

New Tabu Algorithm for The Mining Drill Holes Positioning Problem with Blocks Uncertainty

Authors: Gamache, Guibault, Marcotte, Zagre

Montreal, Canada

Abstract

This paper proposes a new approach to the mining exploration drill holes positioning problem (DPP). This new approach incorporates both geostatistical and optimization techniques: A metaheuristic is developed to solve the DPP taking into account a new geostatistical uncertainty index that quantifies the reliability of the current interpretation of the mineral deposit. The uncertainty index is calculated from multiple deposit realizations obtained by truncated Gaussian simulations conditional to the available drill holes samplings.

Based on the uncertainty index, a linear programming model is defined to select the subset of future drill holes that maximizes coverage of the uncertainty. A Tabu Search algorithm is developed to solve large instances of this set partitioning problem. The proposed Tabu Search algorithm is shown to provide good quality solutions approaching 95% of the optimal solution in a reasonable computing time. This research shows that defining and taking into account the deposit uncertainty adds value and provides better solutions to the DPP. It allows better coverage of uncertainty for a fixed investment in drilling.

Keywords: Geostatistics, Mining Optimization, Tabu Algorithm, Set Covering Problem, Truncated Gaussian Simulation, Gibbs Sampling

1. Introduction

In mining exploration, the drill hole positioning problem (DPP) emerges from the need to accurately plan future definition drill holes whose objective

is to improve the current knowledge on an underground mineral deposit. The solution to this problem requires tools from geostatistics to model our current (lack of) knowledge of the deposit and from optimization to optimally select future drill holes such as to maximize knowledge of the deposit.

The geostatistical approach consists mainly in subdividing the exploration field into blocks that are classified according to their mineral content estimation (and accuracy of that block mineral estimation). Future drill holes are planned with more or less subjectivity, roughly placing new drill holes next to the blocks of interest. Pan [1] subdivides mineral exploration blocks into five categories related to the level of confidence of their mineral content estimation:

- Measured resources: blocks with expected high mineral content as measured per existing drill holes (high confidence).
- Indicated resources: blocks with probable high mineral content (moderate confidence).
- Inferred resources: blocks with possible high mineral content (low confidence).
- Barren: blocks with no mineral value expected.

Future drill holes are then planned targeting areas with moderate or low confidence blocks (indicated or inferred resources). While interesting from a geostatistical standpoint, this approach does not necessarily ensure a cost effective drilling campaigns. The same unique drill hole may simultaneously cover different categories of blocks so that it is hard to assess its real value. New drill holes campaigns will not necessarily bring additional interesting information if not planned methodically using a holistic approach. Furthermore, uncertainty of the blocks or uncertainty on the mineral envelope (surface) needs to be methodically quantified using a reliable measure and a more rational approach. Some authors proposed several approaches related to surface uncertainty in grids but none is actually applicable in the current context to mineral deposits where forms and surfaces are often not smooth and irregular with unexpected or unknown topology.

The optimization approach aims at selecting the best subset of drill holes

(from a predefined set of possible drill holes) that maximizes spatial coverage of the blocks (i.e. overall proximity between drill holes and blocks). However in these approaches the blocks are not differentiated in terms of potential content and the current available information is most of the time ignored or unused. For example, Soltani & al. [2] proposed a metaheuristic (genetic algorithm) to find the position and length of a predefined number of drill holes. In this approach, information from existing samplings are not taken into account and there are other limitations such as the verticality of the drill holes and their predefined quantity. Soltani & Hezarkhani [3] later completed this approach with drill hole inclination capability, but the number of drill holes is still predefined. Bilal and al. [4] reformulate the DPP as a set covering optimization problem and proposed a Tabu Search metaheuristic that searches the best subset of drill holes with the objective of maximizing block coverage. This approach is the most complete from an optimization perspective; however, the current information from existing drill holes is not fully taken into account. Only the proximity of the closest drill hole is used to measure coverage, ignoring joint collaborative information provided by sets of close drill holes possibly just slightly further from the uncovered block.

Saikia & Sarkar [5] in an application for a coal mine exploration, proposed an optimization algorithm to minimize the total kriging variance by increasing/reducing the number of drill holes and the distance between the drill holes. A limitation of this problem is that it considers only vertical and even spaced drill holes. Also, the kriging variance is only dependent on the spatial configuration distance between block and existing drill holes - and not on the actual results observed.

The solution approach that is proposed is based on Bilal & al. [4]. A new uncertainty index is calculated using geostatistical methods. This index permits to discriminate the blocks. More specifically, the uncertainty index assigned to each block is computed using a series of deposit realizations conditional to the available drill holes. Blocks appearing always in the simulated deposit or always out of the simulated deposits receive low uncertainty index. Deposit realisations are obtained using a truncated Gaussian simulation method. The optimization selects by linear programming a subset of drill holes maximizing uncertainty coverage. A Tabu Search algorithm is finally proposed to solve heuristically large instances of this optimization problem.

2. Methodology

2.1. Definition of the Uncertainty Index

The first step consists of calculating, for each block, an uncertainty index representative of the current level of confidence we have on the nature of a block. This is done using geostatistical simulation.

Lets consider an existing (yet unidentified) mineral deposit in an exploration field. The field is subdivided into a set U of n three-dimensional blocks ($n = |U|$). Let x be a variable identifying each block. Lets define $I(x)$ as the deposit indicator function at block x , also called *facies*. We assume, for simplicity, that each block can only be fully inside ($I(x) = 1$) or outside ($I(x) = 0$) the deposit.

A subset Ω ($\Omega \in U$) of m blocks are already sampled by existing drill holes ($m = |\Omega|$). The value of $I(x)$ is known for each of these sampled blocks ($x \in \Omega$) and unknown for all the other blocks. For simplicity we designate blocks belonging to the sampling group by the variable x^* ($x^* \equiv x \in \Omega$). Given the current information from the sampled blocks within Ω , we seek the likelihood of each other block x to be in the deposit. Let $p(x)$ denote that probability: $p(x) = P(I(x) = 1)$. For sampled blocks x^* ($x \in \Omega$), $p(x^*) = 1$ if block x^* is in the deposit, and $p(x^*) = 0$ otherwise. $p(x)$ is initially unknown for all the other non sampled blocks $x \notin \Omega$.

The variable $p(x)$ as defined here is an interpretation of the current expectation that a certain block x is in the deposit or not. Actually we are interested in a variable that would represent the correctness or certainty of that expectation. Let that variable be $u(x)$, the uncertainty index. The certainty of the expectation or prediction $p(x)$ is strong or almost certain for blocks having high $p(x)$ values (close to 100%) or low $p(x)$ values (close to 0%), because in those cases, we are very confident about our interpretation of the blocks facies (inside or outside the deposit). Inversely, the certainty of this expectation or prediction $p(x)$ is minimal for blocks with average $p(x)$ values (close to 50%).

If $p(x)$ represent the success probability of a Bernoulli trial, we can define $u(x)$ as the variance of the associated binomial process: $u(x) = p(x)[1-p(x)]$. $u(x)$ is close to the maximum value of 0.25 for $p(x)$ close to 50%, and $u(x)$ is close to the minimum value of 0.0 for values of $p(x)$ close to 0% or 100%.

The goal is to assess $u(x)$ for all the other blocks based on the knowledge we have on the current sampled blocks. We can intuitively expect that a block which is close to a sampled block will have a similar nature. However the prediction is difficult when a block is far from any existing drill holes or close to two dissimilar blocks (inside and outside the deposit).

The proposed method to calculate $u(x)$, $\forall x \in U$, consists of the following steps:

- Select a theoretical variogram model that is representative of the variability of the facies (variable $I(x^*)$) of the sampled blocks (existing drill holes).
- Encode observed facies of sampled blocks (variable $I(x^*)$) into a Gaussian variable $Z(x^*)$ which values also match the theoretical variogram. This is achieved using a Gibbs sampling algorithm.
- Obtain several independent deposit realisations ($\{I(x), \forall x \in U\}$) using conditioned truncated Gaussian simulations, each realisation is a plausible interpretation of the deposit that is compliant with observed facies at sampled blocks and theoretical variogram. This is achieved using the Turning bands simulation method.
- For each block $x \in U$, calculate $p(x)$ as the mean of the different facies simulated for block x in the different realisations, and $u(x)$ as the variance.

Selection of a Variogram Model

A variogram noted $\gamma(h)$ is a function that characterizes the spatial correlation of variable values observed between two points separated by a distance h . Variogram functions may present a plateau and a range (distance above which two points are no longer correlated and $\gamma(h \geq \text{range}) = \sigma^2$). Variogram types would often define the inner structure of a mineral deposit. There are several theoretical variogram models described in the literature (spherical, cubic, exponential, etc.). For simulation purposes we have to select and use a theoretical variogram model and $\gamma_{th}(h)$ that best fits the correlation presented by the sampling blocks. First, we calculate an experimental variogram $\gamma_{exp}(h)$ from the observed facies in sampling blocks to

which we will match the best fitted theoretical model:

$$\gamma_{exp}(h) = \frac{1}{2|N(h)|} \sum_{\forall(x_i, x_j) \in N(h)} [I(x_i) - I(x_j)]^2 \quad (1)$$

Where:

- h is a distance or range of distances separating two sampling blocks, and
- $N(h)$ is the subset of pairs of blocks matching that range of distances.

The geostatistical simulations are applicable on Gaussian variable (i.e. the underlying variable follows a Gaussian distribution). Next step of our process will delve into how the categorical values $I(x)$ are encoded into a latent Gaussian variable. For the simulation purposes, the theoretical variogram model (γ_{th}) we are interested in is one for a Gaussian variable while our experimental variogram (γ_{exp}) is calculated using categorical values. In order to verify the matching between the selected theoretical model and the experimental variogram, the theoretical model should be converted into an induced categorical (γ_{Ind}) model which is compared to the experimental variogram. The link between the theoretical variogram and its induced categorical variogram for Gaussian variables is given by the following equation:

$$\gamma_{Ind}(h) = \sigma_{Ind}^2 - \frac{1}{2\pi} \int_0^{\rho_{th}(h)} \exp\left(-\frac{c^2}{1+u}\right) \frac{1}{\sqrt{1-u^2}} du \quad (2)$$

where:

- σ_{Ind}^2 is the variance of the categorical variable $I(x)$,
- ρ_{th} is the correlation function of the Gaussian variable (obtained from the theoretical variogram model) and,
- c is a threshold corresponding to the Gaussian-inverse of the proportion of mineral facies observed in the sampling blocks.

At the end of these steps, we have selected a theoretical variogram model which we suppose is representative of the spatial correlation between specific

Gaussian values associated to the blocks.

Encoding Observed Facies into Gaussian Values

The second process is to convert categorical values of observed facies at sampled blocks into Gaussian values. Let $Z(x)$ be the Gaussian variable associated with the blocks facies. Before we perform simulations of Gaussian values for every block of the field, we want to encode observed categorical facies values at sampled blocks $I(x^*) \in \{0, 1\}$ into Gaussian values $Z(x^*) \sim N(\mu, \sigma^2)$. The correlation between these Gaussian values should also satisfy the selected theoretical variogram γ_{th} . In order to achieve this, we use the Gibbs Sampling method which is a Markov Chain Monte Carlo algorithm. Lets suppose that $Z(x^*)$ follows a standard normal (Gaussian) distribution of mean $\mu = 0$ and variance $\sigma^2 = 1$ ($Z(x^*) \sim N(0, 1)$).

First we define a threshold c such that:

$$\forall x \in \Omega (\equiv \forall x^*) \begin{cases} I(x^*) = 1 \Leftrightarrow Z(x^*) \in]-\infty, c] \\ I(x^*) = 0 \Leftrightarrow Z(x^*) \in]c, +\infty[\end{cases} \quad (3)$$

The standard cumulative normal distribution at c corresponds to the proportion of barren blocks ($I(x^*) = 0$) observed in the field (assuming the existing drill holes are well distributed in the field, otherwise a decluttering process would be required). The value of c is given by the inverse cumulative normal function at $\frac{\sum_{x \in \Omega} I(x^*)}{m}$.

We then proceed iteratively to obtain a set of final Gaussian values that satisfy the variogram model (γ_{th}). Let $Y^t = [Z_t(x_1^*), Z_t(x_2^*) \dots Z_t(x_m^*)]$ be a Markov chain representing the Gaussian values assigned to sampled blocks at iteration t of the Gibbs sampling. We can start out with Y^0 by drawing and assigning a random Gaussian value to each block, that respects the threshold corresponding to the blocks facies ($]-\infty, c]$ if $I(x) = 1$ or $]c, +\infty]$ if $I(x) = 0$).

The Gibbs sampling consists, for each block, to draw a value from its conditional distribution. As soon as a new value is obtained, it replaces the old one and the new value is used to estimate the conditional distribution of the next block. An iteration is completed when all blocks have been visited once. Because the field is Gaussian, the conditional distributions are also Gaussian and are easily obtained by simple kriging (Chiles & Delfiner [6]). However, the presence of the threshold c to respect for each block means that we actually have truncated Gaussian distributions. An easy way to sample from a truncated Gaussian distribution is to proceed by acceptance/rejection

(Freulon & de Fouquet [7]). We compute the Gaussian conditional distribution and draw a value from it. If the Gaussian value falls on the right side of the threshold, it is accepted, otherwise we keep the old value. In either case we proceed to the next block.

```

1 :  $t = 0$ ,
2 :  $Y^0 = [Z_t(x_1^*), Z_t(x_2^*) \dots Z_t(x_m^*)]$  (Initial random drawing)
3 :  $Y^{(t+1)} = Y^t$ ;  $t = t + 1$ 
4 :  $k = 1 \dots m$ 
     $Value = RandomDrawing() \sim N(Z_{kr}^*, \sigma_{kr}^{*2})$ ,
    4.1 : If  $[(Value \leq c) \text{ and } I(x_k^*) = 1] \text{ or } [(Value > c) \text{ and } I(x_k^*) = 0]$ 
        Yes :  $Z_t(x_k^*) = Value$ 
        No :  $Z_t(x_k^*) = Z_t(x_k)$ 
5 :  $t > maxiterations$ 
    Yes : STOP.
    No : Back to 3

```

(4)

Where Z_{kr}^* and σ_{kr}^* are the kriging value at x_k^* and the associated kriging variance).

Many iterations are required before convergence to the desired joint distribution (histogram and variogram) is reached. The Gibbs sampling is stopped when a preselected maximum number of iterations is reached. The Gibbs sampling process goes through a burn-in phase when the values are gradually matching the assigned variogram model, as shown by Cuba et al. [8]. Once the burn-in phase is finished, the process enters a stochastic phase or random phase during which there is no substantial improvement in the resulting joint distribution (in terms of matching the desired distribution). The process should always have enough iterations to ensure that the burn-in phase is finished.

Several authors such as Chen & al. [9], Lyster & Deutsch [10] and Reza Najafi & Moradkhani [11] have attempted to find generalized stopping criteria to the Gibbs sampling process. Most were unsuccessful or at best only useful in their specific context. Onibon & al. [12] concluded that the issue with many stopping criteria or techniques is that they are not reliable and

are costly in computational time. Similarly to Emery [13], when using the Gibbs sampling method for conditioning of a Gaussian field, we strongly advise monitoring the convergence by following the evolution of the resulting histograms and variograms after each iteration, and the selected number of iterations should be large enough to ensure that the burn-in phase has passed.

The whole Gibbs sampling process is repeated r times starting from different sets of initial values ($\{Y_1^0, \dots, Y_r^0\}$) to obtain as many independent chains ($Y_1^{final}, \dots, Y_r^{final}$) of $m = |\Omega|$ Gaussian values that are consistent with the assigned theoretical variogram model. Each one of these r independent final chains will serve as a conditioning input for an independent simulation of facies in the field.

Simulation of Facies at Each Block

Having a set of Gaussian values for the sampled blocks, the next step consists of doing a geostatistical simulation of facies for every block in the field. There are several simulation methods applicable to Gaussian fields such as the LU Decomposition, the Sequential Gaussian Simulation (SGS) method, the Fast Fourier Transform (FFT) method and the Turning Band Method. The Turning band method is recommended in this application because of its low algorithmic complexity ($O(n)$) and its applicability on very large grids. Detailed description of that method can be found in Chiles & Delfiner [6] and in Emery [14]. The turning band method creates a Gaussian random field with the desired covariance (related to the variogram) by simulating projections of a two-dimensional (R^2) or three-dimensional (R^3) variable along a finite number of concentric lines.

Lets consider a Euclidean space (R^2 or R^3) with a Cartesian coordinate system in which our blocks are placed. Lets call $G(x)$ the Gaussian value simulated at block $x \in U$ using the turning band method with a desired variogram (or covariance). We suppose k lines L_i passing through the origin of the coordinate system and sweeping the R^3 space. $H_i(x)$ is the simulated value on line L_i ($H_i(x)$ is a projection of $G(x)$ on line L_i). We have:

$$G(x) = \frac{1}{\sqrt{k}} \sum_{i=1}^k H_i(x) \quad (5)$$

The relation between the desired covariance C_G for $G(x)$ (in 3D) and the

covariance C_Y of the one-dimensional fields $H_i(x)$:

$$C_H(h) = \frac{d}{dh}[hC_G(h)] \quad (6)$$

The new simulated Gaussian field ($G(x)|x \in U$) needs to go through a conditioning process in order to be consistent with the value observed at sampling blocks ($Z(x^*)$). This is done using the kriging technique:

$$Z^{sim}(x) = G(x) + (Z_x^{kr} - G_x^{kr}), \forall x \in U \quad (7)$$

Where $Z^{sim}(x)$ is the Gaussian conditioned value at block x , Z_x^{kr} is the Gaussian value at block x calculated by kriging using values at sampled blocks ($\{Z(x^*)\}$) and G_x^{kr} is the Gaussian value at block x calculated by kriging using non conditioned simulated values at sampled blocks ($\{G(x^*)\}$). Lets notice that for sampled blocks $x^* : Z^{sim}(x^*) = Z(x^*)$.

The whole process (turning bands simulation + conditioning) is repeated r times, using each one of the r independent Gibbs chains as conditioning values. The conditioned Gaussian values obtained through simulation ($Z_i^{sim}(x)$) can be converted back into simulated conditioned facies ($I_i^{sim}(x) \in \{0, 1\}$) using the same threshold c and rationale as in equation 3.

In summary, in this step, each one of the r independent chains of conditioning Gaussian values leads to a different set of simulated blocks facies that forms a mineral deposit or realization. All the different realizations simulated are all possible interpretations of the real mineral deposit based on the samplings drill holes.

Calculation of the Blocks Uncertainty Index

For each block x we now have r independent shots at a facies $F(x)$ with only two possible outcomes: 1 (inside deposit) or 0 (outside deposit), as in a binomial process. We can now evaluate the probability $p(x)$ of block x being inside the deposit as the average value of the predicted value for block x in the realizations:

$$p(x) = P(I(x) = 1) = \sum \frac{I_i^{sim}(x)}{r}, \forall x \in U \quad (8)$$

For sampled blocks x^* , the facies probability is either 100% ($I(x^*) = 1$) or 0% ($F(x^*) = 0$) because $I_i^{sim}(x^*) = I(x^*)$. The uncertainty index $u(x)$ can

be calculated as the variance of the binomial process:

$$u(x) = \text{Var}(I(x)) = p(x)[1 - p(x)] \quad (9)$$

This index spans from 0.0 ($p(x) = 100\%$ or 0%) to 0.25 ($p(x) = 50\%$). It measures the chances that the current interpretation of a blocks state is correct. Each block of the field has an assigned uncertainty index. That new information can be used to select future drill holes by favoring drill holes covering the most uncertain zones of the field.

2.2. Optimization for Future Drill Holes Selection

In this second part, an optimization model is proposed for selection of future drill holes. The optimization problem is formulated as an integer linear problem whose objective is to maximize coverage of the blocks uncertainty index. The rationale behind that approach is that from an exploration perspective, drilling in areas with high uncertainty would bring valuable information on the mineral deposit and reduce total uncertainty. A Tabu Search algorithm is proposed for large instances of this optimization problem for which an exact solution is hard to find. This approach is a direct follow-up to the one proposed by Bilal & al. [4] which formulated the DPP as a partial set covering problem (with the objective of maximizing blocks coverage); The improvement being that in the current approach, the blocks are now discriminated using their uncertainty index (there is an incentive to cover uncertain blocks).

Definitions and Decision Variables

- Let UI_i designates the uncertainty index for block i ($u(i)$) as calculated in the first part of this work (calculation of the uncertainty index).
- The set of n blocks is still designated by U .
- Let Ψ designates the set of all possible individual drill holes. We consider a finite number of $q = |\Psi|$ possible drill holes. Drill holes are defined by a collar point (position of the drilling machine) and a terminal point or terminal block. Lets notice that according to this definition, drill holes having the same collar point and orientation but only different length (or depth) are identified different single drill holes which are part of Ψ .

- A solution S is a subset of Ψ and $|S|$ designates the number of drill holes within the solution S .
- A block i is considered covered by a drill hole j if it is located within a certain distance (range) of the drill hole. The coverage index Cov_j^i is defined as:

$$cov_j^i = \begin{cases} 1 & \text{if block } i \text{ is covered by drill hole } j \text{ (distance}(i, j) \leq \text{range}) \\ 0 & \text{otherwise} \end{cases}$$

The block-drill hole coverage matrix $C = [cov_j^i | i \in U \text{ and } j \in \Psi]$ is an input to this optimization problem.

- A block i is considered covered in a solution S when it is covered by at least one drill hole $j \in S$. We denote DH_i the subset of drill holes that covers block i :

$$DH_i = \{j \in \Psi | Cov_j^i = 1\}$$

- X_i is the decision variable associated with the coverage of block i , defined as:

$$X_i = \begin{cases} 1 & \text{if } i \text{ is covered by at least one drill hole of a solution } S: \\ & \Rightarrow \sum_{j \in S} Cov_j^i \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

- Y_j is the decision variable associated with the selection of a drill hole j , defined as:

$$Y_j = \begin{cases} 1 & \text{if } j \text{ belongs to the current solution } S \Rightarrow j \in S \\ 0 & \text{otherwise} \end{cases}$$

- We denote BL_j the subset of blocks covered by drill hole j :

$$BL_j = \{i \in U | Cov_j^i = 1\}$$

- \hat{c}_j is the cost of drill hole j . As with the set of potential drill holes, the drill holes cost array $\hat{c} = [\hat{c}_j | j \in \Psi]$ is an input to the optimization problem.

- \hat{C}_{max} is the maximal total drilling cost allowed (when applicable). A solution is considered acceptable if it satisfies the maximal cost constraint: $\sum_{j \in S} \hat{c}_j \leq \hat{C}_{max}$.

Linear Programming Model

The optimization problem can be formulated as an integer linear program

that selects the subset of drill holes which maximizes coverage of the uncertainty index, for a given cost constraint (maximal drilling cost \hat{C}_{max}):

$$\begin{aligned} & \text{Maximize} \sum_{i \in U} UI_i \times X_i \\ & \text{Subject to: } \begin{cases} \sum_{j \in \Psi} \hat{c}_j \times Y_j \leq \hat{C}_{max} \\ \sum_{i \in BL_j} X_i \geq |BL_j| \times Y_j, \forall j \in \Psi \\ \sum_{j \in DH_i} Y_j \geq X_i, \forall i \in U \\ X_i, Y_j \in \{0, 1\} \end{cases} \end{aligned} \quad (10)$$

The first constraint is the budget constraint. The second set of constraints ensures that if a drill hole j is selected, then all the blocks in BL_j are covered. The third set of constraints ensures that if a block i is covered, then at least one drill hole in DH_i is selected.

We can also define an alternative problem where the goal is to minimize the total drilling costs for a defined level α of uncertainty coverage (for example, we would like to cover $\alpha=90\%$ of the current uncertainty). This problem can be formulated as:

$$\begin{aligned} & \text{Minimize} \sum_{j \in \Psi} \hat{c}_j \times Y_j \\ & \text{Subject to: } \begin{cases} \sum_{i \in U} UI_i \times X_i \geq \alpha \times \sum_{i \in U} UI_i \\ \sum_{i \in BL_j} X_i \geq |BL_j| \times Y_j, \forall j \in \Psi \\ \sum_{j \in DH_i} Y_j \geq X_i, \forall i \in U \\ X_i, Y_j \in \{0, 1\} \end{cases} \end{aligned} \quad (11)$$

Tabu Search Algorithm

Large instances of such integer problem are difficult to solve. In a similar research, Bilal [15] has shown that the *CPLEX* solver was unable to find an optimal solution after several computing days for a similar problem with 20000 blocks. Metaheuristics are then used to find good quality solutions for such problems. The method proposed in this research is a Tabu Search algorithm which explores solutions of a neighborhood while keeping a record (Tabu list) of explored solutions characteristics (drill holes). As demonstrated in further examples, this algorithm provides good quality solutions comparable to the optimal solution for a reasonable computing time.

In this algorithm, a solution S represents a subset of drill holes. The neighborhood of a solution S , noted $N(S)$, is a set of solutions having similar characteristics with S . For each solution S , we define three types of neighborhood sets as shown in Figure 1:

- $N(S)^+$: A solution S' is considered in the neighborhood of the current solution S , if S' has the same drill holes as those in S plus a single additional drill hole. Thus, a solution S has at most $|\Psi| - |S|$ neighbours+ in set $N(S)^+$:

$$N(S)^+ := S' \in \Psi : |S' \cap S| = |S'| - 1$$

- $N(S)^-$: A solution S' is considered in the neighborhood of the current solution S , if S' has the same drill holes as those in S excluding a single drill hole in S . Thus, a solution S has at most $|S|$ neighbours- in set $N(S)^-$:

$$N(S)^- = S' \in \Psi : S' \in \Psi \text{ and } |S' \cap S| = |S| - 1$$

Lets notice that: $S' \in N(S)^- \Leftrightarrow S \in N(S')^+$

- $N(S)j$: A solution S' is considered in the neighborhood $N(S)^*$ of a current solution S , if S and S' has the same number of drill holes and $|S|-1$ common drill holes (S and S' differ by only one drill hole). Thus, a solution S has at most $|S| \times (|\Psi| - |S|)$ neighbours* in set $N(S)^*$:

$$N(S)^* := S' \in \Psi : |S'| = |S| \text{ and } |S \cap S'| = |S| - 1$$

A drill hole is considered nearby another one if they have the same collar point and nearby terminal blocks. In order to reduce the number of neighboring solutions to explore starting from a current acceptable solution S , the considered neighbours^+ and neighbours^* are limited to acceptable neighboring solutions that differ from S by a nearby drill hole.

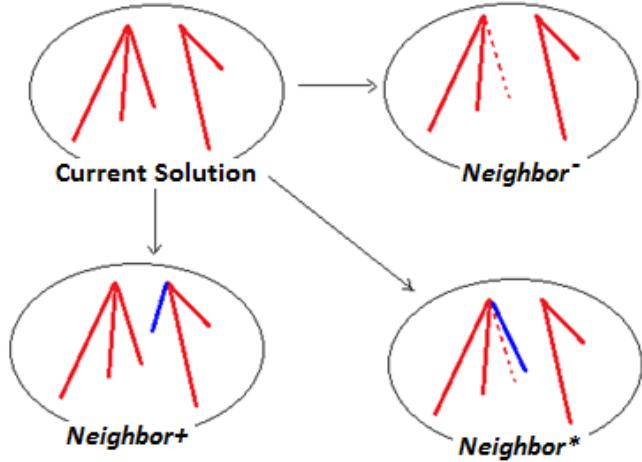


Figure 1: Types of Solution’s Neighborhood

For each acceptable solution explored, a score is calculated which is the evaluation of the objective function for this solution (total uncertainty that is covered by the solutions drill holes). The Tabu Search process (presented in 2) explores the domain of solutions by going from solution to neighbor solution, choosing at each step the local optimum as the current solution (neighbor with the best score to cost ratio, when comparing all types of neighborhood).

At first only neighbors^+ and neighbors^* are considered in the neighborhood search. When the algorithm can no longer find better solution than the current solution in the $\text{neighborhood}^{+/*}$, it considers removing drill holes from the current solution. A score to cost ratio is then used as selection criteria in order to ensure that solution with fewer drill holes (neighbor^-) are considered. The Tabu algorithm can thus remove ineffective drill holes from a current solution and later add more interesting drill holes.

A drill hole added or extracted when passing from the current solution to the overall best neighboring solution is registered in a Tabu list for a certain number of iterations t (actually the length of the Tabu list). A classical and empirical value that is usually recommended for the length of Tabu list is approximately $t \approx \sqrt{n}$. At any time, the overall best solution evaluated is recorded.

An acceptable initial solution can be generated by a simple greedy algorithm, which consists of randomly selecting and adding drill holes (maybe from a predefined list of best single drill holes) until the maximum drilling cost is almost reached and it is no longer possible to add a drill holes to that initial solution. When the Tabu Search algorithm is stuck (i.e. no new best solutions after a certain number of iterations) a temporary violation of the maximal cost constraint can be allowed for a certain number of iterations in order to change some of the current solution features. During this violation time, overall best solutions are not recorded if they are not acceptable solutions.

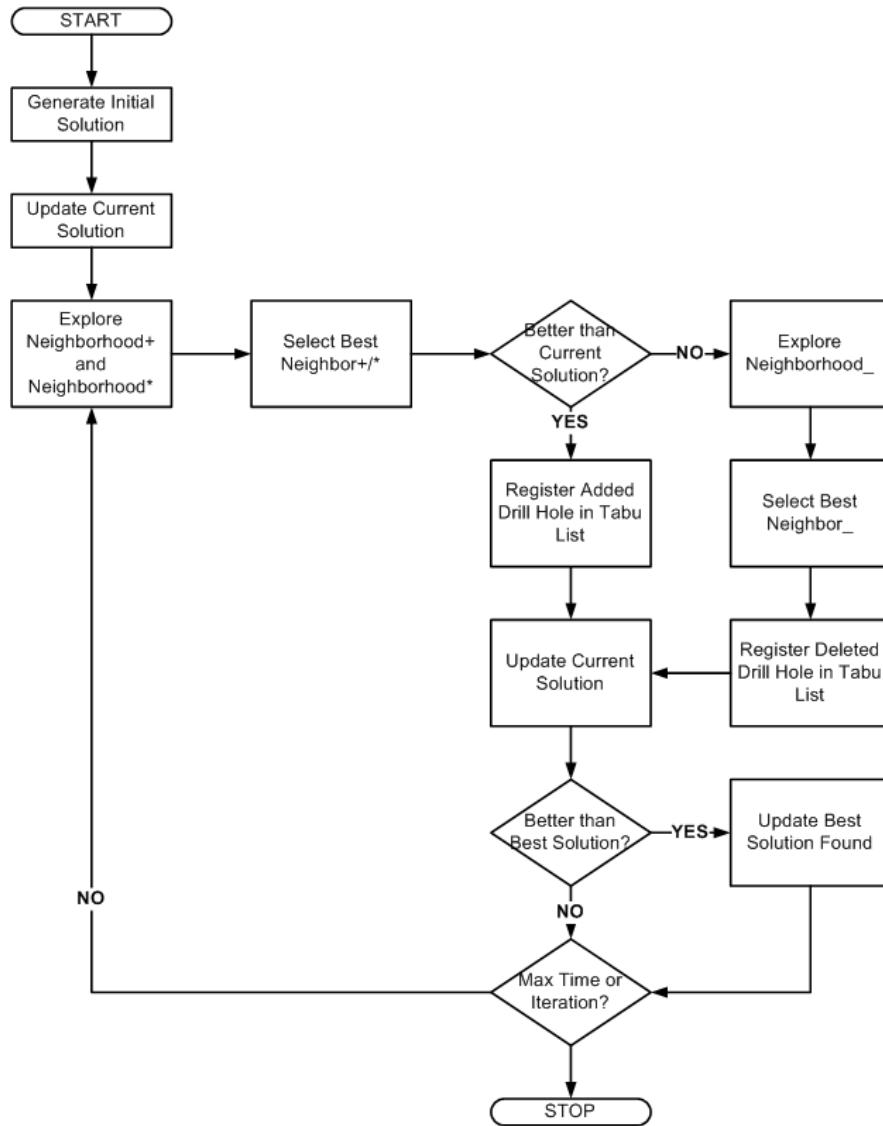


Figure 2: Diagram of the Tabu Process

The Tabu algorithm presented provides good solutions to the optimization problem. We can compare its predictions with optimal solutions for some examples of average-sized problems using test cases.

3. Application on Test Cases

Lets apply the whole process (uncertainty index calculation and drill hole optimization) to test cases.

3.1. Test Case 1: A Two-dimensional Example

Lets consider the mineral deposit presented in Figure 3-a (generated by random unconditioned geostatistical simulation). The field is subdivided into a 300x200 grid (60 000 blocks). Existing drill holes are presented in red lines and results of sampled blocks are shown in Figure 3-b: yellow for inside-deposit facies blocks ($I(x^*)=1$) and blue for outside-deposit facies blocks ($I(x^*)=0$). A total of 1117 blocks are sampled by the existing drill holes.

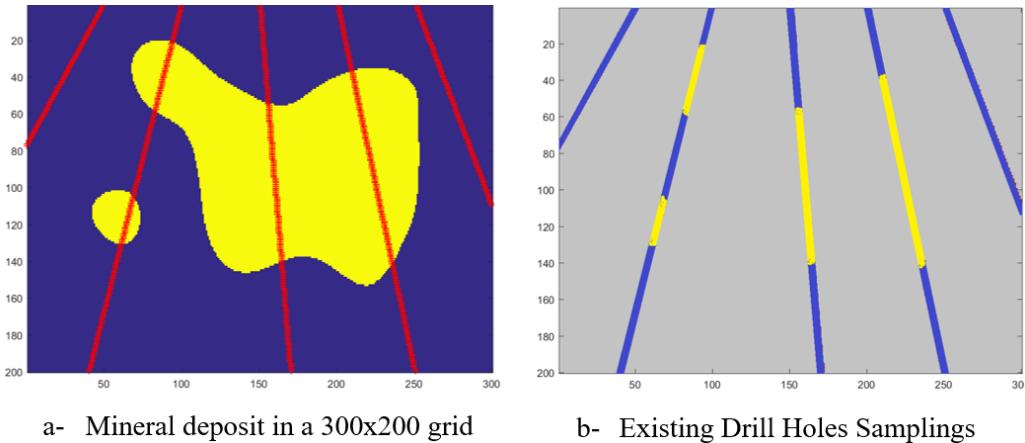
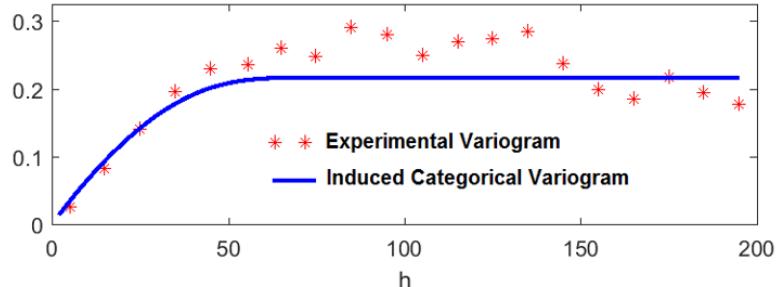


Figure 3: Example of a 2-D Mineral Deposit in a 300x200 Field with Existing Drill Holes

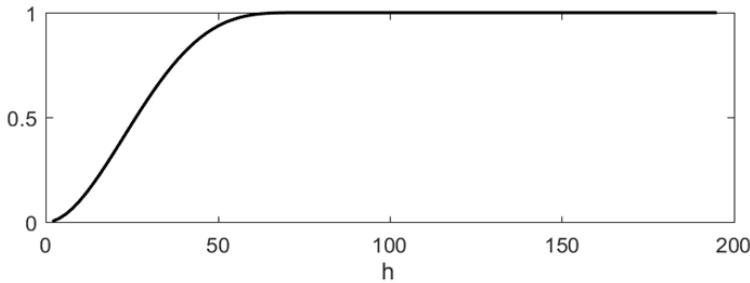
Calculation of an Uncertainty Index

The omnidirectional experimental variogram $\gamma_{th}(h)$ is calculated from those samplings and presented in Figure 4-a (in red dots). The theoretical model selected for the Gaussian variable is an isotropic (omnidirectional) Cubic model with a range of 75.0 distance units presented in Figure 4-b. An isotropic model was selected for simplification purposes (meaning that the variability is considered the same in every direction). The induced categorical variable model $\gamma_{ind}(h)$ and its adjustment to the experimental variogram is shown in

Figure 4-a (blue line).



a- Experimental Variogram and Induced Categorical Variogram



b- Theoretical Variogram Model (Cubic) for the Gaussian Variable

Figure 4: Variogram Model Selection

We can now encode the categorical facies values for sampled blocs $I(x^*)$ into Gaussian values $Z(x^*) \sim N(\mu, \sigma^2)$, using the Gibbs sampling and the theoretical model selected as target variogram. The Gibbs sampling process was applied 100 times to obtain 100 independent sets of Gaussian values. Figure 5 shows an example of the resulting variogram of the final Gaussian values for one of the $r = 100$ independent repetitions. The resulting variogram after 50 iterations (sin blue dash line) shows that the fitting to the target variogram has improved. This test case with 1117 initial values required 30000 iterations to make sure the burn-in phase has passed.

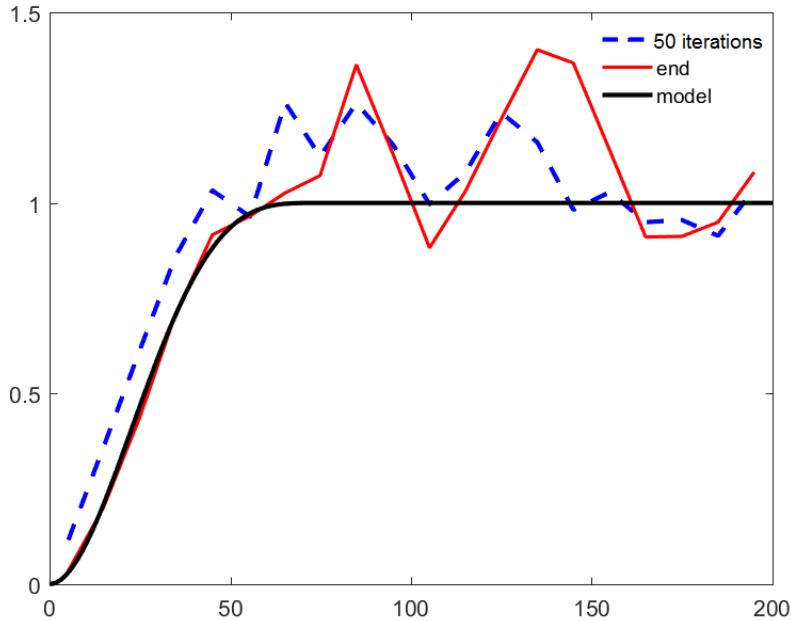


Figure 5: Variograms Produced during the Gibbs Sampling Process

Once the 100 independent sets of Gaussian values are found, they can be used as conditioning values to generate 100 simulations of Gaussian values for the complete field using the Turning Bands Algorithm. These Gaussian values are reconverted into categorical facies values and thus simulated mineral deposits (independent realizations). Figure 6 presents three different independent realizations simulated for the current test case with the Turning Bands method. We can notice that although dissimilar, each one of those realization is consistent with the initial sampling results shown in Figure 3.

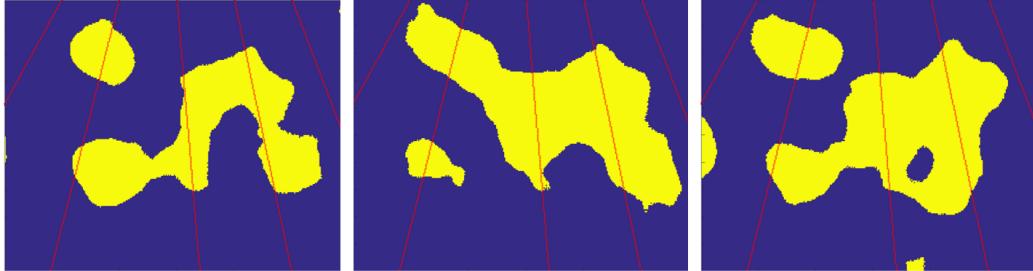


Figure 6: Example of independent realizations generated with the Turning Band Simulation

With 100 independent realizations we can now easily calculate the facies probability and uncertainty index (variance) for each block, illustrated in Figure 7. Figure 7-a shows the facies probability for each block with the values ranging from 0.0 (blue) to 1.0 (yellow). We can distinguish areas where no mineral deposit (blue) is expected from areas likely to contain some mineral deposit (yellow), given the current information provided through existing drill holes. Figure 7-a shows the uncertainty index with values ranging from 0 (blue) to 0.25 (yellow). Blocks nearby existing drill holes are certain (samplings) whereas, generally, blocks far from existing drill holes and blocks that are close to the mineral envelope (facies switch), tend to be the most uncertain.

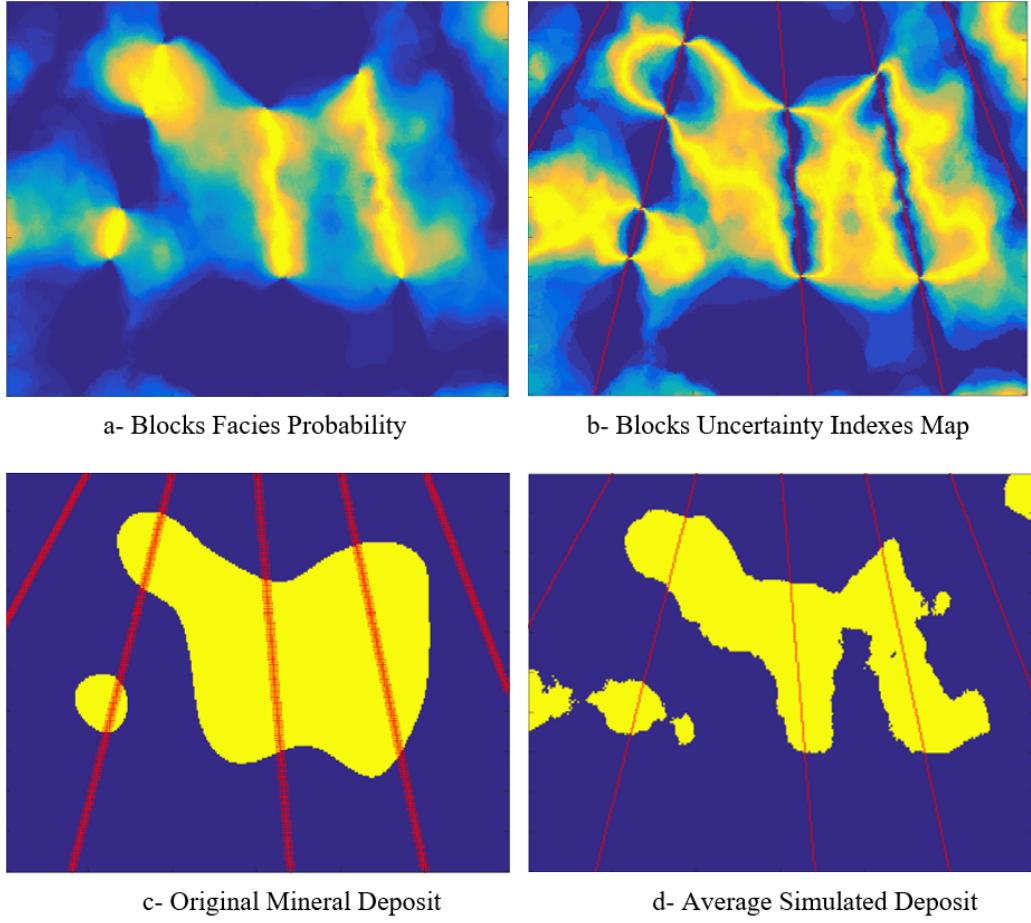


Figure 7: Facies Probability and Uncertainty Index

Figure 7-d shows the average simulated deposit (average of 100 simulations rounded to 0 or 1) and compares it to the original mineral deposit in Figure 7-c. The simulated deposit is comparable to the original deposit. The total facies error (blocks with the wrong interpreted facies) is 8143 (among 60 000), which means that the current facies interpretation (and thus the mineral representation) is 86.4% accurate, which is acceptable given the small number of samplings initially available (1117 blocks).

In order to have reliable final results (uncertainty index), cautiousness is recommended when selecting several parameters of this geostatistical method: choice of variogram theoretical model and range, number of iterations of the

Gibbs sampling process, independence (inter-correlation) of the Gibbs final chains and number of independent realizations. All those parameters have to be selected with particular care, on a case-by-case basis, taking into account the nature and size of the problem. As an example, Figure 8 compares the final realizations generated using a different theoretical variogram models as target: A Spherical model instead of the Cubic model recommended for this test case, with the same range of 75.0. Both models provides similar matching to the experimental variogram, however, the realizations simulated with the spherical model are quite different in structure from the original mineral deposit, despite the fact that it is coherent with the initial sampling.

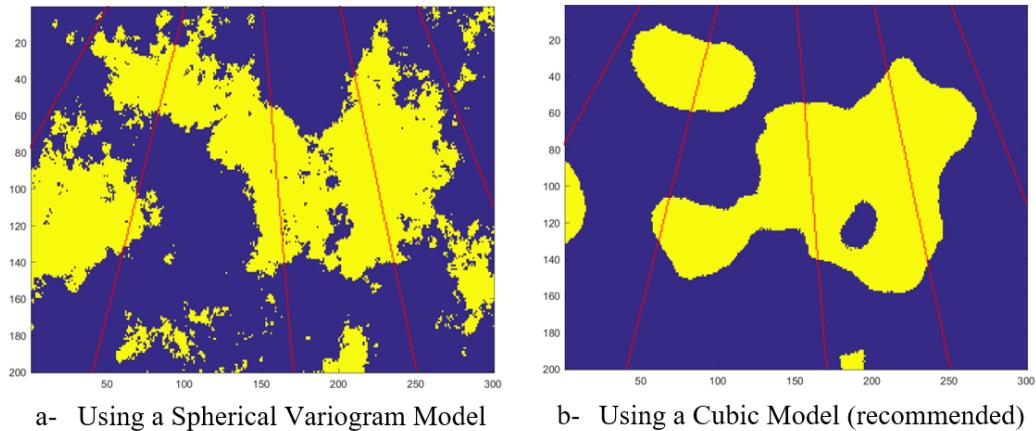


Figure 8: Comparison of Realizations Generated with Different Theoretical variogram Model

Optimization for Future Drill Holes Selection

Once the uncertainty index value for each block of the field is calculated, the optimization problem for selection of future drill holes can be formulated and solved. Figure 9 presents the universe of all possible drill holes considered with three collar points, forming a total of 1800 possible drill holes. For simplification purpose, it is assumed that a drill hole individual cost is proportional to its length. It is assumed that each drill hole has a coverage radius of 10.0 distance units. We want to find the best subset of future drill holes for a total drilling costs of 1000.0 (in distance units), based on the criteria of covering the maximum uncertainty.

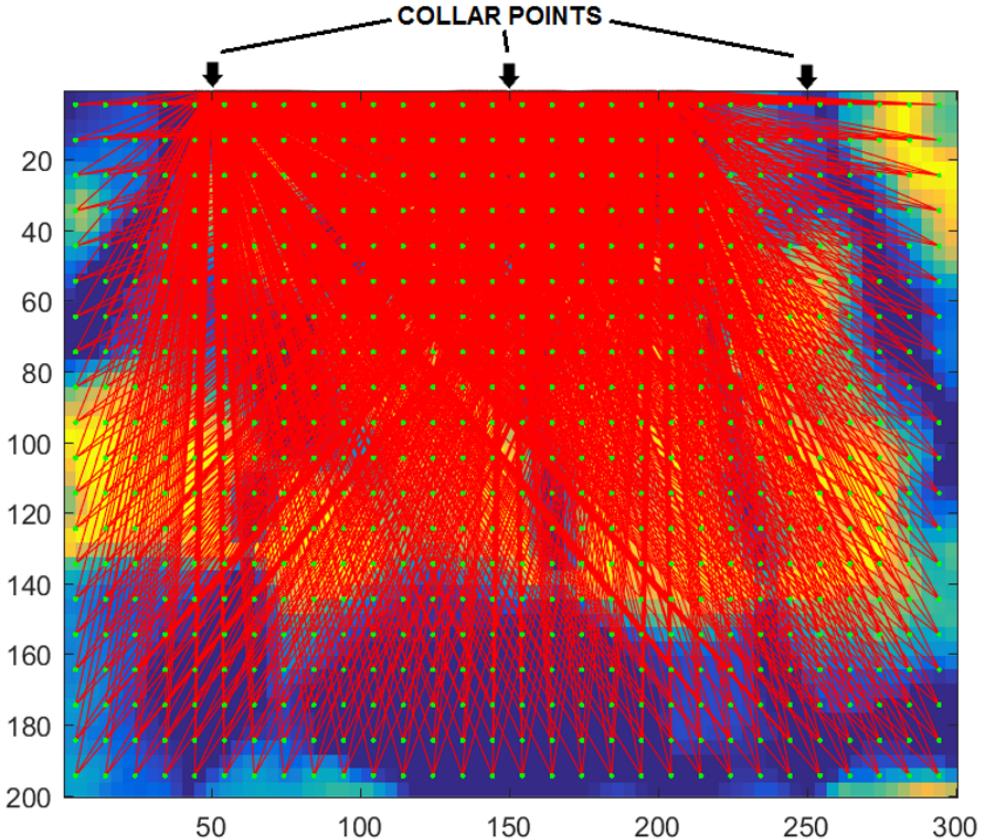
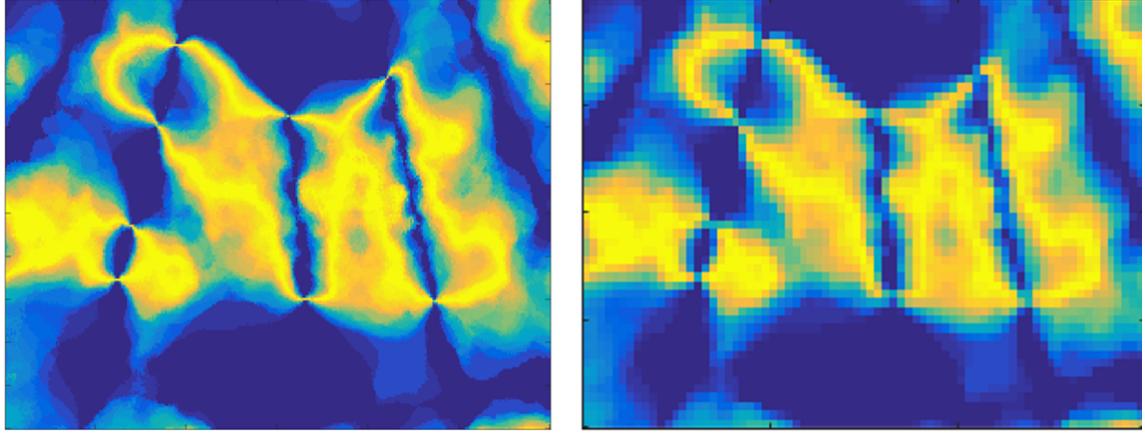


Figure 9: Uncertainty Map and Potential Future Drill Holes

The initial number of blocks is 60000. In order to reduce the number of blocks and thus the number of variables of the problem (to allow optimal resolution), we can astutely aggregate the blocks in the original uncertainty index map. In this example, 300x200 original blocks have been aggregated into 75x50 larger blocks with uncertainty indexes corresponding to the sum of the 4x4 initial blocks. The results are shown in Figure 10, both images are similar (with lower resolution in the second one) but the number of blocks has been reduced to 3750. We have to be cautious with such a reduction: we can reasonably assume that a good solution for the reduced problem will be a good solution for the original set, only if the final aggregated image is similar to the original one.

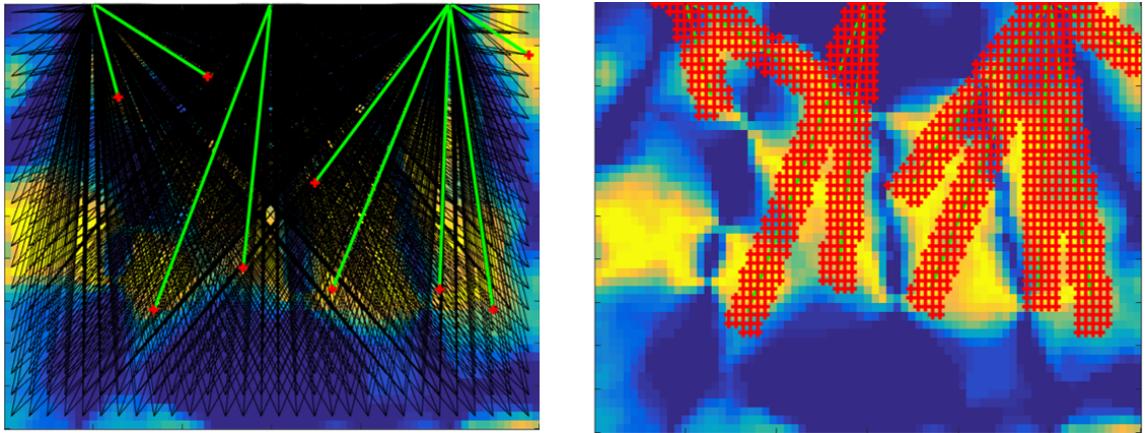


a- Initial Uncertainty Map

b- Aggregated Uncertainty Map

Figure 10: Blocks Aggregation for Variables Reduction

We can now solve the optimization problem presented in 10 with an optimal solver such as CPLEX or MATLAB *Intlinprog* function. The problem is solved within 205.0 seconds of computing time on a *1.70GHz Intel Core i5* processor. The optimal solution found is presented in Figure 11. This optimal solution has a total drilling cost of 998.9 and a coverage score of 2898.5 (total uncertainty covered by selected drill holes).



a- Subset of Selected Drill Holes

b- Covered Blocks and Uncertainty

Figure 11: Optimal Solution found for Test Case 1 with a Maximum Cost of 1000.0

The problem is solved using the proposed Tabu Search algorithm for 1000 iterations. The Tabu Search solution is presented and compared to the optimal solution in Figure 12.

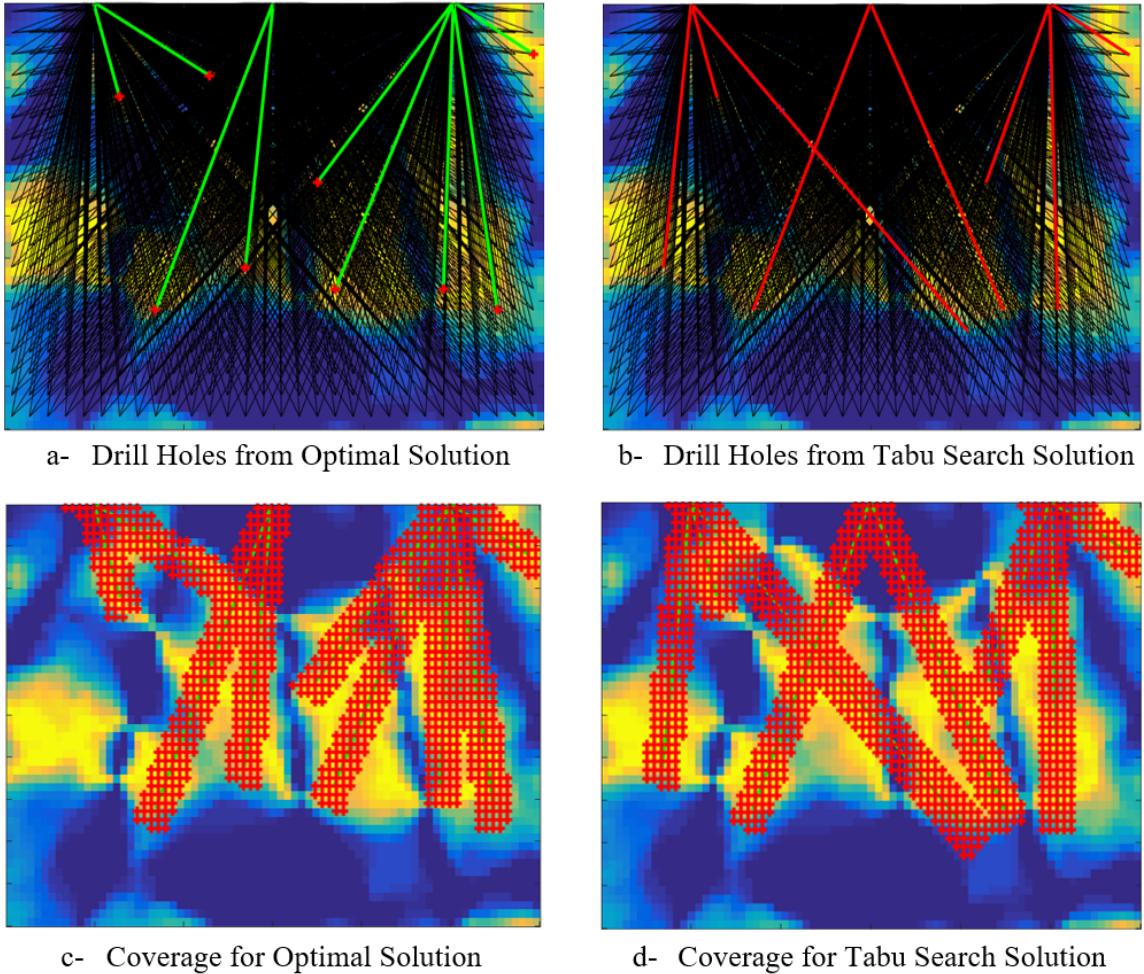


Figure 12: Tabu Search Solution and Optimal Solution for Test Case 1

The Tabu Search solution has a cost of 929.8 and a coverage score of 2852.4 (98.4% of the optimal score). The drill holes and covered uncertainty are very similar to the optimal solution. The best Tabu solution is found after 534 iterations (929.8 seconds on the same *1.70GHz Intel Core i5* processor). Results evolution of the Tabu Search algorithm is presented in Figure 13.

Although slightly less fast than the exact solver (205.0 seconds), the Tabu Search algorithm finds a good solution at 90% of optimality within 37 iterations (94.3 seconds) and most important is not limited in terms of problem size.

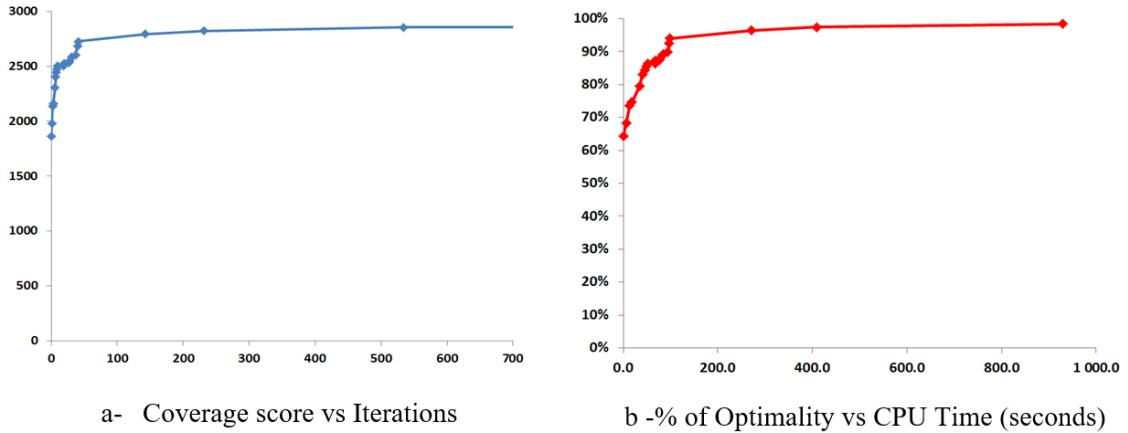


Figure 13: Evolution of the Tabu Search Results

3.2. Test Case 2: A Three-dimensional Example

The proposed method also works for three-dimensional mineral deposits. Lets consider the mineral deposit in a 60x48x24 grid as presented in two views in Figure 14. This mineral deposit was also generated using unconditioned geostatistical simulation. The existing drill holes are shown in red (vertical drill holes are considered in this example for simplification only).

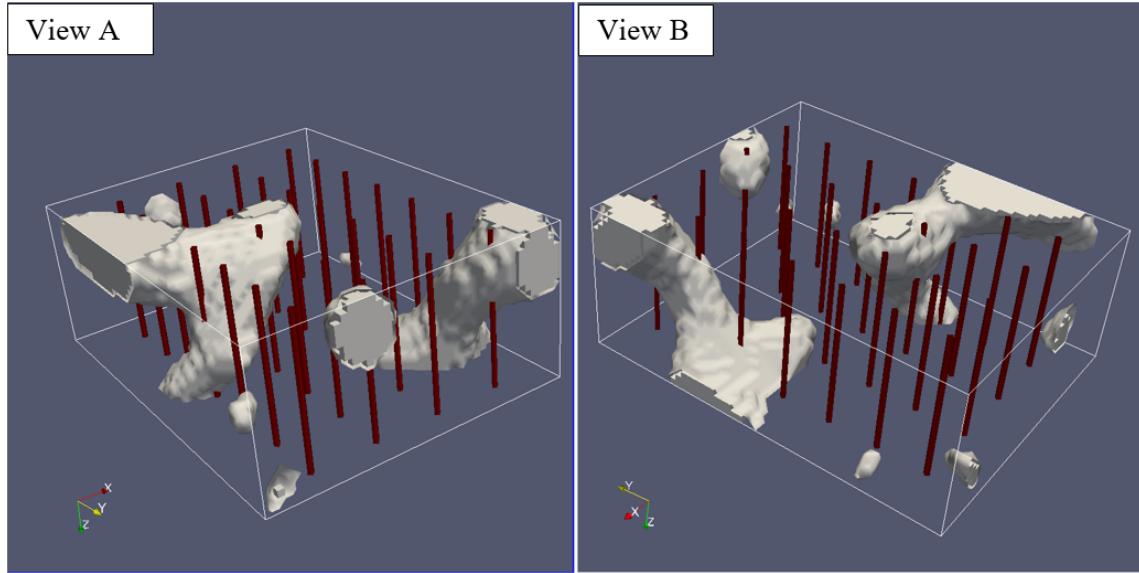


Figure 14: Mineral Deposit and Existing Drill Holes (Test Case 2)

The sampling results from existing drill holes are presented in Figure 15. A total of 672 blocks are sampled by these drill holes (on a total of 69120 blocks). Sampled blocks inside deposit are presented in yellow and blocks outside deposit are presented in dark gray (View D).

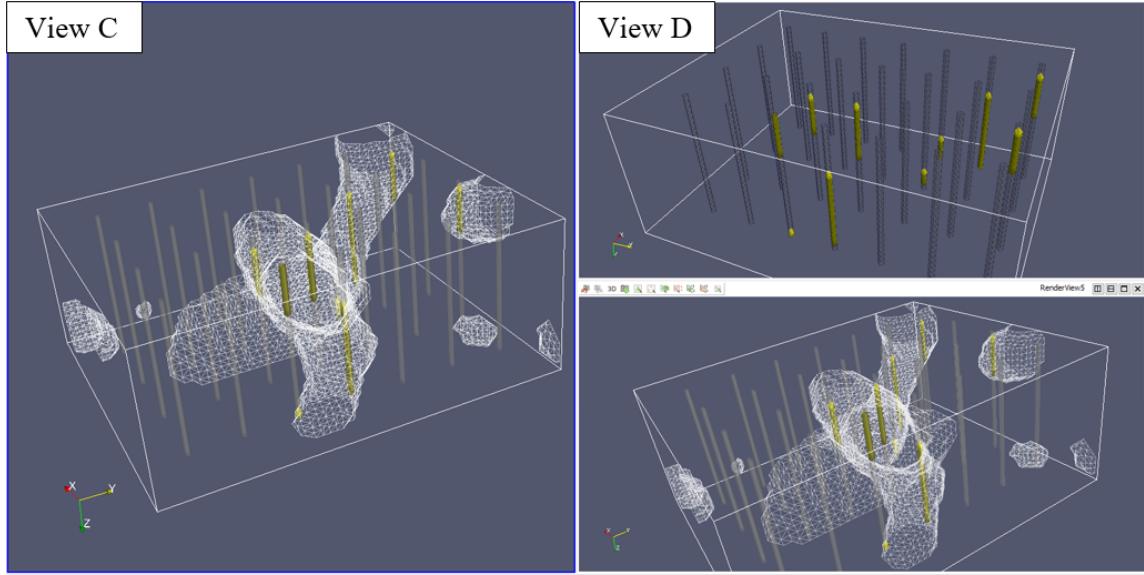


Figure 15: Initial Drill Holes Samplings Results

Calculation of an Uncertainty Index

Figure 16 shows the omnidirectional experimental variogram (in red dots) calculated from the samplings results. The theoretical variogram model selected for the Gaussian variable is an isotropic cubic model with a range of 20.0 units. The induced categorical model is presented in blue, its fitting with the experimental variogram is acceptable.

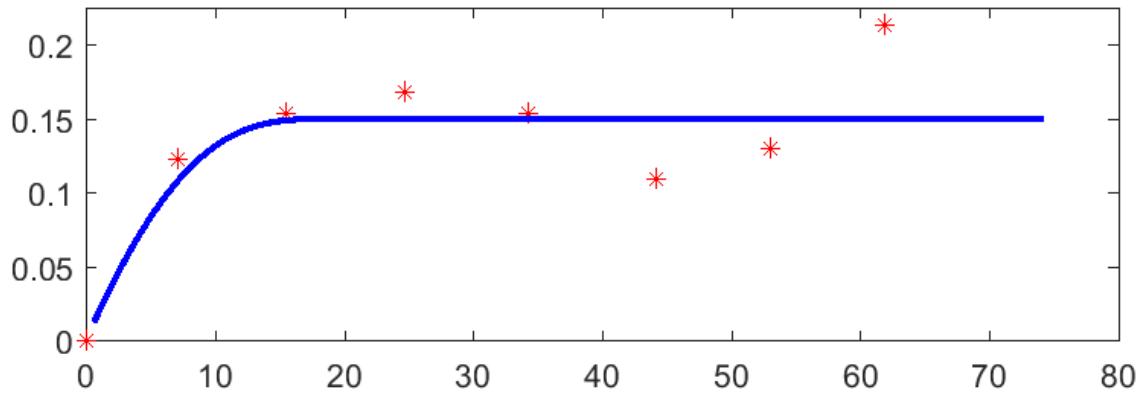


Figure 16: Experimental, Theoretical and Induced Variograms for Test Case 2

The Gibbs Sampling process is repeated 100 times (with 10 000 iterations) to provide 100 sets of independent conditioning Gaussian values. The Turning Bands algorithm is then used to create 100 conditioned simulations of mineral deposits. Figure 17 presents the average simulated mineral deposit (average of 100 simulations rounded to 0 or 1) in View G which can be compared to the original mineral deposit in View F. The simulated deposit is similar to the original deposit. The total facies error is 5987, thus the mineral representation is 91.3% accurate.

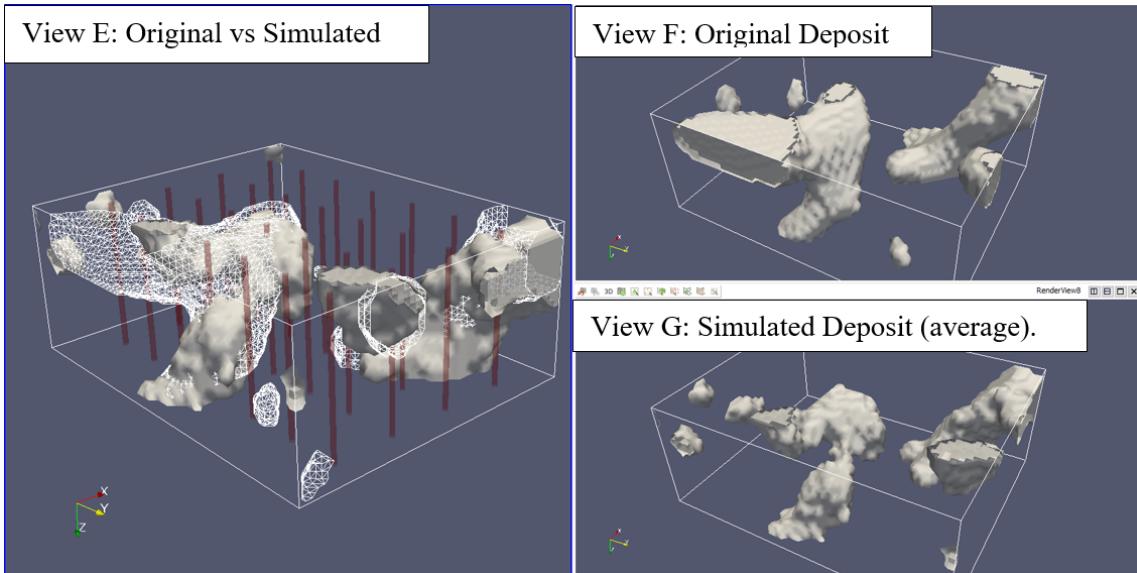


Figure 17: Simulated Mineral Deposit for Test Case 2

Figure 18 shows the blocks uncertainty index with values varying from 0.01 (blue) to 0.25 (red), in order to facilitate the visualization, blocks with uncertainty below 0.01 were not represented. View I show the original deposit in watermark, and confirms that the uncertainty is higher at blocks located nearby the deposit envelope.

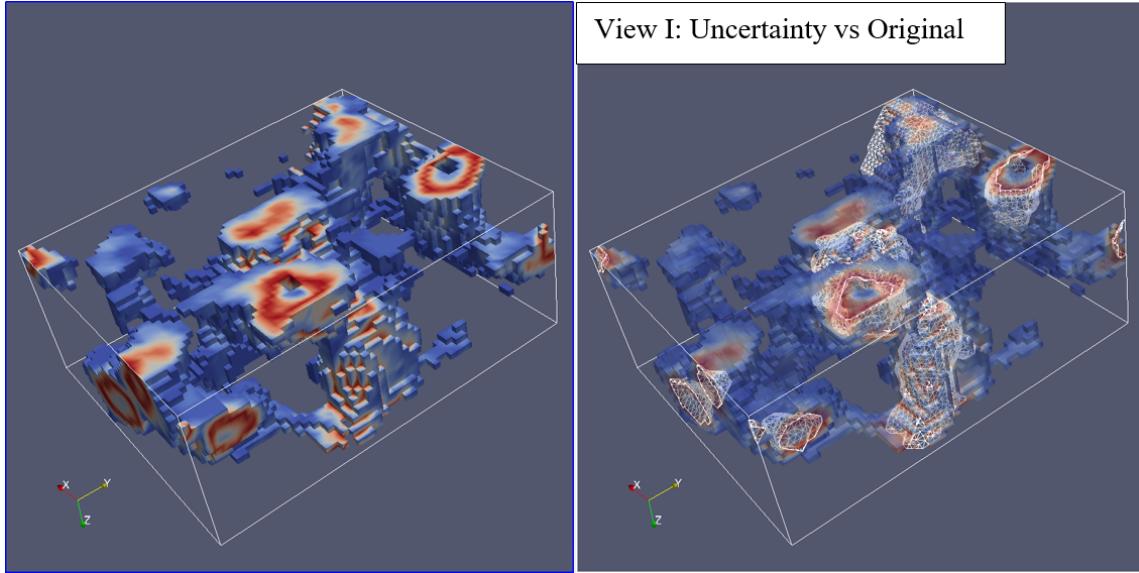


Figure 18: Uncertainty Map for Case test 2

Optimization for Future Drill Holes Selection

The domain of potential future drill holes is presented in Figure 19. Drill holes at the same collar point have six different possible lengths (4.0, 8.0, 12.0, 16.0, 20.0 or 24.0), for a total of 1710 potential drill holes, position and depth combined. We would like to know which subset of drill holes to select for a maximal drilling cost of 200.0 (distance units).

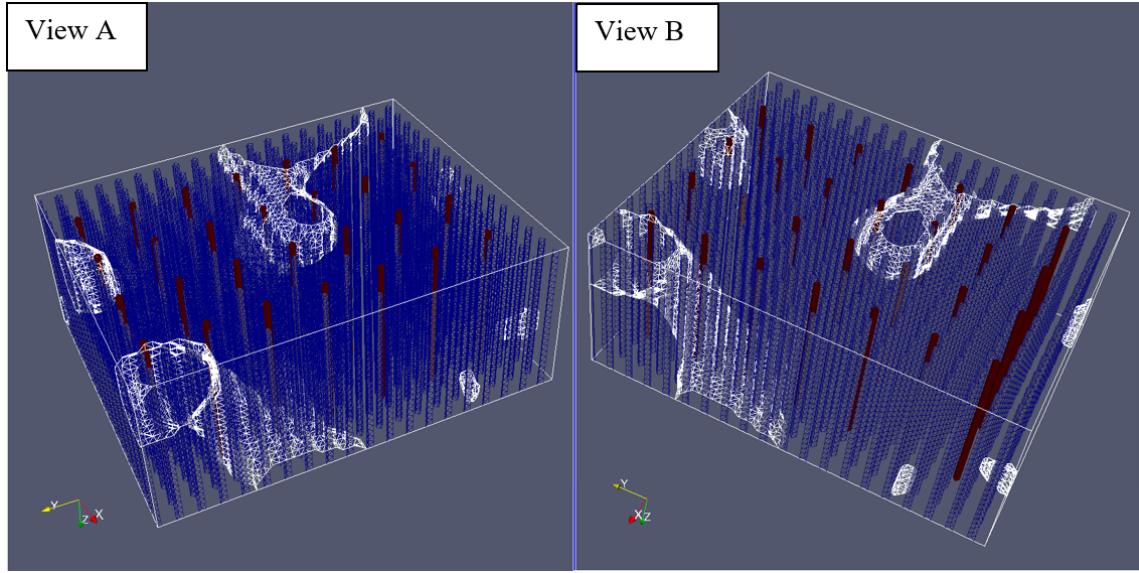


Figure 19: Potential Future Drill Holes for Test Case 2

The optimal solution found for this problem is presented in Figure 20 (green lines). The optimal solution was found in 205.5 seconds of CPU time on a *1.70GHz Intel Core i5* processor using MATLAB *Intlinprog* function. This solution has a cost of 200.0 and a total coverage score of 1379.1 (total uncertainty covered).

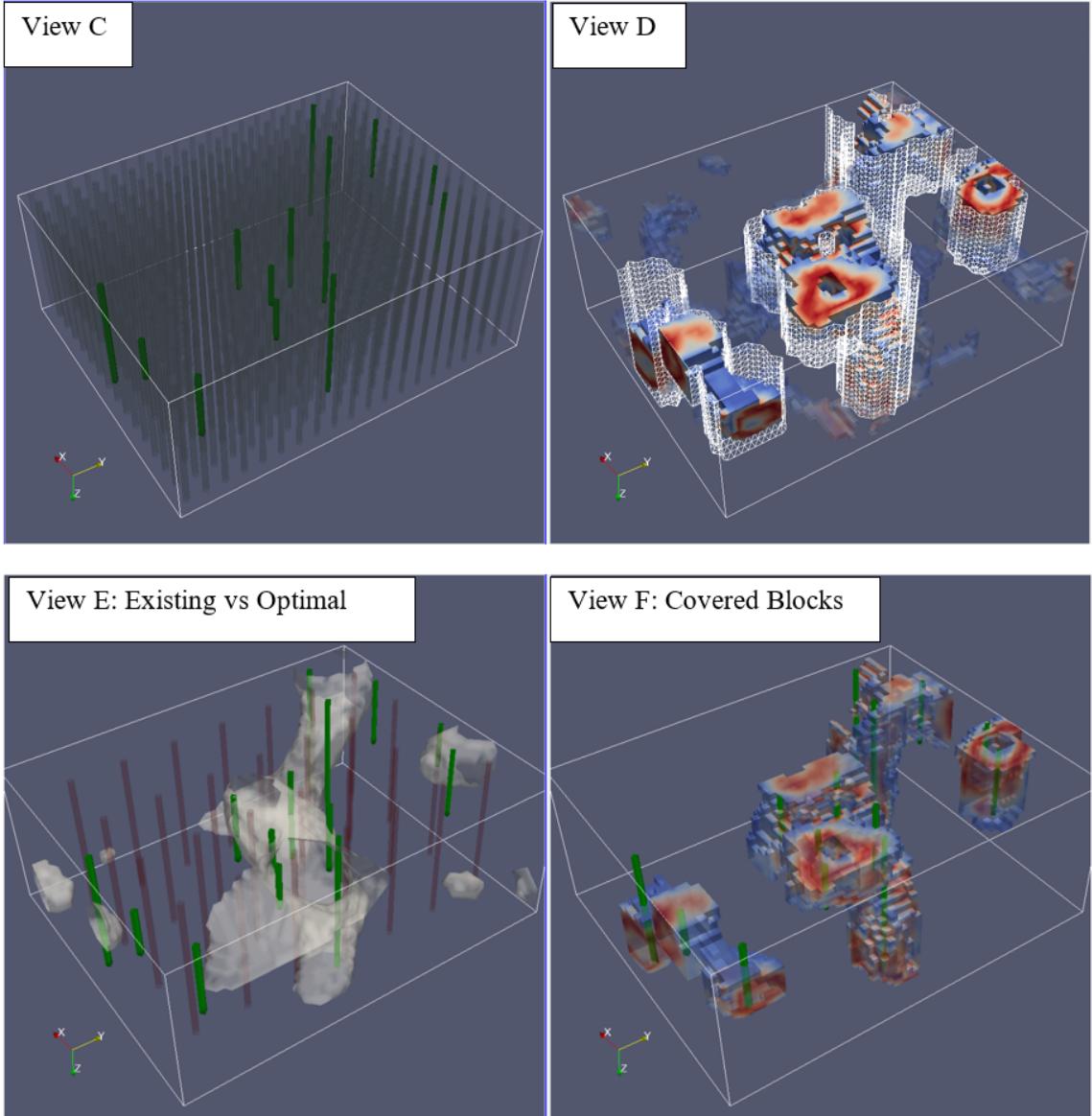


Figure 20: Optimal Drill Holes for Test Case 2

The Tabu Search algorithm applied to this problem for a maximum of 400 iterations, provided the solution presented in Figure 21 (View G). This solution is very similar to the optimal solution (view G). The best Tabu Search solution has a cost of 200 for a maximum coverage of 1243.0 (90.1% of the optimal score) and was found within 288 iterations (5947.1 seconds on

a 1.70GHz Intel Core i5 processor). Although the Tabu Search algorithm is significantly longer than the exact solver, it is less limited in terms of problem size.

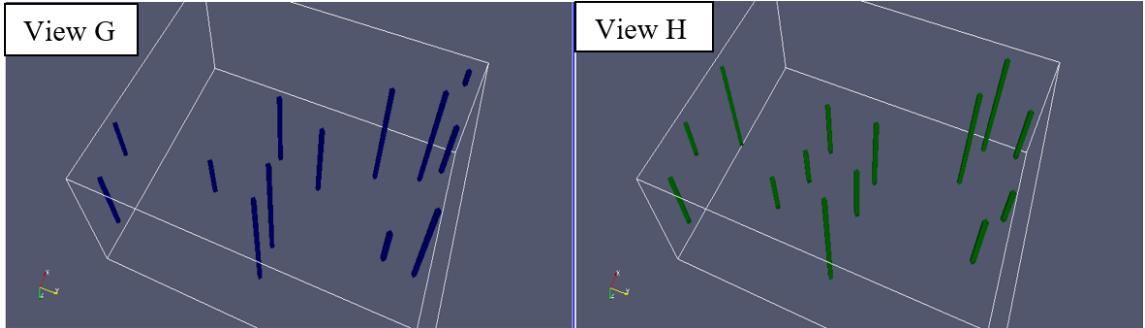


Figure 21: Optimal Solution and Tabu Search Solution for Test Case 2

4. Conclusion

A new approach for the DPP in mineral exploration context is proposed in this paper: an optimization model for maximization of the coverage of a new block uncertainty index. In the proposed approach, blocks are discriminated using an uncertainty index calculated with a geostatistical process based on interpretation of the experimental variogram, Gibbs sampling process and repeated conditioned Gaussian simulations. The uncertainty index (variance of the probability of mineral facies) calculated with this method is proven to be reliable and allows representation of the mineral deposit. Future drill holes are selected from solving an optimization model that maximizes coverage of the uncertainty index. A Tabu Search algorithm is developed to solve this optimization problem and this algorithm is shown to produce good quality results comparable to optimal solutions for small and average of the DPP problem.

5. References

- [1] G. Pan, Geostatistical design of infill drilling programs, Society of Mining Engineers of AIME 142 (1995) 1943–1952.

- [2] S. Soltani, A. Hezarkhani, E. Tercan, B. Karimi, Use of genetic algorithm in optimally locating additional drill holes, *Journal of Mining Science* 47 (2007) 62–72.
- [3] A. Hezarkhani, S. Soltani, Proposed algorithm for optimization of directional additional, *Arabian Journal of Geosciences* 6 (2013) 455–462.
- [4] N. Bilal, P. Galinier, F. Guibault, A new formulation of the set covering problem for metaheuristic approaches, *ISRN Operations Research* (2013).
- [5] K. Saikia, B. Sarkar, Exploration drilling optimisation using geostatistics : a case in jharia coal, india, *Applied Earth Science* 115 (2006) 13–22.
- [6] J.-p. Chiles, P. Delfiner, *Geostatistics: Modeling Spatial Uncertainty*, Wiley, 2012.
- [7] X. Freulon, C. de Fouquet, Conditioning a gaussian model with inequalities, in: A. Soares, *Geostatistics Toria'92*, vol. 1 . The Netherlands, 1993, pp. 201–212.
- [8] M. Cuba, O. Leuangthong, J. Ortiz, Transferring sampling errors into geostatistical modelling, *The Journal of The Southern African Institute of Mining and Metallurgy* 112 (2012) 971–983.
- [9] J. Chen, S. Hubbard, Y. Rubin, C. Murray, E. Roden, E. Majer, Geochemical characterization using geophysical data and Markov chain monte carlo methods: A case study at the south oyster bacterial transport site in virginia, *Water Resources Research* 40 (2004).
- [10] S. Lyster, C. V. Deutsch, Mps simulation with a gibbs sampler algorithm, in: *Proceedings of the 8th International Geostatistics Congress*, p. 107. Chile, 2008.
- [11] M. Reza Najafi, H. Moradkhani, Analysis of runoff extremes using spatial hierarchical bayesian modeling, *Water Resources Research* 49 (2013) 6656–6670.
- [12] H. Onibon, T. Lebel, A. Afouda, G. Guillot, Gibbs sampling for conditional spatial disaggregation of rain fields, *Water Resources Research* 40 (2004).

- [13] X. Emery, Using the gibbs sampler for conditional simulation of gaussian-based random fields, *Computers and Geosciences* 33 (2007) 522–537.
- [14] X. Emery, Tbsim: A computer program for conditional simulation of three-dimensional Gaussian random fields via the turning bands method, *Computers and Geosciences* (2006) 1615–1628.
- [15] N. Bilal, Metaheuristiques hybrides pour les problemes de recouvrement et recouvrement partiel d'ensembles appliquees au probleme de positionnement des trous de forage dans les mines, Ecole Polytechnique de Montreal, These de Doctort, Montreal, 2014.