Natural Resources Research (© 2020)
https://doi.org/10.1007/s11053-020-09625-3

Original Paper

Ore-Waste Discrimination with Adaptive Sampling Strategy

Felipe A. Santibáñez-Leal,[1,4] Julián M. Ortiz,[2] and Jorge F. Silva[3]

Received 8 September 2019; accepted 23 January 2020

Grade control and short-term planning determine the performance of a mining project. Improving this decision, by collecting the most informative samples (data) may have significant financial impact on the project. In this paper, a method to select sampling locations is proposed in an advanced drilling grid for short-term planning and grade control in order to improve the correct assessment (ore-waste discrimination) of blocks. The sampling strategy is based on a regularized maximization of the conditional entropy of the field, functional that formally combines global characterization of the field with the principle of maximizing information extraction for ore-waste discrimination. This sampling strategy is applied to three real cases, where dense blast-hole data is available for validation from several benches. Remarkably, results show relevant and systematic improvement with respect to the standard regular grid strategy, where for deeper benches in the field the gains in ore-waste discrimination are more prominent.

KEY WORDS: Entropy, Uncertainty and information extraction, Multiple-point statistics, Geostatistics, Short-term planning, Advanced sampling.

INTRODUCTION

Short-term planning and grade control aim at determining the optimum assignment of each block of material in a mine, considering the potential economic profit, complying with the mine plan and the constraints in the mine and processing facilities. This assignment can be a specific process, a stockpile or the waste dump. In this paper, the analysis has been limited to the binary decision case of ore sent to the processing plant or waste sent to the dump. In this context, the grade and other properties of each block are estimated from samples located in its neighborhood. In most open pit mines, these samples are taken where blast-holes are drilled. Two important considerations about this estimation procedure are: The block grade estimation involves a change of support, that is the volumetric support of the samples is usually much smaller than that of the block, and the block grade is estimated based on the

The work was supported by the research grants Fondecyt 1170854, CONICYT, Chile. The work of Dr. Silva is supported by the Advanced Center for Electrical and Electronic Engineering (AC3E), Basal Project FB0008. Felipe Santibañez is supported by CONICYT Ph.D. scholarship 21130890 and the Advanced Mining Technology Center (AMTC) Basal project (CONICYT Project AFB180004). Dr. Ortiz acknowledges the support of the Natural Sciences and Engineering Council of Canada (NSERC), funding reference number RGPIN-2017-04200 and RGPAS-2017-507956.

[1]Advanced Laboratory for Geostatistical Supercomputing (ALGES), Department of Mining Engineering, Advanced Mining Technology Center (AMTC), University of Chile, Av. Tupper 2069, Santiago 837-0451, Chile.
[2]The Robert M. Bunchan Department of Mining, Queen's University, Kingston, Ontario, Canada.
[3]Information and Decision System Group (IDS), Department of Electrical Engineering, University of Chile, Av. Tupper 2007, Santiago 837-0451, Chile.
[4]To whom correspondence should be addressed; e-mail: fsantibanez@alges.cl

Published online: 08 February 2020

most current information, which during production is the sampled grades obtained at blast-holes.

Because of these considerations, the estimation technique of the block grade must account for the *support effect* as the block ordinary kriging is typically used to determine the block grade. Secondly, the fact that the assignment of the block to a specific destination is based on an estimation from limited data, means that sampling errors in that information impair the final assignment (Abzalov et al. 2010; Chiles and Delfiner 2012, p. 7), which is known as *information effect* (Chiles and Delfiner 2012, p. 442)

Short-term planning and grade control are almost always based on samples taken from a pseudo-regular grid of blast-holes. These samples suffer from high errors due to numerous factors: poor sample recovery near the collar of the blast-hole, delimitation error by including the subdrill and careless applications of sampling procedures since production has priority and the area must be freed prior to loading explosives (Francois-Bongarcon 1983; Pitard 1993; Ortiz and Deutsch 2004). Poor sample quality leads to extremely large economic losses, which increase when geological domains are poorly understood or when estimation of the block grades suffers from large errors (Magri and Ortiz 2000). These sources of uncertainty are hidden, but have a significant impact in the financial performance of mining projects (Magri et al. 2010, 2011; Ortiz et al. 2012).

Regarding the information used for grade control, most applications in mining deal with a regular grid, and optimization is aimed at determining the spacing of the samples to comply with some measures of quality (Ortiz and Magri 2014). These measures of quality are often linked to the kriging variance, as the measure of performance of the estimation (McBratney and Webster 1981; McBratney et al. 1981; Blackwell 1998; Lloyd and Atkinson 1998; Hassanipak and Sharafodin 2004; Vasát et al. 2010). The use of geostatistical simulation allows to consider uncertainty that may be a function of the block grade, accounting for the *proportional effect* (Journel 1974).

Some authors have proposed other approaches in the optimization of ore-waste discrimination, but most of them attempt to optimize the classification as a post process after the sampled information has been acquired (Ruiseco et al. 2016). Some interesting works have been carried out from the point of view of optimization under process/operational production capabilities constraints (Kumral 2012, 2013).

Furthermore, although in most applications conventional two-point simulation methods, such as sequential Gaussian or sequential indicator simulation, are used, some authors have introduced the use of multiple-point geostatistical simulation in mining applications (Ortiz 2003; Ortiz et al. 2012; Boisvert et al. 2008). These methods require the use of a training image (Mariethoz and Caers 2014), which is commonly built from an analogue, outcrop or a geological interpretation (Boisvert et al. 2008). Data-driven training images can also be used (Ortiz et al. 2012; Silva Maureira 2015) and are precisely the approach taken in this study.

Advanced drilling considering a reverse circulation (RC) drilling rig has been studied as an option to improve the quality of the samples for short-term planning and grade control and to provide grades and other geological information in advance to build a reliable model for decision-making. This approach has led to improved financial returns (Magri et al. 2010, 2011; Ortiz and Deutsch 2004; Ortiz and Magri 2014) and allows the use of more sophisticated tools to build the short-term plan, such as cokriging or cosimulation to account for multivariate relationships. Furthermore, advanced drilling allows a characterization of the geological features of the rock (lithology, alteration, mineralization); hence, these samples can be used to improve and update the geological interpretation. Advanced drilling is normally applied considering a regular drilling grid with a spacing wider than the blast-hole spacing. From the high-quality samples obtained in the advanced drilling grid, a short-term model can be built using geostatistical estimation (kriging or cokriging) or simulation (Journel and Kyriakidis 2004). The simulation approach has usually relied on a multi-Gaussian method, and to the best knowledge of the authors, multiple-point geostatistics has not been applied in grade control. Cost functions that account differently for the costs of misclassification as ore or waste have also been used frequently in mining (Deutsch and Journel 1998; Journel 1974; Verly 2005; Dimitrakopoulos and Godoy 2014).

This work presents a new methodology for short-term planning sampling and grade control based on the selection of the sampling locations that are the most informative in terms of ore-waste discrimination. The proposal aims at characterizing the contact between ore and waste, by learning from the

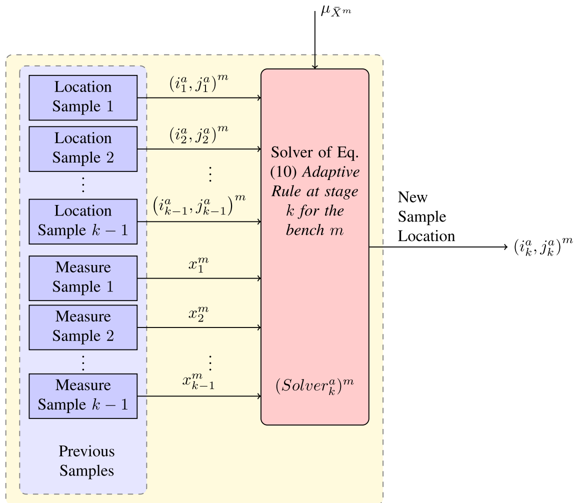

Figure 1. Schematic diagram of the sampling rule (10).

data of previously mined benches the location and geometry of the contact, and determining the best locations to be sampled at the current bench. The implementation of this principle results in an irregular grid of samples. Furthermore, the sample locations adapt to the local conditions of the problem. New samples are drilled at locations that have the highest conditional entropy based on the available samples (within the bench) and on the spatial continuity of the ore and waste blocks. This last component of the model is estimated from previously mined benches. On the specifics, an algorithm is adopted to this problem to select the best locations, previously introduced by the authors (Santibañez et al. 2019). This methodology is summarized (Section Entropy-based Adaptive Sampling Strategy), adapted to the proposed problem (Section Optimal Sampling for Grade Control), applied to three real scenarios and then the results for each case are analyzed, showing the performance improvement with respect to an advanced regular sampling grid (Section Case Study and Experiments). Then, the limitations and potential economic benefits of this approach are discussed and some final conclusions are provided.

ENTROPY-BASED ADAPTIVE SAMPLING STRATEGY

In this section, the proposed method for determining the best sample locations over a binary two-dimensional field is reviewed. This strategy is based on a regularized maximum entropy sampling problem presented in Santibañez et al. (2019), which provides a sampling scheme that maximizes the information extracted from the measurements.

Maximum Entropy Sampling

To formalize the problem, let us consider a *2-D* field with unknown spatial correlation. Notice that this spatial correlation may be fully characterized by a variogram or may require higher order statistics to be described (Mariethoz and Caers 2014; Ortiz 2003). More precisely, and without loss of generality, the regionalized variable X is a *2-D* random array of variables representing a discrete image of finite size $A \cdot B$,

$$X_{u,v} : (\Omega, \mathbb{P}) \to \mathcal{A} = \{0,1\} \quad \forall (u,v) \in [A] \times [B], \tag{1}$$

with values in the finite alphabet $\mathcal{A}$ (limited here to the binary case), and $(\Omega, \mathbb{P})$ describing the sample space and the probability measure for the stochastic regionalized variable.

Adopting the concept of entropy as a measure of uncertainty of a random variable (Cover and Thomas 2006), an algorithm that finds the placement rule f through optimal reduction of the posterior entropy after doing measurements is proposed. More precisely, for any given number $K \leq A \cdot B$ of sample positions to be taken, let $F_K \equiv \{f : \{1,..,k,..,K\} \to [A] \times [B]\}$ be the collection of functions that select K-elements in $[A] \times [B]$. Then for any $f \in F_K$ (sampling rule of size K) let's denote the measured random vector by:

$$X_f \equiv (X_{f(1)}, X_{f(2)}, \ldots, X_{f(K)}), \tag{2}$$

and the remaining non-measured random vector by,

$$X^f \equiv (X_i : i \in [A] \times [B] \setminus f). \tag{3}$$

In this context, the conditional posterior entropy of X^f given X_f measures the remaining uncertainty of the field X after sensing the position in X_f (Cover and Thomas 2006). It can be computed as the joint entropy of the entire process minus the joint entropy of the variables measured by f, as shown in Eq. 4.

$$H(X^f | X_f) = H(X) - H(X_f). \tag{4}$$

Table 1. Summary statistics

	Case study 1		Case study 2		Case study 3	
	Drill-hole Samples	Blast-hole Samples	Drill-hole Samples	Blast-hole Samples	Drill-hole Samples	Blast-hole Samples
Count	2045	19752	747	95815	2368	158772
Mean	1.07	1.18	0.34	0.42	0.57	0.48
Std. dev.	0.67	0.78	0.47	0.56	0.55	0.58
Minimum	0.13	0.01	0.01	0.00	0.00	0.00
Maximum	7.24	9.90	4.04	35.00	4.04	35.00

Table 2. Case study coordinates. Elevations represent the centers of the considered benches

	Case study 1		Case study 2		Case study 3	
	Min	Max	Min	Max	Min	Max
East	24,550	24,730	72,200	72,550	72,600	72,900
North	25,100	25,550	83,100	83,500	83,100	83,600
Elevation	3860	3940	2405	2455	2415	2465

Then, the optimal decision of size K is the solution of

$$f_K^* \equiv \underset{f \in F_K}{\operatorname{argmin}}\, H(X^f | X_f), \tag{5}$$

which minimizes the uncertainty of the field after K measurements. Interestingly, adopting some information theory identities (Cover and Thomas 2006), the optimal decision in (5) is equivalent to the

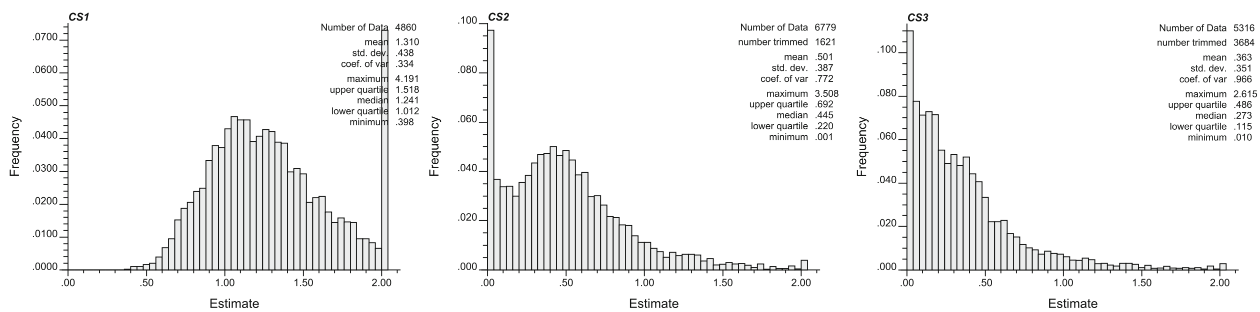

Figure 2. Grade mineral distributions and basic statistics for the available blast-holes. From left to right: CS1, CS2, CS3.

Table 3. Summary parameters SNESIM

	Case study 1	Case study 2	Case study 3
Number of benches (M)	6	6	6
Min data for kriging	4	4	4
Max data for kriging	16	16	16
Max data per octant (0 : not used)	0	0	0
Maximum search radii (x, y, z)	[80,80,10]	[80,80,10]	[80,80,10]
Min data for kriging	10	10	10

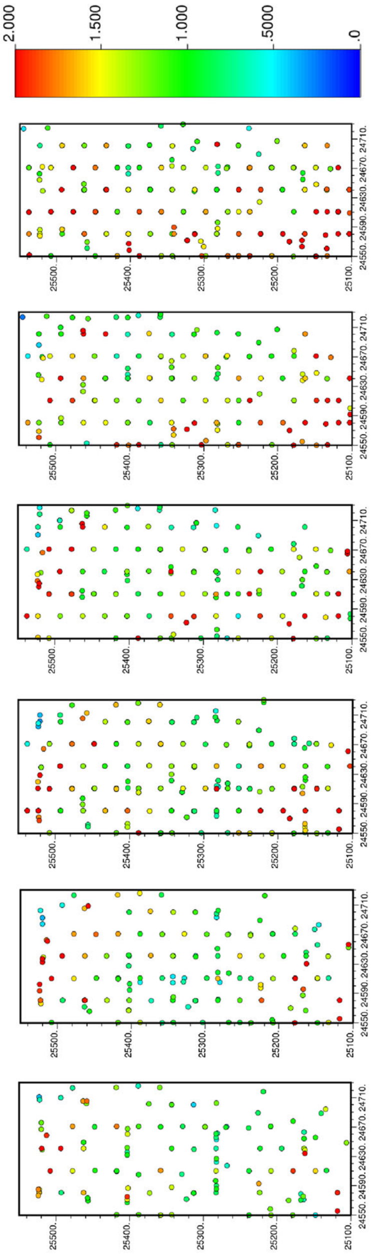

Figure 3. Drill-hole samples data for case study 1. From left to right: Benches 1–6. Colormap denotes the grade of Cu.

solution of the *maximum entropy problem* (Santibañez et al. 2019):

$$f_K^* = \arg\max_{f \in F_K} H(X_f) \tag{6}$$

that finds the K positions that jointly lead to the highest a priori (before the measurements) entropy. This principle is easier to implement as it requires marginal distributions and not the complete joint distribution of X, model for the implementation of Eq. 5.

The Sequential Adaptive Strategy From the general principle in Eq. 6, the focus is on the realistic *sequential problem* where the decision is taken sample by sample, and furthermore, the measurements taken in previous iteration of the algorithm are considered to upgrade the model (or adapt the model to the data) in the next iteration. More precisely, considering the set of previous sensed position $f_k^* = ((i_1, j_1), ..., (i_k, j_k))$ and the measurements of the field X taken at those selected places, i.e., $\bar{x} = (x_1, .., x_k)$, the solution of the $k+1$ position is given by the maximum entropy principle in (6) conditioned on $X_{(i_1,j_1)} = x_1$,.., $X_{(i_{k-1},j_{k-1})} = x_{k-1}$ and $X_{(i_k,j_k)} = x_k$, i.e.,

$$(i_{k+1}, j_{k+1}) = \arg\max_{(i,j) \in [A]\times[B] \setminus \{(i_l,j_l):l=1,..,k\}} H(X_{i,j} | X_{f_k^*} = \bar{x}), \tag{7}$$

where $f_{k+1}^* = (f_k^*, (i_{k+1}, j_{k+1})) \in F_{k+1}$. Then iterating this rule from $k = 1, .., K$ offers an adaptive and sequential solution for the problem of selecting the more informative K positions of the field.

The Regularized Adaptive Strategy The adaptive solution in Eq. 7 requires the knowledge of the statistics of X. In practice, this model is not available requiring to find a way to infer this model from empirical data. In the proposed solution, a training image is required that is deemed to represent the spatial continuity of the random function describing the spatial correlations in the underlying spatial model, and from this a model is estimated using conditional multiple-point simulations (Mariethoz and Caers 2014; Ortiz et al. 2012). In particular, a set of conditional probabilities of the form $\hat{\mu}_{X_{i,j}|X_S}(x_{i,j}|x_S)$ are estimated, where $i, j \in [A] \times [B]$, $S \subset [A] \times [B]$ denotes the conditioning positions (attributed to sensed data), and $x_{i,j} \in A$ and $x_S = (x_{i,j} : (i,j) \in S) \in A^{|S|}$, which is the information needed to implement (7).

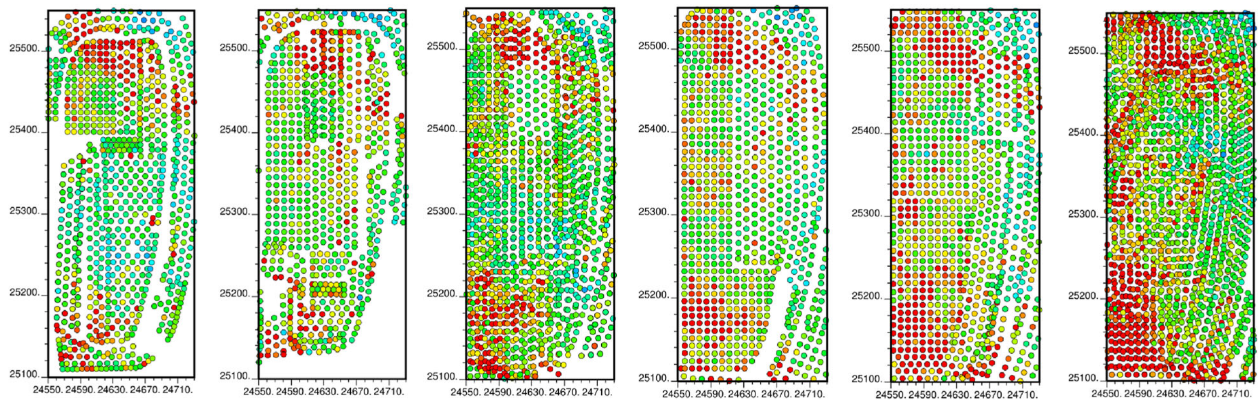

Figure 4. Blast-hole data for case study 1. From left to right: Benches 1–6. Colormap: the same as in Figure 3.

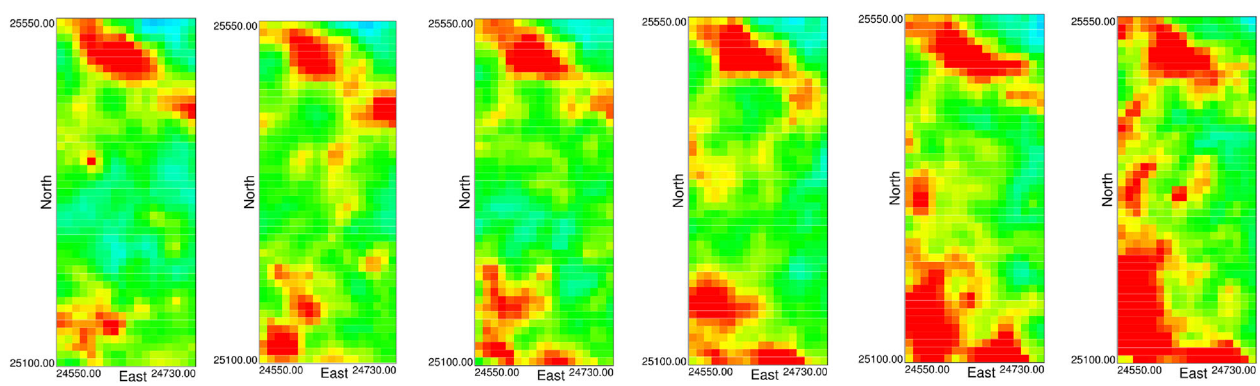

Figure 5. Ground truth estimated from drill-holes and blast-holes samples for case study 1. From left to right: Benches 1–6. Colormap: the same as in Figure 3.

Finally, a convex combination between the maximal entropy adaptive criterion in (7) and a regularization principle that promotes uniform covering of the sampling space is proposed. It has been found that a mixed rule that promotes a compromise between selecting the most informative points (from the conditional model and previous data) and a good cover of the sampling space is better than the rule based on a pure adaptive principle (Santibañez et al. 2019). More precisely, let $S_k = \left\{ (i_l^a, j_l^a) : l = 1, .., k \right\}$ denote the collection of k positions previously obtained. $X_{S_k} = \left\{ X_{i_1^a, j_1^a}, .., X_{i_k^a, j_k^a} \right\}$ denote the random vector at the locations S_k, and $x_{S_k} = \{x_1, .., x_k\}$ be the data collected at S_k. Then, the regularized rule is the solution of

$$\begin{aligned}(i_k^a, j_k^a) \\ &= \arg \max_{(i,j) \in [A] \times [B] \setminus S_{k-1}} \alpha \cdot H(X_{i,j} | X_{S_{k-1}} = x_{S_{k-1}}) \\ &\quad + (1 - \alpha) \cdot D((i,j), S_{k-1}).\end{aligned} \tag{8}$$

Note that the second term promotes a uniform sampling by using a distance criterion

$$D((i,j), S_{k-1}) = \min_{(\tilde{i}, \tilde{j}) \in S_{k-1}} d((i,j), (\tilde{i}, \tilde{j})), \tag{9}$$

where $d((i,j), (\tilde{i}, \tilde{j})) \equiv \sqrt{(\tilde{i} - i)^2 + (\tilde{j} - j)^2}$. Therefore, the proposed regularized rule corresponds to a global maximin strategy.

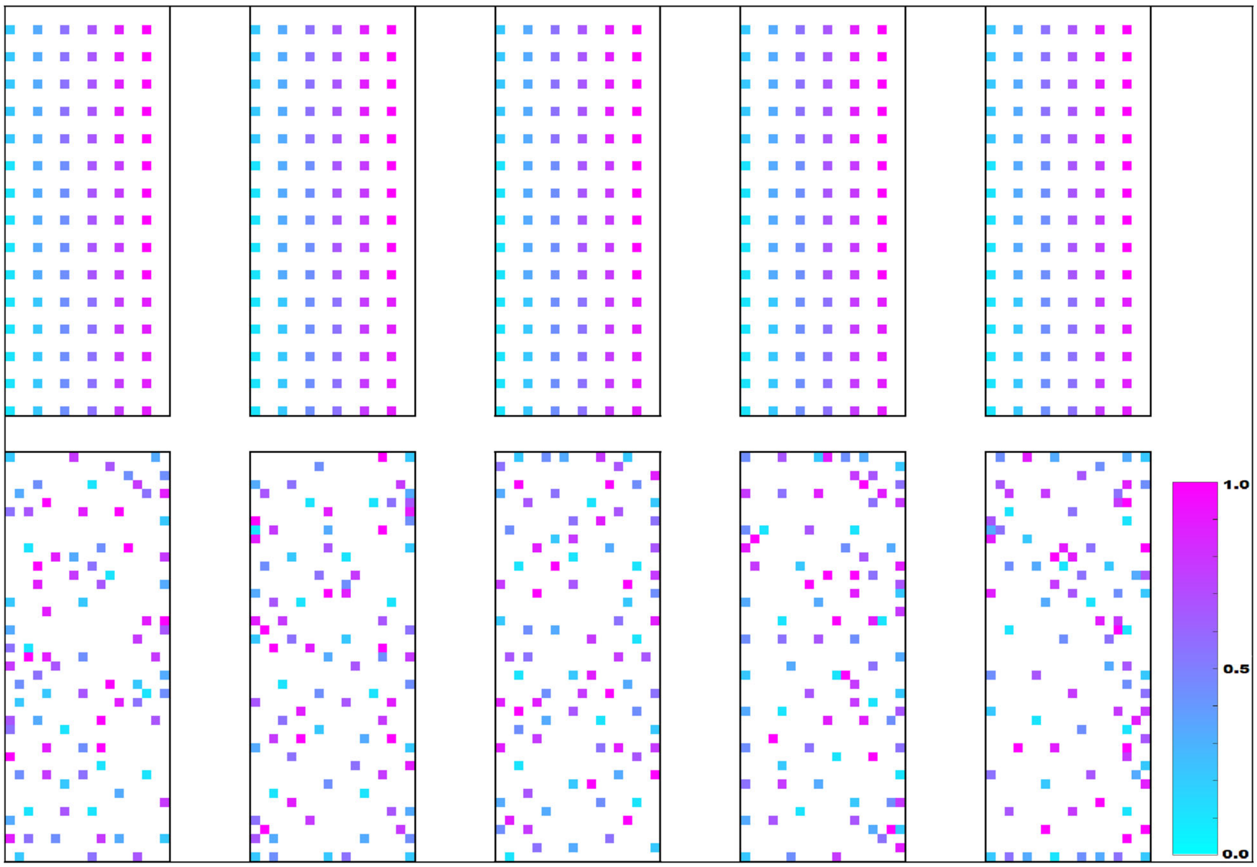

Figure 6. Samples for Case Study 1. From left to right: Benches *2–6*. Top: Samples from structured sampling. Down: Samples from adaptive sampling using *Cut-off* grade 1.012%. Colormap: Describe batch of samples in the order of the performed sampling.

OPTIMAL SAMPLING FOR GRADE CONTROL

In grade control, the target is to define the destination of every single block in the current mining bench. In order to achieve this task, the estimated grade of the block is considered and compared with the *Cut-off* grade. The *Cut-off* grade corresponds to the minimum grade required for a block to be considered for processing (expecting a positive economical benefit). Therefore, a block found to be above this *Cut-off* grade is considered to be ore, while a block below this grade is considered to be waste. Then, the grade control relies on a binary decision even when the value of the block grade corresponds to a continuous variable.

The *Cut-off* grade can be established by various methods. Its selection is related to a certain production objective, such as the use of resources or economic benefit. These objectives give rise to different types of targets, such as maximizing global economic benefits, immediate benefits and so on. The *Cut-off* grade does not have a fixed or predefined value, but instead it corresponds to a strategic variable that has important implications in the design and production of the mine.

The application of the adaptive sampling strategy to grade control is described in this section by considering a scenario where several benches are to be mined in a sequential process. The sampling strategy is used with the objective of improving the prediction of ore-waste contact in successive working benches. For every bench, the method works in two stages: first, the general features of the field are learned through a coarse regular sampling grid; then, the sampling adapts to the features learned from the regular grid and aims at characterizing the contacts between ore and waste for a better characterization of non-sensed blocks as learned from the benches previously mined out. In each bench, the process is

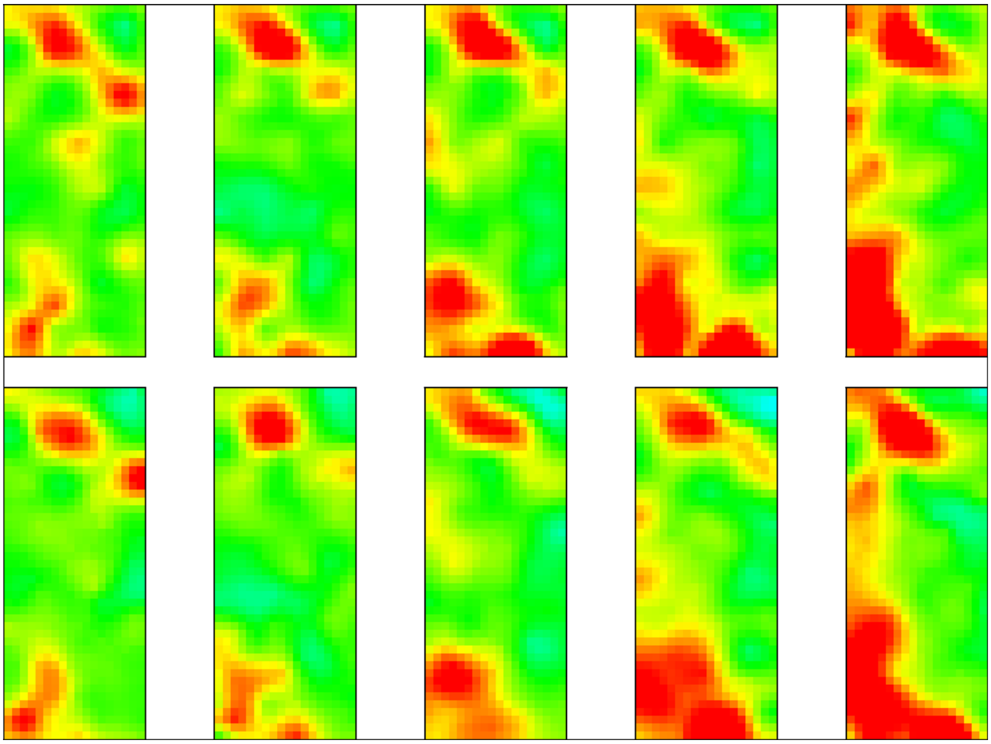

Figure 7. Estimated grade for Case Study 1. From left to right: Benches *2–6*. Top: Samples from structured sampling. Down: Samples from adaptive sampling using *Cut-off* grade 1.012%. Colormap: the same as in Figure 3.

applied sequentially, where current samples are used to make decisions about the subsequent locations to be sampled.

Proposed Methodology for Sampling

More precisely, let consider a collection of M benches, where every bench is indexed by $m \in \{1, \ldots, M\}$ an it is a rectangular grid, i.e., a *2-D* image, denoted by $I^m_{grade} \in \mathbb{R}^{AxB}$ of size A x B. On each bench m, K measurements will be taken to infer a model of the mineral distribution of non-sampled blocks at this bench. Therefore, the main problem is to define the best K locations using the framework presented in Section Entropy-based Adaptive Sampling Strategy. From these K samples, an estimation is made on all the non-sensed blocks of the bench producing what is defined as the block model. Finally, a hard threshold is applied on the block model using a *Cut-off* grade to create a binary *2-D* image denoted by $I^m_{o-w} \in A^{AxB}$ (with $|A| = 2$), where, consequently, non-sensed blocks are classified as ore or waste.

In the initial stage of this process ($m = 1$), the upper bench is considered. As no preliminary data are available from previous benches, the conditional entropies are drawn from an i.i.d distribution of ore and waste for implementing Eq. 8. This worst case scenario in terms of prior information for the inference, reduces Eq. 8 to the classical near regular sampling, since its second term in Eq. 9 dominates the optimization promoting distance as a criterion where a uniform coverage of the space is the optimal solution. Thus for the bench $m = 1$, K samples are distributed in a regular grid. An empirical estimation of the remaining (non-sensed) mineral grades is performed by *2-D* kriging. Finally, after applying the *Cut-off* grade the image I^1_{o-w} is obtained.

For benches $m = 2, \ldots, M$, the use of the previously estimated binary block model I^{m-1}_{o-w} for bench $m - 1$ is used as the training image (model) for inference of the conditional entropies used in Eq. 8 in the current bench. The key assumption made on this selection is that the previous bench $m - 1$ reflects the spatial distribution of the ore and waste blocks more accurately than the i.i.d. assumption made on the initial stage ($m = 1$). Using I^{m-1}_{o-w} as a training image along

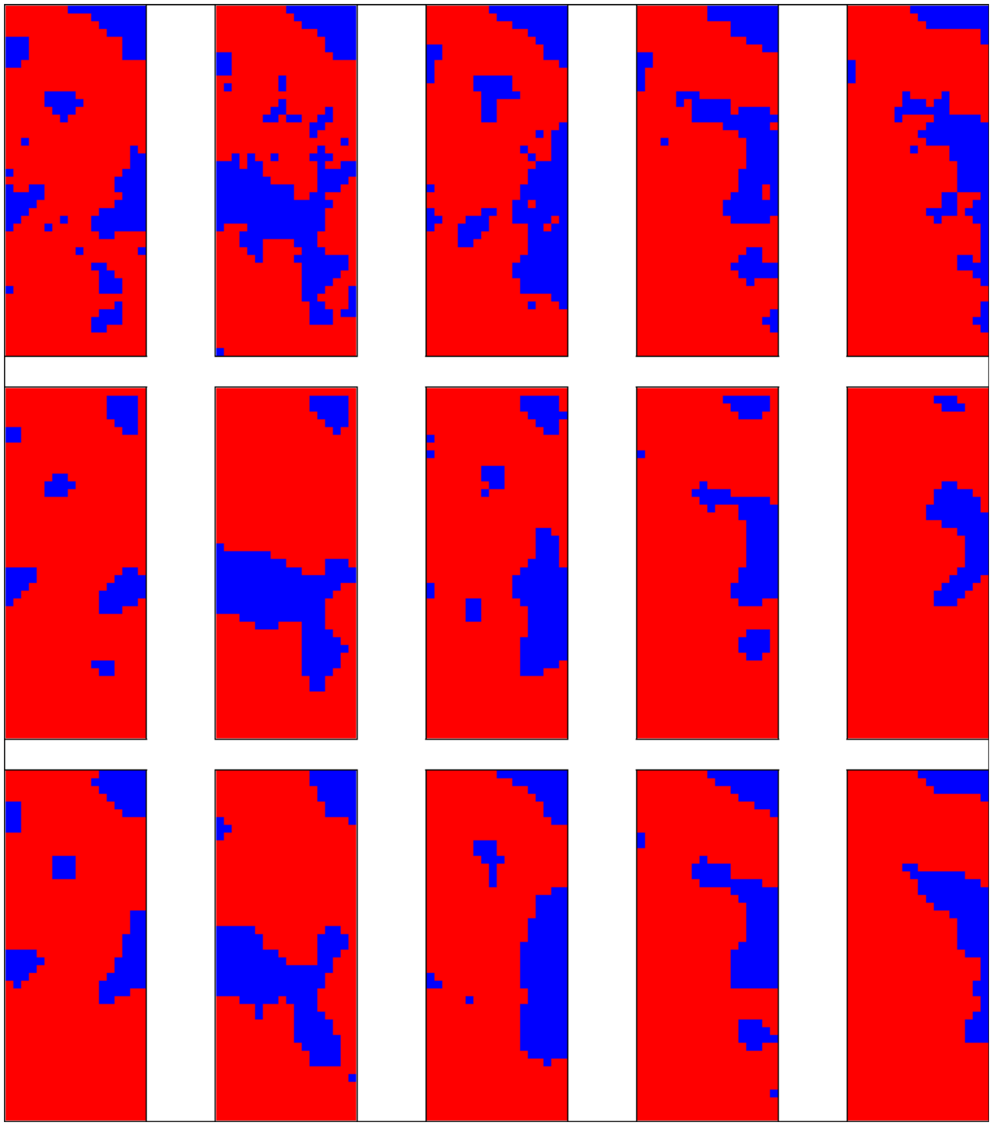

Figure 8. Estimated grade control for Case Study 1. From left to right: Benches *2–6*. From top to bottom: Ground truth, structured sampling, adaptive sampling using *Cut-off* grade 1.012%.

with *MPS*, in particular the *SNESIM* algorithm, and the previous $k-1$ measurements at the current bench m, it is possible to estimate the entropy map, $\hat{H}^{m,k}$, for this bench. This is required for the implementation of the sampling rule in Eq. 8. More precisely, the selection rule for the kth sample at bench m given the previously sampled locations S^m_{k-1} corresponds to the solution of:

$$
\begin{aligned}
(i^a_k, j^a_k)^m &= \arg \max_{(i,j)\in[M]\times[M]\setminus S^m_{k-1}} \alpha \hat{H}^{m,k}(X^m_{i,j} | X^m_{S^m_{k-1}} \\
&= x_{S^m_{k-1}}) + (1-\alpha)
\end{aligned} \tag{10}
$$

The solution of Eq. 10 places the new samples in the locations with an optimal balance between maximum conditional entropy (information criterion) and the maximum distance to previously sampled positions (regularization). It is important to mention that in the proposed practical implementation of Eq. 8, the adaptive sequential sampling strategy selects a new batch of s samples at every sampling step. This iterative strategy is repeated until the K samples are obtained. For completeness, a schematic representation of the sampling rule in (10) is illustrated in Figure 1. Finally, the pseudo-code of the

Figure 9. Confusion Matrix for Case Study 1. From left to right: Benches *2–6*. Top: Structured sampling. Down: Adaptive sampling using *Cut-off* grade 1.012%.

Table 4. Performance error summary for case study 1

	Case study 1					
	Cut-off grade 1.102		Cut-off grade 1.241		Cut-off grade 1.518	
	STR	ADA	*STR*	ADA	*STR*	ADA
Bench 2	0.112	**0.069**	0.133	**0.114**	0.059	**0.055**
Bench 3	0.138	**0.106**	0.104	**0.102**	0.066	**0.064**
Bench 4	0.109	**0.082**	0.091	**0.086**	0.053	**0.051**
Bench 5	0.084	**0.048**	0.086	**0.083**	0.090	**0.083**
Bench 6	0.111	**0.060**	0.127	**0.108**	0.109	**0.097**

implementation of Eq. 10 is presented in Appendix A: Pseudo-code.

Classification Process

Concerning the final ore-waste classification, after the K samples are taken in every bench, ordinary kriging is performed using the routine provided in *GSLIB* (Deutsch et al. 2000). A variogram model is fit in each case over the full bench, performing the estimation by using a minimum of min_S samples and a maximum of max_S, with a search radius of rad_S [m]. A block discretization of $nA \times nB \times nZ$ points per block is used. The resulting estimated block model for the bench is binarized by applying a *Cut-off* grade. The resulting binary model (image I^m_{o-w}) represents the estimated ore and waste blocks that are used in short-term planning.

Table 5. Performance error summary for case study 2

	Case Study 2					
	Cut-off grade 0.220		Cut-off grade 0.445		Cut-off grade 0.692	
	STR	ADA	*STR*	ADA	*STR*	ADA
Bench 2	0.048	**0.043**	0.100	**0.096**	0.112	**0.080**
Bench 3	0.039	**0.032**	0.109	**0.097**	0.091	**0.068**
Bench 4	0.037	**0.034**	0.078	**0.066**	0.057	**0.045**
Bench 5	0.055	**0.038**	0.036	**0.024**	0.036	**0.026**
Bench 6	0.031	**0.015**	0.025	**0.010**	0.013	**0.010**

Table 6. Performance error summary for case study 3

	Case Study 3					
	Cut-off grade 0.115		Cut-off grade 0.273		Cut-off grade 0.486	
	STR	ADA	*STR*	ADA	*STR*	ADA
Bench 2	0.067	**0.048**	0.060	**0.039**	0.067	**0.054**
Bench 3	0.039	**0.031**	0.057	**0.030**	0.030	**0.014**
Bench 4	0.049	**0.029**	0.052	**0.033**	0.054	**0.049**
Bench 5	0.051	**0.037**	0.053	**0.041**	0.053	**0.034**
Bench 6	0.037	**0.021**	0.047	**0.035**	0.038	**0.034**

Table 7. Economical profit estimation for case study 1

	Case Study 1					
	Cut-off grade 1.102		Cut-off grade 1.241		Cut-off grade 1.518	
	STR	ADA	*STR*	ADA	*STR*	ADA
Bench 2	33.574	**34.146**	13.483	**13.726**	**2.675**	1.737
Bench 3	**28.581**	28.462	**10.419**	9.439	1.284	**2.520**
Bench 4	33.562	**34.150**	15.423	**16.114**	**6.295**	5.792
Bench 5	41.065	**41.590**	23.272	**23.814**	11.174	**11.502**
Bench 6	46.401	**47.135**	27.623	**28.165**	**16.420**	15.427
Global	183.183	**185.483**	90.221	**91.257**	**37.849**	36.977

In MM *US*$

Table 8. Economical profit estimation for case study 2

	Case study 2					
	Cut-off grade 0.220		Cut-off grade 0.445		Cut-off grade 0.692	
	STR	ADA	*STR*	ADA	*STR*	ADA
Bench 2	49.281	**49.320**	23.729	**23.992**	6.096	**6.704**
Bench 3	39.309	**39.458**	**16.727**	16.362	−5.363	**−4.555**
Bench 4	**47.299**	47.262	23.257	**23.646**	8.957	**9.425**
Bench 5	36.460	**36.703**	**19.278**	19.181	10.392	**10.714**
Bench 6	9.790	**9.994**	1.145	**1.330**	−4.058	**−3.937**
Global	182.139	**182.735**	84.137	**84.512**	16.023	**18.351**

In MM *US*$

Table 9. Economical profit estimation for case study 3. In MM *US*$

	Case study 3					
	Cut-off grade 0.115		Cut-off grade 0.273		Cut-off grade 0.486	
	STR	ADA	*STR*	ADA	*STR*	ADA
Bench 2	17.685	**17.974**	**17.382**	17.371	**3.059**	2.419
Bench 3	11.577	**11.674**	10.136	**11.130**	−2.159	**−1.943**
Bench 4	10.651	**10.904**	9.772	**10.213**	**−2.440**	−2.553
Bench 5	14.629	**14.859**	**14.111**	13.993	1.192	**1.725**
Bench 6	16.038	**16.288**	**15.316**	15.262	3.136	**3.964**
Global	70.580	**71.699**	66.717	**67.969**	2.787	**3.611**

CASE STUDY AND EXPERIMENTS

The proposed methodology has been applied to three different cases, two of them coming from the same ore deposit. The corresponding databases consist of drill-hole composites widely spaced and denser blast-hole samples, which are used to validate the sampling strategy. The two projects correspond to massive porphyry copper deposits that are currently under operation (more details can be found in supplementary material).

In this section, the implementation parameters are provided and the results of the first case study are described in detail. For sake of space, only the relevant results for the other two cases are presented (additional details and results can be found in supplementary material and Appendix B: Additional Experimental Results).

Database Description

Case studies 1 and 2 come from the same mining project, but from different areas of the open pit. These areas have already been mined out, providing blast-hole samples to test the hypothesis that adaptive sampling can achieve better discrimination between ore and waste than classical regular sampling. For this mining project, the available blast-hole database consists of nearly regularly spaced samples taken in a grid of approximately $10m$ by $10m$. The database contains areas without samples; therefore, it was split into two informed sectors to generate case studies 1 and 2. Finally, case study 3 comes from a different mining project with similar sampling conditions.

Blast-hole data are migrated to a regular grid creating 6 consecutive benches uniformly sampled for each case study. From the selected X–Y section of the raw information provided by the blast-hole samples, a set of dense benches has been created that completely describe the ore distribution of the *3-D* deposit.

The basic statistical information used to build the case studies is summarized in Tables 1 and 2.

The statistical distributions of blast-holes grades present in the analyzed case studies are described in Figure 2, along with their basic statistics.

Construction of Validation Block Model

From the unevenly distributed blast-hole and drill-hole data, a fully informed block model is obtained by performing block ordinary kriging using the kt3d routine of *GSLIB*. Finally, six consecutive benches were considered as ground truth for every case study by considering the blocks inferred from the densely sampled available information. The parameters for the kriging estimation are presented in Table 3.

In order to illustrate the density of available information for every single bench, Figures 3 and 4 show the drill-hole composites and blast-hole samples for the first case study. The block model estimated by ordinary kriging is displayed for these data in Figure 5. Systematic descriptions of the other two case studies can be found in the supplementary material.

Experimental Results: Adaptive Sampling Strategy

As mentioned, the value for the *Cut-off* grade corresponds to an operational decision, then in order to demonstrate the benefit of the proposed sampling approach, results for different *Cut-off* grade values are presented. In this work, the *Cut-off* grade has not been defined under operational or economic principles, instead in the presented results the *Cut-off* grades have been considered from the three quartiles of the distribution of available data (q25, q50 and q75). Thus, the focus of this work is maintained on the effect on ore-waste classification as a function of the sampling strategy.

To illustrate the implementation and outcomes of the adaptive sampling strategy, full details are provided for the first case study. For case studies 2 and 3, only summary figures and results are presented, in the understanding that the procedure is similar to that illustrated for case study 1 (details for case study 2 and 3 can be found in supplementary material).

For each sampling strategy, $\frac{100}{9}\% \approx 11.\bar{1}1\%$ of the available locations for each bench were sampled ($K = \frac{1}{3} * A_m * \frac{1}{3} * B_m$). For case study 1, the proposed samples are presented in Figure 6. The outcomes for kriging estimation from the proposed samples are displayed in Figure 7 for the benches 2, 3, 4, 5 and 6. Three *Cut-off* grade values, corresponding to the three quartiles of the grade distribution, were considered to evaluate the differences between the classical structured sampling (STR) and the proposed adaptive sampling (ADA) approach (Fig. 8).

The results for the case of regular sampling have considered perturbations about the location of the origin of the grid. Thus, the average results are always considered from different origins of the regular grids with the aim to avoid a bias from unsampled locations in the edges of the fields.

From the results presented in Figure 9, it is clear that the number of misclassified blocks is reduced with the adaptive sample in comparison with the structured classical approach.

Performance Assessment

Binary Image Inference Performance We begin by comparing the performances in terms of image recovery achieved by the proposed adaptive sampling strategy with respect to the classical structured sampling. The results are summarized in the Tables 4, 5, and 6 for case studies 1, 2, and 3, respectively. Here the performance is evaluated in the binary ore-waste estimated image for each bench by computing the proportion of misclassified blocks. For all the results Tables 4, 5, 6, 7, 8, and 9 presented in this work, the best performances have been highlighted in bold for each case study and bench.

From the point of view of the binary allocation of blocks (as ore or waste), the adaptive sampling-based method provides a better overall classification of the blocks. This behavior is consistent for the proposed *Cut-off* grades along all three case studies presented.

In general, as the procedure advances down the benches (as m increases), the improvement increases as well in terms of the reduction of error as compared to the classical sampling method. Thus, the adaptive strategies outperform the classical sampling method, achieving a consistent improvement over all benches.

It is worth mentioning that it is possible to improve the performance of the proposed technique by feeding the model for the training image of bench m with all the information available from the previous benches ($1, .., m-1$) instead of just from the bench above (bench $m-1$).

Economic Performance In order to provide a summary of the economic impact related with the sampling strategy in the ore-waste selection process, a brief evaluation of profit is presented taking into consideration some relevant scenarios. The variables and estimations considered for this economic analysis are detailed in Appendix C: Economic Evaluation. The achieved results are summarized in Tables 7, 8, and 9. For the analysis presented, both the price of copper and production costs have been considered as constant. These are only presented to explain the calculation of *Cut-off* grade (for which three values have been studied in order to demonstrate the general improvement provided by the proposed strategy).

Although in some benches the economic results are variable, the overall result obtained shows a systematic improvement over the classic sampling scheme.

Even though the binary assignment as ore-waste is systematically better with the proposed adaptive approach (ADA) in terms of mean global error, the current grades are not considered for this evaluation. Therefore, cases can be found where when block grades are close to the *Cut-off* value, then the classification may fail and the economic value for the bench may decrease.

Negative profits may occur when the processing plants lack ore, and marginal material must be sent to fulfill the production requirements, which may have a negative value, but will likely be higher than dumping these blocks in the waste dump. Now, in most practical cases the *Cut-off* will be defined to generate profit, rather than minimize loss.

CONCLUSIONS

In this work, the problem of optimal sampling in the context of short-term planning and the task of classifying blocks to be processed as waste or ore has been addressed. The problem has been formalized and its validation has been presented through the use of subsets of actual mining data.

The proposed methodology takes advantage of the information available from the locations previously sampled, allowing to improve the performance as compared with the classical non-adaptive sampling schemes that has been used for advanced drilling tasks. The proposed strategy has been validated with real blasting data from the exploitation of two copper mines.

From the results obtained across the three analyzed scenarios, it is possible to see that in terms of both error in image reconstruction and global economic value, the proposed methodology achieves better performance than the regular grid based sampling strategy.

APPENDIX A: PSEUDO-CODE

The implemented pseudo-code that summarizes the proposed framework is shown here.

Initialization

Reference image $\bar{X}$, the regularization term α, the batch size r, the threshold tsh_{grade} for ore-waste selection and the number of samples to take K

```
/* Output: set of sampled positions */
```
$f_K^{bench} = \emptyset$
```
/* Inputs and variables. H^k Estimated Remaining entropy for k-1 samples */
```
$\mathcal{H}^k = ones(size(\bar{X}))$, $\mathcal{D} = inf(size(\bar{X}))$, $distMean = inf$

Computation

for $bench \leftarrow 1$ **do**
- $f_K^1 = Regular\_Sampling\_Approach(K)$
- $\hat{X}^1 = Kriging\_From\_Samples(f_K^1)$
- $\hat{I}^1 = \hat{X}^1 \geq tsh_{grade}$

end

for $bench \leftarrow 2$ **to** B **do**
- `/* Estimate initial TI from the upper bench */`
- $\hat{TI}^{bench} \leftarrow \hat{I}^{bench-1}$
- `/* Init the adaptive sampling process */`
- $f_K^{bench} = \emptyset$, $f_K = \emptyset$
- **for** $k \leftarrow 1$ **to** K **do**
  - $f_{k-1}^{bench} \leftarrow f_K$
  - `/* Update MPS simulations from previous available samples */`
  - **if** $Criterion\_To\_Update\_MPS\_Realizations(k, r)$ **then**
    - $\mathcal{H}^k = Entropy_{Estimated\_by\_MPS\_\&\_Samples}(\hat{TI}^{bench}, X_{f_{k-1}}^{bench}, size(\bar{X}^{bench}), f_{k-1}^{bench})$
    - $\mathcal{D} = Estimated\_Distances\_From\_Sampled\_Locations(size(\bar{X}), f_{k-1}^{bench})$
    - $distMean = Mean(\mathcal{D})$
    - $\hat{X}^{bench} = Kriging\_From\_Samples(f_{k-1}^{bench})$
    - $\hat{I}^{bench} = \hat{X}^1 \geq tsh_{grade}$
    - $\hat{TI}^{bench} \leftarrow \hat{I}^1$
  - **else**
    - $\mathcal{H}^{bench} = \mathcal{H}^{bench} - Local\_Mutual\_Information_{empirical}(f_{k-1}^{bench}(k-1))$
    - $\mathcal{D} = \mathcal{D}. \cdot Radial\_Attenuation\_Centred\_At\_Location(size(\bar{X}^{bench}), distMean, k-1)$
  - **end**
  - `/* Available locations from complement of previous samples */`
  - $\hat{f}_{k-1}^{bench} = Complement\_Set\_Of\_F(f_{k-1}^{bench})$
  - `/* Regularized Criterion by Mixing of Entropy and Distance */`
  - $\mathcal{M} = \alpha \cdot \mathcal{H}^{bench} + (1.0 - \alpha) \cdot \mathcal{D}$
  - `/* Choose a location randomly from the set of non sampled locations with maximal value for the objective function */`
  - $f_K^{bench}(k) = Select\_Random\_Location\_From\_Maximum\_Criterion(\mathcal{M}_{\hat{f}_{k-1}^{bench}})$
- **end**

end

APPENDIX B: ADDITIONAL EXPERIMENTAL RESULTS

Case study 1: *Cut-off* grade 1.241%

With *Cut-off* grade 1.241%, Figures 10, 11, 12, and 13 describe the achieved outcome. Figure 13 provides a summary of the confusion matrices for case study 1 considering a *Cut-off* grade of 1.241%.

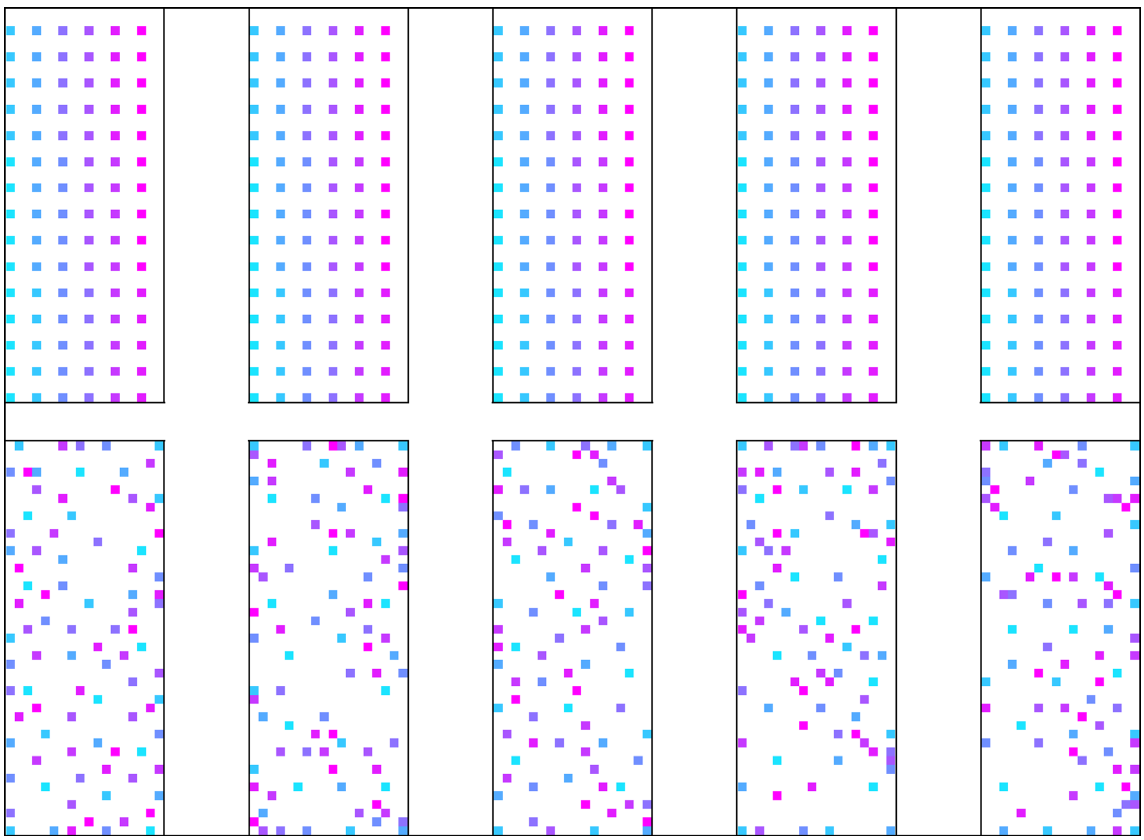

Figure 10. Samples for Case Study 1. From left to right: Benches *2–6*. Top: Samples from structured sampling. Down: Samples from adaptive sampling using *Cut-off* grade 1.241%.

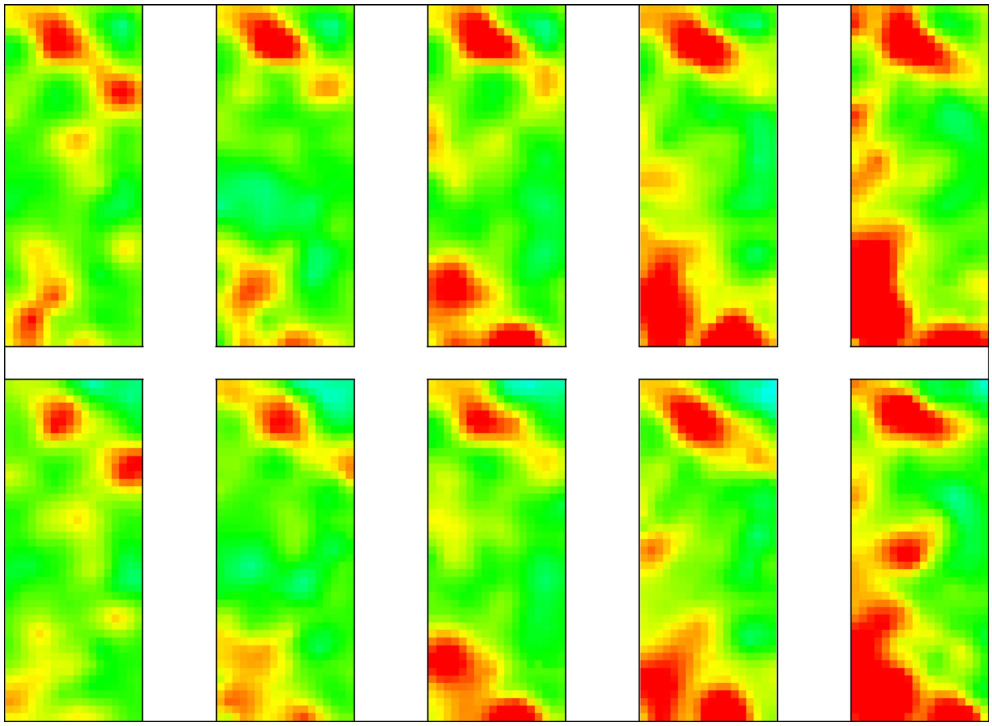

Figure 11. Estimated grade for Case Study 1. From left to right: Benches *2–6*. Top: Kriging from structured sampling. Down: Kriging from adaptive sampling using *Cut-off* grade 1.241%.

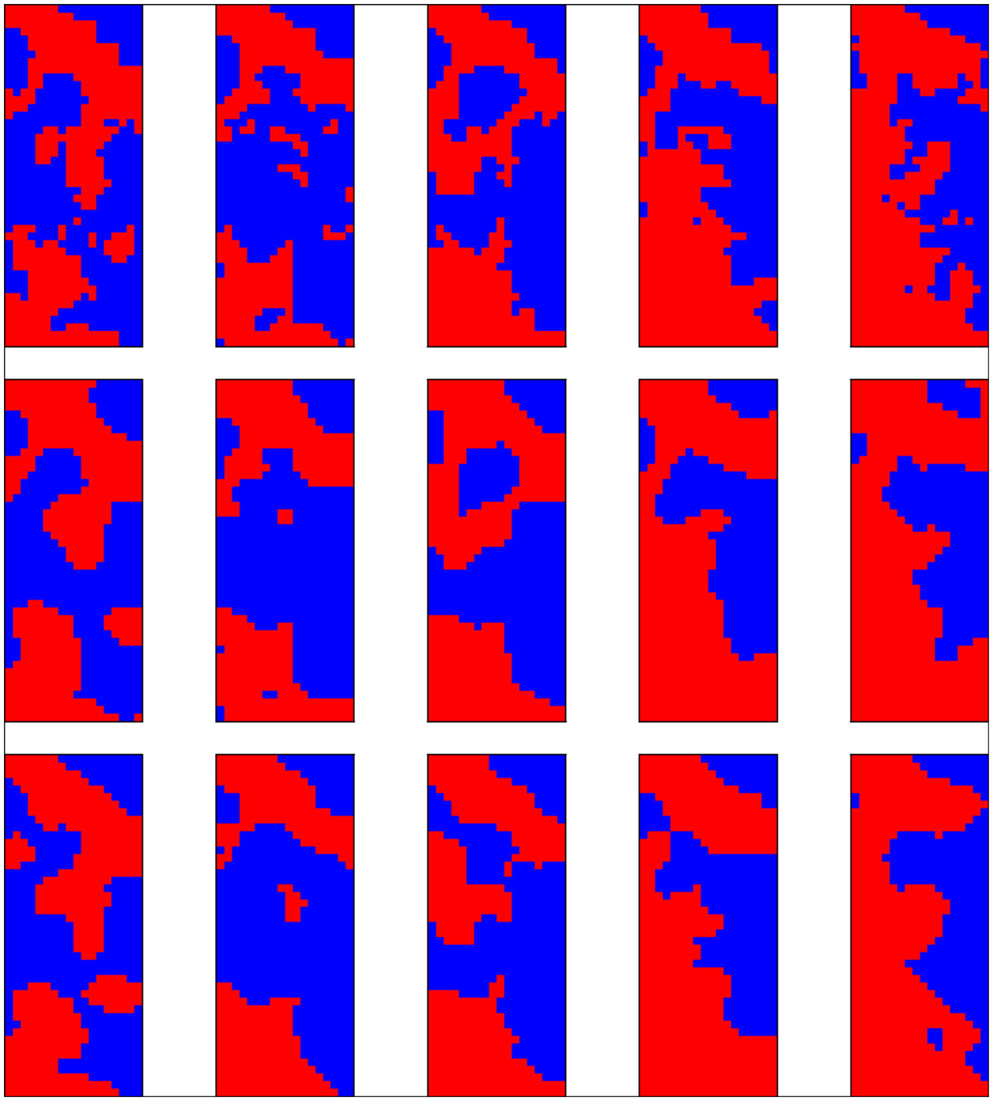

Figure 12. Estimated grade control for Case Study 1. From left to right: Benches *2–6*. From top to bottom: Ground truth, structured sampling, adaptive sampling using *Cut-off* grade 1.241%.

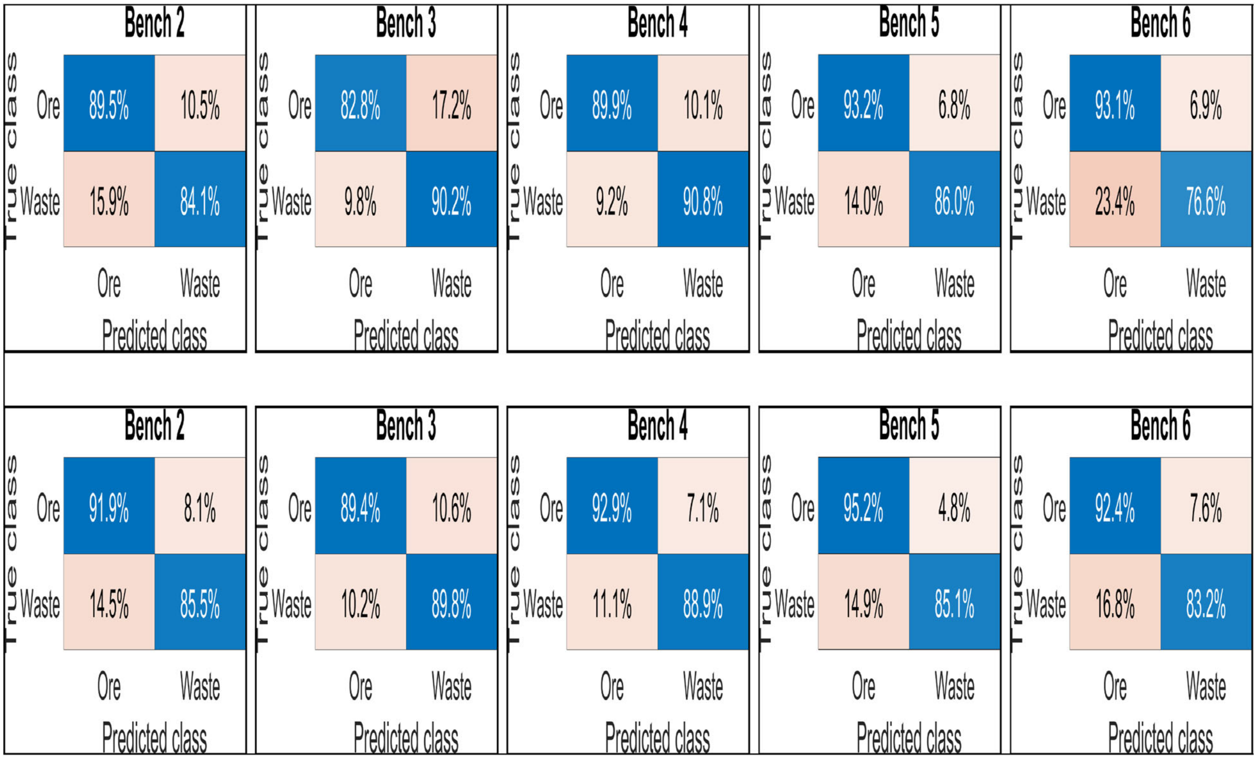

Figure 13. Confusion Matrix for Case Study 1. From left to right: Benches *2–6*. Top: Structured sampling. Down: Adaptive sampling using *Cut-off* grade 1.241%.

Ore-Waste Discrimination with Adaptive Sampling Strategy

Case study 1: *Cut-off* grade 1.518%

Considering the *Cut-off* grade 1.518%, the outcome for the benches 2, 3, 4, 5 and 6 is described in Figures 14, 15, 16, and, 17. The Figure 17 provide the performance summary in terms of the confussion matrices for case study 1 and *Cut-off* grade 1.518%.

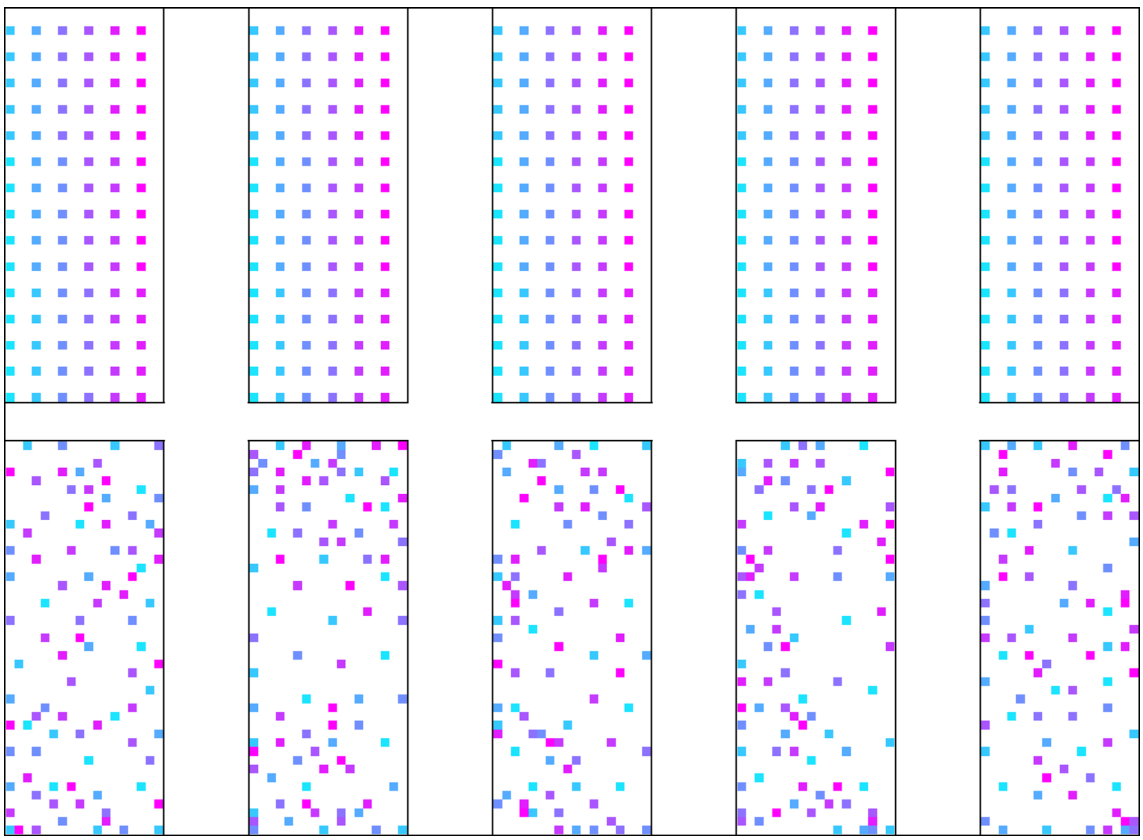

Figure 14. Samples for Case Study 1. From left to right: Benches *2–6*. Top: Samples from structured sampling. Down: Samples from adaptive sampling using *Cut-off* grade 1.518%.

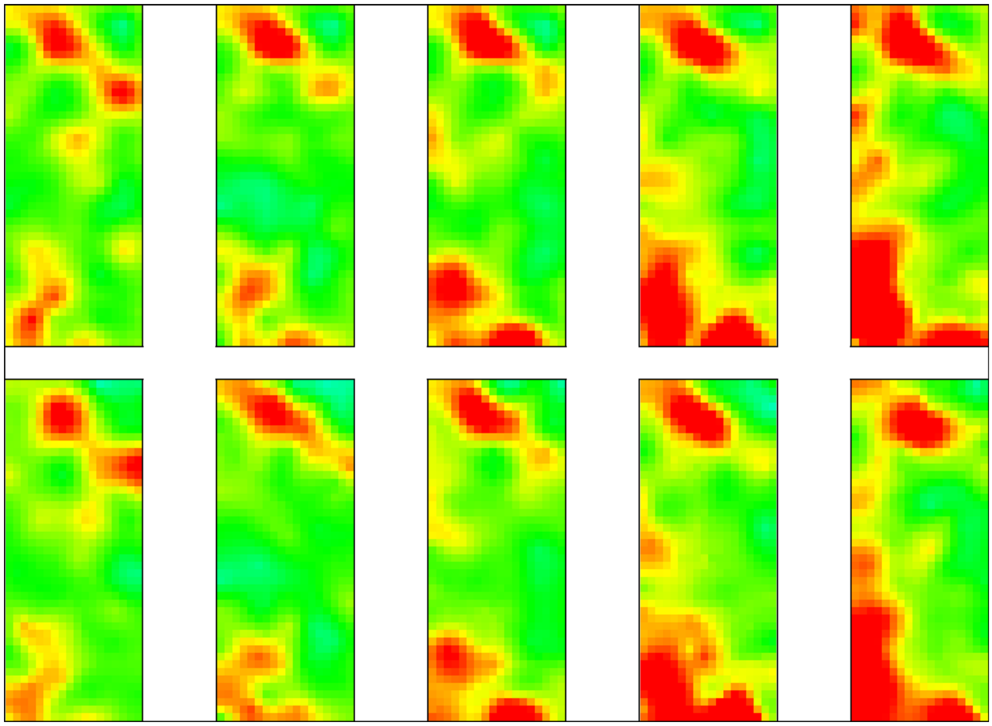

Figure 15. Estimated grade for Case Study 1. From left to right: Benches *2–6*. Top: Kriging from structured sampling. Down: Kriging from adaptive sampling using *Cut-off* grade 1.518%.

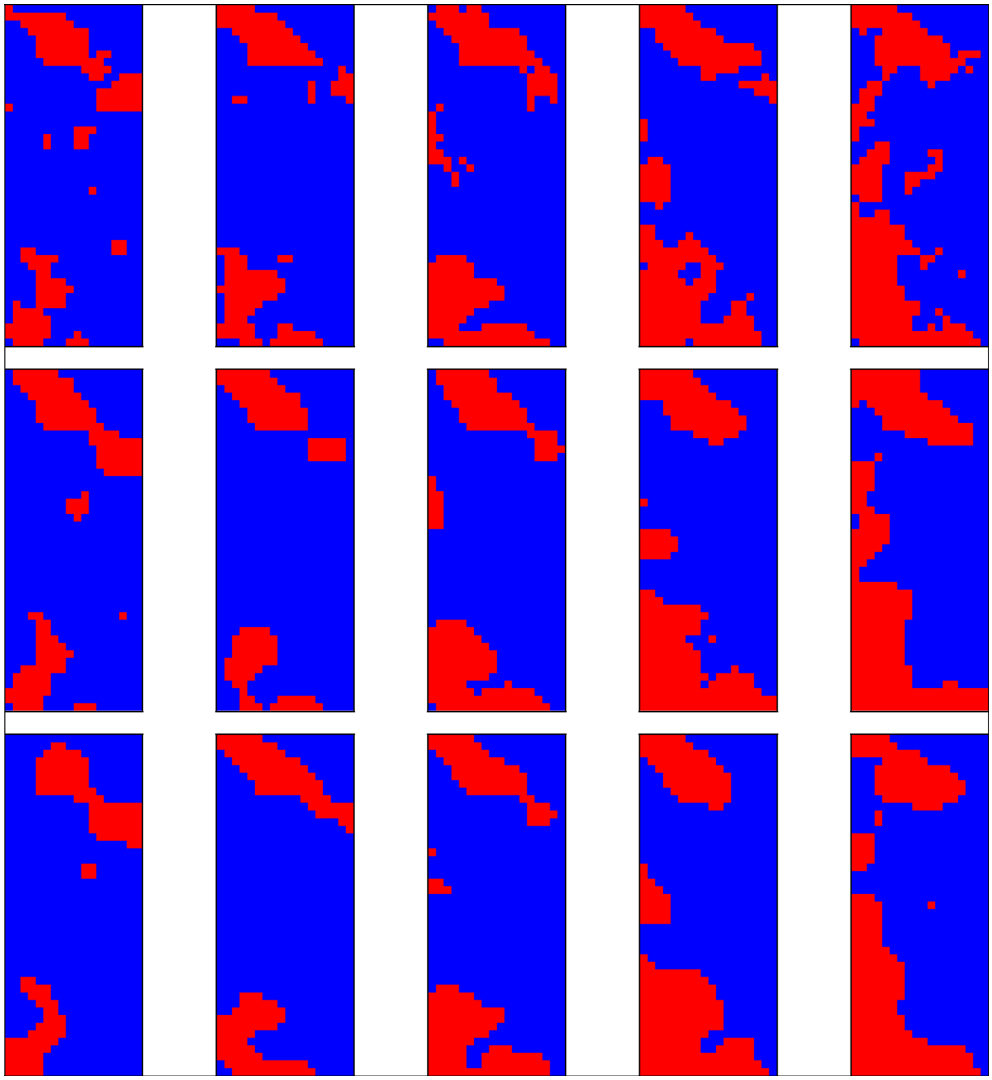

Figure 16. Estimated grade control for Case Study 1. From left to right: Benches *2–6*. From top to bottom: Ground truth, structured sampling, adaptive sampling using *Cut-off* grade 1.518%.

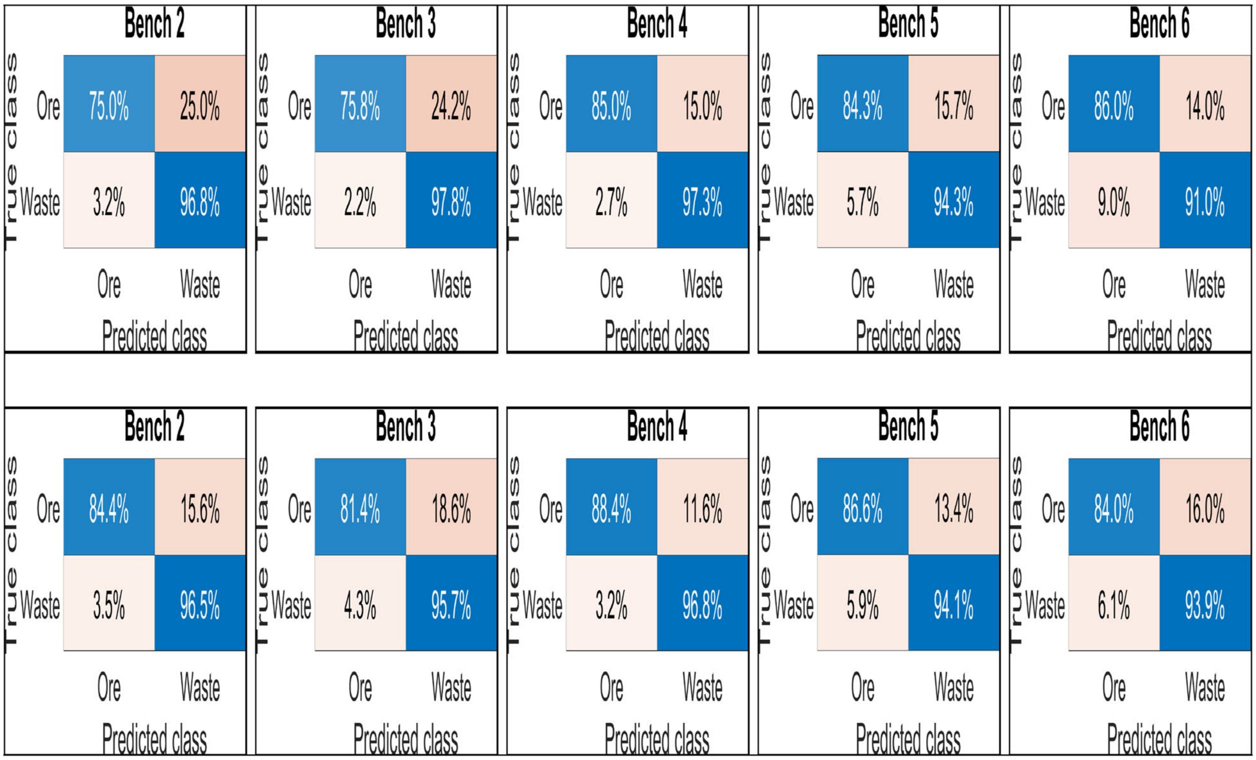

Figure 17. Confusion Matrix for Case Study 1. From left to right: Benches *2–6*. Top: Structured sampling. Down: Adaptive sampling using *Cut-off* grade 1.518%.

APPENDIX C: ECONOMIC EVALUATION

In order to perform a realistic evaluation of cost and profit, several considerations must be defined for the mining and waste processing. In particular, in this work the variables summarized in Table 10 have been considered, where a range of realistic values for these variables is proposed. This analysis takes into account the block size (fixed to 1000 m3, the same block size used in the experimental section), the mining cost by mined ton, the metallurgical recovery, and the stripping ratio (the proportion of tons of expected waste material and ore material). For the purpose of the present analysis, the economic costs considered were the processing cost by processed ton, the price and selling cost per pound of copper.

In practice, the *Cut-off* grade, *Cg*, can be defined by,

$$Cg = \frac{10000 \cdot (Cost^m * (1 + S_r) + Cost^p)}{(Price_{Cu} - Cost^s) \cdot R_m \cdot 2204.6}, \qquad (11)$$

where the value 2204.6 corresponds to the conversion factor from pounds to tons.

Given the set of values for the considered parameters, it is possible to estimate the profit of a block. For a block with an estimated grade under the *Cut-off* grade value, for simplification it has been defined that the cost to process the waste block in the dump facilities is considered equal for each dumped block without taking into account its actual mineral grade or another variables. For this brief analysis, the revenue from dumping the block is zero (in practice if the block has a zero profit, since it has already been mined then it could be considered as stock pile). Thus:

$$Profit_d = -Cost^m \qquad (12)$$

In the case of a block estimated as ore (estimated block grade is above the *Cut-off* grade), it is processed by the mine and its benefit is calculated considering the content of metal in percentage of tons of copper as,

Table 10. Variables considered in the profit analysis

Variable	Symbol	Initial value	Units	Range
Block size	$size_b$	1000	m3	–
Density	d_b	2.70	t/m3	2.3–2.9
Mining cost	$Cost^m$	3.0	US$/$t_{mined}$	1.70–3.50
Stripping ratio	S_r	2.50	-	2.50–3.00
Processing cost	$Cost^p$	10.00	US$/$t_{processed}$	7.00–12.00
Price of a lb of Cu	$Price_{Cu}$	2.30	US$/lb	2.50–2.10
Selling cost by lb_{Cu}	$Cost^s$	0.50	US$/lb	0.30–0.55
Metallurgical recovery	R_m	85.00	%	80.00–90.00

$$Cont_m = \frac{Grade_b}{100.0} \cdot size_b \cdot d_b \qquad (13)$$

The content of recovered metal in tons of copper is estimated as

$$Cont_m^r = \frac{Cont_m \cdot R_m}{100.0} \qquad (14)$$

Then, the revenue from mining the block is estimated by,

$$Rev_b^m = Cont_m^r \cdot 2204.6 \cdot (Price_{Cu} - Cost^s), \qquad (15)$$

while the processing cost of the block is estimated as

$$Cost_b^p = size_b \cdot d_b \cdot (Cost^m + Cost^p). \qquad (16)$$

Finally, the profit of the processed block is defined by,

$$Profit_b^p = Rev_b^m - Cost_b^p. \qquad (17)$$

Therefore, from Eqs. 12 and 17 and the cost of processing ore and waste blocks, it is possible to estimate the profit or loss of any block.

Considering the *Cut-off* grade of the experimental analysis and Eq. 11, the appropriate economic parameters have been estimated. The initial values shown in Table 10 are provided as an example to obtain the *Cut-off* grade as the first quartile in Case Study 1. For each case study and for every empirical *Cut-off* grade, the best set of parameters has been estimated in order to perform the economic analysis.

Then, from the experimental data the profit of every block has been evaluated in the proposed scenarios for the sampling strategies under analysis. Given an estimated bench block model from a specific sampling strategy, the economic profit can be calculated as the sum of the profit for each block conforming this bench.

ELECTRONIC SUPPLEMENTARY MATERIAL

The online version of this article (https://doi.org/10.1007/s11053-020-09625-3) contains supplementary material, which is available to authorized users.

REFERENCES

Abzalov, M. Z., Menzel, B., Wlasenko, M., & Phillips, J. (2010). Optimisation of the grade control procedures at the yandi iron-ore mine, western australia: Geostatistical approach. *Applied Earth Science*, *119*(3), 132–142. https://doi.org/10.1179/1743275811Y.0000000007.

Blackwell, G. (1998). Relative kriging errors—a basis for mineral resource classification. *Exploration Mining Geology*, *7*, 99–105.

Boisvert, J. B., Leuangthong, O., Ortiz, J. M., & Deutsch, C. V. (2008). A methodology to construct training images for vein type deposits. *Computers & Geosciences*, *34*(5), 491–502.

Chiles, J. P., & Delfiner, P. (2012). *Geostatistics: Modeling spatial uncertainty* (2nd ed.). Hoboken: Wiley.

Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory* (2nd ed.). New York: Wiley.

Deutsch, C. V., & Journel, A. G. (1998). *GSLIB—geostatistical software library and user's guide*. Oxford: Oxford University Press.

Deutsch, C. V., Magri, E., & Norrena, K. (2000). Optimal grade control using geostatistics and economics: methodology and examples. *Society of Mining Engineers SME Transactions Transactions*, *308*, 43–52.

Dimitrakopoulos, R., & Godoy, M. (2014). Grade control based on economic ore/waste classification functions and stochastic simulations: Examples, comparisons and applications. *Mining Technology*, *123*(2), 90–106. https://doi.org/10.1179/1743286314Y.0000000062.

Francois-Bongarcon, D. (1983). The practice of sampling of broken ores. *Canadian Institute of Mining, Metallurgy & Petroleum CIM Bulletin*, *86*(970), 75–81.

Hassanipak, A. A., & Sharafodin, M. (2004). GET: A function for preferential site selection of additional borehole drilling. *Exploration & Mining Geology*, *13*(1–4), 139–146.

Journel, A., & Kyriakidis, P. C. (2004). *Evaluation of mineral reserves—a simulation approach*. Oxford: Oxford University Press.

Journel, A. G. (1974). Geostatistics for conditional simulation of ore bodies. *Economic Geology*, *69*, 673–687.

Kumral, M. (2012). Production planning of mines: Optimisation of block sequencing and destination. *International Journal of Mining, Reclamation and Environment*, *26*(2), 93–103. https://doi.org/10.1080/17480930.2011.644474.

Kumral, M. (2013). Optimizing ore-waste discrimination and block sequencing through simulated annealing. *Applied Soft Computing*, *13*(8), 3737–3744. https://doi.org/10.1016/j.asoc.2013.03.005.

Lloyd, C., & Atkinson, P. (1998). Scale and the spatial structure of landform: Optimising sampling strategies with geostatistics. *In Proceedings of the 3rd international conference on geoComputation*. University of Leeds.

Magri, E., & Ortiz, J. (2000). Estimation of economic losses due to poor blast hole sampling in open pits. In *Geostatistics 2000, Proceedings of the 6th international geostatistics congress*. Cape Town, South Africa, vol. 2, pp. 732–741.

Magri, E., Ortiz, J., & Libano, R. (2010). The economic optimisation of advanced drilling grids for short term planning and grade control at El Tesoro copper mine. In *MININ 2010, Proceedings of the 4th international conference on mining innovation*. Santiago, Chile, pp. 89–98.

Magri, E., Ortiz, J., Moya, R., & Salazar, A. (2011). Advanced reverse circulation drilling as a replacement to blasthole sampling: increasing short term planning profitability at cerro colorado copper mine. In *5th world conference on sampling and blending*. Santiago, Chile.

Mariethoz, G., & Caers, J. (2014). *Multiple-point geostatistics: Stochastic modeling with training images*. Hoboken: Wiley.

McBratney, A., Webster, R., & Burgess, T. (1981). The design of optimal sampling schemes for local estimation and mapping of regionalized variables. i. theory and method. *Computers and Geosciences*, *7*(4), 331–334.

McBratney, A., & Webster, R. (1981). The design of optimal sampling schemes for local estimation and mapping of regionalized variables. ii. program and examples. *Computers and Geosciences*, *7*(4), 335–365.

Ortiz, J.M. (2003). *Characterization of high order correlation for enhanced indicator simulation*. Doctoral dissertation, University of Alberta.

Ortiz, J. M., & Deutsch, C. V. (2004). Indicator simulation accounting for multiple-point statistics. *Mathematical Geology*, *36*(5), 545–565. https://doi.org/10.1023/B:MATG.0000037736.00489.b5.

Ortiz, J. M., & Magri, E. J. (2014). Designing an advanced RC drilling grid for short-term planning in open pits: Three case studies. *Journal of the Southern African Institute of Mining and Metallurgy*, *114*, 631–637.

Ortiz, J. M., Magri, E., & Libano, R. (2012). Improving financial returns from mining through geostatistical simulation and the optimized advance drilling grid at el tesoro copper mine. *Journal of the Southern African Institute of Mining and Metallurgy*, *112*(1), 15–22.

Pitard, F. (1993). *Pierre Gy's sampling theory and sampling practice. Heterogeneity, sampling correctness, and statistical process control* (2nd ed.). Boca Raton: CRC Press.

Ruiseco, J. R., Williams, J., & Kumral, M. (2016). Optimizing ore-waste dig-limits as part of operational mine planning through genetic algorithms. *Natural Resources Research*, *25*(4), 473–485. https://doi.org/10.1007/s11053-016-9296-1.

Santibañez, F., Silva, J. F., & Ortiz, J. M. (2019). Sampling strategies for uncertainty reduction in categorical random fields: Formulation, mathematical analysis and application to multiple-point simulations. *Mathematical Geosciences*, *51*(5), 579–624. https://doi.org/10.1007/s11004-018-09777-2.

Silva Maureira, D.A. (2015). *Enhanced geologic modeling with data-driven training images for improved resources and recoverable reserves*. Doctoral dissertation, Ph.D. thesis, University of Alberta.

Vasát, R., Heuvelink, G., & Borúvka, L. (2010). Sampling design optimization for multivariate soil mapping. *Geoderma*, *155*, 147–153.

Verly, G. (2005). Grade control classification of ore and waste: A critical review of estimation and simulation based procedures. *Mathematical Geology*, *37*, 417–476.