Risk Analysis in Short-Term and Medium-Term Mine Planning

Tchipato Amade Karimo

Thesis to obtain the Master of Science Degree in

Mining and Geological Engineering

Supervisor: Prof. Amílcar de Oliveira Soares

Examination Committee

Chairperson: Prof. Maria João Correia Colunas Pereira

Supervisor: Prof. Amílcar de Oliveira Soares

Members of the Committee: Engr. Diogo José D'Hommée Caupers

June 2021

Declaration

I declare that this document is an original work of my own authorship and that it fulfils all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Resumo analítico

O planeamento e sequência de exploração de uma determinada área de um jazigo em ambiente subterrâneo consiste na avaliação de reservas minerais e programação da sua exploração. Parâmetros como os métodos de desmonte em uso, condicionantes da produtividade, custos de exploração e de tratamento, parâmetros geomecânicos e geotécnicos e parâmetros económicos como os preços dos metais são contabilizados na avaliação das reservas e o risco ou a incerteza associada às reservas deveriam tomar em conta a incerteza associada a esses mesmos parâmetros, tal como acontece na quantificação dos recursos (teores estimados do jazigo).

Pretendeu-se com a presente dissertação apresentar uma metodologia em que os parâmetros mais importantes – recuperação mineira, preço dos metais, teores de entrada na lavaria – são considerados variáveis aleatórias com leis de distribuição inferidas a partir de dados experimentais. Para tal, tendo em vista a exploração de zinco na mina de Neves-Corvo, utilizaram-se 181 observações mensais de preços de zinco dos últimos 15 anos e, aplicando o método de alisamento de Holt, obtiveram-se as previsões de preços para os 3 meses seguintes, com MAPE igual a 9.66 %. A lei de distribuição cumulativa de 17 observações de recuperação mineira utilizando o método de desmonte *Bench-and-Fill* foi ajustada a uma equação exponencial cujo coeficiente de determinação (R^2) foi de 96.9 %. Por fim, foi analisada uma amostra de 49 observações de teores de entrada na lavaria de zinco da qual se obteve que a amostra era proveniente de uma distribuição normal com parâmetros média igual a 7.6% e desvio-padrão igual a 0.795%.

O conjunto das leis de distribuição foram simuladas e co simuladas utilizando o método de Monte Carlo. Tendo em conta o princípio de que as minas geram NSRs a volta de 50% para o zinco, obtiveram-se leis de distribuição das reservas minerais, com teor de corte acima de 5.3%, e leis de distribuição do NSR, que podem ser quantificas num risco ao longo dos 3 meses do plano mineiro a curto-médio termo.

Palavras-chave: Planeamento Mineiro, Método de Holt, *Bench-and-Fill*, Risco, Reservas Mineiras, NSR;

Abstract

The planning and sequence of exploration of an area from a mining deposit in an underground environment consists of evaluating mineral reserves and scheduling of their exploration. Parameters such as mining methods in use, productivity constraints, exploration and treatment costs, geomechanical and geotechnical parameters and economic parameters such as metal prices are accounted for in mining reserves evaluations and the risk or uncertainty associated with the reserves should take into account the uncertainty associated with these parameters, as in the quantification of resources (estimated grades of ore in the deposit).

The aim of this dissertation was to present a methodology in which the most important parameters - mining recovery, metal prices, zinc head grades - are considered random variables with probability distribution inferred from experimental data. For this purpose, taking zinc exploration in the Neves-Corvo mine, 181 monthly zinc price observations of the last 15 years were analysed and, applying the Holt exponential smoothing method, the price forecast for the next 3 months was obtained, with MAPE equal to 9.66%. The cumulative distribution of 17 mining recovery observations using the Bench-and-Fill mining method was adjusted to an exponential equation whose determination coefficient (R^2) was 96.9%. Finally, a sample of 49 observations of zinc head grades was analysed, from which it was inferred that the sample came from a normal distribution with parameters mean equal to 7.6% and standard deviation equal to 0.795%.

The set of probability distributions was simulated and co-simulated using the Monte Carlo method. Keeping in mind the rule-of-thumb that mines generate NSRs around 50% for zinc, a probability distribution of mineral reserves was obtained, with cut-off value above 5.3%, and probability distribution of NSR, which can be quantified as risk over the 3 months of the mining plan in the short-medium term.

Keywords: Mine Planning, Holt Exponential Smoothing, Bench-and-Fill, Risk, Mining Reserves, NSR;

Index of Contents

List of tables

List of figures

Chapter 1 – Introduction

1.1 Motivation

Exploitation of natural resources, namely the extraction and processing of raw materials in the mining industry such as metals, used for the production of commodities, plays a major role in the way society has been evolving to complex civilizations during the last centuries and that continues to be the trend. Take zinc for example, for which demand is expected to rise as other industries like civil engineering and automobile will continue to use this product on a large scale associated with steel, Daigo et al (2013), and as countries like China and India continue their expansions towards urbanization. Indeed, zinc is the fourth most widely used non-ferrous metal worldwide only below aluminum, copper and lead and, as table 1.1 shows, the recent zinc consumption between 2016 and 2021 have stabilized around 13 million tons, even during 2020 when the industry was hit by the global COVID-19 pandemic.

Table 1.1 - Zinc Consumption between 2016-2021; Source: International Lead and Zinc Study Group

World Refined Zinc Supply and Usage 2016 - 2021											
000 tonnes						2020	2021	2020/2021			
	2016	2017	2018	2019	2020	Jan		Oct	Nov	Dec	Jan
Mine Production	12668	12681	12820	12892	12145	1017	1042	1120.4	1117.8	1129.9	1042.1
Metal Production	13560	13486	13102	13480	13641	1160	1190	1192.7	1189.6	1212.6	1190.2
Metal Usage	13665	13953	13658	13709	13105	1094	1178	1166.9	1186.0	1189.1	1178.5

Whilst there are good prospects for past, current and future demand for transformed mineral products, the mining industry faces sustainable challenges characterized as the “combination of enhanced socioeconomic growth and development, and improved environmental protection and pollution prevention”, Hilson and Murck (2000). Also, as pointed out during conferences like the Earth Summit (1992) in Rio, addressing general principles to solve this issue, it is necessary to work with local governments, so as to offer them a mining perspective, thereby adopting practices that engage the local communities such as: providing jobs to the residents irrespective of race, gender or religion or using state of the art clean technologies during the extraction, processing and mine closure.

In addition to the social issues addressed above, new technical and financial solutions might become necessary for companies to remain competitive, as the more the industry matures, the more minerals will come from lower grade ore bodies and sometimes in deeper and severe geological conditions. As noted by Albanese and McGagh (2011), higher data density and increased predictive power will enable investors and decision makers to more confidently accept or reject extraction projects in its pre-feasibility stages and, although the scope of this thesis is towards the operational phase of a mining, it is based on this premise that the majority of the work is developed.

1.2 Objectives

The purpose of this thesis is to create a stochastic model of the reserves estimates and profit in a way that can be translated into a risk throughout the mine planning. In the initial stage, it will be a priority to recognize what are the financial and technological parameters that correlates with profit and reserves estimates, such as the mining recovery, processing plant recovery, metal price estimates, mining and processing costs as well as water and tailing costs and General and Administrative (G&A) expenses.

Thereafter, probability density functions (PDFs) of the parameters where historical data is available will be estimated so that a model used to predict different outcomes can be applied, called Monte Carlo simulation. Then, using an equation relating revenues and costs, the PDF of the mine profit will be obtained, therefore enabling mineral reserves evaluation.

The results obtained are critical as they link geostatistics with financial analysis, accounting for risk in the short-term and medium-term horizons, and provide important information to assist in a decision-making process to better act in future events.

1.3 Document Outline

The work presented in this document is divided in five chapters. The first chapter consists of a brief overview of the current situation in the mining industry, major social issues that the sector is facing and will continue to be scrutinized for in the near future. Next, the objective of the thesis is also presented.

The second chapter gives a summary of some simplistic methods for forecasting and the general concept of the Monte Carlo simulations, referring to the inverse transform method. A brief discussion of graphical methods to infer known probabilities from historical data is also presented and finally, a vision on how the risk analysis is applied in the mine planning is addressed.

Chapter 3 describes the methodology used to obtain the final results. It consists of the data analysis that was executed and the mathematical or statistical models that can be applied for each case. It also shows how the simulations were made. Chapter 4 can be viewed as an extension of its preceding chapter, as it presents the results of the methodology and a discussion. Both chapters 3 and 4 make reference to the second chapter.

Finally, chapter 5 gives concluding remarks to the thesis related to literature and describes recommendations and future improvements to be made given the limitations of the present work.

Chapter 2 – State-of-the-art

In this chapter, the knowledge that has been developed in the same working areas of the present thesis, necessary to structure the solutions for the proposed problem, will be presented. Therefore, bibliographic research work was carried out in the following subjects:

- Time series and forecasting methods;
- Monte Carlo simulation;
- Risk analysis in mine planning.

2.1 Time Series and Forecasting Methods

Metal prices (together with knowledge of the structure of the ore deposit to be exploited) represent the factor with the greatest impact on the financial performance of companies in the mining sector, however, the unpredictable changes caused by the fluctuation in the demand for these materials, make this component uncertain and volatile. Eggert (1987) states that changes in metal prices in the present and in the past shape expectations about future prices and influence mining revenues, profits and costs. Against this backdrop, it is essential to use metal price forecasting techniques as an auxiliary tool for mine planning.

Makridakis et al. (1998) classify forecasting methods in two categories:

- Qualitative methods
- Quantitative methods
    - Explanatory: explanatory relationship between two or more variables (e.g., copper price and zinc price).
    - Time series: forecast based on historical patterns.

In the context of this work, quantitative methods were used, which can only be applied when the following conditions exist, identified by Makridakis et al. (1998):

1. Availability of information about the past.

2. This information has to be quantified in the form of numerical data.

3. It can be assumed that patterns observed in the past will continue into the future.

In addition to these aspects, according to Dooley and Lenihan (2005), it is necessary to choose the time scale of the forecast, as well as the appropriate techniques, collect the data to be analysed and test the forecast model.

The choice of forecasting methods presented below is related to their simplicity of use, the fact that they capture trends in metal prices, the availability of inexpensive software both economically and temporally, as well as its effectiveness for short-term forecasts.

To analyse data sets in a simple way, a univariate analysis is used. In this type of analysis, parameters such as mean, standard deviation and variance are evaluated. Data types that show fluctuation patterns around an average value are called stationary in their mean. The concept of Simple Moving Average (SMA), specifies how many observations recorded in the past are desired to be included in the average value, this value being stored in a variable named k. As new observations become available, a new average is computed, discarding the oldest values and adding the most recent values. The SMA forecast (Ft + 1) performed through the sum of the last k observed values (Yi) divided by the number of observations considered (k), translates into equation (2.1):

$$F_{t+1}=\frac{1}{k}\sum_{i=t-k+1}^{t} Y_i \tag{2.1}$$

Makridakis et al. (1998) states that an extension of the method above is prediction by Simple Exponential Smoothing (SES). In this method, the forecast captures the fact that the most recent observations generally offer a more reliable guide for the future, establishing a scheme that gives more weight to the most recent observations and a lesser weight to the older observations, a weight that is transposed to a variable α that ranges between 0 and 1. Equation (2.2) provides the forecast value Ft + 1, obtained from the previously predicted value, Ft, adjusted with a forecast error that is known when the actual value of the respective observation, Yt, becomes available.

$$F_{t+1}= \alpha Y_t+(1-\alpha)F_t \tag{2.2}$$

The use of this simple forecasting model was expanded by Holt (1957) so that the forecast values generated could capture the trend of the data. This improved model, known as the Holt linear method, or double exponential smoothing, is based on two smoothing constants α and β, also with values between 0 and 1, and the number of forecast periods, k, translated into three equations, of level - (2.3), trend (slope) - (2.4), and forecast - (2.5):

$$N_t=\alpha Y_t+(1-\alpha)(N_{t-1}+T_{t-1}) \tag{2.3}$$

$$T_t=\beta(N_t-N_{t-1})+(1-\beta)N_{t-1} \tag{2.4}$$

$$F_{t,k}=N_t+kT_t \tag{2.5}$$

For the forecasting method to be started, Sharif and Hasan (2019) point to the need of initialization of N_0 as the estimate for the level and T_0 for the trend. For the present work, assuming that the data of the prices to be predicted have a regular and slightly erratic behaviour, we take $N_0=Y_0$ and $T_0=Y_1-Y_0$, otherwise, Makridakis et al. (1998) proposes the use of the least number of squares regression of the first values to initialize the variables. For simple exponential smoothing, as well as for double exponential smoothing, the values of α and β can be chosen from the percentage error in a given period (PEt), given by equation (2.6), in order to minimize the value of the equations (2.7) and (2.8), the average percentage error (MPE), the average percentage absolute error (MAPE), respectively, or another error criterion.

$$PE_t=\frac{Y_t-F_t}{Y_t} \tag{2.6}$$

$$MPE=\frac{1}{n}\sum_{t=1}^{n}PE_t \tag{2.7}$$

$$MAPE=\frac{1}{n}\sum_{t=1}^{n}/PE_t/ \tag{2.8}$$

Finally, Makridakis et al. (1984) points out that for situations in which it is difficult to obtain accurate forecasting using smoothing techniques, more sophisticated univariate statistical analysis methods should be used. Diggle (2013) finds that Monte Carlo tests encourage data analysis, since the user can choose any statistical class U (Unbiased) to focus on any “aberrant” characteristic of their data.

2.2 Monte Carlo Simulation

Monte Carlo simulations are part of a computational method originating in the 1940s, developed by scientists who tried an alternative to the trial-and-error method, to solve a problem in the development of nuclear weapons at the Los Alamos National Laboratory. The general idea is to establish a model that generates random values from a variable with known probability distribution. Mohammad et al. (2013) proposed an approach similar to the following steps to carry out a Monte Carlo simulation, in the context of selecting an underground exploration method:

1. Determine the probability density function and cumulative distribution function for each variable to be simulated.
2. Use a pseudorandom number generator between [0,1].
3. For each random number allocate a value, using the cumulative distribution function.
4. Establish a method to combine the results of each simulation and design the resulting probability density function.

As a way of selecting the probability distribution for one variable when historical data is available, Thomopoulos (2013) proposes to use graphical methods such as Q-Q Plot, which consists of comparing quantiles of the data sample to be analysed with the specific theoretical probability distribution quantiles and observing the quality of fit of the data. Therefore, the author advises that we start by confronting the data with more common probability distributions, such as normal, exponential, lognormal, gamma, beta and Weibull. Kolmogorov-Smirnov (KS) test can also be used to compare a set of data with a reference probability distribution. As noted by Hassani and Silva (2015), KS test is used to compare the empirical cumulative distribution function of the data, F_{obs}, with the theoretical cumulative distribution function, associated with the null hypothesis that the sample comes from the theoretical distribution, F_{exp}. The one sample KS statistic is given by equation 2.9 where D_n stands for the maximum absolute distance between the expected and observed distribution functions.

$$D_n = \max_x \left| F_{exp}(x) - F_{obs}(x) \right| \tag{2.9}$$

The variables to be simulated can be considered discrete, if the result set is in a list of possible values, or continuous when the variable takes on any value specified in a range. Thomopoulos (2013) identifies two methods for generating values of a random variable, through a probability distribution:

- Accept-Reject method;
- Inverse transform method.

In the context of this work, the variables to be considered will be continuous and the inverse transform method will be chosen, since it requires fewer steps than the concurrent method and because it is assumed that the probability distributions to be studied will be relatively simple and mathematically easy to apply. In this paradigm, taking x, a continuous variable in the domain [a, b], f (x) is called the probability density function. The cumulative distribution function of x, in domain [0,1], corresponds to equation (2.10):

$$F(x) = \int_a^x f(x)dx \tag{2.10}$$

Considering a number u, from a pseudo-random number generator between 0 and 1, which generates from a standard uniform distribution, u ~ U (0,1), it is possible to obtain the value of x through the inverse function of F(x), represented by equation (2.11):

$$x = F^{-1}(u) \tag{2.11}$$

Geostatistical estimation models the orebody spatial distribution of grades and other characteristics of interest of the deposit that are used by mine production scheduling, applying a set of deterministic assumptions to come up with an “optimal” underground mine design, extraction sequence, and production time schedule, maximizing profits, as noted by Monkhouse and Yeates (2018). From the beginning of the last decade a different framework of the presented above has been developed, using a set of equiprobable models of the deposit as show in figure 2.1, integrating risk and uncertainty.

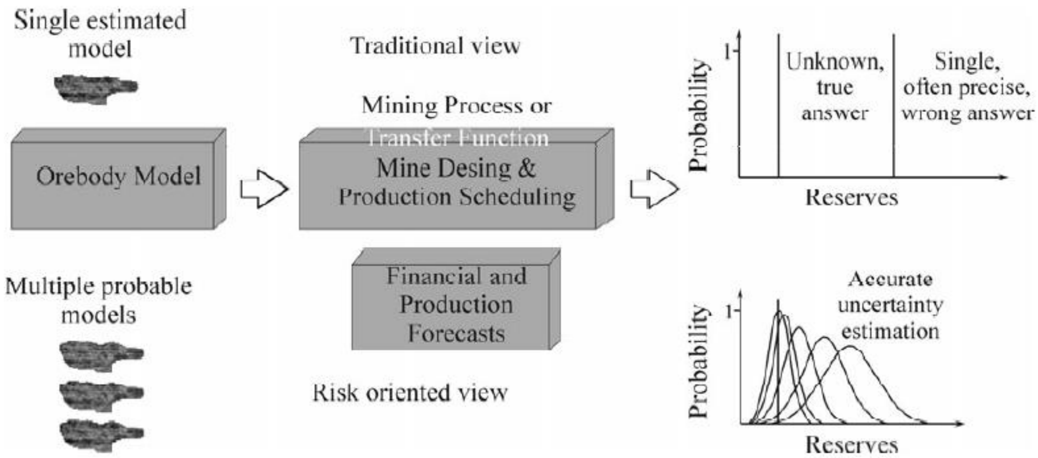

Figure 2.1 - Traditional vs Stochastic approach to reserves evaluation; Source: Strategic Mine Planning Under Uncertainty, Dimitrakopoulos (2011)

2.3 Risk Analysis in mine planning

The Monte Carlo method can be used to assess the uncertainty of future estimates and allows the creation of plans that can mitigate possible risks. Rendu (2002), explains that the risk depends on the parameters that control the value of the project and the uncertainty with which these parameters are known. The objective of risk analysis and management consists firstly of quantifying the uncertainty of a set of factors, characterizing which are the most important factors for the joint uncertainty and in realizing how they can reduce the uncertainties, and what improvements they bring to the value of the project.

Dimitrakopoulos et al. (2002) identify the tonnage and the expected metal content of the deposit to be exploited, as the main technical risk factors in a mining project. With regard to economic parameters, capital costs, operating costs and the price of metals are included as the main factors of uncertainty. In addition to the sources of risk listed, it is worth mentioning that mining projects are subjected to environmental and political challenges, which do not deserve an incisive study in the scope of this dissertation, but can make any exploration unfeasible, even if the geological and economic risks indicate signs of a good investment. Integration in the planning of risk parameters proves to be extremely important, as it allows for more informed decisions in project evaluation, design or management.

Kear (2006) highlights the fact that the mining industry has very long cycles between its different stages, during the useful life. It also compares the aspects between the strategic and tactical planning of a mine, where in the former the objective to be reached is established, usually obtaining the best economic evaluation of a given resource through a detailed plan of operations in the production phase, and in the latter resources are allocated to achieve the respective objective. Neves et al. (2020) point out that many authors consider medium-long-term planning scenarios on a time scale that varies from months to some years. Regarding the optimization of short-term tactical planning, the time scale for precise and detailed solutions, has to be reduced. They also note that, although the two scenarios are complementary, there is no guarantee that the solution found for the short-term horizon will be included in the global optimal solution over a longer horizon and an ideal planning results from a combination of benches to be explored, depending on geotechnical and temporal factors, which are selected according to the maximization or minimization of an objective function, related to metal, tonnage, NSR or NPV.

Maybee and Yana (2017) used techniques imported from the asset management sector, namely the SIM (Single-Index Model), which relates the return on a given asset with the return on the underlying market index. Applied to a case of underground gold mining, the adaptation based on this model relates the internal rate of return for a given exploration area with the change in the gold price in the same period.

For companies, analysts and investors, one of the fundamental parameters for evaluating a project is represented by the NPV (Net Present Value), obtained when discounting future cash flows. Starting from the concept of the value of money over time, in which an investment received today is worth more than an investment received tomorrow, investors are rewarded for accepting that their investment is paid at a later stage of the project. Thus, Gocht et al. (1988) define the discount rate as the cost of capital associated with raising funds from investment outside the company. The NPV is translated into the formula (2.12):

$$NPV=(R_0-C_0)+\frac{R_1-C_1}{1+r}+\frac{R_2-C_2}{(1+r)^2}+\ldots\frac{R_n-C_n}{(1+r)^n} \qquad (2.12)$$

The NPV is the sum of revenues (R) minus costs (C) between year 0 (start of the project) until year n (end of the project), applying a discount rate r. Using this evaluation method, when positive NPV is obtained, it indicates that the estimated revenues are higher than the estimated costs and, therefore, the project must be carried out. Otherwise, the costs will be higher than the revenues and so the project must be abandoned.

In figure 2.2 the evolution of the risk and value of a mining project is represented, highlighting the divergence between these two components in the initial phase. The role of geostatistics and forecasting methods in this phase is essential to bring these two components closer together and to quantify the risk in the best way, optimizing production plans over the life of the project.

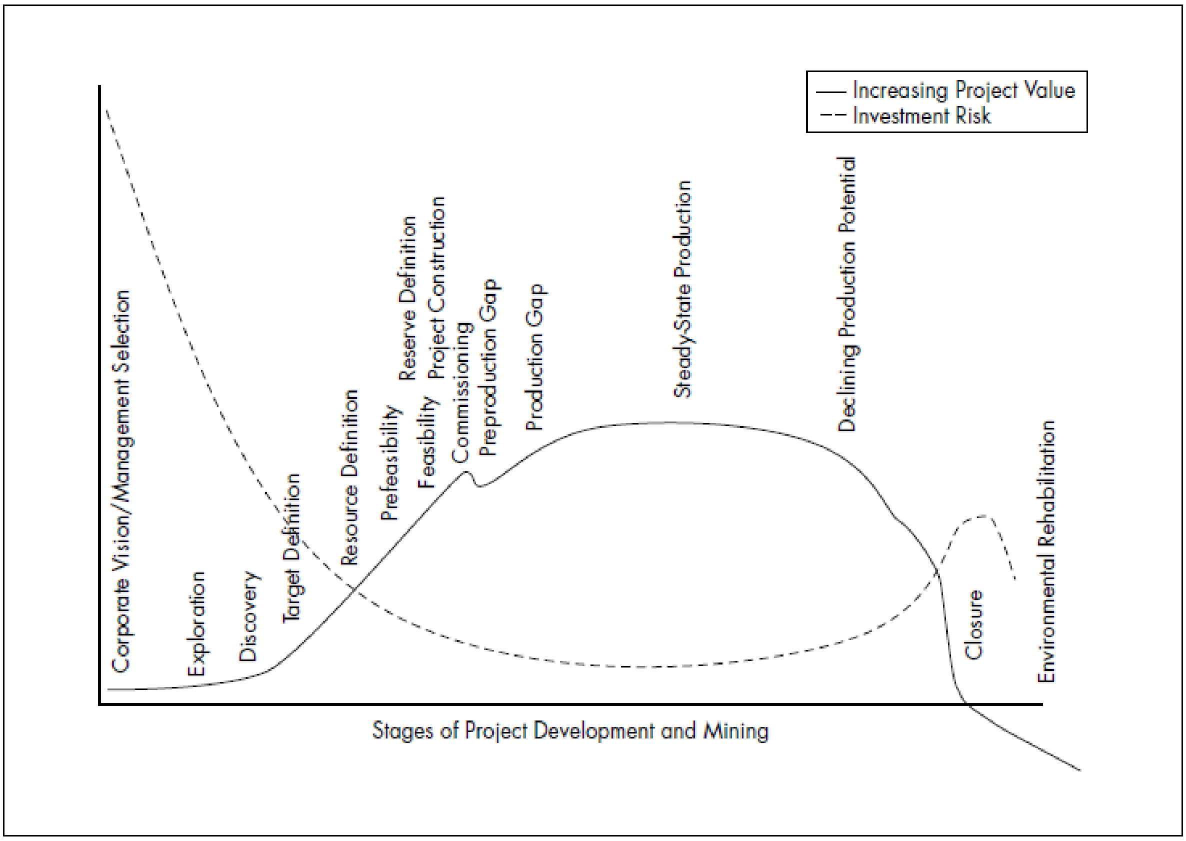

Figure 2.2 - Value of mining project vs investment risk; Source: SME (2011)

Chapter 3 – Methodology

This chapter comprises the methods based on the literature review presented in the previous chapter, to accomplish the objectives proposed in section (1.2). As such, it was divided into the following sections: 3.1 Zinc Price Forecasting; 3.2 Mining Recovery; 3.3 Zinc Head Grade; 3.4 Profit Risk Assessment. It is worth mentioning that for data analysis purposes the open-source cross-platform Integrated Development Environment (IDE) software Spyder for programming in the Python (version 3.8) language was used and that all data used for the present work, except for zinc prices, were obtained through the Neves-Corvo copper, zinc and lead mine that is owned by SOMINCOR (Sociedade Mineira de Neves-Corvo S.A.) which is a subsidiary of Lundin Mining. Despite the fact that copper prices are in almost all-time highs, second only to the price levels in 2011 after the 2008 crisis, the present case-study is focused on zinc as the methodology can also be applied to the remaining metals and based on the assumption that more scientific research for this metal is necessary as the company started zinc production in 2006, compared with copper production initiated in 1989 and, in 2017 an expansion of zinc production denominated ZEP (Zinc Expansion Project) was announced.

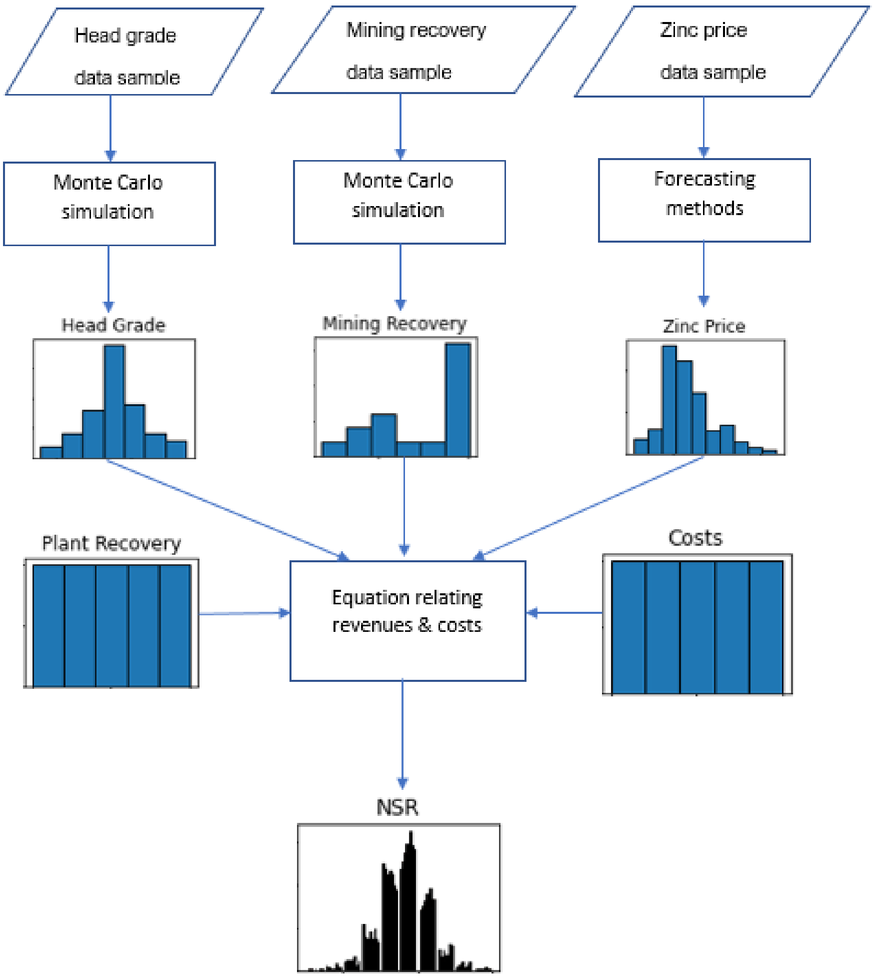

Figure 3.1 - Schematic representation of the methodology

The purpose of this work is to present a method that quantifies the financial risk of a mining investment decision, by calculating the probability distribution of profit distribution, that can be allocated to a tool that translates the performance of a project, Net Smelter Return (NSR) or Net Present Value (NPV). As Armstrong M. (1994) points out, in practice there are several factors that influence profits (e.g., metal price forecasts, reserve estimates) and costs (e.g., fees and labour). The probability distribution for these factors has to be calculated in advance, so that in the end it is possible to combine the various fluctuations of these variables in the calculation of the estimates of possible profits. A diagram summarizing this idea is presented in figure 3.1.

3.1 Zinc Price Forecasting

Forecasting techniques are indispensable for an individual or organizations as addressed by Makiridakis et al. (1998). Dooley and Lenihan (2005) as well as Eggert (1987) took forward this forecasting relevance into the mining context when accounting forecast information for decision making. With this in mind it is important to recognize that forecasting is oftentimes a difficult task and the quantitative methods used in this work should be used in conjunction with each other and with qualitative methods as well.

This section describes the methods used for zinc price forecasting in the short-term horizon, i.e., up until 3 months in the future. For this work, a sample from the Federal Reserve Bank of St. Louis (FRED) comprising monthly prices from February 2006 up until February 2021 was chosen, on the assumption that the data capture enough trend fluctuations and also based on the fact that the methods here presented, give more weight to recent observations when executing forecasts. Figure 3.2 shows zinc price variation in this period.

Figure 3.2 - Zinc Price between Feb 2006 and Feb 2021

Simple Moving Average

Using equation 2.1, moving averages of 3, 10, 30 and 50 periods were performed. Besides not being the main forecast tool for this work context, moving averages smooths the time series resulting in useful information about trend cycles. Using a large number of terms in the moving average increases the chances of randomness being eliminated, the extreme case being SMA (n) where n is the number of observations in the series, which is equivalent to computing the mean of all data. Conversely, SMA (1) is equivalent to using the value from n-1 as a forecast for the period n.

For the realization of this method the data must be stationary, meaning that the process generating the data is in equilibrium around a constant value, the mean, and that the variance around the mean remains constant over time, as noted by Dooley and Lenihan (2005) and Makridakis et al. (1998). The authors also noted that when the dataset is not stationary, day differences and derivatives can be used. For that reason, besides the simple visualization of our dataset that indicates an evidence for stationarity, a test named Dickey and Fuller (DF) statistic was executed.

Testing for the following model in Greene (2003) presented in equation 3.1, where Y_t represents the series in time t and Y_{t-1} in time t-1, γ, a coefficient and ε_t the error term, the test is carried under the null hypothesis of $\gamma=1$, i.e., non-stationarity, against the alternative hypothesis of $\gamma<1$ for stationarity. Calculating a test statistic in equation 3.2, we can compare with the values of the Dickey-Fuller distribution and if the calculated statistic is less than the critical value from the DF distribution, then the null hypothesis is rejected. Python automatically provides the DF statistic and the critical values, using the *adfuller* function under the Statsmodels package.

$$Y_t=\mu(1-\gamma)+\gamma Y_{t-1}+\varepsilon_t \qquad (3.1)$$

$$DF_t=\frac{\hat{\gamma}}{Est.Std.Error(\hat{\gamma})} \qquad (3.2)$$

Given the results in table 3.1 and comparing with a stretch of a Dickey-Fuller table with critical values in figure 3.3, considering 181 observations we can assume with more than 90% confidence that our data is stationary.

Table 3.1 - DF Test Results

Test Statistic	-2.67
P-Value	0.08

	Sample Size			
	25	50	100	∞
AR model with constant (random walk with drift)				
0.01	−3.75	−3.59	−3.50	−3.42
0.025	−3.33	−3.23	−3.17	−3.12
0.05	−2.99	−2.93	−2.90	−2.86
0.10	−2.64	−2.60	−2.58	−2.57
0.975	0.34	0.29	0.26	0.23

Figure 3.3 - Critical Values for DF Statistic; Source: Greene (2013

Simple Exponential Smoothing

Analysing equations 2.1 and 2.2 and given the fact that both the Simple Moving Average (SMA) and the Simple Exponential Smoothing (SES) methods use past observations to forecast, they can be seen as homologous since a low value of α has a similar effect to computing a large moving average and a high value of α is equivalent to taking a low moving average, the main difference being that in SES the weight of each observation decreases exponentially as we move backwards in time, as opposed to SMA where equal weights are given to each observation. Expanding equation 2.2 by replacing F_t with its components as shown in equation 3.3 and conceptualizing that the same exercise can be done for the F_{t-1} term and for the remaining terms until the first observation, it becomes easy to understand that the weights decrease exponentially and in the extreme cases where α equals 0 and 1, the next forecast will account for all the observations or only for the last, respectively.

$$F_{t+1} = \alpha Y_t + (1-\alpha) F_t \tag{2.9}$$

$$F_{t+1} = \alpha Y_t + \alpha(1-\alpha) Y_{t-1} + (1-\alpha)^2 F_{t-1} \tag{3.3}$$

Selecting α can become a cumbersome task as it is usually determined by trial and error and an optimization criterion can be difficult to attain. For this work it was chosen to find the best α minimizing MAPE using the Truncated Newton method.

Double Exponential Smoothing

Contrary to Simple Exponential Smoothing (SES) method explained above which gives a flat forecast, i.e., the forecast made for one period ahead is the same for the subsequent periods, Holt's linear method updates new forecasts with the underlying trend of the data through the constant β in equation 2.4, that controls the speed of adjusting the trend. In other words, if the trend changes very quickly the value of β must be sublime. As in SES method, α and β were chosen to minimize MAPE and using the same nonlinear optimization method (Truncated Newton), of which a stretch of code[1] using Python is shown in figure 3.4. It is worth mentioning that using this way of optimizing can cause the smoother parameters to get to their extremes (0 or 1) only to achieve an optimal solution based on the minimized error, representing a model with poor predictive performance.

[1] Inspired by Andre Queiroz. Available at: https://gist.github.com/andrequeiroz/5888967

```python
#import libraries
from scipy.optimize import minimize

#function that minimizes the type of error
def minimiza(tipo_erro,serie):
    #arbitrary values to initialize error minimization
    alpha = 0.1
    beta = 0.3
    #establish boundaries to the values of alpha and beta (between 0 and 1)
    boundaries = [(0,1),(0,1)]

    #make a copy of the time series values
    Y = serie[:]

    #optimization using Truncated Newtown
    if tipo_erro == 'mape':
        parameters = minimize(MAPE,x0 = (alpha,beta),method = 'TNC',\
                              args = (Y),bounds = boundaries)

    alpha,beta,erro = parameters.x[0],parameters.x[1],parameters.fun

    return alpha,beta,erro
```

Figure 3.4 - Code with function to minimize the MAPE

3.2 Mining Recovery

In mine exploitation, considering a specific area of production, not all the material considered as valuable by the mine planning is extracted. That proportion between the actual quantity extracted and the quantity planned to be extracted is defined as the mining recovery. Sometimes, there is a tonnage increase in that process, given by the dilution parameter, often seen as the mass of waste mined which was not planned. To avoid more valuation fluctuations as this dilution factor may as well contain payable values according to McCarthy (2001), in the context of this work the dilution parameter was considered null, not forgetting that this methodology is not incompatible with a future approach for this problem considering this dilution as one more parameter with uncertainty.

Data analysis was carried out using a sample from 17 observations from different areas of the Neves-Corvo mine where the underground mining method Bench-and-Fill was used.

Univariate Analysis

This simple form of analyzing data reveals data patterns using histograms and describes the sample through measures of central tendency like the mean and median, as well as dispersion measures such as quartiles, minimum, maximum and standard deviation.

Graphical methods of modelling a probability distribution

Selecting the probability distribution when historical data is available can be achieved using graphical ways of analysing the sample. As a histogram of one sample can be very useful in creating a first impression of which probability distribution the data might represent, usually comparing the quantiles of the sample data to the quantiles of a specified distribution and noting if the graphic displays a goodness of fit, can bring more certainty to our first intuition, in a method denominated as QQ (Quantile to Quantile) Plot.

As noted by Wilk and Gnanadesikan (1968), for one-dimensional samples, the empirical distribution function of the sample quantiles is computed using equation 3.4, where n is the number of observations and i the i-th ordered value.

$$F[x(i)]=w_i=\frac{(i-0.5)}{n},\ i=1 \text{ to } n \quad (3.4)$$

For n = 17, the set of probabilities, P_s, is presented in table 3.2. Comparing the fitting of the data with a known probability distribution with cumulative distribution function F(x), each corresponding quantile is obtained using the inverse function as in equation 3.5.

$$q_i= F^{-1}(w_i),\ i=1 \text{ to } n \quad (3.5)$$

Considering the standard normal distribution, using a Z Table or computing using a statistical software, the previous equation gives a set of theoretical quantiles, as shown in the Z column of table 3.2. Creating a scatter plot with the ordered values of the data in the y-axis and the theoretical quantiles in the x-axis, if the scatter plot looks like a 45º line (slope 1) x and y are considered identically distributed variables.

Table 3.2 - Critical values for the cumulative standard normal distribution

Sample	P_s	Z	Sample	P_s	Z
1	0,029	-1,890	10	0,559	0,148
2	0,088	-1,352	11	0,618	0,300
3	0,147	-1,049	12	0,676	0,458
4	0,206	-0,821	13	0,735	0,629
5	0,265	-0,629	14	0,794	0,821
6	0,323	-0,458	15	0,853	1,049
7	0,382	-0,300	16	0,912	1,352
8	0,441	-0,148	17	0,970	1,890
9	0,5	0			

Importing the following modules: *pylab* for plotting and *stats* which contains many statistical functions from the *scipy* open-source software to Python, it is easy to calculate and plot quantiles for a probability plot using the *probplot* function. Figure 3.5 shows the code to achieve these calculations and plot the graph comparing the fitting of the data with the normal distribution, specifying the respective parameters, mean and standard deviation. Keeping in mind Thomopoulos (2013) advice in 2.2, comparing the data with more common distributions such as exponential, lognormal, gamma, beta and Weibull can be achieved with a similar code represented below, modifying the distribution name and finding the respective parameters for each distribution.

```
#import libraries
from scipy import stats
import pylab

#plots the probability plot for the normal distribution
stats.probplot(recuperacao['Recuperacao'],sparams = parameters,\
    dist = 'norm',plot=pylab)
```

Figure 3.5 - Code for the probability plot using the normal distribution

Kolmogorov-Smirnov test to fit the best probability distribution

Selecting a distribution that best fits a sample data can also be examined in conjunction with the Kolmogorov-Smirnov (KS) test that can be viewed graphically as the maximum deviation from the diagonal line on a PP (Probability-Probability) Plot as exposed by Wilk and Gnanadesikan (1968) and in equation 2.9. This test is used to test H_0: the sample comes from a theoretical distribution P, against H_1: the sample does not come from P.

The first step towards a response is to compute the empirical cumulative distribution function given by P_s in table 3.2 and used as $F_{obs}(x_i)$ in table 3.3. Testing again the hypothesis that the data distribution comes from a standard normal distribution, P, it is easy to compute $F_{exp}(x_i)$ following equation 3.6 using a Z normal table or a statistical software. Computing the absolute differences as in equation 2.9, the Kolmogorov-Smirnov statistic gives a maximum of 0.681 as show in red in table 3.3. Finally, the critical value at 95% significance level for a sample of 17 observations is 0.286 as seen in figure 3.6 of a trench of the critical values of the Kolmogorov-Smirnov test statistics. Since 0.681 > 0.286, we reject the null hypothesis, meaning that the data does not derive from a standard normal distribution.

$$F_{exp}(x_i) = P(Z \leq x_i) \qquad (3.6)$$

One-Sided Test: α =	.10	.05	.025	.01	.005
n=1	.900	.950	.975	.990	.995
16	.258	.295	.327	.366	.392
17	.250	.286	.318	.355	.381
18	.244	.279	.309	.346	.371

Figure 3.6 - Critical values for the KS Test; Source: Miller (1956)

Table 3.4 – Absolute differences between the empirical and expected CDFs

Sample	Recovery	$F_{obs}(x_i)$	$F_{exp}(x_i)$	$\lvert D_n \rvert$	Sample	Recovery	$F_{obs}(x_i)$	$F_{exp}(x_i)$	$\lvert D_n \rvert$
1	0,5518	0,029	0,709	0,680	10	0,8762	0,559	0,810	0,251
2	0,7376	0,088	0,770	0,681	11	0,8826	0,618	0,811	0,194
3	0,7835	0,147	0,783	0,636	12	0,8826	0,676	0,811	0,135
4	0,7886	0,206	0,785	0,579	13	0,8830	0,735	0,811	0,076
5	0,8028	0,265	0,789	0,524	14	0,8893	0,794	0,813	0,019
6	0,8114	0,324	0,791	0,468	15	0,8900	0,853	0,813	0,040
7	0,8138	0,382	0,792	0,410	16	0,9006	0,912	0,816	0,096
8	0,8438	0,441	0,801	0,359	17	0,9008	0,971	0,816	0,154
9	0,8597	0,500	0,805	0,305					

Following the above conclusion, we can either repeat the test until we achieve a good fitting of the data, using other parameters such as a different mean and standard deviation for the normal distribution, or testing against other suggested distributions in the literature. Using the function *kstest* from the *stats* model in Python, values for the KS test statistic can be easily computed, as shown in figure 3.7, where the data is fitted to a beta distribution using the maximum likelihood estimation (MLE) method and, subsequently the KS test is performed with the parameters that best fit the data to the beta distribution.

```
#import libraries
from scipy import stats

#Fits the data to a beta distribution and performs the KS Test
dist = getattr(stats,'beta')
parameters = dist.fit(recuperacao['Recuperacao'])
ks_results = stats.kstest(recuperacao['Recuperacao'],'beta',parameters)
```

Figure 3.7 - Code to compute the KS Test using the beta distribution

Analytical methods of modelling a cumulative distribution function

Computing a histogram gives an understanding of the possible distribution that the numerical data might take. In that sense, it is possible to make an estimation of the probability density function (PDF) and if the equation is computed, the cumulative distribution function (CDF) can be obtained using equation 2.10. An alternative to the estimation of the CDF of a variable can also be achieved plotting the graph of the empirical CDF given by $F_{obs}(x_i)$ in table 3.3, on the y axis against the actual sample values as in the second column of table 3.3, on the x axis, as shown in the graph of figure 3.8.

Analysing figure 3.8 gives a sense that the exponential curve fitting might be the equation that best fits the CDF, similar to experiments on radioactive traces in this case rising in time in which one of the most widely accepted forms, Foss (1970), is given in equation 3.7.

$$y = ae^{-bx} \tag{3.7}$$

It is worth noting that one might consider the value for recovery around 0.5518 an anomalous value and disregard it, where in that case polynomial interpolation and smoothing curve fitting, Akima (1970), could be the best approximation for the CDF.

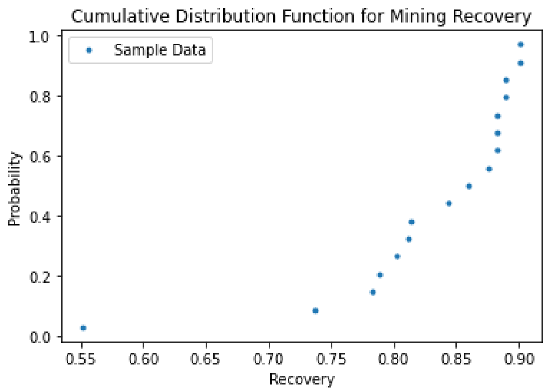

Figure 3.8 - CDF function for the Mining Recovery

Now, it becomes easy to obtain the a and b parameters of equation 3.7 using the *curve_fit* function from the *scipy* open-source library using the non-linear least squares method to fit the function y to data. Next, given the nature of the equation the inverse function is easily obtained. Generating pseudo-random numbers from a standard uniform distribution and, then using equation 2.11 generates a random number from the CDF derived from equation 3.7. The process to attain the Monte Carlo simulations for the mining recovery is shown in figure 3.9.

```python
#import libraries
from scipy.optimize import curve_fit
from sklearn.metrics import r2_score
import numpy as np

#definition of the CDF and the respective inverse
def expo_func(x,a,b):
    return a * np.exp(-b * x)

#order the values of recovery
x = np.sort(recuperacao['Recuperacao'])

y = []

#probability assignement in the empirical CDF
for i in range(1,len(x)+1):
    y.append((i-0.5)/len(x))

#fits the expo_func to the data and gives parameters a and b
popt,pcov = curve_fit(expo_func, x, y)
y_pred = expo_func(x, popt[0],popt[1])

#proportion of data points which lie within the regression equation
r2 = r2_score(y,y_pred)

#set the seed to 15
np.random.seed(15)

#generate pseudo-random values between 0 and 1 where 12 represents the number
#of weeks in 3 months and 400 the number of simulations per week
values = np.random.rand(400,12)

simulated_recovery = []

for i in values:
    simulated_recovery.append(np.log(i/popt[0])/-popt[1]) #inverse function
```

Figure 3.9 - Code to fit the data to an exponential function and generate pseudo-random numbers using the inverse function

3.3 Zinc Head Grade

Taken as one of the riskier factors in a project valuation, head grade can be viewed as the quantity of metal present in the material entering the processing plant during production. When the mine is still in its pre-feasibility and feasibility stages, grade distribution becomes a factor of great uncertainty as our knowledge of the deposit characteristic is limited. As pointed by Rendu (2002), reducing this uncertainty can be accomplished taking additional sampling as well as finding geostatistical methods that use simulations to model the grade distribution within a deposit. For this purpose, Dimitrakopoulos et al. (2019) propose a Stochastic Integer Programming (SIP) approach to manage geological uncertainty. Also, Dimitrakopoulos et al. (2002) appointed the advantages of using Conditional Simulation (CS) for modelling uncertainty. Given the fact that is often difficult to attain samples of a deposit to use the techniques above to estimate its grade distribution, in the context of this work was used historical observations of grades entering the processing plant to model the geological uncertainty.

Univariate Analysis

Assuming that there is uncertainty in all stages of a mine's lifetime as we only know the exact composition of the orebody when it is fully extracted and also given the fact that the geostatistical methods presented above are time consuming and technical challenging, an alternative to model this uncertainty was necessary using the actual values of grade and modelling their uncertainty. For that purpose, 59 observations of head grades feeding the Zn processing plant of the Neves-Corvo mine were obtained. From that, 10 observations with grade of 0% were dumped as in those days when the values where registered, the processing plant was fed with waste to produce material for the backfill.

Graphical methods of modelling a probability distribution

For the choice of the probability distribution that best fits the data acquired, the same methodology used for the mining recovery variable was applied. Due to the fact that the graphical methods and the KS test gave satisfactory results, as discussed in the next chapter, there was no need to apply the analytical methods to perform the Monte Carlo simulations. Instead, using the *numpy* library in Python, it was possible to generate values from the suited distribution specifying the respective parameters. For the purpose of discovering the number of trials for the Monte Carlo simulation, given the fact that the sample for zinc head grades contained a greater number of observations in comparison with the data obtained for the mining recovery, table 3.4 was computed where a number of simulations per week was generated and the respective mean and standard deviation errors were calculated in which μ_s and μ_{MC} represent the mean of the sample and the mean of the Monte Carlo simulations and s_s and s_{MC} the sample standard deviation and Monte Carlo standard deviation, respectively. In that regard, 400 simulations per week were chosen to be performed as it was the only case where an error of less than 5% was registered both for the mean and standard deviation.

Table 3.5 – Mean and standard deviation errors for the Monte Carlo simulations

Number of simulations per week	Mean Error $\left\|\frac{\mu_s-\mu_{MC}}{\mu_s}\right\|$	Std Error $\left\|\frac{s_s-s_{MC}}{s_s}\right\|$
100	0.1339	0.5120
200	0.2279	1.3537
300	0.1213	0.3399
400	0.0216	0.0297
500	0.0105	0.5394
1000	0.0286	0.6197
2000	0.0809	0.1871
3000	0.0487	0.3206
4000	0.00105	0.5044
5000	0.0072	0.6001
10000	0.02205	0.9615

3.4 Profit Risk Assessment

Various works on ore valuation using NSR have been carried out throughout the last three decades, Annels (1991), Hustrulid (2013), Wills and Finch (2016) all of which agree that this concept was created in polymetallic base-metal mines to describe the revenue received from the smelter for the concentrate produced, Bargmann (2000), Goldie and Tredger (1991). However, they differ in accounting for commercial costs, as some include smelter deductions, treatment and refining costs as well as distributions costs on the NSR calculations like Goldie and Tredger (1991), while others take into account only smelting and refining charges ignoring costs such as freight and insurance, Hustrulid (2013).

Adopting Goldie and Tredger (1991) approach NSR can be related to Gross Revenue and Commercial Costs as in equation 3.8:

$$NSR = Gross\ Revenue - CC \tag{3.8}$$

Gross Revenue can in turn be defined as the product of the price of metal and the quantity of concentrate produced in the processing plant, as equation 3.9 shows:

$$Gross\ Revenue = Metal\ Price * Quantity\ of\ Metal \tag{3.9}$$

Finally, quantity of metal can be defined, as in equation 3.10, by the product of the tonnage of mine, the grade z(x), the mining and processing plant recoveries. The operational margin of mine can be attained subtracting the NSR, also known as net revenue, from the operational costs presented in table 3.5, as equation 3.11 shows.

$$Quantity\ of\ Metal = Ton * z(x) * MR * PR \tag{3.10}$$

$$Operating\ Margin = NSR - Operating\ Costs \tag{3.11}$$

It is worth mentioning that due to the complex nature of the subject, time constraints and ultimately diverse information regarding the studies of the NSR a simple explanation was presented herein. For example, aspects of some importance denoted by Wills and Finch (2016) such as impurities in the concentrate which may be penalized by the smelter like arsenic and mercury or metals that may produce a bonus such as silver were overlooked. As a consequence, the intricate nature of contracts with smelters can be affected because a high concentration of impurities will eventually change the strategy of exporting the concentrate product to European smelters to maximize sales reducing freight and logistics cost as the “local” zinc smelters may not have the technology to remove such content.

Although insurance, marketing and transportations costs are part of G&A costs they do not make the whole bulk. It is important to recognize that transportation costs are 100% variable and are dependent on the smelters where the concentrate is sent and are anchored to annual contracts from 1 to 5 years that are mostly confidential. In the absence of more precise data on commercial costs this approach gave results of mine netback with more than 100% error comparing with the results in the literature.

Table 3.6 - Production Costs

Opex	Units	Value
Mine	$/t milled	28.29
Plants	$/t milled	13.55
Water & Tailings	$/t milled	3.05
G&A	$/t milled	8.94
Total	$/t milled	53.83
Forex	EUR/USD	1.176

Additionally, it is important to define the concept of cut-off grade which is the minimum amount of metal that one ton of material must contain before this material is sent to the processing plant, Rendu (2014), helping distinguish ore from waste, seeing as the former profitable to exploit while the latter it is not. For this work the equation below (3.12) relating the Production Costs (PC) with the metal price, mining and plant recoveries (MR and PR respectively) and the NSR in percentage, was used to evaluate this parameter.

$$Z_0 = \frac{PC}{Metal\ Price*MR*PR*NSR} \qquad (3.12)$$

In their study about NSR and its use in polymetallic deposits, Goldie and Tredger (1991) discussed the help of an NSR model in determining cut-off grades in the exploitation phase, the dynamic characteristic of underlying parameters such as fluctuating metal prices and recoveries as well as its unstable impact in mine planning as constant revisions were required in a lengthy manual labour environment. In today's era where every mine has mine planning data stored in a computer, as noted by the authors, this approach became more useful as it gave mine planners the resources to quickly adapt to new paradigms of the market, or technical constraints in the extraction or processing of an ore.

Still in accordance with the aforementioned, a value of 0.5 was used for the NSR in equation 3.12, as Wellmer et.al (2008) presented as a typical value for a Zn mine return. Goldie and Tredger (1991) defined this value as mine netback given, as the ratio of net revenue to gross revenue, and presented a typical Canadian mine netback of 43%. The selection for the former instead of the latter comes with the assumption that Wellmer et.al (2008) achieved this result with studies conducted more recently and with more updated data. More importantly, this choice of taking mine netback constant instead of calculating on a weekly basis, as achieved by the Monte Carlo simulations is justified by the fact that the commercial costs were taken as part of the General and Administrative (G&A) costs.

Chapter 4 - Discussion of Results

This chapter presents the results of the application of the methodology, demonstrated in the previous chapter, to the variables on which the NSR and cut-off grade are dependent and, brief comments on the results that are obtained are also presented.

4.1. Zinc Price Forecasting

Table 4.1 shows the descriptive statistics of the zinc price from February 2006 to February 2021. As it can be seen, 181 observations were made available and using the simple average forecasting technique, we can take the value of mean of 2344.845 USD as the short-term zinc price forecast for the next three months. This result should be critically evaluated as, comparing this value with the last observation available from February 2021, around 2700 USD, as seen in figure 3.2, gives a spread of approximately 350 USD which rarely happens in the data set horizon. Also, conversations about how a possible metals supercycle driven by demand for metals in clean power facilities and transportation, can drive commodities prices high for years or even decades are surging and should not be ignored.

Table 4.1 – Descriptive Statistics of the Zinc Price

Observations	181
Mean	2344.845
SD	600.651
Min	1112.905
25%	1928.011
50%	2273.012
75%	2616.290
Max	4381.447

Simple Moving Average

Using equation 2.1, moving averages with 50, 30, 10 and 3 periods were performed. Although values of forecast for this method are not used because exponential smoothing methods are considered more robust, Simple Moving Averages (SMAs) can still give us some information about trends and higher order forecasts can eliminate some randomness as described in section 3.1. Figure 4.1 shows Simple

Moving Average with 50 terms, SMA (50) represented with a purple dashed line. With this scenario, important cycles can be smoothed out and a lag effect can have a negative influence if we are more interested in short-term forecasts. Contrastingly, SMA (30) and SMA (10), moving averages with 30 and 10 terms, red and green dashed lines respectively, capture the effect of upward prices in 2016 and the followed downturn in 2018 and besides the SMA (30) still not capturing the upward effect in early 2020, SMA (10) and SMA (3) give a good sign that we can have a rise in zinc prices in the near term.

Table 4.2 shows the MAPE error calculated for the training set using equation 2.8 in which each range can be seen in the x axis of figure 4.2 where the plot of the percentage errors is represented using equation 2.6. From that information it is worth noting the case for the SMA (10), where besides having the greater MAPE possibly because of a greater error due to over-forecasting in the beginning of the estimations (2008 to 2010), the most recent forecasts are less biased as their percentage errors do not vary too much around 0%. On the other side, SMA (3) presents the smallest value of MAPE, justified also by the fact that the percentage error does not vary too much around 0%, but it does not add too much information when analysing its importance in figure 4.1 possibly because the algorithm is more sensitive to randomness.

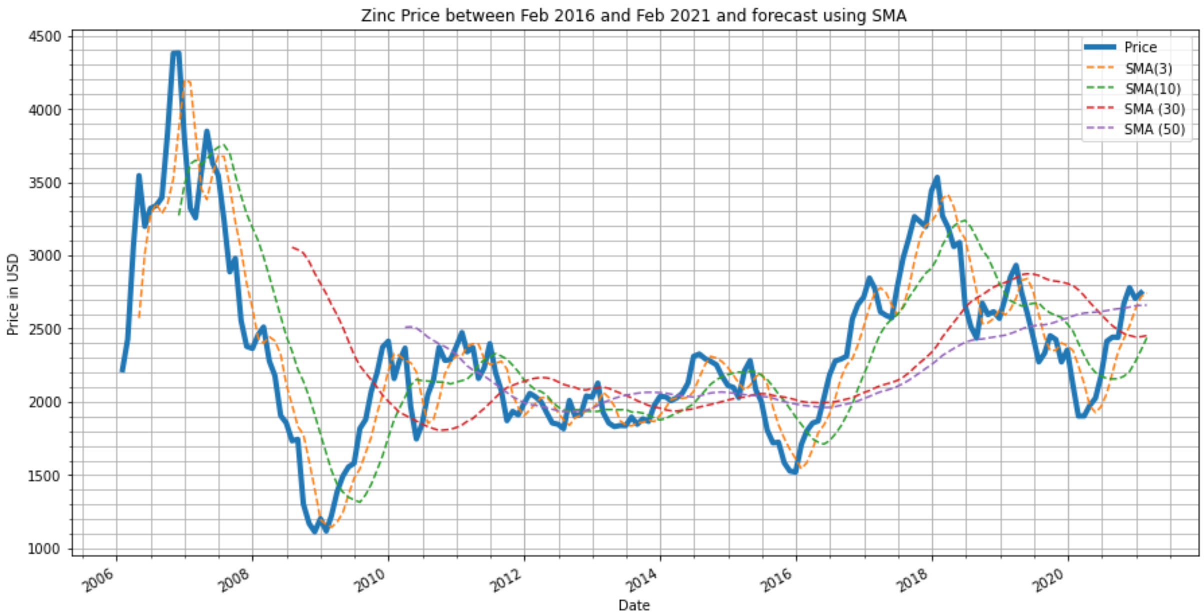

Figure 4.1 - Forecast using Simple Moving Averages

Table 4.2 – Training Error Measures

Moving Average	SMA (3)	SMA (10)	SMA (30)	SMA (50)
MAPE (%)	8.098	13.857	21.054	11.999

Figure 4.2 - Percentage Error Plots for SMA (3), SMA (10), SMA (30) and SMA (50)

Simple Exponential Smoothing

The Simple Exponential Smoothing (SES) method, described in section 3.1, was applied optimizing alpha from equation 2.2 and minimizing MAPE, giving an error of 5.400%, with alpha equal to 1. The smoothing effect and the percentage error plot are shown in figures 4.3 and 4.4, respectively. The results here presented should be interpreted carefully as choosing alpha to minimize an error criterion can result in a smoothing effect where the algorithm gives all the weight to the previous observed value, which happens in this case, potentially due to the fact that rapid fluctuations in the zinc price are present. For that reason, the SES method was also performed with alpha equals to 0.2 which results in a MAPE of 11.781%, where the algorithm adopts a much smoother effect and a slower reaction to prices changes.

Another important factor to be noted in these results is that the algorithm does no project any trend for the future, which means that the last observed value of 2744.503 USD (α=1) or 2546.655 USD (α=0.2) is the forecast for the next month (March) and that value is kept “flat” for the following months (April and May). As pointed by Makridakis et al. (1998), the flat forecast function is used because SES works best with data that have no trend, seasonality or other underlying pattern, which it is not our case.

Figure 4.3 - SES for Zinc Price

Figure 4.4 - Percentage Error Plot for SES

Double Exponential Smoothing

In accordance with the method Double Exponential Smoothing explained in section 3.1, using equations 2.3, 2.4 and 2.5 and following the issues pointed for the SES method in the previous section, a forecast using Holt's method was performed with $N_0 = 2219.725$ and $T_0 = 207.921$. Optimizing alpha and beta using the same non-linear optimization algorithm in SES gave a value for alpha = 1, beta = 0.4 and MAPE of 8.810% for which the graphical representation of the smoothing effect can be seen in figure 4.5. Analysing the results and comparing with the literature, it can be assumed that the error carries an acceptable value as, for a data set of 30 observations, Makridakis et al. (1998) achieved a value around 5% for the MAPE. However, the values of alpha and beta do not represent correctly the data behaviour, as a high value of alpha only takes into account the previous observation for the next prediction and a value of 0.4 for beta is too low and somewhat closer to the value of 0.1 used by Sharif and Hasan (2019) where there was a clear upward trend in the data.

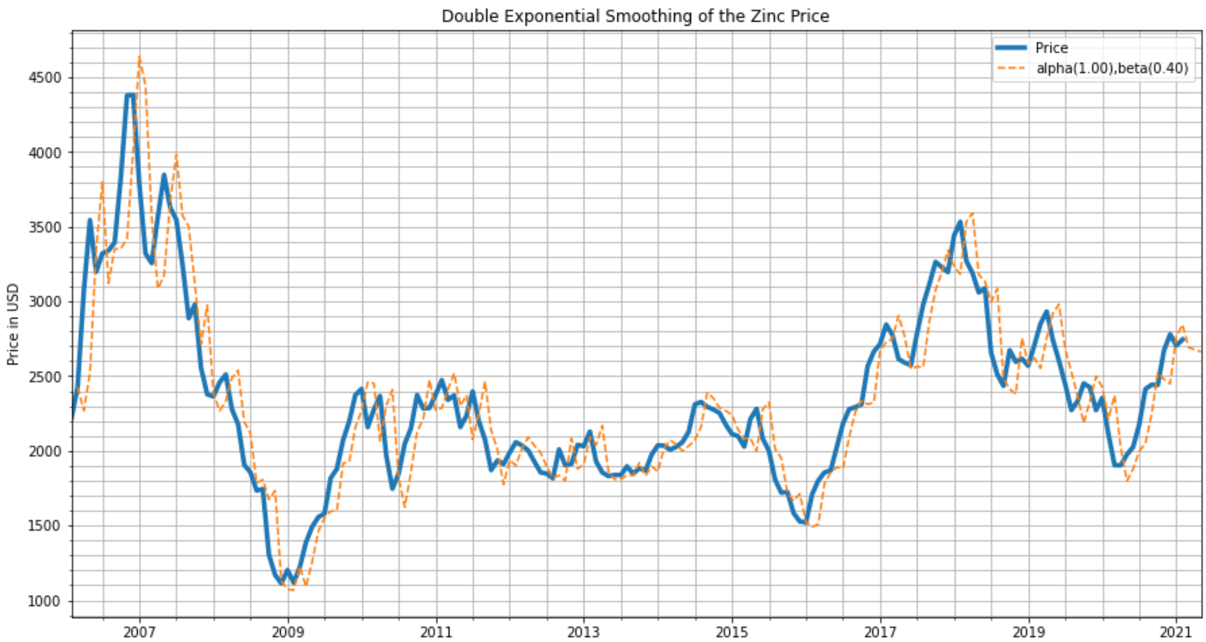

Figure 4.5 - Double Exponential Smoothing with alpha (1) and beta (0.4)

In an attempt to better adjust such parameters to our data, the method was performed in various combinations, as shown in table 4.3, with lower alpha and higher beta. Values of alpha of 0.5 and 0.6 smoothed the data in a similar way and higher values of beta gave lower errors until achieving a minimum around 0.7 and beginning to increase again with higher values. Accordingly, the combination with alpha = 0.6 and beta = 0.7 was chosen as this was the iteration that gave the minimum error and whose values appeared to be best adjusted, as shown in figure 4.6 and in the close-up from January 2020 until May 2021 in figure 4.7.

Table 4.3 – Attempts to better adjust the alpha and beta parameters to the dataset

Combination	α=0.5,β=0.4	α=0.5,β=0.7	α=0.5,β=0.8	α=0.6,β=0.7	α=0.6,β=0.8
MAPE	10.123	9.922	9.929	9.658	9.716

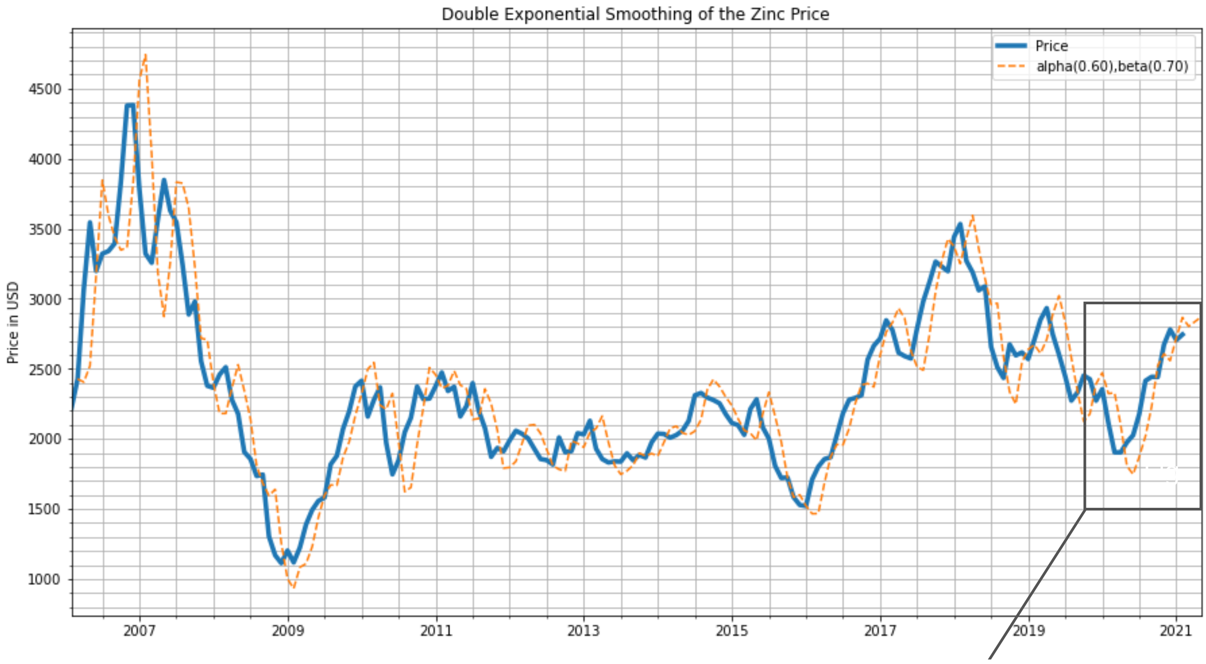

Figure 4.6 - Double Exponential Smoothing with alpha (0.6) and beta (0.7)

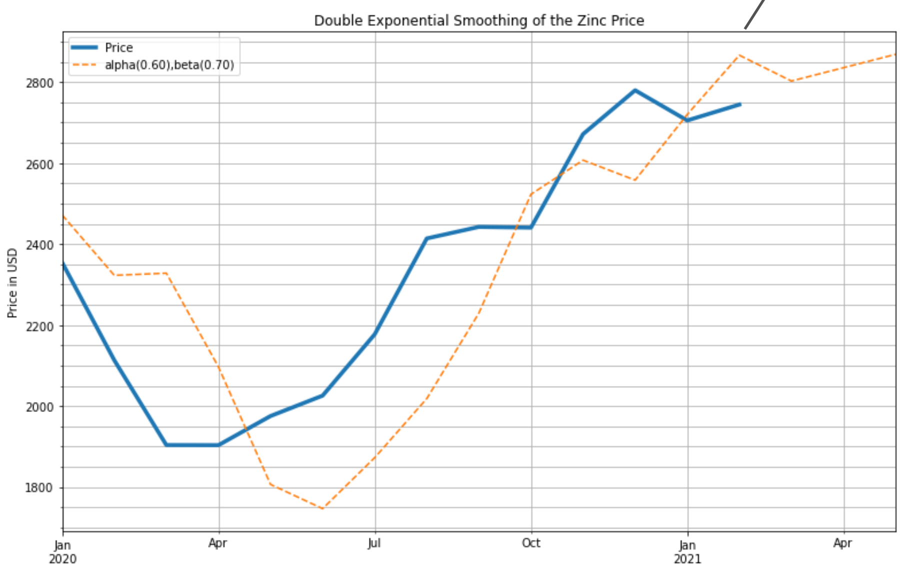

Figure 4.7 - Zinc Price Forecast

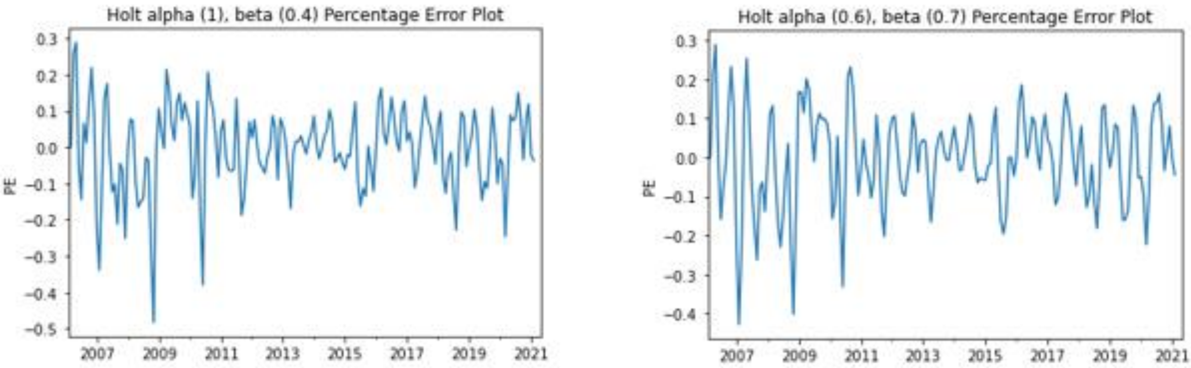

Figure 4.8 - Percentage Error Plots of two scenarios with different smoothing parameters

Observing figures 4.6, 4.7 and the plot of percentage errors for the Holt's method in figure 4.8, it is important to note that the first scenario (alpha = 1 and beta = 0.4) gives a forecast with a downward forecast not conforming to the upward trend in the data started April 2020. In the contrary, the second scenario (alpha = 0.6 and beta = 0.7) gives the correct trend for the forecasts, with more observations captured in the forecast (lower alpha) and more responsiveness with trend fluctuations (higher beta). Finally, the error plot indicates that both scenarios are similarly biased with a difference in MAPE of 0.8 percentage points approximately.

One might consider this to be a high-risk scenario as these predictions are pointing to high values of zinc. With that said, in a business environment it is crucial to consider the uncertainty associated with economic recoveries as some countries are still facing lockdowns due to the COVID-19 pandemic despite the vaccine rollout in developed countries and its possible migration to less developed nations give some optimism about a rebound in zinc demand. Also, as pointed by the S&P Global, an almost 3% percent rise in zinc production to 14 million mt is expected this year, so, with such diversified information about the future, it becomes crucial to have sophisticated judgments of experts in order to adjust current trends for anticipated changes, Eggert (1987).

4.2. Mining Recovery

This section describes the results of analysing mining recovery data, fitting it to a known distribution using graphical methods and its results, as well as adjusting an equation to the empirical cumulative distribution function (CDF) of the data.

Univariate Analysis

To better access the behaviour of the mining recovery data, descriptive statistics and data visualization using a histogram were performed and presented respectively in table 4.4 and figure 4.9. From their analysis, a mean of 0.829 is extracted and a low dispersion of the data represented by a low standard deviation is observed. Also, it is worth noting that although the spread between the minimum value and the mean is of almost 30 percentage points, this is offset in the standard deviation calculation as 75% percent of the data is in the range of [0.803,0.900].

Table 4.4 - Descriptive statistics of the mining recovery

Observations	Mean	SD	Min	25%	50%	75%	Max
17	0.829	0.087	0.552	0.803	0.860	0.883	0.900

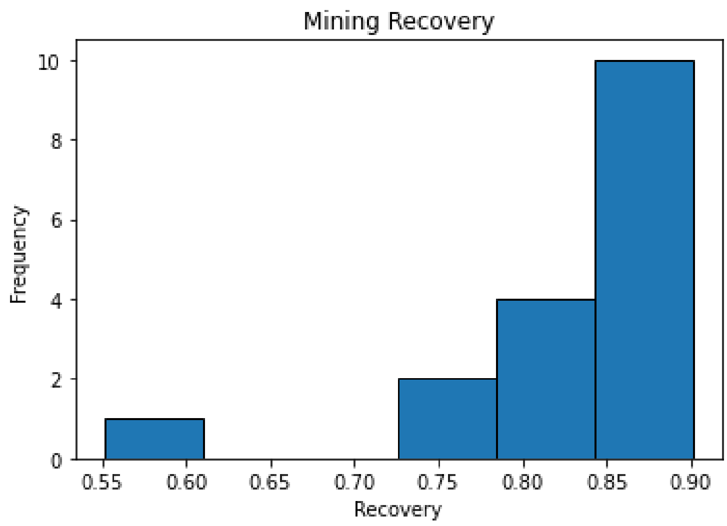

Figure 4.9 - Histogram of the mining recovery

The spread between the minimum value and the 0.25 quantile is also great, around 25 percentage points, which indicates a lack of representativity of the data in this range. Finally, the visualization of the histogram can give an indication of possible distributions functions that can fit the data.

Graphical methods of modelling a probability distribution

In the previous chapter the steps to achieve the plots of the ordered sample values against the quantiles of the reference distribution corresponding to any of the fractions, as stated by Wilk and Gnanadesikan (1968), using the standard normal distribution were presented in section 3.2. Next, still in section 3.2, it was viewed how the Kolmogorov-Smirnov (KS) test can help us accept or reject the hypothesis that our data comes from the chosen distribution, comparing the maximum distance between the theoretical and empirical CDFs, using equation 2.1, with a critical value for the KS Test with 95% confidence level.

This section displays the results for the KS test and probability plot for the common distributions, normal, exponential and beta using its respective parameters, in table 4.5 and figures 4.10, 4.11 and 4.12.

Table 4.5 – KS Test results for the normal, exponential and beta distributions

Distribution	Normal	Exponential	Beta
Statistic	0.1972	0.4485	0.1676

Keeping in mind the critical KS value of 0.286 in figure 3.6 and the literature statement that a good fit is obtained when the probability plot looks like a 45º line from the lower left corner to the upper right corner of the graph, Thomopoulos (2013), the hypothesis that the data comes from an exponential distribution is promptly rejected as a value of 0.4485, greater than the critical KS value, was obtained and its probability plot does not follow the rule in the literature. The rest of the distributions have a statistic value below 0.286, for which the null hypothesis that the data either comes from a normal or beta distributions, can be accepted with 95% confidence.

In fact, one can argue that the beta distribution better fits the data as a lower value of the KS statistic was achieved, in comparison with the normal distribution, and as the line of its probability plot looks more approximate to an angle of 45º.

However, despite the fact that the software used in this work has tools to easily obtain and modify the probability distribution parameters and hence it becomes easy to perform simulations given that distribution, due to the lack of familiarity with the beta distribution and its respective parameters (shape, location and scale), associated with the lost in control that such abstraction could bring, it was decided to progress with the realization of Monte Carlo simulations using the inverse transform method of which the results are discussed in the next section.

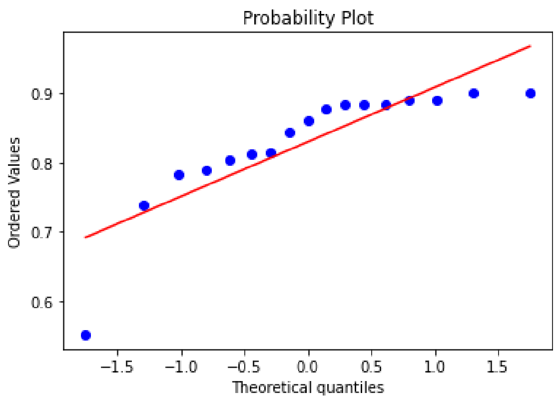

Figure 4.10 - Probability Plot for the normal distribution

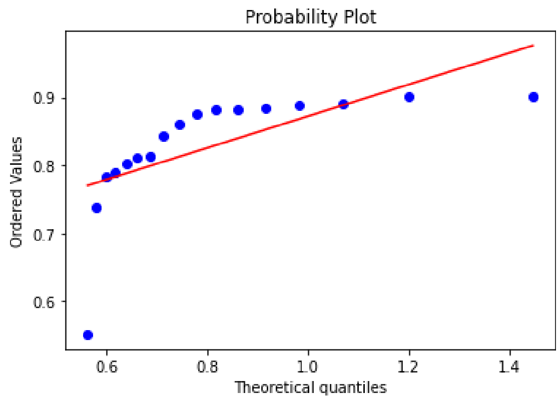

Figure 4.11 - Probability Plot for the exponential distribution

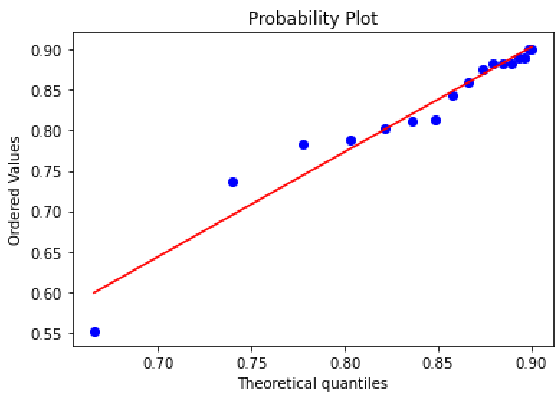

Figure 4.12 - Probability Plot for the beta distribution

Analytical methods of modelling a cumulative distribution function

Following the brief discussion on the analytical methods in the previous chapter and, as a way of asserting that equation 3.7 best fits the cumulative distribution function (CDF) in figure 3.8 as hypothesized, figure 4.13 shows the CDF of the mining recovery and its respective best fitted equation with parameters a = 5.96×10^{-6} and b = -13.25. It is notable the high value of 0.969 for the coefficient of determination (R^2) provides a decent indication of the goodness of fit, i.e., the proportion of data points that lie within the regression equation, giving a measure of how well the samples are predicted by the model.

In that sense, we are more interested in the inverse function of the curve fitting model as it is crucial for obtaining a simulated value of recovery, x, from equation 2.11, given a random number between 0 and 1, as illustrated in figure 3.9. In that regard, table 4.6 and figure 4.14 show, respectively, the mean of the descriptive statistics for each week of simulated results and the histogram of the various simulations for the next three months (400 simulations for each week).

From that information it is fair to claim that the original data and the new simulated values come from the same probability distribution as the histogram of each week and its descriptive statistics are in the most part similar, except for the gap in the range [0.6,0.75] in figure 4.9 that became obsolete as it is not recognized by the CDF equation and the minimum value that varies by approximately 11 percentage points, possibly justified by the fact that the rounding of the CDF equation parameters does not capture very well its lower bound.

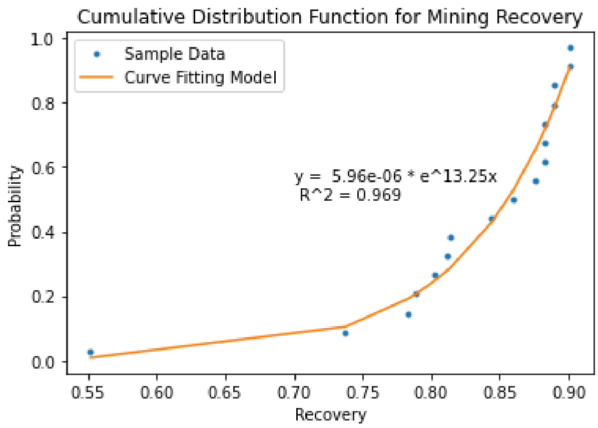

Figure 4.13 - CDF for mining recovery and respective equation

Table 4.6 – Mean of the descriptive statistics for each week of simulated mining recovery

Simulated Realizations	Mean	SD	Min	25%	50%	75%	Max
4800	0.833	0.074	0.442	0.805	0.856	0.886	0.908

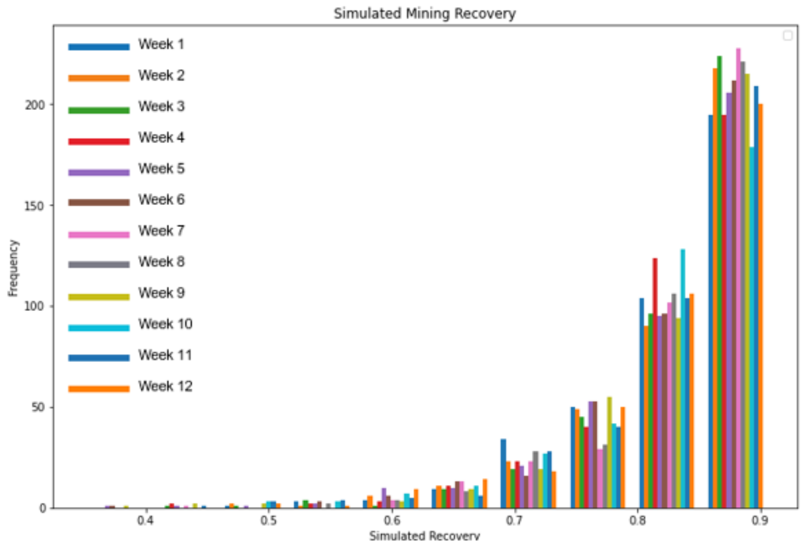

Figure 4.14 - Histogram of Monte Carlo simulations for the mining recovery

4.3. Zinc Head Grade

In this section are presented and discussed the results of the data analysis, distribution fitting using graphical methods, as well as the Monte Carlo simulations for the zinc head grade.

Univariate Analysis

Performing a univariate analysis similarly to the mining recovery data, the descriptive statistics of the head grades in table 4.7 and its visualization, displayed in figure 4.15 in the form of a histogram, were obtained. From this information it is important to note that the values of the mean and the median (quantile 50%) are identical, suggesting that the data comes from a normal symmetric distribution. Also, as no anomalies in the data were found, it can be inferred that the cut-off grade is less than or equal to the minimum value, namely around 5.6%.

Table 4.7 – Descriptive Statistics of the zinc head grade

Observations	Mean	SD	Min	25%	50%	75%	Max
49	7.600	0.795	5.598	7.100	7.658	7.913	9.488

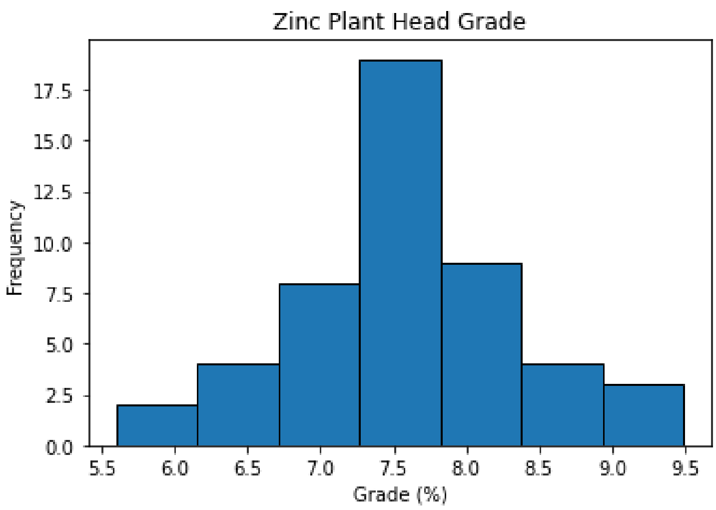

Figure 4.15 - Histogram of the zinc head grade

Graphical Methods

To examine the truth of our assumption that the data comes from a normal distribution, the KS test attributed a test statistic of 0.1349 when testing for the normal distribution. Comparing with the critical value of approximately 0.174 = $\frac{1.22}{\sqrt{49}}$ from the stretch of a KS test table in figure 4.16, as the test statistic is less than the critical value, we can accept with 95% confidence (α=0.05) that our data comes from a normal distribution.

From the observation of figure 4.17 which shows the probability plot with the best fit line for the data when comparing with a normal distribution with parameters mean and standard deviation, as in table 4.7, we can conclude that our assumption is valid, as this line is close to a 45º line intercepting the lower left part of the graph through to the upper right side with coefficient of determination equal to 0.984.

One-Sided Test:					
$\alpha =$	.10	.05	.025	.01	.005
Approximation for $n > 40$	$\frac{1.07}{\sqrt{n}}$	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.52}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$

Figure 4.16 - Critical Values for the KS test for n > 40; Source: Miller (1956)

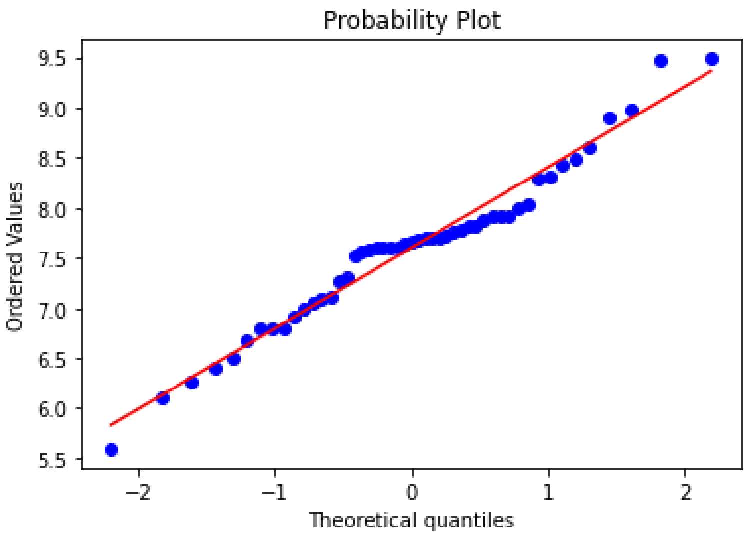

Figure 4.17 - Probability Plot using the normal distribution

Once the Monte Carlo simulations were performed, they presented the results in table 4.8 and figure 4.18. Analysing the table, it is notable that the mean of the summary statistics for each week is in all similar to results given by original data in figure 4.17. From the figure 4.18, the same symmetrical characteristics as for the histogram of the original data in figure 4.15 are noticeable. Moreover, the simulations achieved values that surpassed the minimum and maximum boundaries of the original data but that should not be considered as an issue as these values are still in a range in which this error can be attributed to fluctuation of grades coming from the underground mine, as this is a recurring pattern when we only have an estimation of the orebody grade distribution.

Table 4.8 - Mean of the descriptive statistics for each week of simulated head grades

Simulated Realizations	Mean	SD	Min	25%	50%	75%	Max
4800	7.598	0.795	5.155	7.067	7.604	8.122	9.867

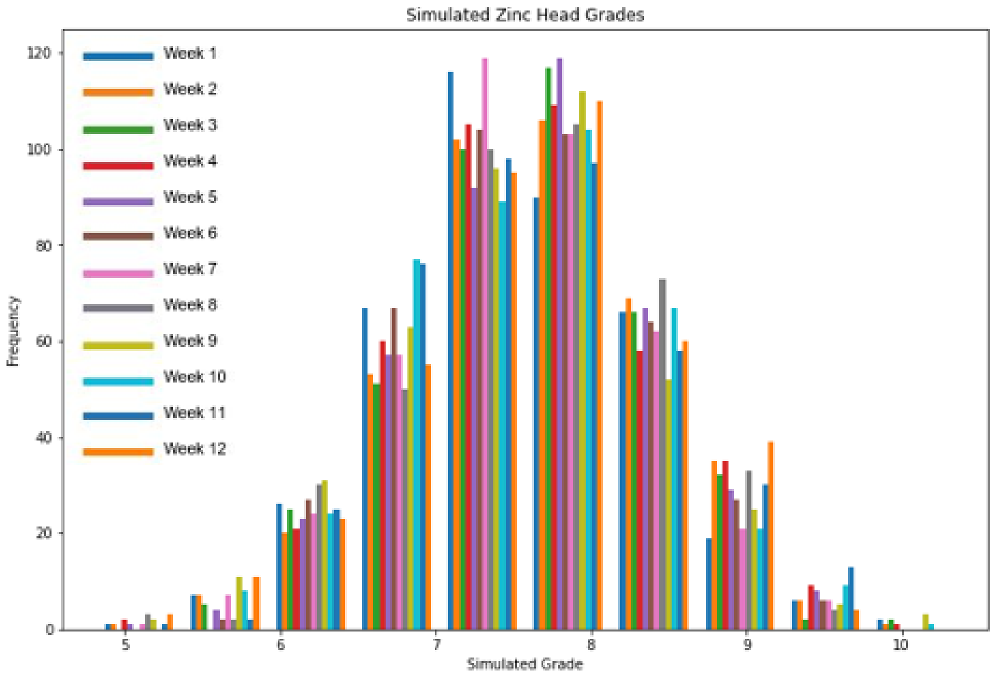

Figure 4.18 - Histogram of Monte Carlo Simulations for zinc head grade

4.4. Profit Risk Assessment

For the calculation of the gross revenue using equation 3.9 it was used the zinc price forecast using the Holt's method with parameters alpha = 0.6 and beta = 0.7 for the months of March (Week 1-4), April (Week 5-8) and May (Week 9 -12) being 2802.713 USD, 2835.642 USD and 2868.572 USD, respectively. Then, for the calculation of the quantity of metal in equation 3.10, processing plant recovery was kept constant at 80% and the values of mining recovery and zinc head grade are each of the simulated for the next three months, represented by the histograms in figures 4.14 and 4.18, respectively. Finally, the value of tonnage of mine is the one that produces one tone to be milled. Using the ratio between net revenue in equation 3.8 and gross revenue in equation 3.9 gives the NSR variation in percentage. Figure 4.19 illustrates these results, for the next three months, in a scenario where it is expected that the zinc metal prices continue to soar, with prices around 2800 USD as figure 4.7 shows.

One could expect a slight increase in the NSR percentage over time due to the increase in prices but that behaviour is not perceptible due to the fact that the forecasting model does not predict values too far off the last observed value in February 2021. Also, it is important to note that the mining recovery and the zinc head grade on which the NSR calculations are highly dependent, do not show a clear ascending or descending pattern overtime, offsetting the expected growth in NSR produced by increased zinc prices.

It can be observed that the values of mine netback are in range of around 93% to 98%, due to the lack of precise data in transportation, freight, insurance and marketing costs that make up the commercial costs. The rule of thumb when dealing with a zinc concentrate NSR is to establish the mine netback in values around 50% as noted by Wellmer et.al (2008).

Keeping the value of mine netback to 50%, calculations of possible cut-off grades for the next three months were achieved using equation 3.12 as figure 4.20 shows. From that, no increasing or decreasing pattern is found over time.

It is notable that a high density of values is slightly less than 6% of grade. Comparing the values of the head grades data and simulated head grades, this is an inferior value as the mean for the original and simulated values is around 7.6%. This factor can be attributed to the fact that no dilution was considered and also by the lack of accuracy when accounting for total operating costs.

The descendent shape of the histogram when looking through higher values of cut-off grade gives a clue that mining recovery plays a major role in its influence as these parameters follow the opposite pattern: the greater distribution of the mining recovery is for the values above 80% recovery and the high values of cut-off grade are possibly due to the calculations that the algorithm achieved using lower values of mining recovery.

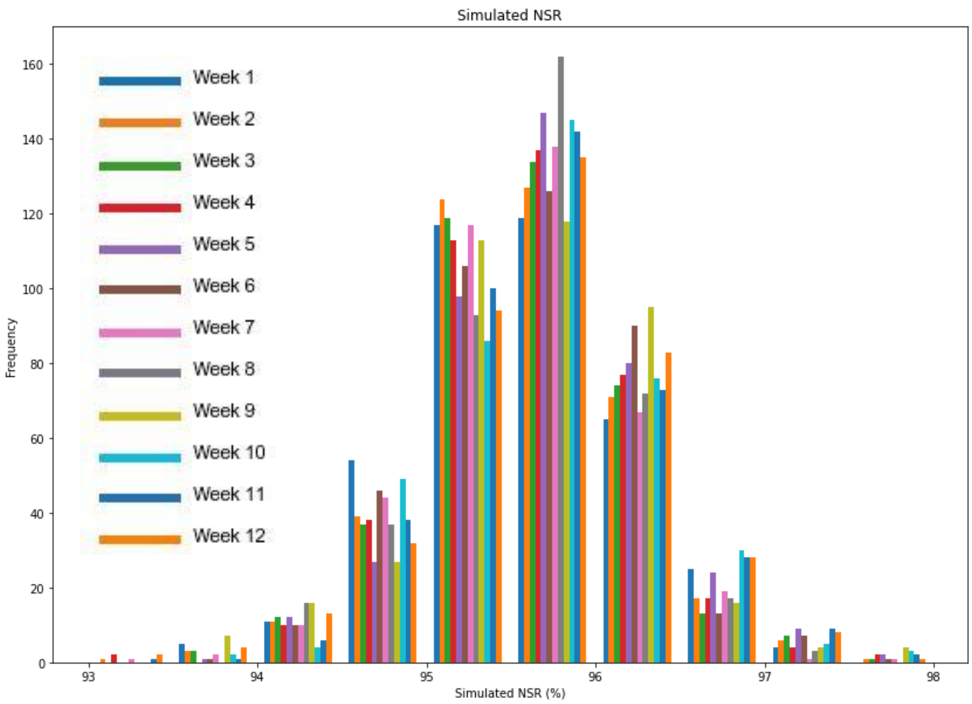

Figure 4.19 - Simulated NSR in percentage

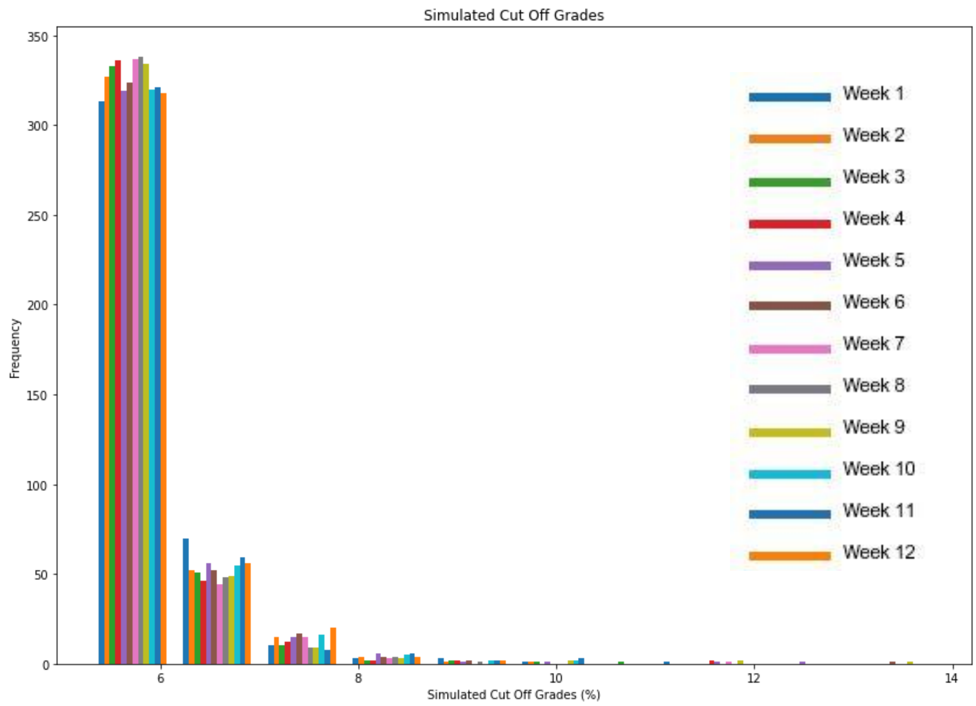

Figure 4.20 - Simulated Cut-Off Grades

Taking into account the information above it is essential that the uncertainty in the NSR calculations and cut-off grades are integrated in the creation of mine plans to better access the risk in the decision-making process, by incorporating ways to reduce this risk and maximizing mining profits.

Annex A contains the operational margin calculated with equation 3.11 where the range in which the mine would profit for the quantity of metal milled given grades, mining and processing recoveries, is observable, if prices forecasted and costs obtained would apply.

Chapter 5 – Conclusions

This final chapter presents the summary of the objectives proposed for this work and how they were accomplished and presented throughout the dissertation. Visualizing a future enterprise framework for the issues discussed in this thesis, some advices and recommendations are also provided.

5.1. Final Remarks

The first major accomplishment of this work was to present a methodology that quantifies the financial risk of a mine planning decision, represented by the probability distribution of the NSR, which was the chosen parameter to quantify the performance of the exploitation of determined area of an orebody. This parameter is influenced by innumerous factors from which the metal price, reserves estimates associated with the zinc head grade and mining recovery were worthy of a special attention.

Upgrading from the assumption that metal prices are deterministic, future metal prices were forecasted in a short-term horizon (3 months) using a dataset of the monthly prices, representative of the global market from the last 15 years. In conjunction with others, the Holt's method achieved the best predictions, coinciding with the underlying trend for the most recent observations.

Afterwards, a calculation for the probability distribution of the mining recovery dataset was achieved. Performing the KS Test and using graphical methods to compare the data distribution with various know distributions, satisfactory results were accomplished. Furthermore, fitting the data distribution with an exponential equation and performing Monte Carlo simulations using the inverse transform method, resulted in a better distribution function that quantifies the mining recovery uncertainty.

Regarding the reserves estimates it is usually assumed that we have perfect knowledge of the orebody but using a stochastic framework adds value to the production schedule. As no model was available, samples of the zinc head grade were obtained and the probability distribution of the data was assertively assumed to have a normal distribution. Recurring to Monte Carlo simulations, future grade estimates were generated for the short-term horizon.

Using the forecasted and simulated results, two models were created: one model of the percentage NSR (mine netback) and another for the cut-off grades that can be quantified in a risk throughout the timeframe of the mine plan. Values of zinc cut-off grade in the range between 5% to 6% are the most expected using 50% NSR and not accounting for dilution.

5.2. Future Work

Another objective of the present work was to develop meaningful ways of interconnecting geostatistics (ore evaluation) with financial analysis.

In accordance with quantification of geological risk using geostatistics and figure 3.1 in mind, the methodology here presented should be incorporated in a corporate environment envisioning a mine plan and production scheduling that takes into account not only one orebody model but multiple probable models and for each model perform an extraction schedule, progressing with the concept presented by Neves et al. (2020). In this approach, the dilution parameter should be integrated as it is a major part of the blasting operation. In that regard, it could also be of great value applying this idea into a panel, obtaining the uncertainty of each block of the panel. Using selective mining strategies, mining recoveries could be diversified using another mining methods in use and choosing the one that obtains the best benefit.

Such as the geology, metal prices play major influence in mine revenues and a simple approach for forecasting was used in the scope of this work. So, in a corporate level it is important to complement the results of quantitative methods with surveys and executives experienced input for forecasting as the forecasts methods here presented are dependent of factors that cannot be predicted. In line with this thought, metal prices can be examined for parameters such as seasonality that can be better addressed using the Holt-Winters method. Also, the gaining popularity of neural networks and its ability to solve complex problems can be of great help in the future.

Finally, the gap between geostatistics and financial analysis can be filled with more research and development in both areas, bringing advantages to a mine in a strategic, tactical and operational fronts as in a longer-term horizon it can influence contractual agreements with smelters, provide higher value to the mine planners in the short-medium term horizon and, last, but not least, efficiently allocate and use resources on a daily basis.

References

Akima H., A New Method of Interpolation and Smooth Curve Fitting Based on Local Procedures. Journal of the Association for Computing Machinery, Vol 17, Nº 4, October 1970, pp. 589-602.

Albanese T, McGagh J, Future Trends in Mining. In SME Mining Engineering Handbook. 3rd Edition. Society for Mining, Metallurgy and Exploration.

Annels A.E., (1991). Mineral Deposit Evaluation – A practical approach. 1st Edition, Chapman & Hall.

Armstrong M. (1994) Is Research in Mining Geostats as Dead as a Dodo? In: Dimitrakopoulos R. (eds) Geostatistics for the Next Century. Quantitative Geology and Geostatistics, vol 6. Springer, Dordrecht. https://doi.org/10.1007/978-94-011-0824-9_34.

Bargmann C.J., (2000). Agua Rica prospect, Catamarca Province, Argentina: an example of deposit evaluation based on net smelter return. The Institution of Mining and Metallurgy 109, May-August 2000.

Daigo R, Osako, S, Adachi Y, Matsuno T (2013), Time-series analysis of global zinc demand associated with steel. Resources, Conservation and Recycling 82 (2014) 35-40. http://dx.doi.org/10.1016/j.resconrec.2013.10.013.

Darling P., (2011), SME Mining Engineering Handbook. 3rd Edition. Society for Mining, Metallurgy, and Exploration.

Diggle P.J., (2013), Statistical Analysis of Spatial and Spatio-Temporal Point Patterns. 3rd Edition, CRC Press.

Dimitrakopoulos R, Farrelly C.T, Godoy M (2002), Moving forward from traditional optimization: grade uncertainty and risk effects in open-pit design. Mining Technology, 111:1, 82-88. https://doi.org/10.1179/mnt.2002.111.1.82.

Dimitrakopoulos R., Rimélé A., Gamache M., (2019). A dynamic stochastic programming approach for open-pit mine planning with geological and commodity price uncertainty. Resources Policy 65 (2020) 101570, https://doi.org/10.1016/j.resourpol.2019.101570.

Dooley G, Lenihan H (2005) An assessment of time series methods in metal price forecasting. Resources Policy 30: 208–217. https://doi.org/10.1016/j.resourpol.2005.08.007.

Eggert, R.G., (1987), Metallic mineral exploration: an economic analysis. Resources for the Future, Washington, DC.

Foss S.D., A Method of Exponential Curve Fitting by Numerical Integration. Biometrics, Vol. 26, Nº 4 (Dec., 1970), pp. 815-821. https://doi.org/10.2307/2528726.

Gocht W.R., Zantop H., Eggert. E.G, (1988), International Mineral Economics. Mineral Exploration, Mine Valuation, Mineral Markets, International Mineral Policies. Springer-Verlag Berlin.

Goldie R. and Tredger P. (1991). Net Smelter Return Models and Their Use in the Exploration, Evaluation and Exploitation of Polymetallic Deposits. Geoscience Canada, 18(4).

Greene W.H., (2003), Econometric Analysis. 5th Edition, Pearson Education, Inc.

Hassani H. and Silva E.S., (2015). A Kolmogorov-Smirnov Based Test for Comparing the Predictive Accuracy of Two Sets of Forecasts. Econometrics 2015, 3, 590-609; doi:10.3390/econometrics3030590.

Hilson G, Murck B (2000), Sustainable development in the mining industry: clarifying the corporate perspective. Resources Policy 26 (2000) 227-238. https://doi.org/10.1016/S0301-4207(00)00041-6.

Hustrulid W., Kuchta M., Martin R., (2013). Open Pit Mine Planning & Design. Volume 1 – Fundamentals. 3rd Edition, CRC Press Taylor & Francis Group.

International Monetary Fund, Global price of Zinc [PZINCUSDM], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/PZINCUSDM, April 1, 2021.

Kear R.M (2006) Strategic and tactical mine planning components. The Journal of the South African Institute of Mining and Metallurgy, vol 106, 93-96.

Leslie H. M., (1956). Table of Percentage Points of Kolmogorov Statistics, Journal of the American Statistical Association, 51:273, 111-121, DOI: 10.1080/01621459.1956.10501314.

Makridakis, S.G., Parzen E., Fildes R., (1984), The Forecasting Accuracy of major time series methods, 1st Edition, Wiley.

Makridakis, S.G., Wheelwright, S.C., Hyndman, R.J., (1998), Forecasting: Methods and Applications, 3rd Edition.

Maybee B., Yana J. (2017). Using the single index model to create short-term mine plan. Underground Mining Technology. First International Conference on Underground Mining Technology.

McCarthy P.L. (2001), Mining Dilution and Losses in Proceedings Underground Mining, in Mineral Resource and Ore Reserve Estimation. The Aus IMM Guide to Good Practice (Ed: A C Edwards), pp333-336 (The Australasian Institute of Mining and Metallurgy: Melbourne).

Mohammad A, Hashem S, Reza M (2013) Monte Carlo Analytic Hierarchy Process (MAHP) approach to selection of optimum mining method. International Journal of Mining Science and Technology 23 (2013) 573–578. http://dx.doi.org/10.1016/j.ijmst.2013.07.017.

Monkhouse P.H.L. and Yeates G.A., Beyond Naïve Optimisation. The Australasian Institute of Mining and Metallurgy 2018 R.Dimitrakopoulos (ed.), Advances in Applied Mine Planning, https://doi.org/10.1007/978-3-319-69320-0_1.

Neves J, Araújo C, Soares A (2020) Uncertainty in Dynamic Mining Reserves. Mathematical Geosciences. https://doi.org/10.1007/s11004-020-09866-1.

Rendu J.M., (2002) Geostatistical Simulations for Risk Assessment and Decision Making: The Mining Industry Perspective. International Journal of Surface Mining, Reclamation and Environment, 16:2, 122-133. https://doi.org/10.1076/ijsm.16.2.122.3399.

Rendu J.M., (2014). An Introduction to Cut-Off Grade Estimation. 2nd Edition, Society for Mining, Metallurgy and Exploration.

Sharif O, Hasan, MZ (2019) Forecasting the stock price by using Holt's method. Indonesian Journal of Contemporary Management Research, 1(1), 15-24. https://doi.org/10.33455/ijcmr.v1i1.83.

Thomopoulos N.T., (2013), Essentials of Monte Carlo Simulation. Statistical Methods for Building Simulation Models. Springer.

Wellmer F-W., Dalheimer M., Wagner M., (2008). Economic Evaluations in Exploration. 2nd Edition, Springer.

Wilk M.B. and Gnanadesikan R., Probability Plotting Methods for the Analysis of Data. Biometrika, Vol 55, Nº1 (Mar., 1968), pp. 1-17, https://doi.org/10.2307/2334448.

Wills B.A., and Finch J.A., (2016). Will's Mineral Processing Technology – An Introduction to the Practical Aspects of Ore Treatment and Mineral Recovery. 8th Edition, Elsevier.

https://www.ilzsg.org/static/statistics.aspx?from=2 – **accessed on 25.03.2021**

https://www.spglobal.com/platts/en/market-insights/latest-news/metals/031021-global-zinc-demand-to-outpace-production-growth-in-2021-narrowing-surplus-ugmk#article0 – **accessed on 22/04/2021**

Annexes

Annex A – Operating Margins for each week in USD/ ton milled.