Improved Estimation With Multiple Data Types For Medium-Term Mine Planning

by

Fernando Agustin Camana Calderon

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

In

Mining Engineering

Department of Civil and Environmental Engineering

University of Alberta

Abstract

Drilling is the primary source of geological information in the form of rock samples for geological logging and chemical assays. There are data from multiple drilling types in an open-pit mining operation, with unique technical features, costs, volume support, and sampling error and bias. The most common drilling types are diamond drilling and blast-holes. Drilling data is classified into primary and secondary according to the confidence and quality of the information. Primary data corresponds to the highest quality and confidence, and usually, diamond drilling is considered primary data. Secondary data presents lower confidence and quality, and usually, the blast-holes are considered in this classification. Access to real data is always quite limited. Besides, many times researching requires data with specific features. Data simulation can provide a solution to these restrictions.

This research has developed a procedure for simulating drilling data, with similar statistics and features to the real one and adjustable to the requirements of the user. The procedure simulates the distribution inside each drill hole, replicates drilling supports and sampling protocols to get data from multiple drilling types, specifically diamond drilling and blast holes.

Simulated drilling data is used to estimate different resources models used in the ore control process, focusing on the medium-term model. The effects of dataset bias in the profits and ore/waste classification are also assessed for different resources models. Besides the ordinary kriging estimation, a cokriging outline is implemented to take advantage of the multiple simulated data types.

The simulation method provides highly realistic data of multiple drilling types. They have been tested several times to check the correct reproduction of input parameters and distribution features. The simulated data have been used in resource estimation and

assessing the performance of the different models used in the ore control process. The medium-term model estimated using cokriging and simulated drilling data provides better results and profits than models estimated using a single dataset type. The bias impacts negatively in the profits as expected, but under some conditions, it can diminish the loss in profits by misclassification.

Dedication

To my family for their patient and love, to my friends and colleagues for their advice and support.

Acknowledgments

I would like to thank my supervisor, Dr. Clayton Deutsch. His enthusiasm, expertise, knowledge, and guidance helped me start and end this master's.

Thanks to my friends and colleagues from the Centre for computational Geo-statistics, the University of Alberta, and Edmonton. Their advice and friendship are priceless. Thanks for making me feel at home.

Thanks to the sponsors of the Centre for computational Geo-statistics by the financial support provided to me and many others of my colleagues.

Finally, an excellent thank you to all my professors; their knowledge and experience helped me be a better professional.

Table of Contents

List of Tables

List of Figures

List of abbreviations

Abbreviation	Description
DH	Drill hole
DD	Diamond drilling
BH	Blast-hole
LTM	Long-term model
STM	Short-term model
MTM	Medium-term modelMedium term model
m	meters
CDF	Cummulative distribution function
PDF	Probabilistic density function
CCDF	Conditional Cummulative distribution function
GSLIB	Geostatistical software library
LMR	Lineal model of regionalization
LMC	Lineal model of coregionalization
NS	Normal scores
RF	Random function
RV	Random variable
ReV	Regionalized variable
Cu	Copper
OK	Ordinary kriging
COK	Ordinary cokriging
SOCK	Standardized ordinary cokriging
NN	Nearest neighbor
ID	Inverse distance

List of equations

1. Introduction

Drill holes are the main source of information in a mining operation; their rock samples are used for many purposes including geological logging, lab assays and metallurgical tests. The focus here is on the application of drill hole data to resource estimation. In an open-pit mining operation, there are drill hole data from different types, each one with different features and qualities. Diamond drilling and blast-holes are common. The main difference between each drilling are the drilling support and quality of the information provided. The data may have different error and bias. Simply combining the data from different drilling types in a single database for estimation would provide poor results. The main drilling data types available in an open-pit mining operation are diamond drilling (DD) and blast-holes (BH). DD is used for exploration and long-term planning models, while BH is used exclusively in short-term models (Ore control).

As access to a real DD and BH database is quite limited, data simulation is considered in this research. A reference true distribution is simulated, and reasonable error and bias can be added to simulate primary or secondary data. For the objectives of our research, it is necessary to develop a simulation method of drilling data that replicates realistic features of DD and BH drilling and sampling.

This research focuses on assessing the performance of the different models used in the ore control process, applying different types of data and estimation techniques. The short-term model, also known as the ore control model, has great importance as it is used to define the production polygons in the ore control. The best estimation of the short-term model is essential for the best economic performance. For this purpose, it is necessary to understand how to manage the bias in the BH data. In addition, the medium-term model is used for planning future mining. This model represents a middle step between the short-term and long-term models. The medium-term model must provide a local estimation with higher resolution and accuracy and consider both DD and BH data. A good solution is to estimate this model using cokriging, with shorter block size. Cokriging provides improved results compared to estimation using a single database and filters the error and mitigating the effects of conditional bias (Minnitt & Deutsch, 2014).

1.1. Problem setting and scope of study

BH data are used almost exclusively for ore control, while the limited but high-quality DD data are useful for long-term planning. Each data type has different advantages and handicaps: DD, sparsely drilled, produces models with low resolution and detail, which is a potential for dilution, besides smooth mineral grades, but provides the best estimation results, unbiased and globally accurate; BH, densely drilled, provides a local high-resolution model, quite useful for ore control, but as BH data have higher bias and errors than DD data, there is a potential for incorrect estimation and imprecise profit estimation.

Practitioners must explore ways to take advantage of both databases, dealing with the bias and error in the secondary data. Medium-term models estimated using cokriging is an excellent way to use both databases, besides its application has a positive impact on planning. Dataset bias affects the model estimation, which affects, directly and indirectly, the profits. It is necessary to understand how exactly bias and error influence profits and which are the best solutions to deal with them. The aims of this research are:

1. The development of a new method to simulate drilling data of multiple types. This method considers the individual simulation of each drill hole, using its specific volume support, and a highly detailed distribution at point-scale. By replicating the drilling support and sampling protocols for specific drilling types, it is possible to get simulated data with differences based on real drilling features.
2. The evaluation of different resources models, using different datasets, and implementing cokriging techniques. The focus will be on the medium-term model, estimated using cokriging, and on the performance of the different models to predict ore zones for planning and mining production.
3. The assessment of the impact of the bias and error in the profits and ore prediction of resource models. By common sense, the bias in the database is prejudicial, because leads to incorrect estimation. But a limited and positive bias could impact in an unexpected way the profits. It is necessary to analyze carefully how the profit is affected by variable bias and which is the most correct approach to deal with biased data.

1.2.Drilling data types

Drilling methods differ in many parameters, but the most important are the support, rock sample type, and the sampling protocol. These factors produce differences in grades, statistics, error, and bias between the drilling data types. In an open-pit mining operation, the most common drilling types are diamond drilling and blast holes. Diamond drilling is used in exploration and infill drilling stages, while the blast holes are used only in drill and blast operations and for short-term models.

DIAMOND DRILLING (DD) provides a continuous rock cylinder (Rock core) that brings detailed information of the underlying geology, besides contamination during drilling is practically null. The sample support is defined by the drilling diameter, which is selected according to the drilling objective: NQ and HQ are common diameters in exploration and infill stages and PQ is used for metallurgical tests and hydrology. The rock core provides accurate geologic information and mineral grades, besides each sample can be linked to a specific point in the space. These features make this drilling type has a very low potential for errors and bias, besides ensuring accurate data. The main con with this drilling type is its expensive costs, which limit the number of drill holes and samples. DD is usually drilled using a large data spacing that can vary from tens to hundreds of meters.

BLAST-HOLES (BH) are used in mining production for blasting the rock mass, gathering geological data, and obtaining samples for chemical assays. The drilling is done by a rotary tri-cone, producing rock cuttings, which are accumulated around the drill hole collar, from which is taken the sample. The contamination is permanent and from different sources: falling walls, exposition to weather elements, loss of fine material by water or wind, etc. The drilling support is not the same for all BH, as there is an over-drilling to improve the blasting, which can vary according to the rock hardness. The BH diameter can vary according to the provider but generally is in a range of 100-250 mm. Another problem is that the cuttings cone is not necessarily representative of the subjacent rock, as this one can be deposited in an asymmetric way, or there could be a low rock sample recovery. The sampling protocol is another source of error, the sampling can be done using a mechanical sampler or even a shovel, which adds extra heterogeneity to the sample. The sample support is variable, and it is based on the sample weight, with a minimum of 6 kg per BH. These factors make that the BH data have a high potential for error and bias, and influence the quality and reliability of this data type. The error and bias are variable and can be

controlled with a strict QAQC protocol, but never eliminated. Unlike DD, the BH drilling is cheaper and uses a short data spacing, usually a few meters. The BH databases are large and can include thousands of samples in the same area where the DD has only a few dozens of samples.

Table 1.1: Drilling support and sampling protocols

Drilling type	Diameter (mm)	Support (m3)	Sampling
DD - NQ	48	0.90	½ core
DD - HQ	64	1.21	½ core
DD - PQ	84	1.58	¼ core
BH	200	3.77	6 kg per hole

Drilling data is classified into primary and secondary based on the quality and confidence of the information they provide. Primary data correspond to the data type with the highest quality and reliability. DD data provides a solid rock sample and has a detailed sampling protocol, ensuring high-quality information, so usually is considered as primary data. BH data have less quality and reliability due to the drilling method and to the lax quality controls in the sampling process, which make them be considered secondary data. Table 1.1 summarizes the drilling supports and sampling protocols for each drilling type.

1.3.Bias and error

The error is the difference between an estimated value and the true one. As the error can be positive or negative, the mean square error (MSE) is the most common metric for this parameter. Bias is the systematic difference between two or more datasets, in a simple way can be calculated as the difference between each dataset means.

Bias and errors change for each drilling data type. DD has the lowest potential for errors by sampling, so is considered as unbiased and the closest values to the unknown true distribution. On the other hand, BH is considered biased, although its degree is variable and depends on the material, mineral grades and the improvement of the sampling protocol. Working with real data makes it impossible to measure error and bias since there is no true data to compare. As DD data are unbiased, and with very low error, they are used as a reference to measure the bias and error of other drilling data types. In the case of using simulated data, the true distribution is known and is possible to measure the error and bias of each drilling data type. In mining

operations, BH data bias must be measured and controlled. This is not always easy, because depends on many factors, like following a strict QAQC protocol, improving sampling methods, selecting an optimal sample weight, etc. In the end, BH data always have a residual bias, and the user must deal with it. The BH bias can have different impacts on the short-term model estimation, a positive bias can originate an over-estimation, and a negative one the contrary effect. Both cases affect the estimated profits of the short-term model.

1.4. Resources models in open pit mining operations

The different stages in a mining operation require specific resources models. For long-term planning, usually from 1 to 3 years, it is required a model that provides an accurate global estimation, while in daily mining operations is necessary a local model with high resolution. Each model requires different databases, estimation parameters, and even estimation methods.

The Long-Term Model (LTM) is built with the perspective to bring acceptable tonnage and grades estimation for the whole life of mine (LOM). It is estimated using DD data and continuously updated with new information from the infill drilling campaigns as the open pit deepens. The primary data used ensures an accurate and unbiased estimation. The cons of the LTM are related to the block size, usually selected according to the selective mining unit (SMU), which can be in the range of tens meters. This results in a low-resolution model, which reflects inaccurately the real mineral distribution, implicating an increase in the dilution and ore loss. Despite this, the LTM is a very useful tool, which brings accurate global predictions for large work scales.

The Medium-Term Model (MTM) has as main objective to provide an accurate local estimation with better resolution than the LTM. As the name suggests, it is used in planning for intermediates periods between the long and short-term planning, usually less than a year. The MTM helps to predict the ore zones in the next stages of mine production, with better accuracy than the LTM. MTM uses drilling data available in mining operations, like DD and BH. Some estimation outlines consider the BH data from upper benches and the DD data integrated into the same dataset. This is not the most recommendable, because both data types have different levels of bias and error, besides the sampling support, so they cannot be integrated into a single dataset without considering a previous treatment to the data. The MTM uses an intermediate block size between the SMU and the ore control model,

regularly less than a couple of tens of meters. Cokriging is the most appropriate method for estimating using both datasets.

The Short-Term Model (MTM), also known as the ore control model, is used in daily mining operations. This model is estimated using only BH data, and its main objective is to provide a high-resolution model for an optimal ore block classification and polygons design. The STM uses a very short estimation grid, just a few meters, being the most recommended a quarter distance of the data spacing. Although BH data are a potential for error and bias, a large number of samples, the dense drilling grid, and the short estimation grid produce a model quite detailed and accurate, very helpful for the objectives of the ore control process. The STM is used on reconciliation, comparing their results versus exploration model, plant reports, dispatch reports, etc.

As each model is built with different objectives, besides different databases and estimation parameters, we cannot expect the same estimation performance for all models. the best way to compare all models is the comparison versus the reference distribution, checking specific metrics (Bias, MSE, ore/waste classification and profits).

1.5.Multi-type drilling data simulation procedure

The mineral deposit can be defined as a set of millions of microscopic particles of mineral and gangue, grouped by a set of common features that control its distribution and magnitude. It is evident that the mineralization just represents a minimum fraction of the whole rock mass, e.g., in a copper porphyry, Cu minerals represent only 2 to 5%. Then, a drill hole sample is essentially a mix of gangue and mineral grains. Depending on the drilling support and rock sample type, this mix can change its composition and with this, its mineral content.

The drilling data simulation procedure developed for this research tries to replicate the features mentioned above, first simulating a highly detailed distribution, where each node corresponding to a single and pure mineral or gangue grain. Then, replicating the drilling supports, we can obtain samples with different volume support, and mainly with different compositions for each drilling type. The final stage is the replication of the sampling protocol and compositing. The simulation procedure has the following stages:

1. True distribution simulation, this is an unconditional Gaussian simulation, which covers the entire work area. The most important parameter is the input

variogram, which must include several short-range structures that ensure an acceptable grade of short-range variability. Each node of this distribution is assigned to a pure mineral species or gangue. As each node has the mineral content of its species, it is necessary to fix up the content of each mineral species to have under control the average grade of the distribution.

2. Drill hole simulation, at this stage, we proceed to simulate each drill hole individually, in the volume occupied by the respective drilling type. This is a Gaussian simulation conditioned to the true distribution, using just the nodes around each drill hole location in a radio of 5 meters. This simulation uses a very short grid, in millimetric scale, ensuring the necessary detail to represents the mineral distribution. The nodes must be assigned to the same mineral distribution of the previous stage.
3. Composition and Sampling protocol, here the simulated drill holes are fitted to the support of the respective drilling type, rejecting the nodes outside of the drilling diameter. Besides, the drill holes are sampled according to the drilling type: DD core is split by half, taking only a half for the respective sampling; BH are sampled taking 4 to 4.5 kg of each drill hole. The final step is the composition to a single length and the addition of a bias and error according to the research requirements.

The resulting datasets of different drilling types are validated, to ensure the correct reproduction of the input parameters, checking the reproduction of the input variogram. Then the simulated drilling data can be used for the different tests of this research.

1.6. Cokriging and its application on resources estimation

Cokriging is an estimation method that minimizes the variance of estimation by exploiting the cross-correlation between several datasets (Isaaks & Srivastavam, 1989). The cokriging lets the estimation using more than one dataset with different qualities, collocated or not (Minnit & Deutsch, 2014). Cokriging presents some advantages over Ordinary or simple kriging, like producing better estimations and unbiased results. Their cons are that requires a complex estimation setup that limits its application, and a certain grade of correlation between the datasets to obtain optimal results. Minnitt & Deutsch (2014) suggest the best performance of cokriging when both data types are unequally sampled and have different grades of quality and reliability. An open-pit mining operation is an adequate scenario to run a cokriging

estimation, as there are abundant drilling data from different types, most of them non-collocated. For this research, we have implemented cokriging in the estimation of the medium-term models, with the objective to use it for predicting mineralization and ore zones in the in-pit operations. As the data in use, DD and BH, are non-collocated, we must use cokriging with full LMC. The linear model of coregionalization or LMC describes the spatial continuity for a set of variables. LMC fitting requires the direct and cross variograms, but the last cannot be calculated using non-collocated data. The solution is calculating it from the cross-covariance and flipping the results, obtaining the cross-variograms.

2. Theoretical background

This research is focused on resource estimation using multiple types of data and implementing cokriging for estimating ore control models in an open-pit mining operation. It is necessary to review estimation for ore control and recall concepts of geology, statistics, and mining needed for this research.

2.1.Resource estimation in mining operations

The resource estimation process involves many stages, from the data acquisition until the validation of the results. Depending on the model in evaluation, the process can include the use of different databases or estimation methods. But, in general, a normal resource estimation process includes data acquisition, exploratory data analysis, spatial variability analysis (variography), estimation, and validation of results.

Ore control process requires the short-term model and long-term models for production scheduling. Long and medium-term models differ in the databases used and estimation setup. Even so, most of them use the same geological and statistical principles.

2.1.1. Random variable and random function

A natural phenomenon, like rock density or mineral grades, can take different values depending on which location is measured. This is called a Regionalized Variable (*ReV*) as the values of the variable change according to the spatial location and is a merely descriptive definition (Journel & Huijbregts, 1978). Mathematically, the ReV would be described as a function that takes values according to their location $\boldsymbol{u}$:

$$ReV = z(\boldsymbol{u}), \qquad \forall \boldsymbol{u} \in D \tag{2.1}$$

Where u represents different locations in the domain D, as the *ReV* is dependent on the location. The *ReV* takes values with some degree of randomness. There could be zones with higher and lower values, e.g., the Cu-sulfides enrichment zone in a porphyry deposit.

The Random Variable (*RV*) takes a value according to a probability distribution (Journel and Huijbregts, 1978). In this way, the *ReV* measured at location $\boldsymbol{u}$ would be the outcome of an *RV* at that point, which is represented by:

$$RV = Z(\boldsymbol{u}), \qquad \forall \boldsymbol{u} \in D \tag{2.2}$$

There are many locations $\boldsymbol{u}$ in the domain *D*, and the *RV* measured at each point $\boldsymbol{u}$ provides a different outcome. This set of *RVs* over a domain *D* is defined as the Random Function (*RF*):

$$\{Z(\boldsymbol{u}), \qquad \forall \boldsymbol{u} \in D\} \tag{2.3}$$

There are two aspects, apparently contradictory, in the *ReV*, the local randomness and the regional spatial correlation. The concept of *RF* addresses both aspects. Considering the *RV* at a location $\boldsymbol{u}$, this is $Z(\boldsymbol{u})$, their probability distribution function would be defined as:

$$F_u(z) = Prob(Z(\boldsymbol{u}) \leq z) \tag{2.4}$$

This formula can be extended to *N* different *RVs*:

$$F_{\boldsymbol{u}_1,\boldsymbol{u}_2\ldots,\boldsymbol{u}_N}(z_1, z_2, \ldots, z_N) = Prob\{Z(\boldsymbol{u}_1) \leq z_1, \ldots, Z(\boldsymbol{u}_N) \leq z_N\}, \qquad \forall u_n \in D \tag{2.5}$$

Where $Z(\boldsymbol{u}_n)$ represents outcomes of *n* different *RVs* at *n* different locations of the domain *D*. This extended probability distribution represents the spatial law of the random function $\boldsymbol{Z(u)}$ (Journel and Huijbregts, 1978).

2.1.2. Assumption of stationarity

The distribution function of the *RF* cannot be inferred using a limited set of samples, but in linear geostatistics, the first two moments of the distribution function are enough to solve most of the problems encountered (Journel and Huijbregts, 1978). The first-order moment is the expected value of the random function $\boldsymbol{Z}(\boldsymbol{u})$, also known as the mean and denoted as *m*:

$$E\{Z(\boldsymbol{u})\} = m(\boldsymbol{u}), \qquad \forall \boldsymbol{u} \in D \tag{2.6}$$

The second-order moments include the variance, covariance and variogram. The variance, denoted as σ^2, is a measure of the dispersion between the samples with respect to the mean. Also, it is equivalent to the square of the standard deviation, SD:

$$\sigma_Z^2 = E\{[Z(\boldsymbol{u}) - m(\boldsymbol{u})]^2\}, \qquad \forall \boldsymbol{u} \in D \tag{2.7}$$

The Covariance, denoted as *C*, is the variability between 02 different *RVs*, over the same domain *D*:

$$C(Z, Y) = E\{[Z(\boldsymbol{u}) - m_Z(\boldsymbol{u})][Y(\boldsymbol{u}) - m_y(\boldsymbol{u})]\}, \qquad \forall \boldsymbol{u} \in D \tag{2.8}$$

The *Variogram*, denoted by γ, is a measure of the variability between two points separated by a vector distance $\boldsymbol{h}$, over the same domain *D*:

$$2\gamma(\boldsymbol{h}) = E\{[Z(\boldsymbol{u}) - Z(\boldsymbol{u}+\boldsymbol{h})]^2\}, \qquad \forall \boldsymbol{u} \in D \tag{2.9}$$

Strict stationarity implies that the distribution function of the *RF* stays the same when a set of *n* samples are translated inside the domain *D*, according to any vector $\boldsymbol{h}$:

$$F_{\boldsymbol{u}_1,\boldsymbol{u}_2,\ldots,\boldsymbol{u}_n}(z_1, z_2, \ldots, z_n) = F_{\boldsymbol{u}_1+\boldsymbol{h},\boldsymbol{u}_2+\boldsymbol{h},\ldots,\boldsymbol{u}_n+\boldsymbol{h}}(z_1, z_2, \ldots, z_n), \qquad \forall \boldsymbol{u} \in D \tag{2.10}$$

Where *n* is a positive integer number, and $\boldsymbol{h}$ a vector of translation inside the domain *D*. This is called *translation invariance*, which means the function produces the same response regardless of the input changes.

Stationarity of second order, is when the mean exists and is constant over domain *D*, not depending on the location $\boldsymbol{u}$.

$$E\{Z(\boldsymbol{u})\} = m, \qquad \forall\, \boldsymbol{u} \in D \tag{2.11}$$

Also, the covariance for any pair of random variables $Z(\boldsymbol{u})$ and $Z(\boldsymbol{u}+h)$ exists and only depends on the vector of separation $\boldsymbol{h}$:

$$C(\boldsymbol{h}) = E\{[Z(\boldsymbol{u}) - m][Z(\boldsymbol{u}+\boldsymbol{h}) - m)]\}, \qquad \forall\, \boldsymbol{u} \in D \tag{2.12}$$

The existence and stationarity of the covariance imply the same for the variance and variogram. The next equations are developed from this premise:

$$\sigma_Z^2 = E\{[Z(\boldsymbol{u}) - m(\boldsymbol{u})]^2\} = C(0), \qquad \forall\, \boldsymbol{u} \in D \tag{2.13}$$

$$\gamma(\boldsymbol{h}) = \frac{1}{2}E\{[Z(\boldsymbol{u}) - Z(\boldsymbol{u}+\boldsymbol{h})]^2\} = C(0) - C(\boldsymbol{h}), \qquad \forall\, \boldsymbol{u} \in D \tag{2.14}$$

In most cases, the stationarity of second order is enough for resource estimation, as it permits the variogram calculation and kriging. The stationarity of a domain is one of the first decisions in the resource estimation process. As it is made over a physical domain, it can be influenced by the number of samples and geological considerations. Analysis and calculations like EDA, variograms, kriging, etc. are made under the assumption that the domains are stationary. If during these processes or others, it is detected some deviation or non-homogeneity, it is possible to redefine the domains.

2.1.3. Data acquisition and error

Drilling data is the main source of information for the subsurface mineral deposit. They provide geological information like rock type, alteration, mineralization, and mineral grades. Although there are multiple types of drilling

data, this research considers diamond drilling (DD) and blast holes (BH) data. Both data types have different drilling conditions and features that condition their levels of error and quality. For this research DD is considered primary data and BH secondary, a classification to indicate which data presents a higher level of quality and confidence.

DD and BH are drilled for different purposes. They have different spacings and drilling depths, determining the number of samples. DD is used in exploration and infill stages, on a wide-spaced grid the entire depth of the deposit. This type of data is useful for the estimation of the whole deposit. BH is used in rock blasting and provides information for the ore control process. BH is drilled for specific zones using a dense drilling grid of just a few meters. The shallow BH data are for the bench in production. BH data is valid only for local estimation of ore control models.

Error refers to the difference between the true and the sample. A constant and accumulative sample error can lead to the bias of the whole dataset. In the case of BH and DD, the error can have multiple sources, mainly during the drilling process and the preparation of the samples. There could be contamination, material losses, oxidation/reduction, leaching, mislabeling, etc. The BHs have more potential for error due to drilling and sampling conditions, besides the large number of samples processed, from tens to hundreds, just in a day. DD has a stricter drilling and sampling protocol, besides chemical analysis is usually done by external labs, which minimizes error. Due to these conditions, DD is considered unbiased, while BH data is biased concerning the DD data. The BH bias degree is quite variable, e.g. in Cu porphyry deposits, it has observed bias from +5% until +25%. These values are not necessarily steady, can vary according to the material type, mineral grades and sampling protocol. Thus, DD is considered as primary data and BH as secondary data. Table 2.1 summarizes the differences between both drilling data types.

Table 2.1: Drilling type features

	DD	**BH**
Grid spacing	Tens of meters	>10 m
Depth	>100 m	1 bench
Assays	External labs	Local lab
Data error	<5%	5-50%
Bias	~1.0%	~1.05%-1.25%
Volume support (vol/m)	0.003-0.006 m^3	0.126 m^3

As access to a real multi-type database is quite limited, data simulation is of common use. For this research, a simulation methodology is developed that includes the replication of drilling features and sampling protocols, resulting in a closer approximation to real drilling data.

2.1.4. Data analysis and treatment

Data are submitted to an exploratory data analysis (EDA) as a key step of any resource estimation process. In this stage, the data is grouped into estimation domains based on geological and statistical features. The most usual domains are based on rock type, alteration, ore type, or even a combination of these features. Each domain must be statistically analyzed and described.

The data samples should have the same sample support. This is accomplished through compositing to an equal length. Concerns arise when there are drilling data with different diameters. Data treatment includes declustering, managing outliers, and joining or separating domains. Statistical analysis may reveal soft and hard contacts, domains with more than one population, or a non-stationarity behaviour.

2.1.5. Spatial variability analysis

Variograms describe the spatial variability of a regionalized variable and are required for estimation methods based on kriging techniques. The experimental variogram is calculated over the estimation domain using drilling data. The semi-variogram is defined as:

$$\gamma(\boldsymbol{h}) = \frac{1}{2N}\sum_{i=1}^{N}[Z(\boldsymbol{u}_i) - Z(\boldsymbol{u}_i + h)]^2\,, \qquad \forall \boldsymbol{u} \in D \tag{2.15}$$

Where N is the number of pairs, $\boldsymbol{u_i}$ are the different locations inside the domain D, and h is the lag vector and distance between the pairs.

As samples are not separated by a fixed distance, and the drilling grid is far from being regular, it is necessary to apply tolerances for the lag distance and the searching angles. The unit lag distance, $\boldsymbol{h}$, is selected according to the average distance between samples and can change depending on the direction of the variogram, e.g., in the vertical variogram, the distance between samples is just a few meters. In contrast, the horizontal distance between drill holes can reach tens or hundreds of meters. The lag tolerance is usually half of the lag distance. Multiple lags of increasing distance in the principal directions are considered.

The variogram must be oriented to the direction of major continuity. This direction can be linked to a geological structure, like a fault or a vein. This direction is defined by two angles, the azimuth and dip. During the variogram calculation, the azimuth and dip are usually fixed with a span of 45º, and tolerances of 22.5º (Deutsch, J., 2015). It is advised to calculate variograms in several directions and choose the principal direction. The use of omnidirectional variograms brings a solution when the anisotropy directions are not clear. This consists of fixing the angles tolerances to include all the sample pairs in all directions for the variogram calculation. This results in an "averaged" experimental variogram for all directions, able for estimation process, although it would not reflect the exact spatial variability of the domain.

The experimental variogram is used to model the spatial variability related to the domain. The variogram model must be defined in 3D, although there are domains that can be defined in 2D, like tabular deposits. The principal major, semi, and minor directions must be orthogonal. If the model has different ranges for each direction, the model is anisotropic. The model is isotropic when all directions have the same range. Variogram models include the combination of several structures, perhaps of different types. A usual variogram model combines at least two structures or models. Figure 2.1 shows an experimental variogram for a DD dataset and the corresponding model. The model includes the nugget effect model and two spherical models. The primary use of the variogram model is to be used for estimation using kriging techniques.

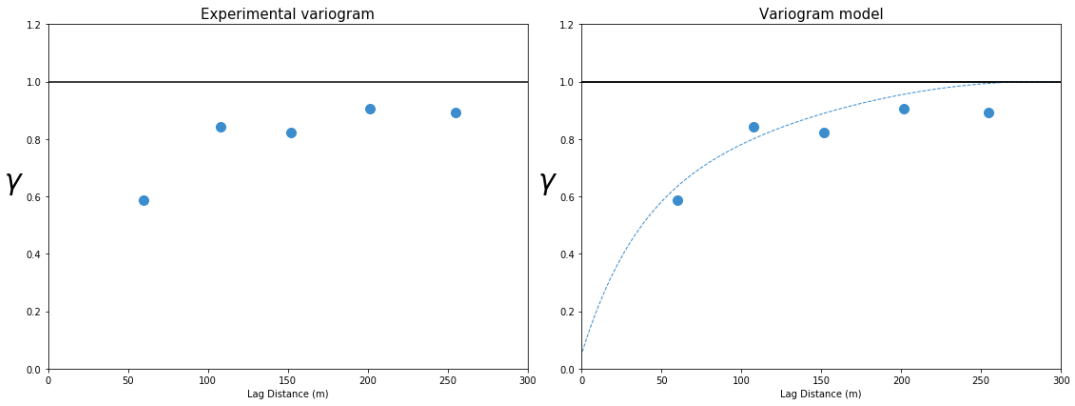


Figure 2.1: a) Experimental variogram and, b) Variogram model

2.1.6. Block models

Estimation results are stored in a 3-dimensional model, which represents the volume of the deposit in evaluation, and is generically called a block model. This is a

geometric arrangement composed of orthogonal blocks representing a specific volume and location in space. The block model must cover the whole deposit and include all the estimation domains. The dimensions of the individual blocks are defined according to the purposes of the model, and they can take from a few to tens of meters in each dimension. Additionally, there could be sub-blocking; this is a minor block length that usually takes place on the borders between domains to improve contact definition. The block model can be rotated to contain the mineral deposit more efficiently.

RESOURCES MODEL FOR PLANNING AND PRODUCTION

In mining operations, the resource models are designed to fulfill the needs of planning and production, in different time periods. Usually, they are divided into long-term, intermediate, and short-term models.

The long-term model (LTM) is used for long periods of scheduling and planning, usually from 12 to 36/48 months. These models are often built using only DD data and a large block size covering the entire mineral deposit. DD data is widely spaced, particularly in the first stages of the mining operation, although this spacing is reduced with successive infill drilling stages. The selective mining unit (SMU) block size is the minimum volume for mineral classification and usually takes several tens of meters. The LTM is updated at least once a year with new DD data collected during the year. This configuration of data and parameters results in a model with grades and tonnages that are globally accurate and suitable for long-term planning but not good enough for local assessment.

The short-term model (STM), also known as the ore control model, is used for daily planning and mining operations. It is often estimated using exclusively BH data drilled in the area scheduled for production and on one bench. The block size is small to obtain a high-resolution model, so it is normal to use lengths of a few meters. The smaller block size produces higher resolution and grades, thereby a better profit. Also, a high-resolution model permits a better definition of dig limits. The estimation is done locally, in the volume defined for production. The results are collected in the STM, which, although occupying the entire mineral deposit, has only estimation results in the zones previously drilled.

Long and short-term models are designed to accomplish different tasks and are not locally comparable. The main issue is the LTM, which is not locally accurate for intermediate and short-term planning. Therefore, it is necessary to use medium

term models (MTM) that provide accurate grades and tonnages for both local and global assessment, besides an acceptable model resolution and updated information. MTM is built using both data types available in the mining operation, DD and BH, and an intermediate size of blocks. There are estimation schemes that use both data without any treatment, which is not recommended considering the differences between DD and BH data types. A cokriging outline, with DD as primary and BH as secondary data, is recommended option.

ESTIMATION GRID SIZE FOR RESOURCE MODELS

The grid size of estimation refers to the distance between each point to estimate. It is commonly confused with the block size, but we must remark that the block size refers to a volume physically occupied by a material, and although both have the same magnitude they must not be considered as synonyms. For this research, we will use the term estimation grid size instead of block size. The grid size varies according to the model, LTM uses the SMU dimensions or similar ones but always above tens of meters, while STM uses grid sizes based on the spacing between BHs, usually a few meters. In the case of MTM, the most advisable is to use a mid-grid size, but closer to the STM grid size, because the BH is the most abundant data type on this model. The grid size spacing (GSS) is the ratio between the estimation grid size and data spacing (Vasylchuk, 2016). GSS is important for STM, where it is necessary to have high resolution. Previous research suggests that a GSS of 0.25 minimizes estimation errors within an acceptable processing time, and above this threshold, the errors are more unpredictable. Figure 2.2 shows a typical BH drilling grid, and the suggested grid size is a quarter of the distance between samples.

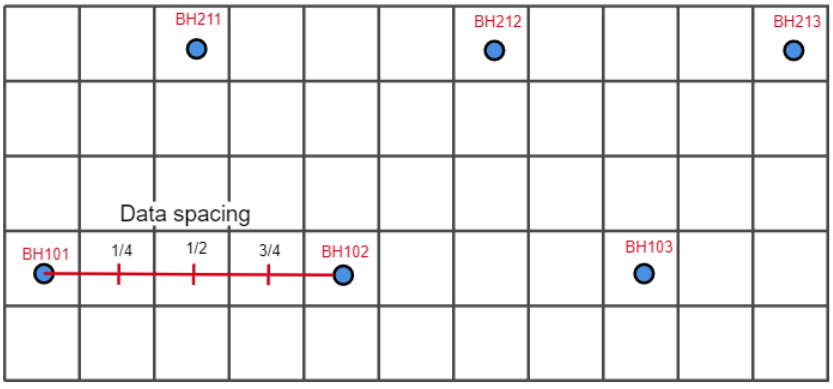


Figure 2.2: Blasthole grid and block grid at ¼ of data spacing for STM.

2.1.7. Ore/waste classification and profits

One of the main steps in ore control is the ore/waste classification. This is made according to the block mineral grades and cut-offs. The blocks are classified as ore or waste, and according to it, their destination can be the processing plant or the waste dump. The classification is critical, as this will condition the block profit. There are several cut-off grades, but the marginal and the break-even cut-off are used for this research. The marginal cut-off is widely used in the ore control process and classifies the blocks between ore or waste. This cut-off is defined using only the metallurgical costs, without considering the mining costs, which are considered sunk costs. This is the typical case of open-pit mining operations, where all blocks are mined, no matter the destination. It is important to remark that this classification does not mean that the ore blocks are necessarily profitable. Other factors needed to define the cut-off are the metallurgical recovery and the mineral price.

$$Cutoff_{Marginal} = \frac{Processing\ Cost}{Recovery \cdot Mineral\ Price} \quad (2.16)$$

The break-even cut-off, also known as economic cut-off, considers mining and processing costs. This ensures that all blocks classified as ore have a revenue equal to or above all costs, meaning that blocks are profitable.

$$Cutoff_{Break\ even} = \frac{Processing\ Cost + Mining\ Cost}{Recovery \cdot Mineral\ Price} \quad (2.17)$$

Roughly, the block profit is the difference between the revenues by selling its mineral content, and the costs of mining and processing it. The revenue is calculated by multiplying the mineral content by the mineral price. The mineral content is defined by the block tonnage, grades, and metallurgical recovery. As the waste blocks are not processed, they do not have any revenue, just costs. The model profit is calculated by adding the positive incomes from the ore blocks and the negative income from the waste.

$$Profit_{Ore} = Revenues - (Processing\ Cost + Mining\ Cost) \quad (2.18)$$

$$Profit_{Waste} = -Mining\ Cost \quad (2.19)$$

$$Revenue_{block} = Mineral\ price \cdot Tonnage \cdot Mineral\ grade \cdot Recovery \quad (2.20)$$

Data simulation opens the opportunity to a real classification of the blocks and permits an assessment of the accuracy of an estimation scheme. Ore/waste

classification with two classes has four possible scenarios according to the real ore/waste classification:

1. **Correct acceptance:** when both classifications, estimated and real coincide, and correspond to the *Ore* class.
2. **Correct rejection:** both classifications, estimated and real coincide. In this case, both classes are *Waste*.
3. **False acceptance:** when the classes do not coincide, and the estimated class is *Ore*. This is the *Dilution* scenario and represents a type I error.
4. **False rejection:** when both classes do not coincide, and the estimated class corresponds to *Waste*. This is the *Ore loss* situation and represent a type II error.

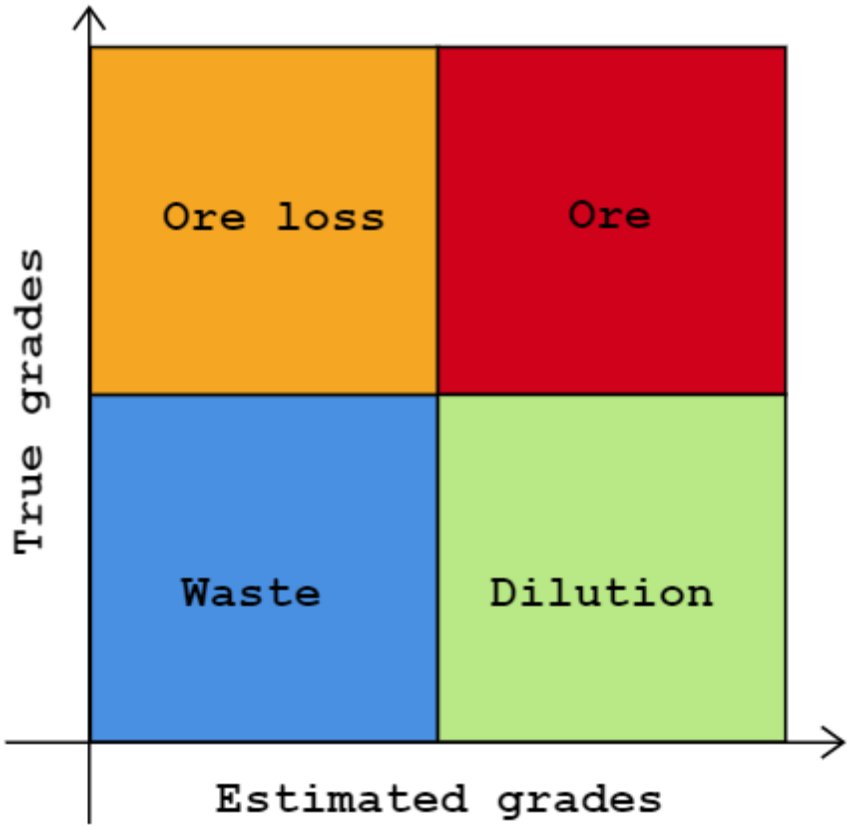


Figure 2.3: Ore/waste classifications

The misclassifications, ore loss and dilution, have a different impact on the model profit. The ore loss generates a double negative impact, the cost of mining as waste and a lost opportunity by not recovering the true mineral content. Depending on the block mineral content, the total loss can be significant. On the other hand, a dilution block can generate some revenues, but less than the minimum necessary to

make it profitable. Block grades below the marginal cut-off generate higher losses. These losses cannot surpass the sum of mining and processing costs, which is the maximum loss of a dilution block. It is common to assign a lower recovery to this class, considering the grades are low and hard to be recovered. Figure 2.4 shows the profit function built using the cut-off and mining parameters. It shows the profits and losses according to the real ore/waste classification. In the cases of correct classification, the profit and loss follow the expected curves (solid lines), but the misclassification (dotted lines) only produces losses. The ore loss (blue dots) is the most harmful because its magnitude increases with grades, so is theoretically unlimited. The dilution (green dots) has less loss, bounded between the mining and processing costs.

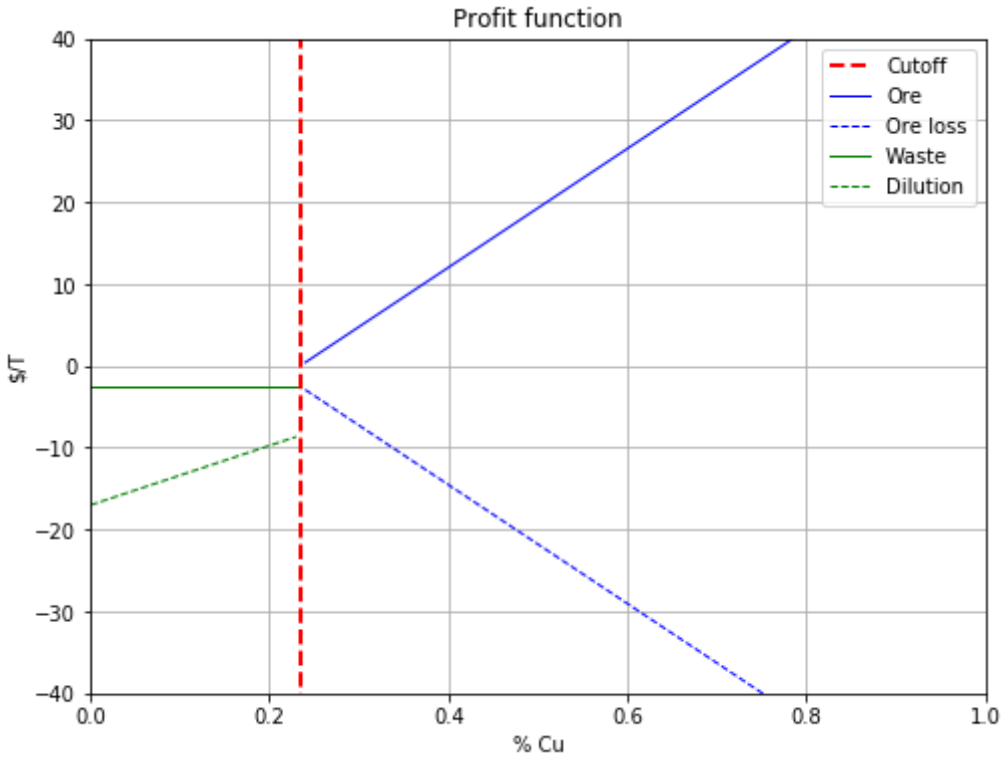


Figure 2.4: Profit function

2.1.8. Estimation techniques

Estimation is the main step in the resource estimation process. There are several methods of estimation, some simple, like polygonal estimation, and others more elaborate, like kriging methods. The user must define which method to use according to their data and objectives. In ore control, due to the big database, an inverse distance or nearest neighbour estimation can provide acceptable results.

Kriging is one of the most used estimation methods, as it ensures the lowest estimation error.

NEAREST NEIGHBOR OR POLYGONAL METHOD

Nearest neighbor is one of the simplest methods, and it is based on polygonal estimation. The grade of a block is the same as the closest sample to the block centroid. This results in a grades model without smoothing, reflecting the sampling grades. It is applicable to model reconciliation and checking.

INVERSE DISTANCE

This is a linear estimation technique that assigns weights to the samples involved in the estimation of the block. The weights are the distance from the sample to the block centroid, inverse and powered. This technique gives more weight to the close samples and less to the farther samples.

$$Z(\boldsymbol{u})^* = \frac{\sum_{i=1}^{n} \frac{z(\boldsymbol{u}_i)}{d_i^p}}{\sum_{i=1}^{n} \frac{1}{d_i^p}} \qquad (2.21)$$

Where n is the number of samples, d is the distance to the sample, and p is the power to use. The user defines the power, although the most common are 2,3 and 5. As the power increases, the farther samples took lower weights. Unlike NN, this method provides smoothed grades, as it uses several samples to estimate a single block. The estimation setup considers minimum and maximum samples, block discretization, limit samples by quadrant, and search ellipsoids. This method honours the estimation domains but does not consider the spatial variability of the domain (variograms) beyond using the anisotropy orientations for the search ellipsoid and distance calculation.

KRIGING

Kriging is an estimation technique that works similarly to inverse distance, assigning estimation weights to the samples involved in the block estimation. The difference is that the weights are based on the covariance between the samples and the block, for this is necessary the variogram model. These are calculated using the variogram and the relation between variogram/covariance. Kriging estimation minimizes the error variance and provides smooth grade models. The most used

kriging methods are simple and ordinary kriging, although ordinary kriging is more stable to variations in stationarity.

1. Simple kriging (SK) requires the stationary mean, calculated using the available samples in the domain *D*.

$$Z_{SK}^{*}(\boldsymbol{u}) = \sum_{i=1}^{n} \lambda_i \cdot [Z(\boldsymbol{u}_i) - m_z] + m_z, \qquad \forall i = 1,2 \dots n \tag{2.22}$$

n is the number of data used to estimate. By minimizing the estimation error, the kriging equations system is developed. The kriging weights are functions of the covariances between the estimated location and the data, and they are not constrained.

$$\sum_{j=1}^{n} \lambda_j \cdot C(\boldsymbol{u}_i, \boldsymbol{u}_j) = C(\boldsymbol{u}, \boldsymbol{u}_i), \qquad \forall j = 1,2 \dots n \tag{2.23}$$

The SK variance is derived from the system of equations:

$$\sigma_{SK}^2 = C(0) - \sum_{i=1}^{n} \lambda_i \cdot C(u, \boldsymbol{u}_i), \qquad \forall i = 1,2 \dots n \tag{2.24}$$

2. Ordinary kriging (OK) is based on the same principles than SK, but do not require the mean. Besides, the OK weights are constrained to sum 1.

$$Z_{OK}^{*}(\boldsymbol{u}) = \sum_{i=1}^{n} \lambda_i \cdot Z(\boldsymbol{u}_i) + \left[1 - \sum_{i=1}^{n} \lambda_i\right] \cdot m_z, \qquad \forall i = 1,2 \dots n \tag{2.25}$$

The OK system equations and variance are defined as:

$$\sum_{j=1}^{n} \lambda_j \cdot C(v_i, v_j) + \mu = \bar{C}(\boldsymbol{V}, \boldsymbol{v_i}), \qquad \forall j = 1,2 \dots n$$

$$\sum_{j=1}^{n} \lambda_j = 1 \tag{2.26}$$

$$\sigma_{OK}^2 = \bar{C}(\boldsymbol{V}, \boldsymbol{V}) - \mu - \sum_{i=1}^{n} \lambda_i \cdot \bar{C}(V, v_i), \qquad \forall i = 1,2 \dots n$$

The kriging estimation setup for both kriging techniques requires the variogram model, min and max data, discretization, and search ellipsoid. This is built using the anisotropy directions and ranges, and depending on its size,

it could be necessary to make consecutive runs to fill the block model. This is more common on LTM estimation, while in STM, often is completed in one-step estimation.

2.2.Simulation

Estimation is a process to produce a single set of values, calculated from a dataset for a particular domain. Simulation uses the same dataset and domain, to reproduce the distribution of the data, honouring its statistics and spatial variability. In this way, simulation produces several sets of values or realizations, each one with the same statistical properties. Simulation is used to model distribution uncertainty, which is necessary for different tasks like data spacing analysis, metallurgical recovery, or profits assessment. In research, simulation is vastly used, to reproduce the conditions and possible outcomes of distribution, and to provide synthetic data. There are different simulation techniques like turning bands, moving average or sequential gaussian. For this research, sequential Gaussian simulation has been applied.

2.2.1. Sequential Gaussian Simulation (SGS)

The SGS algorithm (Isaaks, 1990) is based on the Gaussian conditional distribution at each point to simulate. This conditional distribution is defined by the mean and variance of that point, calculated using simple kriging. From this distribution, a random percentile is drawn. It is good practice to consider one hundred or more realizations. The resulting realizations, which are in gaussian units, must be back-transformed into regular units. SGS can be conditioned to a dataset or not. In the case of conditioned simulation, the process considers the data, honouring its distribution and spatial variability. The SGS algorithm considers the following steps:

1. Apply NS transformation to the data.
2. Variogram calculation and modeling.
3. Follow a random path through the grid.
4. By simple kriging and using the conditioning dataset, estimate the value $m(\boldsymbol{u})$ and kriging variance $\sigma_{sk}(\boldsymbol{u})$ for the point in evaluation. They will be the mean and variance of the conditional gaussian distribution (CCDF).
5. Take a random value between 0 and 1 that represents a percentile of the CCDF and by back-transform get the simulated value $Z_s(\boldsymbol{u})$.

6. The simulated value $Z_s(\boldsymbol{u})$ is added to the conditioning dataset.
7. Repeat steps 4 to 6 until complete the simulation of all points.
8. Back-transform the estimation results.

As all realizations have the same probability, one single realization is not enough to model the uncertainty. It is recommended to consider at least 25 realizations.

2.2.2. Drilling data simulation

The access to real drilling data is quite limited, besides for a deep analysis we require exhaustive data. Simulation is an alternative to real data. Either by conditional or unconditional simulation, it is possible to simulate a distribution with specific features according to the research. The usual drilling data simulation consists of the simulation of a reference true distribution, then it is sampled and can be added a reasonable error and bias to simulate primary and secondary data. Simulated data can be used as any real data, in the estimation processes, model validation, profit evaluation, etc. The use of simulated data lets us know the true mineral distribution, so it is possible to measure error and bias directly and confirm the validity of research results.

2.3. Estimation with multiple data sources

Most of the estimation methods consider using a single drilling data type for the estimation. The use of multiple drilling data types in a single dataset is not recommended, as each drilling data type has its own set of statistical features, errors, and support. Although there are situations where multiple datasets could be combined, this is viable only if both datasets inform about the same variable, present similar distribution and statistics, high correlation, and after a detailed data analysis. Cokriging is a technique that permits the estimation of a point of interest using data from different attributes. This requires primary data with the same attribute as the variable to estimate and secondary data with different attributes, e.g., Au estimation of a block but using Au and Ag samples. Both datasets must be available in the neighbourhood of the estimation, and there must be a spatial correlation between them. The application of cokriging in resources estimation requires a somewhat complicated setup that has restricted its widespread use in the industry. The most optimal use would be the estimation of the grade using DD and BH data which inform the grades with different error and bias. Cokriging requires modelling the joint spatial

correlation between the data involved. This correlation can be modelled through the linear model of coregionalization.

2.3.1. Linear model of coregionalization (LMC)

The linear model of regionalization (LMR) describes the spatial continuity of a variable as the combination of nested structures. The LMC is the natural extension of LMR, but for multiple variables. As the LMR, the LMC assumes a regionalized variable as the sum of multiple independent factors:

$$Z_k(\boldsymbol{u}) = m_k(\boldsymbol{u}) + \sum_{i=0}^{nst} a_{k,k'}^{i} Y_i(\boldsymbol{u}) \qquad \begin{matrix} k = 1,2,..K \\ i = 1,2,..nst \\ E\{Y_i(\boldsymbol{u})\} = 0 \\ E\left\{\left(Y_i(\boldsymbol{u})\right)^2\right\} = 1 \\ E\{Y_i(\boldsymbol{u})Y_i(\boldsymbol{u}')\} = 0 \end{matrix} \qquad (2.28)$$

There are *K* variables, *nst* defines the number of structures or factors, a_k^i is the contribution for each structure, $Y_i(\boldsymbol{u})$ the independent factors, and $m_k(\boldsymbol{u})$ the mean of the *K* variable. The O^{th} factor represents the nugget effect contribution. The mean of each factor is zero and the variance 1, as they are standard variables. The covariance between *K* variables will be 0, as they are independent, for all $\boldsymbol{u} \neq \boldsymbol{u'}$.

The LMC is a linear combination of structures (Γ_i) weighted by its contributions (b_k^i) to the total variance:

$$Y_{k,k'}(\boldsymbol{h}) = \sum_{i=0}^{nst} b_{k,k'}^{i} \cdot \Gamma_i(\boldsymbol{h}) \qquad \begin{matrix} k = 1,2,..K \\ b_{k,k'}^{i} = a_k^i a_{k'}^i \\ \Gamma_i, i = 1,2..nst \end{matrix} \qquad (2.29)$$

LMC is used to model the spatial variability between 2 or more variables and is applicable to the estimation by cokriging. For the case of 2 variables, Z_i and Y_i, the equation system would be (Deutsch, 2002):

$$\begin{aligned} Y_{Z,Z}(\boldsymbol{h}) &= b_{Z,Z}^0 + b_{Z,Z}^1 \cdot \Gamma^1(\boldsymbol{h}) + b_{Z,Z}^2 \cdot \Gamma^2(\boldsymbol{h}) + \cdots \\ Y_{Z,Y}(\boldsymbol{h}) &= b_{Z,Y}^0 + b_{Z,Y}^1 \cdot \Gamma^1(\boldsymbol{h}) + b_{Z,Y}^2 \cdot \Gamma^2(\boldsymbol{h}) + \cdots \\ Y_{Y,Y}(\boldsymbol{h}) &= b_{Z,Z}^0 + b_{Y,Y}^1 \cdot \Gamma^1(\boldsymbol{h}) + b_{Y,Y}^2 \cdot \Gamma^2(\boldsymbol{h}) + \cdots \end{aligned} \qquad (2.30)$$

The LMC models the direct and cross-variograms using the same structures (Γ_i), the contributions (b_k^i) can be variable but their sum must be equal to the total variance. To ensure the model be positive definite the next rules must be followed:

$$b_{ZZ}^{i} \,\&\, b_{YY}^{i} \geq 0$$
$$b_{ZZ}^{i} \cdot b_{YY}^{i} - b_{ZY}^{i} \cdot b_{YZ}^{i} \geq 0 \qquad (2.31)$$

The contributions of the direct variograms must be positive and larger than the cross-variogram contributions. The cross-variogram $\gamma_{Z,Y}(\boldsymbol{h})$ must be defined to fit a valid LMC, but this calculation requires collocated data, which is unusual with different data types. In most of the mining scenarios these can be calculated from the cross-covariance:

$$\gamma_{ZY}(\boldsymbol{h}) = C_{zy}(0) - C_{ZY}(\boldsymbol{h}) \qquad (2.32)$$

The cross-covariance at h=o, is calculated by extrapolating the cross-covariance and using simple calculation obtain the cross-variogram.

2.3.2. Cokriging with full LMC

Cokriging estimates a variable using data from the same variable and from an auxiliary one. There must be a spatial correlation between both variables so that the second variable provides useful information for the estimation. The cokriging technique must be selected according to the primary and secondary data conditions. Minnitt & Deutsch (2014) suggest the best performance of cokriging when both data types are unequally sampled and have different grades of quality and reliability. In a mining operation, DD and BH provide values from the same variable and have an evident spatial correlation between them but differ in the quality assay and the number of samples. Both drilling types are in different grids with different spacing and orientations. Collocated cokriging requires collocated data, this is primary and secondary data measured in the same location. This condition is uncommon, so its application is very limited. In a mining operation, it is more common to have data on different grids, so the cokriging techniques most used are simple cokriging (SCK) and ordinary cokriging (OCK).

2.3.3. Simple cokriging (SCK)

SCK estimates a variable using the primary and secondary datasets in a weighted linear combination derived from the simple kriging equation:

$$Z_{SCK}^{*}(\boldsymbol{u}) - m_z = \sum_{\alpha=1}^{n} \lambda_\alpha(\boldsymbol{u}) \cdot [Z(\boldsymbol{u}_\alpha) - m_Z] + \sum_{\beta=1}^{r} \theta_\beta(\boldsymbol{u}) \cdot [Y(\boldsymbol{u}_\beta) - m_Y] \qquad (33)$$

This is like simple kriging, except for the inclusion of weights (θ_β) and data (Y) from the secondary dataset. m_z and m_y are the stationary means of each variable. In a similar way, the variance is defined by:

$$\sigma^2_{SCK}(\boldsymbol{u}) = C_{ZZ}(0) - \sum_{\alpha=1}^{n} \lambda_\alpha(\boldsymbol{u}) \cdot C_{ZZ}(\boldsymbol{u}_\alpha - \boldsymbol{u}) - \sum_{\beta=1}^{r} \theta_\beta(\boldsymbol{u}) \cdot C_{ZY}(\boldsymbol{u}_\beta - \boldsymbol{u}) \quad (34)$$

The weights λ_α and λ_β are determined by minimizing the estimation variances. The covariances matrix to solve the weights is:

$$\begin{bmatrix} C_{ZZ} & \dots & C_{ZY} \\ \vdots & \ddots & \vdots \\ C_{YZ} & \dots & C_{YY} \end{bmatrix} \begin{bmatrix} \lambda_\alpha \\ \vdots \\ \lambda_\beta \end{bmatrix} = \begin{bmatrix} C_{Z^*} \\ \vdots \\ C_{Y^*} \end{bmatrix} \quad (35)$$

C_{ZZ} and C_{YY} are the direct covariances for each data type, and C_{ZY} and C_{YZ} the cross-covariances between both data types, λ_α and λ_β the data weights, and C_{Z*} and C_{Y*} the covariances between the data and the estimation location. Cokriging requires the covariance matrix be positive definite. Ordinary cokriging (OCK).

2.3.4. Ordinary cokriging (COK)

The ordinary cokriging formula is pretty like ordinary kriging one, but adding the secondary variable Y:

$$Z^*_{COK}(\boldsymbol{u}) = \sum_{\alpha=1}^{n} \lambda_\alpha(\boldsymbol{u}) \cdot Z(\boldsymbol{u}_\alpha) + \sum_{\beta=1}^{r} \lambda_\beta(\boldsymbol{u}) \cdot Y(\boldsymbol{u}_\beta) \quad (2.36)$$

As in the previous SCK, the system of equations requires direct covariances and cross-covariances to solve the weights, and a positive definite covariance matrix. To solve the equations, it is necessary to solve ensuring that the sum of weights be equal to 1, but as there are two sets of weights, this becomes hard to solve. There are different methods to solve:

TRADITIONAL ORDINARY COKRIGING

In this method, the sum of weights is applied only to the primary variable, making the secondary weights be equal to zero:

$$\sum_{\alpha=1}^{n} \lambda_\alpha = 1 \;\&\; \sum_{\beta=1}^{r} \lambda_\beta = 0 \quad (2.37)$$

The main issue with this method, is that the weight for secondary data is reduced to minimum, reducing its importance and contribution in the estimation process.

STANDARDIZED ORDINARY COKRIGING (SOCK)

SOCK considers standardized the secondary variable Y, in that way both variables would have the same mean, besides constraining the sum of weights to 1.

$$\frac{Z^*_{SOCK}(\boldsymbol{u}) - m_Z}{\sigma_Z} = \sum_{\alpha=1}^{n} \lambda_\alpha(\boldsymbol{u}) \cdot \left[\frac{Z(\boldsymbol{u}_\alpha) - m_Z}{\sigma_Z}\right] + \sum_{\beta=1}^{r} \lambda_\beta(\boldsymbol{u}) \cdot \left[\frac{Y(\boldsymbol{u}_\alpha) - m_Y}{\sigma_Y}\right] \quad (2.38)$$

$$\sum_{\alpha=1}^{n} \lambda_\alpha + \sum_{\beta=1}^{r} \lambda_\beta = 1 \quad (2.39)$$

The user must carefully analyze its data and decide which technique is the most appropriate to use. For this research, standardized ordinary cokriging (SOCK) has been employed.

3. Multi type drill hole simulation

Data simulation consists of the simulation of a reference spatial distribution and the sampling of different drilling types. The drilling procedure, volume support and sampling protocol must be considered. The DH simulation outline developed considers a procedure, where instead of the simulation of a whole deposit, it is simulated a set of drill holes. The DH distribution is simulated using a point-scale grid where each node corresponds to a single and pure mineral species. The simulation is limited to a volume that includes the largest drilling diameter and repeated until obtaining the required number of DH. All simulated DH must be conditioned to the same reference distribution. Finally, data are obtained from a specific drilling type by applying the drilling support and sampling protocols. Different degrees of error and bias are considered.

3.1.Simulation procedure

The entire process has been prepared in a python folder, using GSLIB executables and calculation tools. The procedure has been separated into three stages:

1. General setup and reference distribution simulation.
2. Drill hole simulation.
3. DH compositing and datasets

The simulation takes place in a scenario of a copper porphyry, specifically the potassic domain, with mineralization of chalcopyrite (Cpy) and bornite (Bo) in a ratio of 4:1, the average grade of the deposit is 1% Cu. The work area involves a single bench, although, for some cases studies, it has been extended to five benches.

3.1.1. Reference distribution simulation

This stage includes the general setup and the simulation of the reference distribution used as the conditional dataset for further stages.

1. **Work area:** The work area has an extension of 400 by 300 m with a bench height of 10 m.
2. **Grid size:** The simulation grid size is 0.25 m on each axis, resulting in a distribution of 76.8 million points.

3. **Input variogram:** The reference distribution is an unconditional simulation, not having conditioning data, only input variogram. The variogram used considers seven nested structures at different ranges, from mm to hundreds of meters. This is necessary as the input variogram must ensure variability at very short ranges, including the millimetre and centimetre scales to represent different drilling diameters. Setting up the list of structures has been a complicated work that has required multiple tests, tried different numbers and types of structures, ranges and contributions, until obtaining an acceptable result. Table 3.1 summarizes the list of structures selected.

Table 3.1: input variogram parameters

NST	Type	CC	Ranges	Angles
0	-	0.05	-	-
1	Sph	0.15	0.02	0º
2	Sph	0.15	0.05	0º
3	Sph	0.10	1.00	0º
4	Sph	0.10	2.50	0º
5	Sph	0.05	10.0	0º
6	Sph	0.20	120.0	0º
7	Sph	0.20	240.0	0º

Figure 3.1: Input variogram

4. **Simulation:** it is done using the USGSIM tool from the GSLIB library. A single realization is created with 25 previously simulated samples to condition each node. It is important to set up an adequate number of samples to conditioning, as a low number does not guarantee a correct simulation and a high value will increase the processing time with no better results.
5. **Tri-mineral distribution:** The simulation results, in Gaussian units, must be transformed into a tri-mineral distribution (Cpy-Bo-Gangue). The transformation is done by Cu cut-offs to ensure an average distribution grade of 1.0 % Cu. Each node is assigned to a single mineral species and its Cu content. Table 3.2 shows the cut-offs for assigning mineralization.

Table 3.2: Cu mineral species

Mineral specie	Cut-off	Cu
Gangue	=<0.955	0.0%
Cpy	>0.975	35.0%
Bo	>0.955	63.3%

Figure 3.2 shows the histograms for the NS distribution and the tri-mineral distribution. The first plot shows the expected shape of a gaussian distribution. The second one shows the shape for a log-normal distribution with a very high content that correspond to gangue and low content to minerals. This is a realistic mineral composition, where the gangue compounds most of the rock mass and the minerals just a fraction, most of the time below 5%.

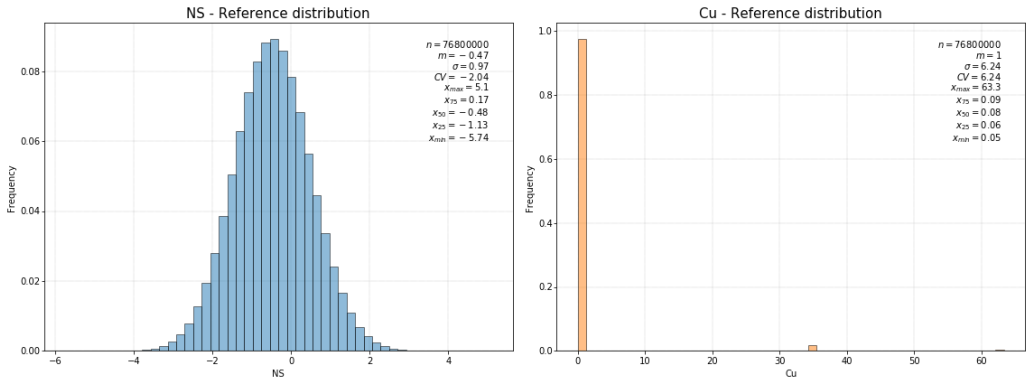


Figure 3.2: Reference distribution histograms in a) NS units and b) Tri-mineral units

As a final step, the reference distribution is resized for use in further evaluations. The original grid size of 0.25x0.25x0.25 m is changed to 1x1x10 m, and 5x5x10 m. The grades for the new block sizes are calculated using volume average. The resized distributions will be used in the case studies. Figure 3.3 shows the Cu grades map for the original grid size and the resized grids.

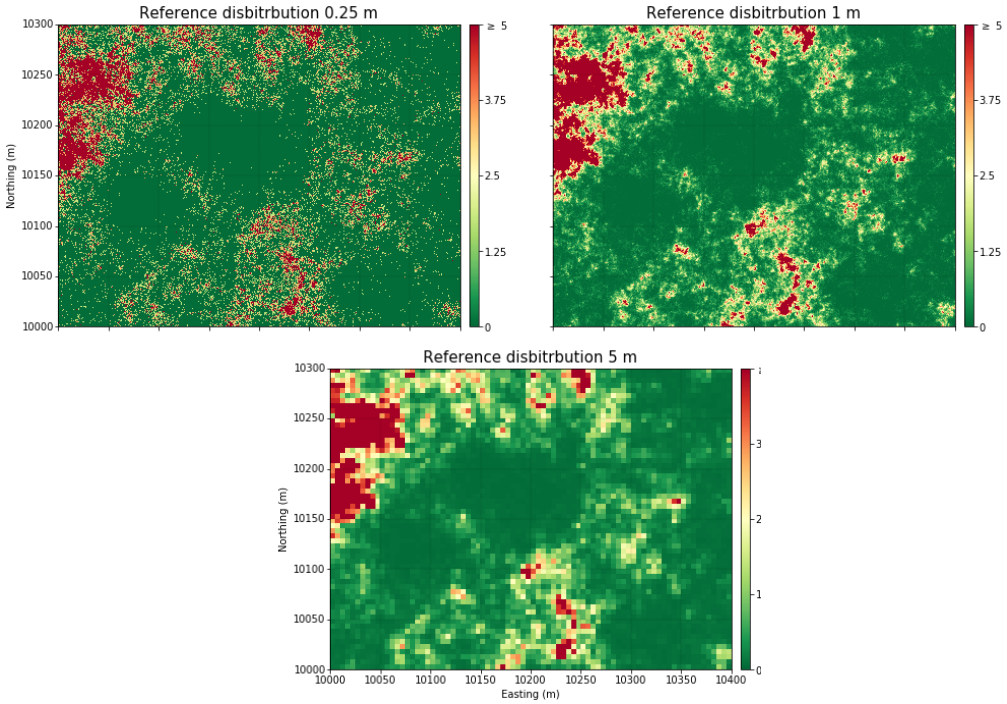


Figure 3.3: Reference distribution with different block sizes a) original grid size, b) 1 m, and c) 5 m.

3.1.2. DH simulation

The second stage involves the simulation of drill holes and includes the steps: DH grid definition, conditioning dataset preparation, DH simulation, validation, tri-mineral distribution, and discretization.

1. **DH grids:** Each drilling type, DD and BH, has a specific drilling grid. DD uses an orthogonal grid of 50 m. The DD collars are randomly moved inside a radius of 5 m concerning the initial position, which brings a realistic degree to the grid. BH drilling uses a spacing of 10 by 12 m and a triangular grid with an orientation of N15W. The BH collars have been randomly moved in a radius of 0.5 m. The difference in drilling spacing significantly impacts the number of drill holes. Figure 3.4 shows the BH grid with +1400 collars, while the DD grid does not reach 50 collars.

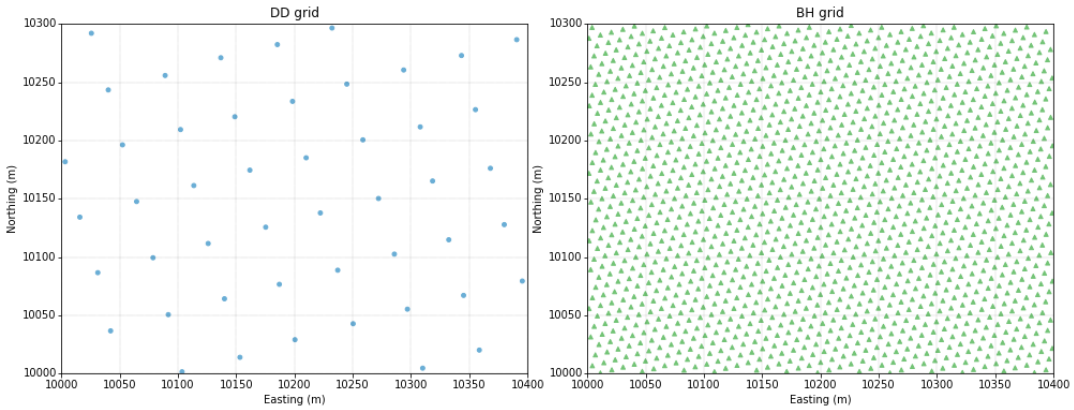


Figure 3.4: Drilling grids for a) DD and b) BH

2. **Conditioning datasets:** The conditioning datasets are prepared from the reference distribution simulated in the previous stage, using the distribution at a grid size of 0.25x0.25x0.25 m. Each DH has a specific conditioning dataset, that includes the nodes located in a 5 m square around the DH collar. Figure 3.5 shows a plan view of the samples of a conditioning dataset for a diamond drill hole.

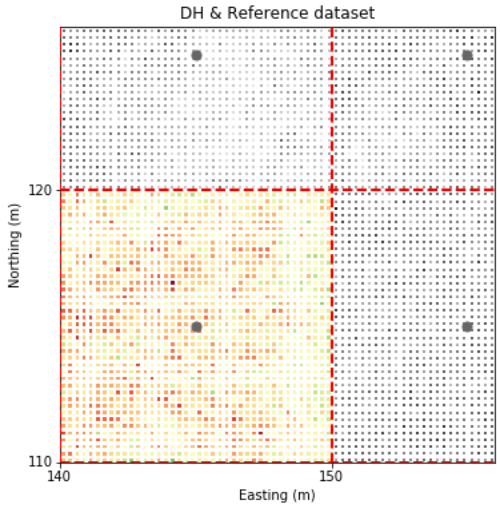


Figure 3.5: DH collar and reference dataset

3. **Simulation:** The DH simulation is done in a volume defined according to the drilling type. DD requires a horizontal area of 0.1x0.1 m and BH of 0.2x0.2 m. Both have the same length at the Z-axis, a single bench or 10 m. The simulation grid sizes are very short in the order of mm. This is to reproduce the distribution at point-scale. The simulation is performed using the USGSIM tool from the GSLIB library. The setup requires the input variogram

used previously in the reference distribution, the conditioning dataset, and the number of samples for conditioning. This last parameter is important because it directly impacts the results. A number below ten samples will produce a fast but low-accurate result, while a number above 50 samples have a better result but increases the processing time significantly. After many trials, this parameter was fixed in 75 samples. From each DH simulation are collected forty realizations. Table 3.3 summarizes these parameters.

Table 3.3: DH simulation setup

	DD	**BH**
Simulation area	10^{-2}x10^{-2}x10 m	10^{-2}x10^{-2}x10 m
Grid size	1^{-2}x1^{-2}x1^{-2} m	2^{-2}x2^{-2}x2^{-2} m
Conditioning samples	75	75
Nº Realizations	40	40
Nodes by DH	100K	50K

4. **Results validation:** After completing the simulation, it is necessary to check that the results are honoring the simulation input parameters. Histograms and variograms are checked, looking for an acceptable reproduction with respect to the input parameters. Besides the results are checked to be in the range of the conditioning datasets and the samples of the reference distribution. Figure 3.6 shows the variogram reproduction for a single DH, and Figure 3.7 the histogram reproduction.

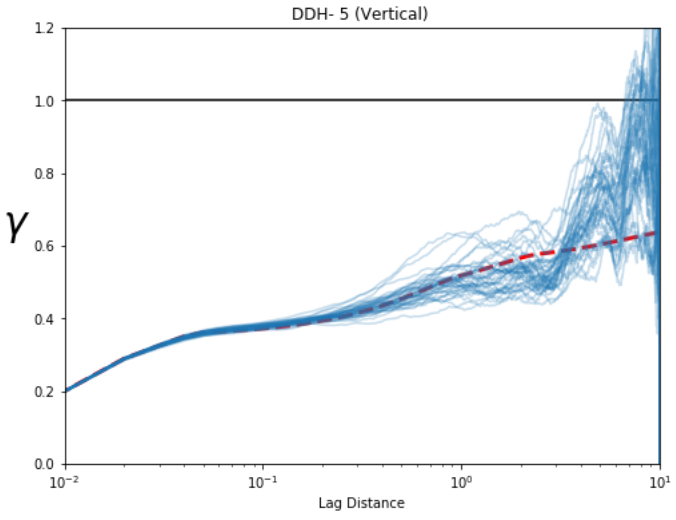


Figure 3.6: Input variogram (red dotted line) versus simulated DDH variogram (Vertical), showing 40 realizations.

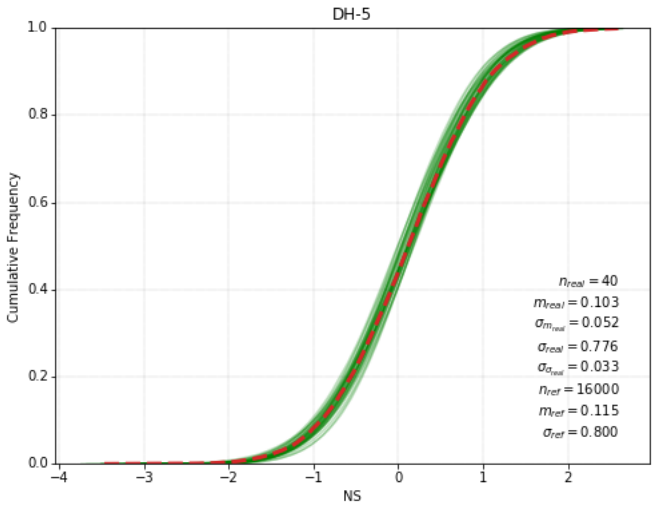


Figure 3.7: Reference distribution histogram (red dotted line) versus simulated DDH histograms, showing 40 realizations

5. **Discretization:** The resulting simulation can have 100 000 nodes per drill hole, which makes processing difficult, so it is necessary to perform a discretization. This reduces the number of nodes, honouring the grades and distribution features. DD holes are discretized into 5 steps of 20 mm for the horizontal direction and 250 steps of 40 mm for the vertical direction. BH holes have a discretization of 5 steps of 40 mm in the horizontal direction and 250 steps of 40 mm in the vertical direction. Figure 3.8 shows the original simulation grid and the discretized one for a single DH.

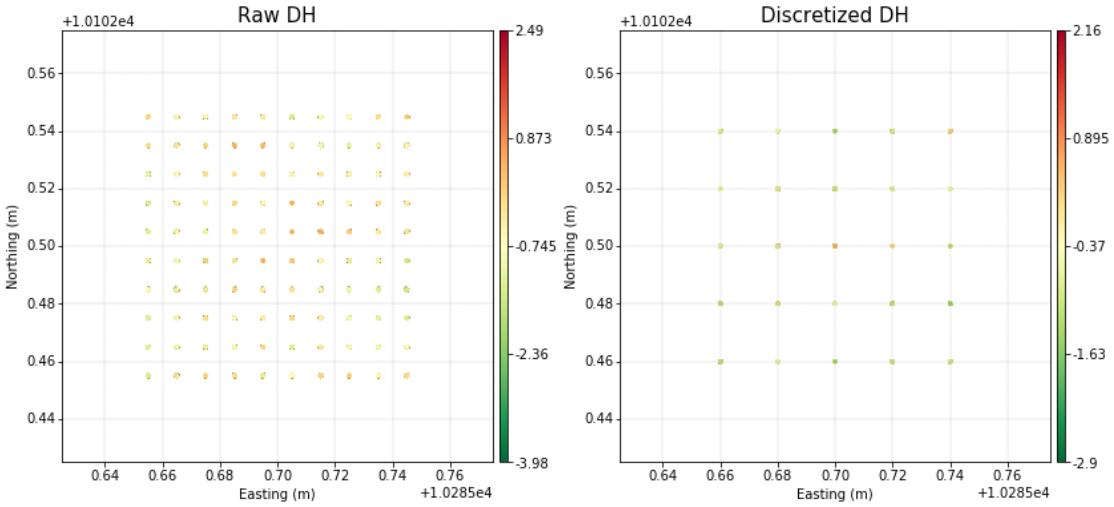


Figure 3.8: a) raw DH and b) discretized DH

6. **Tri-mineral distribution:** The tri-mineral distribution transformation is the last step of this stage. It is applied using the same cut-offs used previously for the reference distribution. This will ensure that all drill holes have the same mineral distribution. Table 3.4 summarizes the main features for both drilling types.

Table 3.4: Drilling type features

	DD	BH
# DH	48	1038
Diameter	64mm	200mm
Data spacing	50x50 m	10x12 m
DH area	0.1x0.1x10 m	0.2x0.2x10 m
# Realizations	40	40
Discretization	5x5x250	5x5x250
# Nodes	100K	50K

3.1.3. Sampling protocol

In this stage, the raw drill holes are transformed into samples datasets. Besides, error and bias are added to BH data, as they are needed for specific tests.

1. **Drilling support:** The support is defined by the drilling diameter, discarding the points outside it. As each drilling type has a different diameter, this will result in drilling samples with a different number of points. Figure 3.9 shows how this process is done.

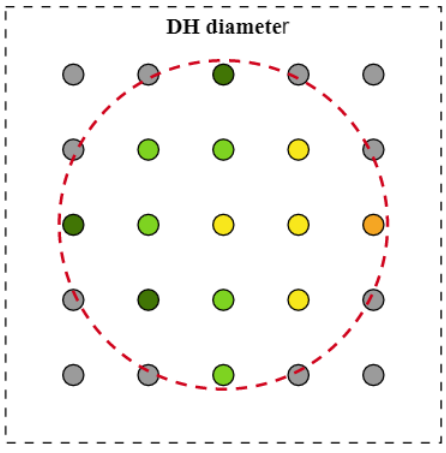


Figure 3.9: Plan view of a DH, the red dotted line indicates the DH diameter. Nodes outside the circle are eliminated from the simulation.

2. **Sampling protocol:** The sampling protocol differs for each drilling type, in the case of DD, the core is cut along the length and half is collected. A half of the nodes are discarded, and the other half represents the sample. Figure 3.10 shows this process. For BH drilling, the sampling is based on weights, in this case, 6 kg are taken per BH. Thus, BH sampling will be a random selection of the nodes, until the sample weight is reached.

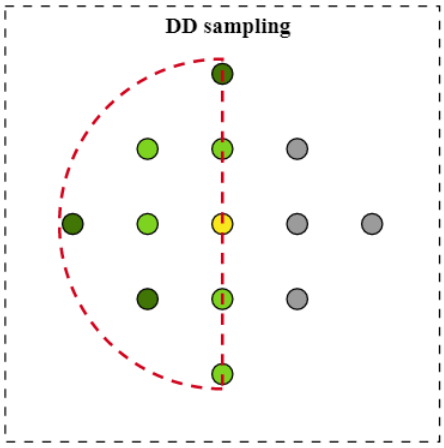


Figure 3.10: DD sampling, only a half of the DH is collected for lab assays.

3. **Compositing:** Both drilling types are composited to the bench height, which is 10 m in this study. The grades are the arithmetic average of the points inside each composite sample.
4. **Error and bias:** Error and bias are added to the BH dataset. These are added by multiplying the original grades and adding a random error between fixed intervals. DD is considered primary data, so no bias is added, just an error of ± 3%. BH is considered secondary data, with a 10% or 25% bias, depending on the database. An error of ±25% is added to all BHs.
5. **Datasets:** The composite samples are collected in datasets according to their drill type (DD & BH) and realization. A third dataset, called combination data (CD), joins both datasets. This last dataset only is used for checking purposes.
6. **Data standardization:** The last step is data standardization, which is necessary as lets us reduce the bias due to different drilling support. This is done for each dataset according to the formula:

$$Z_k^{STD} = \frac{Z_k - m_Z}{\sigma_Z} \qquad (3.1)$$

Where m_z and σ_Z are the mean and standard deviation of dataset k. All variography and estimations are calculated using standardized data, and the results are back-transformed to regular units for the assessments, following the inverse calculation.

3.2. Dataset acquisition

This research has required two databases to be used in the study cases. The areal extent is the same for both databases but differs in height. The first database is for a single bench and the second for five benches. The DH simulation and sampling are done using a similar setup but with minor differences.

3.2.1. Single bench database

This database involves a single bench, and it is used in chapter 4 as data to evaluate the performances of the different resources model. This database has been simulated using the same outline described in subchapter 3.1.1. Tables 3.5 summarize the features of the reference distribution.

Table 3.5: Reference distribution features

Reference distribution	
Plane area	400 x 300 m²
Bench height	10 m
# benches	1
Tonnage	13.5 MT
Cu grade	1.0%
Grid size	0.25 m
# nodes	76.8 M

There are two datasets, diamond drilling (DD) with 48 samples and blast-holes (BH) with 1489 samples, and a third one, that combines both data (CD). There are forty realizations for each dataset. DD and BH have been set as Primary and Secondary data, respectively. DD data is unbiased, and the samples have an aleatory error of ±3%, while BH has a bias of +25% and an error of ±25%. CD dataset is majority formed by BH samples, so both have a similar bias and error. Table 3.6 summarizes the features of the datasets but averaging the forty realizations.

Table 3.6: Dataset features

	DD	BH	CD
# samples	48	1489	1537
Cu grade	1.0%	1.24%	1.23%
Bias	1.0	1.25	1.24
Error	±3%	±25%	±25%

Figure 3.11 shows the reference distribution and the location maps for each drilling type. Visually, BH data shows a better reproduction of the reference distribution than the DD data, this is due to the bigger number of BH samples than DD.

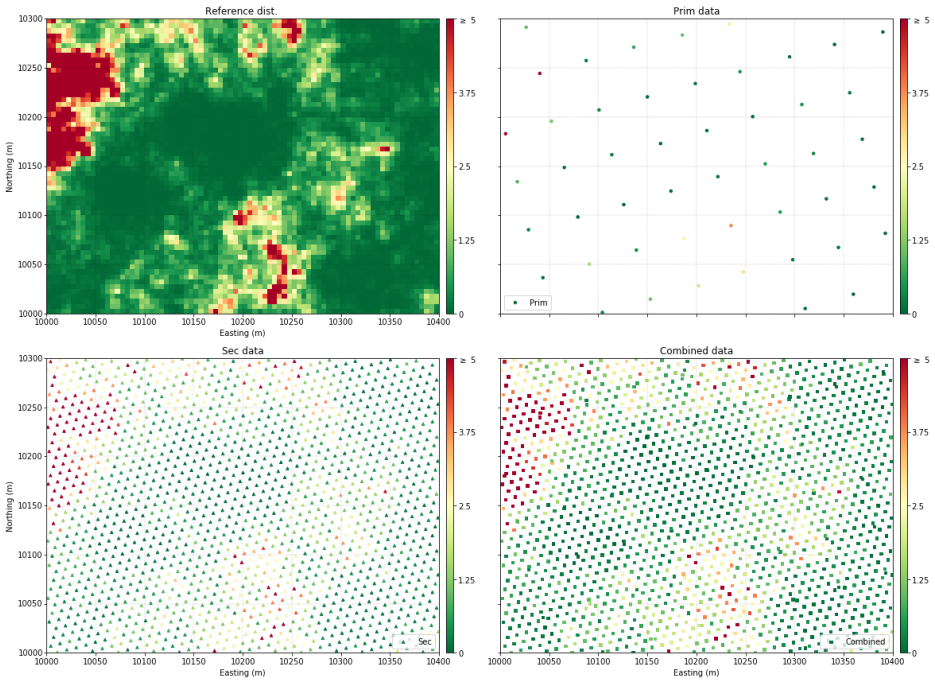


Figure 3.11: a) Reference distribution, b) DD dataset, c) BH dataset and d) CD dataset

Figure 3.12 shows the histograms and statistics for datasets. BH and CD show a remarkable bias, while DD data is unbiased. All datasets have a higher CV than the reference distribution, although the DD data has the closest CV to the reference distribution. Despite the difference in the number of samples, all datasets present similar statistics to the reference distribution.

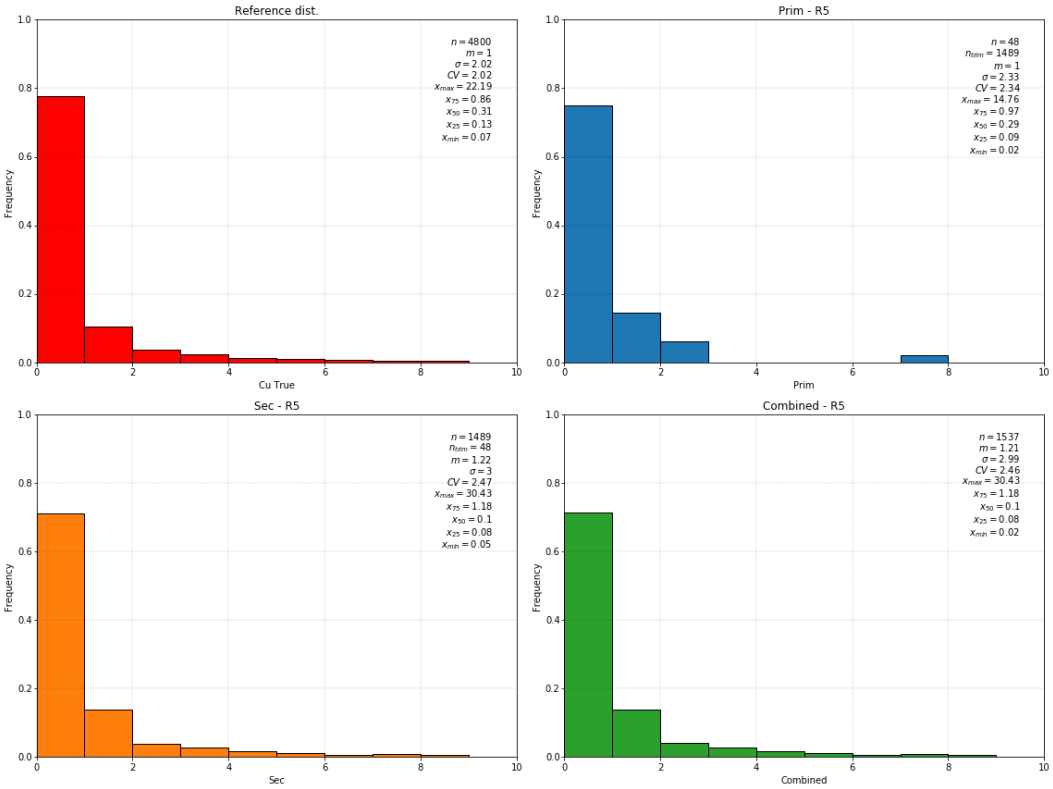


Figure 3.12: Histograms for a) Reference distribution b) DD data c) BH data and d) CD

3.2.2. Multi- bench database

This database is used in Chapter 5 to assess the performance of resources models to predict ore zones during a medium-term production period. It has been used the same outline described in section 3.1.1 but modified to extend the simulation volume to five benches. The grid size has changed from 0.25 m to 0.5 m, because the shorter grid would produce +380 M of nodes, which exceeds the current computational resources. Table 3.7 shows the features of the reference distribution.

Table 3.7: Reference distribution features

Reference distribution	
Plane area	400 x 300 m²
Bench height	10 m
# benches	5
Tonnage	67.5 MT
Cu grade	1.0%
Grid size	0.50 m
# nodes	48.0 M

The datasets are prepared according to the drilling type, DD and BH, and have increased the number of samples. The DD holes were simulated, including all benches, but the BHs were simulated bench by bench, using a different grid on each one. Both drilling types are composited to the same length of 10 m, totalizing 240 DD and 5661 BH samples. Regarding bias and errors, DD has no bias and an error of ±3% and BH has a bias of +10% and an error of ±25%. Table 3.8 summarizes the features of the datasets but averaging the forty realizations.

Table 3.8: Dataset features

Drilling type	Bench	DH grid	# samples	Bias	Error
DD	All	50x50m	240	1.00	±3%
BH	5050	10x10m	1185	1.10	±25%
BH	5040	15x15m	822	1.10	±25%
BH	5030	12x10m	987	1.10	±25%
BH	5020	10x10m	1185	1.10	±25%
BH	5010	10x8m	1482	1.10	±25%

There is no need to prepare a combination dataset for this database. DD and BH will be used separately for ordinary kriging estimation, and jointly for cokriging estimation.

4. Case study I: Ore control model assessment

This first case study evaluates the performance of the models and estimation methods applied in the ore control process. The models being evaluated include a long-term model (LTM), medium-term model (MTM), and short-term model (STM). The combined data model (CDM) is considered for checking purposes. The models are estimated using different methods and datasets but share most estimation parameters. The Bias, MSE, ore/waste classification and profits are used to evaluate the performance of the models.

4.1.Setup

This case study uses the simulated database presented in Section 3.2.1. Sets of exploration (DD) and production (BH) drill holes are simulated, collecting 40 realizations for each drill hole. All drill holes have been composited to a standard length of 10 m.

4.1.1. Data

There are two datasets according to the drilling type (DD & BH) and a third dataset corresponding to the combination of both datasets (CD). DD and BH data are considered Primary and Secondary data, respectively. From this point and forward, DD and BH data will be called primary and secondary data, respectively. Primary data has relatively few samples, just 48, while Secondary data has almost 1500 samples. Secondary data is overrepresented in the CD, representing 97%. Grade overestimation is expected in the models that directly use Secondary data (STM & CDM). Table 4.1 summarizes the features of each data type across the 40 realizations.

Table 4.1: Data summary

	Primary	**Secondary**	**Combined**
# Samples	48	1489	1537
# Reals.	40	40	40
Mean	1.00	1.25	1.24
SD	2.21	3.10	3.08
Min	0.01	0.05	0.01
Max	15.19	44.29	44.29

All datasets have been standardized for the variogram calculation and resources estimation. Once the estimation finishes, the results are back-transformed to regular units to proceed with the evaluation stage.

4.1.2. Variograms calculation

The experimental (semi) variograms are calculated for the three data sets and 40 realizations. For ease of the process, an omnidirectional variogram is used. As each dataset has different grid spacing, the lag distances are fixed at 50, 10, and 20m for Primary, Secondary, and combined data. Figure 4.1 shows how the experimental variogram has minor fluctuations between each dataset. The experimental variogram for secondary and combined data shows a better definition and stability, which is expected considering the larger number of samples available. Primary data has a little noisier variogram due to its smaller number of samples, but in general, both variograms are similar in ranges and shapes.

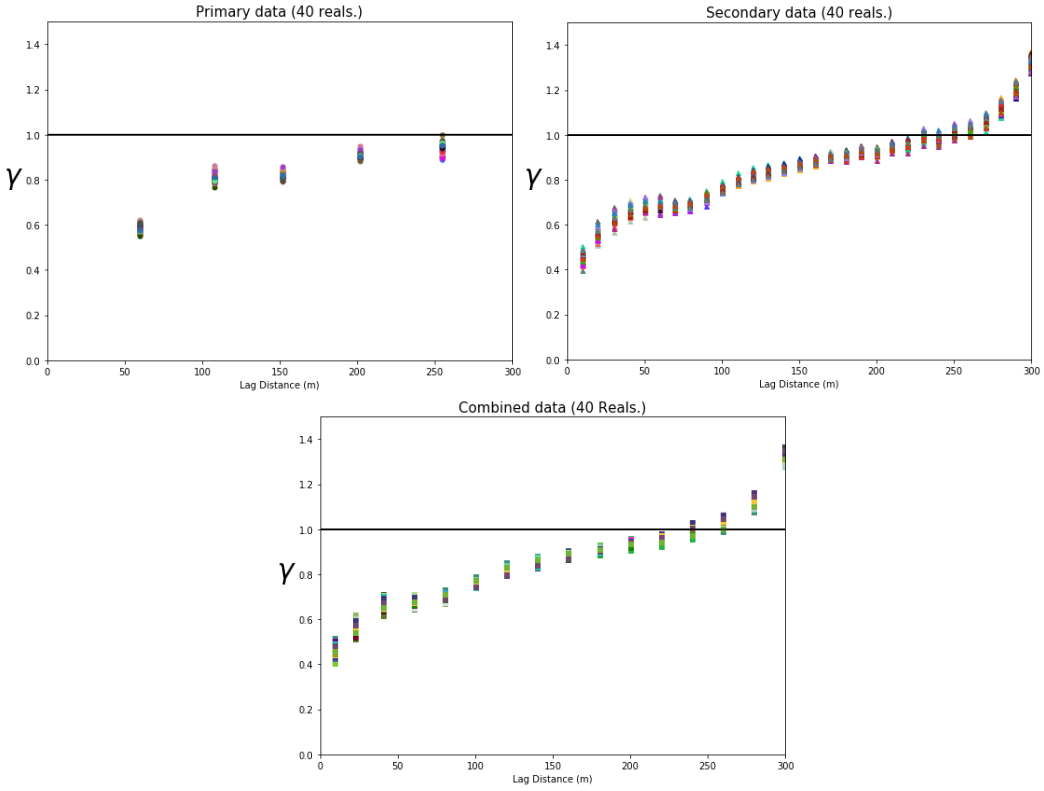


Figure 4.1: Experimental variograms including all realizations for a) Primary data, b) Secondary data and c) Combined data

The variogram is modelled by drilling type and for each realization, but all variograms are expected to have the same structures with minor differences in ranges and contributions. The variogram models must share the same structures for the LMC. As seen in Figure 4.2, the three variogram models present two structures, exponential and spherical, with maximum ranges between 250 to 300m. The nugget effect is fixed with a maximum contribution of 10%, 15%, and 25% for Primary, Secondary, and CD. Table 4.2 summarizes the model parameters.

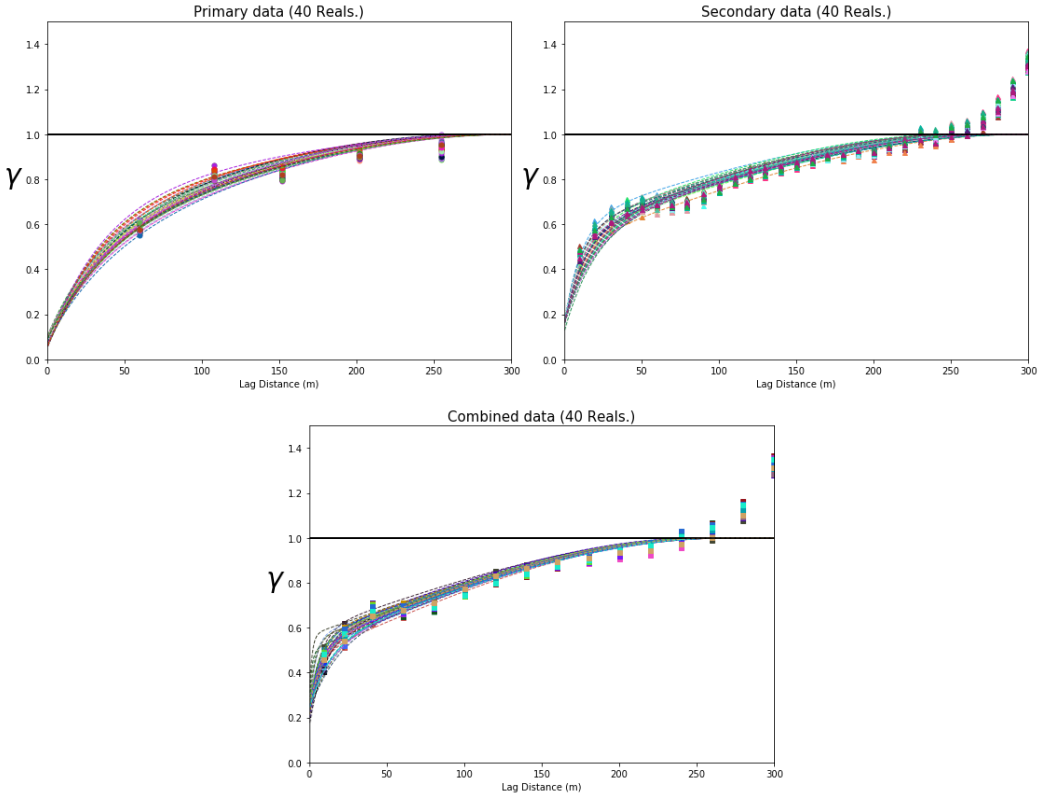


Figure 4.2: Experimental and variogram models including all realizations for a) Primary data, b) Secondary data and c) Combined data

Table 4.2: Variogram model summary

	Primary data	Secondary data	Combined data
# Structures	2	2	2
Nugget	<0.10	<0.15	<0.25
Type 1	Exponential	Exponential	Exponential
Type 2	Spherical	Spherical	Spherical
Contribution 1	0.40-0.60	0.35-0.45	0.25-0.35
Contribution 2	0.30-0.50	0.40-0.50	0.35-0.50
Range 1	100-125	30-60	5-50
Range 2	275-300	250-300	250-275

4.1.3. LMC models

Cokriging estimation requires an LMC of the Primary and Secondary data. This implies the calculation of the cross-variogram, but as both datasets are non-collocated, the direct calculation of cross-variograms is not possible. But the cross-variogram can be calculated indirectly through the cross-covariance. The procedure for this calculation is:

1. Calculation of the experimental variograms for Primary and Secondary variables.
2. Calculation of the cross-covariance between Primary and Secondary data. This is done through the GSLIB tool VARCALC, which uses the same setup as the variogram calc, requiring a lag distance, directions, and the respective tolerances.
3. The cross-covariance at 0 m (C(0)) is defined by extrapolating the cross-covariance curve until it intersects the Y-axis. This point will be used as a pivot to flip the cross-covariance. The nugget effect could be set to zero, as the data is non-collocated.
4. The cross-variogram is calculated by inverting the values of the cross-covariances, using as a pivot point C(0) (see Equation 2.32).

Figure 4.3 shows cross-covariances and the cross-variograms for the 40 realizations, having all of them similar ranges and shapes. Although each one of the realizations has a different cross-covariance curve, it is valid to use the same *C(0)* for all realizations, as this value indicates the degree of correlation between both datasets that seems quite stable. The *C(0)* is fixed in 0.85 by extrapolating the cross-covariance curve and considering the correlation between both datasets.

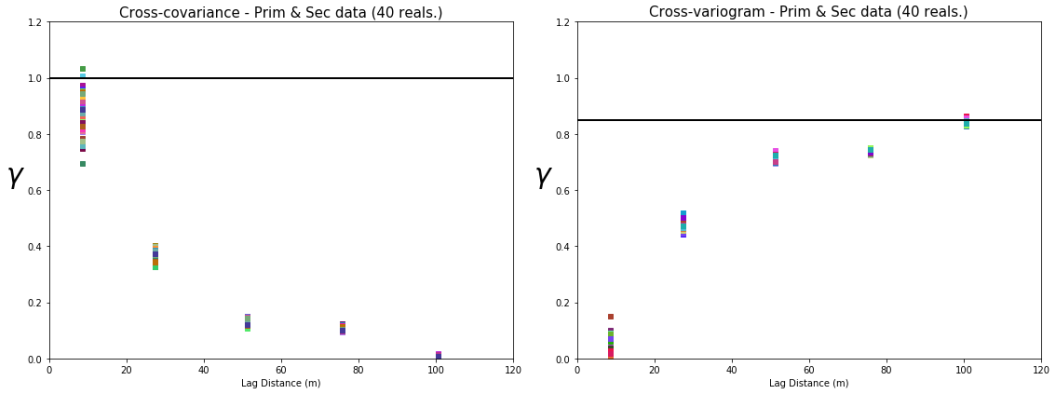


Figure 4.3: Primary-secondary data a) cross-covariances, and b) cross-variograms

The LMC must be fitted to the direct and cross variograms simultaneously. The same structures and ranges are repeated in all models, only varying the contributions. The contributions matrices must be positive definite (see Equation 2.31). The LMC fitting evaluates multiple combinations of structures and contributions for each realization. As there is an evident difference between the direct and cross variograms, the LMC will not necessarily match perfectly to them. In this case, the LMC must first fit the primary and then the cross variogram preferentially. The direct variogram models help to define the LMC. All models must share the same structures but not the contributions. After an in-depth review, exponential and spherical structures are chosen, with max ranges lower to 300 m. The nugget effect of the cross variogram is 0, as the data is non-collocated. The resulting LMC models fulfill the positive definite conditions. Figure 4.4 shows the LMC model for one realization, and table 4.4 summarizes the model parameters for the 40 realizations. Table 4.3 summarizes the LMC parameters.

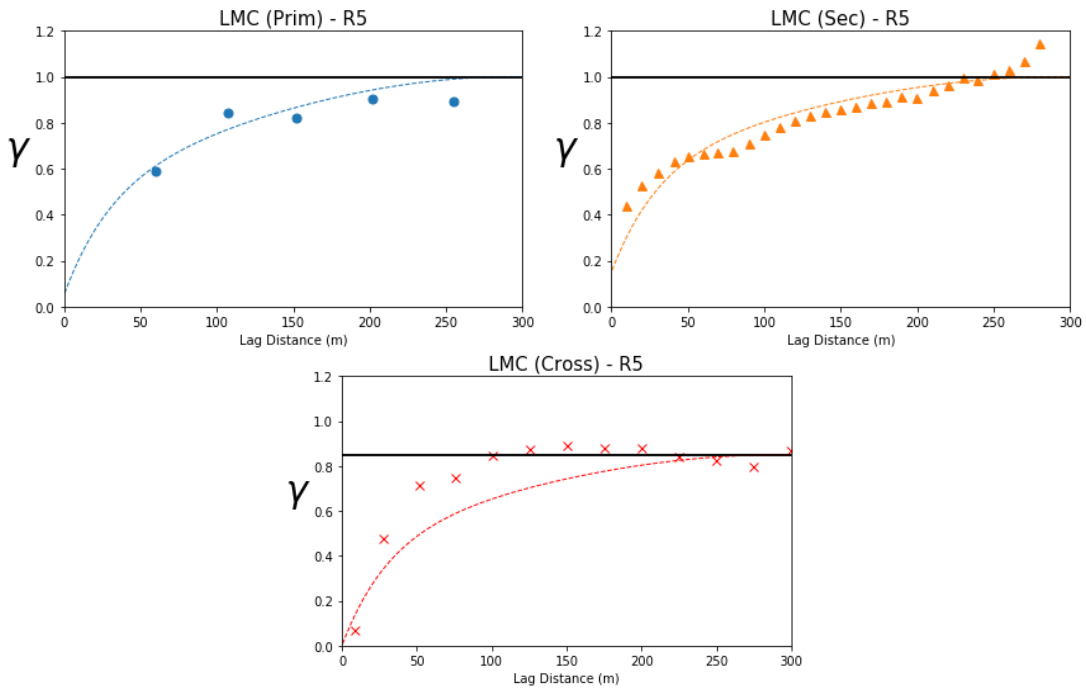


Figure 4.4: Experimental variogram and model for a) Primary data, b) Secondary data and c) Primary-secondary data

Table 4.3: LMC model summary

	Primary data	**LMC**	**Secondary data**
# Structures	2	2	2
Nugget	<0.10	0.0	<0.15
Type 1	Exponential	Exponential	Exponential
Type 2	Spherical	Spherical	Spherical
Contribution 1	0.40-0.60	0.40-0.60	0.40-0.60
Contribution 2	0.30-0.50	0.30-0.50	0.30-0.50
Range 1	75-125	75-125	75-125
Range 2	260-300	260-300	260-300

Fitting the LMC model for each of the 40 realizations requires multiple evaluations of the number of structures, contributions, and ranges. A close fit LMC is not always possible because it depends on the correlation between data types. This complex setup and data requirements make the application of cokriging in resource estimation limited to specific tasks or research.

4.2. Estimation

The estimation is done using the GSLIB tools KT3DN and COKB3D for ordinary kriging and cokriging, respectively. All models use the same estimation grid, fixed at 5x5x10 m, to make the estimation results comparable. Besides, other estimation parameters like the search radius and discretization are shared. Table 4.4 summarizes the estimation parameters of each model.

Table 4.4: Estimation parameters summary

	LTM	STM	CDM	MTM
# samples	48	1489	1537	1537
Method	OK	OK	OK	SOCK
Grid size	5x5x10 m	5x5x10 m	5x5x10 m	5x5x10 m
# Min samples	4	4	4	4
# Max samples	12	12	12	12
Search radio	500 m	500 m	500 m	500 m
Discretization	4x4x1	4x4x1	4x4x1	4x4x1

There are 40 estimation results for each model, the same number as realizations. The estimation results are in standard units, so they are back-transformed by applying the inverse calculation, using the mean and SD of their respective dataset and realization. In the case of the MTM, which uses both data types, the back-transformation uses the mean and variance of Primary data.

4.3. Evaluating the results

The models are compared to the Reference distribution by each of the 40 realizations. As the models have been estimated using a grid size of 5x5x10 m, they must be compared to the reference distribution resized at the same size. The average of all realizations is used to evaluate the performance by model. The parameters to assess are the Cu grades, histograms, MSE and Bias.

4.3.1. General statistics, Bias & MSE

The general statistics for each model are collected in Table 4.5, here are presented the average of the 40 results by model. The over-estimation in grades by STM and CDM is evident due to the use of biased secondary data, and their SD indicates a higher variability than the reference distribution. In the case of LTM, its average grade and SD are below the reference distribution, evidence of smoothing by

using kriging and limited dataset. MTM shows closer results to the reference distribution.

Table 4.5: Result models summary

	Ref. Dist.	LTM	STM	CDM	MTM
Cu	1.00	0.97	1.25	1.24	0.99
SD	2.02	1.39	2.41	2.31	1.66
Min	0.07	0.04	0.07	0.07	0.03
Max	22.19	11.31	21.53	19.89	14.50
Bias	1.00	0.97	1.25	1.24	0.99
MSE	0.00	1.63	1.38	1.22	0.93

Figure 4.5 shows the histograms of each model compared to the Reference distribution. STM and CD have a very similar histogram to the reference distribution, although they are systematically biased relative to the reference distribution. The MTM shows a more accurate histogram, although there is a discrepancy in the 1st interval for grades below 0.15 Cu.

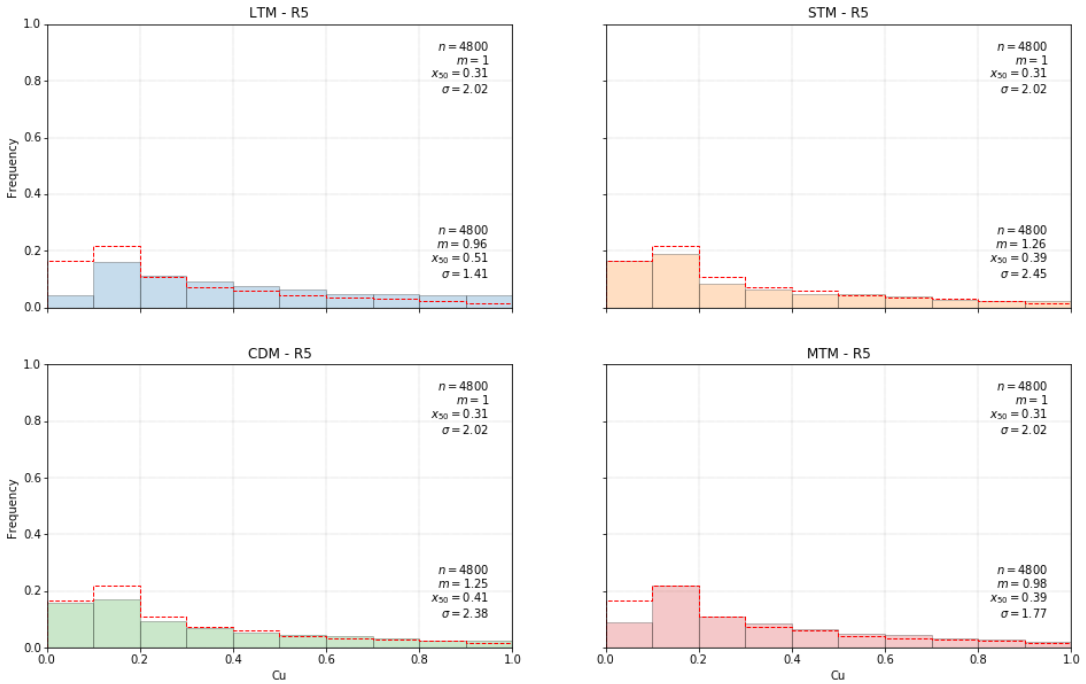


Figure 4.5: Reference distribution histogram (Red dots) vs the histograms of a) LTM, b) STM, c) CDM, and d) MTM.

LTM and MTM present the lowest bias, an average of 0.97 and 0.99 for the LTM and MTM, while the STM and CDM present an average bias of 1.25 and 1.24.

The high bias of these models is due to biased datasets. Besides, although ordinary kriging minimizes the MSE, it does not deal with the bias from the datasets. The MTM uses primary and secondary datasets, but the Cokriging method does not transfer the bias from Secondary data to the estimation. The lowest MSE corresponds to the MTM, then LTM, STM, and CDM. This demonstrates that cokriging is the best-unbiased estimator, providing results with the lowest error possible. Figure 4.6 shows the Bias and MSE by model and realization.

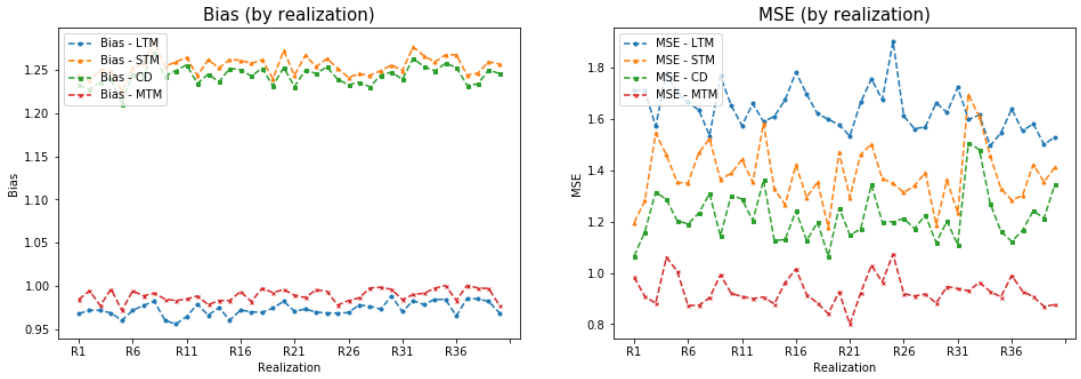


Figure 4.6: a) Bias by model and realization, and b) MSE

These metrics show that MTM has the best performance of all models in evaluation. STM and CDM have poorer performance than the MTM, although in some realizations, the STM/CDM has a better result, but on average, the MTM is the best.

4.3.2. Profit evaluation

The profit is an important metric to evaluate the performance of the models and the estimation methods. The profit is the difference between the revenues by selling the mineral contents of a block and the costs by its mining and processing. The profit is produced only by the ore blocks sent to processing, while the waste blocks, as they are not processed, only represent costs by their mining and hauling to the dumps. This makes the ore/waste classification a critical step in ore control. An incorrect block classification is costly.

The profit calculation requires the ore/waste classification, estimated and real. The estimated classification is done according to the estimated block grades from the models. The real one uses the same estimated classification but identifies the misclassified blocks by comparison with the reference distribution. The

comparison to the reference distribution allows assessment of how much misclassification is included in each model.

The misclassifications are Ore loss and dilution, each with a different frequency and impact on profit. The Ore loss is perhaps the most damaging as it includes the lost opportunity cost, which is equivalent to the profit if the block would have been processed as ore, luckily this misclassification is quite limited. Dilution has a minor economic loss by block, equivalent to the costs of mining and processing the block. Even there could be a limited income by processing the low-grade mineral. Commonly, this misclassification is more abundant than the ore loss.

A different profit calculation is applied to those misclassified blocks (see Equations 2.18, 2.19 and 2.20). The block classification and profits calculation require to define the cut-off grade and other parameters, summarized in Table 4.6.

Table 4.6: Cut-off parameters

Parameter	**Values**
Cu price	9070 $/T
Recovery	80%
Mining costs	2.5 $/T
Processing costs	14.5 $/T
Cut-off	0.23

Figure 4.7 shows the grade maps for the estimated models and reference distribution, just for a single realization. LTM, CDM, and MTM models present a good resolution and accuracy concerning the reference distribution, while the LTM does not. The low resolution of the LTM and the grades over-estimation in STM and CDM are sources of misclassification. Visually, all models reasonably reproduce the reference model.

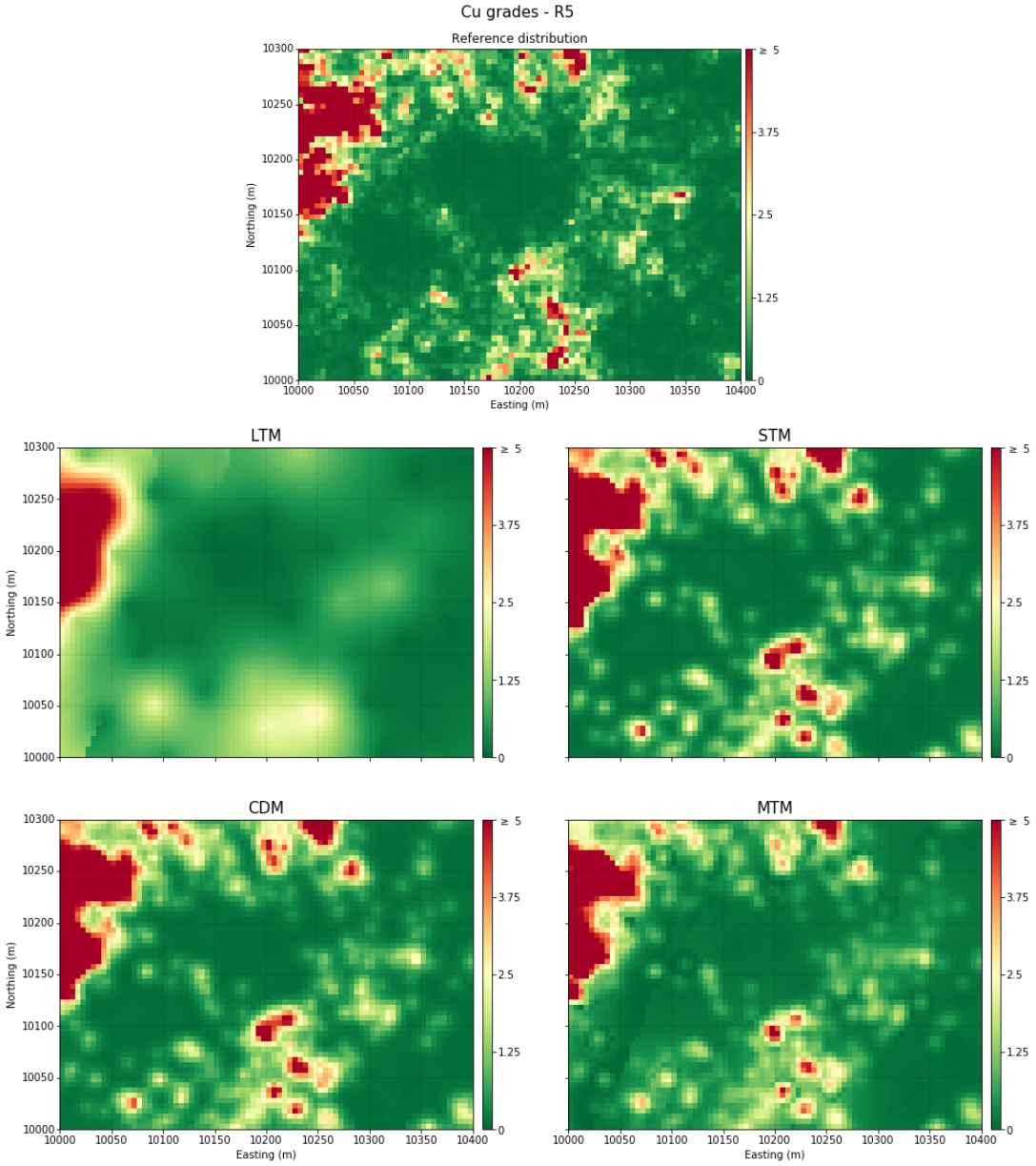


Figure 4.7: Grades map for a) Reference distribution, b) LTM, c) STM, d) CDM, and e) MTM

The estimated and real ore/waste classification maps are shown in Figures 4.8 and 4.9. The significant presence of dilution in all models is evident, while the ore loss is more limited although present in all models.

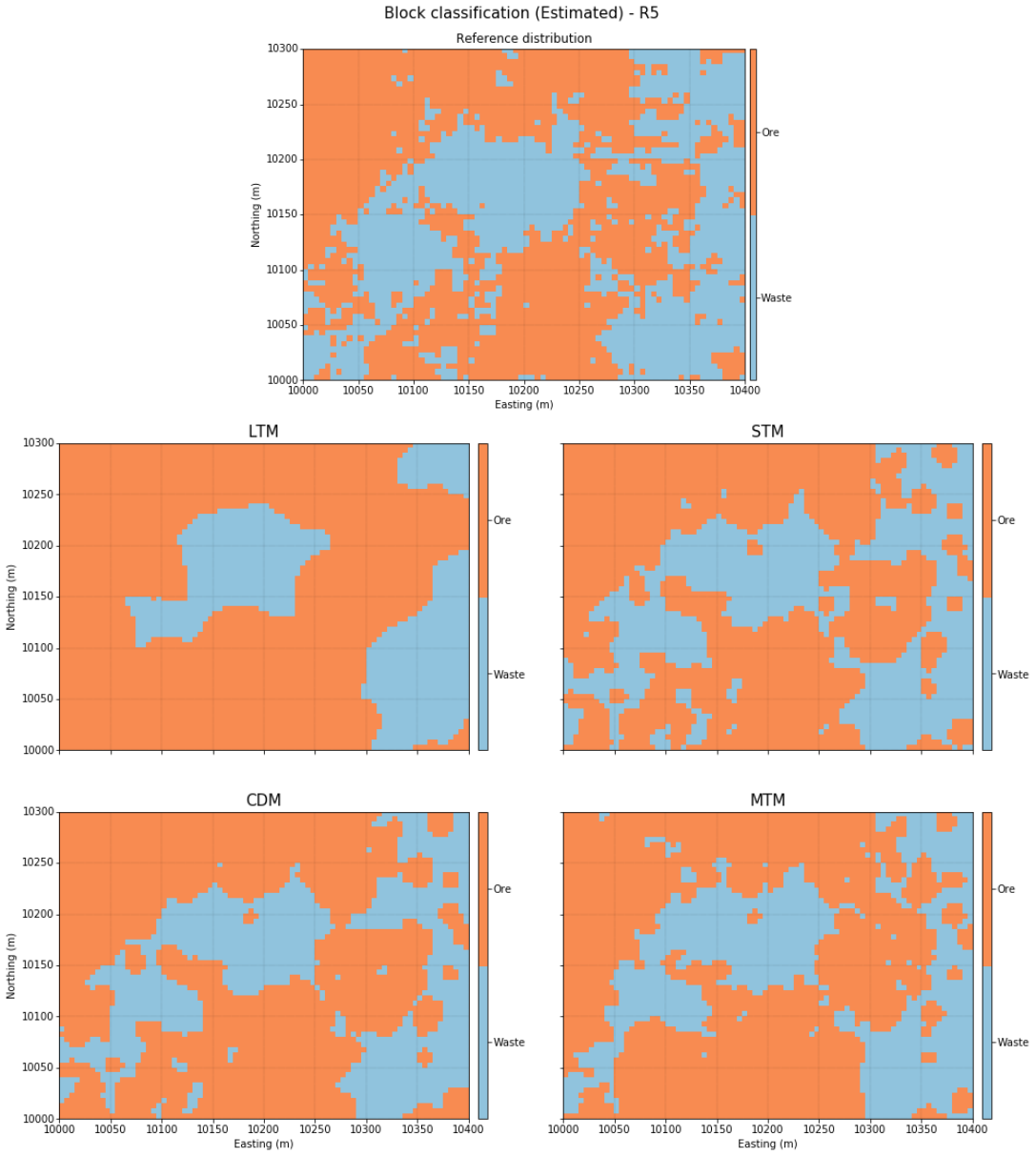


Figure 4.8: Ore/waste classification for a) Reference distribution, b) LTM, c) STM, d) CDM, and e) MTM

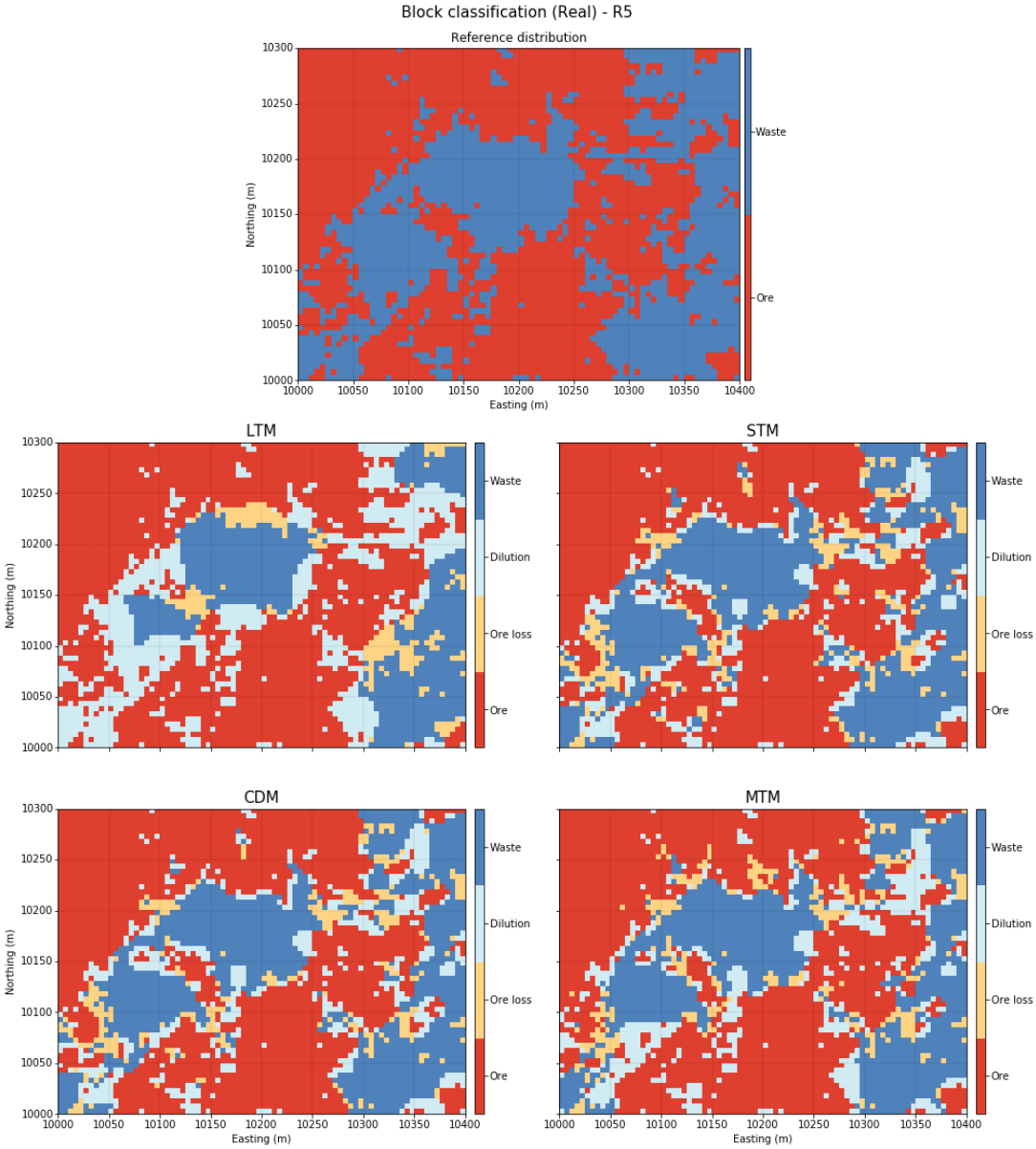


Figure 4.9: Real ore/waste classification for a) Reference distribution, b) LTM, c) STM, d) CDM, and e) MTM

Table 4.7 summarizes the average ore/waste classification according to the model. LTM presents the higher content of dilution, almost 20% of the whole tonnage. Models using secondary data, STM and CDM, have significantly lower dilution, around the 12%. This was expected, as the secondary data is abundant, which improves the resolution and classification, but as this data is positively biased in 25%, an overestimation in grades and additional dilution is expected.

Table 4.7: Ore/waste contents

Model	Estimated		Real			
	Ore	Waste	Ore	Ore loss	Dilution	Waste
Ref. dist	57.29	42.71	57.29	-	-	42.71
LTM	71.91	28.12	51.88	5.41	20.03	22.68
STM	60.11	39.89	48.99	8.30	11.23	31.47
CDM	63.48	36.52	50.70	6.59	12.77	29.94
MTM	68.74	31.26	52.65	4.65	16.47	26.24

Table 4.8 summarizes the results of profits, estimated and real for the average of the 40 realizations, and the respective reconciliation between the real and estimated profit. The profits have been standardized by dividing them for the reference distribution profit. LTM and MTM present the closer estimated profits to the reference distribution one, while the STM and CDM are highly over-estimated, on average a 30%. The real profit of STM and CDM are similar in value, but the MTM presents the highest profit. The reconciliation between the expected and real profit is very good for the MTM and LTM, slightly above 100%, while the STM and CD are far below the objective, around 30%.

Table 4.8: Profits summary

Model	Estimated		Real		Reconciliation
	M$	%	M$	%	
Ref. dist.	175.42	100.00	175.42	100.00	100.00
LTM	171.58	94.66	166.05	91.61	105.34
STM	239.05	131.88	168.11	92.92	68.12
CDM	235.55	129.99	170.04	93.89	70.05
MTM	174.98	96.86	170.70	94.15	103.46

4.4. Conclusions

MTM, estimated using cokriging, has been demonstrated to be the best-unbiased estimation technique providing estimation results with the lowest bias and MSE and ensuring the highest profit and best reconciliation between all models in evaluation.

STM and CDM also present high real profits, but their reconciliation ratio is low, and their estimated profits are over-estimated. The overestimation of estimated

profits can create false expectations, besides directing the planning and production to a wrong objective causing low productivity and economic losses.

Table 4.5 indicates that STM and CDM present a lower dilution but higher ore loss relative to MTM. This difference in the ore loss content makes MTM more profitable than STM and CDM, as the economic magnitude by ore loss is higher than dilution for the problem setup in this thesis.

The high performance in profits for the STM and CDM is counterintuitive, considering the use of biased and highly-error data on these models. The high profit can be related to the combination of the BH dataset, densely drilled and the short estimation grid. These features increase the model resolution and improve ore/waste classification, which directly impacts profits. Indeed, it is observed that STM and CDM present the lowest content of dilution, even lower than MTM, but their contents of ore loss are higher, which in the end, makes them less profitable. MTM uses DD and BH data, but the cokriging technique used assigns to the secondary data a lower weight than primary data and corrects the bias in the standardization.

CDM presents better profits than the STM and below the MTM, but the combined data joins primary and secondary data without any treatment so that both data can take the same weight in the estimation. This is not the best approach, considering the different supports for both datasets. Besides, ordinary kriging does not deal the secondary data bias.

5. Case study II: Resource model evaluation for medium-term planning

This case study evaluates the performance of the LTM and MTM to predict potential ore zones for planning and mining operations. The models are assessed during four months, and their results are compared to the STM and reference distribution. The LTM is estimated using only DD data and is updated once a year, several months after new data collection. Their results are globally accurate and widely used for long-term scheduling (+12 months). The LTM uses large block sizes, usually the SMU, resulting in a low-resolution model, low accuracy at a local scale, and with extensive dilution. These features make LTM not so useful for medium and short-term planning. The MTM is specifically built to provide information for intermediate planning, with improved model resolution and accuracy relative to the LTM. The MTM uses DD and BH data in a cokriging framework and can be updated several times a year as the BH dataset is updated. The STM is used to design mining polygons and for ore control operations. STM could be considered as the closest approximation to the underlying mineralization, as it uses a dense grid of samples, although a cokriging estimation properly set up can provide better results. LTM and MTM are reported by benches, indicating the average Cu grade, the ore/waste classification, and expected profits. These parameters are compared to the STM and the reference distribution to define the best performance.

5.1.Setup

This case study uses the simulated database presented in Section 3.2.2. As in case study I, it takes place in a synthetic Cu-porphyry deposit operated in an open-pit mining setting. In this area, 40 realizations of exploration (DD) and production (BH) drill holes are simulated. The LTM is estimated just one time using all DD data available and embracing the whole work area, the MTM is estimated monthly using the data available at the moment, this means that the database increases with each month. Finally, the STM is estimated by bench and month but using only the data for that specific bench.

5.1.1. Data

The work area occupies a volume of 400x300x50 m, including five benches. There are two datasets, DD and BH, both with 40 realizations. DD data present no bias and has a random error of ±5%, so it is considered the Primary data. BH data has a bias of +10% and an error of ±25% and is regarded as the Secondary data. Primary data uses an orthogonal grid and includes 48 drill holes and 240 samples, composited to 10 m. Secondary data are drilled using different grids according to the bench, making there were a different number of samples per bench, with 4476 samples in total. The ratio of primary:secondary samples is relatively high, 1:22. Table 5.1 summarizes the features of each data by a bench, averaged across the 40 realizations.

Table 5.1: Data summary

	Primary	**Secondary**				
		5040	**5030**	**5020**	**5010**	**All**
# Samples	240	822	987	1185	1482	4476
# Reals.	40	40	40	40	40	40
Mean	0.95	1.10	1.09	1.10	1.10	1.10
SD	2.01	2.12	2.07	2.06	2.10	2.09
Min	0.01	0.01	0.01	0.01	0.01	0.01
Max	18.00	31.48	27.36	28.73	36.35	36.35

The data has been standardized to perform the variogram calculations and estimations. After estimation the results are back-transformed to regular units to proceed with the assessment.

5.1.2. Variogram calculation

The direct variogram models are necessary to run the ordinary kriging for LTM and STM, while the cokriging applied in the MTM requires modelling the LMC of primary and secondary data. Some practitioners consider the use of local variogram models for the estimation of specific areas or benches. This practice improves the results, as there is a better orientation to the local anisotropies. But as this research uses stationary simulated data, the local variograms show just minor changes. Global variogram models, calculated using all data available, are used instead of local variogram models.

The direct variograms for both data types use omni-directional semi-variograms. Figure 5.1 summarizes the experimental variograms for both data types. There is an evident difference between the primary and secondary variograms probably due to bias and error applied on datasets.

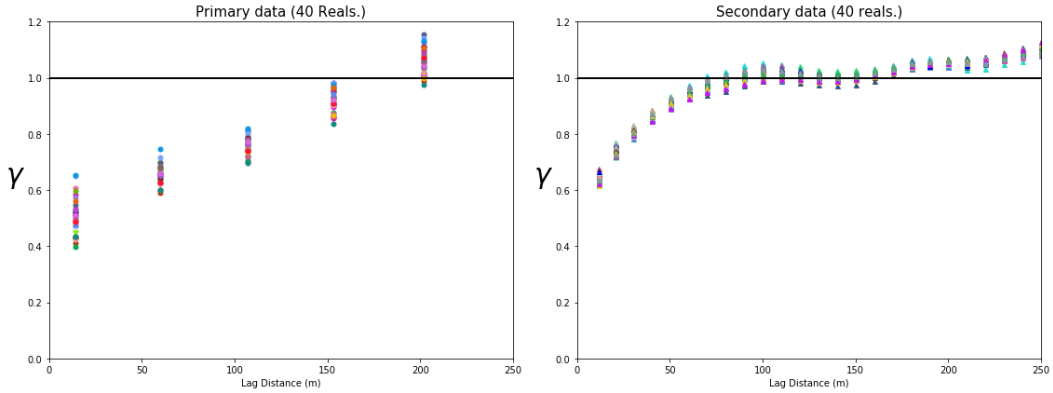


Figure 5.1: Experimental variograms for a) Primary data, and b) Secondary data.

The variogram models have exponential and spherical structures, and the nugget effect is fixed at a maximum of 20% for both models. Figure 5.2 summarizes the model variograms for each data type, and Table 5.2 summarizes the main model parameters.

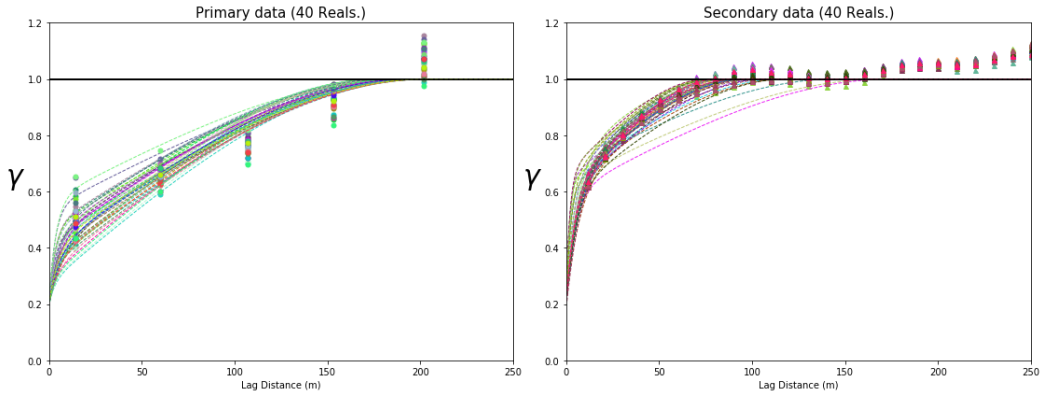


Figure 5.2: Variogram models including all realizations for a) Primary data, and b) Secondary data.

Table 5.2: Variogram model summary

	Primary data	Secondary data
# Structures	2	2
Nugget	<0.20	<0.20
Type 1	Exponential	Exponential
Type 2	Spherical	Spherical
Contribution 1	0.10-0.30	0.35-0.45
Contribution 2	0.60-0.70	0.35-0.45
Range 1	10-20	10-25
Range 2	175-230	90-170

5.1.3. LMC models

A valid LMC is necessary to run the cokriging for MTM. As explained above, the direct calculation of the cross-variogram is not possible, so the cross-covariance is calculated using the same outline as in case study I. The correlation between primary and secondary data is fixed at 0.6. Figure 5.3 shows the cross-covariances and cross-variograms for the 40 realizations.

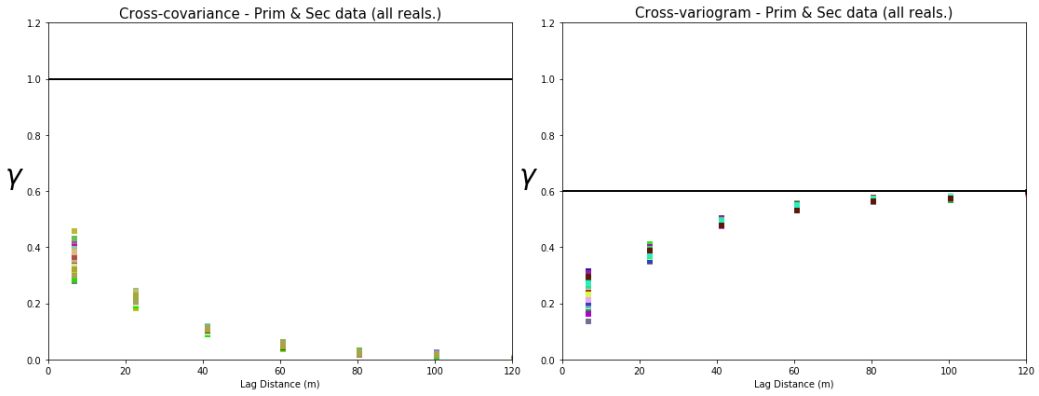


Figure 5.3: a) Cross-covariances, b) Cross-variograms

The LMC requires the variograms to share the same structures and ranges, and the contribution matrices must be positive definite. Two structures, exponential and spherical, with variable ranges and contributions according to the realization are selected. Figure 5.4 shows the LMC model for a single realization, and table 5.3 summarizes the model parameters for the 40 realizations.

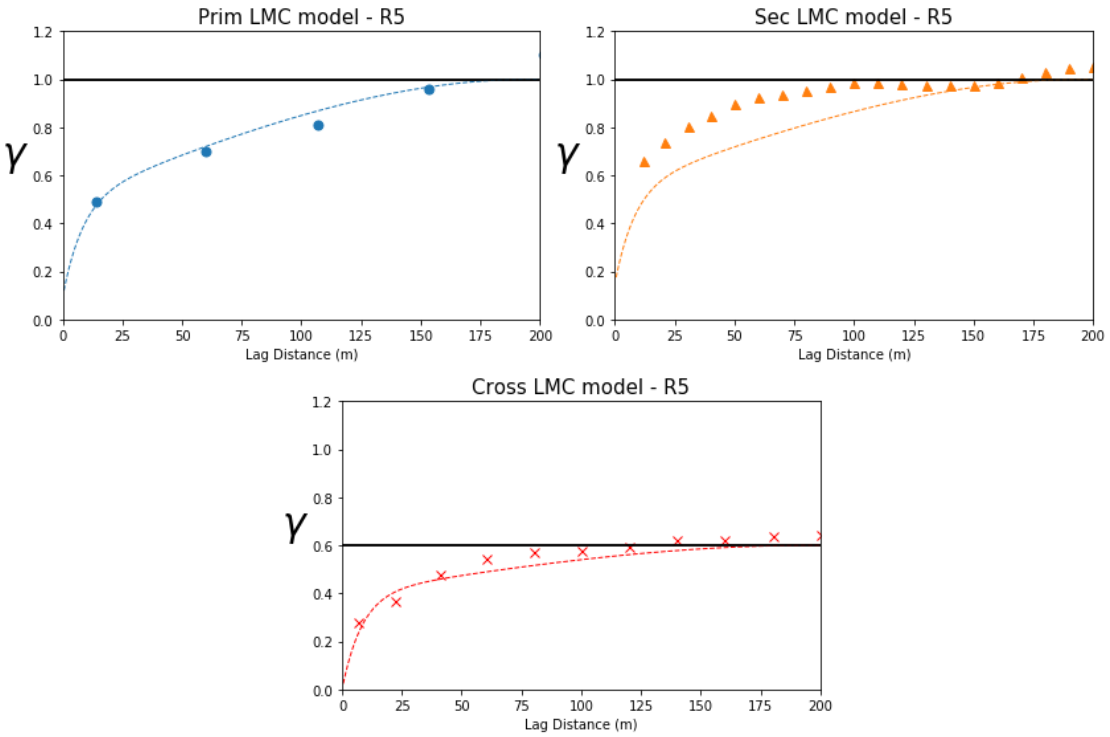


Figure 5.4: a) Primary model, b) LMC model and c) Secondary model

Table 5.3: LMC model summary

	Primary data	**LMC**	**Secondary data**
# Structures	2	2	2
Nugget	<0.20	0.0	<0.20
Type 1	Exponential	Exponential	Exponential
Type 2	Spherical	Spherical	Spherical
Contribution 1	<0.35,0.45>	<0.35,0.45>	<0.35,0.45>
Contribution 2	<0.35,0.45>	<0.35,0.45>	<0.35,0.45>
Range 1	<20,35>	<20,35>	<20,35>
Range 2	<190,210>	<190,210>	<190,210>

The direct secondary variogram model differs from the primary and the LMC. This could be due to the low correlation between the datasets, which is fixed at 0.6. In this case, the LMC must be fitted preferentially to the primary variogram, and then to the cross-variogram.

5.2. Estimation

Each model uses different estimation methods, data, grid size, and volumes, so their estimation setups are different, although they share some parameters. The

estimations are done using the GSLIB tools KT3DN and COKB3D for ordinary kriging and cokriging, respectively. The LTM is estimated using ordinary kriging, with an estimation grid based on the SMU, 50x50m, and a bench height of 10m. The estimation setup for this model is summarized in Table 5.4.

Table 5.4: LTM Estimation parameters summary

	LTM
# samples	240
Method	OK
Grid size	50x50x10
# Min samples	4
# Max samples	6
Search radio	500x500x50
Discretization	4x4x1

The MTM must provide an intermediate resolution and accuracy between the LTM and STM, so its estimation grid is fixed at 25x25x10m. As the MTM is estimated monthly, there are four estimation setups, that differ in the dataset used. The primary dataset is the same in all monthly estimations, but secondary data changes, adding more BH data each month. In this way, the deepest bench will have much more data than the highest one. Table 5.5 summarizes the setup.

Table 5.5: MTM Estimation parameters summary

	MTM			
	5040	5030	5020	5010
# samples	1425	2247	3234	4419
Method	SOCK	SOCK	SOCK	SOCK
Grid size	25x25x10	25x25x10	25x25x10	25x25x10
# Min samples	4	4	4	4
# Max samples	8/8	8/8	8/8	8/8
Search radio	500x500x50	500x500x50	500x500x50	500x500x50
Discretization	4x4x1	4x4x1	4x4x1	4x4x1

The STM requires a high resolution and accuracy, conditioned by the estimation grid size applied. The most recommended is to use a quarter distance of

the data spacing; this is 2.5 m. Like the MTM, the STM is estimated monthly, only for the bench in the evaluation and using the bench data. Table 5.6 summarizes the estimation parameters.

Table 5.6: STM Estimation parameters summary

	STM			
	5040	5030	5020	5010
# samples	822	987	1185	1482
Method	OK	OK	OK	OK
Grid size	2.5x2.5x10	2.5x2.5x10	2.5x2.5x10	2.5x2.5x10
# Min samples	4	4	4	4
# Max samples	20	20	20	20
Search radio	500x500x50	500x500x50	500x500x50	500x500x50
Discretization	4x4x1	4x4x1	4x4x1	4x4x1

5.3.Evaluating the results

The reference distribution has an average Cu grade of 1% over the five benches. As the evaluation only considers the last four benches, the average grade changes to 0.97% Cu. The estimated models are used to calculate the average grades, block classification and profits, per bench. The main evaluation is the performance to predict ore zones, so LTM and MTM results are compared with the reference distribution. Each of the 40 realizations is evaluated, and the average is used to assess the performance by model. As each model uses different grid sizes, resizing them to a standard size is necessary to compare them to the reference distribution. All models are resized to 1x1x10 meters.

5.3.1. Grades, MSE and bias

MTM has the closest average grade to the reference distribution, followed by LTM. The STM differs from them, showing an average Cu grade of 1.10%. The results bench by bench are similar, with the MTM and LTM close to the reference distribution grades, and the STM overestimated around 10%. The bias and MSE have been measured relative to the reference distribution. The MTM presents the lowest bias, followed by the LTM and STM, something expected considering the use of cokriging, which has been demonstrated to be the best-unbiased estimator. The STM has the lowest MSE value. Although this result could indicate a better performance

of the ordinary kriging above the cokriging, it must be remarked that the STM is locally estimated using data from the same bench, and the MTM is estimated using data from the upper benches, decreasing its performance. Table 5.7 summarizes the results.

Table 5.7: Grades. MSE & bias summary

Bench	Model	Cu grades	Bias	MSE
10040	LTM	0.90%	0.90%	3.99
	MTM	1.05%	1.05%	3.19
	STM	1.10%	1.11%	2.26
	Ref. dist	0.99%	1.00%	0.00
10030	LTM	0.89%	0.99%	2.79
	MTM	0.84%	0.93%	2.10
	STM	1.10%	1.22%	1.83
	Ref. dist	0.90%	1.00%	0.00
10020	LTM	0.86%	0.93%	2.59
	MTM	0.86%	0.93%	2.22
	STM	1.10%	1.19%	1.72
	Ref. dist	0.92%	1.00%	0.00
10010	LTM	0.87%	0.81%	3.70
	MTM	0.92%	0.86%	3.56
	STM	1.10%	1.03%	2.11
	Ref. dist	1.07%	1.00%	0.00
Average	**LTM**	**0.88%**	**0.91%**	**3.27**
	MTM	**0.92%**	**0.94%**	**2.77**
	STM	**1.10%**	**1.14%**	**1.98**
	Ref. dist	**0.97%**	**1.00%**	**0.00**

5.3.2. Ore/waste classification

The ore/waste classification is critical in the ore control process and eventually in the profit calculation. In real operations, it is not possible to exactly quantify dilution and ore loss, but simulated data allows the performance of the resource model to be assessed.

The estimated ore/type classification for LTM and MTM differ in tonnage and location. The MTM has the lowest content of waste, around 12%, the LTM and STM have similar contents of 20% and 23%. These values are far from the reference distribution with 51% of waste content. In the case of ore, all models are over-estimated, with contents around 80%. These results are evidence of dilution in all models. Table 5.8 summarizes these results.

Table 5.8: Estimated ore/waste classification

Bench	Model	Ore	Waste
10040	LTM	77%	22%
	MTM	89%	11%
	STM	81%	19%
	Ref. dist	49%	51%
10030	LTM	81%	19%
	MTM	86%	14%
	STM	77%	23%
	Ref. dist	49%	51%
10020	LTM	80%	20%
	MTM	88%	12%
	STM	77%	23%
	Ref. dist	49%	51%
10010	LTM	82%	18%
	MTM	87%	13%
	STM	75%	25%
	Ref. dist	50%	50%
Average	**LTM**	**80%**	**20%**
	MTM	**88%**	**12%**
	STM	**77%**	**23%**
	Ref. dist	**49%**	**51%**

The real ore/waste classification shows the ore loss and dilution quantities for each model, and their results are summarized in Table 5.9. The MTM has the closest content of ore to the reference distribution, followed by the STM and LTM. This indicates that the MTM is the best in ore recovery. MTM also has the lowest content of ore loss, the best misclassification, 2%, which is almost half of the content in LTM. The STM presents an intermediate content of ore loss. Dilution is extensive in all models, on average above 30%, reaching the MTM the 40% of dilution. Figures 5.5 and 5.6 show the estimated and real ore/waste classification; Figure 5.7 and 5.8 shows the block classification maps estimated and real.

Table 5.9: Real ore/waste classification

Bench	Model	Ore	Ore loss	Dilution	Waste
10040	LTM	45%	5%	33%	18%
	MTM	47%	2%	42%	8%
	STM	46%	3%	35%	16%
	Ref. dist	49%	-	-	51%
10030	LTM	45%	4%	36%	15%
	MTM	47%	2%	40%	12%
	STM	45%	4%	32%	19%
	Ref. dist	49%	-	-	51%
10020	LTM	45%	4%	35%	16%
	MTM	47%	2%	41%	10%
	STM	45%	4%	31%	19%
	Ref. dist	49%	-	-	51%
10010	LTM	47%	3%	35%	15%
	MTM	48%	2%	39%	11%
	STM	47%	4%	28%	22%
	Ref. dist	50%	-	-	50%
Average	**LTM**	**45%**	**4%**	**35%**	**16%**
	MTM	**47%**	**2%**	**40%**	**10%**
	STM	**46%**	**3%**	**32%**	**19%**
	Ref. dist	**49%**	**-**	**-**	**51%**

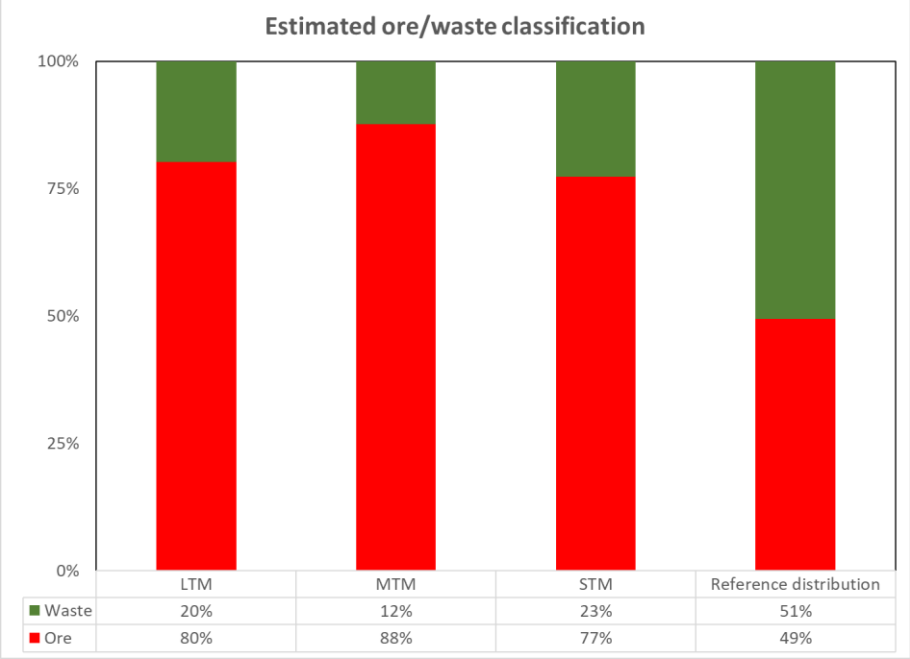
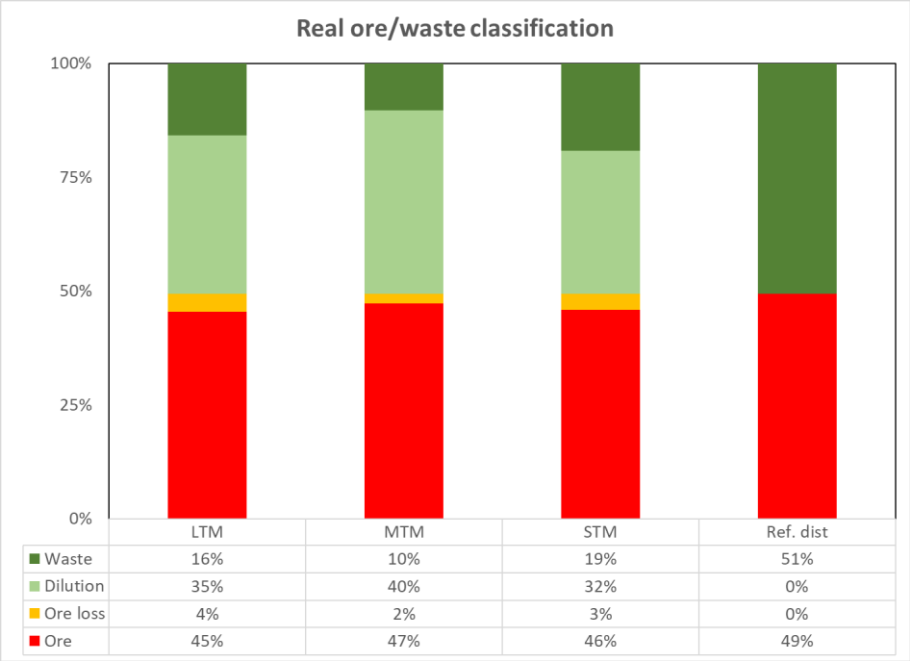


Figure 5.5: Estimated ore/waste classification



	LTM	MTM	STM	Ref. dist
Waste	16%	10%	19%	51%
Dilution	35%	40%	32%	0%
Ore loss	4%	2%	3%	0%
Ore	45%	47%	46%	49%

Figure 5.6: Real ore/waste classification

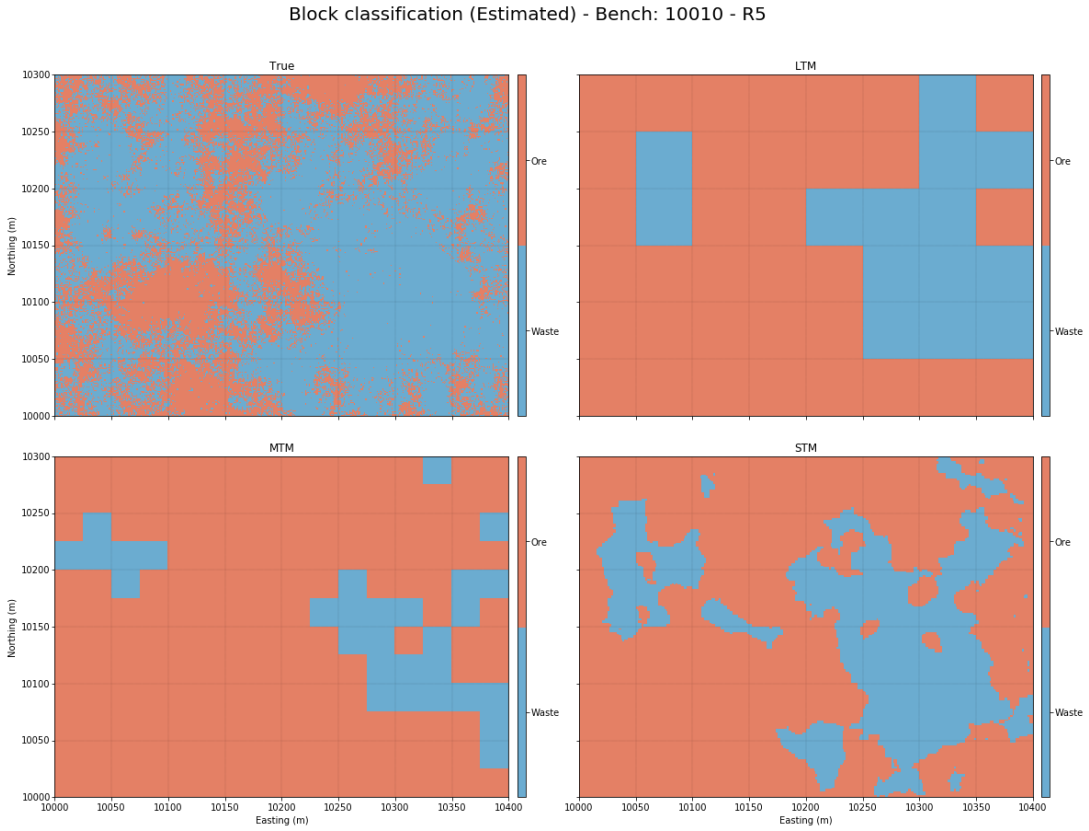


Figure 5.7: Estimated ore/waste classification maps for bench 10010 a) Reference distribution, b) LTM, c) MTM and d) STM.

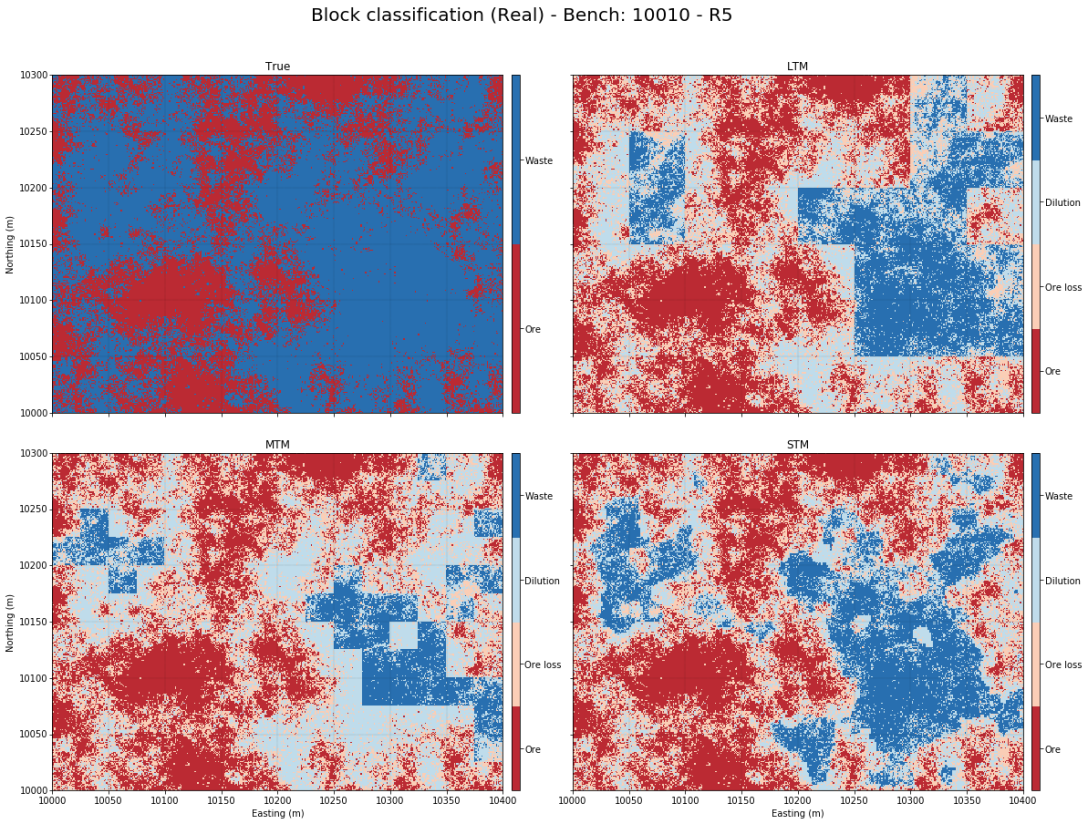


Figure 5.8: Real ore/waste classification maps for bench 10010 a) Reference distribution, b) LTM, c) MTM and d) STM.

5.3.3. Profits evaluation

The profit depends on the ore/waste classification for each model. Bias, error, and grid size influence this classification. The STM presents the highest expected profit, followed by the MTM and LTM. The estimated profit of LTM and MTM are important as they will be used for the planning and mining schedule. The real profit is calculated considering the misclassification content on each model. From Table 5.9, the ore loss is relatively low in all models, although the LTM doubles the content of the MTM. But dilution is the most extended, with contents above 30%, with the STM presenting the lowest content. The low content of both misclassifications in the STM, makes it has the highest real profit, followed again by MTM and then LTM. The expected reconciliation between the expected profit and the real one shows that the STM is over-estimated, with the MTM and LTM having better ratios. Indeed, the MTM has the best reconciliation, around 103%. The real reconciliation, this is the ratio between the real profit and the reference distribution profit, shows that the STM has a good performance, reaching the 88% of the real profit, with the MTM below

87% and then the LTM with 85% of reconciliation. Table 5.10 summarizes the results of estimated and true profits and Figure 5.9 shows the average profits by model.

Table 5.10: Profits & reconciliation by models

Bench	Model	Estimated (M$)	Real (M$)	Estimated Reconciliation	Real Reconciliation
10040	LTM	153.4	158.2	104%	85%
	MTM	186.5	160.9	87%	86%
	STM	199.9	165.7	83%	88%
	Ref. dist	187.2	187.2	100%	100%
10030	LTM	152.1	138.9	92%	84%
	MTM	139.3	144.4	104%	87%
	STM	202.2	143.7	71%	86%
	Ref. dist	166.3	166.3	100%	100%
10020	LTM	145.7	144.3	99%	84%
	MTM	143.3	149.0	105%	87%
	STM	201.4	149.0	74%	87%
	Ref. dist	171.3	171.3	100%	100%
10010	LTM	146.9	180.8	123%	88%
	MTM	157.7	181.1	115%	88%
	STM	202.6	184.1	91%	90%
	Ref. dist	205.1	205.1	100%	100%
Average	**LTM**	**149.5**	**155.6**	**104%**	**85%**
	MTM	**156.7**	**158.8**	**103%**	**87%**
	STM	**201.5**	**160.6**	**80%**	**88%**
	Ref. dist	**182.5**	**182.5**	**100%**	**100%**

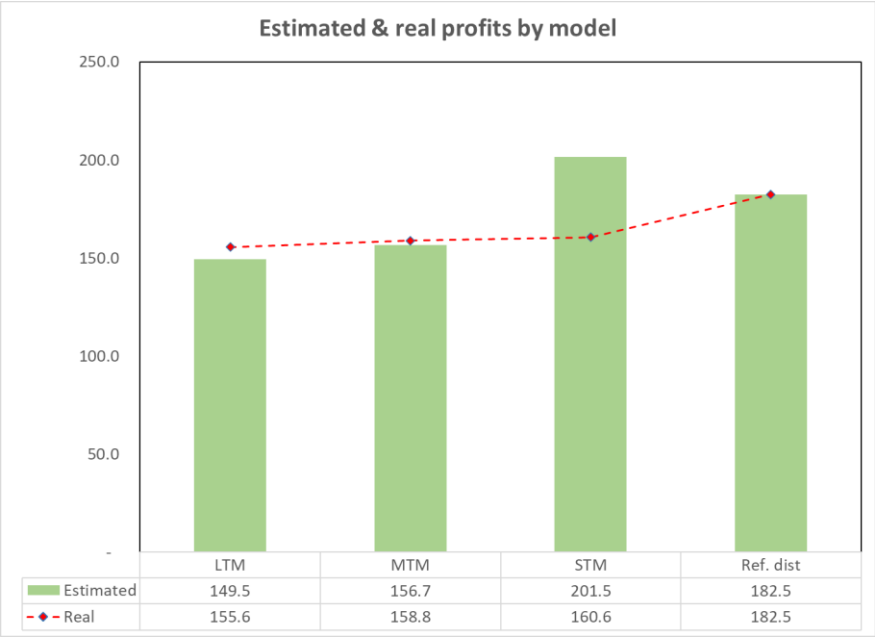


Figure 5.9: Profits by models

5.4. Conclusions

The MTM has a better performance than LTM to predict ore zones in a scenario of medium-term planning. The ore prediction for the MTM is more accurate than the LTM, showing the lowest ore loss content, although its dilution content is the highest. Besides, the real profit is higher for MTM than for LTM. The best-estimated reconciliation corresponds to MTM, too. This is possible for two main reasons: the use of a cokriging outline, which takes advantage of the primary and the secondary data available, and the use of smaller estimation grid sizes. The performance of MTM is just below the STM, even in some benches, the MTM profit is the same or higher than STM, which confirms its accuracy for ore prediction in short to middle planning.

Cokriging is accurate when there are available primary and secondary data, but its performance decreases as the estimation gets far from the secondary data. The MTM is estimated monthly, using the data available from upper benches, not from the bench in estimation, which makes the MTM performance and accuracy not necessarily be the best. This low performance is detected in the MSE, which lowest value, by benches and on average, corresponds to the STM, not the MTM.

Table 5.9 shows that dilution is more abundant misclassification in all models and is above 30%, reaching 40% for the MTM. The high content of dilution in the MTM can be reduced by incrementing the resolution model, which is done by lowering the estimation grid size. The massive occurrence of dilution in all models is worrying, no matter it has a limited impact on profits. Besides, the negative effect of dilution on planning and the processing plant is not quantified. The processing plant is not prepared to receive an overload of ore/dilution, so delays and economic losses could happen in refined mineral production. Finally, highly optimistic planning results in poor reconciliation and low productivity.

MTM estimated with cokriging is the best option for predicting mineral zones. It provides more accurate ore/waste contents than the LTM, and their profits are very similar to the STM and the Reference distribution model. The LTM has a good performance and is quite accurate considering its estimation setup. Even so, the LTM has a lower performance in predicting ore zones, bias, MSE and profits than the MTM.

6. Conclusions and future work

Data simulation for research is practical when access to real data is not possible or data with specific features are required. Many applications of simulation aim to simulate relatively widely spaced discretization points to assess resources at a large scale. This research is focused on the distribution inside a drill hole. Data are obtained from a specific drilling type by explicitly considering the drilling support and sampling protocols. According to the research requirements, different degrees of bias and error can be applied to data. Simulated DH data have been tested to reproduce the input parameters like histograms and variograms.

Simulated data show acceptable results, reproducing the spatial variability (variogram) of the reference distribution. The point-scale simulation and drilling support have a moderate impact on grades. The sampling protocol and error/bias also significantly impact grades and distribution. After applying the sampling protocol, DD and BH distributions differ more in most statistical parameters like histogram, mean, SD, or variograms. The sampling protocols imply a significant change in the drilling support, mainly for the BH data. After applying the sampling protocol, the estimation with BH data increments its MSE and bias.

The diamond drilling (DD) and blast hole (BH) data are classified as primary and secondary based on their quality and confidence. The best performance was achieved with cokriging with multiple drilling data types. Cokriging uses primary and secondary data and has a lower MSE and bias than models estimated using single datasets of primary and secondary data and even combined data from both datasets. The profits evaluated from the models show that the MTM has the highest and most accurate of all assessed models. The reconciliation is critical as it indicates the accuracy of the model to predict ore composition or profits. MTM has the most acceptable and realistic reconciliation ratio.

The bias can impact the estimation in unexpected ways. It is known that a positive bias over-estimates grade, distorting the ore/waste classification, which has a direct impact on the profit calculation. The misclassifications are dilution and ore loss. Ore loss is of critical economic importance in this study because, apart from the production costs, it includes the economic loss of not recovering the mineral content in the block. In contrast, the economic loss of dilution is limited to production costs

by block. Bias affects the estimation increasing average grades, which increases dilution but has an unexpected secondary effect of ensuring the recovery of more ore blocks. It has been observed that only a 1% difference in ore loss can make a model more profitable than others, but a difference of 5% to 10% in dilution, not necessarily imply a better profit. In this study, ore loss is economically more damaging than dilution.

The performance of medium and long-term resources models has been compared to assess which one better predicts ore zones and profits. The MTM estimated with cokriging demonstrated a better result. The MTM presents unbiased results, lower MSE and, importantly, an accurate reconciliation. The reconciliation ratio is critical as it is an important indicator for future planning and mining schedule.

FUTURE WORKS

The tests run in this research present acceptable results that confirm the validity of the simulated DH data. But these results are not definitive. The main observation relates to the input variogram, specifically in the setup of structures. This includes seven structures defined by trial and error. It is necessary to manually adjust this setup to obtain an acceptable outcome, valid for any mineral distribution or element.

The cokriging outline has provided better results than ordinary kriging, but this improvement is of modest magnitude. The setup of cokriging is more complicated. It is necessary to improve the LMC fitting.

Bibliography

Babak O. and Deutsch, C. V. (2007). *An Intrinsic Model of Coregionalization that Solves Variance Inflation in collocated cokriging*. CCG Annual Report 09, Centre for Computational Geostatistics, University of Alberta, Edmonton, Canada.

Babak O. and Deutsch, C. V. (2007). *Comparison of Cokriging with a LMC Versus the Intrinsic Model.* CCG Annual Report 09, Centre for Computational Geostatistics, University of Alberta, Edmonton, Canada.

Barnett, R.M. and Deutsch, C.V. (2018). *Straightforward Modeling of a Linear Model of Coregionalization: The Very Large LMC Alternative*. CCG Annual Report 20, Centre for Computational Geostatistics, University of Alberta, Edmonton, Canada.

Cuba, M. A., Ortiz, J. M. and Leuangthong, O. (2008). *Accounting for Different Sampling Errors in Exploratory Drilling Campaigns in Resource/Reserves Modelling*. CCG Annual Report 10, Centre for Computational Geostatistics, University of Alberta, Edmonton, Canada.

Cuba, M., Babak, O., and Leuangthong, O. (2009). *On the selection of secondary variables for cokriging and co-simulation*. CCG Annual Report 11, Centre for Computational Geostatistics, University of Alberta, Edmonton, Canada.

Cabral Pinto, F. (2020). *Independent Factor Simulation for Improved Multivariate Geostatistics.* Ph.D. Thesis, University of Alberta, Edmonton, Canada.

Deutsch, C.V. and Journel, A.G. (1997). *GSLIB Geostatistical Software Library and User's Guide*. Oxford University Press, New York, second edition.

Deutsch, C. V. (2007). *The Slope of Regression for Kriging Estimators*. CCG Annual Report 09, Centre for Computational Geostatistics, University of Alberta, Edmonton, Canada.

Deutsch, C. V. (2010). *Display of Cross Validation / Jackknife Results*. CCG Annual Report 12, Centre for Computational Geostatistics, University of Alberta, Edmonton, Canada.

Deutsch, C.V. (2016). *Workflows in Multivariate Mining Geostatistics*. CCG Annual Report 18, Centre for Computational Geostatistics, University of Alberta, Edmonton, Canada.

Deutsch, J. L. (2015). *Experimental Variogram Tolerance Parameters*. In J. L. Deutsch (Ed.), Geostatistics Lessons. Retrieved from http://geostatisticslessons.com/lessons/variogramparameters

Deutsch, J., Szymanski J., Deutsch, C. V. (2015). *Checks and Measures of Performance for Kriging Estimates*. CCG Annual Report 15, Centre for Computational Geostatistics, University of Alberta, Edmonton, Canada.

Donovan, P. N., (2015). *Resource Estimation with Multiple Data Types*. M.Sc. Thesis, University of Alberta, Edmonton, Canada.

Donovan, P. N., and Deutsch, C. V. (2014). *Cokriging with Multiple Data Types - Theor*y. CCG Annual Report 16, Centre for Computational Geostatistics, University of Alberta, Edmonton, Canada.

Donovan, P. N., and Deutsch, C. V. (2014). *Case Study in Cokriging with Multiple Data Types*. CCG Annual Report 16, Centre for Computational Geostatistics, University of Alberta, Edmonton, Canada.

Frykman, P., & Deutsch, C. V. (1999). *A Short Note on Geostatistical Scaling Laws Applied to Core and Log Data*. CCG Annual Report 01, Centre for Computational Geostatistics, University of Alberta, Edmonton, Canada.

Glacken, I. M. and Snowden, D. V. (2001). *Mineral Resource Estimation*, in *Mineral Resource and Ore Reserve Estimation – The AusIMM Guide to Good Practice* (Ed: A C Edwards), pp189–198. The Australasian Institute of Mining and Metallurgy: Melbourne.

Isaaks, E. H. & Srivastava, R. M. (1989). *An introduction to applied geostatistics*. New York: Oxford University Press.

Journel, A. G., Huijbregts, Ch. J. (1978). *Mining geostatistics*. Academic press.

Leuangthong, O., Neufeld C. and Deutsch C. V. (2003). *Optimal Selection of Selective Mining Unit (SMU) Size*. CCG Annual Report 05, Centre for Computational Geostatistics, University of Alberta, Edmonton, Canada

Manchuk, J. G., and Deutsch, C. V. (2015). *Latest SGSIM Program*. CCG Annual Report 17, Centre for Computational Geostatistics, University of Alberta, Edmonton, Canada

Manchuk, J., Leuangthong, O. and Deutsch, C. V. (2004). *A New Approach to Direct Sequential Simulation that Accounts for the Proportional Effect: Direct Lognormal Simulation*. CCG Annual Report 06, Centre for Computational Geostatistics, University of Alberta, Edmonton, Canada

Manchuk, J. G., and Deutsch, C. V. (2016). *Revisiting the Intrinsic Cokriging Model for Integrating Secondary*. CCG Annual Report 18, Centre for Computational Geostatistics, University of Alberta, Edmonton, Canada.

Matheron, G. (1971). *The theory of regionalized variables and its applications*. Ecole national superieure des mines.

Minnitt, R.C., & Deutsch, C. V. (2014). *Cokriging for optimal mineral resource estimates in mining operations*. The Journal of The Southern African Institute of Mining and Metallurgy, vol. 114.

Samson, M., & Deutsch, C. V. (2020). *Collocated Cokriging*. In J. L. Deutsch (Ed.), Geostatistics Lessons. Retrieved from http://geostatisticslessons.com/lessons/collocatedcokriging

Sinclair A., Blackwell G. (2002). *Applied Mineral Inventory Estimation*. Cambridge University Press.

Sunday, C.A., & Deutsch, C.V (2020). *Choosing the Discretization Level for Block Property Estimation*. In J. L. Deutsch (Ed.), Geostatistics Lessons. Retrieved http://geostatisticslessons.com/lessons/collocatedcokriging

Rossi, M. E., Deutsch, C. V. (2013). *Mineral resource estimation*. Springer Science & Business Media.

Wackernagel, H. (2003). *Multivariate geostatistics*. Springer-Verlag Berlin Heidelberg

Wawruch, T. M., Deutsch, C. V. and J. A. McLennan (2002). *Geostatistical Analysis of Multiple Data Types that are not Available at the Same Locations*. CCG Annual Report 04, Centre for Computational Geostatistics, University of Alberta, Edmonton, Canada

Vasylchuk, Y. & Deutsch, C. V. (2015). *A Short Note on Optimal Kriging Grid Size Relative to Data Spacing*. CCG Annual Report 17, Centre for Computational Geostatistics, University of Alberta, Edmonton, Canada

Vasylchuk, Y. (2016). *Integrated System for Improved Grade Control in Open Pit Mines*. Master's. Thesis, University of Alberta, Edmonton, Canada.

Vasylchuk, Y. V., 2019, *Advanced Grade Control with Multivariate Geostatistics, Blast Movement Modeling, and Optimized Dig Limits,* Ph.D. Thesis, University of Alberta, Edmonton, Canada