A NOVEL APPROACH FOR GEOSTATISTICAL MODELING AND PRODUCTION SCHEDULING OF IRON DEPOSITS WITH INEQUALITY CONSTRAINTS

by

Sultan Abulkhair

2021

Thesis submitted to the School of Mining and Geosciences of Nazarbayev University in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Mining Engineering

Nazarbayev University
April 16th, 2021

EXAMINATION COMMITTEE MEMBERS:

Main supervisor:	Dr. Nasser Madani Nazarbayev University
	_____
Co-supervisor:	Dr. Nelson Morales Universidad de Chile
	_____
Internal reviewer:	Dr. Zauresh Atakhanova Nazarbayev University
	_____
External reviewer:	Dr. Mohammad Maleki Universidad Católica del Norte
	_____
GPC representative:	Dr. Lei Wang Nazarbayev University
	_____

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my Principal Supervisor, Dr. Nasser Madani, and Co-supervisor, Dr. Nelson Morales, for their valuable guidance throughout this project. Continuous support and excellent mentoring from Dr. Nasser Madani were crucial in determining the topic for this thesis, and his supervising helped to complete the geostatistics part of this work. The invaluable advice and patience from Dr. Nelson Morales played a big part in completing the mine planning part of the thesis.

The financial support from Nazarbayev University through Faculty Development Competitive Research Grants for 2018–2020 under Contract No. 090118FD5336 is most appreciated. Finally, I would like to acknowledge Delphos Mine Planning Laboratory for giving access to MineLink mine planning library.

ORIGINALITY STATEMENT

I, Sultan Abulkhair, hereby declare that this submission is my own work, and to the best of my knowledge, it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at Nazarbayev University or any other educational institution, except where due acknowledgment is made in the thesis.

Any contribution made to the research by others with whom I have worked at NU or elsewhere is explicitly acknowledged in the thesis.

I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.

Signed on April 16th, 2021

ABSTRACT

The modern mining industry employs plenty of exploration data digitization and utilizes computational resources for future forecasting and production scheduling. In that regard, geostatistics and mine planning as disciplines are critical parts of the mining business. However, traditional mine planning does not allow the risk management associated with geological uncertainty due to using a single mineral resource model as an input. Integration of stochastic geostatistical realizations into mine planning can help to minimize the risk of not meeting the production targets. Nevertheless, there are still many limitations in commercially used stochastic geostatistical algorithms, particularly multivariate methods. For example, it is common to deal with challenging datasets containing bivariate complexities among variables, such as inequality constraints. The poor reproduction of this feature can lead to an overestimation of the secondary variable, affecting the processing cost. Therefore, this work's motivation is to propose a way to model datasets containing inequality constraints between variables and generate production schedules considering geological uncertainty.

This study proposes an algorithm based on a hierarchical cosimulation framework integrated with inverse transform sampling, which is designed to model variables within thresholds derived from a linear inequation. The proposed algorithm is applied to a real case study from an Iron deposit with a sharp inequality constraint between Iron and Silica. Results are assessed based on validation of marginal distribution, spatial correlation and bivariate relationship. Overall, the proposed algorithm demonstrates simulations of similar quality as the conventional approach, which can be observed through histogram and variogram validations. Unlike conventional cosimulation, the integration of inverse transform sampling helps to reproduce an inequality constraint in the bivariate relationship and increases the correlation coefficient.

Next, a two-stage stochastic production scheduling approach is used as a replacement for deterministic mine planning. In the first stage, extraction periods are obtained using an e-type model and remain fixed for all realizations. The second stage decisions re-evaluate block destinations based on a set of realization. This approach minimizes the risk of sending an extracted material to the wrong destinations while producing optimal production schedules. Its NPV is only 17% lower than the best possible NPV for each scenario and 59% higher than the deterministic plan. Moreover, the proposed cosimulation algorithm increases the NPV by 0.76% due to the reproduction of an inequality constraint.

PREFACE

The thesis is the original work by Sultan Abulkhair. Parts of the research documented throughout the thesis were either previously published or submitted to journals and conferences.

Cosimulation algorithm proposed in this study is the original work of myself and Dr. Nasser Madani. The idea behind the mine planning methodology described in the thesis is given by Dr. Nelson Morales and completed by me under his supervision.

The cosimulation methodology described in chapter 2.1 is inspired by the paper published as Madani and Abulkhair (2020) and proposed in the other journal and conference papers submitted recently.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

This chapter defines problems addressed in this thesis, discusses limitations of conventional geostatistical and mine planning methodologies, motivates the study, identifies the significance and primary contribution to the industry.

1.1 Problem definition and motivation

The most common method for mineral resource modeling is kriging (Krige, 1951; Matheron, 1963), which is a deterministic geostatistical approach. There are three main reasons why deterministic methods are still preferred in the mining industry:

1. It takes less computation time because it produces one result instead of multiple realizations in stochastic algorithms.

2. It is based on the assumption that there can only be one singular geological model.

3. Standard mine planning programs use only one block model as an input.

However, earth science is dealing with the complex subsurface phenomenon, which is an uncertain medium. Most reliable data come from boreholes that sample only a fraction of the deposit. Therefore, integration of probability is crucial and can be used for risk and economic assessment. Stochastic geostatistical methods provide multiple unbiased scenarios that reproduce the spatial correlation of original data and can be used for uncertainty quantification (Goovaerts, 1997).

Conventional mine planning produces scheduling and pit constraints based on a single block model. For that reason, lack of risk management in conventional long-term production scheduling approaches can lead to potential losses due to deviations from production targets. Furthermore, the application of deterministic integer programming methodologies for multiple realizations obtained from stochastic simulations produces inefficient results (Ramazan and Dimitrakopoulos, 2004). Therefore, such limitations indicate that there is a need for stochastic production algorithms capable of optimally integrating geological uncertainties. In that regard, stochastic production scheduling aims to maximize net present value while minimizing the risks associated with uncertainty. Although recent research contributions demonstrate promising

results, the efficiency of some stochastic production scheduling approaches is still an issue due to the high number of binary variables and constraints.

Another limitation is the ability of conventional stochastic geostatistical methodologies to deal with multivariate relationships. It is common practice in the mining industry to have different types of complexities between variables of interest (i.e., metal grades, geometallurgical variables and geophysical properties), such as nonlinearity, heteroscedasticity and geological constraints. In this context, geological constraints, particularly inequality and equality constraints, can be observed between coregionalized grade variables (Madani and Abulkhair, 2020) and between geomatallurgical variables (Abildin et al., 2019). Unfortunately, traditional geostatistical approaches cannot reproduce these constraints, decreasing the validity of produced models. As a result, poor reproduction of geological constraints leads to an overestimation of secondary variables and affects the NPV.

For this study, several problems identified in mine planning and geostatistics are listed as follows:

- Conventional mine planning produces scheduling and pit constraints based on a single block model, making it suboptimal as it does not manage the risk of not meeting production targets.

- Cokriging, a common multivariate modeling technique, suffers from a smoothing effect in the presence of an unequal sampling pattern.

- Stochastic mine planning highly depends on the quality of simulation/cosimulation results, which is indicated by the reproduction of histograms and variograms.

- Reproduction of inequality constraints in bivariate relationships is challenging for current cosimulation techniques.

Therefore, motivation to conduct this study comes from the following issues that the industry is facing: the need for unbiased uncertainty quantification, stochastic production scheduling and reproduction of inequality constraints between variables.

1.2 Thesis objectives

The primary objective of this thesis is to develop a stochastic geostatistical algorithm based on hierarchical cosimulation to model variables with inequality constraints. Models obtained from this algorithm are going to be integrated into two-stage stochastic long-term production scheduling. As a result, this thesis's primary contribution will be a mine planning that implements the multiple grade uncertainties and respects inequality constraints between variables. The completion of the following sub-goals is required in this study:

- To conduct an in-depth literature review on mine planning and geostatistics, paying particular attention to multivariate geostatistics for deposits with complex bivariate relationships and stochastic production scheduling for open-pit deposits;
- To develop a code for geostatistical cosimulation of cross-correlated variables with inequality constraints;
- To prepare a framework for two-stage stochastic long-term open-pit production scheduling.
- To test proposed methods on a real Iron deposit, where Iron is the main element, silica is a disturbing element, and a sharp inequality constraint exists in the bivariate relationship.
- To compare the simulated results with the traditional geostatistical cosimulation algorithm that cannot reproduce an inequality constraint;
- To compare the proposed two-stage stochastic production scheduling with upper and lower bounds.
- To assess how much the proposed geostatistical cosimulation algorithm can increase/optimize the NPV of the project for a long-term production plan.

1.3 Thesis statement and contribution to the industry

Thesis statement: *A new two-stage stochastic long-term production scheduling and novel hierarchical cosimulation algorithm integrated with an inverse transform sampling to model variables with inequality constraints increase economic profit and minimize the deviations from the production targets.*

The importance of this project is that it provides a workflow for stochastic strategic mine planning in Iron deposits, where the Iron shows a complex bivariate relationship with covariates such as Silica. This thesis offers a robust geostatistical methodology designed explicitly for deposits with inequality constraints between variables. The methodology can be applicable to many mining companies dealing with the modeling of Iron deposits. Furthermore, mine planning results will be more valid if the disturbing element is not overestimated and the primary element is accurately modeled. The benefit of this project will add more value to the company and eventually lead to maximizing the NPV of their Iron projects. The proposed algorithm can be provided to the companies for geostatistical modeling because the commercial software packages cannot deal with inequality constraints. Aside from tedious variogram modeling, the developed algorithm is straightforward and can be used by geostatisticians and mine engineers through simple instructions.

1.4 Thesis outline

Chapter 2 demonstrates a comprehensive literature review on geostatistics and mine planning. This review is mainly focused on common univariate and multivariate geostatistical algorithms, methods designed for the modeling of datasets with complex bivariate relationships, deterministic and stochastic mine planning approaches.

Chapter 3 describes proposed geostatistical modeling and production scheduling methodologies. Hierarchical sequential Gaussian cosimulation integrated with inverse transform sampling is proposed for geostatistical modeling of variables with inequality constraints. The production scheduling method in this study is based on a two-stage stochastic approach. The first stage decisions are based on extraction periods from average information and the second stage uses multiple scenarios to re-evaluate block destinations.

Chapter 4 demonstrates the performance of the proposed methodologies on a real case study from an Iron deposit with a sharp inequality constraint between Iron and Silica. The comparison between proposed and conventional hierarchical cosimulation algorithms is made using geostatistical validations. Meanwhile, two-stage stochastic production scheduling is assessed by comparing it with upper (i.e., best NPV) and lower bounds (i.e., deterministic case).

Chapter 5 provides a summary of results, limitations and recommendations.

2 LITERATURE REVIEW

A comprehensive literature review of information relevant to this study is provided in the following chapter. This review consists of three sections: 1) estimation and simulation, 2) multivariate geostatistics, 3) open-pit mine planning.

2.1 Estimation and simulation

Resource estimation is the initial step in the exploration and production stages of mineral extraction in the mining and petroleum industries. Geostatistical estimation and simulation algorithms aim to use limited exploration information from boreholes and geophysical investigations to produce unbiased and spatially correct models (Goovaerts, 1997; Pyrcz and Deutsch, 2014). In practice, mineral resource estimation consists of several steps (Rossi and Deutsch, 2014):

- Collection of borehole samples used as input conditioning data;
- Interpretation or modeling of geological or lithological domains (i.e., the definition of estimation domains);
- Modeling grades of minerals, petrophysical properties, and other types of parameters inside each estimation domain for further stages in a mining project.

Geostatistical algorithms are divided into two main groups closely related to each other: deterministic (estimation) and stochastic (simulation) methodologies. The former focuses on finding one unique result, while the latter uses random number generation to produce different yet similar realizations. The following subsections will review traditional univariate geostatistical algorithms.

2.1.1 Kriging

Kriging is a geostatistical tool developed in the middle of the 20th century by Daniel Krige and used as a basis for most geostatistical algorithms (Krige, 1951; Matheron, 1963). It predicts values in unsampled locations using spatial continuity data represented by an experimental variogram, a measure of variability over certain distances (Chilès and Delfiner, 2012). The

theoretical variogram model can be fitted manually or by using semi-automated (Larrondo et al., 2003) and automated (Emery, 2010) approaches. This estimator's crucial advantage is calculating kriging weights that reduce expected error variance, making this method acceptable in terms of the least-squares criterion. However, there are limitations in this method's performance: its deterministic behavior, overestimation of low-grade and underestimation of high-grade values (smoothing effect), and screening out the influence of one data by another (screening effect). The most popular types of kriging are (Mata-Lima, 2005):

- Simple Kriging (SK): an unbiased kriging algorithm that uses the mean of original samples to predict variables.
- Ordinary kriging (OK): calculation of mean is not necessary because the summation of all the weights is assumed to be equal to 1.
- Universal kriging (UK): used in the presence of a trend that slightly changes mean values.
- Non-linear kriging: linear kriging algorithms (SK and OK) that are used on original samples being transformed into non-linear transforms, such as gaussian-based, lognormal and indicator kriging (IK).

2.1.2 Simulation

The smoothing effect problem of the kriging algorithm can be overcome by using stochastic simulation algorithms to produce multiple unbiased realizations. Validation of correctness of produced realizations is often based on the ability to reproduce histogram and variogram of original data, which are first and second-order statistics, respectively (De Iaco and Maggio, 2011). Simulation techniques are not totally unbiased but less biased than estimation methods, enabling them to reproduce marginal distribution and local variability.

The most common geostatistical simulation methodologies are Gaussian-based algorithms, which use multivariate Gaussian assumption on grades transformed into normal scores. For example, sequential Gaussian simulation (SGSIM) uses mean and variance obtained from kriging methods (usually SK) and a random number generated from conditional distribution (Isaaks, 1990). However, because of the maximum entropy, connectivity between the minimum and maximum values is either low or disconnected. Another popular algorithm is turning bands simulation (TBSIM) that generates the trend along one-dimensional lines and uses random

numbers similar to SGSIM (Matheron, 1973; Emery and Lantuéjoul, 2006). Thanks to computational resources available today, which enables simulation of a significant number of lines, this method is still used after nearly 50 years. Other methods for simulating continuous variables include LU decomposition (Luster, 1985), direct sequential simulation (Soares, 2001) and simulated annealing (Deutsch, 1992).

2.2 Multivariate geostatistics

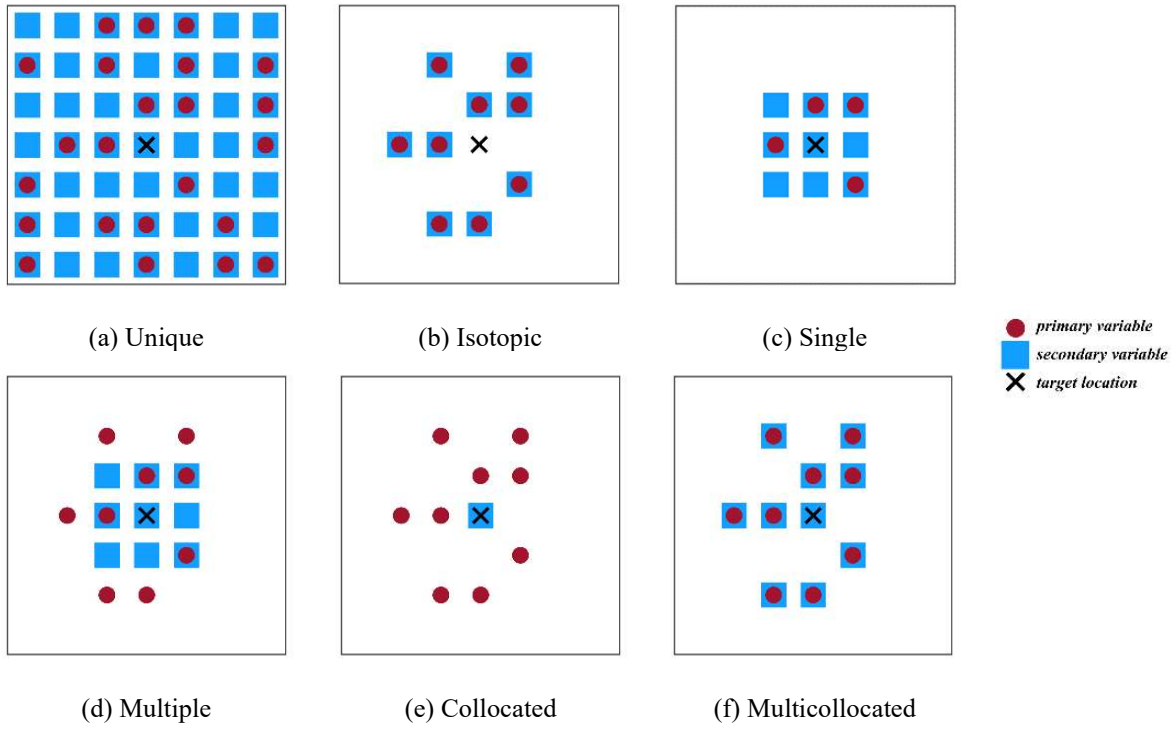
A multivariate extension of common univariate geostatistical algorithms, described in the previous section, is a way to model two or more variables and reproduce the relationship between them (Wackernagel, 2003; Chilès and Delfiner, 2012). Together with variograms or covariances of each variable, multivariate methods use cross-covariances or cross-variograms as a measure of inherent correlations among variables. It defines the coregionalization model that can be taken into account by conventional cokriging and cosimulation techniques, particularly a linear model of coregionalization (LMC) and the Markov model are popular options today. Another alternative is to use factorization or transformation-based methods for cases that involve complexities between variables, such as heteroscedasticity, nonlinearity and geological constraints. The following subsections discuss common cokriging neighborhood strategies, conventional cokriging and cosimulation algorithms, factorization-based methods and special approaches for modeling variables with inequality constraints.

2.2.1 Cokriging neighborhood configurations

Similar to univariate geostatistics, there are both deterministic and stochastic multivariate methods. However, multivariate datasets are larger and require more computational resources, considering that there are two or more variables that need to be processed (Emery, 2009). Another difficulty is the sampling type, variables can be entirely heterotopic, partially heterotopic or isotopic, depending on how they are scattered in space (Wackernagel, 2003):

- Isotopic: variables are sampled from the same locations.
- Entirely heterotopic: variables are sampled from entirely different locations.
- Partially heterotopic: there are both isotopic and heterotopic sets of data.

These challenges can be settled using a limited number of data points in the cokriging system and implementing various moving neighborhood strategies. Therefore, based on the neighborhood strategies, new cokriging methods emerged. Hence, the integration of neighborhood strategies into cokriging systems gave rise to strictly collocated, multicollocated and heterotopic cokriging. Furthermore, selecting an appropriate searching strategy affects the screening effect (Rivoirard, 2004; Subramanyam and Pandalai, 2004). The examples of common neighborhood strategies are demonstrated in Figure (1).



(a) Unique (b) Isotopic (c) Single

(d) Multiple (e) Collocated (f) Multicollocated

Figure 1. Schematic illustrations of the unique neighborhood (a); isotopic (b), single (c), multiple (d), collocated (e) and multicollocated (f) search strategies of moving neighborhood for 9 closest data points

For the bivariate cases, the following neighborhood strategies are usually incorporated into cokriging and cosimulation algorithms (Madani and Emery, 2019):

1. Unique neighborhood: Every data point that belongs to primary and secondary variables is selected in the cokriging system (Figure 1a).

2. Isotopic: A moving neighborhood that searches for n closest data points shared by both variables (Figure 1b).

3. Single: A heterotopic moving neighborhood that searches for n closest data locations, whether the primary or secondary variables are known at these locations (Figure 1c).

4. Multiple: A heterotopic moving neighborhood strategy that first searches for n closest data of the primary variable; then searches for n closest data of the secondary variable, irrespective of the first search (Figure 1d).

5. Collocated (Xu et al., 1992): A moving neighborhood strategy that searches for n closest data of the primary variable and the secondary data at the target location (Figure 1e).

6. Multicollocated (Rivoirard, 2001): A moving neighborhood that searches for n closest data points shared by both variables and the secondary data at the target location (Figure 1f).

Comparison of all neighborhood configurations from Figure (1) suggests that the two best options in terms of cokriging variance are multicollocated and multiple searching strategies, in which the former is slightly more accurate (Madani and Emery, 2019).

2.2.2 Cokriging and cosimulation

This thesis focuses on stochastic multivariate geostatistics, which means that cokriging methods are not reviewed as potential alternatives. However, simple cokriging approaches based on different neighborhood configurations are usually incorporated in various cosimulation methodologies. Therefore, these methods are discussed before cosimulation approaches.

Simple cokriging is a multivariate extension of simple kriging, thus it considers the mean values of the primary and secondary variables as well as spatial correlation structure (Myers, 1982; Wackernagel, 2003). Fitting covariance matrices is usually done by a linear model of coregionalization using nested structures (Goovaerts, 1997). However, it is common to have multiple variables that do not share sample locations, for which conventional simple cokriging might not be the best choice. Heterotopic simple cokriging is a way to model such datasets by using single or multiple moving neighborhood strategies. In the case of entirely heterotopic sampling, where the calculation of cross-variograms is not possible, either direct and cross-covariances or pseudo-cross variograms (Myers, 1991) can be used instead.

Using the Markov model (i.e., type I and type II) to define the spatial correlation is an alternative for LMC, which is suitable for collocated cokriging (Journel, 1999). A strictly collocated search

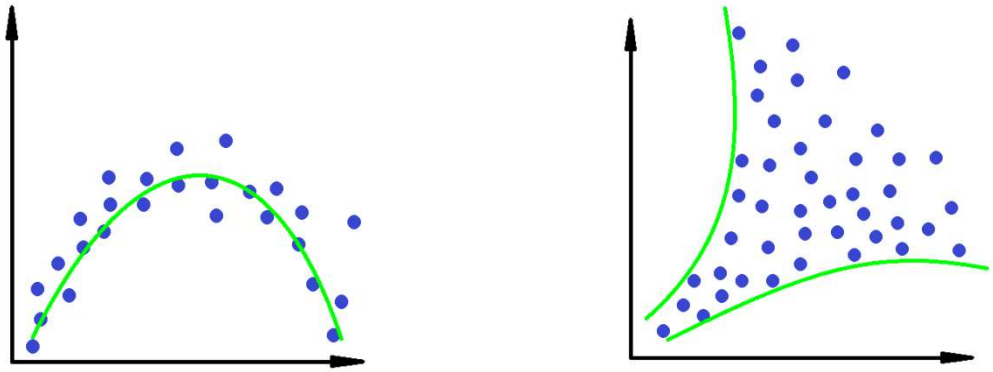
strategy is used for the moving neighborhood in this type of cokriging (Xu et al., 1992). The literature extensively stated that collocated secondary data at the target location screens out the secondary data located further away (Journel, 1999). However, various studies show that this statement may be misleading. For example, tests on the screening effect in collocated cokriging demonstrate that it is not equivalent or even close to full cokriging with a unique neighborhood (Madani and Emery, 2019). Rivoirard (2001) suggests that strictly collocated cokriging is an approximation to cokriging and proposes multicollocated cokriging as a better option. For this cokriging system, using the intrinsic correlation models is a replacement for LMC and Markov models (Wackernagel, 2003). Multicollocated cokriging has been proven to be a better cokriging estimator in other studies as well (Paravarzar et al., 2015; Madani and Emery, 2019).

Cokriging methods have the same disadvantages as kriging and are inferior to cosimulation algorithms. Notation of uncertainty quantification in multivariate resource modeling is as crucial or perhaps even more critical compared to univariate geostatistics. For example, SGSIM and TBSIM algorithms are extended to model multiple variables by using different cokriging methodologies (typically simple cokriging) as an estimator. Among many others, the following conventional cosimulation algorithms are popular today: sequential Gaussian cosimulation (SGCOSIM) (Verly, 1993) and turning bands cosimulation (TBCOSIM) (Emery, 2008). A hierarchical SGCOSIM (Almeida and Journel, 1994) simulates variables in a hierarchical manner: first primary variable using simple kriging; then secondary variable using collocated cokriging conditional to the previous simulation. Multicollocated cokriging can also be used in the second run of the simulation, while simple kriging can be replaced with simple cokriging (Madani and Abulkhair, 2020). Paravarzar et al. (2015) compared SGCOSIM and TBCOSIM approaches using the real case study and concluded that the sequential algorithm is slightly more biased, especially when collocated cokriging is used as an estimator.

2.2.3 Factorization methods

Conventional cosimulation methods, which employ a rather dreary fitting of direct and cross-covariances, are constantly compared to factorization-based approaches. These are advanced techniques for decorrelation of variables and then their independent simulation. The following normal transformations and decorrelations can be used in multivariate modeling: principal component analysis (PCA) (Davis and Greenes, 1983), minimum/maximum autocorrelation factors (MAF) (Desbarats and Dimitrakopoulos, 2000; Rondon, 2012), enhanced

coregionalization analysis (ECA) (Emery and Ortiz, 2012). Although these algorithms seem to be great alternatives to conventional cosimulation, they have their drawbacks as well. For example, they are not adequate to model complex multivariate datasets because they are still based on linear models of coregionalization. This means that PCA, MAF and ECA cannot reproduce nonlinearity and heteroscedasticity complexities between variables. Figure (2) shows schematic representations of these two complexities.



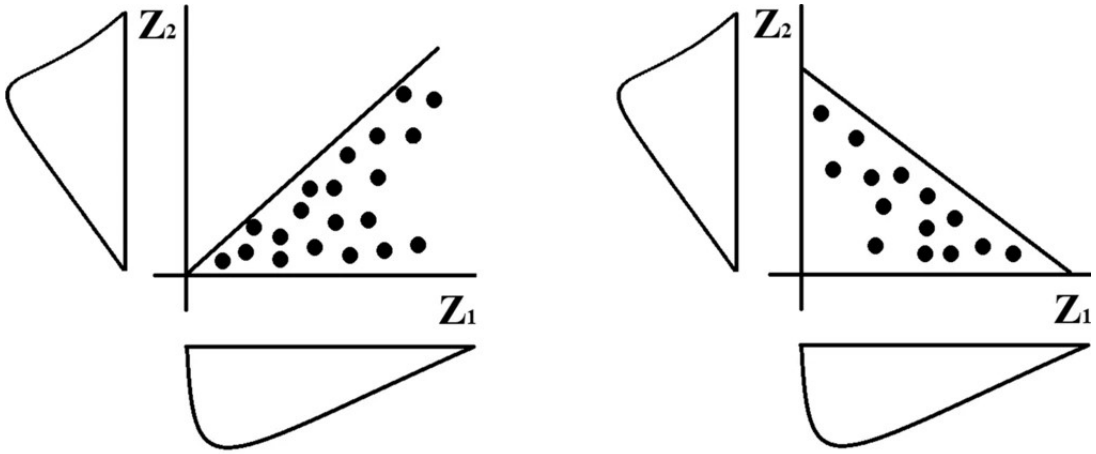
(a) Nonlinearity (b) Heteroscedacity

Figure 2. Common bivariate complexities, namely nonlinearity (a) and heteroscedasticity (b) (Battalgazy and Madani, 2019b)

There are other factorization-based geostatistical approaches, such as stepwise conditional transformation (SCT), which transforms original variables into uncorrelated multiGaussian variables (Leuangthong and Deutsch, 2003). Because transformed variables do not have any multivariate complexity and correlation, their simulation and then back transformation ensures the reproduction of bivariate relationships. Removing complexity and spatial correlation between variables at lag zero does that for other lags as well, which is something both SCT and PCA share. Another alternative is projection pursuit multivariate transform (PPMT) that utilizes projection pursuit density estimation developed by Friedman (1987) and can reproduce nonlinearity and heteroscedasticity (Barnett et al., 2014, 2016). In theory, this transformation can consider any type of bivariate complexity. In Battalgazy and Madani (2019a), PPMT is compared with TBCOSIM and TBSIM, which shows that the transformation-based method is considerably better in reproducing correlation coefficients and improves tonnage estimation. Nevertheless, Van den Boogaart et al. (2017) developed a flow anamorphosis algorithm based on an affine-equivariant transformation, showing that it can be more robust than PPMT.

2.2.4 Methods for geological constraints reproduction

Multivariate modeling of coregionalized variables deals with other types of complexities, aside from Gaussianity, nonlinearity and heteroscedasticity. One of those complexities that are challenging for algorithms reviewed so far is inequality constraint. It is a geological constraint that can be expressed by strict/non-strict inequalities or characterized by an inequation (i.e., linear inequation with slope and intercept). Schematic representation of inequality constraints in the presence of positive and negative correlations is shown in Figure (3).



(a) Positive correlation (b) Negative correlation

Figure 3. Examples of bivariate relationships expressed by inequality constraints with positive (a) and negative (b) correlations (Madani and Abulkhair, 2020)

Several authors tried to develop methodologies for modeling multivariate datasets with inequality, sum and fractional constraints. Older algorithms for such cases include quadratic programming (Mallet, 1980) and interpolated functions taking into account inequality and equality constraints (Dubrule and Kostov, 1986). SCT developed by Leuangthong and Deutsch (2003) can be used to model and reproduce inequality constraints between geometallurgical variables (Hosseini and Asghari, 2015). Another way of integrating transformation-based methods into modeling inequality constraints is to transform variables into Gaussian random fields through SCT and then cosimulate them using LMC (Emery, 2012). Arcari Bassani et al. (2018) propose implementing PPMT (Barnett et al., 2014) on variables changed into ratios, which can reproduce this type of complexity. Transformation of original variables to log-ratio (Pawlowsky-Glahn and Olea, 2004; Pawlowsky-Glahn and Egozcue, 2006), using stoichiometric relations (Mery et al., 2017; Adeli et al., 2018) and Gibbs Sampling (Emery et al., 2014) are another set of algorithms that are worth mentioning. Furthermore, Abildin et al.

(2019) proposed a methodology that implements MAF (Desbarats and Dimitrakopoulos, 2000) to change original data to variables free of inequality constraints (Figure 4).

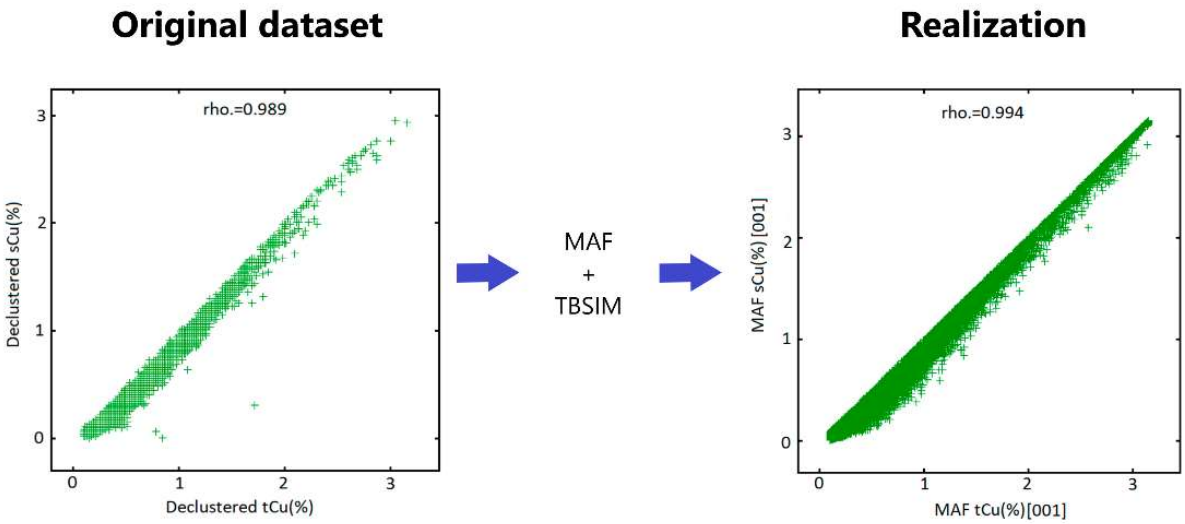


Figure 4. Reproduction of bivariate relationship using an approach proposed by Abildin et al. (2019)

2.2.5 Hierarchical cosimulation with an acceptance-rejection method

All of the methods that involve transformation-based algorithms to convert original variables to correlation-free variables have two drawbacks: restriction to homotopic sampling pattern and identical marginal distributions (Madani and Abulkhair, 2020). Alternatively, inequality constraints can be reproduced by a novel algorithm that is based on a hierarchical cosimulation framework (Almeida and Journel, 1994) integrated with an acceptance-rejection method, which rejects secondary data located outside the inequation (Madani and Abulkhair, 2020). On top of that, there is a methodology to obtain the minimum support line for an acceptance-rejection method, which identifies the regression line and shifts it to include all data points. However, this algorithm has two unresolved issues. Firstly, an acceptance-rejection method is an iterative algorithm, thus it significantly reduces the simulation speed. Secondly, this method is based on single searching strategy for the first simulation and multicollocated for the second, meaning that partially heterotopic cases may cause some issues. The solution for the first issue is replacing an acceptance rejection technique with methods that re-simulate variables in one iteration. Finally, an original algorithm can be employed with a multiple searching strategy that has a better performance than a single search (Madani and Emery, 2019).

2.3 Open-pit mine planning

According to Dagdelen (1985), the standard procedure for open-pit mine planning is as follows:

1. Construct a resource block model using geostatistical methods (i.e., kriging).
2. Generate the economic block model from the resource model, predefined prices and costs.
3. Identify the ultimate pit limit (UPL) based on the economic block model.
4. Within this pit limit, compute nested pits (mining pushbacks).
5. Generate a production schedule for the project's life of mine (LoM).

After almost two decades, the general procedure has not changed much (Hustrulid et al., 2013). Production of accurate resource models and challenges in geostatistics are discussed in sections 2.1 and 2.2. The next stage is to convert the modeled parameters into an economic variable, determining the necessity of mining for each produced block. That is conducted by identifying the revenue based on the amount of ore, its grade and its market price, then subtracting operational and production costs. This determines each individual block's profitability, and then, based on economic evaluation methods, such as net present value (NPV), an optimal solution can be found.

However, the fact that this approach has not changed since the 20th century is not an indication of stability. In traditional mine planning, information is incorporated in an indirect or limited way. More importantly, they require deterministic models generated by geostatistical estimation methods that tend to produce unreliable results for reasons discussed in the earlier sections. This section provides the history of open-pit mine planning and reviews current methodologies for ultimate pit limit identification, deterministic and stochastic production scheduling techniques.

2.3.1 Ultimate pit limit

The idea of using economic values of mining blocks to identify the boundary of an open-pit was first introduced as an optimization algorithm based on the graphic theory (Lerchs and Grossmann, 1965), moving cone algorithm (Pana, 1965), math programming (Johnson, 1968) and network-flow algorithm (Picard, 1976). The former, for example, weights all blocks in the

model based on their economic value and then generates pit constraints. These solutions provide optimal results that can be applied to three-dimensional models. The idea was later called ultimate pit limit or UPL, which basically means boundaries of the part of the deposit that is planned to be mined.

The ultimate pit limit (UPL) problem is one of the well-researched topics in the mine planning community compared to production scheduling (Espinoza et al., 2013). Among many solutions, Lerchs-Grossman (LG) algorithm (Lerchs and Grossmann, 1965) is a popular method used in commercial software. The original algorithm can be applied to both two-dimensional and three-dimensional cases, using the graph method in the latter. The implementation is based on the economic block model that contains the value of each block and geometrical restriction represented by slope precedence. Ultimately, the main objective of computing UPL is to maximize the NPV of the project by finding the most feasible final pit contour. LG algorithm connects nodes in a tree form using a graph method according to block values and pit design. The approach can be expressed as follows:

$$\max \sum_{\alpha \in M} v_\alpha p_\alpha \tag{1}$$

$$subject\ to \quad p_\alpha \leq p_\beta \qquad \forall \alpha \in A, \beta \in B \tag{2}$$

$$p_\alpha \in \{0,1\} \qquad \forall \alpha \in A \tag{3}$$

where v_α is block value, p_α is an integer variable that indicates if a block is extracted ($p_\alpha = 1$) or left untouched ($p_\alpha = 0$), A is a set of blocks α, B is a set of blocks β that are predecessors of blocks α.

However, the LG algorithm can be considered a maximum closure problem that can be reduced to a minimum cut network flow problem, significantly reducing the computation time (Deutsch et al., 2015). One of the best efficient ways to solve a network flow problem is using the pseudoflow algorithm (Hochbaum and Chen, 2000; Hochbaum, 2001). For more details, readers are referred to Hochbaum (2008) that demonstrates different methods, such as pseudoflow, pseudoflow simplex and parametric implementations of the two algorithms.

2.3.2 Deterministic methods

Deterministic, or conventional, production scheduling techniques produce results based on a single geological model, which has the following shortcomings: 1) deterministic block models may be unrealistic and are not able to predict the truth with high accuracy; 2) these methods poorly reproduce the in-situ variability; 3) time value is not considered (Dowd et al., 2016). Lerchs and Grossmann (1965) enabled many earlier methodologies to be developed (Johnson, 1968; Gershon, 1983; Dagdelen, 1985; Hoerger et al., 1999; Whittle, 1999). These methods are extensively reviewed by Osanloo et al. (2008) with pros and cons, recommendations and comments on uncertainty-based algorithms. The scheduling can be done by using heuristic algorithms (Dagdelen and Francois-Bongarcon, 1982; Gershon, 1987), dynamic programming (Tolwinski and Underwood, 1996), mixed-integer programming (MIP) (Gershon, 1983) and Lagrangian relaxation (Dagdelen and Johnson, 1986).

Integration of the branch and cut method into an integer programming algorithm is another possible solution (Caccetta and Hill, 2003). In the branch and cut method, a binary variable representing the mining of blocks in a certain period is introduced in the formulation to solve the linear programming relaxation problem. More recently, Ramazan (2007) used the tree algorithm to propose a block aggregation method of production scheduling, in which the number of binary variables is reduced to be applicable for large projects. Bienstock and Zuckerberg (2010) presented another method of solving linear programming relaxation in cases involving arbitrary side constraints. This algorithm uses Lagrangian relaxation in the linear programming problem and produces promising results. Among others, the knapsack problem (Ibarra and Kim, 1978) is another crucial problem because it involves maximizing profits and satisfying precedence constraints. Moreno et al. (2010) developed an algorithm to solve it in multiple periods for open-pit production scheduling based on linear programming.

However, Ramazan and Dimitrakopoulos (2004) demonstrated that traditional MIP that uses a single estimated model as an input is not practical in terms of equipment movement and smoothness, while incorporating simulated models leads to the production of suboptimal results. Suppose traditional MIP models cannot produce feasible solutions with and without accounting on grade variability. In that case, there is a need for new stochastic methodologies capable of optimally integrating geological and grade uncertainties.

2.3.3 Stochastic methods

One of the main advantages of stochastic mine planning is its possible application for risk assessment that was first researched by Ravenscroft (1992) and Dowd (1994), and later continued by Dimitrakopoulos et al. (2002). These research contributions suggest that using stochastic geostatistics is a better option than kriging methods. The idea behind stochastic production scheduling is to use many possible scenarios of the economic block model to produce more robust schedules by maximizing NPV and minimizing risks (Dimitrakopoulos, 2011). Long-term production scheduling that takes uncertainty into account can be achieved by simulated annealing (Godoy and Dimitrakopoulos, 2004; Albor Consquega and Dimitrakopoulos, 2009) and stochastic integer programming (SIP) (Ramazan and Dimitrakopoulos, 2007, 2013). Furthermore, stochastic mine planning implemented by either of the above methods can increase NPV up to 25%, while using the stochastic approach for optimal pit limits can also produce 15% higher tonnage and larger pits (Dimitrakopoulos, 2011). Even though risk-based methodologies, such as simulated annealing, can significantly increase NPV up to 26-28% (Godoy and Dimitrakopoulos, 2004; Leite and Dimitrakopoulos, 2007), grade variability is not directly embedded in those methods. On top of that, another difference between simulated annealing and SIP is that the latter makes optimal waste removal decisions (Dimitrakopoulos, 2011).

SIP is mathematical programming that considers multiple realizations (i.e., stochastic geostatistical models) to produce optimal production schedules, in which risks are managed and NPV is maximized (Ramazan and Dimitrakopoulos, 2007). Another definition of SIP is MIP's probabilistic extension, which has variables containing uncertainty (Escudero et al., 1993). In SIP, the risk is defined as the loss of not meeting production targets. To minimize the loss and maximize NPV, SIP can be equipped with a two-stage recourse model containing both anticipative and adaptive models of stochastic optimization (Dimitrakopoulos and Ramazan, 2008; Ramazan and Dimitrakopoulos, 2013). Many stochastic production scheduling methods have way more binary variables than MIP, which makes them inefficient. However, SIP approaches can manage these issues by not using binary variables for waste blocks and defining deviation as a continuous variable.

Even though stochastic mine planning algorithms still did not replace conventional methods, many research outputs and multiple unbiased comparisons with statistical evidence suggest that they produce more accurate schedules and higher economic values. Table (1) shows the

economic benefits of simulated annealing and SIP reported in research outputs with real case studies, in which stochastic production scheduling methods are compared to conventional approaches.

Table 1. Economic benefits of stochastic mine planning

Method	Reference	Findings
Simulated Annealing	(Godoy and Dimitrakopoulos, 2004)	28% increase of NPV compared to traditional production scheduling in a gold mine.
	(Leite and Dimitrakopoulos, 2007)	26% higher NPV than deterministic methodology in a copper deposit.
	(Albor Consquega and Dimitrakopoulos, 2009)	Showed 25% higher NPV than the conventional method by using conventionally optimized pit limits and an additional increase of 10% by using stochastically optimized pit limits.
Stochastic integer programming	(Ramazan and Dimitrakopoulos, 2007)	When compared to conventional methodology, up to a 25% NPV increase is reported.
	(Dimitrakopoulos and Ramazan, 2008)	In comparison with a conventional approach, 10% higher NPV in a gold mine and 25% higher NPV in a copper deposit.
	(Ramazan and Dimitrakopoulos, 2013)	Demonstrated a 10% increase of NPV compared to the deterministic method in an Australian gold deposit.
	(Maleki et al., 2020)	Comparison with deterministic integer programming showed a 6% increase in cumulative cash flow when applied to an iron deposit.

Integration of multiple components of mining complex into a single model while using two-stage SIP can also help to meet production targets within the mineral value chain (Goodfellow and Dimitrakopoulos, 2017). Another source of uncertainty that affects mine planning is commodity price. One way to account for commodity price uncertainty is to use two-stage SIP combined with stochastic dynamic programming that identifies the best policy for produced subsets of blocks depending on commodity price evolution (Rimélé et al., 2020). However, some of these methods are only applicable to smaller projects. One scalable alternative capable of integrating a grade uncertainty is the two-stage stochastic approach, in which geostatistical scenarios are used during the second stage to assign destinations of each mining block (Moreno et al., 2017). Although many of the recent SIP methodologies outperform this method in terms of maximizing NPV, it is an efficient method capable of handling large projects and providing optimal results with good NPV maximization and risk minimization.

3 METHODOLOGY

This study demonstrates a workflow, which consists of geostatistical cosimulation of coregionalized variables with inequality constraint and adaptive open-pit production scheduling. The first part focuses on hierarchical sequential Gaussian cosimulation integrated with inverse transform sampling technique to model secondary variable according to thresholds derived from inequation. The second part introduces techniques used for two-stage stochastic long-term production scheduling of Iron deposits, which can incorporate uncertainty and inequality constraints from stochastic modeling.

3.1 Cosimulation of variables with an inequality constraint

The idea of using hierarchical cosimulation for the modeling of Iron deposits with inequality constraints between variables was first introduced by Madani and Abulkhair (2020). In that algorithm, integration of an acceptance-rejection technique in the second simulation helped to re-simulate rejected faulty values (i.e., values outside the interval derived from inequation). In this study, an acceptance-rejection technique is replaced by inverse transform sampling that is able to re-simulate faulty values within truncated thresholds. Before describing an algorithm, the basic ideas and methods behind inequality constraint, fitting an inequation and inverse transform sampling are going to be explained in detail.

3.1.1 Inequality constraints

Inequality and equality constraints are common to deal with in multivariate geostatistical modeling. There are three types of this bivariate complexity between primary Y_1 and secondary Y_2 variables. The classification of inequality constraint is based on the strictness of the relationship between variables.

- Strict inequality constraint can be expressed by $Y_2 > Y_1$ or $Y_2 < Y_1$. In other words, it is a type of inequality constraint when the secondary variable is either greater than or less than the primary variable throughout the whole deposit.

- Similar to strict inequality constraint, the following expressions can show a non-strict type: $Y_2 \geq Y_1$ or $Y_2 \leq Y_1$. This time the bivariate relationship is not restricted so that the secondary variable can be greater than, less than or equal to the primary variable.

- In this study, the primary focus is on the third type: inequality constraint expressed by linear inequation. Similar to the first two types above, it can be either non-strict, $Y_2 \geq aY_1 + b$ or $Y_2 \leq aY_1 + b$, or strict, $Y_2 > aY_1 + b$ or $Y_2 < aY_1 + b$. In these inequations, a is a slope and b is an intercept of a linear function.

3.1.2 Inverse transform sampling

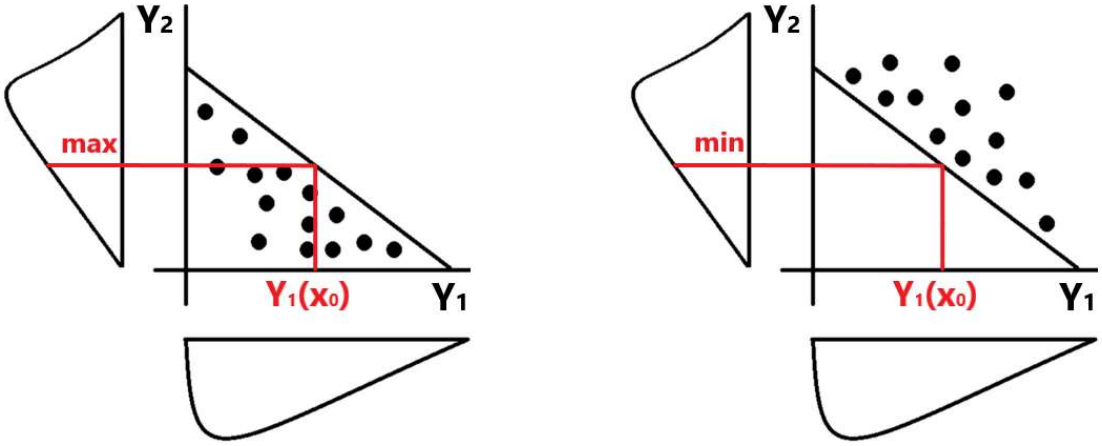
The significant problem with conventional cosimulation algorithms is that they are unable to reproduce inequality constraints. As a result, some values are simulated outside the linear inequation. An inverse transform sampling technique is integrated into the proposed hierarchical cosimulation algorithm to prevent this problem. This method is based on random number generation: computing inverse cumulative distribution function (cdf) of the random numbers generated uniformly between 0 and 1 (Devroye, 1986). Moreover, this method can be used for generating random numbers within truncated thresholds (Burkardt, 2014).

An interval of two truncated thresholds $[min, max]$ representing the minimum and maximum possible values can be used in inverse transform sampling, so that $min \leq Y \leq max$. The following expression represents inverse transform sampling within truncated thresholds (Devroye, 1986):

$$V = F^{-1}\big(F(min) + \big(F(\max) - F(min)\big) \cdot U\big) \tag{4}$$

where, V is random inverse transform, F is the conditioned cdf, F^{-1} is quantile function, and U is an independent random value distributed uniformly between 0 and 1.

In the case of inequality constraint, $[min, max]$ interval of the secondary variable is obtained based on identified linear inequation for each value of the primary variable $Y_1(x_0)$. Figure (6a) demonstrates an example of negative inequation $Y_2 \leq aY_1 + b$, for which $min = \min(Y_2)$ and $max = \mathrm{a}Y_1(x_0) + b$. On the contrary, $min = \mathrm{a}Y_1(x_0) + b$ and $max = \max(Y_2)$ in the case of positive inequation $Y_2 \geq aY_1 + b$ (Figure 6b). The procedure of the hierarchical cosimulation algorithm integrated with this technique is explained later in section 3.1.4.



(a) $Y_2 \leq aY_1 + b$ (b) $Y_2 \geq aY_1 + b$

Figure 5. Obtaining minimum and maximum thresholds $[min, max]$ of the secondary variable conditional to the simulated value of the primary variable $Y_1(x_0)$ in the case of negative (a) and positive (b) inequations

3.1.3 Cokriging configurations

Another advantage of the proposed methodology is that it can incorporate different moving neighborhood configurations (i.e., searching strategies). Cosimulation in a hierarchical way by modeling the primary variable first and then modeling the secondary variable conditional to the first simulation makes this algorithm considerably slower than traditional cosimulation techniques. The proposed methodology also has additional inverse transform sampling integrated into the algorithm, which slows down the simulation process. In that regard, a moving neighborhood is a wise choice in comparison with a unique one. Although the proposed algorithm is equipped with single, multiple, collocated and multicollocated neighborhood strategies, this study focuses only on multiple for the first simulation and multicollocated for the second. Both of these strategies produce the closest results in terms of estimation variance to the unique search (Madani and Emery, 2019).

For the general information about searching strategies, refer to section 2.2.1. The schematic representations of unique, single, multiple, collocated and multicollocated strategies are presented in Figure (1). Simple cokriging $Z^*_{1_{SCK}}(x_0)$ and simple multicollocated cokriging $Z^*_{2_{SMCK}}(x_0)$ estimators together with their estimation variances $\sigma^2_{SCK}(x_0)$ and $\sigma^2_{SMCK}(x_0)$ for both variables are expressed as follows:

$$Z_{1_{SCK}}^{*}(x_0) = m_1 + \sum_{\alpha=1}^{n_1} w_\alpha^1 [Z_1(x_{1,\alpha}) - m_1] + \sum_{\beta=1}^{n_2} w_\beta^2 [Z_2(x_{2,\beta}) - m_2] \quad (5)$$

$$Z_{2_{SMCK}}^{*}(x_0) = m_2 + \sum_{\alpha=1}^{n_1} w_\alpha^{1'} [Z_1(x_{1,\alpha}) - m_1] + \sum_{\beta=1}^{n_1} w_\beta^{2'} [Z_2(x_{2,\beta}) - m_2] \quad (6)$$

$$\sigma_{SCK}^2(x_0) = C_{11}(x_0 - x_0) - \sum_{\alpha=1}^{n1} w_\alpha^1 C_{11}(x_{1,\alpha} - x_0) - \sum_{\beta=1}^{n2} w_\beta^2 C_{21}(x_{2,\beta} - x_0) \quad (7)$$

$$\sigma_{SMCK}^2(x_0) = C_{22}(x_0 - x_0) - \sum_{\alpha=1}^{n1} w_\alpha^{1'} C_{12}(x_{1,\alpha} - x_0) - \sum_{\beta=1}^{n1} w_\beta^{2'} C_{22}(x_{2,\beta} - x_0) \quad (8)$$

where x_0 is the target location; $x_{1,\alpha}$ $(\alpha = 1, \dots n_1)$ and $x_{2,\beta}$ $(\beta = 1, \dots n_2)$ are the data locations of primary Z_1 and secondary Z variables, respectively; w_α^1 and w_α^2 are the weights assigned to Z_1 and Z_2 at the data locations α and β; m_1 and m_2 are the mean values of both variables; C_{11} is the direct covariance of Z_1, C_{22} is the direct covariance of Z_2, C_{21} and C_{12} are the cross covariances between Z_1 and Z_2.

3.1.4 Proposed hierarchical cosimulation algorithm

The proposed hierarchical cosimulation algorithm is an updated version of conventional hierarchical sequential Gaussian cosimulation (Almeida and Journel, 1994), which is integrated with inverse transform sampling for modeling variables with inequality constraints and different configurations of moving neighborhood. Like the conventional methodology, variables are simulated in a hierarchical order: the primary variable is simulated in the first stage, followed by the simulation of the secondary variable conditional to the primary variable's simulated values. The proposed algorithm proceeds as follows:

1. Transform the primary Y_1 and secondary Y_2 variables into their normal scores Z_1 and Z_2 after defining the hierarchical order of variables. Transformation of each variable must proceed independently and the primary variable must have the best autocorrelation.

2. Define the simulation path, which can be either regular or random. However, the use of random sequence is considered as a better option (Isaaks, 1990). Another important detail is that each grid node x_0 must be visited only once.

3. First simulation: determine a Gaussian conditional cdf (ccdf) in order to obtain global statistical parameters of the primary variable Z_1 at each target node x_0. In this step, a simple cokriging estimator is used to infer mean and variance parameters and proceed to the simulation of n realizations of the primary variable $Z_1^n(x_0)$:

$$Z_1^n(x_0) = Z_{1\,SCK}^*(x_0) + \sqrt{\sigma_{1\,SCK}^2(x_0)} \cdot U^n \tag{9}$$

where U^n is an independent random value uniformly distributed between 0 and 1.

This step should be applied to all target grid nodes for simulated values of the primary variable to be available in the entire region.

4. Second simulation: at each node x_0, obtain global statistical parameters of the secondary variable from the Gaussian ccdf using simple multicollocated cokriging. In this step, the estimator considers the hard data of both variables and the simulated values of the primary variable. The simulation of n realizations of the secondary variable $Z_2^n(x_0)$ proceeds as follows:

$$Z_2^n(x_0) = Z_{2\,SMCK}^*(x_0) + \sqrt{\sigma_{2\,SMCK}^2(x_0)} \cdot U^n \tag{10}$$

Loop until all target grid nodes are simulated.

The inverse transform sampling technique presented in section 3.1.2 is implemented in this step to re-simulate faulty values that do not respect the existing inequality constraint. The following steps demonstrate the procedure in detail:

i. Back-transform the simulated values $Z_1^n(x_0)$ to the original scale of the primary variable $Y_1^n(x_0)$.

ii. Determine the minimum and maximum thresholds $[min, max]$ of the secondary variable for all simulated values of the primary variable $Y_1^n(x_0)$ using the method explained in section 3.1.2 (see Figure 6).

iii. Transform thresholds obtained in step ii into normal scores and identify faulty values of the secondary variable $Z_2^m(x_0)$ that lie outside the interval.

iv. Re-simulate the faulty values of the secondary variable $Z_2^m(x_0)$ using random numbers V^m generated within truncated thresholds $[min, max]$ for m realizations:

$$Z_2^m(x_0) = Z_{2_{SMCK}}^*(x_0) + \sqrt{\sigma_{2_{SMCK}}^2(x_0)} \cdot V^m \tag{11}$$

v. Loop until all nodes with faulty values are re-simulated.

5. Back-transform realizations of the primary Z_1^n and secondary Z_2^n variables to their original scales Y_1^n and Y_2^n, respectively.

Note that either simple cokriging or simple collocated cokriging can be used instead of simple multicollocated cokriging in the second simulation. Moving neighborhood configurations required for these types of cokriging are explained in section 2.2.1 (see Figure 1).

3.1.5 Program code

The proposed algorithm is implemented by using SGCo_IC MATLAB code. This code allows using the following geostatistical algorithms: traditional sequential Gaussian cosimulation, conventional hierarchical cosimulation and proposed algorithm for variables with inequality constraints. SGCo_IC program code is integrated with single and multiple moving neighborhood strategies for the first simulation; single, multiple, collocated and multicollocated strategies for the second simulation; and only applicable for bivariate cases. This program uses 8 subroutines:

- BACKTR: Back transformation from normal scores to original scale (Emery, 2008).
- COVA: compute covariance values (Emery, 2008).
- CREATE_PARAMFILE_INPUT: create default parameter file SGCOSIM.PAR.
- IMPORTFILE: read data from GSLIB file.
- EXPORTFILE: print output into GSLIB file.
- SEARCHDATA: search for data according to moving neighborhood (Emery, 2008).
- SETROT: set up a matrix for rotation and reduction of coordinates (Emery, 2008).

- WEIGHTCAL: calculation of weights.

SGCo_IC has the following features:

- An option to model a single plane (number of planes equals the number of Z nodes).
- Ability to handle heterotopic datasets by defining trimming limits for Gaussian data. Data value is considered missing if it is a non-number (NaN) or number outside trimming limits.
- An option to back transform simulated values to the original scale. Variogram reproduction can be checked on untransformed values.
- The number of nested structures in the linear model of coregionalization is not restricted. Subroutine COVA uses the following structure types: spherical, exponential, cubic and Gaussian. Geometric anisotropy is defined by three ranges along rotated x, y and z axes (SETROT subroutine); and three angles (azimuth, dip and plunge) according to Deutsch and Journel (1998).
- An option to apply one of three cosimulation algorithms: 1) conventional sequential Gaussian; 2) hierarchical sequential Gaussian; and 3) proposed algorithm for variables with inequality constraints. Latter is implemented by defining coefficients of linear restriction: slope and intercept.
- Only moving neighborhood is integrated and defined by an ellipsoid with three values for maximum radii, three rotation angles and the maximum number of data points (SEARCHDATA subroutine). The algorithm searches for original data and simulated nodes separately.
- Either single or multiple moving neighborhood strategy can be used in the first simulation. For the second simulation, there are four options: single, multiple, collocated and multicollocated strategies. In the case of hierarchical cosimulation, neighborhood parameters for each simulation run must be defined.

Table (2) demonstrates a default ASCII parameter file SGCo_IC.PAR generated using CREATE_PARAMFILE_INPUT subroutine. Alternatively, input parameters can be used directly in the MATLAB environment as workspace variables.

Table 2. Default parameters file for SGCo_IC

```
                 Parameters for SGCo_IC
                  **********************

START OF PARAMETERS:
Grid.out          % target location
1 2 3             % columns for location coordinates
3 33 113 37       % index of plane (0=all the planes), number of nodes
nscore.out        % file with conditioning data
1 2 3             % columns for data coordinates
7 8               % columns for Gaussian data values
Fe Al2O3          % original variable names
-10.0 10.0        % trimming limits for Gaussian data
nscoreFe.trn      % variable 1: file with conversion table (raw-Gaussian)
25.1 69.17        %             minimum and maximum values for raw variable
1.0 1.0           %             parameters for lower-tail and upper-tail extrapolation
nscoreAl2O3.trn   % variable 2: file with conversion table (raw-Gaussian)
0.1 32.1          %             minimum and maximum values for raw variable
1.0 1.0           %             parameters for lower-tail and upper-tail extrapolation
0                 % Back-transformation? (0=yes, 1=no)
vargfit           % basename for files with variogram models
2000 2000 2000    % original data: maximum search radii in the rotated system
0 0 0             % angles for search ellipsoid
0                 % divide into octants? 1=yes, 0=no
80                % number of data per octant (if octant=1) or in total
2000 2000 2000    % simulated nodes: maximum search radii in the rotated system
0 0 0             % angles for search ellipsoid
0                 % divide into octants? 1=yes, 0=no
80                % number of data per octant (if octant=1) or in total
1                 % co-simulation: 0=Conventional, 1=Hierarchical
0                 % if co-simulation=0 --> searching strategy: 0=single; 1=multiple
1                 % if co-simulation=1 --> 1th searching strategy: 0=single; 1=multiple
2                 % if co-simulation=1 --> 2th searching strategy: 0=single; 1=multiple;
                  2=multicollocated; 3=collocated
0                 % if co-simulation=1 --> inequality constraint: 0=yes; 1=no
-0.89 62          % if inequality constraint=0: linear coefficients: slope, intercept
100               % number of realizations
9784498           % seed for random number generation
sgco_ic.out       % name of output file
3                 % number of decimals for values in the output file
```

3.2 Long-term open-pit production scheduling

This study proposes a two-stage stochastic production scheduling approach that aims to integrate geostatistical simulation results into mine planning. The proposed methodology uses average information for the first stage decisions, extraction periods, and multiple equally probable scenarios for the second stage decisions, block destinations. The following sections describe the fundamental methodologies and formulations involved in this approach, such as economic model calculation, computation of ultimate pit limit and nested pits by pseudoflow algorithm, production scheduling by mixed integer programming and the workflow of the proposed two-stage stochastic production scheduling.

3.2.1 Economic block value

The block's economic value depends on the mining, processing, and selling costs dictated by the chosen mining method. The primary variable is not always the only element that matters in the mine planning procedure because impurity content can also affect the processing cost. Therefore, the following equation of block value v_α considers the penalization for cases when the secondary variable exceeds a certain threshold:

$$v_\alpha = \max\left((P - SC) \cdot R \cdot TON_\alpha \cdot Y_{1_\alpha} - (MC + PC) \cdot TON_\alpha - (Y_{2_\alpha} > Y_{2_{TH}}) \cdot H \cdot (Y_{2_\alpha} - Y_{2_{TH}}) \cdot TON_\alpha, -MC \cdot TON_\alpha\right) \quad (12)$$

where SC, MC and PC are the selling, mining and processing costs (\$/ton), respectively, P is the Iron price (\$/ton), TON_α is the total block tonnage (ton) for each block α, H is the penalization on the exceeding amount of disturbing element (\$/ton), R is the recovery rate, Y_{1_α} and Y_{2_α} are grades of the primary and secondary variables, respectively, and $Y_{2_{TH}}$ is the maximum threshold of the secondary variable.

3.2.2 Pit limits

Ultimate pit limit and nested pits are generated using a pseudoflow algorithm (Hochbaum and Chen, 2000; Hochbaum, 2001). Background on the working principle of UPL identification is discussed in section 2.3.1. To implement a pseudoflow nested pit solver, revenue factors are applied to the metal price to find different nested pits that will be used as an input in production scheduling. The value formula is changed to have different values for a different pit in the following way:

$$v_{\alpha_\gamma} = \max\left(\left(P \cdot rf_\gamma - SC\right) \cdot R \cdot TON_\alpha \cdot Y_{1_\alpha} - (MC + PC) \cdot TON_\alpha - (Y_{2_\alpha} > Y_{2_{TH}}) \cdot H \cdot (Y_{2_\alpha} - Y_{2_{TH}}) \cdot TON_\alpha, -MC \cdot TON_\alpha\right) \quad (13)$$

$$rf_\gamma = rf_0 + \gamma / N_{pits} \qquad \forall \gamma \in N_{pits} \quad (14)$$

where rf_γ is a revenue factor applied to each pit, $\gamma \in N_{pits}$ is a set of different revenue factors, and rf_0 is starting revenue factor (usually $rf_0 = 0$).

3.2.3 Mixed integer programming

The mixed integer programming (MIP) algorithm is used in this study to generate a production plan by sequentially extracting nested pits. The objective is to find an optimal solution by maximizing a net present value (NPV) to specify the extraction sequence (Morales et al., 2019). The following equation characterizes the objective function of the production scheduling:

$$max \sum_{\rho \in P} \sum_{\alpha \in A} \frac{v_\alpha^\rho}{(1+r)^\rho} S_\alpha^\rho \tag{15}$$

where S_α^ρ is an integer variable that indicates if a block is extracted ($S_\alpha^\rho = 1$) in period ρ or not ($S_\alpha^\rho = 0$), r is a discount rate, and $\rho \in P$ is a set of periods.

The next step is to apply several constraints on the production scheduling problem:

1. Grade blending constraints: the average grade of the block should be greater than the desired minimum grade:

$$\sum_{\alpha \in A} (g_\alpha - g_{min}) PROCTON_\alpha S_\alpha^\rho \geq 0 \tag{16}$$

   where g_α is an average grade of block α, g_{min} are the minimum acceptable average grade, $PROCTON_\alpha$ is a tonnage of ore material that is going to be sent into the processing plant.

2. Mining and processing capacity constraints: total extracted material and ore tonnage should be lower than minimum available mining and processing capacities:

$$\sum_{\alpha \in A} TON_\alpha S_\alpha^\rho \leq MC_{max} \tag{17}$$

$$\sum_{\alpha \in A} PROCTON_\alpha S_\alpha^\rho \leq PC_{max} \tag{18}$$

   where MC_{max} and PC_{max} are maximum thresholds for mining and processing capacities.

3. Reserve constraints: all the blocks within the UPL should be mined once.

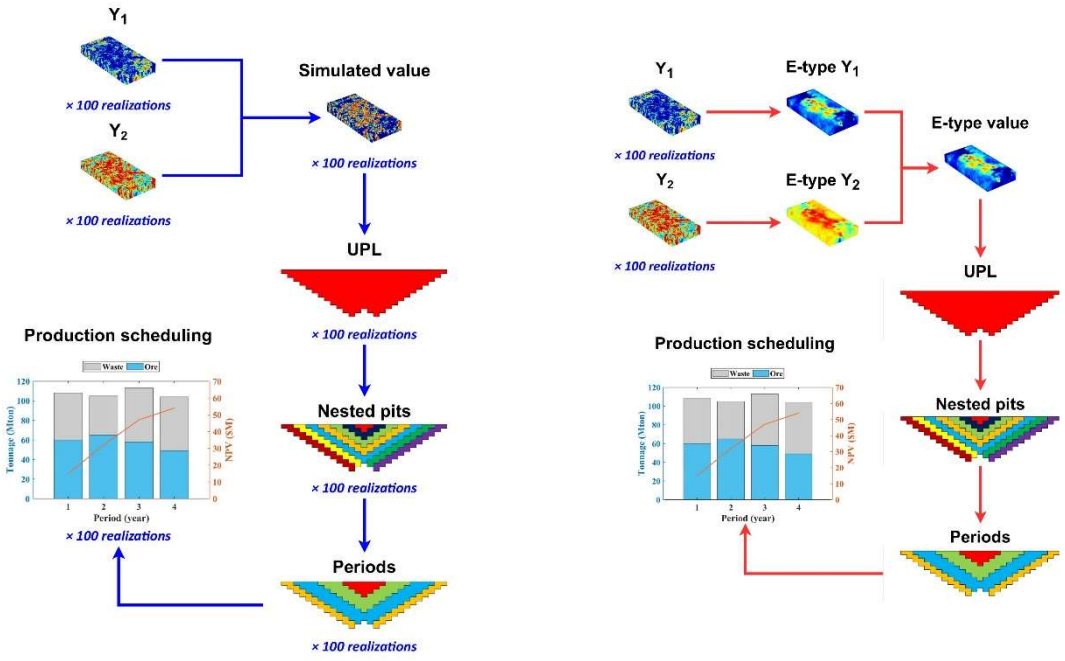
$$\sum_{\alpha \in A} S_{\alpha}^{\rho} = 1 \tag{19}$$

For the more detailed MIP workflow, the readers are referred to Ramazan and Dimitrakopoulos (2004) and Askari-Nasab et al. (2011).

3.2.4 Two-stage stochastic production scheduling approach

Similar to Moreno et al. (2017), this study proposes a two-stage production scheduling approach. First stage decisions correspond to extraction periods, while block destinations (i.e., waste dump or processing plant) are assigned in the second stage. The proposed method is compared to upper and lower bounds defined as follows:

- Upper bound: Both 1st stage, production periods, and 2nd stage decisions, block destinations, are coming from realizations (Figure 6a). The traditional deterministic methodology is applied to all geostatistical realizations, which results in identifying the highest possible NPV for each scenario. However, this method is not applicable in actual practice because it produces different scheduling and optimal pit limits depending on the actual scenario.



(a) Upper bound (b) Lower bound

Figure 6. Schematic representation of the upper and lower bounds of production scheduling

- Lower bound: Both 1st stage and 2nd stage decisions are obtained from e-type models, which is an average of all geostatistical realizations (Figure 6b). The main issue with this approach is that it does not consider the risks of not meeting production targets. Notably, if a block is sent to a plant but turns out to be a waste according to the realization, its extraction and processing cost will negatively affect the NPV. Alternatively, valuable blocks can be sent to a waste dump.

In the proposed approach, extraction schedules are fixed and obtained based on e-type models, but block destinations are adapted to each equally probable scenario informed by geostatistical realizations. The grade uncertainty is taken into account in the second stage decisions to minimize potential losses of not meeting production targets. Depending on the available information about actual grades, the plan can be adapted using a set of simulated scenarios. The general workflow is the following:

1. Obtain e-type models for n realizations of the primary Y_1^n and secondary Y_2^n variables.

2. From each of n realizations of geostatistical modeling calculate economic block models. Generate a separate economic block model for e-type results.

3. 1st stage decisions: evaluation of extraction periods

    i. Compute UPL and nested pits from the e-type economic block model.

    ii. Define operational, capacity, grade blending and reserve constraints.

    iii. Select the pushbacks by grouping pits using a deterministic model, so that selected groups have similar tonnages.

    iv. Compute production periods.

    v. Extract and store production periods, tonnages and destinations.

4. 2nd stage decisions: a re-evaluation of block destinations.

    i. Re-evaluate the tonnages and economic values using n realizations of the primary Y_1^n and secondary Y_2^n variables.

ii. Re-evaluate the block destinations according to a set of realizations for economic values and grades of each block.

iii. Store new results for block destinations, economic values, ore and waste capacities by replacing the values obtained during the 1st stage.

Figure (7) shows the general scheme of the proposed production scheduling methodology.

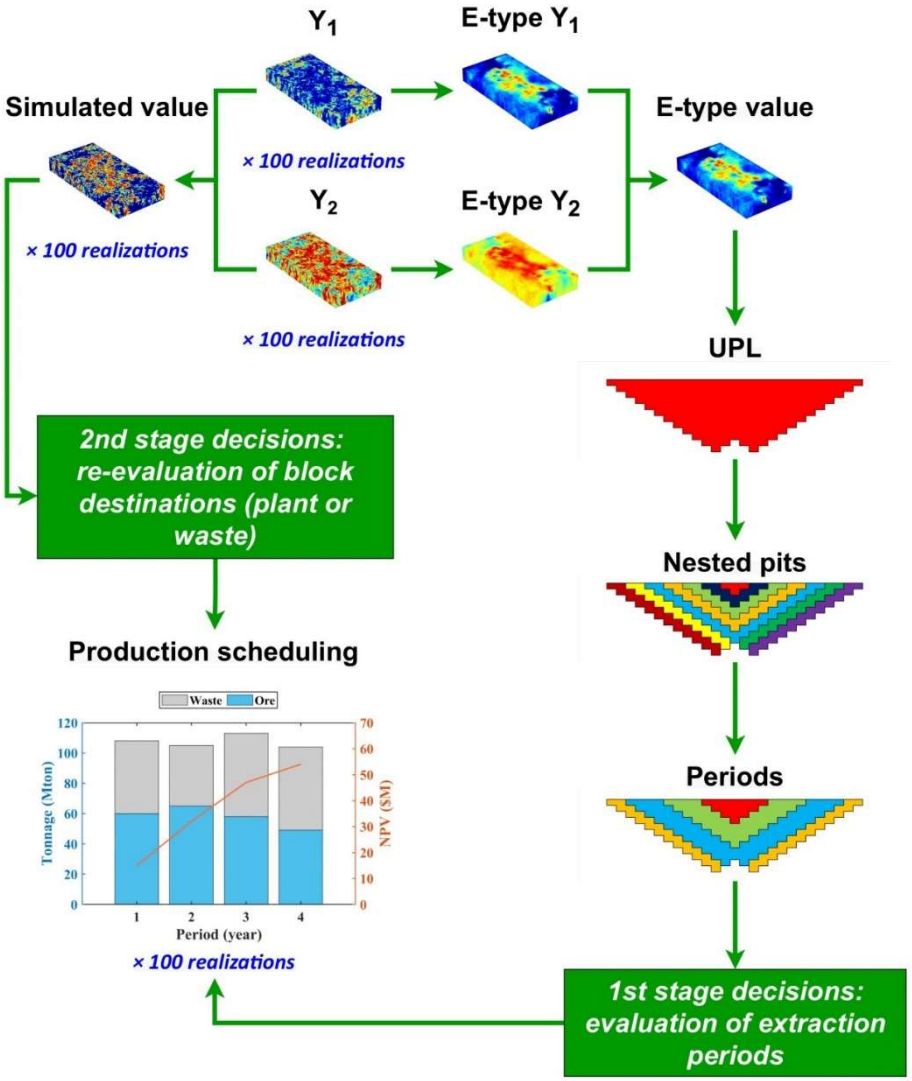


Figure 7. Schematic representation of the proposed production scheduling approach

4 RESULTS

4.1 Case study

Conventional and proposed cosimulation algorithms are applied to a real case study from an Iron deposit. The dataset consists of borehole information coming from an exploration project. To keep the data confidential, coordinates are altered and grades are multiplied by a scale factor. It is a homotopic dataset composed of Iron and Silica grades scattered along 2137 sample points. The center part of the deposit is highly concentrated with Iron and has low Silica content, which means that there is a negative correlation between these two variables (Figure 8). Iron is the primary element in this deposit, while Silica can be considered a disturbing element that might decrease the ore value.

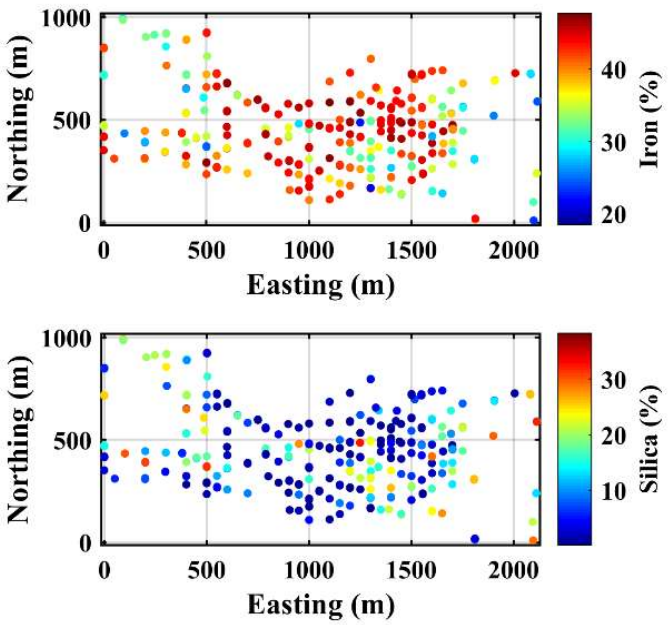


Figure 8. Location map of the Iron dataset showing the spatial distribution of coregionalized variables: Iron and Silica

4.2 Exploratory data analysis

The sampling pattern in this deposit does not appear to be preferential, but it is pretty irregular. The irregularity can be seen in the location map (see Figure 8), where samples are mostly taken from the center part of the deposit. To alleviate this issue, cell-declustering is performed (David, 1977; Deutsch, 1989), reducing the effect of sampling irregularity. A 50 m × 50 m × 10 m cell

dimension is chosen after checking different sizes and their effect on global statistical parameters. Global statistical parameters of the dataset before and after cell-declustering are shown in Table (3). Because weights of the samples located further from the center are higher, the decrease of mean Iron grade and increase of mean Silica grade can be observed.

Table 3. Statistical parameters of Iron dataset before and after declustering

Statistical parameter	Before declustering		After declustering	
	Iron (%)	Silica (%)	Iron (%)	Silica (%)
Mean	37.71	11.71	37.54	11.96
Standard deviation	6.50	9.71	6.55	9.80
Coefficient of variation	0.17	0.83	0.17	0.82
Correlation coefficient	-0.95		-0.95	

The bivariate distribution between Iron and Silica shows a sharp inequality constraint observed in the scatter plot (Figure 9). The marginal distribution of Silica shows that most of the samples have low Silica content, while the histogram of Iron has a more uniform shape. Linear inequation that expresses the inequality constraint is obtained by fitting the linear function to the scatter plot. The methodology of obtaining the minimum support line for inequality constraint was introduced by Madani and Abulkhair (2020): 1) apply Huber loss (Huber, 1964) to fit the line through regression analysis; 2) move the line upwards to find the minimum support line (red line in Figure 9).

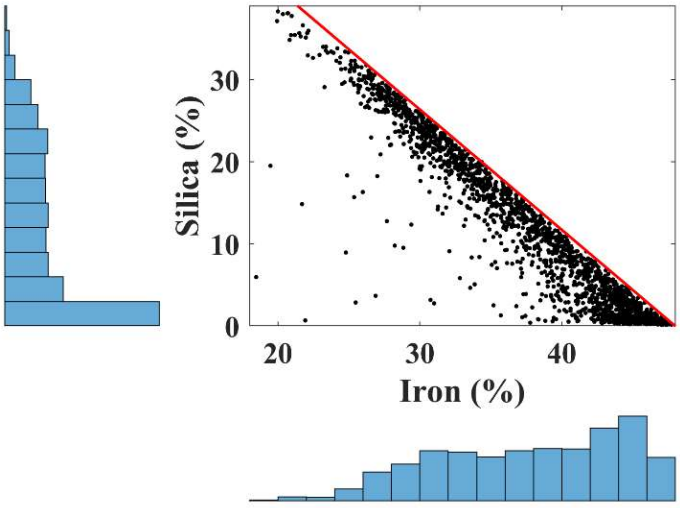


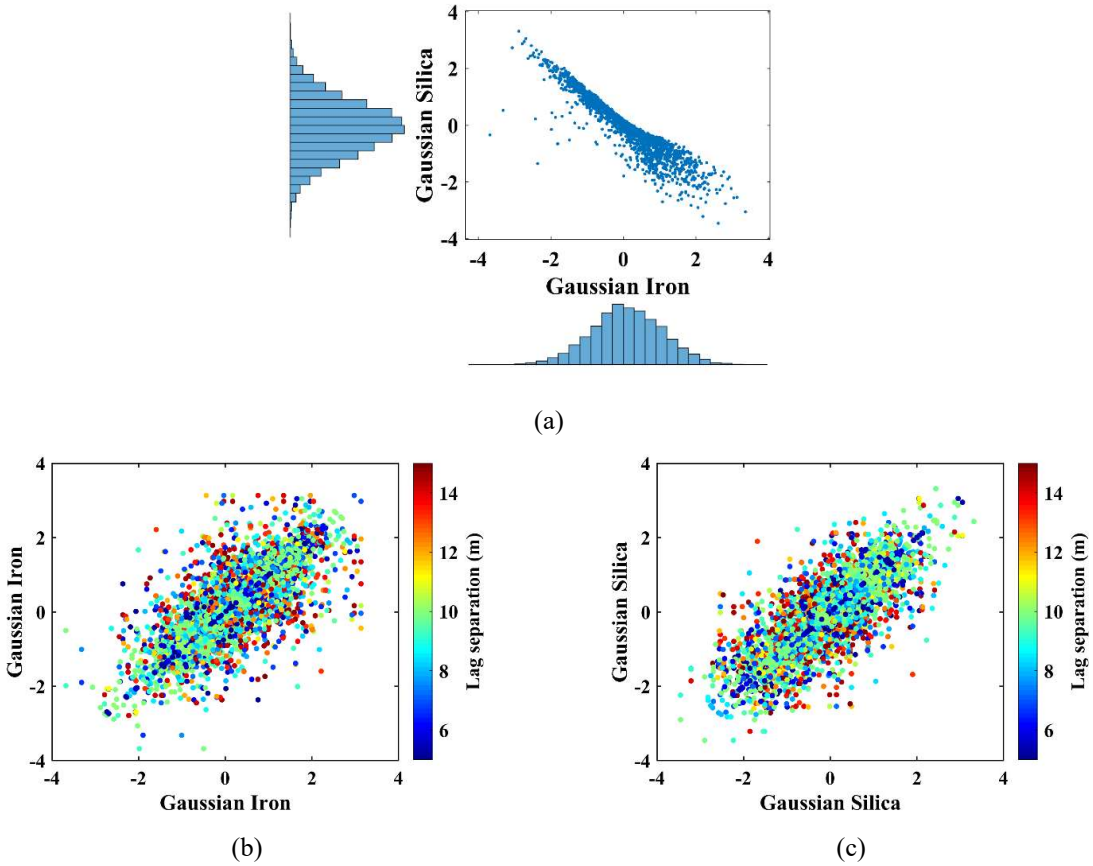
Figure 9. Scatter plot between Iron and Silica with their corresponding marginal distributions. Red line is an inequality constraint between both variables.

In this case, the slope of the line $a = -1.465$ and intercept $b = 70.316$, so linear inequation is expressed as follows:

$$SiO_2 \leq -1.465Fe + 70.316 \quad (20)$$

4.3 Variogram analysis

The hierarchical cosimulation algorithm requires all the variables to be transformed into their normal scores taking into account their declustering weights (Deutsch and Journel, 1998). However, the dataset should follow the multivariate Gaussian assumption before proceeding to the simulation. In order to identify if the dataset is suitable for cosimulation, the multivariate Gaussianity is checked first by investigating the collocated scatter plot between normal scores of Iron and Silica (Figure 10a). The shape of the bivariate distribution between Gaussian Iron and Gaussian Silica is consistent with a correlation coefficient of -0.93.



(a)

(b)

(c)

Figure 10. Scatter plot between normal scores of Iron and Silica with their marginal distributions (a); lagged scatter plots of normal scores of Iron (b) and Silica (c)

A bivariate Gaussianity is checked by producing lagged scatter plots for each normal score, where the elliptical shape is considered suitable (Rivoirard, 1994). To do so, a 10 m lag with 50% lag tolerance is chosen to check this assumption for Iron and Silica. Figure (10b) shows that lagged scatter plot of Gaussian Iron at a lag of 10 m resembles an elliptical shape. Similarly, the shape of the bivariate distribution for Gaussian Silica is close to an ellipse (Figure 10c). Therefore, it can be concluded that both multivariate and bivariate Gaussian assumptions are reasonable, thus hierarchical cosimulation can be implemented in this study.

The next step is to examine the anisotropy by deriving sample variograms in different directions. Through rigorous investigation, no anisotropy is found in the horizontal direction, meaning that the variogram is going to be isotropic horizontally. However, there is a vertical anisotropy, which means that the experimental direct and cross variograms are calculated in two directions: vertical and horizontal. Theoretical variograms are fitted manually to obtain a two-structured linear model of coregionalization that is expressed as follows (Figure 11):

$$\begin{pmatrix} \gamma_{Fe} & \gamma_{Fe/SiO_2} \\ \gamma_{SiO_2/Fe} & \gamma_{SiO_2} \end{pmatrix} = \begin{pmatrix} 0.25 & -0.17 \\ -0.17 & 0.18 \end{pmatrix} Exp(35m, 35m, 35m) + \begin{pmatrix} 0.65 & -0.66 \\ -0.66 & 0.73 \end{pmatrix} Exp(180m, 180m, 120m) \quad (21)$$

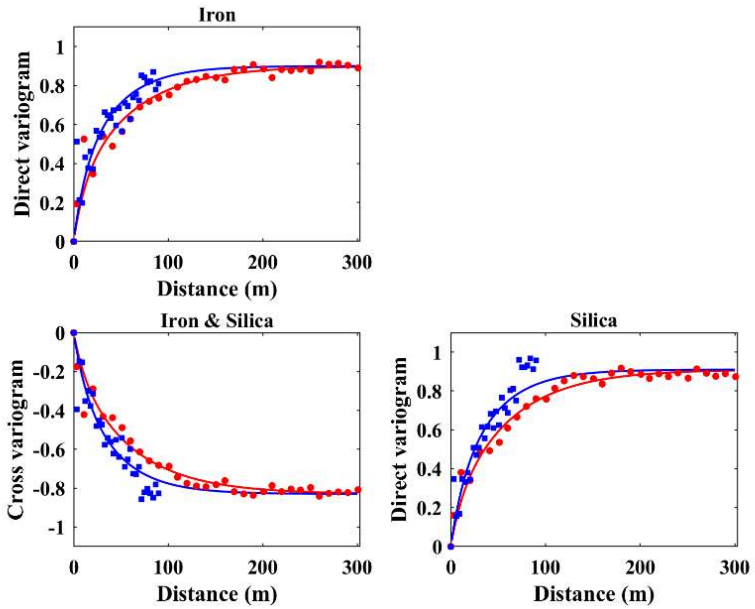


Figure 11. Direct and cross-variograms of the normal scores of Iron and Silica in the vertical (blue) and horizontal (red) directions. Points represent experimental variograms, and lines demonstrate theoretical direct and cross-variograms.

4.4 Conventional and proposed cosimulation results

The modeling of Iron and Silica grades is done by the proposed hierarchical cosimulation algorithm integrated with inverse transform sampling to ensure the reproduction of an inequality constraint. According to Madani and Emery (2019), multicollocated and multiple searching strategies are the closest to unique neighborhood configuration in terms of accuracy and estimation variance. Therefore, simple cokriging with a multiple search is used for the first simulation, while multicollocated cokriging is implemented during the second simulation. Although multicollocated search is proven to be more effective than multiple, it cannot be applied during the first simulation. A regular grid with 20 m × 20 m × 10 m block dimension is selected for cosimulation of 100 realizations. Figure (12) shows 4 out of 100 realizations of Iron and Silica in the same plain. It can be noticed that high Iron content is concentrated in the center part of the deposit, but the southeast and northwest regions have high Silica grades. These results are compatible with the location maps from Figure (8).

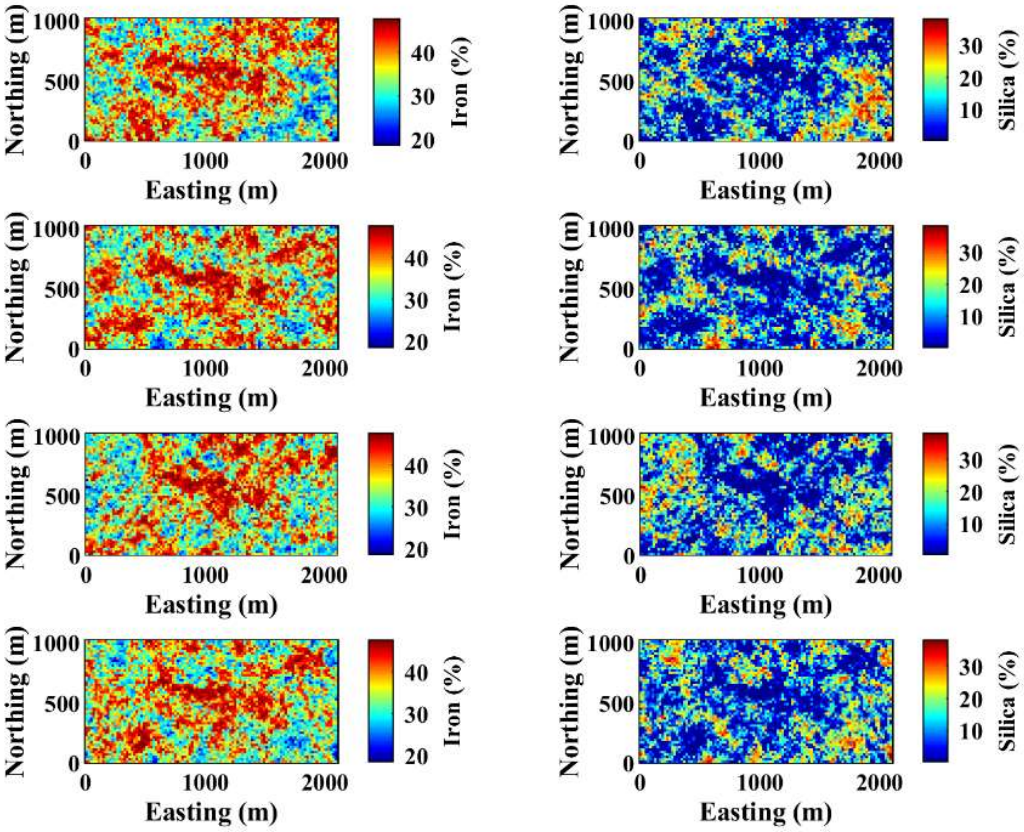


Figure 12. Four random realizations of Iron and Silica produced by the proposed algorithm

To showcase the effectiveness of the proposed algorithm, these realizations are compared to the simulations produced by conventional hierarchical cosimulation with the same parameters applied. Conventional hierarchical cosimulation is implemented by the same SGCo_IC computer program (see Table 2) with the same parameters (i.e., variogram, grid dimensions, conditioning data and neighborhood configurations). Because both algorithms are precisely the same until the simulation of the secondary variable, it is pointless to compare Iron realizations. Thus E-type maps of Silica produced by conventional and proposed cosimulation algorithms are compared to demonstrate the difference between them (Figure 13). Considering that conventional cosimulation algorithms are not able to reproduce inequality constraints, secondary variables are often overestimated. The proposed algorithm can overcome this limitation by re-simulating values that do not lie within the inequation intervals. As a result, the E-type map from the proposed algorithm has slightly lower Silica grades throughout the whole region, except the center part.

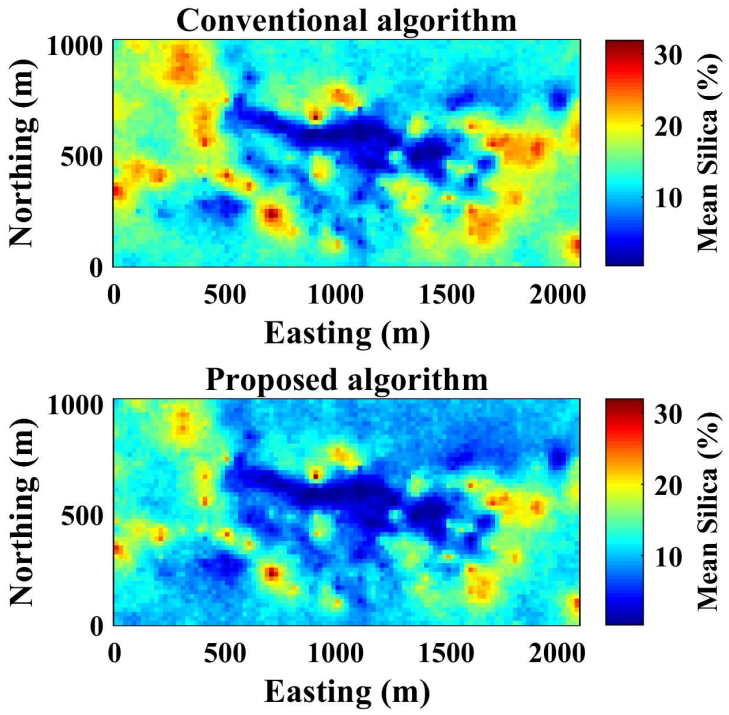


Figure 13. Comparison of E-type maps of Silica obtained from conventional and proposed algorithms

4.5 Comparison of algorithms through geostatistical validation

Examination of geostatistical results usually involves histogram and variogram validation, which are first-order and second-order statistics. In this study, the proposed algorithm must

produce realizations with a similar quality to the conventional approach, but with a perfect reproduction of an inequality constraint between Iron and Silica. Therefore, the validation is demonstrated by comparing both algorithms in terms of reproduction of global statistics, local statistics and bivariate relationships.

4.5.1 Reproduction of global statistical parameters

The first stage of validation is checking the resemblance between mean, standard deviation and correlation coefficient of realizations and original dataset. To do so, the average mean grade, standard deviation, coefficient of variation for both Iron and Silica together with the average correlation coefficient between them are calculated from realizations produced by conventional and proposed cosimulation algorithms. Table (4) compares calculated statistical parameters with declustered statistics of the original dataset. As expected, Iron statistics are the same in both conventional and proposed algorithms, which means that the comparison is fair and unbiased because of the same parameters used in both algorithms. However, the results of Silica are vastly different. Because the integration of inverse transform sampling prevents the overestimation of Silica, the mean grade is lower and closer to the original value. Standard deviation, however, shows that the conventional algorithm is closer to the original. The issue with the reproduction of variance is noticed by Madani and Abulkhair (2020), but the exact reason is still not identified. Finally, the proposed algorithm produces higher correlation coefficient results thanks to the reproduction of an inequality constraint (Figure 14). More details about the reproduction of an inequality constraint between Iron and Silica are discussed in one of the following sections.

Table 4. Reproduction of the average statistical parameters over 100 realizations

Statistical parameter	**Conventional algorithm**		**Proposed algorithm**		**Original dataset**	
	Iron (%)	**Silica (%)**	**Iron (%)**	**Silica (%)**	**Iron (%)**	**Silica (%)**
Mean	36.50	13.40	36.50	11.06	37.54	11.96
Standard deviation	6.67	10.04	6.67	9.48	6.55	9.80
Coefficient of variation	0.18	0.75	0.18	0.86	0.17	0.82
Correlation coefficient	-0.91		-0.93		-0.95	

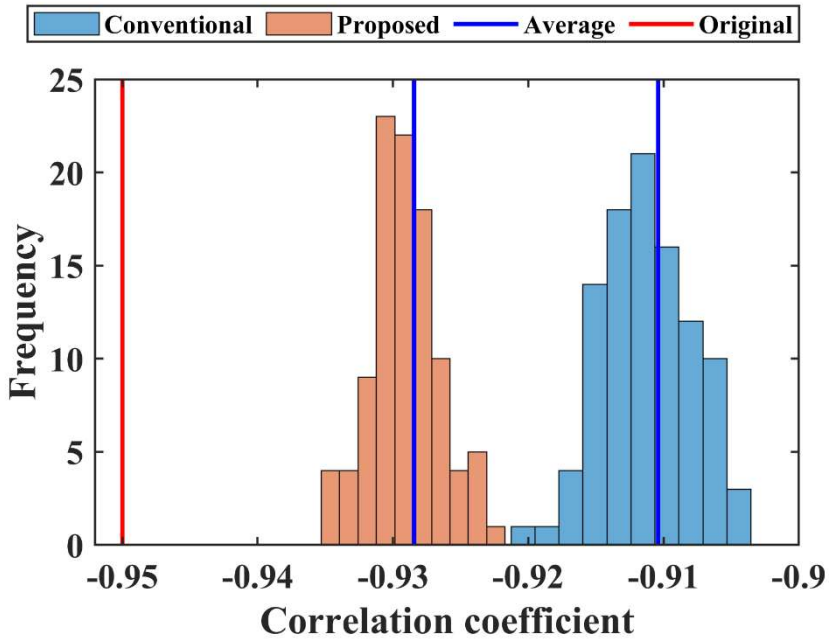


Figure 14. Histogram of correlation coefficients between Iron and Silica from realizations produced by conventional and proposed cosimulation algorithms. The black line is the original correlation coefficient after declustering, and the blue lines show average correlation coefficients over realizations.

One of the concerns with the proposed algorithm is its ability to reproduce the marginal distribution of the secondary variable. The reason for that is the re-simulation of values that do not respect an inequality constraint. A histogram plot of one realization for each variable is a way to examine the reproduction of the marginal distributions by comparing them to the original histograms. Similar to the average statistics, Iron is precisely the same from both algorithms, which further proves that the comparison is valid (Figure 15). The histogram's shape computed from one Iron realization is similar to the original marginal distribution, showing the effectiveness of cosimulation methodologies.

In terms of Silica histograms, results are also very compatible with the original marginal distributions (Figure 16). The histogram's shape from the proposed algorithm resembles the original plot slightly more, demonstrating that the proposed method did not overestimate Silica grades. However, the comparison between algorithms through histogram validation of one realization is not enough. Histogram validation is required to show that the proposed algorithm is able to reproduce the marginal distribution of the secondary variable, despite the re-simulation using inverse transform sampling.

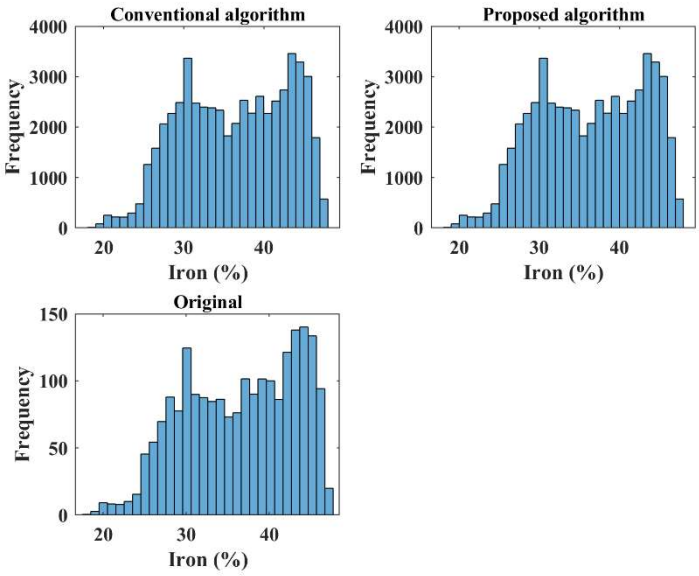


Figure 15. Comparison of Iron histogram reproduction by conventional and proposed algorithms for realization #1

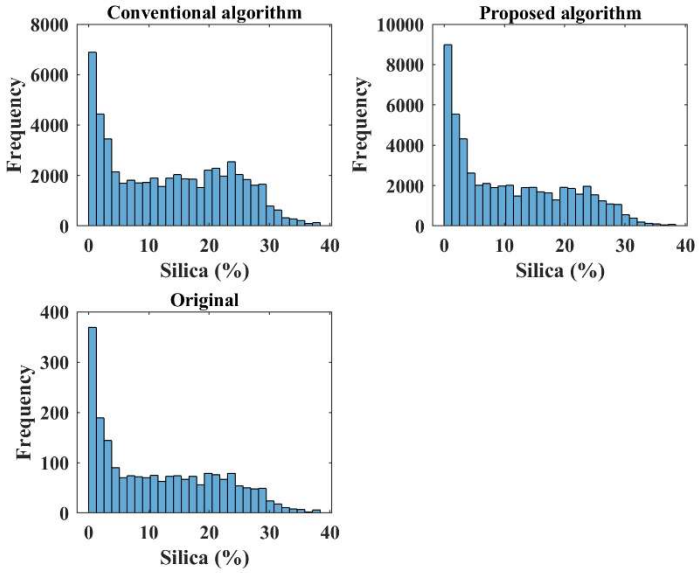


Figure 16. Comparison of Silica histogram reproduction by conventional and proposed algorithms for realization #1

4.5.2 Reproduction of local statistical parameters

In geostatistics, local statistical parameters are sometimes more important than global. Variogram validation is a popular method of checking the reproduction of local statistics. To do so, variograms should be calculated over the block model in the directions used during

variogram analysis. The level of resemblance between original variograms and variograms over simulations signifies the reproduction quality of local statistics.

Figure (17) shows the direct and cross-correlation spatial structures of the realizations from the conventional algorithm and compares them to the experimental direct and cross-variograms used for variogram analysis before simulation. The results are compatible with the original variograms. However, the vertical variogram of Iron (blue color) is slightly worse than the horizontal variogram (red color). Considering that the deposit is shallow, the reproduction of smaller ranges is more crucial. The conventional cosimulation algorithm demonstrates acceptable quality in that regard. Direct variograms of Silica and cross-variograms between both variables over 100 realizations fit with the experimental points in terms of range and sill, proving the effectiveness of the conventional cosimulation.

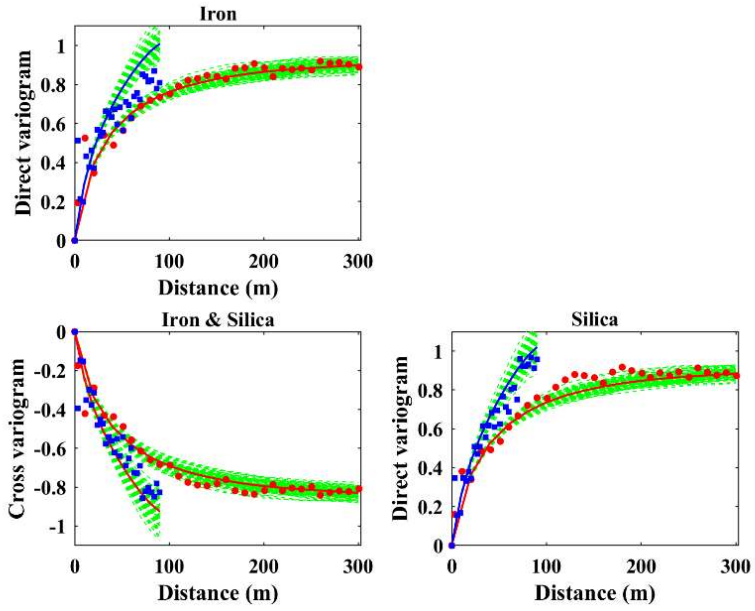


Figure 17. Direct and cross variogram reproduction of Iron and Silica obtained by the conventional cosimulation algorithms. Green: variogram of each realization; red: average horizontal variogram over 100 realizations; blue: average vertical variogram over 100 realizations; red and blue points: horizontal and vertical experimental variograms, respectively.

Realizations from the proposed algorithm are analyzed in the same way (Figure 18). As stated earlier, realizations of Iron are entirely the same in both algorithms, thus direct variograms of Iron do not need to be compared. The comparison between conventional and proposed methodologies can be conducted through visual analysis of the direct variograms of Silica and

cross-variogram between both variables. The simulated results are compatible with the experimental points in terms of direct and cross-variograms. The difference between both algorithms is tiny and can be noticed only by computing the average plot over 100 realizations. For example, the cross-variogram structure is better reproduced by the proposed cosimulation algorithm, particularly in the vertical direction (blue color). Nevertheless, in the conventional algorithm, the average direct variogram of Silica in the horizontal direction is slightly closer to the experimental points. Both methods demonstrate how effective they are in reproducing global and local statistics, and the difference between them is not significant. Therefore, it can be concluded that the integration of inverse transform sampling for modeling variables with inequality constraints does not have a severe effect on the simulation quality.

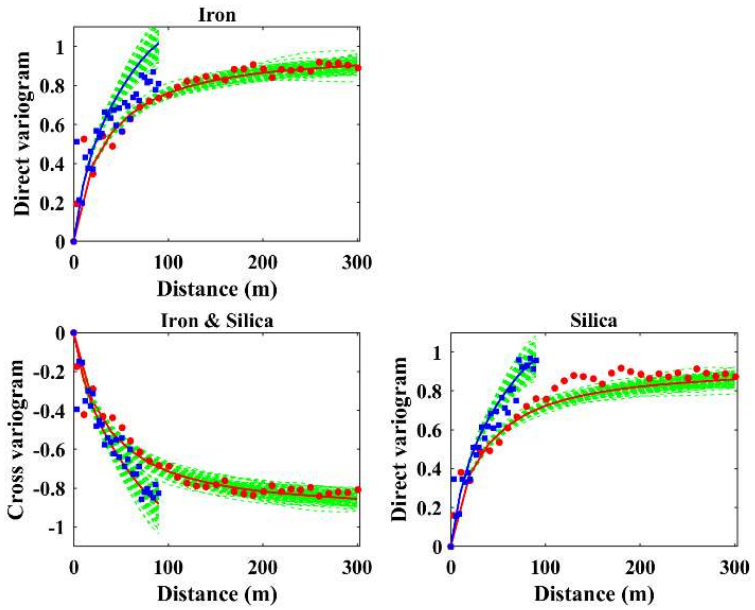
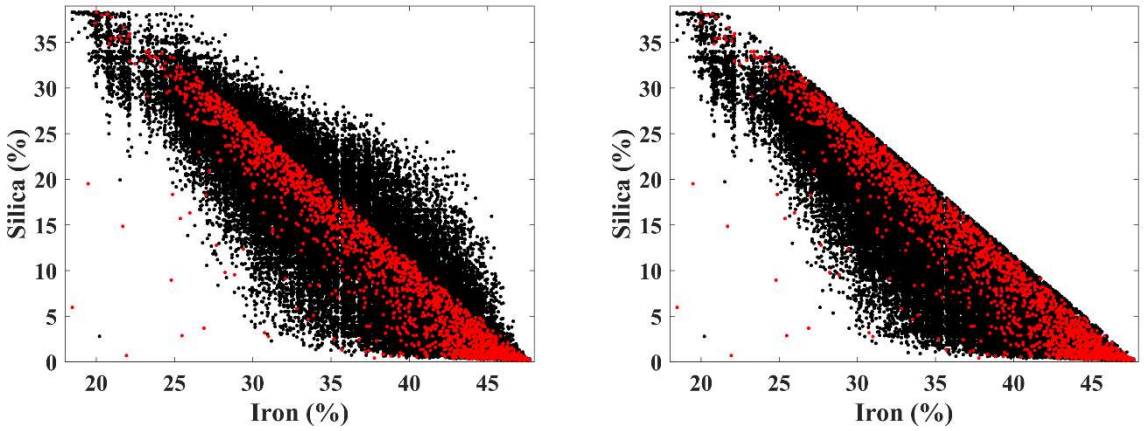


Figure 18. Direct and cross variogram reproduction of Iron and Silica obtained by the proposed cosimulation algorithms. Green: variogram of each realization; red: average horizontal variogram over 100 realizations; blue: average vertical variogram over 100 realizations; red and blue points: horizontal and vertical experimental variograms, respectively.

Overall, the simulation results by both methodologies are compatible with the original dataset. The accuracy of simulations can be improved by increasing the number of conditioning data, which in turn decreases the simulation speed. Therefore, the size of neighborhood ellipsoid and conditioning data must be selected accordingly to find the balance between computational time and simulation quality.

4.5.3 Reproduction of a bivariate relationship

This study aims to present a methodology for modeling variables with an inequality constraint in their bivariate relationship. It can be crucial for Iron projects, where the modeling of disturbing elements is as vital as Iron grade itself. Although the average correlation coefficient from the conventional cosimulation algorithm is high, the scatter plot of individual realization shows the vast difference with the proposed algorithm (Figure 19). The conventional method shows poor reproduction of the bivariate relationship between Iron and Silica (Figure 19a). Some of the simulated values (black) went above the sharp inequality constraint of original values (red). Conventional cosimulation methodologies consider only the cross-correlation structure between Gaussian random fields to reproduce the bivariate relationship. Depending on the marginal distributions and correlation coefficient, the reproduction of bivariate relationships can have even lesser quality. In this case, the correlation coefficient is high and points are well distributed, thus the simulated correlation coefficient is not affected. In the proposed algorithm, however, there is an additional restriction applied by inverse transform sampling, which re-simulates all points that went above an inequality constraint. As a result, the proposed cosimulation approach perfectly reproduces an inequality constraint, which can be seen in Figure (19b), where scatter plots between simulated (black) and original (red) values are shown together.



(a) Conventional algorithm (b) Proposed algorithm

Figure 19. Reproduction of the bivariate relationship between Iron and Silica with conventional (a) and proposed (b) algorithms in one realization. Black points: realization no. 1; and red points: original data.

4.6 Mine planning

4.6.1 Parameters for mine planning

Table (5) shows the set of parameters used for mine planning. The objective of this study is to demonstrate the performance of the proposed methodology. Therefore, prices and costs are not indicated according to this particular deposit but rather based on popular practices from iron projects. The slope angle for computing slope precedents is selected to be 50 ˚, which is later used to create arcs for the pseudoflow algorithm. Furthermore, thresholds for Iron and Silica are conceptual, because the aim is to assess the performance of the proposed production scheduling and compare it to the upper and lower bounds using the same parameters.

Table 5. Set of parameters for long-term open-pit production scheduling

Parameter	**Value**	**Parameter**	**Value**
Price PC (\$/ton)	80	Recovery rate R	0.8
Mining cost MC (\$/ton)	3	Iron threshold Fe_{TH} (%)	30
Processing cost PC (\$/ton)	10.5	Silica threshold $SiO_{2_{TH}}$ (%)	5
Selling cost SC (\$/ton)	30	Mining capacity Cap_M (ton/day)	153500
Penalization H (\$/ton)	1.5	Processing capacity Cap_P (ton/day)	130000

The set of parameters and realizations produced by the proposed cosimulation algorithms were used to obtain economic block models that are later used for computing optimal pit limits and production schedules. In this study, a two-stage stochastic production scheduling approach is used to obtain the feasible solution, which is an intermediate plan. This intermediate solution is compared to the upper and lower bounds to analyze the performance of the proposed production scheduling approach. All three methods and the notion of decision stages are discussed in section 3.2.4. Furthermore, the selection of pushbacks for the intermediate, upper and lower solutions are based on the deterministic e-type model. As a result, three pushbacks are selected using deterministic results by grouping pits and ensuring that each group has a similar tonnage. This is not ideal for stochastic mine planning, but it helps to ensure that all three solutions are produced based on the same number of pushbacks.

4.6.2 Ultimate pit limit

Pseudoflow algorithm is used to find 100 realizations of UPL for each geostatistical realization, which are later used for the upper bound. Figure (20) demonstrates the cross-sectional map that

shows the probability of UPL for each block. It can be observed that deviations are considerably higher away from the center, while the central part shows an almost 100% probability of being extracted. The average ore tonnage is 476 Mton and waste tonnage is 49 Mton.

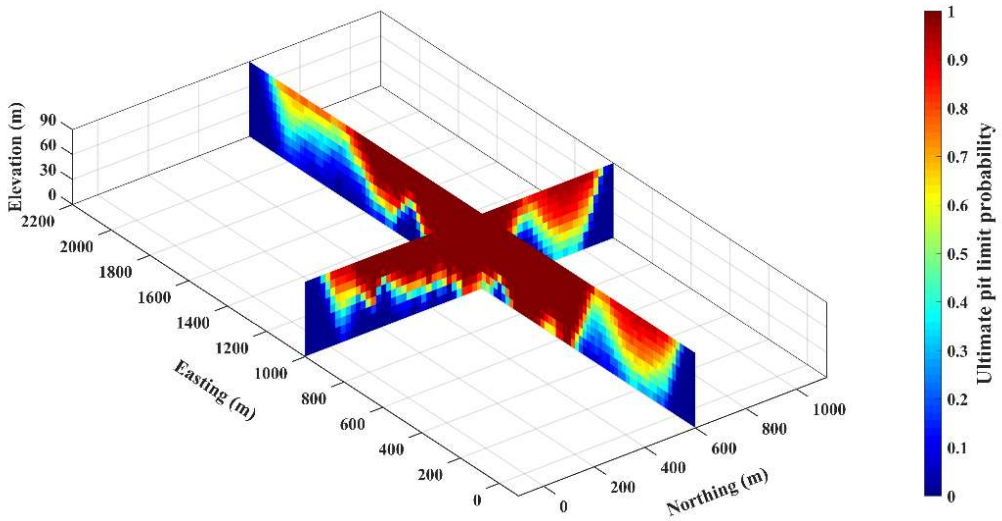


Figure 20. Ultimate pit limit probability over 100 realizations for the upper bound (cross-sectional view)

UPL of the lower bound is computed using an e-type model without any variability (Figure 21). There is a clear resemblance between the UPL of the lower and upper bounds, especially in the central part. Obtained UPL contains 457 Mton of ore and 69 MTon of waste, a lower amount of ore and higher waste tonnage than the upper bound. This fixed UPL is used in the proposed intermediate solution as well.

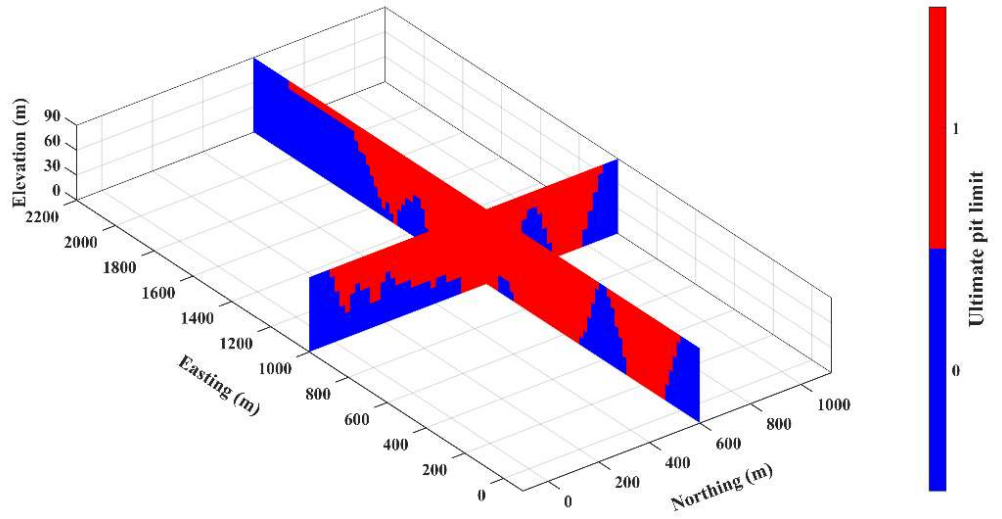


Figure 21. Ultimate pit limit for the lower bound (cross-sectional view)

4.6.3 Production scheduling

The upper bound defines the optimistic strategy with the best NPV for each scenario, which is also impossible considering that it produces different pit limits and extraction sequences depending on realization. For the upper bound, 100 production scheduling results are produced using MIP. The aim, however, is not to restrict the algorithm to produce a certain number of periods. It is done to find the best possible solution for each realization. As a result, each realization has a different pit boundary and a different number of periods that vary from 9 to 11. Figure (22) shows the upper bound results of NPV after each year of production (right axis), ore and waste tonnages in each production period. NPV is calculated using a discount rate of 10%. Moreover, realizations do not exceed the maximum mining (gray dashed line) and processing (blue dashed line) capacities. 95% confidence interval is used to demonstrate the deviations between realizations. 95th and 5th percentiles of ore tonnages are shown as blue squares and orange squares are for NPV. All realizations produce 47.45 Mton/year of ore in the first eight years, which is equal to the maximum processing capacity. This further proves that the upper bound demonstrates the optimal solution to get the project's maximum NPV.

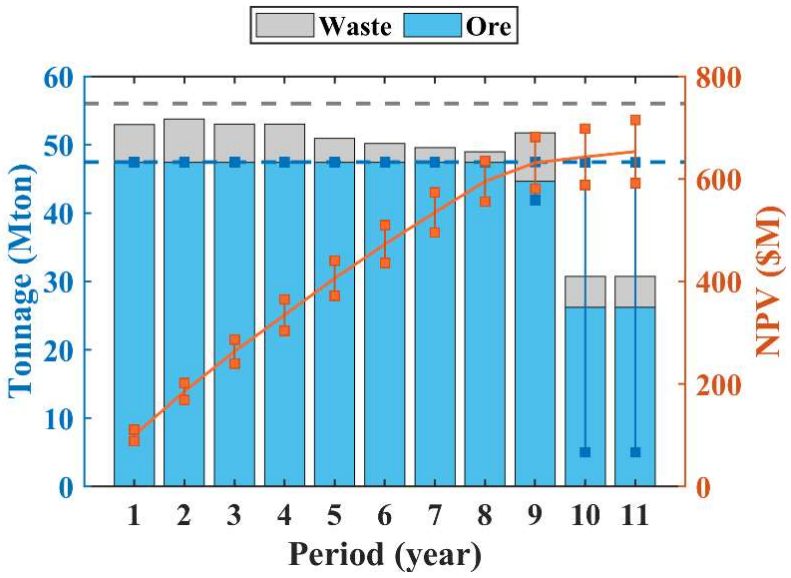


Figure 22. NPV and Ore/Waste tonnages from production scheduling of the upper bound. Grey and blue dashed lines represent maximum mining and production capacities, respectively.

As stated earlier, the deterministic production scheduling using the e-type model as an input is used to obtain the extraction periods, which define the 1st stage decisions for the proposed intermediate solution. The 2nd stage decisions involve re-evaluating block destinations

depending on realizations. As a result, the 2nd stage decisions introduce the uncertainty from realizations to the production scheduling problem. Figure (23) demonstrates ore and waste tonnages and NPV in each production period of the intermediate solution. Although the ore tonnages decreased compared to the upper bound, the production is consistent through the years. 95th and 5th percentiles of ore tonnage do not exceed the maximum available processing constraint, while the deviations in the NPV plot are fairly distributed.

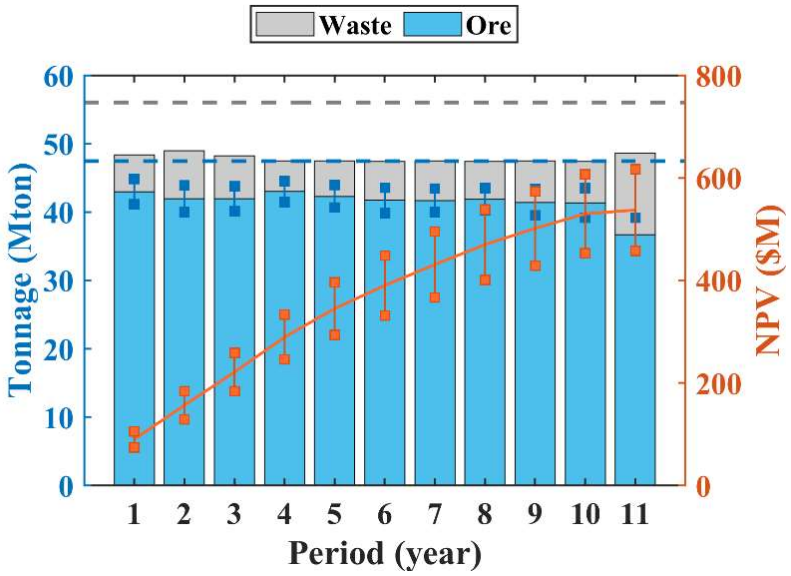


Figure 23. NPV and Ore/Waste tonnages from production scheduling of the intermediate solution. Grey and blue dashed lines represent maximum mining and production capacities, respectively.

4.6.4 Analysis of the production scheduling results

One way to analyze the effect of incorporating 2nd stage decisions for the proposed production scheduling approach is to compare the NPV values of the lower bound, upper bound and intermediate solution. Figure (24) shows boxplots of NPV results after 11 years of mining operations for the upper bound, lower bound and intermediate feasible solution. The intermediate boxplot is significantly closer to the upper bound than to the lower one. Moreover, it can be observed that even the worst scenario from the proposed intermediate solution has a higher NPV than the best scenario from the lower bound. Another important detail is that deviations in the lower bound are between 300 $M and 0 $M. Therefore, if 100 realizations were actual scenarios, it would mean that there will be a significant probability of not gaining any profit by using the deterministic mine planning.

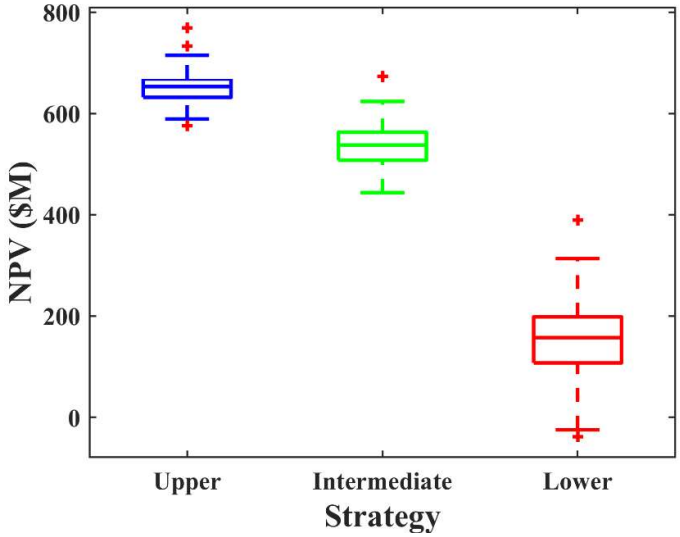


Figure 24. Boxplot of cumulative NPV results for the upper bound, lower bound and deterministic strategies applied on 100 realizations from the proposed cosimulation

Rather than comparing actual values, it is of interest to use the normalization technique to a reference plot. Figure (25) shows the comparison between NPV plots of the upper bound, lower bound and the proposed intermediate mine planning solution, where they are normalized to the reference plot, the upper bound. In this figure, the difference between the intermediate solution and the upper bound is around 17%. Meanwhile, the difference between the intermediate strategy and the lower bound is 59%.

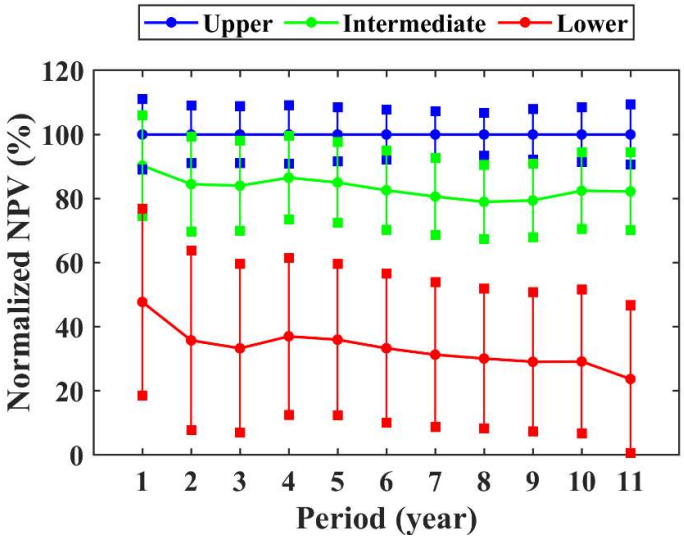


Figure 25. Average NPV plots of the lower bound and deterministic mine planning normalized to the upper bound (reference case of 100%). Squares represent 95th and 5th percentiles from realizations.

Another crucial detail is that the reproduction of an inequality constraint between variables helped to gain more profit. The summary of NPV results and final minable ore tonnages using conventional and proposed cosimulation results are given in Table (6). The proposed cosimulation algorithm increased the NPV of the project and produced a higher amount of ore. The upper bound NPV showed the benefit of 0.94% and the lower bound – 0.76% by using the proposed cosimulation algorithm.

Table 6. Comparison of NPV and ore tonnage values for conventional and proposed cosimulation results

Parameter	Conventional cosimulation	Proposed cosimulation	Average difference
Upper:			
NPV ($M)	647.20 ± 30.30	653.33 ± 30.72	~ +6.13
Ore tonnage (Mton)	471.23 ± 33.83	476.68 ± 32.81	~ +5.45
Intermediate:			
NPV ($M)	533.23 ± 39.67	537.31 ± 39.74	~ +4.08
Ore tonnage (Mton)	448.83 ± 7.30	457.22 ± 7.44	~ +8.39

5 CONCLUSIONS AND RECOMMENDATIONS

The thesis accomplished the objectives that were set before conducting the experiments. A new geostatistical algorithm for variables with inequality constraints was developed and its performance is shown in this study. Moreover, the thesis explains the computer code of the proposed algorithm made according to GSLIB standards (Deutsch and Journel, 1998). The second part of the thesis explains the two-stage stochastic production scheduling that can be used to obtain the optimal stochastic solution for open-pit mine planning.

A comprehensive literature review on the current research state in mine planning and geostatistics identified potential gaps explored in this thesis. The mining industry still employs deterministic approaches for mineral resource estimation and long-term production scheduling, despite the apparent drawbacks of these methods. Moreover, mineral resource estimation traditionally uses kriging algorithms, which cannot reproduce grade variability and generate smooth results that are not realistic. Considering that conventional mine planning techniques use unreliable estimated models, the risk of not meeting production targets is not managed. Alternatively, the implementation of multiple realizations generated by geostatistical simulation techniques into mine planning algorithms is a heavily researched topic. Stochastic production scheduling methods show many benefits in terms of profit, production volumes and design optimization. It is especially true in open-pit mine planning, where smooth kriging results negatively affect the optimal pit limits. Geostaticians, on the other hand, work on multivariate algorithms capable of reproducing complexities present in the bivariate relationships. Particular examples of complex relationships are inequality constraints that are common in geometallurgical variables and iron deposits.

This thesis introduces the hierarchical cosimulation algorithm integrated with inverse transform sampling for modeling variables within the boundaries imposed by an inequality constraint in the bivariate relationship. A new geostatistical approach is based on the updated hierarchical cosimulation integrated with an acceptance-rejection technique developed by Madani and Abulkhair (2020). The main difference between algorithms is that inverse transform sampling re-simulates values that lie outside an inequality constraint only once, instead of making several iterations by an acceptance-rejection method. On top of that, the algorithm incorporates different neighborhood searching strategies, including heterotopic (i.e., single and multiple), collocated and multicollocated. The proposed algorithm is applied to a real case study from Iron

deposit with a sharp inequality constraint between Iron and Silica variables. In the analysis, the algorithm demonstrates acceptable results in histogram and variogram validation without any drawbacks from integrating inverse transform sampling into the second simulation. Compared to the conventional cosimulation, the proposed methodology perfectly reproduces an inequality constraint and shows a higher correlation coefficient.

However, the proposed cosimulation algorithm is not yet completed as it can be further improved in the future. Among many possible recommendations, extending the algorithm to model more than two variables and incorporating a multiple grid technique into the cosimulation can be of interest. Moreover, there are also the following limitations: 1) the methodology is based on sequential Gaussian cosimulation, in which simulation quality highly depends on the predefined neighborhood; 2) the computer code is still slower than other commercial software and needs further improvement; 3) the problem with underestimating the standard deviation of the secondary variable is not resolved.

The second part of the thesis demonstrates the two-stage stochastic long-term production scheduling for open-pit mining of Iron deposits. The methodology is based on two groups of decisions related to mine planning. The 1st stage decisions involve computing the extraction sequence, while the 2nd stage decisions are focused on block destinations. The production scheduling is performed using an e-type model during the first stage, producing fixed extraction periods. The destinations are then re-evaluated in the second stage based on information from realizations. This approach is applied to 100 realizations obtained by the proposed cosimulation algorithm, generating promising results in terms of capacities and NPV. The proposed intermediate solution is compared to the upper and lower bounds. The upper bound uses MIP for each realization individually, making it an impossible strategy. Alternatively, the lower bound is a deterministic production scheduling applied to the e-type model, thus it is an unfavorable strategy that cannot manage the risk of sending extracted material to the wrong destinations. Based on the analysis, the intermediate solution is closer to the upper bound, with only a 17% difference. On the other hand, it has a 59% higher NPV compared to the lower bound. Moreover, the NPV benefit of using the proposed cosimulation algorithm is 0.94% for the upper and 0.76% for the intermediate solution compared to the conventional cosimulation. Although integrating inverse transform sampling into hierarchical cosimulation does not affect the simulation of the primary variable, preventing overestimation of the secondary variable helps to generate higher NPV.

6 REFERENCES

Abildin, Y, Madani, N and Topal, E, 2019. A Hybrid Approach for Joint Simulation of Geometallurgical Variables with Inequality Constraint. *Minerals*, 9(1):24.

Adeli, A, Emery, X and Dowd, P, 2018. Geological Modelling and Validation of Geological Interpretations via Simulation and Classification of Quantitative Covariates. *Minerals*, 8(1):7.

Albor Consquega, F and Dimitrakopoulos, R, 2009. Stochastic mine design optimisation based on simulated annealing: pit limits, production schedules, multiple orebody scenarios and sensitivity analysis. *Mining Technology*, 118(2):79-90.

Almeida, A S and Journel, A G, 1994. Joint simulation of multiple variables with a Markov-type coregionalization model. *Mathematical Geology*, 26:565-588.

Arcari Bassani, M A, Costa, J and Deutsch, C V, 2018. Multivariate geostatistical simulation with sum and fraction constraints. *Applied Earth Science*, 127(3):83-93.

Askari-Nasab, H, Pourrahimian, Y, Ben-Awuah, E and Kalantari, S, 2011. Mixed integer linear programming formulations for open pit production scheduling. *Journal of Mining Science*, 47:338.

Barnett, R M, Manchuk, J G and Deutsch, C V, 2014. Projection Pursuit Multivariate Transform. *Mathematical Geosciences*, 46:337-359.

Barnett, R M, Manchuk, J G and Deutsch, C V, 2016. The Projection-Pursuit Multivariate Transform for Improved Continuous Variable Modeling. *Society of Petroleum Engineers*, 21(6):2010-2026.

Battalgazy, N and Madani, N, 2019a. Categorization of Mineral Resources Based on Different Geostatistical Simulation Algorithms: A Case Study from an Iron Ore Deposit. *Natural Resources Research*, 28:1329–1351.

Battalgazy, N and Madani, N, 2019b. Stochastic Modeling of Chemical Compounds in a Limestone Deposit by Unlocking the Complexity in Bivariate Relationships. *Minerals*, 9(11):683.

Bienstock, D and Zuckerberg, M, 2010. Solving LP relaxations of large-scale precedence constrained problems. In *Proceedings of the 14th conference on Integer Programming and Combinatorial Optimization (IPCO)* (eds: F Eisenbrand and F B Shepherd), pp 1-14 (Springer: Berlin).

Burkardt, J, 2014. The Truncated Normal Distribution, [online]. Available from: < https://people.sc.fsu.edu/~jburkardt/presentations/truncated_normal.pdf > [Accessed: 29 March 2021].

Caccetta, L and Hill, S P, 2003. An Application of Branch and Cut to Open Pit Mine Scheduling. *Journal of Global Optimization*, 27:349-365.

Chilès, J P and Delfiner, P, 2012. *Geostatistics: modeling spatial uncertainty*, 734 p (Wiley: New York).

Dagdelen, K, 1985. Optimum multi-period open pit mine production scheduling, PhD thesis, Colorado School of Mines, Golden.

Dagdelen, K and Francois-Bongarcon, D, 1982. Towards the complete double parameterization of recovered reserves in open pit mining. In *Proceedings of the 17th International Symposium on the Application of Computers and Operations Research in the Mineral Industry*, pp 288–296 (Society of Mining Engineers: New York).

Dagdelen, K and Johnson, TB, 1986. Optimum open pit mine production scheduling by Lagrangian parameterization. In *Proceedings of the 19th International Symposium on the Application of Computers and Operations Research in the Mineral Industry*, pp 127–142 (Jostens Publications: State College).

David, M, 1977. *Geostatistical ore reserve estimation*, 384 p (Elsevier: New York).

Davis, B M and Greenes, K A, 1983. Estimation using spatially distributed multivariate data: An example with coal quality. *Journal of the International Association for Mathematical Geology*, 15:287-300.

De Iaco, S and Maggio, S, 2011. Validation Techniques for Geological Patterns Simulations Based on Variogram and Multiple-Point Statistics. *Mathematical Geosciences*, 43:483-500.

Desbarats, A J and Dimitrakopoulos, R, 2000. Geostatistical Simulation of Regionalized Pore-Size Distributions Using Min/Max Autocorrelation Factors. *Mathematical Geology*, 32:919-942.

Deutsch, C V, 1989. DECLUS: a fortran 77 program for determining optimum spatial declustering weights. *Computers & Geosciences*, 15(3):325-332.

Deutsch, C V, 1992. Annealing techniques applied to reservoir modeling and the integration of geological and engineering (well test) data, PhD thesis, Stanford University, Stanford.

Deutsch, C V and Journel, A G, 1998. *Geostatistical software library and users guide*, 369 p (Oxford University Press: New York).

Deutsch, M, González, E and Williams, M, 2015. Using simulation to quantify uncertainty in ultimate-pit limits and inform infrastructure placement. *Mining Engineering*, 67(12):49-55.

Devroye, L, 1986. *Non-Uniform Random Variate Generation*, 843 p (Springer: New York).

Dimitrakopoulos, R, 2011. Stochastic optimization for strategic mine planning: A decade of developments. *Journal of Mining Science*, 47:138-150.

Dimitrakopoulos, R, Farrelly, C T and Godoy, M, 2002. Moving forward from traditional optimization: grade uncertainty and risk effects in open-pit design. *Mining Technology*, 111(1):82-88.

Dimitrakopoulos, R and Ramazan, S, 2008. Stochastic integer programming for optimising long term production schedules of open pit mines: methods, application and value of stochastic solutions. *Mining Technology*, 117(4):155-160.

Dowd, P A, 1994. Risk assessment in reserve estimation and open-pit planning. *Transactions of the Institution of Mining and Metallurgy*, 103:148-154.

Dowd, P A, Xu, C and Coward, S, 2016. Strategic mine planning and design: some challenges and strategies for addressing them. *Mining Technology*, 125:22-34.

Dubrule, O and Kostov, C, 1986. An interpolation method taking into account inequality constraints: I. Methodology. *Mathematical Geology*, 18:33-51.

Emery, X, 2008. A turning bands program for conditional co-simulation of cross-correlated Gaussian random fields. *Computers & Geosciences*, 34(12):1850-1862.

Emery, X, 2009. The kriging update equations and their application to the selection of neighboring data. *Computational Geosciences*, 13:269-280.

Emery, X, 2010. Iterative algorithms for fitting a linear model of coregionalization. *Computers & Geosciences*, 36(9):1150-1160.

Emery, X, 2012. Co-simulating Total and Soluble Copper Grades in an Oxide Ore Deposit. *Mathematical Geosciences*, 44:27-46.

Emery, X, Arroyo, D and Peláez, M, 2014. Simulating Large Gaussian Random Vectors Subject to Inequality Constraints by Gibbs Sampling. *Mathematical Geosciences*, 46:265-283.

Emery, X and Lantuéjoul, C, 2006. TBSIM: A computer program for conditional simulation of three-dimensional Gaussian random fields via the turning bands method. *Computers & Geosciences*, 32(10):1615-1628.

Emery, X and Ortiz, J M, 2012. Enhanced coregionalization analysis for simulating vector Gaussian random fields. *Computers & Geosciences*, 42:126-135.

Escudero, L F, Kamesam, P V, King, A J and Wets, R J-B, 1993. Production planning via scenario modelling. *Annals of Operations Research*, 43:309-335.

Espinoza, D, Goycoolea, M, Moreno, E and Newman, A, 2013. MineLib: a library of open pit mining problems. *Annals of Operations Research*, 206:93–114.

Friedman, J H, 1987. Exploratory Projection Pursuit. *Journal of the American Statistical Association*, 82:249-266.

Gershon, M E, 1983. Optimal mine production scheduling: evaluation of large scale mathematical programming approaches. *International Journal of Mining Engineering*, 1:315-329.

Gershon, M E, 1987. An open-pit production scheduler: Algorithm and implementation. *Mining Engineering*, 39:793–796.

Godoy, M and Dimitrakopoulos, R, 2004. Managing risk and waste mining in long-term production scheduling of open-pit mines. *SME Transactions*, 316:43-50.

Goodfellow, R and Dimitrakopoulos, R, 2017. Simultaneous Stochastic Optimization of Mining Complexes and Mineral Value Chains. *Mathematical Geosciences*, 49:341–360.

Goovaerts, P, 1997. *Geostatistics for Natural Resource Evaluation*, 483 p (Oxford University Press: New York).

Hochbaum, D S, 2001. A new-old algorithm for minimum-cut and maximum-flow in closure graphs. *Networks*, 37(4):171-193.

Hochbaum, D S, 2008. The Pseudoflow Algorithm: A New Algorithm for the Maximum-Flow Problem. *Operations Research*, 56(4):992–1009.

Hochbaum, D S and Chen, A, 2000. Performance analysis and best implementations of old and new algorithms for the open-pit mining problem. *Operations Research*, 48(6):894-914.

Hoerger, S, Seymour, F and Hoffman, L, 1999. Mine planning at Newmont's Nevada operations. *Mining Engineering*, 51:26-30.

Hosseini, S A and Asghari, O, 2015. Simulation of geometallurgical variables through stepwise conditional transformation in Sungun copper deposit, Iran. *Arabian Journal of Geosciences*, 8:3821–3831.

Huber, P J, 1964. Robust Estimation of a Location Parameter. *Annals of Mathematical Statistics*, 35(1):73-101.

Hustrulid, W A, Kuchta, M and Martin, R K, 2013. *Open Pit Mine Planning and Design*, 1308 p (CRC Press: Boca Raton).

Ibarra, O H and Kim, C E, 1978. Approximation Algorithms for Certain Scheduling Problems. *Mathematics of Operations Research*, 3(3):197-204.

Isaaks, E H, 1990. The application of Monte Carlo methods to the analysis of spatially correlated data, PhD thesis, Stanford University, Stanford.

Johnson, T, 1968. Optimum open pit mine production scheduling, PhD thesis, University of California, Berkeley.

Journel, A G, 1999. Markov Models for Cross-Covariances. *Mathematical Geology*, 31:955–964.

Krige, D G, 1951. A statistical approach to some basic mine valuation problems on the Witwatersrand. *Journal of the Southern African Institute of Mining and Metallurgy*, 52(6):119-139.

Larrondo, P F, Neufeld, C T and Deutsch, C V, 2003. VARFIT: a program for semiautomatic variogram modelling, [online]. Available from: < http://www.ccgalberta.com/ccgresources/report05/2003-122-varfit.pdf > [Accessed: 29 March 2021].

Leite, A and Dimitrakopoulos, R, 2007. Stochastic optimisation model for open pit mine planning: application and risk analysis at copper deposit. *Mining Technology*, 116(3):109-118.

Lerchs, H and Grossmann, L, 1965. Optimum design of open pit mines. *Canadian Institute of Mining Transactions*, 68:17-24.

Leuangthong, O amd Deutsch, C V, 2003. Stepwise Conditional Transformation for Simulation of Multiple Variables. *Mathematical Geology*, 35:155–173.

Luster, G R, 1985. Raw materials for Portland cement: Applications of conditional simulation of coregionalization, PhD thesis, Stanford University, Stanford.

Madani, N and Abulkhair, S, 2020. A hierarchical cosimulation algorithm integrated with an acceptance–rejection method for the geostatistical modeling of variables with inequality constraints. *Stochastic Environmental Research and Risk Assessment*, 34:1559-1589.

Madani, N and Emery, X, 2019. A comparison of search strategies to design the cokriging neighborhood for predicting coregionalized variables. *Stochastic Environmental Research and Risk Assessment*, 33:183–199.

Maleki, M, Jélvez, E, Emery, X and Morales, N, 2020. Stochastic Open-Pit Mine Production Scheduling: A Case Study of an Iron Deposit. *Minerals*, 10(7):585.

Mallet, J L, 1980. Régression sous contraintes linéaires : application au codage des variables aléatoires. *Revue de statistique appliquée*, 28(1):57-68.

Mata-Lima, H, 2005. Geostatistic in Reservoir Characterization: from estimation to simulation methods. *Estudios Geológicos*, 61(3):135-145.

Matheron, G, 1963. Principles of geostatistics. *Economic Geology*, 58(8):1246-1266.

Matheron, G, 1973. The Intrinsic Random Functions and Their Applications. *Advances in Applied Probability*, 5(3):439-468.

Mery, N et al, 2017. Geostatistical modeling of the geological uncertainty in an iron ore deposit. *Ore Geology Reviews*, 88:336-351.

Morales, N et al, 2019. Incorporation of Geometallurgical Attributes and Geological Uncertainty into Long-Term Open-Pit Mine Planning. *Minerals*, 9(2):108.

Moreno, E et al, 2017. A two-stage stochastic model for open pit mine planning under geological uncertainty. In *Proceedings of the 38th International Symposium on the Application of Computers and Operations Research in the Mineral Industry*, (Colorado School of Mines: Golden).

Moreno, E, Espinoza, D and Goycoolea, M, 2010. Large-scale multi-period precedence constrained knapsack problem: A mining application. *Electronic Notes in Discrete Mathematics*, 36:407-414.

Myers, D E, 1982. Matrix formulation of co-kriging. *Journal of the International Association for Mathematical Geology*, 14:249–257.

Myers, D E, 1991. Pseudo-cross variograms, positive-definiteness, and cokriging. *Mathematical Geology*, 23:805-816.

Osanloo, M, Gholamnejad, J and Karimi, B, 2008. Long-term open pit mine production planning: a review of models and algorithms. *International Journal of Mining, Reclamation and Environment*, 22(1):3-35.

Pana, M T, 1965. The simulation approach to open-pit design. In *Proceedings of the 5th International Symposium on Application of Computers and Operations Research in the Mineral Industry*, pp 1–14 (CRC press: Boca Raton).

Paravarzar, S, Emery, X and Madani, N, 2015. Comparing sequential Gaussian and turning bands algorithms for cosimulating grades in multi-element deposits. *Comptes Rendus Geoscience*, 347(2):84-93.

Pawlowsky-Glahn, V and Egozcue, J J, 2006. Compositional data and their analysis: An introduction. *Geological Society Special Publications*, 264(1):1-10.

Pawlowsky-Glahn, V and Olea, R A, 2004. *Geostatistical analysis of compositional data. Studies in mathematical geology*, 304 p (Oxford University Press: Oxford).

Picard, J C, 1976. Maximum Closure of a Graph and Applications to Combinatorial Problems. *Management Science*, 22(11):1268-1272.

Pyrcz, M J and Deutsch, C V, 2014. *Geostatistical Reservoir Modeling*, 376 p (Oxford University Press: New York).

Ramazan, S, 2007. The new Fundamental Tree Algorithm for production scheduling of open pit mines. *European Journal of Operational Research*, 177(2):1153-1166.

Ramazan, S and Dimitrakopoulos, R, 2004. Traditional and New MIP Models for Production Scheduling With In-Situ Grade Variability. *International Journal of Surface Mining*, 18(2):85-98.

Ramazan, S and Dimitrakopoulos, R, 2013. Production scheduling with uncertain supply: a new solution to the open pit mining problem. *Optimization and Engineering*, 14:361-380.

Ramazan, S and Dimitrakopoulos, R, 2018. Stochastic optimisation of long-term production scheduling for open pit mines with a new integer programming formulation, in *Advances in Applied Strategic Mine Planning* (ed: R Dimitrakopoulos), pp 139-153 (Springer: Cham).

Ravenscroft, P, 1992. Risk analysis for mine scheduling by conditional simulation. *Transactions of the Institution of Mining and Metallurgy*, 101:A104-A108.

Rimélé, A, Dimitrakopoulos, R and Gamache, M, 2020. A dynamic stochastic programming approach for open-pit mine planning with geological and commodity price uncertainty. *Resources Policy*, 65:101570.

Rivoirard, J, 1994. *Introduction to disjunctive kriging and non-linear geostatistics*, 192 p (Clarendon Press: Oxford).

Rivoirard, J, 2001. Which Models for Collocated Cokriging? *Mathematical Geology*, 33:117-131.

Rivoirard, J, 2004. On Some Simplifications of Cokriging Neighborhood. *Mathematical Geology*, 36:899–915.

Rondon, O, 2012. Teaching Aid: Minimum/Maximum Autocorrelation Factors for Joint Simulation of Attributes. *Mathematical Geosciences*, 44:469–504.

Rossi, M E and Deutsch, C V, 2014. *Mineral Resource Estimation*, 332 p (Springer: Berlin).

Soares, A, 2001. Direct Sequential Simulation and Cosimulation. *Mathematical Geology*, 33:911-926.

Subramanyam, A and Pandalai, H S, 2004. On the Equivalence of the Cokriging and Kriging Systems. *Mathematical Geology*, 36:507–523.

Tolwinski, B and Underwood, R, 1996. A scheduling algorithm for open pit mines. *IMA Journal of Management Mathematics*, 7(3):247-270.

van den Boogaart, K G, Mueller, U and Tolosana-Delgado, R, 2017. An Affine Equivariant Multivariate Normal Score Transform for Compositional Data. *Mathematical Geosciences*, 49:231–251.

Verly, G W, 1993. Sequential Gaussian Cosimulation: A Simulation Method Integrating Several Types of Information, in *Geostatistics Tróia '92. Quantitative Geology and Geostatistics, vol 5* (ed: A Soares), pp 543-554 (Springer: Dordrecht).

Wackernagel, H, 2003. *Multivariate Geostatistics*, 388 p (Springer: Berlin).

Whittle, J, 1999. A decade of open pit mine planning and optimisation - the craft of turning algorithms into packages. In *Proceedings of the 28th International Symposium on Application of Computers and Operations Research in the Mineral Industry*, pp 15–24 (Colorado School of Mines: Golden).

Xu, W, Tran, T T, Srivastava, R M and Journel, A G, 1992. Integrating Seismic Data in Reservoir Modeling: The Collocated Cokriging Alternative. In *SPE Annual Technical Conference and Exhibition*, pp 833-842 (Society of Petroleum Engineers: Washington).