades are independently distributed, a risk measure ρ_α is* risk nested *for the UP problem if for $\alpha_1 > \alpha_2$ the set of extracted blocks in the optimal solution of problem* (2.2) *obtained by using ρ_{α_1} is contained in the set obtained using ρ_{α_2}.*

The CVaR is a commonly used risk measure thanks to its tractability. Following [11], the CVaR can be defined as

$$\mathrm{CVaR}_\alpha[X] = \min_{\eta\in\mathbb{R}} \left\{ \eta + \frac{1}{1-\alpha}\mathbb{E}(-X-\eta)_+ \right\}, \tag{3.4}$$

where X is a random variable and $(a)_+ := \max(a,0)$. Remark that $\alpha \in [0,1[$. A value of zero corresponds to the risk-neutral case and as α approaches one the risk measure protects against the worst-case realization. The next proposition shows that it may not be the ideal risk measure for stochastic UP problems under our requirements.

Proposition 1. *The CVaR risk measure is not risk-nested for the UP problem.*

Proof. We will prove that the CVaR is not risk-nested with a counter-example: let three blocks be positioned as in Figure 2.

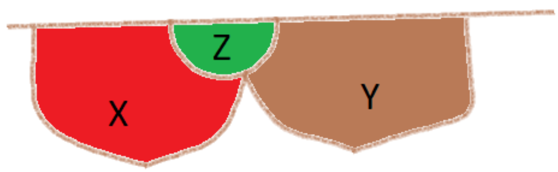


Figure 2: Three blocks.

In order to extract block X or Y, we must extract block Z (i.e. Z is a precedence of X and Y). Let the ore grade of the three blocks be independent of each other and normally distributed. If $W \sim N(\mu,\sigma^2)$ is normally distributed with mean μ and variance σ^2, then

$$\mathrm{CVaR}_\alpha(W) = \mu + \frac{\sigma^2}{(1-\alpha)\sqrt{2\pi}} e^{-z_{1-\alpha}^2/2}, \tag{3.5}$$

where $z_{1-\alpha}$ is the $(1-\alpha)$-quantile of the standard normal distribution. Let $\tilde{X} \sim N(-2,18)$, $\tilde{Y} \sim N(1,0)$ and $\tilde{Z} \sim N(-4,10)$ be the random variables that represent the negative profit of each block.

Table 1 shows the results of computing the CVaR using (3.5) for different values of α. It can be seen that with $\alpha = 0.05$ it is optimal to extract blocks X and Z, but with $\alpha = 0.1$ the optimal solution is to extract Y and Z. Finally with $\alpha = 0.5$ the preferences are such that it is better not to extract any blocks. Since $0.1 > 0.01$, extracting Y and Z and then X and Z shows that the CVaR is not risk nested. □

A second desirable property is related to the consistency of extracting decisions. Current geostatistical techniques assume a dependence over ore grades within blocks of the mine that are physically near, but this dependence diminishes as we select blocks that are far from each other. In other words, assuming two blocks of the mine could be represented as X and Y, the decision of mining

Combination of blocks	α 0.05	0.1	0.5
Nothing	0	0	0
$X+Y+Z$	-1.96	0.46	17.34
$X+Z$	-2.96	-0.54	16.34
$Y+Z$	-1.91	-1.05	4.98

Table 1: Comparison of mining preference under CVaR.

X instead of Y should not depend whether we are now considering another block Z which could be far from blocks X and Y, or of a different geological composition, acting as an independent random variable. In summary, we believe it is desirable that a risk-measure used in mining satisfies the *additive consistency* property:

Definition 2. *Let X, Y and Z be random variables where Z is independent of both X and Y, and $\rho_\alpha(\cdot)$ be a risk measure where α is the risk level of the decision maker. If $\rho_\alpha(X) < \rho_\alpha(Y)$ and ρ is additive consistent, then*

$$\rho_\alpha(X+Z) < \rho_\alpha(Y+Z).$$

Definition 2 states that decisions over a block, area, phase or cluster of blocks should not depend on the uncertainty of independent elements within the mine. In [3] the authors proved that the only risk measure that complies with additive consistency is the entropic risk measure. Given the uniqueness showed in [3], we will focus our attention in the entropic risk measure and examine its suitability for risk-averse UP problem (2.1).

3.2 The Entropic risk measure

Let X be a random variable. The entropic risk measure of X at level α is defined as

$$\rho_\alpha^{Ent}[X] := \begin{cases} \frac{1}{\alpha}\log \mathbb{E}[e^{\alpha X}] & \text{if } \alpha \neq 0, \\ \mathbb{E}[X] & \text{if } \alpha = 0. \end{cases}$$

If $\alpha > 0$, the decision maker is risk-averse, $\alpha < 0$ means risk-seeker, and $\alpha = 0$ corresponds to the risk-neutral (expected value) case. We will focus on non negative values of α. The entropic risk measure (Ent) is convex, but is not *coherent* [1,3] as shown in Appendix B, Lemma 5.

Our objective is to show that Ent is a viable risk-averse tool for the stochastic UP problem:

Lemma 1. *The entropic risk measure is risk nested when $\rho_\alpha = \rho_\alpha^{Ent}$ in problem* (2.2).

Proof. Proof is available in Appendix A.2 □

As a convention, if we are *indifferent* to whether extract a block or not, we will always choose to extract the block to avoid issues that could potentially yield non-nested pits solutions in problem (2.2).

In the next section, we will exemplify our approach using different configurations of uncertainty in the ore grade of the blocks of a mine, and compare the pits obtained to the same solution we would obtain by using the commonly used nested pits approach fixing the ore grade to its mean value.

4 Computational results

We present two sets of results: first, we consider a small-sized mine which was designed to illustrate the properties and methodologies we propose in this paper. The second mine is a real life instance with a large number of blocks to test the applicability of our methodology in a realistic environment.

We will provide a comparison of the results using different risk measures, and discuss the practical implications and managerial insights of the pits obtained with each method. We implemented all models in Python 3.7.2 and Gurobi 8.1, running on an Intel(R) i7 CPU 7700HQ @ 2.80 GHz and 12 GB of memory over Windows 10.

4.1 Small mine

This mine has 36 blocks in a two dimensional configuration, and is defined in terms of its topology and ore grade configuration. Figure 3 shows the topology of the mine, and the precedences were configured as follows: if a block is to be extracted, then the blocks on top, right and left of the block on top must be extracted, emulating a 45-degree slope angle.

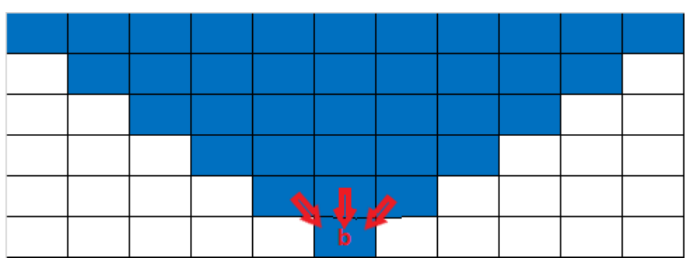


Figure 3: Topology of the small mine: blue means the block is part of set of precedent blocks of block b.

One of our goals is to study the behavior of using Ent as a risk measure versus the solutions obtained by using the classical nested pit methodology. In the next Lemma we show that, under normality, we can solve the problem directly, avoiding the use of sampling.

Lemma 2. *If the ore grade of the blocks follows a multivariate normal distribution $\tilde{g} \sim N(\mu, \Sigma)$ where μ is the vector of mean values and Σ is the variance-covariance matrix, then problem (2.2) for the entropic risk measure can be formulated as a mixed integer program with a convex quadratic objective function (MIQP)*

Proof. See Appendix A.3 for the proof. □

Using Lemma 2, we use a multivariate normal distribution for the ore grades obtaining the following equivalent formulation for UP with Ent:

$$\min_{x^e, x^p \in X^{EP}} \sum_{b \in B} c_b^e x_b^e + \sum_{b \in B} (c_b^p - r_b \bar{g}_b) x_b^p + \frac{1}{2} \alpha \sum_{b \in B} \sum_{b' \in B} r_b x_b^p r_{b'} x_{b'}^p \Sigma_{bb'}, \tag{4.6}$$

where X^{EP} is the same feasible set of solutions for problem (2.2). Problem (4.6) is convex given that matrix Σ is positive semidefinite, and it is amenable to be solved by Gurobi.

4.1.1 Results

We present the results for the following choice of parameters: $c_b^e = 0.5, c_b^p = 0, r_b = 1 \ \forall \ b \in B$. Since we have the closed quadratic formulation (4.6), we can calculate the exact distribution of the negative profits for each tested model.

The small mine was created as shown in Figure 4: the mean negative profit of extracting and processing any of the colored blocks as yellow (1) and red (2) always *pays* its extraction and its precedences. The red blocks have a variance greater than 0 and all possible pairs of red blocks have a positive covariance, while the yellow block has a variance of 0. We included the yellow block because we want to study the behavior of the models where there are blocks with no uncertainty to be extracted. All other blocks (blue) are waste, with grade $g_b = 0$.

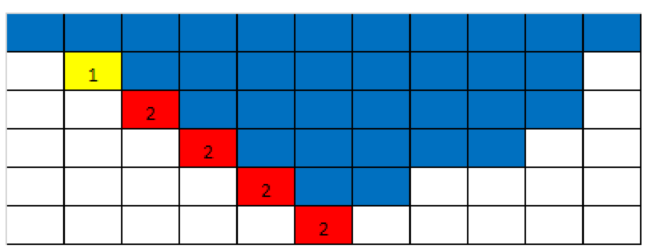
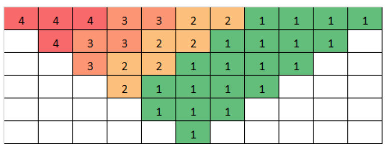


Figure 4: Configuration of the blocks in the small mine.

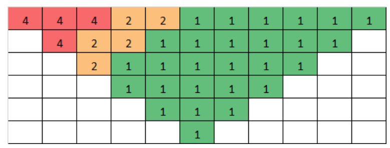
The different pits for Ent are shown in Figure 5. It is interesting to note the evolution of the pits when we multiply the parameter Σ by some positive constant: as variability grows, the optimal pit configuration changes for the same values of α. The sub-figures show the evolution of the pits using the following values:

- Green (1): $\alpha = 0$, risk neutral case.
- Yellow (2): $\alpha = 0.1$.
- Orange (3): $\alpha = 0.5$.
- Red (4): $\alpha = 1$.

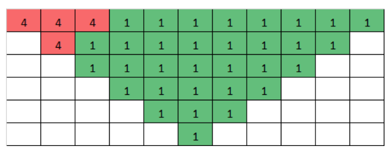
Note that if any of these values (or colours) are missing, is because the pit obtained is the same as the pit obtained by using a higher value of α, i.e., in Figure 5b the optimal solution using 0.5 and 1 are the same.



(a) Σ.



(b) 2Σ.



(c) 10Σ.

Figure 5: Configuration of the pits obtained with different values of α and changing the variance-covariance configuration.

As uncertainty becomes larger, the pits obtained show a conservative evolution: as α grows the optimal solution pit obtained becomes smaller. In Figure 5c we see that for any value of $\alpha > 10$ all pits obtained extract the yellow block. Therefore, if variability in the ore grade is high then Ent will provide pits with small differences between the solutions when changing the value of α. Please note that, while our objective was to observe problem (2.2) using Ent under different conditions of uncertainty, we observed that the solutions obtained were nested, even if the covariance of the ore grades between some of the blocks are not 0.

The same analysis can be performed using the current practice in the industry, namely solving problem (2.1) and using different values of β (the classical nested pits methodology) to this mine. Note that this methodology doesn't consider the variance of the grades, only its mean values. Since the mean value is constant, we get the same pits even in scenarios of high uncertainty. Figure 6 shows this, the configuration of matrix Σ does not change the optimal solutions found by using different values of β as described in Section 2.2. The evolution of these pits is obtained using the following values:

- Green (1): $\beta = 0$, risk neutral case.
- Red (2): $\beta = 0.5$.

- *No blocks are extracted*: $\beta \geq 0.7$.

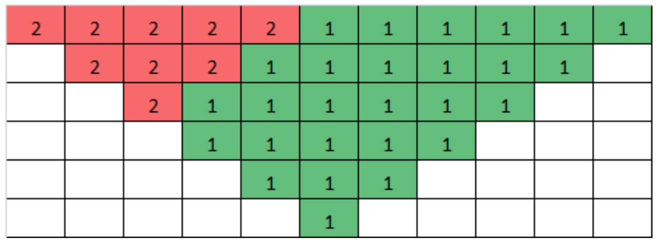


Figure 6: Configuration of the pits obtained with different values of β.

The definition of *risk* between models differs: Ent is aiming at controlling variability, and its most conservative result has variance 0. For the nested pit case, since we are manipulating the price, for values of β close to 1 the mining *profit* is small compared to the costs of extraction and processing, which means the corresponding pit will be small, and in the limit will correspond to not extracting any block.

In summary, for a self-constructed small mine the classical approach has the tendency of being overly conservative for some values of β, avoiding extraction altogether. Moreover, the classical approach is completely insensitive to changes in the correlation between blocks, which is alarming. We can clearly see how variability becomes an important factor with the Ent approach, and how it is controlled by the choice of α: conservative values aim for small but less variable expected profits.

4.2 Andina case

The Andina mine is an open pit mine that belongs to Codelco, Chile's state-owned company that controls around 20% of global reserves. Andina is located in Rio Blanco (approximately 80 km NE of Santiago) and is still active after 82 years. A sector of the mine is represented by 26,400 blocks, and the ore distribution is approximated by 120 scenarios generated from a set of drilling holes via conditional simulations using the turning bands algorithm [5].

Since this mine does not have a multivariate normal distribution of ore grades we cannot use the closed quadratic formulation (4.6). We tried to solve the problem directly on a smaller version of Andina (less than 100 aggregated blocks) using non-linear optimization solvers (AMPL/MINOS) to check if we could handle the real mine with its 26,400 blocks. The solver was not able to close an optimality gap of 99.9% after 5 hours running, which means it is hopeless to try to attack the problem directly.

Ent is a non-linear function, and since we have integer variables the resulting risk-averse UP problem is extremely challenging to solve for larger mines. However, given the convexity of the exponential function, we can use a piecewise linear approximation of the objective function. In order to accomplish this, we modify the objective function of (2.2) thanks to the translation invariance of

Ent (please refer to Lemma 6) and the monotonicity of the log function. All necessary proofs and final model which approximates the value of problem (2.2) using Ent can be found in Appendix C.

4.2.1 Results

We divide the set of 120 ore grade scenarios in a subset of 20 samples which is used to solve the problem, and the remaining 100 are used for the out-of-sample analysis.

Our piecewise linear approximation was designed to be an uniform grid between integers -20 and 20 with steps of 0.01 (4,000 steps in total), because Gurobi treats values below 10^{-8} as 0, and $\exp(-20) \approx 2.06 \cdot 10^{-9}$. Therefore, we calculated a scaling factor and we changed the `NumericFocus` parameter of Gurobi to its maximum value in order to give more processing power in the numerical calculations, and to avoid numerical issues given the small numbers that will be used by our model. We use the mean absolute deviation (MAD) to check the performance of our approximation.

We denote by NP the results obtained by the nested pit methodology. In each case, we evaluate the total profit of the resulting ultimate pit on each of the out-of-sample scenario independently. This give us a set of 100 values for each solution. We also compare these values with the *lower bound* (LB) obtained solving problem (2.1) fixing g_b to each scenario. LB will have a different pit and negative profit per scenario, serving as a benchmark of the best possible result. Also, we present the variance coefficient (VC), which is the ratio between standard deviation and average of the profits.

Table 2 shows the results for NP, while changing the value of β. If $\beta \geq 0.7$, the solutions obtained have less than 100 blocks extracted and for $\beta \geq 0.9$ it doesn't extract blocks, therefore these results are omitted. This behavior was already shown in the small mine example in the previous section: conservative values of β will incur in avoiding extraction altogether. However, as risk-averseness grows, the VC becomes larger (6.9% in average).

Model	LB	NP 0	NP 0.1	NP 0.2	NP 0.3	NP 0.4	NP 0.5	NP 0.6
Upit blocks	-	15,775	12,738	7,303	4,218	1,963	816	310
Average	-132.1	-121.4	-110.7	-86.5	-64.5	-43.2	-24.6	-12.4
VC	-	5.9%	5.6%	5.3%	5.6%	6.3%	8.1%	11.0%
% vs LB		91.9%	83.8%	65.5%	48.8%	32.7%	18.6%	9.4%

Table 2: Table of descriptive statistics for NP of the out of sample results.

Table 3 shows the results for Ent. We observe a larger amount of blocks extracted in the pit of the neutral case, and using larger values of α the pits obtained have less blocks. Please note, as risk-averness grows, the value of the VC becomes smaller (5.6% in average). The MAD for different values of α are of order $O(10^{-5})$ in the worst case, showing that our linear approximation is effective. Values for $\alpha > 12.5$ are not present in the table since the model incurs in numerical errors: the solution obtained was an empty pit.

Model	LB	Ent 0	Ent 7.5	Ent 10	Ent 12.5
Upit blocks	-	21,770	19,545	14,810	11,273
MAD	-	-	2.62E-07	2.55E-07	2.92E-07
Average	-132.1	-128.8	-125.6	-110.7	-98.6
VC	-	5.9%	5.7%	5.3%	5.4%
% vs LB	-	97.5%	95.1%	83.8%	74.6%

Table 3: Table of descriptive statistics for Ent of the out of sample results.

Figure 7 shows the negative profit obtained per scenario of Ent, NP and LB on each of the testing set ore grade samples, sorted by the results of LB and using different value of α and β respectively.

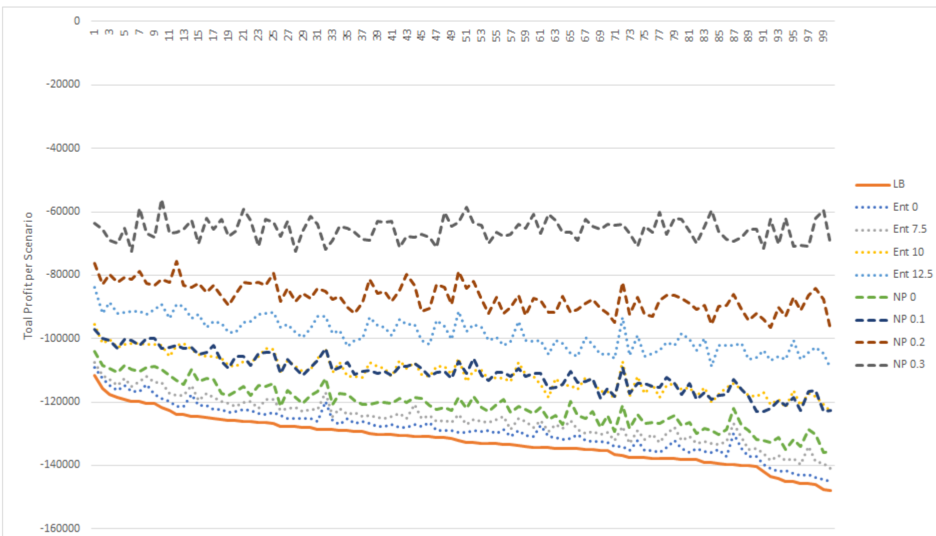


Figure 7: Negative profits per scenario for the Andina case.

Figure 8 shows a cross sectional view of the pit, where we can observe the differences on the optimal solutions of Ent and NP. We used $\alpha = 10$ and $\beta = 0.03$ to obtain pits with a similar number of blocks extracted (14,810 in Ent versus 14,983 in NP.). We can see both models aim for different sections of the mine.

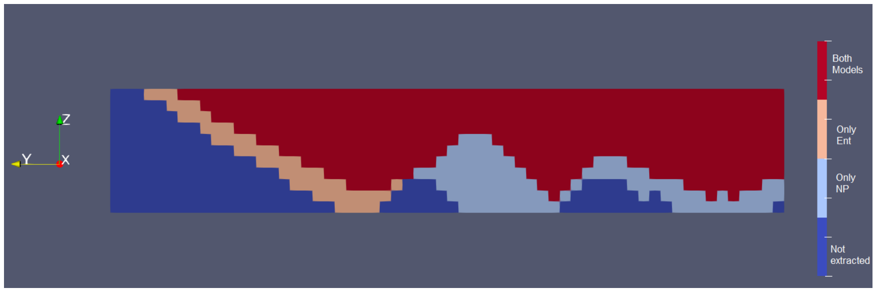


Figure 8: Comparing pits of both models.

The importance of these results are the effective application of an approximation method to prove Ent as a viable tool for risk-aversion in real life instances of mining optimization problems. The error obtained in the approximation to the real value of the exponential function is of order $O(10^{-5})$, which was within our acceptance tolerance and Gurobi was able to solve the approximation formulation in an average time of under 5 minutes for all the values of α.

5 Conclusions

In this paper, we formulate and discuss how to solve the risk-averse UP problem. We start by describing two properties that a risk measure should satisfy in this context: *risk nestedness* and *additive consistency*. Risk nestedness is a fundamental property since modern mine scheduling algorithms use the idea of nested pits as an input for generating a sequence of block extractions over time. Additive consistency is a desirable property that is intuitive in the sense that it preserves preferences in the presence of independent waste blocks. An interesting consequence of our work is that one of the most popular risk measures, the Conditional Value-at-Risk, fails to satisfy both properties even in the case of blocks with independent distribution, calling into question its use in mining.

In a small simulated mine we contrast results for the current methodology of changing sale prices to obtain nested pits, the classical nested pit (NP) approach, and the proposed Entropic risk measure methodology (Ent). The former behaves as expected: we obtain different pits that vary from mining everything up to avoid working on the mine altogether in the most conservative cases. The Ent approach shows a different behavior: we obtain a different pit transition as we vary the level of risk-aversion, more focused on avoiding variance within the results, thus high uncertainty scenarios will have smaller pits than low uncertainty cases for the same value of α.

We apply our method to a real-life mine (Andina, in Chile) and use a linear approximation to efficiently solve the risk averse problem with the entropic risk measure. Our results show the approximation errors were within tolerable margins, validating the linear approximation scheme. Future work includes extensions to our numerical algorithm to be able to cope with larger mines with millions of blocks, and the study of dynamic risk-averse models for mine scheduling problems.

A Mathematical proofs

A.1 Risk-nestedness for the UP problem of the entropic measure

Lemma 3. *Let $\alpha_1, \alpha_2 \in \mathbb{R}$, with $\alpha_1 > \alpha_2$. For any random variable X we have that $\rho^{Ent}_{\alpha_1}(X) \geq \rho^{Ent}_{\alpha_2}(X)$.*

Proof. We will show that the objective function of (2.2) increases monotonically with $\alpha > 0$ (since $\alpha = 0$ is the expected value case and $a < 0$ is the risk seeking case). Let us define $f(X, \alpha)$ as:

$$f(X, \alpha) = \frac{1}{\alpha} \log \mathbb{E}[e^{\alpha X}],$$

where X is a random variable. Now, let us study the behavior of the first partial derivative over α:

$$\frac{\partial}{\partial \alpha} f(X, \alpha) = \frac{1}{\alpha^2} \left(\frac{\mathbb{E}[\alpha X e^{\alpha X}]}{\mathbb{E}[e^{\alpha X}]} - \log \mathbb{E}[e^{\alpha X}] \right).$$

Let $Y = e^{\alpha X}$, then:

$$\begin{aligned} \frac{\partial}{\partial \alpha} f(X, \alpha) &= \frac{1}{\alpha^2} \left(\frac{\mathbb{E}[Y \log Y]}{\mathbb{E}[Y]} - \log \mathbb{E}[Y] \right) \\ &= \frac{1}{\alpha^2} \left(\frac{\mathbb{E}[Y \log Y] - \mathbb{E}[Y] \cdot \log \mathbb{E}[Y]}{\mathbb{E}[Y]} \right). \end{aligned}$$

Since $\frac{1}{\alpha^2} > 0$ and $\mathbb{E}[Y] = \mathbb{E}[e^{\alpha X}] > 0$, we just need to study the sign of the numerator in the partial derivative. Let $h(Y) = Y \log Y$, then:

$$\mathbb{E}[Y \log Y] - \mathbb{E}[Y] \log \mathbb{E}[Y] = \mathbb{E}[h(Y)] - h(\mathbb{E}[Y]).$$

Finally, since $h(Y)$ is a convex function ($\frac{\partial^2 h(Y)}{\partial Y^2} = \frac{1}{Y} > 0$) then by using Jensen's inequality we know that $\mathbb{E}[h(Y)] - h(\mathbb{E}[Y]) \geq 0$, therefore $f(X, \alpha)$ increases monotonically. □

Lemma 3 ensures the monotonic behavior of Ent w.r.t. the value of α.

Lemma 4. *Let X, Y are two independent random variables and $0 < \alpha < \infty$, then*

$$\rho^{Ent}_{\alpha}(X + Y) = \rho^{Ent}_{\alpha}(X) + \rho^{Ent}_{\alpha}(Y)$$

Proof. Please refer to [3] for the proof. □

A.2 Proof of Lemma 1

Proof. Let $\tilde{g}_b$ be the random variable of the ore grade for each block $b \in B$ and assume they are independently distributed, and let $\alpha_1, \alpha_2 \in \mathbb{R}$ where $0 < \alpha_2 < \alpha_1 < \infty$. Suppose that a certain block b is in the pit obtained by solving (2.2) using Ent with α_1, but is not present in the pit under the same problem solved by using α_2. Let U_1 be the pit obtained in the first case and U_2 the second case.

Since extracting 0 blocks is a feasible solution and $\rho_{\alpha}^{Ent}(0) = 0$ for $\alpha > 0$ then we know that $\rho_{\alpha_1}^{Ent}(U_1) \leq 0$. Furthermore, $\rho_{\alpha_1}^{Ent}(U_1 \backslash U_2) \leq 0$, else $U_1 \cap U_2$ would obtain a better value for α_1 but U_1 is the optimal solution which would be a contradiction.

By using Lemma 3, since $\alpha_2 < \alpha_1$, then $\rho_{\alpha_2}^{Ent}(U_1 \backslash U_2) \leq \rho_{\alpha_1}^{Ent}(U_1 \backslash U_2) \leq 0$. But then the pit $U_2 \cup (U_1 \backslash U_2)$ is also feasible and by Lemma 4 we have that $\rho_{\alpha_2}^{Ent}(U_2 \cup (U_1 \backslash U_2)) = \rho_{\alpha_2}^{Ent}(U2) + \rho_{\alpha_2}^{Ent}(U_1 \backslash U_2) \leq \rho_{\alpha_2}^{Ent}(U2)$, which is a contradiction of the optimality condition of U_2, finishing the proof. $\square$

A.3 Proof of Lemma 2

Proof. Let $\tilde{g} \sim N(\mu, \Sigma)$ then let Z be as follows:

$$Z := \left(\sum_{b \in B} c_b^e x_b^e + \sum_{b \in B} (c_b^p - r_b \tilde{g}_b) x_b^p \right) \rightarrow N(\mu', \sigma'^2),$$

where

$$\mu' = \sum_{b \in B} c_b^e x_b^e + \sum_{b \in B} (c_b^p - r_b \bar{g}_b) x_b^p,$$

and

$$\begin{aligned} \sigma'^2 &= (r \circ x^p)^T \Sigma (r \circ x^p) \\ &= \sum_{b \in B} \sum_{b' \in B} r_b x_b^p r_{b'} x_{b'}^p \Sigma_{bb'}, \end{aligned}$$

where operator $\circ$ is the Hadamard (element-wise) matrix product.

We can verify that $\mathbb{E}[e^{\alpha Z}]$ is the moment generating function of a univariate normal distribution, which in turn corresponds to:

$$\exp \left(\alpha \mu' + \frac{1}{2} \alpha^2 \sigma'^2 \right).$$

Therefore, we obtain the following equivalent formulation:

$$\min_{x^e, x^p \in X^{EP}} \rho^{Ent} \left(\sum_{b \in B} c_b^e x_b^e + \sum_{b \in B} (c_b^p - r_b \tilde{g}_b) x_b^p \right) = \min_{x^e, x^p \in X^{EP}} \mu' + \frac{1}{2} \alpha \sigma'^2,$$

where X^{EP} is the same feasible set of solutions for problem (2.2), and since Σ is s.d.p., the proof is finished. $\square$

B Other Lemmas and Propositions of interest

Lemma 5. *The entropic risk measure is not a coherent risk measure.*

Proof. Let X be a random variable and $a \geq 0$. We have:

$$\rho_\alpha^{Ent}(aX) = \frac{1}{\alpha}\log \mathbb{E}[e^{\alpha a X}].$$

If $\beta = \alpha a$, then

$$\begin{aligned}\rho_\alpha^{Ent}(aX) &= \frac{a}{\beta}\log \mathbb{E}[e^{\beta X}] \\ &= a\rho_\beta^{Ent}(\beta X) \neq a\rho_\alpha^{Ent}(X).\end{aligned}$$

Therefore, Ent is not positive homogeneous, finishing the proof. □

Lemma 6. *The entropic risk measure is a translation invariant risk measure.*

Proof. Let X be a random variable and $a \geq 0$. We have:

$$\begin{aligned}\rho_\alpha^{Ent}(X+a) &= \frac{1}{\alpha}\log \mathbb{E}[e^{\alpha(X+a)}] \\ &= \frac{1}{\alpha}\log \mathbb{E}[e^{\alpha X}e^{\alpha a}] \\ &= \frac{1}{\alpha}\log e^{\alpha a}\mathbb{E}[e^{\alpha X}] \\ &= \frac{1}{\alpha}(\log e^{\alpha a} + \log \mathbb{E}[e^{\alpha X}]) \\ &= \frac{1}{\alpha}(\alpha a + \log \mathbb{E}[e^{\alpha X}]) \\ &= \frac{1}{\alpha}\log \mathbb{E}[e^{\alpha X}] + a \\ &= \rho_\alpha^{Ent}(X) + a.\end{aligned}$$

Therefore, Ent is translation invariant, finishing the proof. □

Proposition 2. *If* $\rho(\cdot) = \mathbb{E}[\cdot]$, *then formulation* (2.1) *is equivalent to* (2.2).

Proof. It can be seen that in formulation (2.2) we are using two sets of variables: x^e set of blocks to be extracted and x^p set of blocks to be processed. In (2.3) we have the constraint $x_b^p \leq x_b^e \; \forall b \in B$, which limits the possibility to process a block only if it is extracted in the first place, then for a certain block b the cost in the objective function can be one of three possible values:

1. 0 if the block is not extracted, i.e. $x_b^e = 0$.
2. c_b^e if the block is extracted but not processed, i.e. $x_b^e = 1$ and $x_b^p = 0$.
3. $c_b^p - r_b g_b^\omega + c_b^e$ if the block is extracted and processed, i.e. $x_b^e = 1$ and $x_b^p = 1$.

As (2.3)'s objective is to minimize the negative profit of processing the blocks given $\tilde{g}$, then the decision can be stated as:

$$x_b^p = \begin{cases} 1, & \text{if } c_b^p < r_b g_b^\omega, \\ 0, & \text{if } c_b^p \geq r_b g_b^\omega. \end{cases}$$

On the other hand, the expression $c_b^e - \mathbb{E}[(r_b \tilde{g}_b^\omega - c_b^p)^+]$ can assume exactly the same three previously described options. Since we can pre-compute the value of $\mathbb{E}[(c_b^p - r_b g_b^\omega)^+]$, $\forall b \in B$, we can eliminate the use of a second set of variables. If $x_b = 1$ in (2.1) it means that the block must be extracted whether is waste or is going to be processed. □

C Entropic risk objective function linearization

The risk-averse UP problem using Ent can be written as

$$\min_{x^e, x^{p,\omega} \in X^{EP\Omega}} \frac{1}{N} \sum_{\omega \in \Omega} \exp\left(\alpha \left(\sum_{b \in B} c_b^e x_b^e + \sum_{b \in B} (c_b^p - r_b g_b^\omega) x_b^{p,\omega}\right)\right), \tag{C.7}$$

where Ω is the domain of the random variable g_b and

$$\begin{aligned} X^{EP\Omega} :=& \{x^e \in \{0,1\}^{|B|}, x_b^{p,\omega} \in \{0,1\}^{|B \times \Omega|} : x_{b'}^e \leq x_b^e \ \forall (b,b') \in P, \\ & x_b^{p,\omega} \leq x_b^e \ \forall b \in B, \forall \omega \in \Omega\}. \end{aligned}$$

We will use two auxiliary variables: v_ω and z_ω, where

$$\sum_{\omega \in \Omega} v_\omega \geq \frac{1}{N} \sum_{\omega \in \Omega} \exp(z_\omega).$$

The strategy here is to use an approximation of the exponential function for each term in the sum of the objective function in (C.7), using the auxiliary variable v_ω. Problem (C.7) is equivalent to the following problem:

$$\begin{aligned} \text{UP}_e := \min_{x^e, x^p \in X^{EP\Omega}} \quad & \sum_{\omega \in \Omega} v_\omega \\ \text{s.t.} \quad & z_\omega \geq \alpha \left(\sum_{b \in B} c_b^e x_b^e + \sum_{b \in B} (c_b^p - r_b g_b^\omega) x_b^{p,\omega}\right) \ \forall \omega \in \Omega, \\ & N v_\omega \geq \exp(z_\omega) \ \forall \omega \in \Omega, \\ & v_\omega \geq 0 \ \forall \omega \in \Omega. \end{aligned} \tag{C.8}$$

The non-linear constraint of (C.8) will be replaced by a piecewise linear approximation, detailed in the work of [10], for each constraint represented by z_ω. For differentiable convex functions (such as $f(x) = \exp(x)$), we can use a lower approximation using the gradient at a given point t_i: $f(x) - f(t_i) \geq$

$\nabla f(t)(x - t_t)$. We will select a set of K points $U := \{t_1, \ldots, t_K\}$ to calculate the value of $\nabla f(i)\ \forall i \in U$ and approximate the exponential function using linear functions as shown in Figure 9.

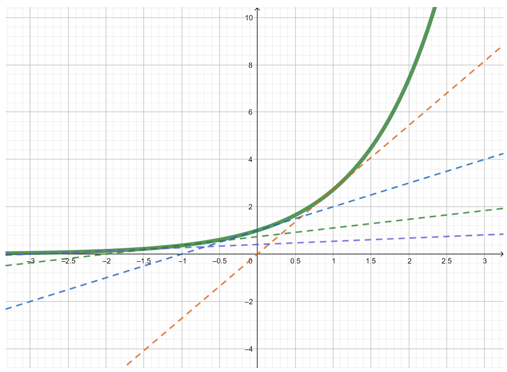


Figure 9: Subgradient linear approximation of $\exp(x)$ using $K = 4$ support points where $U = \{-2, -1, 0, 1\}$.

Using this in (C.8), we get the following constraints:

$$v_\omega \geq \exp(u_\omega - t_i) + \exp(t_i) \qquad \forall \omega, i \in \Omega \times U,$$

which in turn can be used to construct our approximated problem:

$$\begin{aligned} \mathrm{UP}_A := \min_{x^e, x^p \in X^{EP\Omega}} \quad & \sum_{\omega \in \Omega} v_\omega \\ \text{s.t.} \quad & z_\omega \geq \alpha \left(\sum_{b \in B} c_b^e x_b^e + \sum_{b \in B} (c_b^p - r_b g_b^\omega) x_b^{p,\omega} \right) \ \forall \omega \in \Omega, \\ & v_\omega - e^{t_i} z_\omega \geq (1 - t_i) e^{t_i} \ \forall \omega, i \in \Omega \times U, \\ & v_\omega \geq 0 \ \forall \omega \in \Omega. \end{aligned} \tag{C.9}$$

References

[1] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent Measures of Risk. *Mathematical Finance*, 9(3):203–228, 1999.

[2] Louis Caccetta and Stephen P. Hill. An application of branch and cut to open pit mine scheduling. In *Journal of Global Optimization*, 2003.

[3] Roberto Cominetti and Alfredo Torrico. Additive Consistency of Risk Measures and Its Application to Risk-Averse Routing in Networks. *Mathematics of Operations Research*, 41(4):1510–1521, 2016.

[4] Noel Cressie. The origins of kriging. *Mathematical Geology*, 1990.

[5] Xavier Emery and Christian Lantuéjoul. Tbsim: A computer program for conditional simulation of three-dimensional gaussian random fields via the turning bands method. *Computers & Geosciences*, 32(10):1615–1628, 2006.

[6] D Espinoza, G Lagos, E Moreno, and J P Vielma. Risk Averse Approaches In Open-Pit Production Planning under ore Grade uncertainty: a Ultimate Pit Study. 2013.

[7] Marcos Goycoolea, Daniel G. Espinoza, Eduardo Moreno, and Orlando Rivera Letelier. Comparing new and traditional methodologies for production scheduling in open pit mining. *Application of Computers and Operations Research in the Mineral Industry - Proceedings of the 37th International Symposium, APCOM 2015*, pages 352–359, 2015.

[8] H. Lerchs and F. Grossman. Optimum design of open-pit mines. In *Operations Research*, 1965.

[9] Denis Marcotte and Josiane Caron. Ultimate open pit stochastic optimization. *Computers and Geosciences*, 51:238–246, 2013.

[10] G. Parra. Generación de benchmark de fondos para el sistema de pensiones en chile: Un enfoque basado en optimización estocástica. Master's thesis, Universidad de Chile, 2016.

[11] R. Tyrrell Rockafellar and Stanislav Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–41, 1997.