STOCHASTIC CHARACTERISATION OF A MINING PRODUCTION SYSTEM

Nancy Selemagae Magagula

A dissertation submitted to the Faculty of Engineering and the Built Environment, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science in Engineering.

Johannesburg, 2016

DECLARATION

I declare that this dissertation is my own unaided work. It is being submitted to the degree of Master of Science in Engineering to the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination to any other University.

· ·

(Signature of candidate)

Signed on the · · · day of · year · · · · · · · · ·

at ·

ABSTRACT

There are discrepancies between deterministic mine planning and the actual mining process due to geological uncertainties associated with mineral deposits and inherent production system variabilities. The misalignment between the planning process and the actual mine production process often leads to non-achievement of production outcomes. Stochastic mine planning has been developed to minimise these misalignments but it is computationally intense and requires constraint functions to operate effectively. However, the stochastic mine planning approaches in literature do not have an embedded process analysing the interactions between the Key Performance Indicators (KPIs) and the mine production activities.

This dissertation proposes an approach to study the interactions/correlations between KPIs used to measure the progress of a mining operation and the mining activities. The Multinomial Logistic Regression (MLR) approach is a non-linear and non-normal measurement method which can assist in understanding the behaviour of mine production activities when compared to assessed KPIs. The MLR model can also assist in establishing which production activities require maximisation or minimisation in attaining the desired KPIs.

This study shows that 71% of the KPIs for a case study in mining production system are influenced by the movements of the production activities in the mining process and the level of uncertainty on the forecasted KPIs is reduced through applying the MLR model. This method will help mining companies in assessing in the initial stages of mine planning the mine production activities that management should focus on to achieve desired KPIs by directing more effort and resources to these statistically significant activities.

DEDICATION

To Sibusiso Vusi and Busisiwe Phola Magagula & Mashane Rosina Mampholo

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor Prof Cuthbert Musingwini and co-supervisor Prof Montaz M. Ali for their guidance and supports. Without their extensive comments, corrections and suggestions this dissertation would not be possible.

I would like to acknowledge Mr Tinashe Tholana for reviewing my work and Mr Tshele Sebutsoe from Impala Platinum mine for explaining some of the mining concepts.

Last but not least, I would like to thank my husband, Sibusiso and my daughter, Busisiwe for their constant encouragement, support and understanding.

Contents

List of Figures

List of Tables

LIST OF SYMBOLS

Alternative hypothesis $\cdots$ H_1

Correlation $\cdots$ ρ

For all $\cdots$ $\forall$

"*is an element*" $\cdots$ ϵ

Kendall tau $\cdots$ τ_k

Null hypothesis $\cdots$ H_0

Planned targets $\cdots$ α

Planned versus actual ratio $\cdots$ γ

standard deviation of x $\cdots$ σ_x

standard deviation of y $\cdots$ σ_y

LIST OF ACRONYMS

AIC - Akaike Information Criterion

BIC - Bayesian Information Criterion

EVT - Extreme Value Theory

GDP - Gross Domestic Product

GL - Generalised Linear

GLM - Generalised Linear Models

KPI - Key Performance Indicator

LR - Logistic Regression

LSM - Least Square Method

MLE - Maximum Likelihood Estimation

MLR - Multinomial Logistic Regression

NPV - Net Present Value

PDF - Probability Density Function

SAS - Statistical Analysis Software

Chapter 1

INTRODUCTION

1.1 Chapter overview

This chapter introduces the mine planning concept, its limitations and a possible solution which is discussed throughout this dissertation. Section 1.2 outlines the background information to the problem researched. Section 1.3 discusses the problem in depth. Section 1.4 provides a brief explanation on the process planned by the mine planning method. Section 1.5 adds on the information provided in Section 1.3 by outlining further the importance of this dissertation in resolving existing limitations stated in Section 1.2. Section 1.6 discusses the objective of the study and Section 1.7 defines all definitions used in the dissertation. Finally, Section 1.8 concludes the chapter by providing, the structure of the dissertation.

1.2 Background

Mine planning plays a critical role in activities preceding the successful operation of a mine production system. Inputs into the planning process have traditionally been average values that result in a deterministic mine production plan. This is in contrast to the actual production process where temporal values of both inputs and outputs are actually stochastic in nature. This dichotomy between plans and actual activities has often led to non-achievement of production outcomes. Consequently, there has been gradual developments within the minerals industry where mining companies have to face litigations or lawsuits from investors for non-delivery of promised outcomes. For example, in recent times financiers of mining projects have started resorting to litigation against the project proponents, claiming that they were misled into investing in a project that would neither be completed on time nor within budget. A case in point is the shareholder class action lawsuit led against NovaGold over the Galore Creek copper-gold project in which costs were revised to 127% greater than the initial estimates and the project was two and half years behind schedule (Mineweb (2008a), Mineweb (2008b), Mineweb (2009)).

Kwok and Roantree (2014) stated that Hong Kong's securities regulator was suing the Chinese conglomerate CITIC Ltd. over misconduct linked to the $2bn FX loss suffered in 2008. This loss was suffered due to an investment in a troublesome Australian iron mine which did not meet its planned target and failed to provide returns. Kosich (2013), on the other hand suggested that after the recent 2009 financial crisis, investors have been pulling away from risky investments. Hence junior mining companies are fighting for survival since capital funding of mining projects is limited. These instances cited by Kosich (2013) and Kwok and Roantree (2014) illustrate the severity under which the mining industry is operating, as well as the consequences of failing to meet planned targets.

In order to respond to the challenges described above, there have been attempts to develop ways in creating more robust mine plans that recognise the mining process as a stochastic process. This is an emerging discipline broadly referred to as stochastic mine planning. Examples of such studies include Mustafa (2010), Dimitrakopoulos (2011), Newby et al. (2012) and Ramazan and Dimitrakopoulos (2013).

1.3 Problem statement

Literature suggests that there has been a developing paradigm shift from deterministic mine planning towards stochastic mine planning. This research study explores how interdependent mining activities that are characteristically stochastic in nature interact to result in the production Key Performance Indicators (KPIs). Such KPIs include tonnage and grade of milled ore and the resultant mineral product produced. The key question is *"how do stochastic activities in a mining production system interact to produce the resultant KPIs?"* In order to answer this question, realistic hypothetical probability distributions for mining activities are assumed for a mining production system to establish questions such as "if the distribution for a particular activity is modified with a view to reduce uncertainty, how does this affect the overall production system?". For example, "how would a 10% tightening in the distribution of a production activity within a production system impact on the overall production system?"

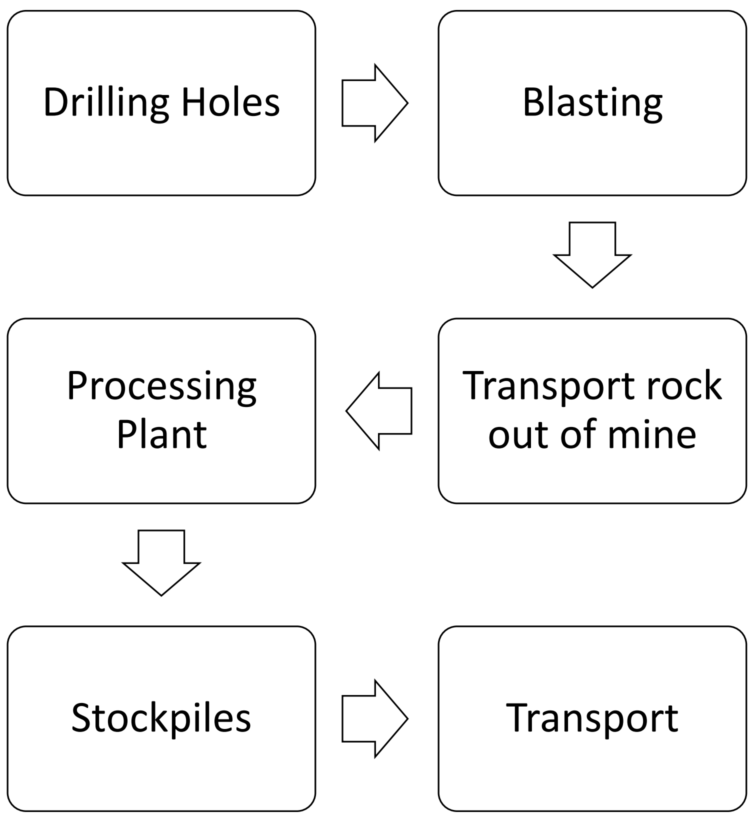
1.4 Description and analysis of a typical mine production system

The mine production process explained and outlined in Figure 1.1 represents a typical iron ore mining process. An assumption is made that all mining processes used in extracting the commodity are typically the same excluding the refining process. Hence this study will assume that the process outlined below is a generic mining process and using Figure 1.1 it is broken down into four broad steps (drill, blast, load, haul, process):

- This is an open pit mining process which implies that the first step involves drilling and blasting holes in the ground as outlined in Figure 1.1;
- Subsequent to the initial step is blasting which disintegrates the intact rock into fragmented rock or ore that can be easily loaded onto transport units;
- The fragmented rocks are then transported out of the mine to a processing plant; and
- In the processing plant the mineral product is recovered and waste fines are then dumped onto a slimes dam.

All of these processes defined above inform the mine production activities and various data from these steps is then collected. The collected data is then used in an algorithm to

Figure 1.1: An example of a typical mine production system



optimise the process further. This section aids the conceptualisation of mine production activities and how they interact to produce the KPIs used in monitoring the progress of mining. The first step is the exploration process where geologists explore the area to be mined in relation to the ore available. The second step involves mine design, where mine planning is conducted and other related issues such as environmental studies are done. Subsequent to this step, "mine construction" is conducted which includes building of the infrastructure that will facilitate easier extraction and production. Thereafter, is the mine production process, as discussed in the bullet points above. Lastly, when mining stops permanently final rehabilitation of the environment is done.

1.5 Significance of the dissertation

McCulloch (2009) studied the impact of litigations to the mining industry because of work-related diseases. McCulloch (2009) stated that over 250 000 jobs were lost due to regulatory reforms on disease compensations and due to lawsuits induced onto the gold mining industry. An important element to take from McCulloch (2009) is that mining litigations can wipe out attained returns if they are not prevented. Indeed, as a result of litigations, four major mining players were left operating namely; AngloGold, AvGold, Goldfields and RandGold (McCulloch, 2009). Similar inferences, if investigated, can be drawn for other mining operations other than gold mining as discussed by McCulloch (2009).

Mining is very sensitive to costs, therefore minimisation of negative influences can improve its sustainability in the South African economy. Another example is the recent wage-strike in the mining sector in 2014. Mining companies such as Amplats and Lonmin threatened to cut jobs. This is due to the losses incurred from the suspended mine production activities caused by the strike. The mining industry is severely affected by operational and regulatory risks due to the mining operations. Therefore, any form of understanding relating to inputs feeding into the mine plan and outputs set as targets is beneficial in reducing some of the inefficiencies currently existing in the mining process.

Studying the interactions between the mine production activities and resultant KPIs from a stochastic mine planning perspective might provide insightful information. This can improve on over-runs and minimise litigations due to such over-runs in an overall mining process. Constant improvements on the mine planning processes are important. This will ensure that all preliminary budgeting figures and mining targets governing the mine production activities are close to the actual realizations yield by the mining process. Therefore, a study looking into the correlation with an aim of improving on inefficiencies in mining operations is useful in guaranteeing sustainable mining operations. Lastly, stochastic mine planning in the current literature does not account for these interactions as discussed in Section 1.2.

1.6 Objective of the dissertation

The objective of this research is to:

- establish interactions between interdependent production activities in a hypothetical mining production system considering the stochastic nature of the mining production activities. Such understanding of stochastic interactions of variables will improve the understanding of stochastic mine planning.

1.7 Important terms used throughout the dissertation

This section outlines and defines some of the key terms used in this research, although some of the definitions are defined in the dissertation as it progresses. Some definitions are explained below:

Definition 1 *Actual targets are mined mineral tonnage or ore grade produced in a mining process.*

Definition 2 *Constraint function(s) are secondary functions in an optimization model which are used in solving the optimization problem by narrowing the feasible region of the solution set.*

Definition 3 *Correlation is the degree of association between two or more random variables. For example, if random variables X and Y have a correlation of -0.70 then this implies that an increase in X will result in a 70% decrease in Y.*

Definition 4 *A correlation matrix is a mathematical representation of a correlation structure between two or more random variables.*

Definition 5 *Dependent variables are mining production activities measurements in the mining process. For example, ore grade collected in mine production is a dependent variable in mining.*

Definition 6 *Deterministic mine planning is a premeditated course of action based on average values collected from historical mine production figures, adjusted as necessary.*

Definition 7 *A finite method is a method of finding a solution to a particular problem where problem boundaries and assumptions are predefined, and the method terminates in a finite number of iterations.*

Definition 8 *Independent variables are variables which are manipulated in order to determine the corresponding values of dependent variables in a model setting.*

Definition 9 *Key Performance Indicators (KPIs) are output variables of quantities used in measuring the progress of a mining process.*

Definition 10 *A mine plan is a complete schedule that is used by a mine to ensure that it is within the limit in terms of project timelines, budget and financial objectives of the mining project. A mine plan is the end-product of a feasible mine planning and effective scheduling processes.*

Definition 11 *Mine planning, according to Bhattacharya (2007), is a premeditated course of action of extracting natural minerals resources optimally.*

Definition 12 *Mine production activities are courses of actions driving the mining process.*

Definition 13 *Mining is the extraction of mineral resources and the refinement of these minerals to sell in the market for positive economic returns.*

Definition 14 *Multinomial Logistic Regression (MLR) is a mathematical model used to analyse categorical independent variables with more than two levels.*

Definition 15 *Open pit mining is the extraction of mineral resources through drilling starting from the surface.*

Definition 16 *Optimisation is an iterative mathematical process used to determine the "best" or optimum solution.*

Definition 17 *A planned target is an estimated pre-planned target of mineable mineral resources in a particular mine production process.*

Definition 18 *Stochastic mine planning is a premeditated course of action based on the data collected from historical mine production figures and as well incorporating uncertainty into the modelling processes.*

Definition 19 *Underground mining is when mineral resources are extracted from the earth by excavating openings or shafts leading below ground level. These shafts are used to transport miners to the location of the mineral resources where the extraction takes place.*

1.8 Structure to the dissertation

In addition to this chapter there are five more chapters to give a total of six chapters. Chapter 1 (Introduction) provides background information to the problem statement, articulates the problem statement and research question, and provides an outline of the dissertation. Chapter 2 (Literature review) provides an analysis of literature on the subject matter, highlighting why the research study is important. Chapter 3 (Correlation methods) discusses correlation methods in general. Chapter 4 (MLR methodology) details the methodology followed and provides the hypothetical data used in the study. Chapter 5 (Hypothetical platinum mine case study and model analysis) provides detailed empirical analysis based on the problem presented in Chapter 2 as well as outlining briefly the mining production scenario described by the data investigated in this dissertation. Conclusions and recommendations are presented in Chapter 6 (Conclusions and Recommendations). The dissertation ends by referencing all sources of information used in the study and appendices supporting the work in the body of the dissertation.

Chapter 2

LITERATURE REVIEW

2.1 Chapter overview

This chapter presents a literature survey of all discussions drawn from the studies about the research question highlighted in Section 1.3. Section 2.2 outlines the historical origin of mining in South Africa. Section 2.3 defines challenges facing the mining sector. Section 2.4 outlines the differences between deterministic and stochastic mine planning approaches. Section 2.5 provides a discussion on research conducted on stochastic mine planning. Section 2.6 discusses the limitations on stochastic mine planning and motivates the need for this research. Section 2.7 concludes the chapter's intent.

2.2 Mining overview

Modern mining operations started in the 1850s in South Africa, according to Tempelhoff et al. (2014). Ever since, South Africa has been a major producer of precious minerals such as diamonds, gold and platinum as well as strategic minerals, such as iron ore, coal and uranium. PROJECTS IQ (2015) stated that, the discovery of the first diamond on the banks of the Orange River in 1867 sparked mining exploration activities in South Africa. The exploration activities in South Africa catalysed the discovery of the world largest gold deposit in 1886 in the Transvaal area.

Mining has driven and stimulated economic growth in South Africa. McCulloch (2009) stated that mining in general contributed to the economic growth and development observed in the country. In the 1970s, gold mining in South Africa contributed about 68% to the global gold production (PROJECTS IQ, 2015). According to PROJECTS IQ (2015) currently mining in South Africa contributes:

- On average 20% to the Gross Domestic Product (GDP) directly and indirectly;
- R330 billion to the total annual income in 2014; and
- Majority of employment to the country's citizens and non-citizens with more than one million jobs being mining-related.

Its contributions have been vast but in recent times, this *"golden goose"* has been encountering countless challenges threatening its development and survival. Therefore, robust mine planning is important, especially during this difficult period. The next section will discuss some of the challenges faced by mines in South Africa.

2.3 Major mining events threatening its existence

The modern mining practice as known today has been in existence for about 164 years in South Africa, according to Tempelhoff et al. (2014). As stated in Section 2.2, the discovery of mining came with consequences as well, for example, the Anglo-Boer War. However, in the past these cons were overshadowed by the pros the mining industry brought into the South African economy. That being said, in recent times the mining industry is sometimes challenged by lawsuits due to poor mine planning, negative environmental impacts and numerous mining strikes.

According to SAHO, S.A.H.O (2014), the platinum strike during 2014 was the longest wage strike in South Africa which stretched from the 23^{rd} of January 2014 to 23^{rd} of June 2014. More than $70,000$ platinum mine workers were requesting a minimum wage of $R12,500.00$

per month. Although their intentions of the strike were legitimate, it was costly to the South African mining industry and the economy. Findings by Singh (2014) suggested that over R16-billion was lost which resulted in a GDP of less than a percent in the 4^{th} Quarter (Q4) of 2014. Mining strikes have negative impacts against mine production and KPIs. Such occurrences are among major factors motivating this research. It is important to note which of the stochastic mining activities are influential on the measured KPI and what is the correlation between these mining activities?

Mining is not affected by strikes only; Teckcominco was sued by indigenous people in Alaska for mining without rights (Nelsen, 2007). Lawsuits are among other external factors influencing the mining processes and affecting the final KPIs. In recent times mining is also affected by market fluctuations. Nattrass (1995) stated that, mining has long been the backbone of South Africa's economy. However due to rising costs, falling ore grades and a depressed gold price[1], the mining industry is struggling. According to Nattrass (1995) gold price started falling from 1987 onwards which resulted in retrenchments. Other factors stated by Nattrass (1995) which significantly affected the mining industry are:

- high effective rate of taxation;
- devaluation of the rand in the 1980s onwards; and
- wage demands.

These aforementioned challenges amplify the necessity of conducting robust mine planning and understanding which mine production activities are statistically significant in the attainment of the desired KPIs.

2.4 Deterministic versus stochastic mine planning

Hajdasinski (1988) stated that optimization of the mine size and life of mine have been studied thoroughly in the 1950s and 1960s, but with the advancements of computers, these

[1]However in recent times the gold price has significantly improved when compared to the time stated by Nattrass (1995).

methodologies need reviewing. Optimization techniques are the core fundamentals of mine planning. Numerical techniques use inputs into the mathematical model used by the optimization algorithm to derive the best results as outputs. Optimization is used in mine planning to obtain optimal mine production outcomes within a planning horizon. There are two approaches of conducting mine planning namely; deterministic and stochastic approaches. Deterministic mine planning use historical data collected from previous mine production activities and assumes that this trend will continue in future processes. For example, suppose the statistics in Table 2.1 are from a gold mine,

Table 2.1: Hypothetical gold mine statistics

Month	**Blasted Tonnes**	**Ore Tonnes**	**Gold price($/oz)**
Jan-12	752	552	1,689.52
Feb-12	963	845	1,789.22
Mar-12	845	742	1,546.63
Apr-12	600	600	1,326.45
May-12	851	851	948.78
Jun-12	745	645	645.23
Jul-12	956	742	525.12
Aug-12	852	549	1,478.68
Sep-12	450	352	1,569.75
Average	779	653	1,279.93

In a deterministic mine planning approach the average values are used to compile and forecast production and price for October 2012. There are few variations that are ignored by this approach, hence stochastic mine planning is promoted because it incorporates the variation noted in Table 2.1. Table 2.1 also shows that mine production activities are stochastic in nature and not static as assumed by the deterministic mine planning approach.

The stochastic mine planning approach was mainly designed at minimising all shortcomings relating to the deterministic mine planning process. Stochastic mine planning is a complex scheduling process due to its ability to incorporate uncertainties. Sabour and Dimitrakopoulos (2011) used a risk-based optimisation model which incorporates uncertainties from both the geological and economic factors while minimizing cost. However, they indicated the difficulties surrounding creating such planning process. Dimitrakopoulos and Ramazan (2008) argued that the difficulties are stemming from the calibration of the orebodies. Dimitrakopoulos (2011) agreed with the observation since in most cases stochastic mine planning is unable to account for the *in-situ* spatial variability of the deposit grades.

In literature, stochastic mine planning is conducted using optimisation tools. Kortelev et al. (1992) suggested that a feasibility plan will only include the optimal decision functions when applicable in the simulation. However, Dimitrakopoulos et al. (2007) stated that stochastic mine planning is better in calibrating the Net Present Value (NPV) of the project despite its shortcomings echoed above.

Uncertainty is introduced into the modelling process in studies by Dimitrakopoulos and Ramazan (2004), Dimitrakopoulos et al. (2007), Dimitrakopoulos and Ramazan (2008) and Dimitrakopoulos (2011). Stochastic mine planning is more aligned to actual temporal values assigned to both inputs and outputs of the actual mine production process. The preceding statement implies that, the use of a stochastic mine planning approach minimises model risks present in the deterministic approach due to the dichotomy between the estimated and actual data; see Section 1.1 in the chapter overview.

Based on the discussions presented above the deterministic mine planning approach is inappropriate in constructing feasible mine plans due to its assumptions. However, by incorporating uncertainties into the modelling of a mine plan an optimization algorithm can minimise the risks present in the deterministic approach. Flaws in deterministic mine planning are the main reasons fostering the development of more sophisticated mathematical models and optimisation techniques known as stochastic mine planning.

2.5 Shortcomings of deterministic mine planning approaches

Deterministic mine planning does not highlight an optimal sequencing of the mine production activities and fails to accurately estimate the completion time. This is because deterministic approaches use historical data as inputs and the assumption driving the process, is that, mining activities are deterministic. These shortcomings are mainly due to its inability to incorporate uncertainties and the randomness of mine production activities which are calibrated in the deterministic mine planning methods. Table 2.1 shows that, by taking the deterministic mine approach the variation in the "gold price" during May 2012 to August 2012 is ignored.

To overcome the shortcomings of the deterministic mine planning, various researchers developed stochastic optimisation algorithms incorporating random behaviour in some of the mine production activities. For example, Dimitrakopoulos (2011) developed a stochastic optimisation model which randomises the orebodies extracted in various mining blocks by assuming various paths per mining block modelled. Kortelev et al. (1992) looked at the feasibility of using mine plans in the decision making process and monitoring of a mining production process. On the other hand, King (2011) evaluated the complexity behind the concept of "optimality" to various mining engineers within the field. This is because some of the mine planners will consider incorporating uncertainties into their plans by using economic data and others by modelling uncertainties from the mining process itself.

Although stochastic mine planning incorporates uncertainties and minimizes risk when compared to deterministic mine planning, people are more important than software or algorithms due to their capacity to add or destroy project value (King, 2011). Therefore, Sabour and Dimitrakopoulos (2011) focused on the procedure integrating uncertainty and operational flexibility into open pit mine design selection. This model is different from other models where geological and market uncertainties are taken into account (Sabour and Dimitrakopoulos, 2011) . There is various literature supporting the development and modelling of the stochastic mine planning approach, namely, Mustafa (2010), Dimitrakopoulos (2011), King (2011) and Opoku and Musingwini (2013).

2.6 Limitations of stochastic mine planning approaches

Stochastic mine planning approaches are primarily geared to realistic optimization. In a two dimensional platform, an optimization method is solvable manually but still demanding, see appendix A. Optimization techniques are generally complex since they require an establishment of initial values (these values kick-start the optimization process), thresholds used as decision rules, tolerance levels and furthermore optimization techniques are computationally intense. For example, consider the stochastic model with a cluster of mine blocks by Dimitrakopoulos (2011). The cluster of mine blocks use multiple possible economic values and the optimization method maximizes

$$
\begin{gathered}
\left(s_{11}x_1^1 + s_{21}x_2^1 + \cdots\right) \\
\text{Subjected to} \\
s_{i1}x_1^i + s_{i1}x_2^i + \cdots \geq b_i \\
\vdots \\
s_{p1}x_1^p + s_{p2}x_2^p + \cdots \geq b_p
\end{gathered}
$$

Here s_{ip} is the economic value of block i in period j, x_i^p is the binary decision variable related to block i in period p and b_j is a given arbitrary value of the deposit in period j. In order to solve the above optimisation problem, one requires complex and powerful computational tools.

Many researchers mentioned that stochastic mine planning is critical in developing feasible mine production scheduling processes. Both the mining and mine planning processes are of importance in completing a mining project in time, within budget and with high returns. However, Dimitrakopoulos (2011) noted that stochastic mine planning approaches are unable to account for the *in-situ* spatial variability of the deposit grades.

According to Journel (1983) commodity modelling data has long-tailed distributions with the coefficient of variation in the range of 2 - 5. These effects mentioned by Journel (1983) cause the stochastic model to be unable to detect variability of deposit grades in various sections of the mine block developed in the modelling process. Journel (1983) noted that, these variations caused by the spatiality of the deposit grades cause some outliers. Furthermore, Journel (1983) suggested that this situation can be fixed by:

- truncating the high-valued data, usually called outliers in statistical theories; or
- smoothing out the data by working on some smoothing function, for example, their square roots, or natural logarithms.

The risk of the above mentioned remedial measure is losing the important information required in establishing a robust mine production scheduling process.

Stochastic approaches provided in literature do not assess the correlations between KPIs and mining activities. Mineable resources across the mining industry have reduced significantly when compared to the 1800s era. The economic factors driving the mining industry such as exchange rates, currency, country rating and other factors are volatile due to unstable circumstance across the global financial markets. In such circumstances it is of essence that key mining activities are identified to reduce mining costs thereby increasing the returns. By identifying the key mining activities, mine planners and other stakeholders involved in the mining process/project can focus on these variables in terms of optimizing mining returns or KPIs. A correlation study would be able to provide such analysis that would aid in the identification of key mining production activities versus assessed KPIs.

2.7 Summary

This chapter highlighted that research done to date on stochastic mine planning lacks descriptive statistics such as correlation which can be beneficial to the modelling process, especially in subdue economic climates. The next chapter discusses correlation methods in detail.

Chapter 3

CORRELATION METHODS

3.1 Chapter overview

This chapter reviews the importance of correlation studies in improving the understanding of a stochastic mine production system. Section 3.2 provides an introductory overview to correlation analysis. Section 3.3 discusses the pitfalls of correlation. Section 3.4 discusses different methods of modelling correlation. Section 3.5 links correlation to a stochastic mine production system.

3.2 Correlation analysis - mine planning perspective

Correlation studies are a field which investigates the association of various variables in a particular model. These studies depend on the joint distribution of the measured variables and their individual standard deviation. Therefore, in mine planning such a measure is helpful because it can provide an insight as to how movements in the mine production activities will translate to the KPIs monitored monthly. The problem of calculating correlation lies in the determination of the joint distribution, especially if the individual distributions are unknown.

Over the years various stochastic methods have been developed since most of the empiricists in science and engineering have realised their problems were not centralised in the normal distribution theory. Kutner et al. (2005) noted that, the development of correlation or association measurement originated from normal distribution and linear modelling theories. Furthermore, the problem deviates from the normal distribution and linear modelling theories since the error term grows enormously depending on the alternative method selected. Most modellers leverage on the fact that, the rejection region can be expanded beyond the traditional 5% depending on the limitations of the data and method used. Hence the use of non-conventional methods is permitted.

3.3 Shortcomings in correlation analysis or modelling

Correlation studies conducted for medical studies have limitations outlined below. Grimes and Schulz (2012) stated, the following shortcomings are associated with correlation analysis:

- Correlation studies fail to answer the five "W": Who? What? Why? When? Where?;
- Correlation studies are used as a supplement to a more sophisticated model. This implies that correlation alone cannot solve the problem but is an addition to an existing method. For example, this correlation study can be an addition to the existing stochastic mine planning method developed by Dimitrakopoulos (2011); and
- Hypotheses and theorems cannot be formulated from a correlation analysis.

Therefore, the correlation method developed can be seen as supplementing and not replacing the existing mine planning approach implemented in mining at present.

3.4 Correlation methods

In many related fields, correlation has been vastly investigated. For example, in medical science, Schrauzer et al. (1977) investigated the correlation effects of taking selenium in cancer patients. The results obtained when using Pearson correlation coefficient is negative and close to -0.80 for female patients; this suggests that taking selenium reduces cancer.

Huff and Shipp (1969) in the field of meteorology discussed the spatial correlation analysis for weather pattern involving storms. Spatial correlation is the multivariate of the Pearson correlation study. This correlation study in mining will assist by highlighting which of the mine production activities have either negative or positive impact on the identified KPIs. The correlation study will show which activities or variables are important to the KPIs. The following sub-sections discuss possible correlation methods that can be used in resolving the problem statement mentioned in Section 1.3.

3.4.1 Pearson correlation

The Pearson correlation is used widely in linear models. The Pearson correlation coefficient indicates the relationship of two random variables where the underlying assumption is that, these variables are linear and follow a normal distribution. Given X and Y as random variables, the Pearson correlation is calculated as follows:

Step 1. INPUT

X_i and Y_i are observed explanatory and response variables, respectively, $\forall\ i = 1, \cdots, n$.

Step 2. OUTPUT

ρ is the correlation measure.

Step 3. Pearson correlation method

Calculate the interaction term, that is COV(X,Y)$= \sum_{i=1}^{m} (Y - \mathbb{E}(Y))(X - \mathbb{E}(X))$ where,

$$\mathbb{E}(X) = \frac{1}{m}\sum_{i=1}^{m} X_i$$

and

$$\mathbb{E}(Y) = \frac{1}{m}\sum_{i=1}^{m} Y_i$$

are sample means. The modulus of the calculated quantity COV(X,Y) can be greater than 1, hence it is normalized as

$$\rho = \frac{\text{COV(X,Y)}}{\sigma_x \sigma_y},$$

where $\rho\ \epsilon\ [-1, +1]$, σ_x and σ_y are the respective standard deviations of the two random variables.

Pearson correlation method was developed as a supplementary approach in understanding the variables modelled under linear regression models. The entire correlation calculation is based on the assumption that, the response variable Y is linearly associated to the explanatory variable X. This correlation is deemed computable if the following assumptions hold in the modelled linear function:

- $Y \sim N(\mu, \sigma^2)$ since the error term $\epsilon_i \sim N(0, \sigma^2)$ and the linear combination of the normal random variable follows a normal distribution;
- $\lim_{m\to\infty}\left(\sum_{i=1}^{m}(Y_i - \mathbb{E}(Y))^2\right) = a$ or $\lim_{m\to\infty}\left(\sum_{i=1}^{m}(X_i - \mathbb{E}(X))^2\right) = b$, where "(a,b)" are convergent values, $(a, b)\ \epsilon\mathbb{R}^2$ and m is the sample size. This implies that, the scatterplot of the variance should approximate a uniform distribution pattern (known as homoscedasticity); and
- Y should be written as a linear combination of X_i.

The Pearson correlation function shown in Section 3.4.1 returns values between -1 and +1. $\rho>0$ indicates a positive correlation (i.e. an increase in X leads to an increase in Y) and $\rho<0$ indicates a negative correlation (i.e. an increase in X leads to an decrease in Y).

3.4.2 Spatial correlation

Spatial correlation is a multivariate correlation structure which assumes that the various random vectors compared within the matrix follow a normal distribution. This method follows a similar analogy to Section 3.4.1 but COV(X,Y) is not linear since $\mathbb{E}(X)$ and $\mathbb{E}(Y)$ are variant. The spatial correlation at the concentrated areas is very small. The methodology of this approach is not outlined due to its complexity.

3.4.3 Kendall-Tau correlation

Kendall Tau correlation is a nonparametric correlation method that tests rank correlation coefficients. Kendall Tau (i.e. τ_k) is usually used for discrete circumstances where the tested variable pairs are ranked and their ranking is used to decide whether they are concordant or discordant. A pair of variables is concordant if and only if x_i and y_i have the same rank at level i and discordant if otherwise. Numerically Kendall Tau correlation is calculated as follows:

Step 1. INPUT

X_i and Y_i are observed explanatory and response variables, respectively $\forall\ i = 1, \cdots, n$.

Step 2. OUTPUT

τ_k is the correlation measure.

Step 3. Kendall-Tau approach

Create pairs of X and Y,

$$F = \{(X_1, Y_1), (X_2, Y_2), (X_3, Y_3), \cdots, (X_j, Y_j)\}.$$

Then sort $F \uparrow$. Calculate

$$\tau_k = \frac{P(A) - P(B)}{\frac{1}{2}N(N-1)}.$$

where $P(A)$ is the number of concordant pairs, $P(B)$ is the number of discordant pairs and N is the total sample size modelled.

However, nonparametric tests are less powerful in predicting correlation than conventional methods since a lot of variation in the variables modelled is lost in the ranking and normalising processes. Other tests such as the Mann-Whitney follow the same approach as outlined in the Kendall-Tau correlation approach.

3.5 Application of correlation modelling in mining

As discussed in Chapter 1 that the stochastic mine planning approach operates under the assumption that the KPIs and mine production activities are stochastic, these can be defined as follows,

$$\{X_i(t) : -\infty < t < \infty\}, \forall i = 1, \cdots, n(\text{denotes the studied mine production activities}).$$

and

$$\{Y_i(t) : -\infty < t < \infty\}, \forall i = 1, \cdots, n(\text{denotes the studied KPIs}).$$

Since $X_i(t)$ and $Y_i(t)$ are random variables, then from preceding sections these are also nonlinear. Furthermore, in the correlation approach, dependency between variables has to be assumed because if not, correlation equates to zero. Modelling dependency is another stumbling block in the numerical analysis, especially if the data available is insufficient to infer the relationship between the factors to be tested. The copulas method[1] can be used in this case. The copulas method works best in stable functions but, mining data is not predictable due to its inherent uncertainties. From these discussions, it can be gauged that modelling correlation under the assumptions outlined in Chapter 1 will be complex and difficult to compute.

To address the shortcomings outlined, correlation is modelled under linear assumptions. Kutner et al. (2005) stated that MLR models can be translated into a linear model using log transformation. Magagula et al. (2015) proved that the KPIs modelled against the mine production activities under a linear function follow the MLR model. This implies that the Pearson correlation method can be used and inferred on the MLR model. The method below briefly describes the correlation computation under linear assumptions for mine planning.

Let $W_i(t)$ be the mine production activities and $R_i(t)$ be the KPIs.

$W_i(t) = \left(X_i(t_k) - X_i(t_{k-1})\right)/X_i(t_k)$ such that $W_i(t) \epsilon \mathbb{Q}$.

and

$R_i(t) = \left(Y_i(t_k) - Y_i(t_{k-1})\right)/Y_i(t_k)$ such that $R_i(t) \epsilon \mathbb{Q}$.

Then create $G_i(t)$ that is partitioning $R_i(t)$ into positive natural numbers which implies $G_i(t) \epsilon \mathbb{Z}^+$.

Then

$$G_i(t) = W_i(t) \cdots \ (G_i(t) \text{ has } m \text{ levels, } m > 2)$$

The model outlined above is a Multinomial Logistic Regression (MLR) model which is dis-

[1] Joint distribution creator method given unrelated distributions.

cussed in detail in Chapter 4. Using this notation it follows that the MLR model is defined as,

$$G_i(t) = 1 + \exp(-\beta_0 - \beta_j W_i(t)) + \epsilon_i,$$

where $G_i(t)$ is the transformed KPIs, $W_i(t)$ is the transformed explanatory variables, (β_0, β_j) are unknown parameters and ϵ_i is the error rate taken.

Taking natural logarithms on both sides transforms this into a linear function which implies one can calculate the correlation using the Pearson methodology. It follows that,

$$\underline{\rho} = \left\{ \sum_{i=1}^{k} (G_i(t) - \mathbb{E}(G_i(t)))\,(W_i(t) - \mathbb{E}(W_i(t))) \right\} / \sigma_{W_i(t)}\sigma_{G_i(t)}.$$

Since natural logs were applied under the nonlinear environment, the correlation is calculated as,

$$\underline{\theta} = \exp(\underline{\rho}) - 1.$$

3.6 Summary

Stochastic mine planning incorporates uncertainties from geological phenomena or economic perspectives or both (Dimitrakopoulos, 2011). Data from the stochastic event is usually heavy tail distributed, see (Journel, 1983). Due to the fact that distributions are skewed either to the left or right, they do not follow the normal distribution. Hence Pearson and Spatial correlation methods cannot be used without modification. The Kendall-Tau correlation method can be implemented instead. Kutner et al. (2005) further proved that nonparametric tests are less powerful in predictive power when compared to parametric methods such as the Pearson correlation method. This is among many reasons for modifying the Pearson correlation method to fit the problem outlined in Chapter 1. The next chapter presents the proposed methodology in answering the research question.

Chapter 4

MLR METHODOLOGY

4.1 Chapter overview

This chapter discusses the Multinomial Logistic Regression (MLR) model in relation to the correlation method discussed in Section 3.5.1. Section 4.2 discusses the MLR method in general. Section 4.3 outlines all tests that should be conducted to assess data fit to the MLR model. Section 4.3 discusses the MLR method in the context of modelling correlation. Section 4.4 outlines all estimation methods used in obtaining the results shown in Chapter 5.

4.2 MLR model in a mine planning context

There are several ways of conducting the correlation analysis as discussed in preceding chapters but, these methods are primarily driven by the assumption of the association between KPIs and mine production activities. These models are both linear models and non-linear models. Kutner et al. (2005) stated that if correlation or association is assumed to be linear, a regression model can be used in calibrating this correlation analysis in a functional form. This is normally defined as, Y $= \sum_{i=1}^{n} \beta_i x + \epsilon_i$, where Y is the dependent variable (i.e. KPIs), x's are the independent variables (i.e. mine production activities), β's are the unknown parameters estimated by the *least square methods*, and ϵ(s) is the error term which is normally distributed. The above model falls in a group of models called linear regression models or Generalised Linear (GL) models. However, Sabour and Dimitrakopoulos (2011) noted that economic data is an essential input into the mine planning process. Economic data are random parameters that do not follow a normal distribution. This implies that nonlinear models, namely, Logistic Regression (LR), Multinomial Logistic Regression (MLR) models, stochastic models (i.e. Levy-process models, Brownian motion models, etc.), Exponential models, etc. can be considered. However, these models are fairly complex to determine numerically due to model assumptions which must hold for the model to be employable. Based on the data description denoted in studies mentioned earlier, a nonlinear model (MLR approach) will be used in attaining the covariance structure presented.

The previous section outlined the main reasons behind selecting a correlation-based method in answering the research problem. The correlation analysis approach proposed can assist in understanding the behavior and influences of certain mine production activities with regards to the measured KPIs. The model can also assist in establishing which mine production activities require minimisation or maximisation in attaining the desired KPIs.

Let γ be the ratio between planned targets (denoted by α) and actual mine outputs or outcomes (denoted by θ). Then γ will be the KPI calculated as follows:

$$\gamma = 1 - \left[(\alpha - \theta)/\alpha \right]. \tag{4.1}$$

When γ is 1, it implies that the mine meets its target as stipulated in the plan. Various open pit mines use different scales in measuring actual outputs against targets. For example, an AngloGold Ashanti mine in Carletonville will differ when compared to a GoldFields mine in the Free State because their capacity, resources, funding from investors, mine design and geographical issues are completely unique. Before proposing the model, one needs to discretize γ so that the final model developed is applied across all mines with few modifications on the resultant KPI (i.e. γ). Table 4.1 describes the discretizing approach for γ.

Table 4.1: Discretisation of γ

Level	γ's interval	Meaning of the levels
1	$0.80<\gamma \leq 1$	Mine planning model adequately calibrates planned targets and incorporates uncertainties and risks in the mining activities.
2	$0.75 \leq \gamma<0.80$	Mine planning model calibrates planned targets accordingly. However, there are few deviations between the historical data and current mining outcomes. For example, if the mining activities used in modelling planned targets are the availability of loading equipment; there might have been a sudden delay in loading by the loading equipment for a few hours which impaired on the actual transported tonnages.
3	$0.50 \leq \gamma<0.75$	Differences between planned and actual are near critical levels and therefore the planning process needs recalibrating.
4	$0.25 \leq \gamma<0.50$	Differences between planned and actual are at critical levels and therefore the planning process needs recalibrating.
5	$0.00<\gamma \leq 0.25$	The planned targets are not attained in the actual mining production system. Therefore, there may be a sudden major move in one of the mining activities used in determining planned targets which was not originally included in the planning process.
6	$\gamma = 0.00$	The planned targets are not attained in the actual mining production system as there was no production due to such factors as the legally prescribed Section 54 safety stoppages.

This discretization of KPIs allows the research question to be solved using the Multinomial

Logistic Regression (MLR) model. Mathematically the MLR model is generally defined as:

$$\gamma = \left(1/\left\{1 + \sum_{i=1}^{N-1} e^{\beta_k X_i}\right\}\right) \tag{4.2}$$

where, N is the number of level in the response variable and β_k is the unknown parameter of X_i (which are the mine production activities).

4.3 Multinomial Logistic Regression properties and tests

The MLR model is an easily interpretable model, although it operates under non-linear assumptions. Bayaga (2010) stated that the MLR model has the following assumptions which makes the modelling process not to be constructive:

- It does not assume any normal distribution of variables involved;
- The relationship between the response variable and independent variables is nonlinear;
- It does not assume homoscedasticity (constant variance); and
- The independent variables need to be discrete or continuous variables.

Kutner et al. (2005) stipulated that if the nonlinear model can be translated into a linear model, the same linear principles and tests associated with it can be used. These models are known as the Generalised Linear Model (GLM)[1]. Hair, Jr. et al. (2010) proved that under logit link function or log transformation the MLR model can be re-written as a GLM approach which is defined as,

[1]Hair, Jr. et al. (2010) defined GLM as linear model which are based on three components: (1) a variate formed by the linear combination of independent variables (i.e. the mine production activities), (2) a probability distribution specified by the researcher based on the characteristics of the dependent variables, and (3) a link function that denotes the connection between the variate and the probability distribution.

$$
\begin{aligned}
\text{Logit}_i = \ln\left\{\frac{\gamma}{1-\gamma}\right\} &= \ln\left\{\frac{\left(\frac{1}{1-\sum_{i=1}^{N-1} e^{\beta_k X_i}}\right)}{\left(1-\frac{1}{1-\sum_{i=1}^{N-1} e^{\beta_k X_i}}\right)}\right\} \\
&= \ln\left\{\frac{1}{\cancel{1-\sum_{i=1}^{N-1} e^{\beta_k X_i}}} \times \frac{\cancel{1-\sum_{i=1}^{N-1} e^{\beta_k X_i}}}{-\sum_{i=1}^{N-1} e^{\beta_k X_i}}\right\} \\
&= \ln\left\{\frac{1}{-\sum_{i=1}^{N-1} e^{\beta_k X_i}}\right\} \\
&= \sum_{i=1}^{N-1} \beta_k X_i \\
&= (\beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n)\,.
\end{aligned} \tag{4.3}
$$

Since this MLR model can be translated into GLM measure by using logit link function or log transformation, this implies that $\exp(\rho)$ is suitable to model correlation in nonlinear scenario articulated by equation (4.2). However, $\exp(\rho)$ where ρ is the "Pearson correlation" does not translate accurately from the linear into the nonlinear environment due to the following limitations:

- For $\rho>0$, it follows that $e^{(\rho)}>0$;
- For $\rho = 0$ (i.e. suggesting that there is no correlation) translates into a correlation of 1 (i.e. $e^0 = 1$) in the nonlinear setting, that is a perfect positive correlation; and
- Only negative "Pearson" correlation translates accurately into the nonlinear scenario inferred by equation (4.3).

Boundaries are required for the nonlinear correlation measure (denoted by θ) to mimic a similar structure outlined by "Pearson" formulation that is,

$$\theta = \begin{cases} \rho<0, & \theta<0 \\ \rho = 0, & \theta = 0 \\ \rho>0, & \theta>0 \end{cases} \tag{4.4}$$

where $\theta = e^{\rho} - 1$. This nonlinear "Pearson correlation" holds if the data is modelable under the MLR scenario defined in equation (4.2).

According to Bayaga (2010) the MLR approach is semi-nonparametric since it does not have to adhere to the assumptions stipulated in linear models or any other parametric models. The only assumption required is that, the response variable must be nonmetric. Therefore, the normal model validation process followed in the MLR modelling framework will be used in proving the data is multinomial before using the nonlinear correlation defined above.

4.4 Goodness-of-fit test

Goodness-of-fit tests are methods used in assessing whether the data fitted in the model adheres to the underlying assumptions. This is an important step because if the data passes all tests defined below, it implies that the MLR model is an appropriate model to use and furthermore equation (4.3) holds.

4.4.1 Akaike Information Criterion (AIC) test

The AIC test measures how well data fits into the MLR model. This test is calculated as:

Step 1. **INPUT**

X_i, Y_i as defined in Section 4.2.1

Step 2. **OUTPUT**: AIC

Step 3. **AIC goodness-of-fit method**

Let k = number of parameters estimates (i.e. $\beta_j \ \forall \ j = 1, \cdots, m$) and N = number of observations.

3.1 The likelihood function is given by $Y_i = \left(1 + e^{-\left(\beta_0 + \sum_{i=1}^{N} (\beta_j X_j)\right)}\right)$ and natural log is the solution in equation (4.3) (i.e. denoted by ln (likelihood)).

3.2 $\therefore AIC = -2 \times \ln(\text{likelihood}) + 2 \times k$.

3.3 The larger the AIC, the poor the fit. For this research a threshold is set at 120 which implies that anything greater than 120 indicates poor fit.

4.4.2 Bayesian Information Criterion (BIC) test

The difference between AIC and BIC is that BIC incorporates N which is problematic to determine if complete data is not provided. Hence in literature there are fewer modifications of BIC with an appropriate N. A similar method to the AIC approach is used to determine BIC. BIC is calculated as

$$BIC = -2 \times \ln(\text{likelihood}) + \ln(n) \times k$$

Most computer packages assume $n = e^N$ unless provided.

4.4.3 -2 Residual Log Likelihood test

It is similar to AIC and BIC in functionality. This test is not used in interpreting the goodness-of-fit tests due to its inherent biasness. Biasness of this measure is out-of-scope for this research.

4.4.4 Error Rate or Residual Analysis

Error rate analysis observe the data against the fitted likelihood function and deviation from the fitted pattern. Kutner et al. (2005) stated that the residual can be defined as,

$$\text{Residual} = Y_i - \hat{Y}_i,$$

where Y_i is the actual response and $\hat{Y}_i$ is the predicted response.

$$\text{Error rate} = \left(\frac{Y_i - \hat{Y}_i}{Y_i}\right) \times 100.$$

The rejection area set at 35% which implies that if this error rate is greater than this, it indicates a poor fit.

4.4.5 Graphical test - Natural log function transformation

Myers and Montgomery (1997) stated that in recent years there is an exponential demand for statistical models to solve complex problems that are based on non-normal errors. Due to uncertainties in the mining process, the function modelable is usually nonlinear and fits the modelling description outlined by Myers and Montgomery (1997). Most theories and estimation methods are based on the assumption of normal distribution. Myers and Montgomery (1997) proved that if a nonlinear function can be rewritten as GLM, one can leverage on the normal distribution characteristics-and-linear model, see (Rice, 2007; Hair, Jr. et al., 2010). Fahrmier and Kaufmann (1985) stated that most statistical analyses and models are based on underlying assumption of asymptotic properties associated with Maximum Likelihood Estimate (MLE) methods. Myers and Montgomery (1997) stated that GLM models operate under the same model assumptions scenario as linear models, namely;

- linearity;
- normal distribution; and
- independence.

Equation (4.3) has shown that the MLR when using natural log is transformed into a linear function which implies that the transformed response variable $\hat{Y}_i$ follows a normal distribution. To plot the normal distribution as outlined in Chapter 5, various intervals are created, that is, $\{[Y_1, Y_{10}), \cdots, [Y_{N-p}, Y_N)\}$ which are called "bins". Thereafter, one can calculate the count per interval. Let R be the number of bins or intervals created. One then plots the count on the x-axis against the corresponding number of bins on the y-axis. The smoothing of this scatter plot will show a distribution close to a normal distribution if the data follows an MLR modelling process, see Chapter 5 for the graphs. For example BIC $\geq$ 120 indicates a bad fit and one would expect the error rate to be more or less 0.25. Kutner et al. (2005) suggested that the goodness-of-fit tests and the error rate are functional when the sample used is large enough.

4.5 Maximum Likelihood Estimate (MLE) method

MLE is a gradient-based methodology which is used to estimate the unknown parameters. This analysis provides important information pertaining to the significance of the independent variables in the model versus response variable. In equation (4.3) there are unknown parameters which need to be estimated. The Maximum Likelihood Estimation (MLE) method is used in estimating these β_k as outlined earlier. Kutner et al. (2005) proved that these estimates are accurate since they comply to the Cramer-Rao inequality theorem. Using Bohning (1992)'s algorithm the method shown below can be extended to fit the MLR scenario under the Newton-Raphson method. Furthermore, this MLE method employs a known approach called Least Square Method (LSM) in calculating the unknown parameters. Traditionally, the LSM has been used to minimize the squared residual function Q for the linear regression function and then using the gradient of Q in solving for the unknowns. The following algorithm can be used.

Algorithm 1 Least-Square Method (LSM) for estimating unknown parameters

1. **INPUT**

X_i, Y_i as defined in Section 4.2.1

2. **OUTPUT**

β_k as defined in equation (4.3)

3. **Optimisation LSM method**

3.1 Let $Y_i = \left(1 + e^{-\left(\beta_0 + \sum_{i=1}^{N} (\beta_k X_i)\right)}\right)$ and then $\ln(Y_i) = -\sum_{k=1}^{N-1} e^{\beta_k X_i}$.

3.2 Using Kutner et al. (2005) formulation it follows that, the squared residual function Q is $(Y_i + \beta_0 + \beta_1)^2$ for k=2 under logistic regression.

3.3 The following are the constraints applied on the objective function in this minimisation problem, that is,

$$\begin{aligned}\frac{\partial Q}{\partial \beta_0} &= 2\sum_{i=1}^{N}(Y_i - \beta_0 - \beta_1 X_i) \cdots (\beta_0 \ \& \ \beta_1 \text{ are estimates of the unknowns}) \\ &= 2\left(\sum_{i=1}^{N} Y_i + n\beta_0 + \beta_1 \sum_{i=1}^{N} X_i\right),\end{aligned}$$

and

$$\begin{aligned}\frac{\partial Q}{\partial \beta_1} &= 2\sum_{i=1}^{N} X_i(Y_i - \beta_0 - \beta_1 X_i) \cdots (\beta_0 \ \& \ \beta_1 \text{ are estimates of the unknowns}) \\ &= 2\left(\sum_{i=1}^{N} X_i Y_i + \beta_0 \sum_{i=1}^{N} X_i + \beta_1 \sum_{i=1}^{N} X_i^2\right).\end{aligned}$$

3.4 Equate these partial derivatives to zero. 3.5 Using simultaneous equation method solve the two unknowns it follows that,

$$\beta_1 = -\left(\frac{\sum_{i=1}^{N} (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{N} (X_i - \bar{X})^2}\right),$$

and

$$\beta_0 = -\left(\bar{Y} + \beta_1 \bar{X}\right),$$

where

$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} (X_i)$ and $\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} (Y_i)$.

3.6 Since natural logs were applied to remove the exponential function, it follows that the unknowns under logistic are,

$$\acute{\beta}_0 = e^{\beta_0},$$

and

$$\acute{\beta}_1 = e^{\beta_1}.$$

Let the p-value be the error taken in estimating the unknowns above and let α be the significance level that error taken should not pass.

4. **Intrepretation**

All unknowns with p-value $\leq \alpha$ are assumed to follow the MLR model (reject $H_0 : \beta_k = 0$) refer to the next chapter.

4.6 Summary

This chapter presented the methodology, mathematical attributes of the model and validation tests needed to assess the validity of the measure. The central conclusion by various researchers including Kutner et al. (2005) is that, when building prediction models, it is important that the measure predicts the desired target value for new data accurately. Empirical results are discussed in Chapter 5.

Chapter 5

MODEL ANALYSIS

5.1 Chapter overview

The formulations developed in previous chapters are tested in this chapter through using empirical data from a hypothetical platinum mine production system based on real data. Section 5.2 discusses the extracted data and defines various parameters used in the model. Section 5.3 discusses modifications implemented on γ in equation (4.3) due to the limited data sample size being used in modelling. Section 5.4 discusses the results obtained from methodologies outlined in Chapters 3 and 4.

5.2 Data

The data used in this dissertation was extracted from a platinum mine in South Africa. The recent strikes in the latter part of 2014 to early 2015 meant that some of the data points were not usable since there were no production activities. Seven KPIs were selected from the sampled variables and thereafter nine independent variables were used. The extracted data stretches from the 2008 to 2015 financial years; the number of usable data points are limited. Tables 5.1 and 5.2 highlight the variables used in the modelling process described in Chapter 4.

Table 5.1: Explanatory variables

Variable Name	**Parameter**
Face_Length (m)	X_1
Blast	X_2
Teams	X_3
Offreef (%)	X_4
Allow_OB (tonnes)	X_5
Error_OB (tonnes)	X_6
Channel_Width (tonnes)	X_7
Weighline (tonnes)	X_8
Replacement factor (m^2/m)	X_9

The detailed explanations on the variables in Table 5.1 are:

- Face_Length mined is the area mineable to access the platinum.
- Blast is the number of blast planned.
- Teams is the number of teams planned to mine.
- An Allow_O_B (allowable overbreaks) is a process where the mine authorizes further blasting and mining over the agreed upon area in the planning phase.
- An Error_O_B stands for error overbreaks where the mine realizes that the conducted blasting and mining were done incorrectly or excessively based on the mine planning conducted.
- Offreef is the area that is mined but has no platinum mineralization.
- Weighline is a mine production activity where the official measurement of the ore extracted is done (Bartlett and Liebenberg, 2014).
- Replacement factor is the ore reserves kept for sustainability purposes (Musingwini (2009), Musingwini (2010), Ferreira (2012)).

- According to Wimberger (2004) "channel_width" is the confidence interval of grades of platinum acceptable as valuable ore in a mining process and any ore below this encourages allowable overbreaks.

Table 5.2: Response variables

Variable Names	**Parameter**
Centares (m^2)	Y_1
On_reef_dev (m)	Y_2
Off_reef_dev (m)	Y_3
Channel_Dilution (tonnes)	Y_4
Survey_Call (tonnes)	Y_5
Current_Sweepings (m^2)	Y_6
Mill (g/t)	Y_7

Then, these variables defined in Table 5.2 are defined as follows:

- Centares (Total_Eq) is the total area blasted in the mine production process.
- On_reef_dev is the on_reef development for establishing mining areas.
- Off_reef_dev allows access to the ore reserves.
- Current_Sweeping is a process where the blasted ore is re-checked through sweeping for any platinum that might have been missed and remaining behind in fines from the mined-out areas;
- Mill_grade is the quality of the platinum ore extracted in the mining production process.
- Survey_Call is a process where mine surveying is done to check whether the mining process is progressing as stipulated in the planning process.
- Channel dilution is a quantity used to classify the extracted ore in the mining process. This channel dilution is used to quantify overbreaks in the mining process according

to Wimberger (2004).

5.3 Model analysis

The results attained when γ is partitioned into six levels were not conclusive. This is understandable since the sample data is less than hundred. The inconclusive results translated into a Type I error[1]. This high degree of Type I error can be fixed through implementing the following modelling initiatives:

- Increasing the sample size by using simulation techniques;
- Reducing the discretization levels shown in Table 4.1 in Chapter 4 since the data extracted is small; and
- Use an alternative algorithm in SAS[2] to execute the MLR approach outlined in Chapter 4 and thereafter compute the correlation when data fits the MLR model.

5.3.1 Simulation approach

The sample size used in this research is very small to implement a multivariate technique such as the MLR model comprehensively. Hence inconclusive results will be attained and this theory is supported by Hair, Jr. et al. (2010). Another shortcoming is that, the small sample is further truncated into six levels based on the definitions stipulated in Table 4.1 which reduces modellable data points per block. Another observation is that, the segregation into these blocks is not equally distributed, for example level 4, level 5 and level 6 in Table 4.1 for the extracted data have close to zero data points whereas the mode of these distributions are in level 1 or level 2 in the data integrity test conducted. These unequal allocations among the six blocks translate into inconclusive results.

[1] "Type I" is the probability that the null hypothesis is rejected when it is true. The opposite of this error is "Type II", that is, the probability of accepting the null hypothesis when it is false.

[2] SAS is the statistical tool used for data mining & modelling.

This shortcoming of having insufficient data points can be fixed by an "importance sampling" approach widely used in recent Extreme Value Theory (EVT) studies which were dedicated at solving unbalanced problems in natural sciences. Importance sampling is operative when or where sampled observations are evident in the process modelled. The "importance sampling" technique operates by drawing an alternative distribution whose support is significantly concentrated in the truncation region (i.e. level 4, level 5 and level 6 in Table 4.1 in this case). In principle, numerical importance sampling is,

$$\begin{aligned} \int_{\mathbb{F}} sf\left(s\right)ds &= \int_{\mathbb{G}} s\frac{f\left(s\right)}{g\left(s\right)}g\left(s\right)ds \\ &= \int_{\mathbb{G}} s(\omega(s))g\left(s\right)ds. \end{aligned} \tag{5.1}$$

where $\mathbb{F}$ is the sampling region,
$f\left(s\right)$ is the density of s over $\mathbb{F}$,
$g\left(s\right)$ is the new density of s over $\mathbb{G}$,
and $\omega\left(s\right) = \frac{f(s)}{g(s)}$ is the sampling weight pre-defined in the simulation process.

The definition provided in equation (5.1) is similar to the Bayesian statistic discussed by Rice (2007). Operationalizing a Bayesian algorithm numerically is computationally demanding and overlaying such processes on the MLR approach might present computational difficulties. Therefore, this importance sampling approach is not considered in minimising model risks seen in the above analysis conducted thus far.

5.3.2 Redefining γ's levels in Table 3.2

Based on the data extracted six levels do not provide large datasets and these levels are reduced to three levels. The γ levels in Table 4.1 are redefined as follows:

Table 5.3: Redefining the discretisation of γ

Level	**γ's interval**	**Meaning of the level**
1	$0.75<\gamma \leq 1$	Mine planning model calibrated planned targets accordingly and incorporated uncertainties and risks into the mining activities hence minimal deviation between planned targets and actual outputs.
2	$0.50<\gamma \leq 0.75$	Differences between planned and actual are at intermediary levels and when below 0.50, the planning process needs to be re-calibrated.
3	$0.00 \leq \gamma \leq 0.50$	The planned targets were not attained in the actual mining production system. Therefore, there might be a sudden major move in one of the mining activities used in determining the planned targets which was not originally included in the planning process hence the misalignment.

The MLR approach outlined in Chapter 4, under the new levels as shown in Table 5.3 is still applicable because the model is MLR if and only if the response variables or KPIs have three or more levels. The main reason for the truncation of the response variable into three levels is that, under the equal segregation assumption each block will have more or less than 32 data points which might result in better output than the one yielded previously. In sections to follow Table 5.3 levels will be used.

5.4 MLR validation results

In Section 3.5.1 it was defined that correlation under the nonlinear scenario θ is calculated as $\theta = \exp(\rho) - 1$. For this assertion to be true the data needs to be modellable under the MLR approach defined in equation (4.3). Three empirical tests were conducted using SAS, namely:

- The Maximum Likelihood Estimation (MLE) analysis;
- Goodness-of-fit tests;
- The final test to assess the error rate taken in estimating the unknown parameters.

However, the Pearson correlation method should adhere to linearity, normal distribution characteristics and independent assumptions. The linearity and independent assumptions are explained by the transformation from MLR to GLM, see equation (4.3). Kutner et al. (2005) and Hair, Jr. et al. (2010) proved that GLM approaches have similar behaviour to traditionally linear models. The normal distribution assumption will be tested using the natural logarithm function transformation and inferring the normal Probability Density Function (PDF). If all of these checks adhere to the set threshold the correlation analysis computed will be assumed to be accurate from a numerical and mathematical perspective.

5.4.1 Maximum Likelihood Estimation (MLE) analysis

Section 4.5.1 outlined, how the unknown parameters are calculated. The MLE analysis tests whether these unknowns for the respective explanatory variables are influential using the following hypotheses.

$$H_0 : \beta_k = 0, \forall k = 1, \cdots, k-1 \cdots \text{(null hypothesis), and}$$

$$H_1 : \beta_k \neq 0, \forall k = 1, \cdots, k-1 \cdots \text{(alternative hypothesis).}$$

For all tests conducted in Section 5.4, it can be assumed that a significance level $\alpha = 0.10$ is acceptable. The null hypothesis is rejected when the calculated p-value is less than α which suggests that the explanatory variable under the MLR model contributes to γ. Otherwise accept the null hypothesis. The analysis conducted below assesses the data fit to the MLR model in all tested KPIs in this study. The significance level for all tests is set at 10% because the sample size is less than 100.

The analyses conducted in Table 5.4 to 5.10 are sufficient to conclude that the data follows an MLR approach.

Table 5.4: MLE analysis for Centares (Y_1)

Parameter	**p-value in %**
Intercept	<0.1%
X_1	<0.45%
X_2	46.23%
X_3	<0.10%
X_4	19.85%
X_5	9.30%
X_6	10.00%
X_7	26.47%
X_8	6.33%
X_9	77.15%

The p-values for Face_Length (X_1), Teams (X_3), Allow_OB (X_5), Error (X_6) and Channel_Width (X_7) are smaller than the 10% significance level. The null hypothesis is rejected which means that these five production activities have a significant effect on the Centares mined.

Table 5.5: MLE analysis for Onreef_dev (Y_2)

Parameter	**p-value in %**
Intercept	<0.1%
X_1	<8.34%
X_2	59.38%
X_3	85.57%
X_4	3.58%
X_5	8.84%
X_6	<0.1%
X_7	<0.1%
X_8	<0.1%
X_9	6.60%

The p-values for Face_Length (X_1), Offreef (X_4), Allow_OB (X_5), Error_OB (X_6), Channel_Width (X_7), Weighline (X_8) and Replacement_factor (X_9) are smaller than 10%. These tests shows that these the seven production activities have a significant effect on the Onreef development. X_2 and X_3 have no effect on Y_2.

Table 5.6: MLE analysis for Offreef_dev (Y_3)

Parameter	**p-value in %**
Intercept	<0.1%
X_1	<2.63%
X_2	<0.1%
X_3	73.64%
X_4	<0.1%
X_5	21.64%
X_6	<0.1%
X_7	24.85%
X_8	<0.1%
X_9	0.19%

The p-values for Face_Length (X_1), Blast (X_2), Offreef (X_4), Error_OB (X_6), Weighline (X_8) and Replacement_factor (X_9) are statistically significant. The null hypothesis is rejected which implies that the six production activities have a significant effect on the Offreef_dev.

Table 5.7: MLE analysis for Channel_Dilution (Y_4)

Parameter	**p-value in %**
Intercept	<0.1%
X_1	28.51%
X_2	1.87%
X_3	6.02%
X_4	3.18%
X_5	0.41%
X_6	1.65%
X_7	96.50%
X_8	5.93%
X_9	13.14%

The p-values for Blast (X_2), Teams (X_3), Offreef (X_4), Allow_OB (X_5), Error_OB (X_6) and Channel_Width (X_7) are statistically significant.The null hypothesis is rejected. This means that blast, number of teams, Offreef, allowable overbreaks, error overbreaks and channel width have an effect on the Channel_Dilution.

Table 5.8: MLE analysis for Survey_Call (Y_5)

Parameter	**p-value in %**
Intercept	<0.1%
X_1	42.14%
X_2	16.99%
X_3	67.20%
X_4	28.19%
X_5	0.05%
X_6	0.06%
X_7	53.47%
X_8	3.86%
X_9	67.61%

The p-values for Face_Length (X_1), Blast (X_2), Teams (X_3), Offreef (X_4), Channel_Width (X_7) and Replacement_factor (X_9) are greater than the 10% significance level. One cannot reject the null hypothesis. This implies that none of these production activities have a significant effect on Y_5. It can be argued that since the MLE analysis conducted above shows that only three of the explanatory variables reject the null hypothesis, it is not sufficient to conclude that the data under this KPI follows an MLR. MLR is a multivariate method and Hair, Jr. et al. (2010) stated that its precision is heavily reliant on the sample size. Therefore if the sample size is increased adequately; an assumption is made that X_5, X_6 and X_8 p-values will be less than the significance level of 10%. On this basis the MLR assumption is proven to hold. This as well suggests the need to study the tolerance level of the inference testing in MLE for the MLR model as the sample size increases.

Table 5.9: MLE analysis for Current_Sweepings (Y_6)

Parameter	**p-value in %**
Intercept	<0.1%
X_1	76.20%
X_2	38.29%
X_3	29.08%
X_4	5.87%
X_5	5.35%
X_6	6.83%
X_7	0.81%
X_8	<0.1%
X_9	24.52%

Offreef (X_4), Allow_OB (X_5), Error_OB (X_6), Channel_width (X_7) and Weighline (X_8) have a significant effect on Current_Sweepings. The null hypothesis for these tests is rejected. These test results imply that the correlation can be modelled using the formulation in Chapter 3.

Table 5.10: MLE analysis for Mill (Y_7)

Parameter	**p-value in %**
Intercept	<0.1%
X_1	33.58%
X_2	<0.1%
X_3	0.89%
X_4	10.36%
X_5	0.99%
X_6	2.54%
X_7	73.56%
X_8	7.71%
X_9	14.26%

The p-values for Blast (X_2), Teams (X_3), Allow_OB (X_5), Error_OB (X_6) and Weighline (X_8) are statistically significant. The null hypothesis is rejected. These means that the following production activities Blast, Teams, Allow_OB, Error_OB and Weighline have a significant effect on the Mill. These results show that the data follows an MLR model and correlation can be modelled using the formulation in Chapter 3.

5.4.2 Goodness-of-fit test - MLR approach

The previous tests highlighted which explanatory variables (i.e. mine production activities) contribute the most to the modelled KPIs. Sections 4.4.1 to 4.4.4 outlined all formal-based goodness-of-fit test methodologies, whose results are discussed below. Normally small results from goodness-of-fit tests indicate a good fit. For the tests suggested in Table 5.11 to Table 5.17; less than 120 indicate good fit whereas greater than 120 indicate poor fit. The following are the interpretations of the results shown in the tables.

Table 5.11: Goodness-of-fit tests for Centares (Y_1)

Test/Criterion	**Statistic**
-2 Res Log Likelihood	82.0
AIC	84.0
BIC	86.4

AIC and BIC results are less than 120 which implies that, the data fit the MLR approach. This conclusion drawn in Table 5.11 is in agreement with the outcome in Table 5.4.

Table 5.12: Goodness-of-fit tests for Onreef_dev (Y_2)

Test/Criterion	**Statistic**
-2 Res Log Likelihood	142.2
AIC	144.2
BIC	146.6

Table 5.12 results are greater than 120. This indicates poor fit to MLR which is in contrast to conclusions drawn in Table 5.5. Therefore, further tests had to be conducted.

Table 5.13: Goodness-of-fit tests for Offreef_dev (Y_3)

Test/Criterion	**Statistic**
-2 Res Log Likelihood	128.1
AIC	130.1
BIC	132.5

The results in Table 5.13 indicate a poor fit for the MLR approach and are in contrast with the conclusion taken in Table 5.6. Additional tests were done as all results are greater 120.

Table 5.14: Goodness-of-fit tests for Channel_Dilution (Y_4)

Test/Criterion	**Statistic**
-2 Res Log Likelihood	55.4
AIC	57.4
BIC	59.8

The results in Table 5.14 indicate a good fit for the MLR model which is in agreement with the MLE analysis in Table 5.7.

Table 5.15: Goodness-of-fit tests for Survey_Call (Y_5)

Test/Criterion	**Statistic**
-2 Res Log Likelihood	106.0
AIC	108.0
BIC	110.4

The AIC and BIC results in Table 5.15 for the Survey_Call indicate that the data fits the MLR approach.

Table 5.16: Goodness-of-fit tests for Current_Sweepings (Y_6)

Test/Criterion	**Statistic**
-2 Res Log Likelihood	124.5
AIC	126.5
BIC	128.9

The AIC and BIC results in Table 5.16 are greater than 120. This implies that the data does not fit the MLR model which is in contrast with the MLE analysis in Table 5.9.

Table 5.17: Goodness-of-fit tests for Mill (Y_7)

Test/Criterion	**Statistic**
-2 Res Log Likelihood	7.1
AIC	9.1
BIC	11.5

The results in Table 5.17 indicate a good data fit. In conclusion Y_1, Y_4, Y_5 and Y_7 show that the data fits the MLR model.

5.4.3 Error/Residual rate analysis

Another test complementing the analyses done from Table 5.11 to Table 5.17 is the error or residual analysis as outlined in Section 4.4.4. The residual analysis conducted was used to decide whether the suggested models by the MLE analyses are modellable under the MLR approach. The following is the threshold used for the error rate analysis:

- $\leq 10\%$ indicates that the MLE analysis is accurate and since the error rate is small it implies that the data fits the MLR modelling framework;
- >10% but smaller than 50% the error rate is still moderate. Although the fit is not ideal since the data sample used is small it can be assumed that the "data fit to the

MLR" is satisfactory; and

- >50% implies that the error rate is big which is an indication of a poor fit.

Observing the previous two analysis discussed above with the residual analysis shown below all KPIs are modellable under the MLR model framework which implies that the correlation measure defined in Chapter 3 can be used to calculate non-linear correlation. The error rate in Table 5.21 and Table 5.24 are less than 10% which implies that the data fit the MLR model. The error rate in Table 5.18, Table 5.19, Table 5.20 and Table 5.23 are greater than 10% but less than 50% which is moderate. This implies that the data fit the MLR model.

Table 5.18 Error rate analysis for Centares

Test/Criterion	**Statistic**
Error rate	12.830%

Table 5.19 Error rate analysis for Onreef_dev

Test/Criterion	**Statistic**
Error rate	27.510%

Table 5.20 Error rate analysis for Offreef_dev

Test/Criterion	**Statistic**
Error rate	23.010%

Table 5.21 Error rate analysis for Channel_Dilution

Test/Criterion	**Statistic**
Error rate	9.165%

Table 5.22 Error rate analysis for Survey_Call

Test/Criterion	**Statistic**
Error rate	17.390%

Table 5.23 Error rate analysis for Current_Sweepings

Test/Criterion	**Statistic**
Error rate	21.980%

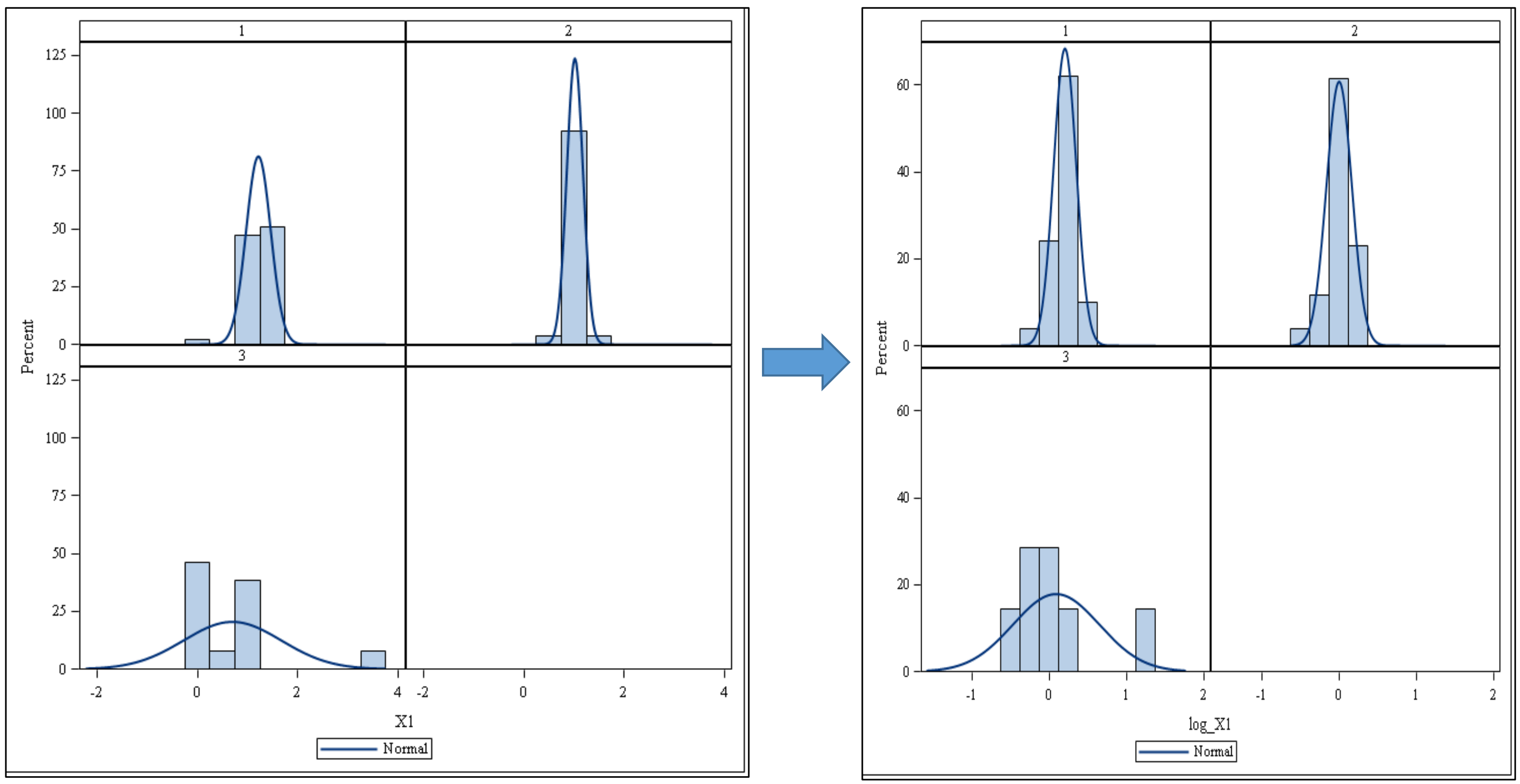
Table 5.24: Error rate analysis for Mill

Test/Criterion	**Statistic**
Error rate	4.972

5.4.4 Normal distribution graphical tests

Equation (4.1) proved the linearity and independent assumptions whereas Figure 5.1 and Figure 5.2 show the normal distribution characteristics. The left side graph shows the data before GLM and the right side shows the data after the logit transformation. The normal Probability Density Function (PDF) fitted to the GLM data on the right fits the mode of the data segmented as per Table 4.3. Figure 5.1 and Figure 5.2 show that the GLM models are compliant to the normal distribution assumption. Figure 5.1 shows that the distribution fit after the applying the log transformation improves. Figure 5.2 shows that the data before log transformation does not rank but after applying the log transformation the distribution fit is improved. This means that the normal distribution assumption holds. Other similar analyses are outlined in Appendix C.

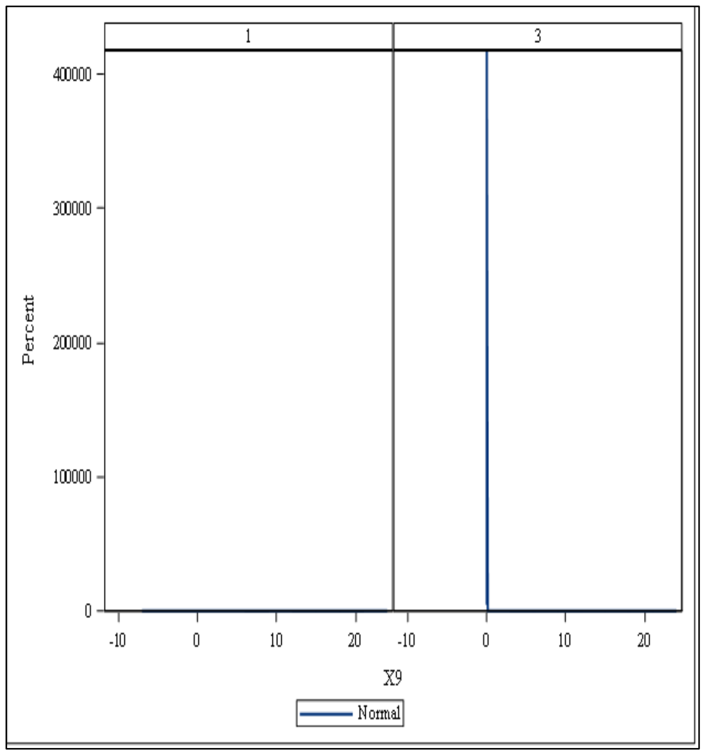
Figure 5.1: Log transormation predictive power analysis 1



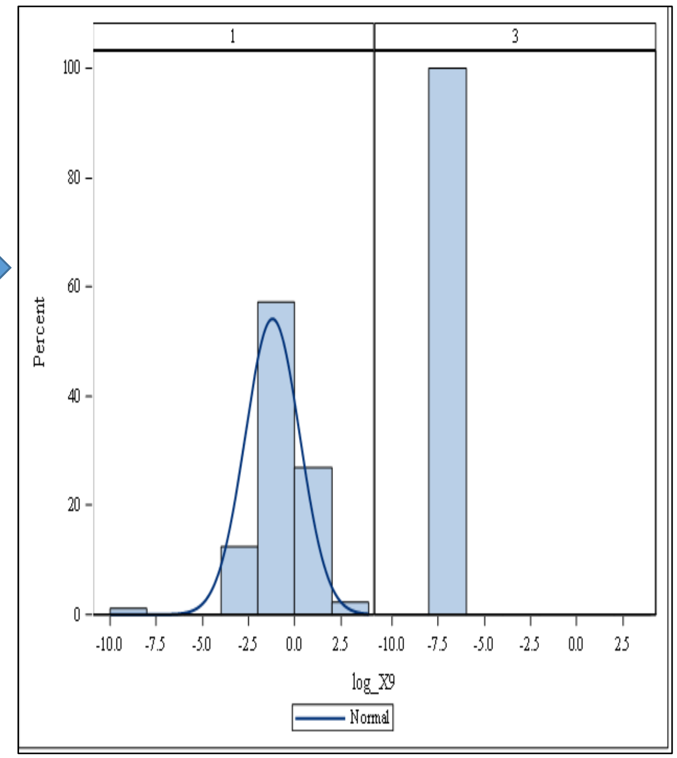
Graph1: Face_Length before log transformation under Centares discrete distribution

Graph2: Face_Length after log transformation under Centares discrete distribution

Figure 5.2: Log transormation predictive power analysis 2



Graph1: Off_reef before log transformation under the Mill discrete distribution



Graph2: Off_reef after log transformation under Mill discrete distribution

5.4.5 Correlation analysis

Previous analyses validate that six of the seven modelled KPIs are modellable under the MLR approach defined in equation (4.3). This MLR approach belongs to the exponential model family which can be re-defined as a GLM model by using the link function (i.e. logit link function, see Table 3.1). Since all assumptions hold, the correlation in this study as discussed in Chapter 3 will be calculated as indicated in Table 5.25.

All statistical significance correlation shown in Table 5.25 are negative correlations except for Offreef (X_4) with Y_1. This implies that any increment in the explanatory variable will lead to a decrease in the "Centares" mined. The correlations are interpreted as follows:

- The correlation between the Centares and Face_Length (X_1) mined is -33% and is statistically significant. This implies that, increasing the face length mined will lead to a reduction of -0.33 in the Centares mined.
- Although the correlation between mineable area and Allow_OB (X_5) is significant, it is close to zero. There is a -55% chance that authorizing more breaks will lead to a decrease in the Centares mined.
- According to this correlation, the higher the weight of the ore extracted the smaller the chance of recovering platinum in the mineable area.
- An increase in the Channel_width by 1 unit will lead to a -0.34 decrease in the mineable area.

The central theme in the correlation shown in Table 5.25 is that, increasing face length, allowable overbreaks, error overbreaks, channel width and weighline lead to minimisation of producing platinum ore in the mineable area. The null hypothesis (i.e. H_0) analysed by Table 5.25 is that, there is correlation between mine production activities and KPI modelled.

Table 5.25: Correlation analysis for Centares (Y_1)

Centares	$\mathbf{X_1}$	$\mathbf{X_2}$	$\mathbf{X_3}$	$\mathbf{X_4}$	$\mathbf{X_5}$	$\mathbf{X_6}$	$\mathbf{X_7}$	$\mathbf{X_8}$	$\mathbf{X_9}$
Y_1	-33%	-15%	-8%	+5%	-55%	-52%	-34%	-57%	-7%
p-value	<10%	>10%	>10%	>10%	<10%	<10%	<10%	<10%	>10%

"<10%" - indicate statistical significance whereas " >10%" - indicate rejection of H_0

Table 5.26 shows that four correlations are statistically significant with this KPI, that is:

- Onreef_dev negatively correlates with Teams (X_3), Error_OB (X_6) and Channel_Width (X_7). This means any increment in these attributes leads to a decrease in the Onreef development metres to open ore reserves for mining.
- There is a positive correlation between the Onreef_dev and Replacement_factor. This implies that any increase in the ore reserves will lead to an increase for the Onreef development metres. The metres required to open the ore reserves for mining can be obtained.

Table 5.26: Correlation analysis for Onreef_dev (Y_2)

Onreef_dev	$\mathbf{X_1}$	$\mathbf{X_2}$	$\mathbf{X_3}$	$\mathbf{X_4}$	$\mathbf{X_5}$	$\mathbf{X_6}$	$\mathbf{X_7}$	$\mathbf{X_8}$	$\mathbf{X_9}$
Y_2	-16%	-6%	-27%	+6%	-1%	-27%	-43%	-10%	+46%
p-value	>10%	>10%	<10%	>10%	>10%	<10%	<10%	>10%	<10%

"<10%" - indicate statistical significance whereas " >10%" - indicate rejection of H_0

Offreef development allows access to the ore reserves. Seven explanatory variables correlate with this KPI as shown in Table 5.27. The correlations are interpreted as follows:

- Offreef_dev negatively correlates with Face_Length (X_1), Teams (X_3), Allow_OB (X_5), Error_OB (X_6), Channel_Width (X_7) and Weighline (X_8). This implies that any 1 unit increase in these attributes will lead to a decrease in the development metres to access the ore reserves.

- This KPI is positively correlated with Replacement_factor (X_9), which implies that any increase in the replacement factor will result in an increment for the Offreef development metres. This means that there is higher chance to access the ore reserves for mining.

Table 5.27: Correlation analysis for Offreef_dev (Y_3)

Offreef_dev	$\mathbf{X}_1$	$\mathbf{X}_2$	$\mathbf{X}_3$	$\mathbf{X}_4$	$\mathbf{X}_5$	$\mathbf{X}_6$	$\mathbf{X}_7$	$\mathbf{X}_8$	$\mathbf{X}_9$
Y_3	-28%	-7%	-30%	+12%	-20%	-34%	-40%	-27%	+47%
p-value	<10%	>10%	<10%	>10%	<10%	<10%	<10%	<10%	<10%

"<10%" - indicate statistical significance whereas " >10%" - indicate rejection of H_0

Table 5.28 shows that 67% of these tests are negative correlations and statistically significant. The correlations are interpreted as follows:

- Channel_dilution negatively correlates with Face_Length, Teams, Allow_OB, Error_OB, Channel_Width and Weighline. This means that any 1 unit increase in these attributes will lead to a decrease in Channel_dilution. In a practical sense this implies that, the availability of different platinum ore grades are mostly impacted by number of teams, allowable overbreaks, weight of the ore mined, the cut-off grades and permissible mining process.

Table 5.28: Correlation analysis for Channel_Dilution (Y_4)

Channel_Dilution	$\mathbf{X}_1$	$\mathbf{X}_2$	$\mathbf{X}_3$	$\mathbf{X}_4$	$\mathbf{X}_5$	$\mathbf{X}_6$	$\mathbf{X}_7$	$\mathbf{X}_8$	$\mathbf{X}_9$
Y_4	-48%	-15%	-43%	-1%	-46%	-43%	-33%	-46%	-8%
p-value	<10%	>10%	<10%	>10%	<10%	<10%	<10%	<10%	>10%

"<10%" - indicate statistical significance whereas " >10%" - indicate rejection of H_0

Table 5.29 shows that seven correlations are statistically significant with Survey_Call, that is

- Survey_Call negatively correlate with Face_Length, Blast, Teams, Allow_OB, Error_OB, Channel_Width and Weighline. Any increment in these attributes will lead to a decrease in survey call. Intuitively it makes sense that, if the face length, blasting, number of teams, Allowable overbreaks, Error overbreaks and weighline are increased, the planned target might be reached quickly. This means that there is no need for survey in recalibrating the mining process to improve output results.

Table 5.29: Correlation analysis for Survey_Call (Y_5)

Survey_Call	$\mathbf{X}_1$	$\mathbf{X}_2$	$\mathbf{X}_3$	$\mathbf{X}_4$	$\mathbf{X}_5$	$\mathbf{X}_6$	$\mathbf{X}_7$	$\mathbf{X}_8$	$\mathbf{X}_9$
Y_5	-35%	-22%	-20%	-2%	-44%	-40%	-36%	-44%	-0.1%
p-value	<10%	<10%	<10%	>10%	<10%	<10%	<10%	<10%	>10%

"<10%" - indicate statistical significance whereas ">10%" - indicate rejection of H_0

All statistical significance correlation shown in Table 5.30 are negative correlations except for Offreef (X_4). This implies that any increment in the explanatory variable will lead to a decrease in the "Sweepings" process. The correlations are interpreted as follows:

- The correlation between the Current_Sweepings and Allow_OB (X_5) is -47% and it is statistically significant. In a practical sense, this implies that, authorizing more overbreaks will lead to more platinum ore being produced in the normal process which suggests that the team might reach their target. Then using this relationship outlined above its intuitively arguable that, if more overbreaks are made in the mining process, chances of reaching target increases. This scenario implies that, the Current_Sweeping process is minimised since there will be no need to collect more platinum from using this approach since target set is met.

- Current_Sweepings negatively correlate with Face_Length (X_1), Error_OB (X_6), Channel_Width (X_7) and Weighline (X_8). This implies that any increase in face length mined, allowable overbreaks, error overbreaks, the tolerance level of the cut off grades permissible and the weight of the mineable ore lead to minimisation of the sweeping process.

This minimisation is motivated by the fact that each team in the mining process is given targets to meet and once the targets are met there is no need to mine further to ensure continuous deliverance on the planned target. The null hypothesis analysed in Table 5.30 is that, there is correlation between the mine production activities and KPI modelled.

Table 5.30: Correlation analysis for Current_Sweepings (Y_6)

Current_Sweepings	$\mathbf{X_1}$	$\mathbf{X_2}$	$\mathbf{X_3}$	$\mathbf{X_4}$	$\mathbf{X_5}$	$\mathbf{X_6}$	$\mathbf{X_7}$	$\mathbf{X_8}$	$\mathbf{X_9}$
Y_6	-27%	-8%	-3%	+53%	-47%	-42%	-22%	-52%	-9%
p-value	<10%	>10%	>10%	>10%	<10%	<10%	<10%	<10%	>10%

"<10%" - indicate statistical significance whereas " >10%" - indicate rejection of H_0

Mill is the quality of the platinum ore extracted in the mining production process through concentration. Table 5.31 shows that two correlations are statistically significant with this KPI, that is:

- The correlation between the Mill grade and Weighline is negative. This implies that the larger the weight of the ore mined there is a smaller chance of producing the more platinum in the mining process.

- The Mill grade negatively correlate with Face_Length (X_1), Blast (X_2), Teams (X_3), Offreef (X_4), Allow_OB (X_5), Error_OB (X_6) and Channel_Width (X_7). Similar to other interpretation above an increase in these attributes leads to a decrease in the Mill KPI since the correlation is negative. This implies that if an increase in the allowable overbreaks, number of teams, blast and face-length will lead to an increase in waste area and circumstantial to all actions discussed previous it will result in a decrease in the Mill grade . All of these occurrences lead to a decrease in the Mill grade since extracted platinum is impacted by the tolerance level of the cut off grades permissible in the production process.

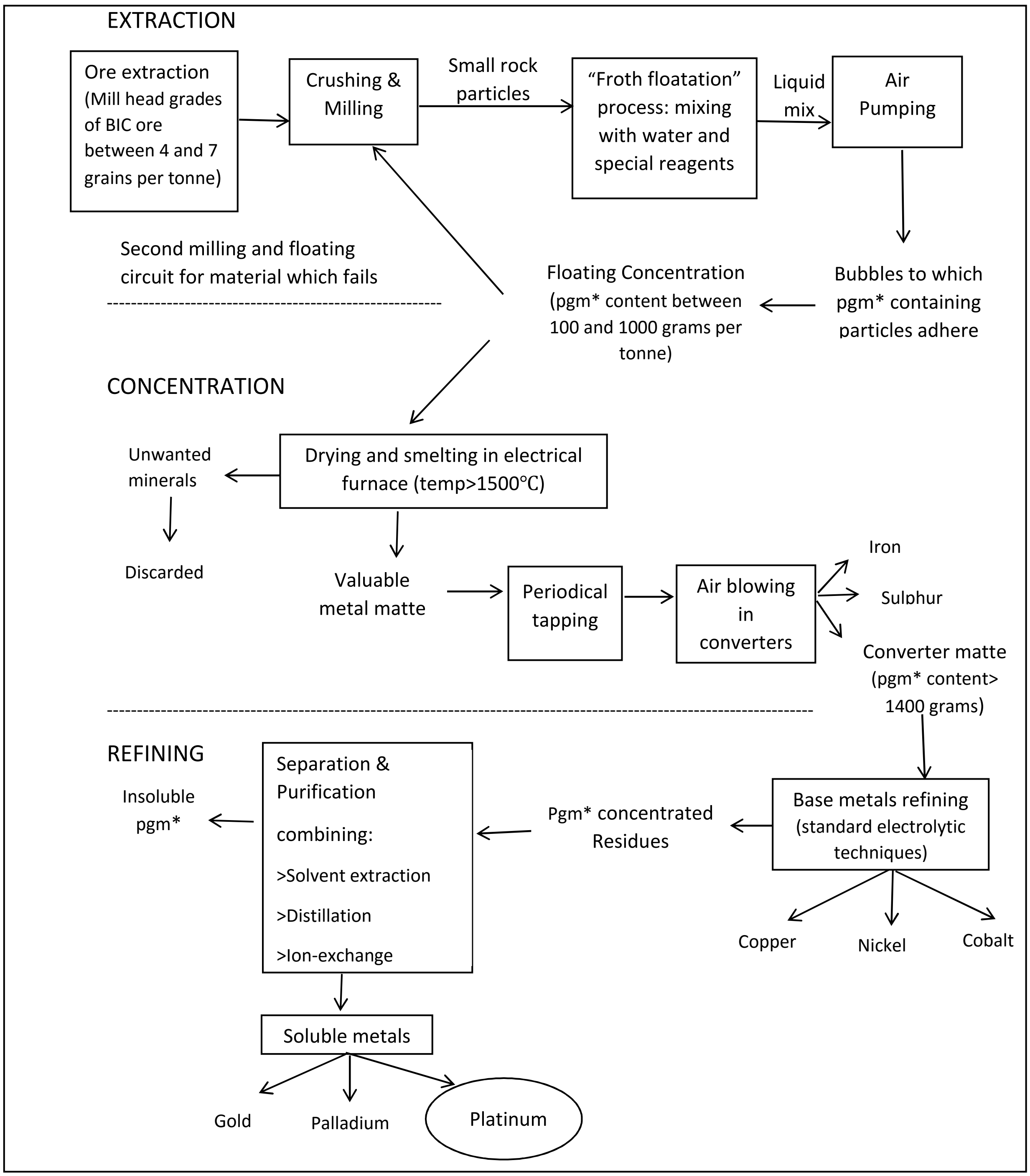
Table 5.31: Correlation analysis for Mill (Y_7)

Mill	$\mathbf{X}_1$	$\mathbf{X}_2$	$\mathbf{X}_3$	$\mathbf{X}_4$	$\mathbf{X}_5$	$\mathbf{X}_6$	$\mathbf{X}_7$	$\mathbf{X}_8$	$\mathbf{X}_9$
Y_7	-44%	-41%	-40%	-32%	-43%	-40%	-40%	-42%	-7%
p-value	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	>10%

"<10%" - indicate statistical significance whereas " >10%" - indicate rejection of H_0

The correlation analysis discussed above outlined that there is a correlation between KPIs and mine production activities. In Appendix C, there is a correlation between the explanatory variable and response variable, refer Section C.1. Then Figure 5.3 outlines the important mine production activities based on Table 5.25 to Table 5.31 in a platinum mine production process. Figure 5.3 also outlines under which section the assessed KPIs fall in a normal platinum mine process, see below.

Figure 5.3: Platinum mine process (Wealth Management Group, 2015)



Based on a generic mine production process as suggested by Figure 5.3 the platinum mine production comprises of three main stages namely, extraction, concentration and refinement. The extraction[3] process is measured by Centares, On_reef_development, Off_reef_development and Survey_Call KPIs studied in this dissertation. This KPIs are negatively correlated to the Face_Length (X_1), Allow_OB (X_5), Weighline (X_7), Channel_Width (X_8) and it is positively correlated to the Replacement factor (X_9). This implies that the extraction process is prolonged if the facelength and the size of the ore extracted is decreased since the mining team needs to ensure that valuable resources are extracted and targets are met. If the facelength is increased this process is shorter since the planned target is easily met provided that valuable ore is extracted. Therefore, the negative correlations outlined above makes intuitive sense. The positive correlation suggests that there is a 10% probability that an increase in reserves implies an increase in the extraction. These inferences can be made since "Centares", "On_reef_development" and "survey_call" are the KPI measuring the extraction process.

The second main process in the mine production process is the concentration process[4]. Current_sweeping and Channel_dilution are the KPIs used to measure the concentration process. These KPIs are negatively correlated with Face_Length (X_1), Allow_OB (X_5), Error_OB (X_6), Weighline (X_7) and Channel_Width (X_8). This implies that if they are used to measure the progress in the concentration process outlined in Figure 5.3, any increase in these variables will lead to a decrease in the concentration of the platinum ore. Therefore, to manipulate the concentration process in a mine production system managing X_1, X_5, X_6, X_7 and X_8 might yield the desired quality of platinum ore according to the correlation analyses in Table 5.23 and Table 5.26.

The final process in the mine production process as shown in Figure 5.3 is the refining[5] pro-

[3]Extraction procedure in the mining process involves drilling, blasting, loading, hauling crushing and milling which leads to the extraction of mineral resources from the earth.

[4]Concentration is the process in the mine production process where the extracted ore is re-processed to extract the mined commodity and discard the waste.

[5]Refining process is the transformation of the mined platinum into the final product sellable in the market.

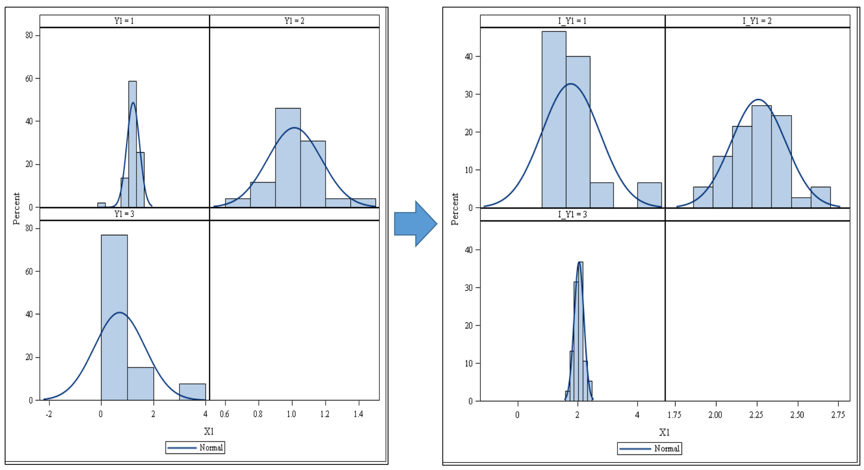
cedure. In this research Mill is the KPI used to assess progress in this process. According to the correlation analyses, the larger the Face_Length(X_1), Blasted area(X_2), Teams(X_3), Waste (X_4), Allow_OB (X_5), Error_OB (X_6), Weight of mined ore (X_7) and Cutoff grades (X_8) implies that there is a smaller chance of producing the high quality platinum grade in the refining process. Therefore minimisation of these attributes in the mine production process can lead to the maximisation of the quality of platinum ore in the mine production process. These correlation analyses provide a snapshot on the behaviour of the measuring agents (i.e. KPIs) used to assess the progress in the mining process. However, this information in the traditional mine planning approach is not modellable.

5.4.6 Hypothetical probability distributions for mining activities

The question driving this research is *"how do stochastic activities in a mining production system interact to produce the resultant KPIs?"* In order to answer this question, probability distributions for mining activities are assumed for the mining production system. The left side graph shows the distribution before the production activity is modified whereas the right side shows after the activity is modified with the view of reducing the uncertainty, see Figure 5.4 to Figure 5.13.

Figure 5.4 shows that the distribution fit for Centares after modifying reduces uncertainty as it improves the fit. This implies that to attain the required Centares the Face_length needs to be maximised in the mining process. The Face_length has an impact on the overall Centares mined.

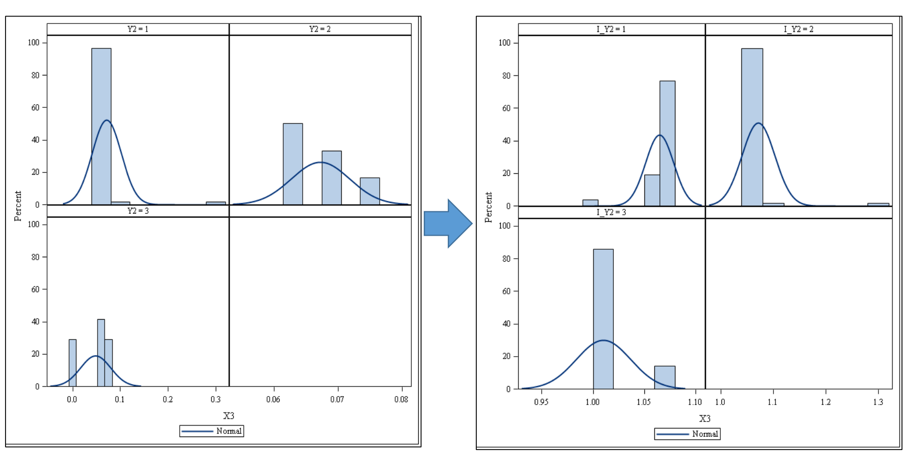
Figure 5.4: Probability distribution analysis for Centares (Y_1)



Graph1: Centares Before **Graph2: Centares after Face length is increased**

Figure 5.5 shows that the distribution fit for the On_reef_development meters after modifying the distribution fit reduces uncertainty. Increasing the number of teams improves the On_reef_development to open the ore reserves. This implies that teams have an impact on the overall On_reef_development metres in the production system.

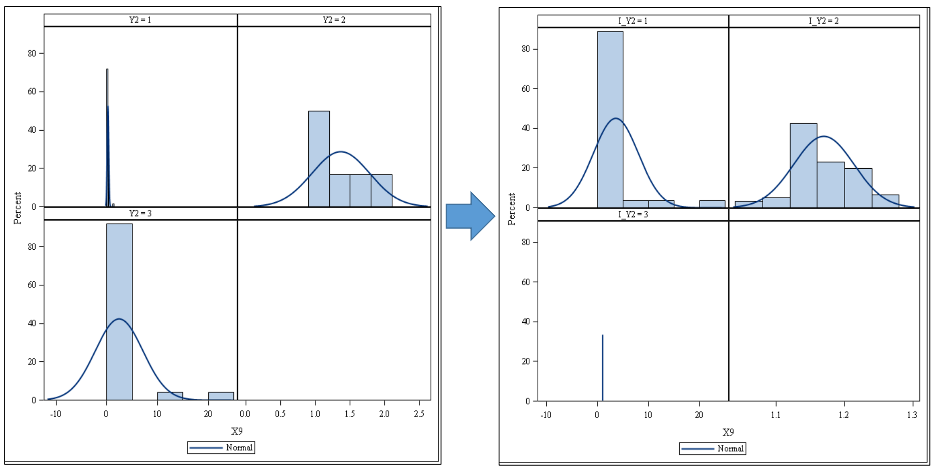
Figure 5.5: Probability distribution analysis1 for On_reef_dev (Y_2)



Graph1: On-reef-development before **Graph2: On-reef-development after modification**

Figure 5.6 shows that modifying the On_reef_dev distribution fit reduces uncertainty by increasing the replacement factor. This means that the both Replacement_factor and On_reef_development metres need to be maximised to extract the ore reserves.

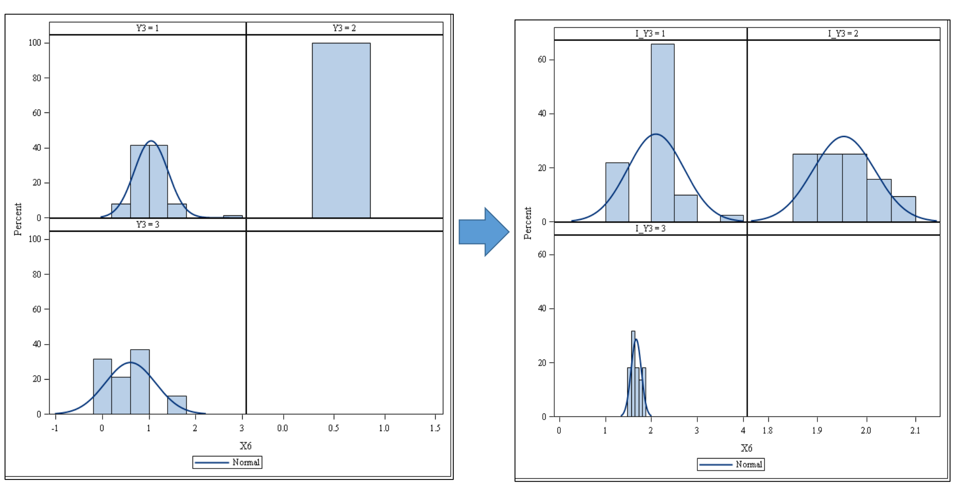
Figure 5.6: Probability distribution analysis2 for On_reef_dev (Y_2)



Graph1: On-reef-dev before **Graph2: On-reef-dev after modification**

Figure 5.7 shows that the distribution fit for On_reef_dev after modification reduces uncertainty. This means that increasing Error_OB in the production system has an impact on the Off_reef_development metres. Error_OB needs to be maximised to access the ore reserves.

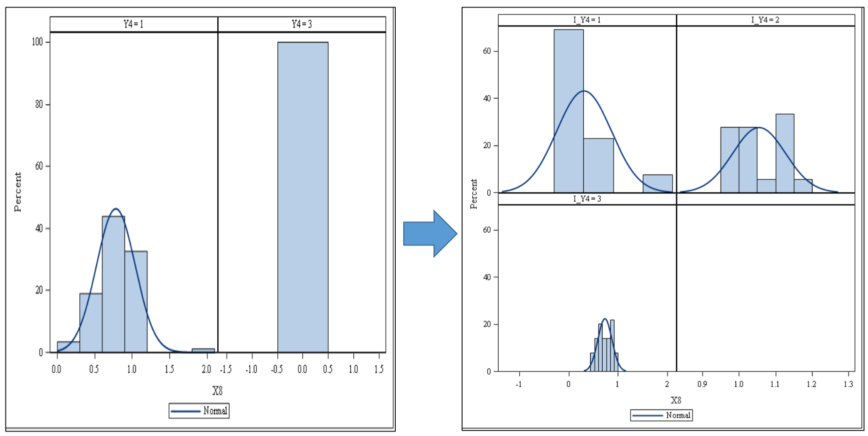
Figure 5.7: Probability distribution analysis for Off_reef_dev (Y_3)



Graph1: Off-reef-dev before **Graph2: Off-reef-dev after modification**

Figure 5.8 shows that the distribution fit for Channel_dilution after modification reduces uncertainty. Increasing the Weighline in the production process improves the distribution fit for Channel_dilution. This implies that the weight of the ore extracted has an impact on the overall quantity used to classify the ore extracted in the production system.

Figure 5.8: Probability distribution analysis for Channel_dilution (Y_4)

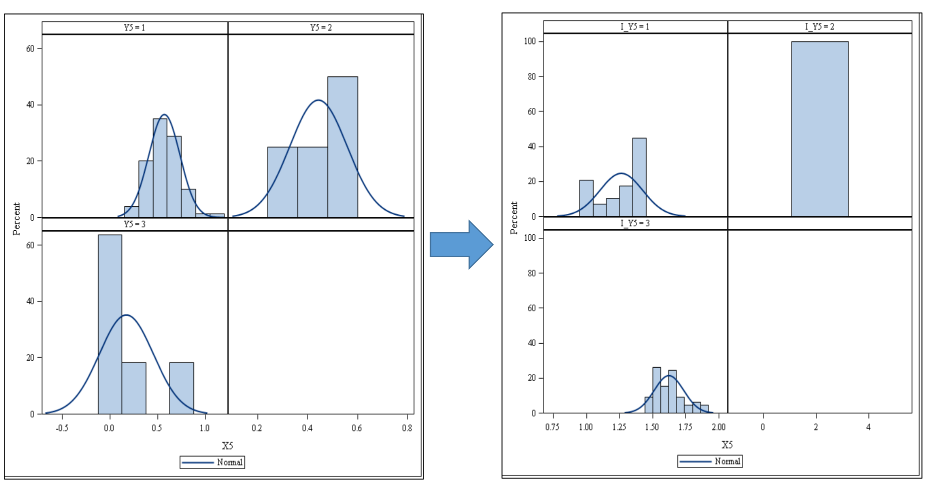


Graph1: Channel_Dilution before

Graph2: Channel_Dilution after modification

Figure 5.9 shows that the distribution fit for Survey_call before modification is a better fit. This means that increasing the allowable overbreaks does not reduce the uncertainty. Allow_OB does not have an impact on survey call in the overall production system.

Figure 5.9: Probability distribution analysis for Survey_call (Y_5)

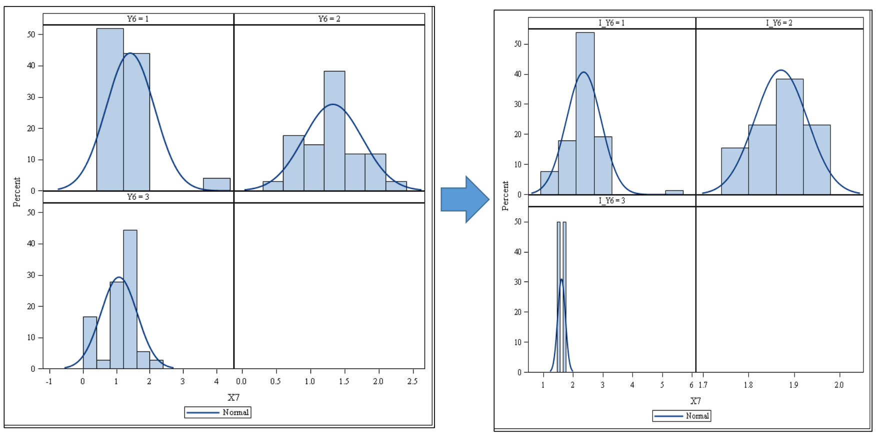


Graph1: Survey_Call before

Graph2: Survey_Call after modification

Figure 5.10 shows that the distribution fit for Current_sweepings before modification is a better fit. This implies that increasing or decreasing the Channel_width does not reduce the uncertainty. Channel_width does not have an influence in the sweepings process.

Figure 5.10: Probability distribution analysis for Current_Sweepings (Y_6)

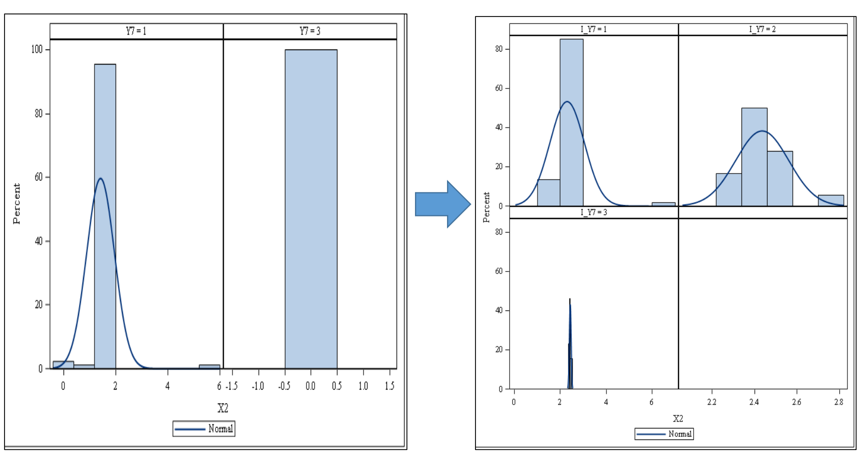


Graph1: Current_Sweepings before

Graph2: Current_Sweepings after modification

Figure 5.11 show the distribution fit for the Mill KPI before and after blasting is modified and increasing blasting reduces uncertainty. This means that the larger the blasting area the smaller the chance of producing high quality platinum grade. This activity needs to be minimised which can lead to the maximisation of extracting quality ore in the production process.

Figure 5.11: Probability distribution analysis1 for Mill (Y_7)

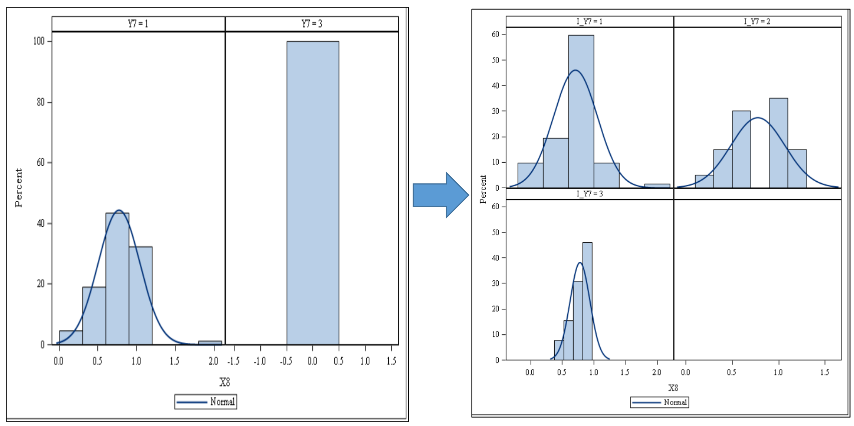


Graph1: Mill before

Graph2: Mill after modification

Figure 5.12 shows that by modifying the distribution for the Mill KPI by increasing Weighline reduces uncertainty. This implies that the weight of the ore extracted needs to be maximised to produce high grade of platinum. Weighline has a big impact on the grade of platinum produced.

Figure 5.12: Probability distribution analysis2 for Mill (Y_7)

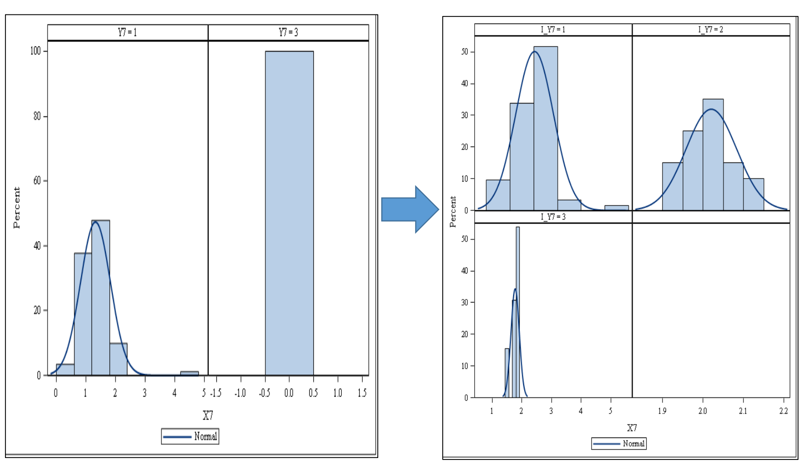


Graph1: Mill before

Graph2: Mill after Weighline is increased

Figure 5.13 shows that the distribution fit for the Mill_grade after modification reduces uncertainty. Increasing the Channel_width in the production process improves the level of grade of platinum produced. This means that the cut-off grade permissible in the mining process has an impact in the overall production system for the Mill_grade.

Figure 5.13: Probability distribution analysis3 for Mill (Y_7)



Graph1: Mill before

Graph2: Mill after modification

5.5 Summary

This chapter presented the model analysis in answering the proposed question. Five out of seven KPIs are highly impacted by the behaviour of the mining activities in the production process. The next chapter presents conclusions and recommendations.

Chapter 6

CONCLUSIONS AND RECOMMENDATIONS

6.1 Chapter overview

This chapter provides concluding remarks on the analyses and modelling conducted in the previous five chapters in quantifying correlation in a mine production operation. Section 6.1 provides preliminary findings whereas Section 6.2 outlines the contributions and limitations made by the research. Section 6.3 concludes the chapter by indicating possible future research on this topic.

6.2 Conclusions

The correlation method developed in this research assumes that the mine production data is modellable under the MLR approach. Due to the MLR model's ability to be transposed into a GLM model by applying natural logarithm function transformation, allows the correlation to be computable under the Pearson correlation method. However few adjustments had to be made since the extracted data in terms of the sample size was small. These adjustments were namely:

- Significance level for all hypothesis testing method was assumed to be 10% because the sample size is less than 100;
- The MLE method and other tests assessing assumptions assumed by the MLR model, their acceptance or tolerance level were increased; and
- The response variables (i.e. γ in Table 5.3) level were reduced from six to three to allow MLR modelling.

The main observations in the correlation method is that, all seven KPIs modelled concluded that data fits the MLR model. Based on this, the correlation can be calculated and the following inferences were drawn:

- Centares mined are affected by the facelength, allowable overbreaks, error overbreaks, weighline and channel width. This implies that increasing these attributes will lead to more Centares being achieved which means that the teams might reach their target. Based on the correlation computed one can find that the ratio between the facelength and allowable overbreaks against the ratio between facelength and grades of the mined ore are crucial in determining the available ore in the mined block (i.e. in the mine design process).
- Correlation analysis further outlined that the quality of the ore depends on the ratio between the size of tonnages and the grades of mined ore.

These are among many interpretations that can be drawn in the correlation analysis in Section 5.5.4. The main finding in this research is that, the ratio of mine production activities against the observed KPIs can be used in setting the restriction functions in the stochastic mine planning approaches. These ratio analysis can be used to forecast the behaviour of KPIs, which is important for appropriate management actions.

This study shows that by applying the MLR approach, it is possible to infer what would be the forecasted KPIs using the correlation analysis. About 71% of the KPIs are greatly influenced by the movements of the production activities in the mining process. This implies that the level of uncertainty on the forecasted KPIs is reduced through applying the MLR model. The following four KPIs are key in the mine production system, namely; Centares mined, tonnage, grade produced and Platinum kilograms produced. These KPIs are calculated as follows;

$$\begin{aligned}
\text{Centares mined}(m^2) &= \text{Face length mined}(m) \times \text{advance per blast}(m) \times \text{number of blast} \\
&\quad \times \text{number of teams} \\
\text{Grade produced}(g/t) &= \text{This is the inherent quality of the ore-body} \\
\text{Tonnage}(t) &= \text{Centares mined}(m^2) \times \text{stoping width}(m) \times \text{specific gravity}(t/m^2) \\
\text{Metal content produced}(Kg) &= \text{Volume of ore mined}(m^3) \times \text{specific gravity}(t/m^2) \times \text{grade}(g/t)
\end{aligned}$$

6.3 Research contributions and limitations

The research contributions are as follows:

- The correlation model developed in this study under the MLR approach should be used as a preliminary analysis tool in the mine planning process.
- Correlation analysis in this study should be used as an indicative measure on deciding thresholds, tolerance levels and risk appetite of the stochastic mine planning algorithm as discussed in previous chapters.
- The Pearson correlation model reviews mine planning in a risk management perspective. The reason for this statement is that mine planning methods in isolation provide information on how to execute the mine process optimally minimising costs but correlation analysis defined in Chapter 3 provides information on the behaviour of factors informing mining which can influence risk management since relationships and underlying patterns are known.

Furthermore, this research assumed that:

- The correlation is computable only under instances where data fits the MLR model. All seven models analysed in this research, fit the MLR modelling assumption as stipulated in the simplistic correlation model.
- The model is functional under a simplified mine production process. This research discounts the existence of dependency between assessed KPIs.

6.4 Recommendations for future research

Following on from discussions arising from the research findings, the possible future area of research should be to extend to correlation modelling when data is not modellable under the MLR modelling process. This correlation model should be embedded into the stochastic mine planning method in a decision-tree based modelling incorporating forecasting, constraint functions setting and operational risk management in the mine's operation.

REFERENCES

BARTLETT, H. E. AND LIEBENBERG, M. J. (2014). Mine to metal: a practical balance for a large platinum producer. *Journal of the Southern African Institute of Mining and Metallurgy*, **114** (1), 103 – 108.

BAYAGA, A. (2010). Multinomial logistic regression: Usage and application in risk analysis. *Journal of Applied Quantitative Methods*, **5** (2), 288–297.

BHATTACHARYA, J. (2007). *Principles of mining planning*. Allied Publishers Private Limited, New Delhi.

BOHNING (1992). Multinomial logistic regression algorithm. *Annals of the Institute of Statistical Mathematics*.

DIMITRAKOPOULOS, R. (2011). Stochastic optimization for strategic mine planning: a decade of developments. *Journal of Mining Science*, **47** (2), 139–150.

DIMITRAKOPOULOS, R., MARTINEZ, L., AND RAMAZAN, S. (2007). A maximum upside/minimum downside approach to the traditional optimization of open pit design. *Journal of Mining Science*, **43** (1), 81–90.

DIMITRAKOPOULOS, R. AND RAMAZAN, S. (2004). A stochastic optimization model for open pit mine planning: Application and risk analysis at a copper deposit. *Transactions of the Institution of Mining and Metallurgy: Mining Technology*, **116** (3), 109–118.

DIMITRAKOPOULOS, R. AND RAMAZAN, S. (2008). Stochastic integer programming for optimizing long term production schedules of open pit mines. *Transactions of the Institution of Mining and Metallurgy: Mining Technology*, **117** (4), 155–160.

Fahrmier, L. and Kaufmann, H. (1985). Consistency and asymptotic normality of the maximum likelihood estimator in generalized linear models. *The Annals of Statistics*, **13** (1), 342–368.

Ferreira, S. J. S. (2012). A trade-off study between down-dip and breast mining on the Merensky reef at Khomanani 2. *The Journal of The Southern African Institute of Mining and Metallurgy*, **112** (4), 257–266.

Grimes, D. A. and Schulz, K. F. (2012). Descriptive studies: what they can and cannot do. *The Lancet*, **359** (9301), 145–149.

Hair, Jr., J. F., Black, W. C., Babin, B. J., and Anderson, R. E. (2010). *Multivariate Data Analysis: A Global Perspective*. 7th edition. Prentice Hall (Pearson), New Jersey.

Hajdasinski, M. M. (1988). Optimization of mine size and mine life in a historical perspective. *Mining Science and Technology*, **7**, 305–310.

Huff, F. A. and Shipp, W. L. (1969). Spatial correlation of storm, monthly and seasonal precipitation. *Journal of Applied Meteorology*, **8**, 542–550.

Journel, A. G. (1983). Nonparametric estimation of spatial distributions. *Mathematical Geology*, **15** (3), 445–468.

King, B. (2011). Optimal mining practice in strategic planning. *Journal of Mining Science*, **47** (2), 247–253.

Kortelev, O. B., Kuznetsov, A. S., Sychev, V. A., and Ivanov, Y. V. (1992). Decision making and monitoring in mining. *Journal of Mining Science*, **28** (3), 286–294.

Kosich, D. (2013). Lack of capital forces junior miners to fight for survival – e&y. Internet.
URL: *http://www.mineweb.com/archive/lack-of-capital-forces-junior-miners-to-fight-for-survival-ey [Accessed 8 October 2014].*

Kutner, M. H., Nachtsheim, C. J., and Li, J. N. N. (2005). *Applied Linear Statistical Models*. 5th edition. McGraw-Hill, Singapore.

Kwok, D. and Roantree, A. M. (2014). Hong Kong regulator sues CITIC over $2bn fx losses in 2008. Internet.
URL: *http://www.mineweb.com/hong-kong-regulator-sues-citic-over-2bn-fx-losses-in-2008 [Accesssed 15 October 2015].*

Magagula, N., .Musingwini, C., and M.M.Ali (2015). Multinomial logistic regression analysis of a stochastic mine production system. *Proceedings of the 23rd International Symposium on Mine Planning and Equipment Selection (MPES). The Southern African Institute of Mining and Metallurgy (SAIMM), Johannesburg*, 331–339.

McCulloch, J. (2009). Counting the cost: Gold mining and occupational disease in contemporary South Africa. *African Affairs*, **108** (431), 221–240.

Mineweb (2008a). Pension fund sues NovaGold over Galore Creek copper-gold project. Internet.
URL: *http://www.mineweb.com/mineweb/view/mineweb/en/page66?oid=62068&sn=Detail [Accessed 17 August 2014].*

Mineweb (2008b). Novagold hit by class action lawsuit over Galore Creek copper-gold project. Internet.
URL: *http://www.mineweb.com/mineweb/view/mineweb/en/page34?oid=59099&sn=Detail [Accessed 17 August 2014].*

Mineweb (2009). Novagold Galore Creek class action suits multiply. Internet.
URL: *http://www.mineweb.com/mineweb/view/mineweb/en/page66?oid=90134&sn=Detail [Accessed 2 September 2014].*

Musingwini, C. (2009). *Techno-economic optimisation of level and raise spacing range in planning a Bushveld Complex platinum reef conventional breast mining layout.* Ph.D. thesis, University of the Witwatersrand.

Musingwini, C. (2010). Techno-economic optimisation of level and raise spacing range in Bushveld Complex platinum reef conventional breast mining. *in The Journal of the Southern African Institute of Mining and Metallurgy,*, **110** (8), 425–436.

Mustafa, K. (2010). Robust stochastic mine production scheduling. *Engineering Optimization*, **42** (6), 567–579.

Myers, R. H. and Montgomery, D. C. (1997). Atumodels on generalized linear models. *Journal of Quality Technology*, **29** (3), 274–291.

Nattrass, N. (1995). The crisis in South African gold mining. *World Development*, **23** (5), 857–868.

Nelsen, J. L. (2007). *Social license to operate : Integration into mine planning and development*. Masters dissertation, The University of British Columbia.

Newby, E., Fomeni, F. D., Ali, M. M., and Musingwini, C. (2012). A method for stochastic estimation of cost and completion time of a mining project. *In Proceedings of the Mathematics in Industry Study Group 2010*. pp. 13–20.

Opoku, S. and Musingwini, C. (2013). Stochastic model of the open pit to underground transition interface for gold mines. *International Journal of Mining, Reclamation and Environment*, **27** (6), 407–424.

PROJECTS IQ (2015). Mining in South Africa. Internet.
URL: *http://www.projectsiq.co.za/mining-in-south-africa.htm [Accessed 20 June 2015].*

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

Rice, J. A. (2007). *Mathematical Statistics and Data Analysis*. Duxbury,, United States of America (USA).

Sabour, S. A. A. and Dimitrakopoulos, R. (2011). Incoporating geological and market uncertainity and operational flexibility into open pit mine design. *Journal of Mining Science*, **47** (2), 191–201.

SAHO, S.A.H.O, S. A. H. O. (2014). 2014 South African platinum strike: long wage strike in South Africa. Internet.

URL: *http://www.sahistory.org.za/article/2014-south-african-platinum-strike-longest-wage-strike-south-africa [Accessed 6 February 2015].*

Schrauzer, G. N., White, D. A., and Schneider, C. J. (1977). Cancer mortality correlation studies-iii: Statistical associations with dietary selenium intakes. *Bioinorganic Chemistry*, **7**, 23–34.

Singh, S. (2014). Effects of the mining strikes on the South African economy. Internet.
URL: *http://www.miningreview.com/effects-of-the-mining-strikes-on-the-south-african-economy [Accessed 5 April 2015].*

Tempelhoff, J. W. N., Ginster, M., Motloung, S., Gouws, C. M., and Strauss, J. S. (2014). The 2012 acid mine drainage (amd) crisis in Carolina's municipal water supply. *African Historical Review*, **46** (2), 77–107.

Wealth Management Group (2015). Precious Metals.
URL: *http://www.wmgmtllc.com/products/hard-assets-metals [Accesssed 8 November 2015].*

Wimberger, M. G. (2004). Breakeven extraction factors for the Merensky reef using stope width and market price simulation. *In International Platinum Conference 'Platinum Adding Value', The South African Institute of Mining and Metallurgy.* pp. 37–46.

Appendix A

Numerical example of the optimisation process

This is a hypothetical situation to illustrate the involvement of calibrating an optimised solution and thought process that a computer undergoes when solving a matrix of unknown quantities. Suppose Mine X in Alaska wants to assess maximum loading capacity of its mining process during the cold season. However, some engineers in this mine have gathered intelligence that maximising $P = xy$ will help to solve the problem and where x is the diameter of the loading shuttle and y is the frequency of the loading. A similar mine in Alaska has studied these patterns recently and discovered that $2x + y^3 = 50088$. This finding is used as a proxy in the optimisation algorithm below:

Optimisation algorithm

Input: $P = xy$ and $2x + y^3 = 50088$

Output: $\widehat{P}$

Solve for x

return $\therefore x = 25044 - y^3/2$

Plug this equation into the P function

return $P = 25044y - y^4/2 \cdots\cdots\cdots (a)$

Differentiate equation (a) and equate it to zero (i.e. solve for y)

return $\therefore 25044 - 2y^3 = 0$

then $y = 23.2216$ and $x = 18782.961$

Substitute x and y into the "$P = xy$" function

return $P = 436170.40$

$\therefore$ Maximum extractable deposit in an Alaskan mine during a cold day is 436170.40 tons.

Optimisation algorithms are quite involved calculations requiring intensive computational power. The hypothetical situation described above shows numerical decisions involved when maximising an element. The hypothetical case solved above involves two unknowns and a single constraint function. However, optimisation algorithms considered in studies such as in Dimitrakopoulos (2011) is more complex in a sense that these methodologies involve multiple unknowns and constraint functions. Therefore, computational steps followed in such optimisation algorithms are more involved and complex to solve. In some instances, the maximised function might not be differentiable which is important in optimising a function by obtaining its minimum or maximum. None differentiability of the optimised function is problematic as it introduces a new dimension into the problem solving algorithm, for example, using diffusion processes to obtain differentiability of this none differentiable function.

The above defined algorithm does not include correlation analysis between x and y. Inclusion of correlation analysis in the algorithm intensifies the modelling process even more and might therefore present some technical difficulties in the execution process, that is, taking longer to compute, etc.

Appendix B

SAS code used

In SAS a library is created to store all data files created under a particular analysis. If a specific library has not been created all work will be saved under the default library but, the problem with this library in a SAS application should not be closed or else all work done is lost. The following syntax was used in creating the library.

```
Libname Nancy 'C: \ Users\ SAS output';
```

A SAS application is divided into two main compartments that is, the "Data step" and "proc step". The "Data step" is used for formatting the imported or created dataset in the SAS environment whereas the "proc step" is primarily used for importing, exporting, analysing and modelling of data. The data extracted is subdivided into planned and actual datasets of variables discussed in Table 5.1 and Table 5.2. Therefore two "proc IMPORT" procedures were run in extracting the Excel files which were in .csv format (i.e. comma separated value format or comma delimited format). Comma delimited format means the entries per line are separated using commas. See the code below:

```
proc import datafile="C:\ Users\ SAS output\ Raw_Data 2.xlsx"
out=Nancy.Actual
dbms=xlsx
```

```
replace;
getnames=yes;
run;

proc import datafile="C:\ Users\ SAS output\ Planned Data.xlsx"
out=Nancy.Planned
dbms=xlsx
replace;
getnames=yes;
run;
```

The original data file names were lengthy and this sometimes presented problems in the modelling since some of the sections are hard-coded in the modelling process. Therefore shorter name conversions were adopted. The following is the SAS code used in changing the naming conversions of the data files as well merging the two datasets imported in the above steps. As per Table 5.3 in Chapter 5, the ratio of planned versus actual was discretised into three levels. The motivation on this is that for the data extracted, the sample size is 96 which, is small for a multivariate approach. The "data step" below is used to reformat the levels as re-defined in Table 5.3 before applying the "proc MIXED" algorithm.

```
data data Nancy.Mine(keep= X1 X2 X3 X4 X5 X6 X7 X8 X9 Y1 Y2 Y3 Y4 Y5 Y6 Y7
log_X1 log_X2 log_X3 log_X4 log_X5 log_X6 log_X7 log_X8 log_X9);

merge Nancy.Planned1 Nancy.p2;

X1=1-((Face_Length1- Face_Length)/Face_Length1);
X2=1-((Blast1-Blast)/Blast1);
X3=1-((Teams1-Teams)/Teams1);
X4=1-((Offreef1 - Offreef)/Offreef1);
```

```
X5=1-((Allow_OB1 - Allow_OB)/Allow_OB1);
X6=1-((Error_OB1 - Error_OB)/Error_OB1);
X7=1-((Weighline1 - Weighline)/Weighline1);
X8=1-((Channel_Width1 - Channel_Width)/Channel_Width1);
X9=1-((Replacement_factor1 - Replacement_factor)/Replacement_factor1);
Y1=1-((Centares1 - Centares)/Centares1);
Y2=1-((On_reef_dev1 - On_reef_dev)/ On_reef_dev1);
Y3=1-((Off_reef_dev1 - Off_reef_dev)/Off_reef_dev1);
Y4=1-((Channel_Dilution1 - Channel_Dilution)/Channel_Dilution1 );
Y5=1-((Survey_Call1 - Survey_Call)/Survey_Call1);
Y6=1-((Current_Sweepings1 - Current_Sweepings)/Current_Sweepings1);
Y7=1-((Mill1 - Mill)/Mill1);
log_X1=log(X1);
log_X2=log(X2);
log_X3=log(X3);
log_X4=log(X4);
log_X5=log(X5);
log_X6=log(X6);
log_X7=log(X7);
log_X8=log(X8);
log_X9=log(X9);
if 0 <Y1 = <0.50 then Y1=3;
else if 0.50 <Y1 = <0.75 then Y1=2;
else Y1=1;
if 0 <Y2 = <0.50 then Y2=3;
else if 0.50 <Y2 = <0.75 then Y2=2;
else Y2=1;
if 0 <Y3 = <0.50 then Y3=3;
else if 0.35 <Y3 = <0.75 then Y3=2;
else Y3=1;
```

```
if 0 <Y4 = <0.50 then Y4=3;
else if 0.50 <Y4 = <0.75 then Y4=2;
else Y4=1;
if 0 <Y5 = <0.50 then Y5=3;
else if 0.50 <Y5 = <0.75 then Y5=2;
else Y5=1;
if 0 <Y6 = <0.50 then Y6=3;
else if 0.50 <Y6 = <0.75 then Y6=2;
else Y6=1;
if 0 <Y7 = <0.50 then Y7=3;
else if 0.50 <Y7 = <0.75 then Y7=2;
else Y7=1;
run;

quit;
```

The "proc MIXED" approach as per Chapter 4's discussion have been deemed accurately to model the MLR scenario defined in Chapter 3, see the SAS codes below.

```
ODS _ALL_ CLOSE;
ODS RTF file="C:\ Users\ SAS output\ results.doc";

proc mixeddata=Nancy.Mine;
class Y1;
model Y1= X1 X2 X3 X4 X5 X6 X7 X8 X9/s chisq;
run;
quit;

proc mixeddata=Nancy.Mine;
class Y2;
```

```
model Y2= X1 X2 X3 X4 X5 X6 X7 X8 X9/s chisq;
run;
quit;

proc mixeddata=Nancy.Mine;
class Y3;
model Y3= X1 X2 X3 X4 X5 X6 X7 X8 X9/s chisq;
run;
quit;

proc mixeddata=Nancy.Mine;
class Y4;
model Y4= X1 X2 X3 X4 X5 X6 X7 X8 X9/s chisq;
run;
quit;

proc mixeddata=Nancy.Mine;
class Y5;
model Y5= X1 X2 X3 X4 X5 X6 X7 X8 X9/s chisq;
run;
quit;

proc mixeddata=Nancy.Mine;
class Y6;
model Y6= X1 X2 X3 X4 X5 X6 X7 X8 X9/s chisq;
run;
quit;

proc mixeddata=Nancy.Mine;
```

```
class Y7;
model Y7= X1 X2 X3 X4 X5 X6 X7 X8 X9/s chisq;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y1 / novarname;
histogram X1;
density X1;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y1 / novarname;
histogram log_X1;
density log_X1;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y2 / novarname;
histogram X1;
density X1;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y2 / novarname;
histogram log_X1;
```

```
density log_X1;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y3 / novarname;
histogram X1;
density X1;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y3 / novarname;
histogram log_X1;
density log_X1;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y4 / novarname;
histogram X1;
density X1;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y4 / novarname;
histogram log_X1;
density log_X1;
```

```
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y5 / novarname;
histogram X1;
density X1;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y5 / novarname;
histogram log_X1;
density log_X1;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y6 / novarname;
histogram X1;
density X1;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y6 / novarname;
histogram log_X1;
density log_X1;
run;
```

```
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y7 / novarname;
histogram X1;
density X1;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y2 / novarname;
histogram log_X1;
density log_X1;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y1 / novarname;
histogram X2;
density X2;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y1 / novarname;
histogram log_X2;
density log_X2;
run;
quit;
```

```
proc sgpanel Data=Nancy.Mine;
panelby Y2 / novarname;
histogram X2;
density X2;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y2 / novarname;
histogram log_X2;
density log_X2;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y3 / novarname;
histogram X2;
density X2;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y3 / novarname;
histogram log_X2;
density log_X2;
run;
quit;
```

```
proc sgpanel Data=Nancy.Mine;
panelby Y4 / novarname;
histogram X2;
density X2;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y4 / novarname;
histogram log_X2;
density log_X2;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y5 / novarname;
histogram X2;
density X2;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y5 / novarname;
histogram log_X2;
density log_X2;
run;
quit;

proc sgpanel Data=Nancy.Mine;
```

```
panelby Y6 / novarname;
histogram X2;
density X2;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y6 / novarname;
histogram log_X2;
density log_X2;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y7 / novarname;
histogram X2;
density X2;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y7 / novarname;
histogram log_X2;
density log_X2;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y1 / novarname;
```

```
histogram X3;
density X3;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y1 / novarname;
histogram log_X3;
density log_X3;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y2 / novarname;
histogram X3;
density X3;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y2 / novarname;
histogram log_X3;
density log_X3;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y3 / novarname;
histogram X3;
```

```
density X3;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y3 / novarname;
histogram log_X3;
density log_X3;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y4 / novarname;
histogram X3;
density X3;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y4 / novarname;
histogram log_X3;
density log_X3;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y5 / novarname;
histogram X3;
density X3;
```

```
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y5 / novarname;
histogram log_X3;
density log_X3;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y6 / novarname;
histogram X3;
density X3;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y6 / novarname;
histogram log_X3;
density log_X3;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y7 / novarname;
histogram X3;
density X3;
run;
```

```
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y7 / novarname;
histogram log_X3;
density log_X3;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y1 / novarname;
histogram X4;
density X4;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y1 / novarname;
histogram log_X4;
density log_X4;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y2 / novarname;
histogram X4;
density X4;
run;
quit;
```

```
proc sgpanel Data=Nancy.Mine;
panelby Y2 / novarname;
histogram log_X4;
density log_X4;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y3 / novarname;
histogram X4;
density X4;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y3 / novarname;
histogram log_X4;
density log_X4;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y4 / novarname;
histogram X4;
density X4;
run;
quit;
```

```
proc sgpanel Data=Nancy.Mine;
panelby Y4 / novarname;
histogram log_X4;
density log_X4;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y5 / novarname;
histogram X4;
density X4;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y5 / novarname;
histogram log_X4;
density log_X4;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y6 / novarname;
histogram X4;
density X4;
run;
quit;

proc sgpanel Data=Nancy.Mine;
```

```
panelby Y6 / novarname;
histogram log_X4;
density log_X4;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y7 / novarname;
histogram X4;
density X4;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y7 / novarname;
histogram log_X4;
density log_X4;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y1 / novarname;
histogram X5;
density X5;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y1 / novarname;
```

```
histogram log_X5;
density log_X5;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y2 / novarname;
histogram X5;
density X5;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y2 / novarname;
histogram log_X5;
density log_X5;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y3 / novarname;
histogram X5;
density X5;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y3 / novarname;
histogram log_X5;
```

```
density log_X5;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y4 / novarname;
histogram X5;
density X5;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y4 / novarname;
histogram log_X5;
density log_X5;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y5 / novarname;
histogram X5;
density X5;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y5 / novarname;
histogram log_X5;
density log_X5;
```

```
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y6 / novarname;
histogram X5;
density X5;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y6 / novarname;
histogram log_X5;
density log_X5;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y7 / novarname;
histogram X5;
density X5;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y7 / novarname;
histogram log_X5;
density log_X5;
run;
```

```
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y1 / novarname;
histogram X6;
density X6;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y1 / novarname;
histogram log_X6;
density log_X6;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y2 / novarname;
histogram X6;
density X6;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y2 / novarname;
histogram log_X6;
density log_X6;
run;
quit;
```

```
proc sgpanel Data=Nancy.Mine;
panelby Y3 / novarname;
histogram X6;
density X6;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y3 / novarname;
histogram log_X6;
density log_X6;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y4 / novarname;
histogram X6;
density X6;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y4 / novarname;
histogram log_X6;
density log_X6;
run;
quit;
```

```
proc sgpanel Data=Nancy.Mine;
panelby Y5 / novarname;
histogram X6;
density X6;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y5 / novarname;
histogram log_X6;
density log_X6;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y6 / novarname;
histogram X6;
density X6;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y6 / novarname;
histogram log_X6;
density log_X6;
run;
quit;

proc sgpanel Data=Nancy.Mine;
```

```
panelby Y7 / novarname;
histogram X6;
density X6;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y7 / novarname;
histogram log_X6;
density log_X6;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y1 / novarname;
histogram X7;
density X7;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y1 / novarname;
histogram log_X7;
density log_X7;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y2 / novarname;
```

```
histogram X7;
density X7;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y2 / novarname;
histogram log_X7;
density log_X7;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y3 / novarname;
histogram X7;
density X7;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y3 / novarname;
histogram log_X7;
density log_X7;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y4 / novarname;
histogram X7;
```

```
density X7;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y4 / novarname;
histogram log_X7;
density log_X7;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y5 / novarname;
histogram X7;
density X7;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y5 / novarname;
histogram log_X7;
density log_X7;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y6 / novarname;
histogram X7;
density X7;
```

```
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y6 / novarname;
histogram log_X7;
density log_X7;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y7 / novarname;
histogram X7;
density X7;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y7 / novarname;
histogram log_X7;
density log_X7;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y1 / novarname;
histogram X8;
density X8;
run;
```

```
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y1 / novarname;
histogram log_X8;
density log_X8;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y2 / novarname;
histogram X8;
density X8;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y2 / novarname;
histogram log_X8;
density log_X8;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y3 / novarname;
histogram X8;
density X8;
run;
quit;
```

```
proc sgpanel Data=Nancy.Mine;
panelby Y3 / novarname;
histogram log_X8;
density log_X8;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y4 / novarname;
histogram X8;
density X8;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y4 / novarname;
histogram log_X8;
density log_X8;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y5 / novarname;
histogram X8;
density X8;
run;
quit;
```

```
proc sgpanel Data=Nancy.Mine;
panelby Y5 / novarname;
histogram log_X8;
density log_X8;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y6 / novarname;
histogram X8;
density X8;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y6 / novarname;
histogram log_X8;
density log_X8;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y7 / novarname;
histogram X8;
density X8;
run;
quit;

proc sgpanel Data=Nancy.Mine;
```

```
panelby Y7 / novarname;
histogram log_X8;
density log_X8;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y1 / novarname;
histogram X9;
density X9;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y1 / novarname;
histogram log_X9;
density log_X9;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y2 / novarname;
histogram X9;
density X9;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y2 / novarname;
```

```
histogram log_X9;
density log_X9;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y3 / novarname;
histogram X9;
density X9;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y3 / novarname;
histogram log_X9;
density log_X9;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y4 / novarname;
histogram X9;
density X9;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y4 / novarname;
histogram log_X9;
```

```
density log_X9;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y5 / novarname;
histogram X9;
density X9;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y5 / novarname;
histogram log_X9;
density log_X9;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y6 / novarname;
histogram X9;
density X9;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y6 / novarname;
histogram log_X9;
density log_X9;
```

```
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y7 / novarname;
histogram X9;
density X9;
run;
quit;

proc sgpanel Data=Nancy.Mine;
panelby Y7 / novarname;
histogram log_X9;
density log_X9;
run;
quit;

ODS RTF off;
listing ;
```

Appendix C

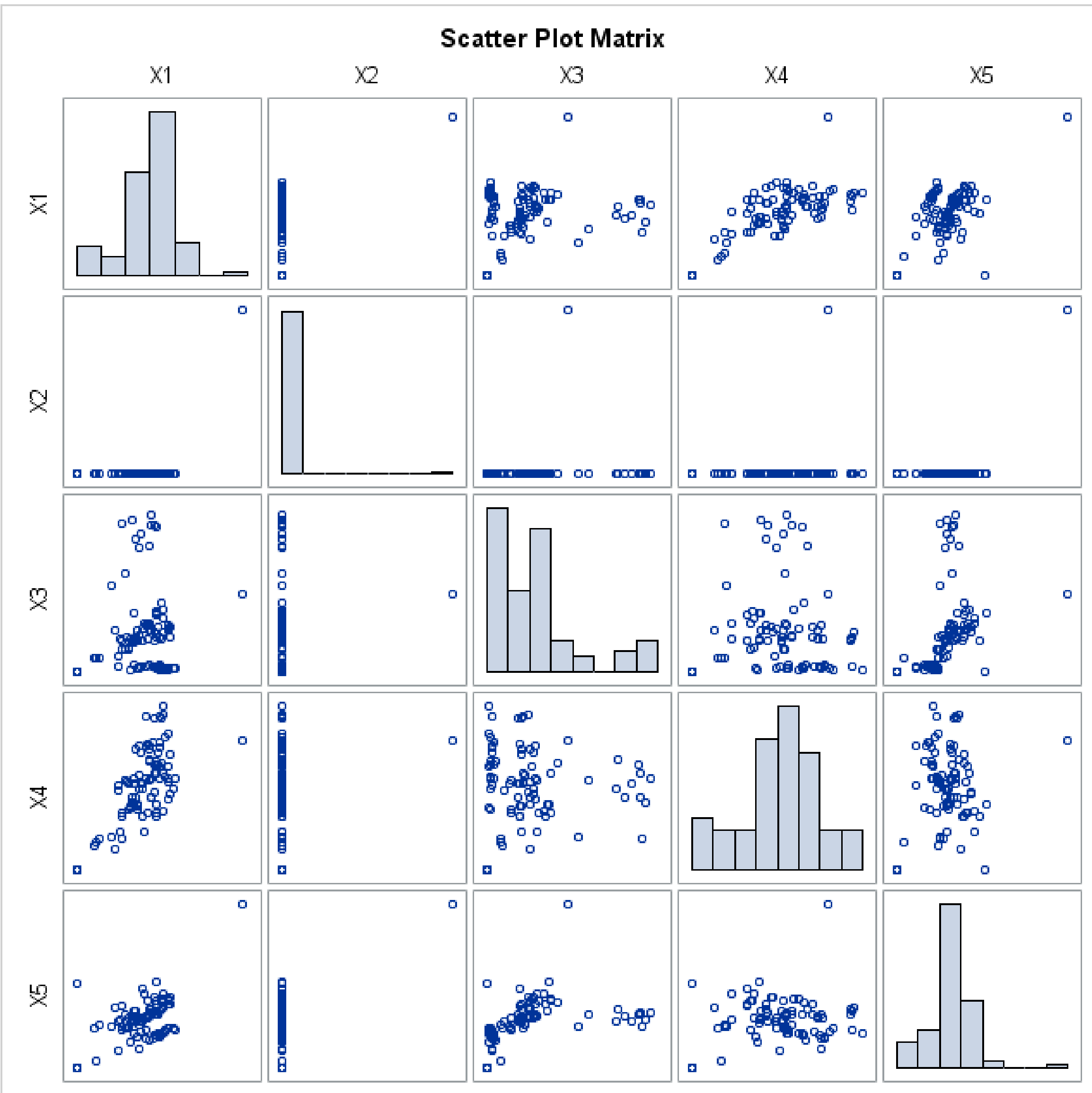
Detailed results not shown in the report

This appendix shows all results not shown in the main report.

C.1 Correlation analysis

The correlation analysis between the explanatory variables assessed suggests that some of the variables are linearly correlated; see Figure C.1 (i.e. X_1 and X_4, X_3 and X_5). That being said, X_2 is non-linear as shown by the scatter plots and this is a similar relationship that would have been outlined if the response variables against the explanatory variable were graphed as shown below. Section 4.4, referred to this analysis.

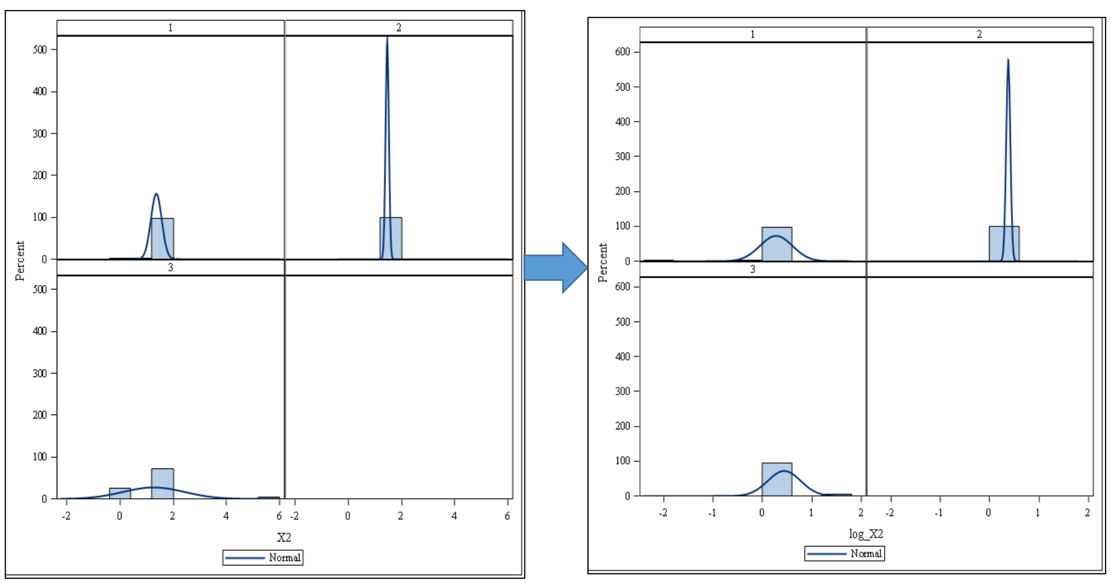
Figure C.1: Correlation analysis of the explanatory variable



C.2 Log-transformation predictive power graphical test

The "proc MIXED" approach implemented in Chapter 4 suggests that the MLR approach is nothing but a generalised linear model under natural log function transformation. This formulation continues and suggests that the generalised linear model follows a normal distribution and these graphical tests assess the ability of the natural log function's transformation in improving normal distribution characteristics in the assessed variables.

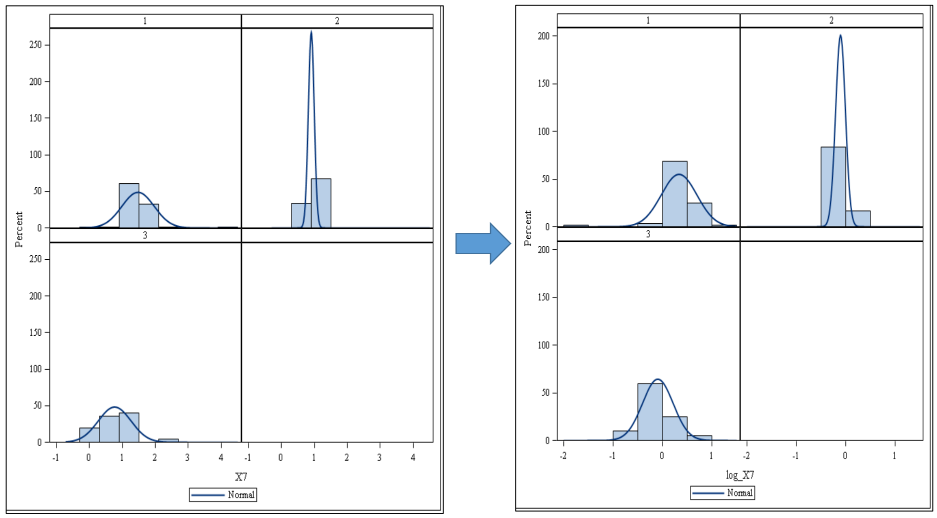
Figure C.2: Log transformation predictive power analysis 6



Graph1: Blast before log transformation under On_reef_dev discrete distribution

Graph2: Blast after log transformation under On_reef_dev discrete distribution

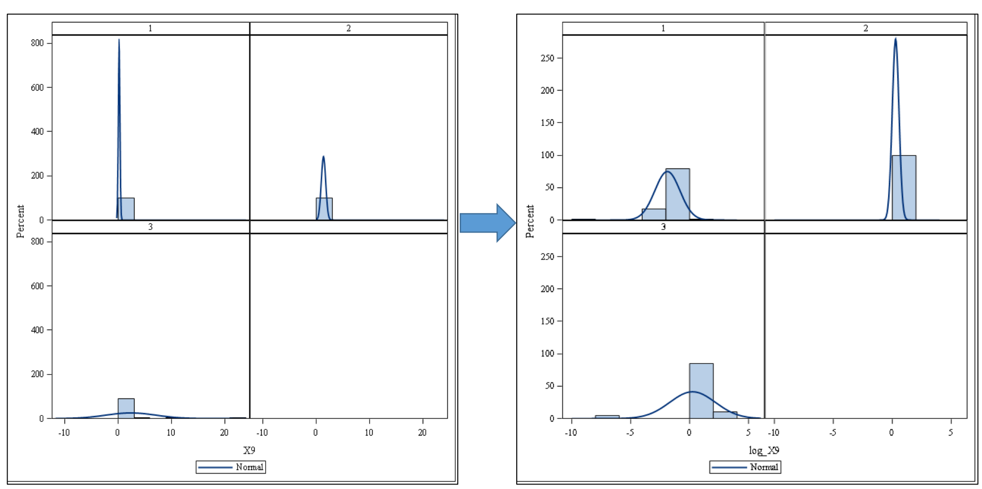
Figure C.3: Log transformation predictive power analysis 7



Graph1: Channel_Width before log transformation under On_reef_dev discrete distribution

Graph2: Channel_Width after log transformation under On_reef_dev discrete distribution

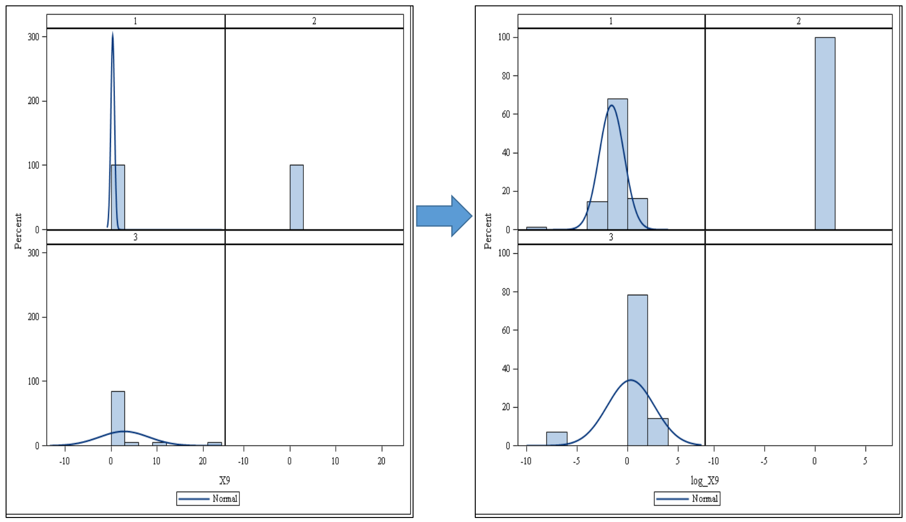
Figure C.4: Log transformation predictive power analysis 8



Graph1: Replacement_factor before log transformation under On_reef_dev distribution

Graph2: Replacement_factor after log transformation under On_reef_dev distribution

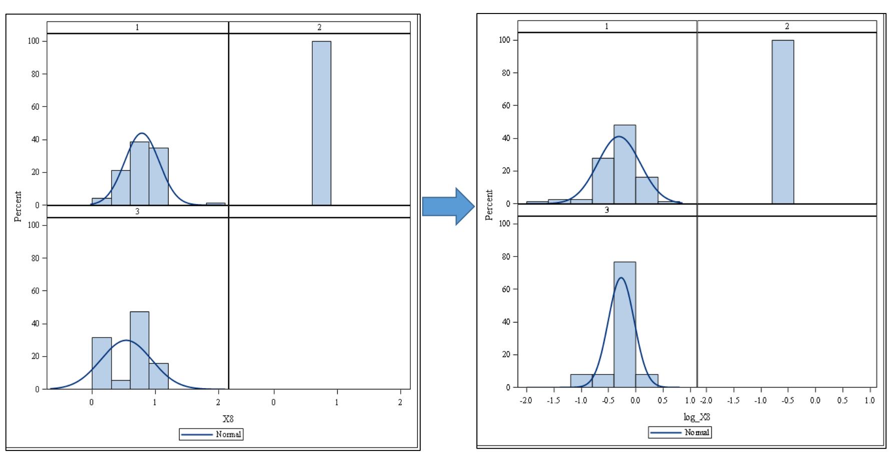
Figure C.5: Log transformation predictive power analysis 9



Graph1: Replacement_factor before log transformation under Off_reef_dev

Graph2: Replacement_factor after log transformation under Off_reef_dev

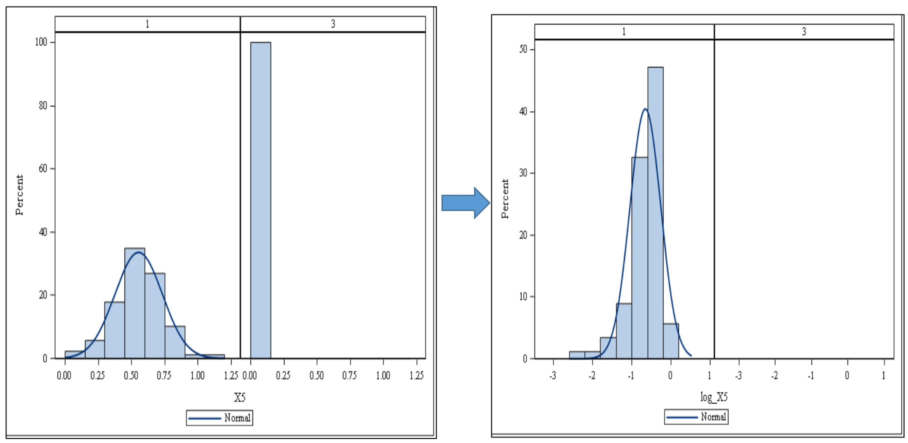
Figure C.6: Log transformation predictive power analysis 10



Graph1: Weighline before log transformation under Off_reef_dev

Graph2: Weighline after log transformation under Off_reef_dev

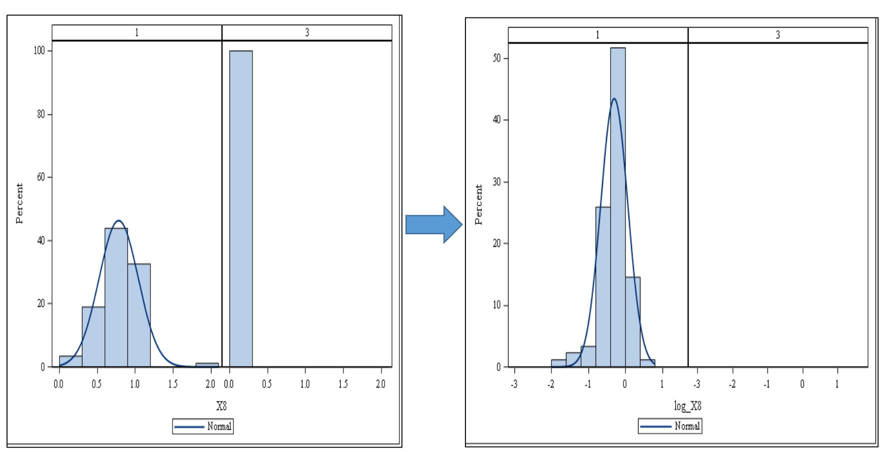
Figure C.7: Log transformation predictive power analysis 11



Graph1: Allow_OB before log transformation under Channel_Dilution

Graph2: Allow_OB after log transformation under Channel_Dilution

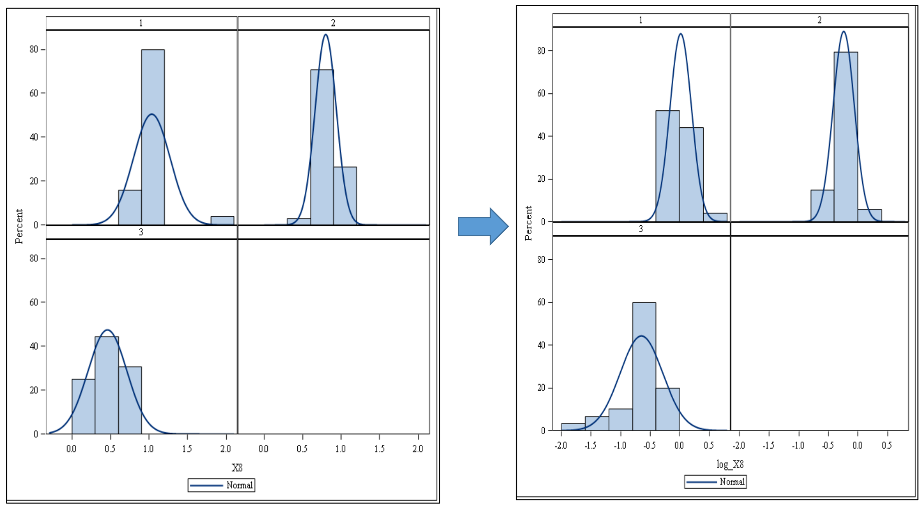
Figure C.8: Log transformation predictive power analysis 12



Graph1: Weighline before log transformation under Channel_Dilution

Graph2: Weighline after log transformation under Channel_Dilution

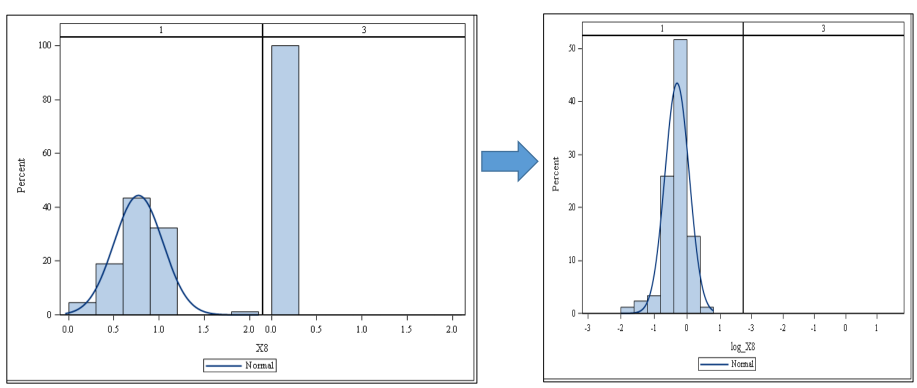
Figure C.9: Log transformation predictive power analysis 13



Graph1: Weighline before log transformation under Current_Sweepings

Graph2: Weighline after log transformation under Current_Sweepings

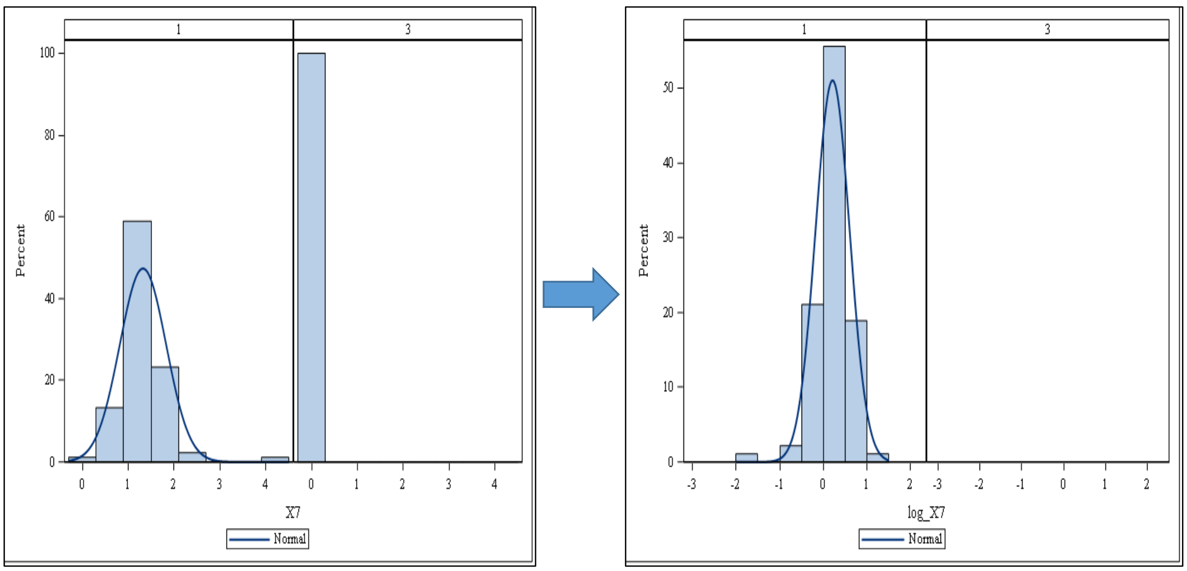
Figure C.10: Log transformation predictive power analysis 14



Graph1: Weighline before log transformation under Mill discrete distribution

Graph2: Weighline after log transformation under Mill discrete distribution

Figure C.11: Log transformation predictive power analysis 15



Graph1: Channel_Width before log transformation under Mill discrete distribution

Graph2: Channel_Width after log transformation under Mill discrete distribution

Appendix D

Data used in modelling

Due to the proprietary nature of the data used, the source of the information is not revealed. The data is subdivided into two sections, that is, the "planned" and "actual". Using equation (3.1) the two sub-datasets are combined into a workable dataset used in this research.

Table D.1: Planned dataset I

No	Face_L	Blast	Teams	Offreef	Allow_OB	Error_OB	Channel_W	Weighline	Replace_factor
1	2822	1352	98	6.28	9988	22506	92851	170841	26.66
2	2823	1352	98	6.24	10094	22433	92907	167127	27.00
3	2765	1352	96	6.25	9868	22404	92399	178269	26.38
4	2765	1352	96	6.20	9827	22380	92195	167127	26.4

Continues next page...

Table D.1 – *Continued from previous page*

No	Face_L	Blast	Teams	Offreef	Allow_OB	Error_OB	Channel_W	Weighline	Replace_factor
5	2794	1339	97	6.25	9446	22369	92644	178269	26.1
6	2765	1352	96	6.14	8746	23006	97839	174555	27.25
7	2448	1173	85	6.08	8854	22924	97804	128131	28.79
8	2506	1201	87	6.01	8854	22891 7 98397	185697	29.42	
9	2477	1187	86	5.94	8732	22873	98457	167127	29.51
10	2794	1339	97	5.91	8863	22915	98253	168984	30.07
11	2794	1339	97	5.96	8785	22618	97064	155985	30.72
12	2794	1339	97	6.01	9438	22497	94870	159699	28.77
13	2822	1352	98	6.01	10981	23738	107689	184891	32.01
14	2909	1394	101	6.00	10934	23720	108197	183285	31.23
15	2909	1394	101	5.98	10908	23784	108713	189680	30.94
16	2938	1408	102	6.00	10536	23798	108594	181963	30.01
17	2938	1408	102	6.03	10552	23867	109130	191054	30.39
18	2938	1408	102	6.03	10573	23764	108945	172562	31.58
19	2938	1408	102	6.03	10558	23823	109244	126116	31.96
20	2880	1380	100	6.03	10508	23773	108963	189247	32.26
21	2851	1366	99	6.03	10467	23820	109084	178006	31.03
22	2851	1366	99	6.06	10460	23909	109245	171842	29.49

Continues next page...

Table D.1 – *Continued from previous page*

No	Face_L	Blast	Teams	Offreef	Allow_OB	Error_OB	Channel_W	Weighline	Replace_factor
23	2880	1380	100	6.06	10460	23909	109245	165847	29.49
24	2909	1394	101	6.07	10439	23892	109244	165485	30.87
25	2822	1352	98	9.27	14181	14802	95161	175626	27.41
26	2736	1311	95	9.24	14166	14814	95820	135889	27.45
27	2736	1311	95	9.27	14106	14784	96228	85054	28.19
28	2678	1283	93	9.29	14047	14798	96616	159980	29.84
29	2678	1283	93	9.30	13778	14774	97436	165859	29.09
30	2678	1283	93	9.28	13331	14706	98060	164011	31.63
31	2880	1380	100	9.27	13731	14796	98310	102880	31.97
32	2851	1366	99	9.31	13776	15087	101033	168597	33.18
33	2851	1366	99	9.32	13323	15276	102913	160315	35.9
34	2851	1366	99	9.32	13329	15068	102063	150152	35.32
35	2822	1352	98	9.34	13318	14923	102103	152082	36.38
36	2794	1339	97	9.36	13347	14753	101885	156909	36.64
37	2822	1352	98	10.03	13185	15355	97237	174878	24.55
38	2938	1408	102	10.02	13192	15365	97342	173808	25.44
39	2966	1421	103	10.02	13156	15340	97371	177123	25.57
40	2938	1408	102	10.02	13205	15312	97108	163884	26.25

Continues next page...

Table D.1 – *Continued from previous page*

No	Face_L	Blast	Teams	Offreef	Allow_OB	Error_OB	Channel_W	Weighline	Replace_factor
41	2938	1421	102	10.06	13165	15311 7 97018	174369	28.47	
42	2966	1408	103	10.05	13163	15188	97086	161954	30.76
43	2938	1421	102	10.03	13003	15468	98882	119909	33.8
44	2938	1408	102	10.03	12841	15541	99456	175415	34.96
45	2966	1408	103	10.02	12830	15494	99527	152415	36.13
46	3082	1421	107	10.01	12769	15403	99887	172865	36.98
47	3197	1477	111	10.01	12641	15389	100087	127641	36.88
48	3197	1532	111	10.01	12262	15295	99765	163364	36.1
49	3226	1532	112	10.82	12303	13573	84503	163716	31.43
50	3226	1546	112	10.82	12464	13636	84872	152396	29.76
51	3226	1546	112	10.82	12534	13648	84873	172192	29.09
52	3226	1546	112	10.78	11600	14052	86457	157150	27.55
53	3254	1546	113	10.78	11874	14062	86613	169176	28.08
54	3254	1559	113	10.78	12066	14042	86227	153933	26.54
55	1958	1559	68	10.76	12351	14017	85741	100822	25.8
56	634	938	22	10.77	12502	14162	86643	165480	25.73
57	2477	304	86	10.76	12421	14085	86261	146645	28.09
58	3110	1187	108	10.78	12463	14026	86386	143021	29.02

Continues next page...

Table D.1 – *Continued from previous page*

No	Face_L	Blast	Teams	Offreef	Allow_OB	Error_OB	Channel_W	Weighline	Replace_factor
59	3110	1490	108	10.78	12641	14026	86730	138418	32.82
60	3082	1490	107	10.78	12583	13986	86730	166390	34.13
61	3197	1477	111	12.00	7959	13915	74570	155376	18.19
62	2966	1532	103	12.00	8909	15590	83572	152654	20.6
63	2909	1421	101	12.00	9459	16512	88681	171855	22.98
64	2880	1394	100	12.00	10082	17553	94047	162542	27.11
65	2909	1380	101	12.00	10106	17657	94382	191025	28.17
66	2909	1394	101	12.00	10259	17709	94941	168687	27.12
67	2909	1394	101	12.00	10283	17798	94788	76283	30.15
68	2938	1394	102	12.00	10117	17895	95257	187157	32.58
69	2938	1408	102	12.00	9965	17847	94524	157567	33.9
70	2880	1408	100	12.00	9962	17859	94376	134425	34.38
71	2851	1380	99	12.00	9990	17887	94195	150746	33.79
72	2822	1366	98	12.00	10011	17893	94140	173283	33.93
73	2851	1352	99	14.89	13301	12242	67237	135731	31.36
74	2851	1366	99	14.89	14271	13326	73370	143084	36.29
75	2880	1366	100	14.89	13793	12610	69521	146963	34.51
76	2880	1380	100	14.89	13589	12188	67077	131795	30.25

Continues next page...

Table D.1 – *Continued from previous page*

No	Face_L	Blast	Teams	Offreef	Allow_OB	Error_OB	Channel_W	Weighline	Replace_factor
77	2880	1380	100	14.89	13534	12175	66857	148572	30.03
78	2880	1380	100	14.89	11736	10584	58091	126936	27.19
79	2880	1380	100	14.89	8092	7301	40124	61031	22.8
80	2851	1380	99	14.96	14252	12808	70488	155155	27.05
81	2765	1366	96	15.00	12993	11667	64245	128454	26.84
82	2678	1325	93	15.00	10526	9563	52171	123460	24.88
83	2592	1283	90	15.00	13672	12307	67512	120734	27.27
84	2563	1242	89	15.00	13641	12281	67390	139780	25.72
85	2650	1228	92	0	0	0	0	0	0
86	2650	1270	92	15.00	5309	4213	25627	45532	4.4
87	2650	1270	92	15.00	11193	8895	53724	97798	67.5
88	2650	1270	92	15.00	14856	11775	70935	141259	44.43
89	2650	1270	92	15.00	16744	12839	78478	155039	22.97
90	2650	1270	92	15.00	16084	12519	75730	130012	22.18
91	0	1270	0	15.00	8881	6881	41752	72571	23.39
92	0	0	0	14.94	15991	12396	74819	148929	21.83
93	0	0	0	14.88	15313	11830	71901	130692	22.35
94	0	0	0	14.82	12541	9625	58649	129419	22.81

Continues next page...

Table D.1 – *Continued from previous page*

No	**Face_L**	**Blast**	**Teams**	**Offreef**	**Allow_OB**	**Error_OB**	**Channel_W**	**Weighline**	**Replace_factor**
95	0	0	0	14.76	13929	10683	64871	124438	23.08
96	0	0	0	14.77	16405	12943	77703	138593	23.34

Continues next page...

Table D.2: Planned dataset II

No	**Centares**	**On_reef_dev**	**Off_reef_dev**	**Channel_Dilution**	**Survey_Call**	**Current_Sweepings**	**Mill**
1	38738	965.1	488.2	53.21	165900	36061	3.83
2	38748	974.6	460.5	53.05	166248	36070	3.83
3	38753	957.7	511.1	53.25	165277	36075	3.82
4	38792	966.7	502.7	53.27	165437	36111	3.81
5	38821	978.4	509	53.09	165220	36139	3.83
6	40259	918.7	559	51.75	169552	37468	3.88
7	40320	911	489.5	51.58	169061	37526	3.89
8	40461	869.7	505.6	51.22	168992	37660	3.90
9	40320	853.8	512.6	51.04	168155	37526	3.90
10	40358	854.4	487.5	51.24	168094	37562	3.89
11	39901	810.6	488.1	51.20	165642	37128	3.91
12	39401	842.6	526.9	52.12	164228	36653	3.87

Continues next page...

Table D.2 – *Continued from previous page*

No	Centares	On_reef_dev	Off_reef_dev	Channel_Dilution	Survey_Call	Current_Sweepings	Mill
13	42673	804.9	528.2	46.96	174382	39780	3.88
14	42705	826.3	541.2	46.74	175080	39811	3.87
15	42837	765.8	618.8	46.64	175078	39935	3.87
16	42692	801	621.7	46.72	174635	39835	3.87
17	42855	787.9	622.3	46.69	175374	39990	3.87
18	42714	787.3	565.1	46.57	175320	39856	3.87
19	42805	782.8	556.6	46.56	175456	39943	3.87
20	42717	763.2	560.8	46.58	174649	39859	3.87
21	42769	787.2	591.2	46.61	174816	39908	3.86
22	42745	780.6	668.9	46.68	174880	39904	3.85
23	42745	780.6	668.9	46.68	174880	39904	3.85
24	42707	794.8	588.6	46.65	175100	39867	3.84
25	41161	973.8	527.7	41.67	158920	38392	3.64
26	41171	903.7	596.2	41.32	158900	38402	3.65
27	41167	855.1	605.3	41.10	158812	38399	3.65
28	41134	813.4	564.9	40.92	158426	38402	3.66
29	41123	744.8	669	40.56	157974	38392	3.68
30	40916	665.8	628	40.14	157166	38177	3.71

Continues next page...

Table D.2 – *Continued from previous page*

No	Centares	On_reef_dev	Off_reef_dev	Channel_Dilution	Survey_Call	Current_Sweepings	Mill
31	41081	656.3	628.5	40.29	157983	38317	3.67
32	42041	668.8	598.1	40.17	161653	39161	3.67
33	42492	608.5	575.1	39.95	162558	39556	3.69
34	42081	583.5	608	39.76	161143	39165	3.69
35	42183	594	565.4	39.63	161488	39262	3.68
36	42073	595.5	552.8	39.41	161068	39158	3.68
37	40666	874.7	781.8	43.23	162280	38357	3.67
38	40692	846.2	753.3	43.21	162358	38381	3.67
39	40676	845.9	745.2	43.17	162187	38367	3.67
40	40687	787.9	762.3	43.22	161164	38377	3.68
41	40691	753.3	676.1	43.32	160825	38362	3.68
42	40690	717.1	605.6	43.26	161016	38362	3.68
43	41494	700.5	527	43.47	162284	38927	3.70
44	41443	663	522.6	43.49	161857	39077	3.71
45	41445	632	515.1	43.45	161619	39078	3.71
46	41447	612	508.8	43.21	161470	39080	3.72
47	41430	594.3	529.2	43.10	160777	39064	3.72
48	41194	625.2	515.8	42.96	160020	38840	3.73

Continues next page...

Table D.2 – *Continued from previous page*

No	Centares	On_reef_dev	Off_reef_dev	Channel_Dilution	Survey_Call	Current_Sweepings	Mill
49	36232	577.2	575.7	47.47	149316	33086	4.21
50	36136	586.7	627.7	47.37	150024	33257	3.69
51	36206	580.7	664.2	47.42	149936	33323	3.69
52	36421	609.4	712.7	47.78	150819	33528	3.70
53	36482	683.4	615.6	47.76	152083	33586	3.68
54	36429	822.2	550.6	47.91	153850	33536	3.64
55	36351	835	574.1	48.07	153183	33462	3.63
56	36738	796.1	631.8	48.08	153630	33829	3.64
57	36598	721.1	582	48.07	152487	33696	3.63
58	36681	611	652.9	47.97	151580	33775	3.64
59	36872	501.7	621.8	47.82	152001	33957	3.64
60	36758	462.7	614.4	47.78	151399	33848	3.63
61	30501	896.8	780.2	51.09	132938	27894	3.82
62	34212	831.2	829.8	51.29	146022	31420	3.82
63	36250	847.4	730.1	51.29	154326	33341	3.82
64	38507	771.6	649	51.49	162788	35485	3.85
65	38713	698.6	675.9	51.66	162700	35681	3.85
66	38873	646.4	787.2	51.54	161787	35832	3.85

Continues next page...

Table D.2 – *Continued from previous page*

No	Centares	On_reef_dev	Off_reef_dev	Channel_Dilution	Survey_Call	Current_Sweepings	Mill
67	38982	493.2	799.7	51.95	160138	35936	3.85
68	39121	458.9	741.9	51.98	160523	36068	3.85
69	38933	501.2	647.3	52.23	159656	35889	3.85
70	38914	512.5	619.5	52.31	159757	35871	3.85
71	38910	436.9	714.5	52.33	159495	35868	3.85
72	38957	451.7	696.6	52.39	160539	35913	3.85
73	28305	453.7	448.8	56.90	136263	26036	3.47
74	30687	440.6	405.1	56.66	146234	28232	3.47
75	29207	437.4	409	56.73	139405	26856	3.47
76	28367	554.9	382.8	56.95	136296	26083	3.47
77	28316	576.2	366.6	57.05	136364	26056	3.47
78	24577	579.1	324.7	57.04	120388	22613	3.47
79	16958	476.1	267.5	56.97	87104	15595	3.47
80	29653	592.2	503.9	56.89	142097	27289	3.47
81	26968	539.7	464.9	56.85	130825	24835	3.47
82	21943	520.4	361.6	57.18	109193	20218	3.47
83	28338	590.3	448.7	57.00	137396	26099	3.47
84	28305	597.7	503	57.08	137113	26070	3.47

Continues next page...

Table D.2 – *Continued from previous page*

No	Centares	On_reef_dev	Off_reef_dev	Channel_Dilution	Survey_Call	Current_Sweepings	Mill
85	0	0.1	0	0	1	0	0
86	10486	0.1	0	55.45	45543	9692	3.75
87	22063	197.9	129	55.74	97799	20414	3.71
88	29180	421.2	235.6	55.85	131761	26906	3.64
89	32293	784.4	621.6	55.53	149195	29863	3.64
90	31093	803.6	598.3	55.70	143498	28723	3.64
91	17163	407.6	326.2	55.58	81819	15490	3.61
92	30993	727.3	692.2	55.96	142814	28628	3.61
93	29857	677.5	658.2	55.75	136935	27549	3.61
94	24500	552.8	521.1	55.74	112636	22460	3.61
95	27320	577.3	606.3	55.98	124289	25139	0
96	32679	685.5	714.8	56.25	147368	30149	0

Continues next page...

Table D.3: Actual dataset I

No	Face_L	Blast	Teams	Offreef	Allow_OB	Error_OB	Channel_W	Weighline	Replace_factor
1	3601	2093	6.73	7.31	9016	26421	104063	174801	37.14
2	4307	1932	6.75	6.70	8231	24765	105692	171965	39.1

Continues next page...

Table D.3 – *Continued from previous page*

No	**Face_L**	**Blast**	**Teams**	**Offreef**	**Allow_OB**	**Error_OB**	**Channel_W**	**Weighline**	**Replace_factor**
3	4034	1913	6.71	8.66	6423	21120	90009	160144	32.99
4	3552	1968	6.61	8.51	7786	26548	96694	175128	32.89
5	3685	2010	6.53	8.32	6443	24567	93856	177048	28.08
6	3891	1890	6.61	9.90	7146	21481	88923	163224	30.87
7	3480	1684	6.53	9.02	4883	14394	59949	90876	35.76
8	4039	1792	6.48	10.32	5611	21495	90759	159145	35.43
9	4023	1742	6.45	10.53	6409	24887	99136	162148	33.32
10	3836	1952	6.53	8.95	5973	22526	94204	156447	39.81
11	4118	2086	6.68	10.81	7769	25751	99563	160761	32.68
12	3535	2005	6.67	14.20	8327	21818	87632	143793	35.75
13	3129	1961	6.60	15.09	3103	11856	46894	87621	325.79
14	3410	1926	6.58	12.94	6021	19076	79248	139310	711.84
15	3875	1964	6.42	10.89	6767	22995	99220	164529	127.48
16	3652	2008	6.63	10.29	6242	20739	92610	159063	60.72
17	3569	1964	6.51	13.30	7366	20850	86228	152219	52.04
18	3954	2008	6.46	11.61	8786	20557	96389	140347	64.33
19	3329	2025	6.26	8.01	5138	14376	74175	104639	62.75
20	3328	1949	6.38	12.18	4765	13642	67327	171202	32.26

Continues next page...

Table D.3 – *Continued from previous page*

No	Face_L	Blast	Teams	Offreef	Allow_OB	Error_OB	Channel_W	Weighline	Replace_factor
21	3153	1886	6.47	16.29	5814	16261	75276	176750	37.48
22	3380	1878	6.55	15.60	5765	13939	68926	86559	38.4
23	3181	1856	6.63	13.73	6049	12102	73826	158552	40.48
24	3482	1967	6.57	11.89	8019	13154	85038	117846	40.59
25	3577	1856	6.52	16.14	7363	14832	86963	248465	6.16
26	3509	1792	6.48	17.88	7113	13592	76775	185604	8.01
27	2434	1743	6.61	16.65	2645	3734	23970	98873	5.9
28	3615	1841	6.38	15.42	8169	15928	91716	275445	5.16
29	3102	1766	6.61	13.63	7043	11971	60747	252684	3.36
30	3660	1865	6.72	14.11	10831	17177	90044	299088	4.75
31	2835	1896	6.81	14.16	6553	8832	47405	169997	5.05
32	3966	1899	6.78	19.90	8860	13233	74165	240169	5.16
33	3897	1968	6.71	15.56	8499	13731	77010	218176	5.38
34	3462	1878	6.72	12.87	6791	14155	68303	196904	5.00
35	3377	2192	6.81	14.38	7809	11914	75037	202 136	5.43
36	3917	1926	6.72	14.69	8551	16444	80133	212050	5.83
37	3988	1986	6.66	16.80	8235	16380	84741	246087	6.85
38	3581	2076	6.58	11.93	8780	15270	82325	251397	5.35

Continues next page...

Table D.3 – *Continued from previous page*

No	Face_L	Blast	Teams	Offreef	Allow_OB	Error_OB	Channel_W	Weighline	Replace_factor
39	3411	2119	6.66	15.05	8977	18969	82149	253957	5.99
40	3148	2043	6.75	14.75	7569	15424	81219	240391	5.7
41	2951	1842	6.72	9.26	4039	9354	54815	191894	3.77
42	3067	1924	6.64	15.34	5234	11963	60522	208780	4.43
43	2504	2010	6.70	12.34	4302	8128	39278	130883	5.00
44	3377	2021	6.75	15.19	7038	14885	67971	218846	6.1
45	3047	2036	6.70	15.20	6422	13849	65213	208124	5.45
46	3385	2169	6.81	16.65	8387	14290	67111	216002	5.72
47	3534	2251	6.72	22.05	5185	11416	48400	179582	5.44
48	4021	2220	6.71	17.74	7511	19077	73803	254119	6.29
49	4203	2221	6.72	15.82	7964	18478	76508	243896	6.02
50	4408	2259	6.60	15.98	3609	23596	95677	260244	6.28
51	3627	2275	6.62	22.07	5204	20314	80021	265918	6.01
52	3234	2236	6.81	24.33	5338	19368	68140	223786	6
53	3178	1919	6.86	22.75	4364	17867	63379	226802	5.42
54	3038	2108	6.78	22.65	2834	12717	44883	161526	4.95
55	2397	1166 c	6.60	23.11	2766	9524	37184	126073	6.32
56	2231	112	6.60	19.03	1146	4018	18562	28518	5.29

Continues next page...

Table D.3 – *Continued from previous page*

No	Face_L	Blast	Teams	Offreef	Allow_OB	Error_OB	Channel_W	Weighline	Replace_factor
57	0	1775	0	100.00	0	0	0	328038	0.02
58	2358	2140	6.76	21.63	4048	11612	47205	156445	6.07
59	2509	2082	7.12	19.52	4653	14386	54837	207533	4.22
60	2483	2037	6.93	26.28	5274	14372	48207	222797	4.64
61	2989	2188	7.04	25.59	5456	17779	61632	247433	4.98
62	3423	1966	6.80	26.44	5516	18681	66888	228891	4.83
63	2888	1803	6.84	22.02	4936	14066	51257	208937	4.42
64	3202	1783	6.91	18.93	6413	17106	65247	199297	5.04
65	2702	1777	6.90	25.84	4347	12549	46279	199026	4.08
66	2839	1802	6.86	22.18	4802	15947	61277	195246	4.59
67	1849	1935	6.71	21.16	1482	4510	16331	81720	2.97
68	2833	1912	6.69	15.12	5426	18014	64758	236688	4.51
69	2826	1898	6.84	19.48	4251	16001	55359	195975	4.3
70	2632	1769	6.97	17.58	3537	14940	44733	133621	4.32
71	2633	1792	6.90	16.11	4145	17234	54192	197557	5.36
72	3013	1795	6.87	16.28	5364	17209	54754	207466	4.32
73	3114	1849	6.83	19.85	5719	19976	66807	244844	5.02
74	3458	1856	6.79	17.54	7316	17081	68310	270719	4.22

Continues next page...

Table D.3 – *Continued from previous page*

No	Face_L	Blast	Teams	Offreef	Allow_OB	Error_OB	Channel_W	Weighline	Replace_factor
75	3626	1797	6.72	17.59	8657	17139	77273	260230	4.73
76	3783	1875	6.92	14.96	8153	17031	78648	246297	4.58
77	3108	1857	6.76	14.72	8526	17150	76500	254021	4.57
78	2541	1812	6.88	11.82	6364	15066	58537	214608	3.54
79	1880	1918	6.70	15.37	2909	4856	21396	77139	3.25
80	0	0	0	0	0	0	0	0	0
81	0	0	0	0	0	0	0	0	0
82	0	0	0	0	0	0	0	0	0
83	0	0	0	0	0	0	0	0	0
84	0	0	0	0	0	0	0	0	0
85	0	0	0	0	0	0	0	0	0
86	2330	1870	6.48	14.43	6083	11953	52610	197233	7.08
87	2416	1632	6.64	14.56	5101	9871	46879	173722	4.31
88	2143	1660	6.76	12.26	5991	6740	32898	225720	2.7
89	2704	1666	6.77	13.52	9751	17508	76224	272344	4.77
90	3104	1641	6.94	15.58	10587	15678	69572	298214	4.06
91	2216	1792	6.75	14.87	4516	7859	31836	151996	4
92	3015	1708	6.76	13.11	10147	18195	79112	295062	5.39

Continues next page...

Table D.3 – *Continued from previous page*

No	Face_L	Blast	Teams	Offreef	Allow_OB	Error_OB	Channel_W	Weighline	Replace_factor
93	3034	1673	6.92	10.29	8157	15735	61492	255576	4.02
94	2890	1644	6.92	16.52	7375	12823	51409	198949	4.14
95	3334	1656	6.82	15.06	9615	19158	73736	141671	41.31
96	3206	1770	7.00	12.90	12493	22051	75632	140747	35.77

Continues next page...

Table D.4: Planned dataset II

No	Centares	On_reef_dev	Off_reef_dev	Channel_Dilution	Survey_Call	Current_Sweepings	Mill
1	44348	387.2	807	52.76	180133	35564	3.67
2	43816	330.6	789.9	49.59	175556	28316	3.78
3	37786	380.9	764.6	49.36	160441	30226	3.62
4	41788	449	821.7	54.95	175015	30500	3.65
5	39802	329.3	1088.3	56.75	170676	32556	3.75
6	38276	370.6	869.3	57.18	163251	31636	3.74
7	25570	203.6	511.5	55.55	108789	20108	3.51
8	37973	365.7	706.2	56.57	154841	25367	3.62
9	41812	466.5	788.2	55.27	179911	40189	3.53
10	39154	331.5	652	51.64	171946	34574	3.46

Continues next page...

Table D.4 – *Continued from previous page*

No	**Centares**	**On_reef_dev**	**Off_reef_dev**	**Channel_Dilution**	**Survey_Call**	**Current_Sweepings**	**Mill**
11	42398	468.3	828.9	56.3	186206	37229	3.40
12	39485	467	637.6	59.5	171231	33575	3.42
13	20851	64	0	67.1	96002	18455	3.45
14	34401	48.3	0	58.21	140500	24475	3.35
15	41708	236.5	90.7	51.22	169660	36665	3.45
16	38686	394.5	242.6	49.92	159598	30700	3.53
17	38155	443.6	289.6	55.86	159019	28396	3.41
18	41450	422.7	221.6	49.13	168310	33301	3.51
19	30062	304.9	174.2	42.59	124615	23595	3.65
20	28781	567.7	324.4	49.89	124061	20610	3.58
21	34579	310.7	185.3	59.12	148423	27539	3.54
22	30819	248.2	248.2	58.47	133805	22749	3.71
23	32067	227.4	156.6	51.28	136394	22568	3.53
24	36417	641.7	255.5	43.06	144102	25584	3.63
25	38753	3745.4	2549.6	53.73	251775	29474	3.66
26	35096	2493	1887	60.51	192146	23966	3.58
27	10982	956.2	906.2	50.94	64735	1921	3.50
28	40930	4719.7	3209.8	51.69	307488	34917	3.53

Continues next page...

Table D.4 – *Continued from previous page*

No	Centares	On_reef_dev	Off_reef_dev	Channel_Dilution	Survey_Call	Current_Sweepings	Mill
29	27279	5288	2829.9	50.25	256312	17306	3.54
30	40976	5546.5	3087.6	48.62	320702	28508	3.61
31	21813	2474.7	1847.9	49.13	172700	12764	3.59
32	35872	4258.1	2690.8	58.34	254588	25073	3.73
33	34617	3790	2643.9	49.05	242156	25534	3.74
34	29899	3023.7	2954.7	49.59	210607	16330	3.85
35	32727	3238.2	2790.7	48.7	218239	15447	3.54
36	35815	3592.7	2547	54.42	235320	22009	3.48
37	38100	3024.3 2537.8	57.6	232454	23392	3.57	
38	35600	3819.9	2839.7	51.13	255042	23949	3.75
39	37290	3769.7	2457.2	58.77	253152	24585	3.77
40	35435	3452.2	2769.4	54.36	250764	27790	3.73
41	21709	3414.3	2346.7	42.32	188373	9377	3.70
42	26635	3418.2	2591.9	54.92	216126	16365	3.50
43	16809	1789.5	1573.7	48.55	132573	12380	3.68
44	30420	2554.1	2429.1	55.73	206566	16528	3.73
45	28790	2785.9	2498.2	55.05	210990	19063	3.54
46	30953	2574.1	2837.9	59.88	216971	17480	3.75

Continues next page...

Table D.4 – *Continued from previous page*

No	Centares	On_reef_dev	Off_reef_dev	Channel_Dilution	Survey_Call	Current_Sweepings	Mill
47	23730	2429	1936.4	78.52	179775	11329	3.54
48	34626	2892.9	2611.4	72.68	237558	18910	3.75
49	33939	3042.3	2590.7	67.64	230886	17535	3.63
50	39394	3008.4	3267.1	70.89	265704	30163	3.60
51	36061	3038.8	2965.4	76.63	244788	25795	3.37
52	33345	2718.8	2842.3	83.38	225272	25228	3.48
53	30206	2588	2981	79.47	208945	13203	3.25
54	20958	2089.8	2143.3	81.01	161140	6828	3.38
55	18014	1405	1445.9	79.25	119967	12384	3.18
56	8249	754	805.6	65.33	64146	3812	3.76
57	3	79.4	119.5	0	2250	0	3.09
58	22346	1546.2	2135.4	73.21	162232	10647	3.26
59	25651	3138.6	2933.8	72.48	219250	17742	3.23
60	25425	2830.8	2645	92.79	195558	13805	3.11
61	30859	2847.7	3354.1	88.03	238441	21529	3.28
62	32931	3122.9	3688.9	92.56	231484	20646	3.24
63	24189	2610	2861.3	78.72	190695	12841	3.44
64	29693	2538.9	3357.4	73.88	208522	20129	3.16

Continues next page...

Table D.4 – *Continued from previous page*

No	**Centares**	**On_reef_dev**	**Off_reef_dev**	**Channel_Dilution**	**Survey_Call**	**Current_Sweepings**	**Mill**
65	23074	2469.2	3187.3	87.63	183893	13791	3.54
66	28370	2716.4	3463.4	75.41	211379	14939	3.52
67	7641	1055	1517.5	74.98	76491	3135	3.47
68	27875	3092.7	3092.9	69.24	228280	23686	3.60
69	25017	2468.8	3343.6	75.88	198883	15411	3.85
70	20197	2060.5	2619.4	78.7	163594	11935	3.68
71	23557	2028.2	2365.2	76.93	182598	17684	3.70
72	24456	2929.7	2736.9	81.9	208625	14831	3.64
73	30073	3097.1	2897.5	80.88	257922	22790	3.59
74	28165	3583.3	3086.3	74.04	275411	18095	3.56
75	31260	3415.8	3187.2	72.9	259973	21278	3.75
76	30772	3495.9	3218.7	62.85	270138	22161	3.36
77	30226	3313	3303.4	61.58	244159	20506	3.77
78	22686	3151.3	3262	59.41	227306	14679	3.51
79	8611	1242.4	1404.3	66.75	86918	4367	3.58
80	0	0	0	0	0	0	0
81	0	0	0	0	0	0	0
82	0	0	0	0	0	0	0

Continues next page...

Table D.4 – *Continued from previous page*

No	Centares	On_reef_dev	Off_reef_dev	Channel_Dilution	Survey_Call	Current_Sweepings	Mill
83	0	0	0	0	0	0	0
84	0	0	0	0	0	0	0
85	0	0	0	0	0	0	0
86	20319	1602.7	1266.9	64.34	161156	12381	3.73
87	18282	2196.1	2050.5	61.08	192534	3617	3.65
88	13069	2619.5	2229.5	52.54	203643	14817	3.71
89	29924	3589.1	2680	59.19	322246	21284	3.61
90	29000	4023	3128.2	60.65	318444	18176	3.77
91	13169	1782.3	1512.7	63.98	137265	8714	3.51
92	31200	3034	2749.8	59.2	300027	24195	3.67
93	23467	3102.9	2729.8	60.86	265011	15276	3.40
94	21478	2495.1	2692.3	71.83	224381	13464	3.57
95	29891	0	0	69.73	0	17921	3.48
96	31343	0	0	69.61	0	20497	0

Continues next page...

The zeros in Table D.4 highlight the period when the platinum mines were on strike.