A Refined Resource Model for Tshepong Mine.

Ruan du Toit

Student number: 395098

School of Engineering and the Built Environment

University of the Witwatersrand

Johannesburg, South Africa

Supervisor: Prof CE Dohm

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

May 2018

DECLARATION

I declare that this research report is my own unaided work. Where use was made of the work of others, it was duly acknowledged. It is being submitted for the Degree of Masters of Science in Mining Engineering at the University of the Witwatersrand, Johannesburg.

It has not been submitted before for any degree or examination to any other University.

Ruan du Toit

16th day of MAY 2018 in WELKOM

ABSTRACT

The purpose of this research is to generate a refined Mineral Resource model for Tshepong mine by validating and cleaning the sampling data base that has historically been subjected to hard coded grade capping and by harnessing the assay data of check samples. It is standard practice to chip a check sample at the bottom contact of each underground sample section. An additional experimental variogram value at a lag of 8cm was calculated and used in the modelling of the variograms, improving the estimation of the nugget effect.

No recorded reason or justification for the capping was found and the capped values did not have context in current estimations.

It was necessary to revert to the original values prior to the application of hard-coded values resulting in a validated, raw and error free sampling database.

The geological domains were updated based on updated facies plans and a value trend analyses. Exploratory data analyses were performed within the six newly defined geozones confirming that the domaining was effective and the evidence of stationarity within the geozones was deemed acceptable.

Variogram contours maps, at a sampling grid scale as well as at two regularised block grid scales, were used in the modelling of the spatial continuity of the mineralisation within each geozone. Ordinary Kriging is used for estimation into 30m x 30m blocks and kriging neighbourhood analyses were carried out using the sampling grid variogram models. Simple Macro Kriging is used for the estimation into 60m x 60m and 120m x 120m blocks.

The refined estimation model was validated by cross-validations (jack-knifing), and comparisons of the grade distributions and mean grades. Reconciliation between the old and the refined model highlighted differences that are interpreted and accounted for.

This research contributed towards improving the quality of Mineral Resource model for Tshepong mine by delivering a validated assay database and exploiting the available information and knowledge and the inclusion of recommendations for future Mineral Resource model updates.

ACKNOWLEDGEMENTS

I would like to thank my wife, Terryanne, for her understanding, support and patience during the time that I prepared this research report.

I would also like to thank Harmony Gold Mining Company Limited for providing me with data and software that was necessary for this research.

I would also like to thank to thank the following individuals:

- Professor Christina Dohm (University of Witwatersrand), my supervisor, for her support, challenges, encouragement and knowledge that she provided me with;
- Dr. Craig Morgan (Ore Reserve Manager: Geostatistics) for the useful discussions and Datamine® support;
- Andrew Louw (Ore Reserve Manager) for providing me with workspace and time to complete this research;
- Conrad Pienaar (Senior Geologist) for his assistance during the Facies plan update process;
- Dries Fourie (Geostatistician HOD) for his advice and support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

Centimetre grams per ton ……..cmg/t

Easting direction……..X

Grams per ton……..g/t

Lag distance……..h

Metres……..m

Metre grams per ton……..mg/t

Northing direction……..Y

Kilometres……..km

LIST OF ACRONYMS

BCF……Black Chert Facies
CRM……Certified Reference Material
EDA……Exploratory Data Analysis
ISO……International Organisation of Standards
LoM……Life of Mine
LF……Loraine Facies
MF……Melkkraal Facies
NE……North East
NW……North West
PTO……Planned Task Observation
QAQC……Quality Assurance and Quality Control
QKNA……Quantitative Kriging Neighbourhood Analysis
SAMREC……The South African Mineral Resource Committee
SMU……Smallest Mining Unit
SE……South East
SW……South West
UF1……Upper Footwall 1
GZ#……Geozone Number

1 INTRODUCTION

Tshepong mine is one of the Harmony administered gold mines that operates in the Free State Goldfields. The mine was initiated as a high tonnage, low cost operation that employed a semi-mechanised method to mine the Basal reef. In the early 1980's the results of 28 surface boreholes and the information of a 2D seismic survey aided in the decision of where to sink the mine shaft. In the early 1990's the mine layouts had to change due to unforeseen geological faulting found in the primary access development that moved out from the shaft. In 1993, the mining strategy was re-scoped and it was decided to use conventional mining techniques.

Tshepong mine is located in the northern part of the Free State near Odendaalsrus, 20km north of Welkom (Figure 1). Tshepong has an expected life of 19 years (2016) and is in the process of extending the decline to access high grade ore at deeper levels.

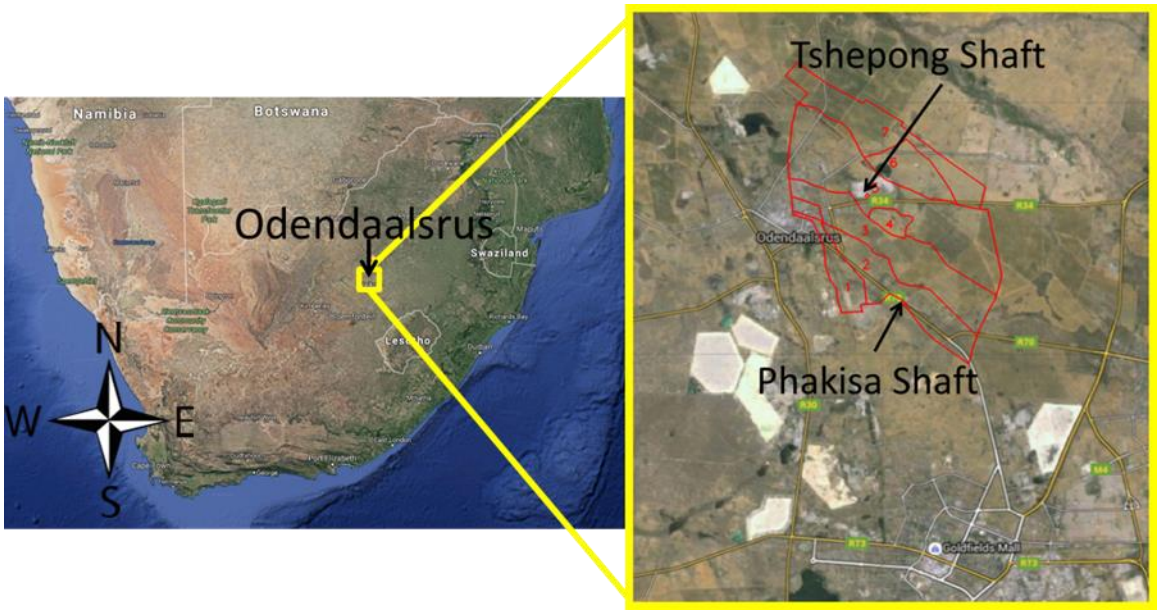


Figure 1: *Location of Tshepong and Phakisa mine*

Tshepong currently mines at depths of between 1600m and 2200m below the surface. The reef horizon is accessed by conventional grid development. Tshepong's Basal reef was deposited in a marine environment with an

upwards fining sequence and has two main facies namely the Black Chert Facies (BCF) and the Loraine Facies (LF). Generally, the grade continuity is in the north-west south-east direction. The primary economical reef of the shaft is the Basal reef that is extracted by undercut mining and leaving a quartzite beam in the hanging wall to support the overlying shale and to reduce dilution.

A good Mineral Resource evaluation programme using geostatistics among other techniques is currently in place. The objective of this research is to further improve on the quality of the estimation process and the resulting estimates by validating the sampling database and using additional available information in the estimation of the nugget effect.

Harmony complies with the SAMREC code in their approach to the classification of Mineral Resources. The mineral deposit within the lease area of the mine is divided into blocks. The geological model and the mine design determine the shape and the orientation of these blocks. The blocks are classified into Measured, Indicated and Inferred categories in accordance with the SAMREC code definitions. On Tshepong mine the category is related to the range of the sample variogram and the estimation is done into a 30m x 30m block model. The Indicated and Inferred Mineral Resource blocks which are related to the ranges of the variograms of the Indicated (60m x 60m) and Inferred (120m x 120m) block models. The Mineral Resource classification methodology used at Tshepong does not form part of this research.

An in-depth understanding of the data used for Mineral Resource evaluation is fundamental to reduce uncertainty in the Mineral Resource estimate and the success of this research will add value through the impact that the outcomes thereof will have on enhancing the quality of the mine plan. The optimal extraction of the Mineral Resource is extremely reliant on the validity of the Mineral Resource and geological models.

Increasing the confidence attached to the geological and Mineral Resource models will result in an improved management of the mining extraction mix

and consequently the production of gold. Capital investment projects like the extension of the decline will also benefit from less uncertainty and reduced risk regarding the projected financial returns.

1.1 Background Information

Research opportunities were triggered after interesting facts were revealed by the exploratory data analysis (EDA) of the historical data of Tshepong mine in 2015. This EDA was carried out as part of a course the Practical Implementation of Geostatistical Mineral Resource Evaluation (MINN7043). Data validations performed in the 2015 project, highlighted concerns with the data itself and with the predetermined scripted applications of geostatistical procedures. Scripts are automatic processes that are executed during the Mineral Resource evaluation process.

At Tshepong there is a strong need for a Mineral Resource model that can reliably predict gold grades within the geologically complex geozones. The comprehensive database covers a broad geographical spread and has been subjected to a consistent Quality Assurance and Quality Control (QAQC) program.

The mine has 7 geozones, the data issues identified in the one geozone analysed in the MINN7043 project, will by default also occur in the other 6 geozones of Tshepong mine, as it is believed that these issues are, to some extent, a function of data treatment and scripted rules that were historically applied to the sampling database.

Tshepong's sampling data is not stored in a single file, to give context to the data issues it is important to explain the composition of the data files. The sampling data has been captured in three different files namely the historical data file, the borehole file and the export file. The historical data file contains data captured prior to 6 December 2004, in a different data base and using a different software package than what is currently in use. The borehole file contains all the information of drilled boreholes and the export file contains

all the sampling data captured from 6 December 2004 onwards using the current software in place.

In the EDA of the 2015 project, two spikes in the tail of the centimetre grams per ton (cmg/t) histogram were observed, resulting in two flat areas occurring in the percentile plot of cmg/t leading to the conclusion that the historical data file was capped at two separate times, once at 3094cmg/t and once at 3697cmg/t. These values differ by about 600cmg/t.

As capping is carried out in the model creation process, it means that the capped historical data is again capped when the estimation scripts are run. It also means that the values used in the capping process can be affected by previously hard-coded capped values in the historic data base. The original database should be secure and untouched. The original data should never be over written automatically. Fortunately, it was possible to retrieve the original historical data from backup files.

A second important aspect is that a large amount of the available data can be used for improving the estimation of the nugget effect, i.e. close-range variability. It is standard practice on the mine to sample on a 5m x 5m sample grid; and to, at the bottom contact of the reef, chip two samples right next to one another, these samples are called check samples and are 8cm x 8cm each.

Currently, the closest point to the origin of the variogram is at a lag distance of 5 metres and the nugget is then modelled from the experimental 5 metre lag distance towards the origin of the variogram. Analysis of the semi-variance of these check samples provides an additional experimental semi-variogram value for samples that are separated by 8cm. It is expected that having this additional semi-variance data point at 8cm on the experimental semi-variogram will assist in improving the estimation accuracy of the nugget effect.

1.2 Research Objectives

The ultimate objective of this research is to create a refined Mineral Resource evaluation model with an improved Mineral Resource estimation. The accuracy of the model can be defined as the ability of the model to reliably predict the grade of the ore that will be recovered in future mining.

To create this refined model, an improvement of the data quality to be used in the estimation, was required, and was obtained through thorough data analyses and validations. All the parameters of the kriging neighbourhood were quantitatively determined. A small constant was added to all data points to keep the true values confidential.

1.3 Research Questions considered in this research

- Does updating the facies plan have a substantial effect on the 2017 Mineral Resource and the consequential Mineral Reserve blocks?
- Can the unsubstantiated historic data capping be resolved?
- How does an uncapped database influence the global mean values?
- Can the check sampling data provide an improved estimate of the nugget effect of the variogram?
- What is the impact of the improvements on the Mineral Resource estimate?

1.4 Literature Review

Domaining on the mine currently entails a visual check of the colour coded cmg/t plot of the sampling data and the geozones. The correctness of the geozones was improved by making EDA part of the domaining process. The spatial distribution of grades in a deposit is often related to lithological characteristics. By using geostatistical techniques to assist with the delineation of geological domains, uncertainty was reduced (Emery & González, 2007).

Domains can have hard or soft boundaries, hard boundaries occur where there is an abrupt change in grade, and soft boundaries where the change in

grade has a transition zone. Grades from both sides of the domains are used when estimating over soft boundaries. The data populations of Tshepong differ immensely from geozone to geozone and soft boundaries would have a negative influence on the accuracy of estimations that cater for transitional zones. The characteristics of data within a geozone should be more similar than to what it is outside of the geozone. Hard boundaries were used for the Tshepong model to ensure that no interpolation takes place across the boundaries because of the abrupt changes in grade when moving across boundaries. (Ortiz & Emery, 2006).

The geological model must be created to be consistent and to be able to explain the observed distribution of mineralisation (Glacken & Snowden, 2001).

It is natural to mine more high-grade portions of an ore body than low-grade portions as it is common to mine to an average above cut-off grade, inevitably causing more high-grade data than low-grade data to occur in the database due to the gathering of sampling information during the mining process.

Over-sampling of high grade areas causes a sampling bias. The preferential clustering can be corrected with domaining or declustering to prevent biasedness. Preferential clustering of data was analysed before data analysis took place to prevent biases in the Variography. (Glacken & Snowden, 2001).

In their book 'Geostatistics Explained' the authors introduce essential techniques of spatial data analysis, variograms and its application in kriging. As a confirmation of domain boundaries, the variances and the means of data at different locations were analysed at required levels of significance to determine similarity between domains (McKillup & Darby Dyar, 2010).

A spatial analysis of the data reveals continuity, range of influence and anisotropy. A display of the contour map (Varmap) of the sample variogram surface reveals the directional anisotropies and creates a clear picture that is used to understand how the semi-variances change in different directions.

The variogram contour maps must be kept in mind when looking at the experimental semi-variograms in all different directions to establish anisotropy (Isaaks & Srivastava, 1989) (Hengl, 2009).

To create a kriging estimate one must decide on search parameters and a maximum of local data to consider. To determine the optimum search range and the maximum number of samples the current methodology used on the mine was compared to the kriging neighbourhood analysis as proposed by Vann, et al (2003).

Behaviour of variograms at large distances is important, but the study of the behaviour near the origin is even more important (Armstrong, 1998). Armstrong gives a view on how to apply linear geostatistics and provides the theoretical knowledge required, which are not shown here but which were considered in this research. Studying the behaviour for small lag distance (h) distances will reveal the spatial regularity and continuity of the variable. This research includes a thorough investigation of behaviour near the origin by making use of the check sample data.

Modelling of structure and continuity in the deposit and problem areas like strong trends, random phenomena and proportional effect can be handled with techniques described by Clark (1979). The author dealt with trend by only modelling the variogram to the distance where the trend starts, changing the domain boundaries. To correct the proportional effect, individual experimental semi-variograms were divided by the square of the average of the samples that went into the calculation which resulted in a relative semi-variogram.

The current method of estimation and model validation used on Tshepong mine occurs in a process where the Mineral Resource and Mineral Reserve blocks are valuated with two different models and then comparing the results. The models used in this exercise are the Life of Mine (LoM) models, which are created once a year, resulting in a comparison of the current year's

results to the previous year's results. This method gives differences in the bottom line (i.e. in a global sense) in areas that are currently being mined.

The kriging weights are directly related to the variogram model, it is therefore important to perform a kriging neighbourhood analysis to prevent the acceptance of default software parameters.

The Quantitative Kriging Neighbourhood Analysis (QKNA) techniques discussed by Vann et al (2003) assisted in ensuring minimal conditional biases. Kriging can only be the minimum variance estimator if the neighbourhood is defined properly. The benefit of applying a QKNA is summarised in the results of case studies that they undertook, by comparing results for different simulations. They show and discuss what effect the amount of information has on the estimation statistics, and how the slope of regression, the correlation coefficient, the weight of the mean, the sum of negative weights and the kriging standard deviation changes with different scenarios.

By considering the estimated and true value of blocks, the slope of regression line is often used to measure conditional bias. The slope of the line of conditional unbiasedness is equal to one, and the slope of regression is always less than one. The slope of regression is minimised by using the correct data configuration (Deutsch, 2007).

According to Krige the evaluations of individual blocks that are done with limited data will be subject to error. The minimum error variances are influenced by the valuation technique that includes appropriate data search routines. Limiting a search neighbourhood will result in samples being uncorrelated to the true grades of the blocks which results in conditional biases, therefore a balance is required (Krige, 1996).

The number of samples used for estimations were optimised for each geozone in this research.

To test the effectiveness of the variogram used for estimations, the cross-validation method used by Armstrong (1998) was applied to different areas within the geozones. The method consists of temporarily eliminating some of the data and then kriging the value of the block using the remaining data and comparing the result. This method can highlight local estimation problems compared to the current method that only highlights a global problem.

Deutsch and Szymanski (2014) mentions the importance of site-specific techniques of kriging and that a block model must suit its purpose. They also recommend that cross-validations and assessments of conditional bias must form part of the implementation of kriging, which was done in this research. The cross-validation (Jack-knife) technique assisted with achieving the aim to reduce over- and under valuation of un-mined areas of Tshepong.

1.5 Research Methodologies

The research methodology applied in this research project is largely influenced by the Mineral Resource evaluation processes and techniques currently in use at Tshepong mine and is supported by the evaluation experience of the author and the literature research carried out.

An outline of the methodology and procedures for the creation of the Mineral Resource model of Tshepong mine is highlighted below, where the following steps were sequentially executed:

a. Creation of a raw, error free, uncapped data base of sampling information
    - Validation of the available data sources, including an investigation on the sampling support at Tshepong mine vs. Phakisa mine
    - Creation of an uncapped database
    - Removal of the Phakisa data due to the incompatibility of the sampling support
    - Exploratory data analysis (EDA) of the total uncapped database
b. Identification of geozone boundaries for accurate domain delineation and validation

- Value (cmg/t) trend analysis of gold grades on the appropriate sample support
- Updating of the facies plan based on information from the geologist

c. Geostatistical exploratory data analysis to improve the estimation of the nugget effect within the geozones
   - Distributions of differences of the check samples
   - The means and variances of the differences
   - Calculation of the semi-variance at 8cm to be used as an additional point on the experimental semi-variogram

d. Variography and spatial correlation analysis of the of the values in the geozones
   - Varmaps or contour maps for analysis of spatial continuity
   - Identification of the anisotropy i.e. directions of continuity of the mineralisation
   - Variogram modelling based on the varmap information and the nugget effect estimations from the variance of the check samples

e. Determination of the parameters to be used in the kriging process through QKNA
   - Search and estimation parameters
   - Maximum and minimum number of samples
   - Block discretisation
   - Conditional bias analysis: slope of regression and kriging efficiency

f. Mineral Resource block modelling at 30m x 30m, 60m x 60m, and 120m x 120m
   - Global mean determination
   - Grade capping – at the 95^{th} percentile of the values
   - Ordinary kriging
   - Simple macro kriging

g. Model validation
   - Jack-knifing
   - Block factor

- Distribution comparison
- Mean grade comparison

h. Reconciliation and comparison with previous model and the mining results leading to the conclusions of this research

2 GEOLOGICAL MODELLING

Geological modelling is the interpretation of geological data, a process that precedes the geo-domaining process used in Mineral Resource estimation.

2.1 Regional Geology

Tshepong is situated on the south-western corner of the Witwatersrand Basin. The Basin is situated on the Kaapvaal Craton which has been filled by a series of sedimentary rocks of about 6km thick and extends for hundreds of kilometres (Russel and Associates, 2014). The regional geology of the Free State Goldfield is shown in Figure 2.

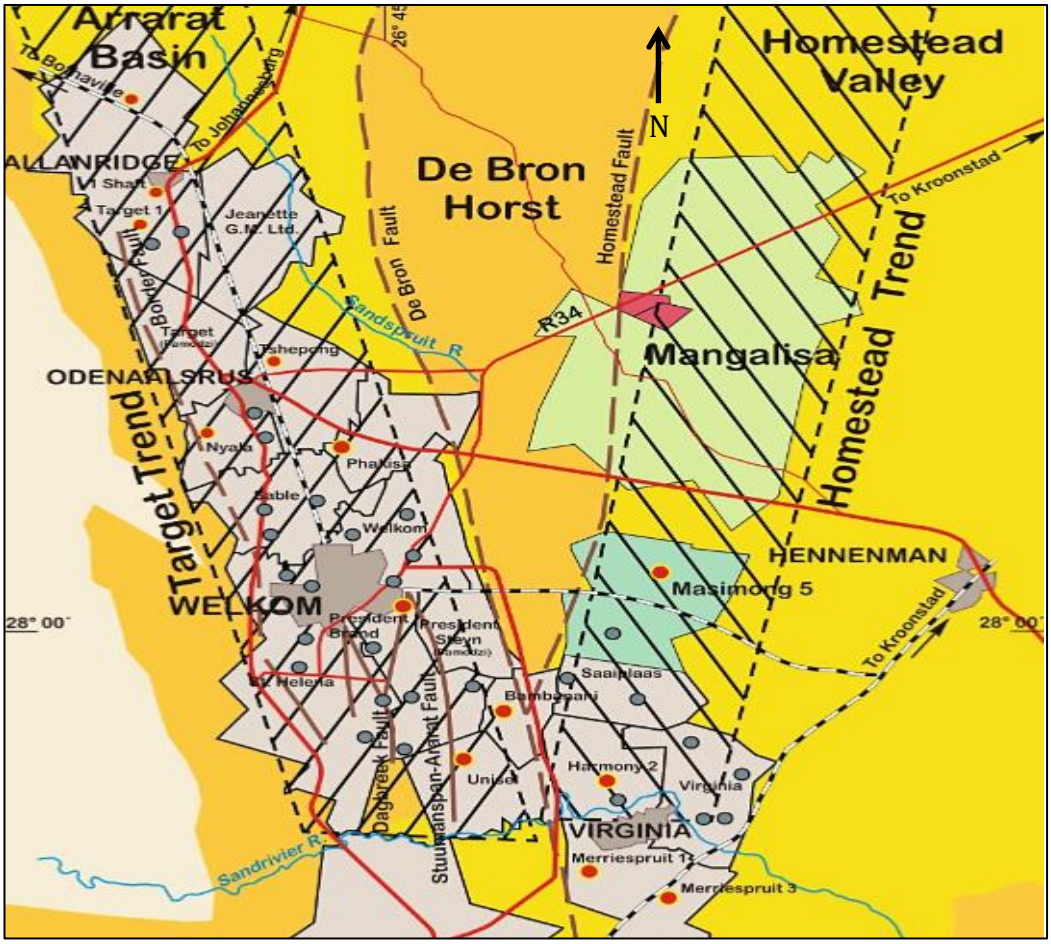


Figure 2: *Regional Geology of the Free State Goldfield. (Superior Mining International Corporation, 2011)*

The Free State Goldfield is divided into two parts by the De Bron fault which dips at an angle of 65˚ to the west. The fault strikes from north to south and

has a vertical downward displacement to the west of 1.5km and has a dextral shift of 4km (Ridge, 2013).

Other major faults that lie parallel to the De Bron fault are the Ararat, Dagbreek, Eureka and the Struismanspan faults (Superior Mining International Corporation, 2011).

The reefs to the west of the De Bron fault mostly dips at an angle of 30˚ to the east. Closer to the De Bron fault the dip of the reefs becomes a lot steeper. The reefs east of the De Bron fault mostly dip at 20˚ to the west. Between the De Bron and the Homestead faults lies an uplifted horst block of the West Rand Group where no reef has been preserved.

The mineralised zone to the west of the De Bron fault is the Ararat Basin, which is bounded to the west by synclines and complex structures. Towards the south and the east, the reefs sub-crops against the Karoo super group. The most common reef that is mined at most of the mines is Basal reef which varies in channels of between a single pebble lag and up to two metres thick (Tankard et al., 1982).

2.2 Mine Scale Geology

The Dagbreek fault strikes in a north-south direction and splits the mine into a western and eastern domain. The Dagbreek fault dips at an average angle of between 30˚ and 45˚ and has a variable down throw of approximately 300 metres to the western side. The main infrastructure of the shaft itself was developed in the loss zone of the Dagbreek fault as shown in Figure 3.

Tshepong's mineralisation model report indicates that the chemistry of the dykes, the distribution of the alteration, underground exposures, and a boundary analysis showed no evidence of lateral displacement by the Dagbreek fault, and that it should not be used as a break in the kriging estimation process. (Freeman et al., 1999).

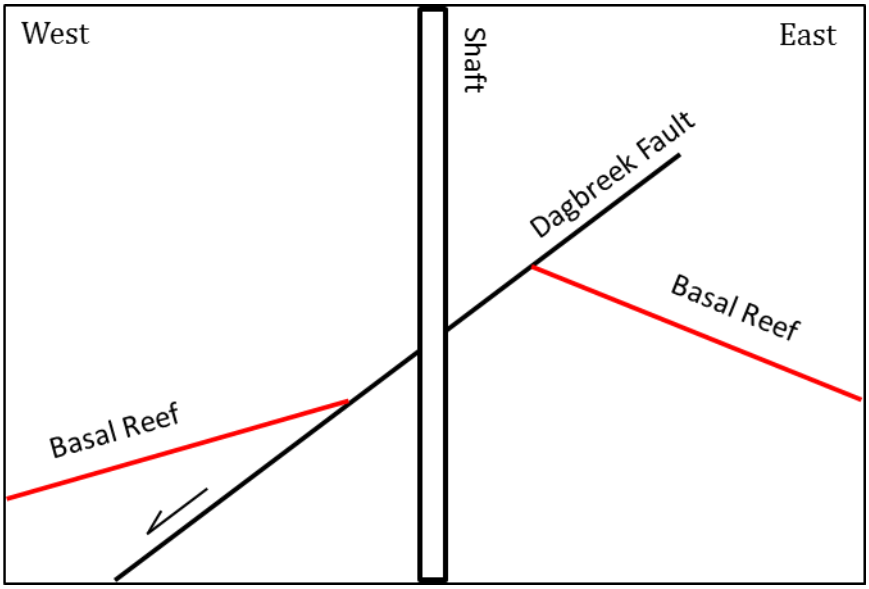


Figure 3: *Conceptual section view of Tshepong mine looking north (figure not to scale, for illustration purposes only)*

The domain to the west of the Dagbreek fault is structurally a lot more complex than the eastern domain. The western domain experienced major deformation, particularly in the northern part where block rotation has taken place resulting in a change of reef dip direction.

2.2.1 Stratigraphy

The mine-scale stratigraphy of the Khaki Shale, Waxy Brown Leader Quartzites and the Basal reef is displayed in Figure 4 below. Directly on the UF1 (Upper Footwall) quartzites lies the BCF and the LF of the Basal reef.

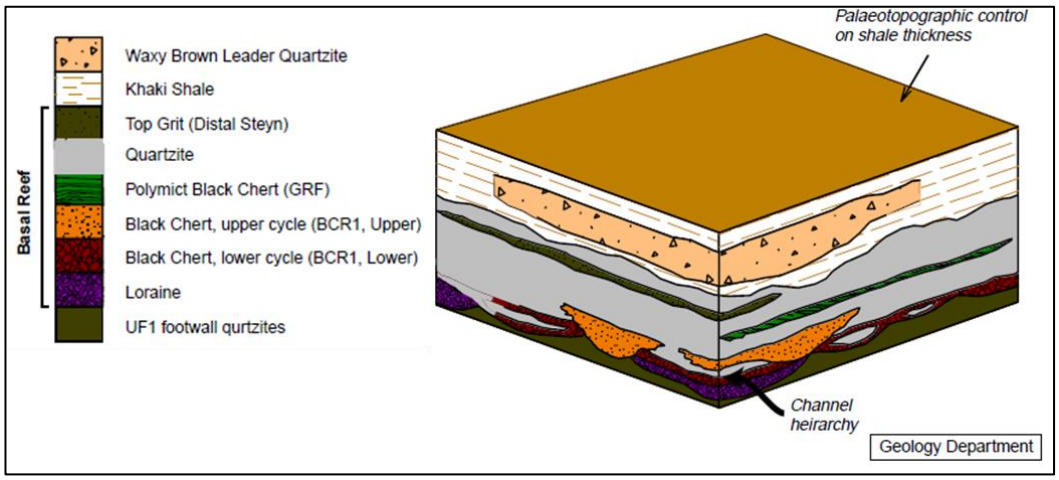


Figure 4: *Stratigraphy model of Tshepong mine (Freeman et al., 1999)*

The BCF always occurs above the LF. The Basal quartzite occurs above the Basal reef and is used as a beam to support the shale.

2.2.2 Deposition type

The Basal reef of Tshepong mine was deposited in a marine environment where finer sediments were transported away from high energy environments and settled in calmer environments. The sediments comprise of an upwards fining sequence, coarser sediments at the bottom and finer sediments at the top (personal observations).

2.2.3 Definition of geological domain

Two main facies of Basal reef exist on Tshepong mine, the LF to the north, and the BCF to the south. The make-up of the BCF contains up to 25% black chert clasts. The LF can also contain black chert clasts but only up to 5% of the make-up. There are small differences between the facies and this can complicate the identification of a facies from boreholes, but it is relatively easy from mined exposures due to the change in scale of observation (Freeman et al., 1999).

Looking at boreholes one can sometimes mistake LF for BCF. The LF comprises of a single, clean, oligomictic pebble lag. In the northern parts of the BCF the quartzite above the reef is cleaner and better sorted. The reef width is also thinner in the northern part of the BCF. The distribution of gold values (cmg/t) is cyclical between higher and lower values across the zone altering along strike with a frequency of roughly 100 metres.

A very small portion of Melkkraal Facies (MF) occurs at the south-western part of the mine in GZ1.

2.2.4 Sample support size

Tshepong uses an undercut mining technique to mine the Basal reef as displayed in Figure 5. The data for Tshepong is composited to this specific mining cut. A quartzite beam is left in the hanging wall to support the Khaki shale. The thickness of the quartzite beam differs over the lease area and for

safety reasons must be at least 60cm thick to provide sufficient support and prevent shale exposures.

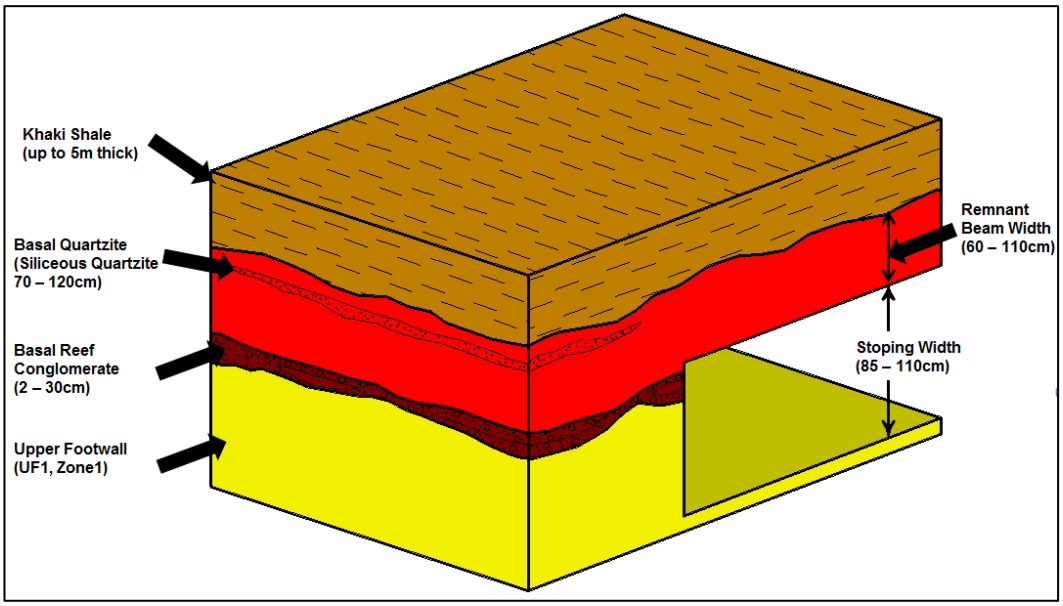


Figure 5: *Under-cut mining method illustration (Tshepong Geology Department, 2005)*

Tshepong mine and Phakisa mine are exploiting the Basal reef using two different mining techniques. The difference in the samples is that Phakisa mine does not use the undercut technique like Tshepong mine. This means that where Tshepong mine only has grades for the bottom contact of the reef, Phakisa mine has grades for the full channel. The channel width plot with geozones can be seen in Figure 6 below.

Both mines exploit GZ2 and GZ3 and consequentially it also means that some geozones contains areas where only the bottom contact was sampled, and the database contains grades for the bottom contact only, and other areas where the full channel was sampled grades for the full channel were captured. Samples that do not have the same support size are not compatible. This research focuses on the former data with bottom contact only.

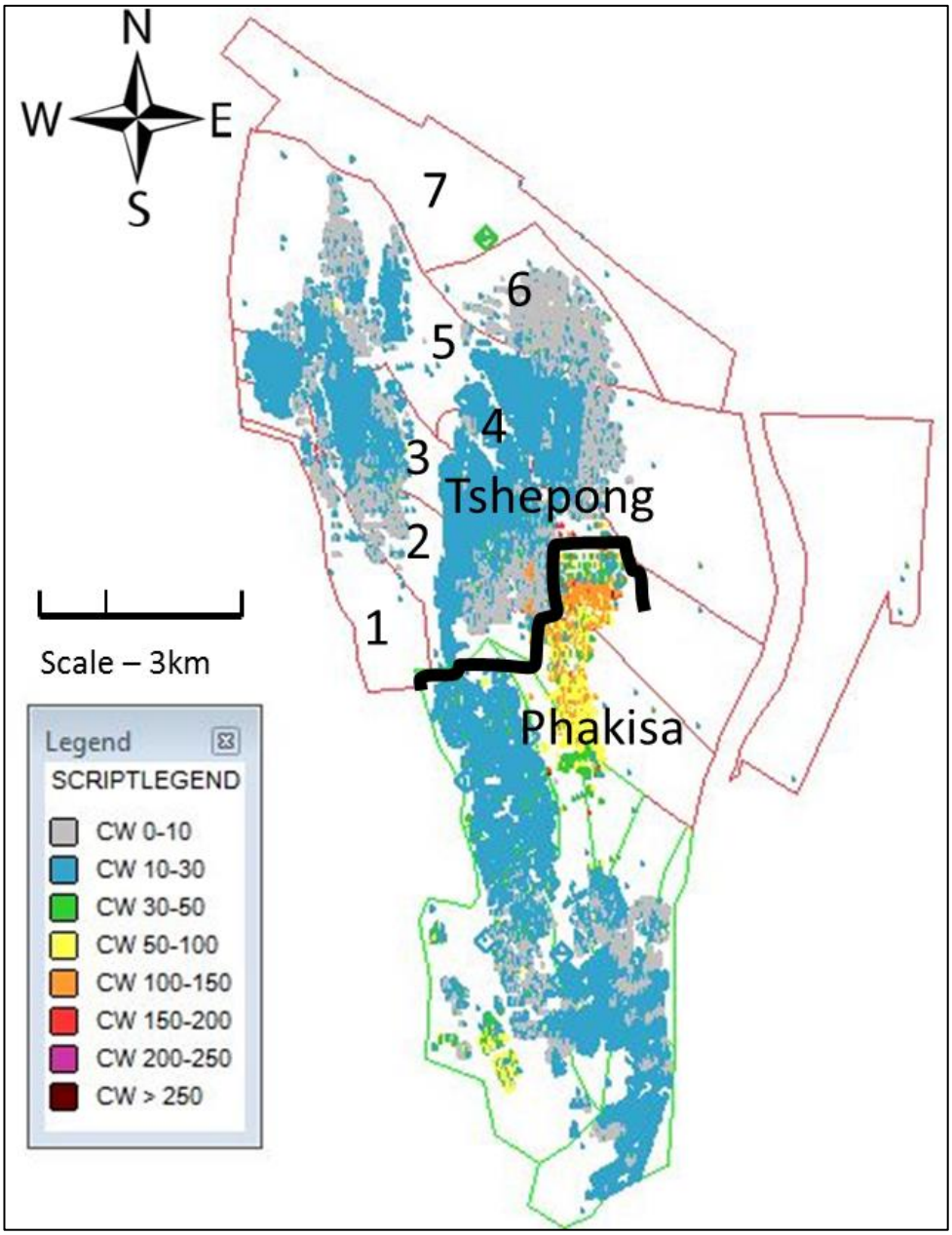


Figure 6: *Channel width plot of Tshepong and Phakisa data with geozone numbers for Tshepong mine*

2.3 Data Collection

In practice, it is not possible to have exhaustive sampling of any mineral deposit. Tiny fractions are sampled and through statistical and geostatistical analysis the properties of the samples are then inferred onto the population or estimated at unknown locations (Dohm, 2015 A).

Samplers are responsible for the sampling of stope panels and on-reef development ends. Basal stopes are sampled on a 5m x 5m grid and the development on 4m intervals. The aim is to maintain a 100% sampling coverage which has consistently been achieved for many years.

Sampling sections are determined by measuring from survey pegs. All sampling sections are washed with clean water to prevent contamination.

Sample positions are demarcated with chalk and are segregated according to the mineralisation (Figure 7). Clino rules are held parallel to the strike of the reef and measurements are taken at right angles to the dip. Check samples are chipped on the bottom contact of all sample sections. Sample widths are measured and recorded accurately in the sampler's fieldbook.

Figure 7: *Sample demarcation, sample measuring and sample chipping*

Samples are numbered with barcoded tickets corresponding to the assigned numbers in the sampler's fieldbook. Samplers record all measurements, geological features and mineralisation descriptions in their fieldbooks.

Samples are chipped by using a hammer and a chisel, a process that requires extreme caution to prevent contamination and the introduction of errors and bias. The aim is to extract the sample in the delineated area at uniform depth without any spillage or contamination. The samples are transferred from the sample dish to plastic bags and then transported to surface. The samples are transported to the assay laboratory in locked containers accompanied by a waybill and an analysis request form.

Samplers capture the geological and field measurement information from their fieldbooks in MineRP software before the assay results are electronically loaded. The process of sample collection and data capturing is checked on a regular basis by means of a Planned Task Observation (PTO) that is done by seniors in the department. A PTO covers the following stages of a sample:

- The definition of a sample's mass and shape
- The extraction of a sample
- Sample preparation, drying, crushing and splitting
- Sample analysis

Validating the above ensures the correctness of samples and can be used to control and reduce biasedness. Despite all the protocols human error is a possibility that cannot be ruled out and database validations before Mineral Resource modelling is crucial.

2.4 Database Validation

Inferences made from data to populations can only be as good as the quality of the original data. Checking for errors in data is the most time-consuming task in geostatistical modelling. The data must be examined for extreme low and high values. Coordinates should fall within expected limits and sampling data must be representative of the zone of interest and free of any errors.

The first validation point is where the Valuator checks that the submitted QAQC samples are within specified limits. Assay results will only be loaded into the database if the QAQC validation check is passed. The QAQC process as applied at Tshepong is discussed in section 2.5. Once the assay results are loaded the valuator will check each sample sheet individually to ensure the following:

- Data points must plot on the correct coordinates
- Footwall, reef, hanging wall, channel width and stoping width must correspond with the sampler's fieldbook

- Workplace names and sample dates
- Sampling intervals are correct

Additional database validations that are done before creating the Mineral Resource models include:

- Checking for duplicate coordinates within a specified radius
- Verify zero or blank gold grades, zero implies waste and blank implies no assay value
- Channel width less or bigger than specified widths
- Stoping width less or bigger than specified widths
- Visual inspection
- Visual review of a cmg/t histogram per geozone

The database validation process ensures that the Mineral Resource estimate is not influenced by potential biases or sampling errors.

2.4.1 Historical Data File

The sampling data of Tshepong mine is not stored in a single file, it occurs in three separate files namely the historical data file, the borehole file and the export file. The historical data file contains historical data that was captured in different software than what is currently in use and was captured prior to December 2004.

The borehole file contains all the borehole information and the export file contains all the sample data captured from December 2004 onwards.

The three separate files are merged to create one data file that is used for the Mineral Resource model creation. If any one of these three files, contains incorrect data before the merging process, this incorrect data will be merged into the final data file. It is thus mandatory to ensure that the data in the three separate files is correct before one proceeds with the merging process.

If more than one peak on the histogram is visible the data must be rechecked to ensure that the data comes from a homogeneous population (Armstrong, 1998).

The EDA of the historical data revealed many peaks in the tail of the distribution as shown in Figure 8. Due to the small bin size of 1cmg/t, it may appear in Figure 8 and later Figure 11 that some intervals have roughly twice the frequency of their neighbours. The histogram up to the first 500cmg/t in the 1cmg/t bin classes (Appendix A) shows that there is no preference of odd and even numbers, and that phenomenon observed is only an artefact of the small bin size and show the natural randomness of cmg/t values to be expected.

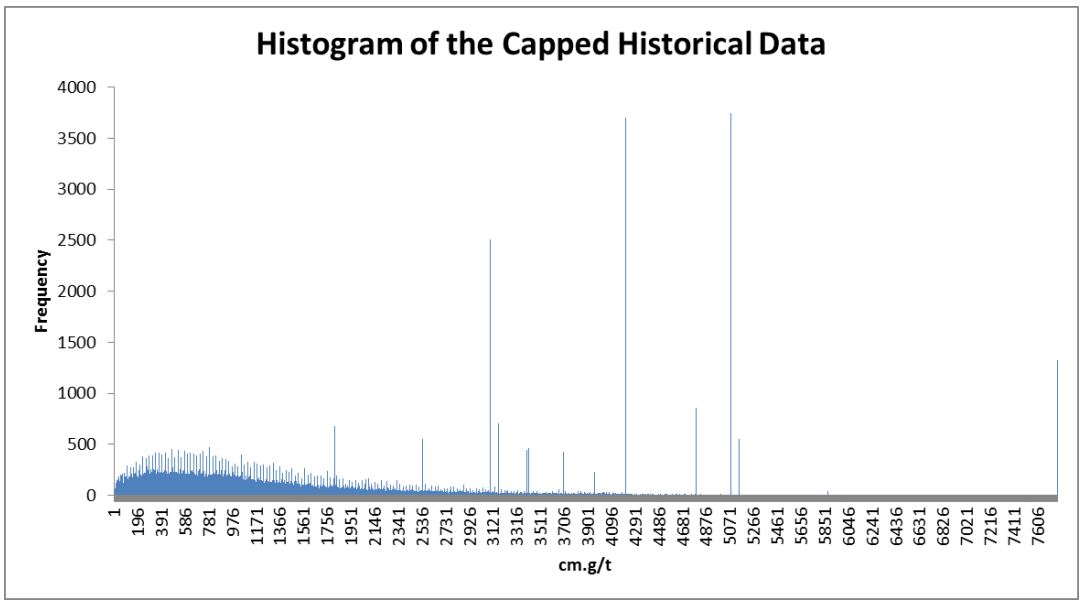


Figure 8: *Histogram of the capped historical data file*

In the 2015 EDA process, it was established that the data in the historical data file was capped on different values for each geozone. Capping was done by an external contractor, who was responsible for the Mineral Resource models a few years back. It is unclear how the cut-off values for the capping were chosen and what techniques were used. The data in the historical data file should not be capped and this data base should be validated and represent the assay results as received from the laboratory.

The capped values were identified, and each value was analysed individually. One of the methods used was to look at their spatial distribution on the map of Tshepong and Phakisa mines and within the previously defined geozones superimposed.

The plots in Figure 9 show that different capped values appear in different geozones, which means that whoever capped these values applied a different capping value for each geozone and that the Tshepong and Phakisa datasets were combined at that time.

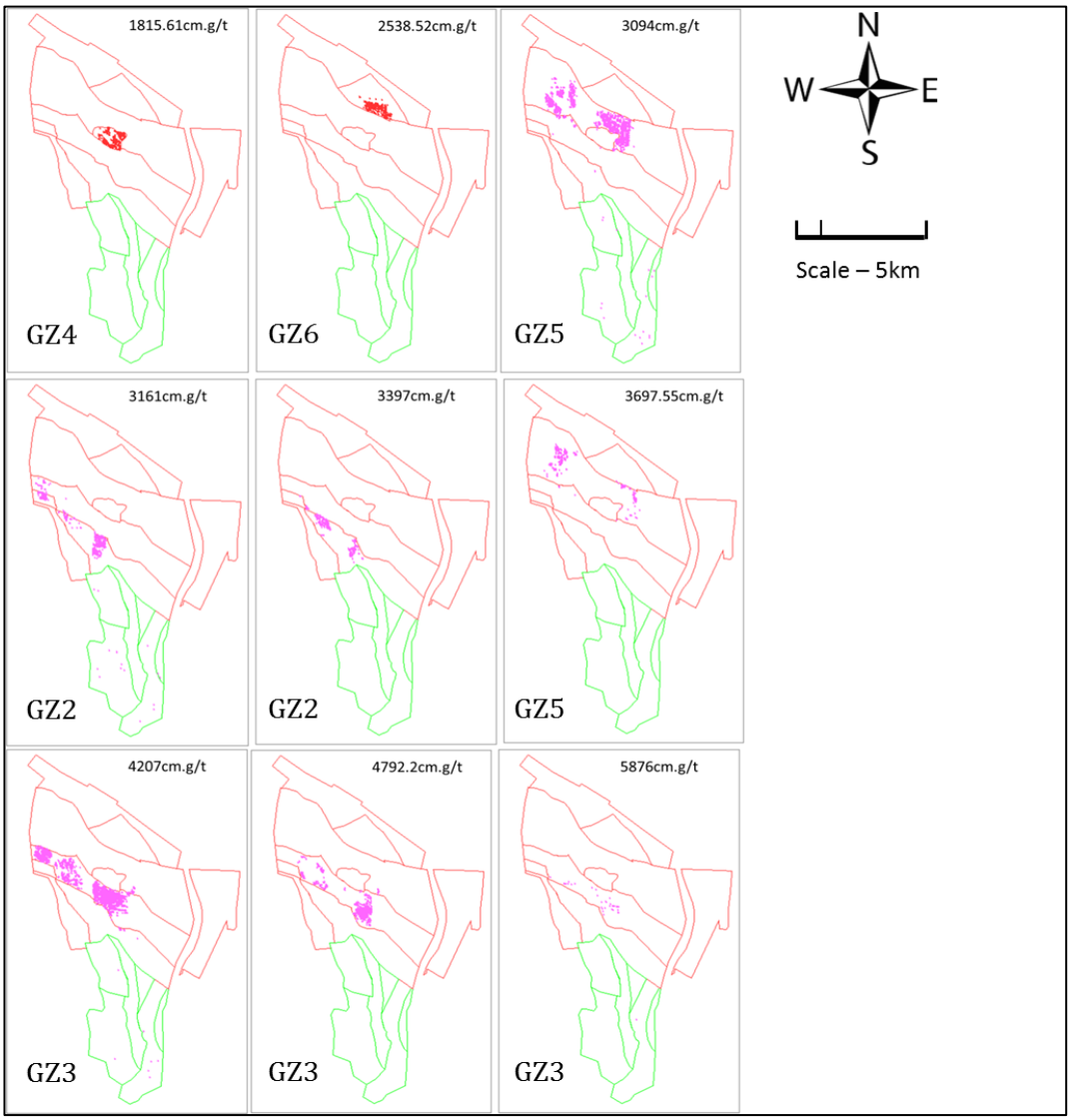


Figure 9: *Plan views of the location of all the capped values*

Geozones change because as mining progresses additional information becomes available and the geozone boundaries are reinterpreted. Data was capped on a specific value in a certain geozone, but after geozone changes the values could fall in neighbouring zones. It is clear that from the above and the evidence in Figure 9, that the hard-coding of capped values had a far-reaching influence on the interpretation of geozone boundaries as well as the base data in the geozones. It is important to clarify and confirm that hence forth the research is based on only the Tshepong mine sampling data.

To obtain the uncapped values, two old uncapped sampling data files from 2007 and 2010 were used to replace all the capped data with original data. The historical data file was corrected and validated to confirm that all the capped data was replaced with its original uncapped values.

Figure 10 below gives a summary of the specific capped cmg/t values and the percentage of the historical data that were capped to those values.

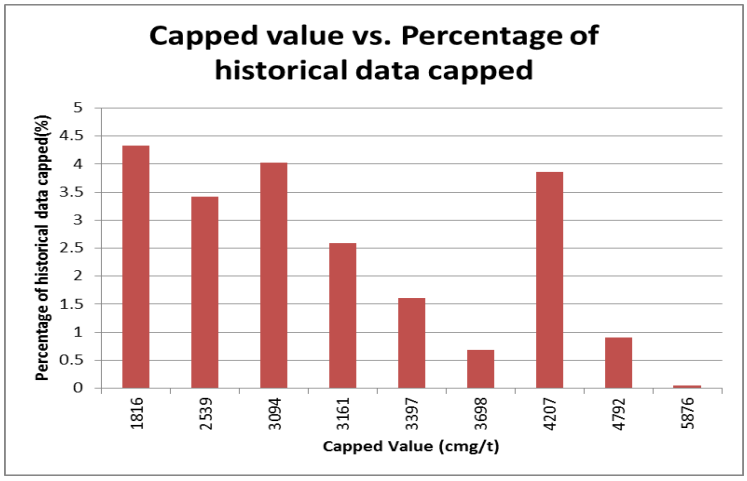


Figure 10: *Capped value (cmg/t) vs. percentage of historical data capped*

All the capped values were replaced by original values except for those highlighted in Table 1.

Table 1: *Summary of the capped values per geozone*

Geozone number	Capped Value (cmg/t)	Number of capped values	Percentage capped data %	Comments All original values were found bar the exceptions
2	3161	707	2.6	
2	3397	440	1.6	7 not found
3	4207	3699	3.9	
3	4792.2	860	0.9	4 not found
3	5876	39	0.04	1 not found
4	1815.62	590	4.3	2 not found
5	3094	2505	4.0	
5	3697.55	424	0.7	
6	2538.52	523	3.4	

In the few cases where original values were not found the capped value was retained as an original value, as it is assumed that it was not capped.

The previous histogram in Figure 8 shows many spikes in the tail of the distribution identified as capped values. The cmg/t histogram after replacing the capped values with the original values is shown in Figure 11. Both histograms were based on a bin size of 1cmg/t. To reveal the detail in each bin and to investigate the possibility of a preference for odd or even numbers, the distribution in Figure A-1 (Appendix A) was created up to 500cmg/t in bins of 1 cmg/t. The zoomed in detail show that there is a randomness in the values and the possibility of preference or bias has been ruled out.

A comparison of the two histograms reveals that issues with all the capped values have been addressed; the maximum frequency in Figure 11 is now around 450 whereas in Figure 8 it was around 3700, and all the spikes in the tail of the distribution have disappeared.

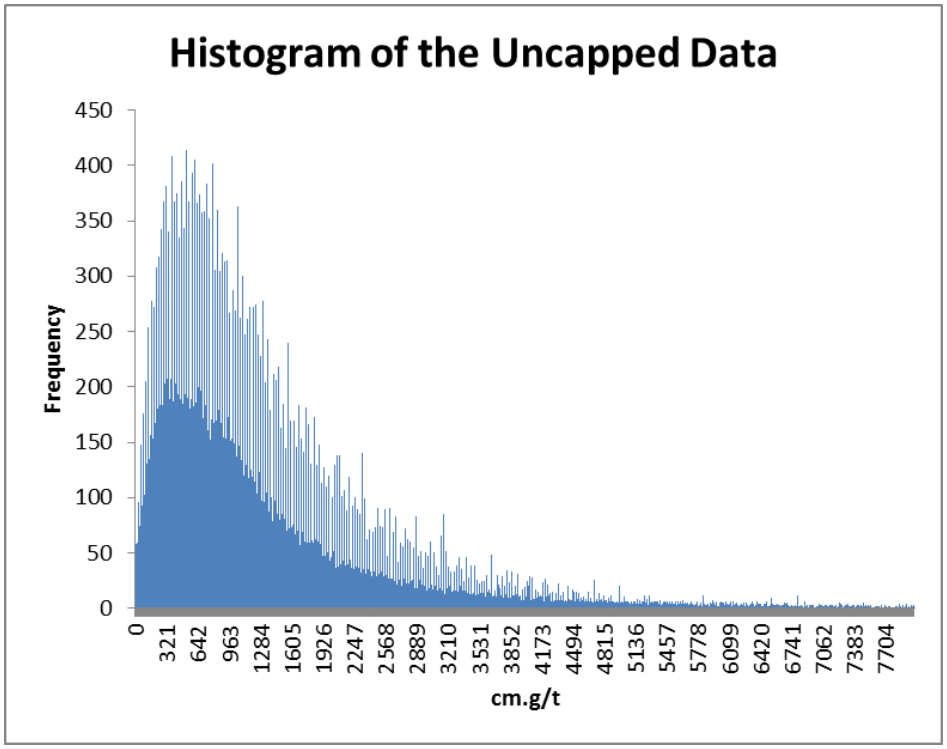


Figure 11: *Histogram of the uncapped historical data*

2.4.2 Duplicate samples

The historical data file contained some duplicate data that might have been introduced by digitising the same data from adjacent assay tracings sheets. The duplicate samples were removed before doing EDA by using a function in Datamine® called "FILTPO", which filters points, a radius of 50cm was used. The same function and a radius of 50cm was also applied to the export and borehole data files.

2.5 Quality Assurance and Quality Control (QAQC)

Mineral Resource estimation uses sampling information as a key input and thus the sampling data should be error free. Analytical results must be reproducible and unbiased. To ensure that Harmony complies with the International Organisation of Standards (ISO) 170025, a competent supervisor checks all the steps and procedures of the analytical process.

The author scrutinised the assay QAQC procedures in place at Harmony to gain an understanding thereof. There after he validated that the protocols are not only in place but are also being followed.

Assay data is checked for accuracy, precision and contamination. All the Harmony mines run their own QAQC programs in conjunction with the normal QAQC insertions done by the laboratory.

Accuracy is checked by submitting Certified Reference Material (CRM) which has a known certified grade. A CRM sample fails if the assayed value is out by more than three standard deviations.

Precision is checked by submitting duplicate samples that have been previously assayed, for another identical analysis. Failure of precision is when less than 80% of the duplicate samples falls within a 20% absolute relative difference value.

Contamination is checked by submitting blank samples. Silica quartz is used to ensure there are no gold grades in blank samples. If a blank sample has a value of more than three times the detection limit of 0.063g/t it will fail. Three types of blank samples are used, a 4mm course blank, a 19mm course blank and a pulverised blank. The 19mm sample enters the analytical process right at the start where samples are crushed by a jaw crusher. The 4mm blank enters the process at the pulverisation process. The pulverised blank sample enters the analytical process at the fluxing process. By using the three different blank samples it is easier to pinpoint which analytical process might have caused any contamination.

If a QAQC sample fails the criteria, a re-assay of the whole tray is requested. If the QAQC sample fails the second time, the whole tray of samples will be discarded and none of the assay results will be added to the data base. This prevents any biased or contaminated assay results to enter the data base.

A summary of the QAQC results of 2016 is shown in Table 2. The QAQC graphs for each (CRM) and the blanks are shown in Appendix B.

Table 2: *Summary of the QAQC results of 2016*

QAQC Results for the past year.			
CRM	Submitted	Failures	Comments
AMIS0108 (0.063g/t)	349	7	2.0% failure
AMIS0244 (6.77g/t)	120	1	0.8% failure
AMIS0245 (88.42g/t)	471	6	1.2% failure
AMIS0303 (8.77g/t)	195	7	3.6% failure
AMIS0369 (26.36g/t)	178	1	0.6% failure
AMIS0412 (5.74g/t)	32	8	2.5% failure
AMIS0428 (43.42g/t)	115	2	1.7% failure
AMIS0429 (22.93g/t)	97	0	0% failure
Blank 4mm	695	15	2.2% failure
Blank 19mm	619	6	1.0% failure
Duplicates	623	103	16.5% failure

The QAQC process starts with the chains of custody where lab personnel checks if the delivered samples correspond with the analysis request sheets from the mines. The condition of sample containers, sample bags and sample identification bar codes are checked, and any anomalies are recorded.

At the sample preparation section, all the drying temperatures are checked with a calibrated thermometer. Sample dryness is checked before a sample goes to the crushing process. Sample crushers and pulverisers are checked for sample loss and efficiency. Grading tests are done after samples are crushed to check if the sample fineness is to standard (80% of the sample should pass through a 75 micron sieve). Crushers and pulverisers are cleaned at intervals per the set standard (once before each tray of samples).

Balances at the fluxing section are checked daily with certified mass pieces to ensure that sample aliquot weights are correct. When required, additional chemicals can be added to specifically identified samples to ensure good fusion.

In the fusing section the temperatures of all cupellation furnaces are checked and logged to ensure that operational temperatures are within set limits. The completeness of fusion and cupellation are visually checked by a competent supervisor.

Proficiency testing is done by participating in a round robin where samples are exchanged between assay laboratories to compare assay results.

This program prevents assaying errors of entering the data base and ensures that assay values are accurate and precise and can be used for Mineral Resource modelling.

Together, all the different validations ensured that the data was error free, representative, accurate and precise before creating the Mineral Resource model.

3 DOMAINING

Before mining commenced on Tshepong, the total area was macro-kriged, as it was established that the Dagbreek fault does not have any significant lateral offset (Freeman et al., 1999).

One of the objectives of this research was to use the existing data to create a mineralisation model that would be capable of delineating zones with geological and geostatistical homogeneity.

3.1 Value Trend Analysis

A swath analysis or value trend analysis was used to identify trends in gold value in different directions across the deposit. A feeling for the different components of the cmg/t variability is obtained in this process. In this research the regional structured component, the systematic component, and the random component (check sample variance) are all considered.

The cmg/t values are more continuous along the transport flow direction than across it. The values in the direction across the transport flow direction changes cyclically between lower and higher values in bands that are roughly between 500 and 1000 metres wide (Figure 12).

The geozones fall close to the edges of these individual bands. Within the geozones the value populations are relatively homogeneous. Moving across these bands one can see that the value populations vary substantially. On the other hand, the variability in the cmg/t values within or along these bands is much less. This is clearly an early indication that the long axis of the anisotropy is likely to be in the NW/SE direction.

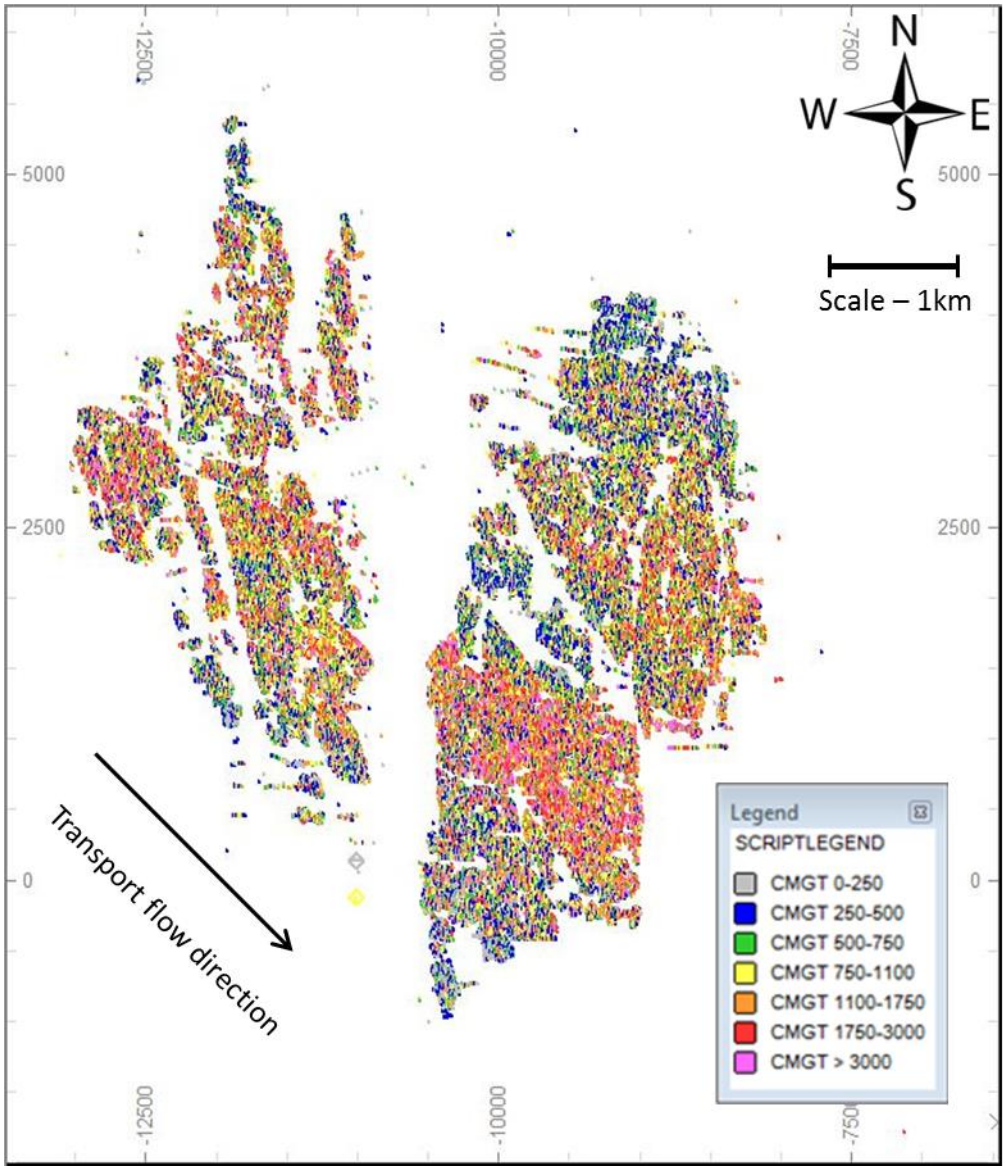


Figure 12: *Plan view of the cmg/t values*

From experience on the mine and the scrutiny of underground observations the author identified the direction of continuity in the cmg/t value to be broadly NW/SE which is parallel to the mineralisation transport direction. That is the gold values appear more continuous in the NW/SE direction. The main directions of the swaths were thus chosen so that the NW/SE swaths are parallel to this mineralisation transport direction. Swath strings were created in Datamine® for the directions, NW/SE and the perpendicular NE/SW direction. The "SELEXY" function of Datamine® was used to assign a zonal attribute to the data points to identify the swaths zones. This function

was also used to copy the data of each swath zone to a separate file. Statistical analyses were then carried out on the cmg/t data in each swath file and the resultant parameters for example the means and standard deviations were then used to create the value trend plots of the average cmg/t at the two scales chosen for the swaths; 500m and 90m.

In this analysis, the mine lease area was divided into equidistance 500m slices along the NW/SE direction as shown in Figure 13 and the mean and standard deviation of each slice was plotted along the slices as shown in Figure 14.

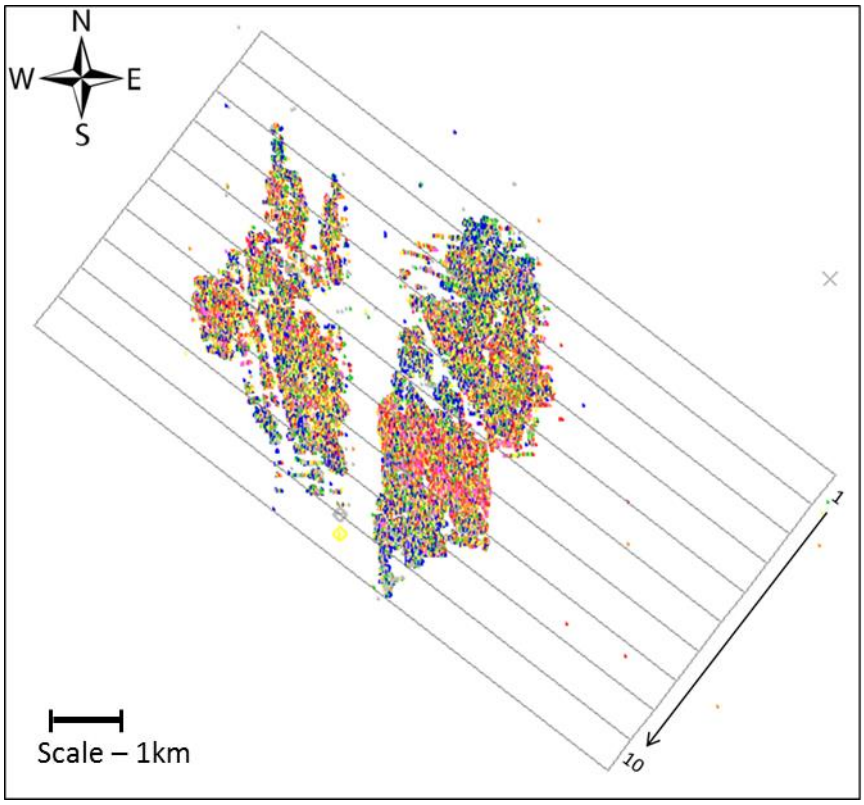


Figure 13: *cmg/t plot of the structured component of variability in a NW/SE direction – 500m swaths*

The change in average cmg/t values are reflected in Figure 14 moving from Swath1 to Swath10. From Figure 13 one can clearly see that the first and last swaths have fewer data than the other swaths in between and one must take this into consideration when interpreting the results. Swath1 has a low mean cmg/t value as there are mostly low (blue) values in the swath. The cmg/t

values will generally increase when moving towards Swath7 which has the highest mean cmg/t value. From Swath7 to Swath10 the cmg/t values again drop off. Another observation from this plot is that overall the average values vary roughly between 800cmg/t and 1600cmg/t, ignoring Swath1 and Swath10.

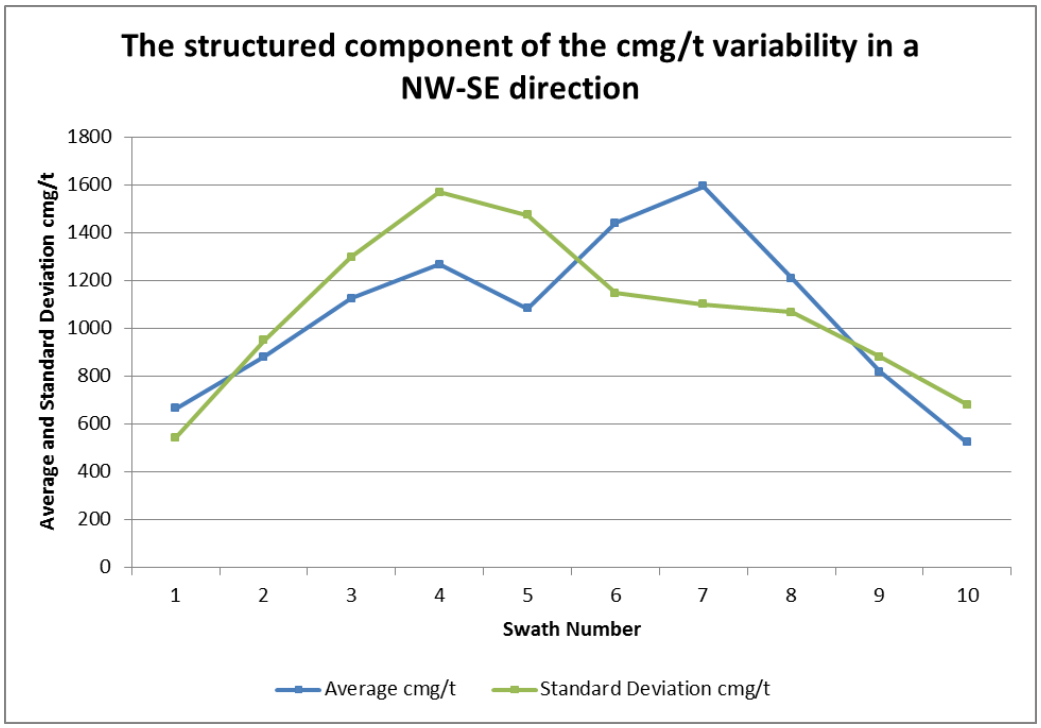


Figure 14: *Structured component of variability in a NW/SE direction – 500m swaths*

What is also interesting about the swath plot in Figure 14 is that it highlights the positive relationship between the averages and the standard deviations within the swaths. This is to be expected of the gold values which usually follow a positive skew distribution. The global cmg/t distribution previously shown in Figure 11 confirms a positively skewed shape.

The two anomalies are in Swath6 and Swath7 where the standard deviation did not increase with the increase in the mean, in fact it decreased. However, a closer look at Swath6 and more so Swath7 in Figure 13 reveal a solid high cmg/t portion (violet) with little other colours being present, which explains both the high cmg/t average and the relatively low standard deviation in those swaths because the values are similar.

Differences in results appear at different scales; the size of the swaths can be changed to look at different components of variability. For the structured component of variability, swaths were taken every 500 metres to reflect the large-scale cmg/t tendencies and provide a picture of the cmg/t distribution on a global scale. For the systematic component of variability, swaths were taken every 90 metres to reflect the low and high-grade changes on smaller scale than the global.

Smaller scale swaths at 90m in the NW to SE direction were considered to get a feel for the variability at closer intervals and are shown in Figure 15.

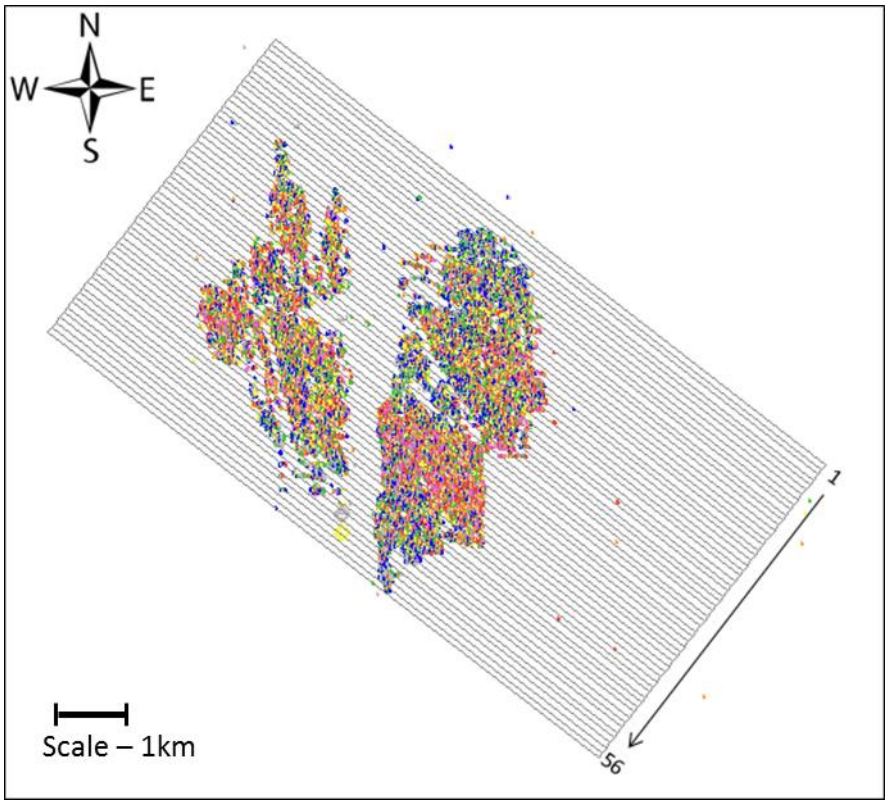


Figure 15: *cmg/t plot of the systematic component of variability in a NW/SE direction at 90m intervals*

The smaller scale displays the cmg/t changes at shorter distances and highlights the systematic changes in the average cmg/t values. Changes in the average and standard deviation in cmg/t moving from Swath1 to Swath56 can be observed in Figure 16. The smaller scale swaths (90m) have the same general trend as the larger scale but provide the detail to be taken into

account when considering geozone boundaries. As before the swaths on the edges, Swath1 to Swath5 and Swath50 to Swath 56, have significantly less data and the interpretations recognised this observation. The 90m swaths also highlights a range in average grades roughly 500cmg/t to 1800cmg/t as opposed to the previously observed range from 800 to 1600cmg/t for the 500m swaths. This observation also confirms the change of support effect. The larger support 500m swaths vary less than the smaller support 90m swaths. For the smaller scale the positive relationship between average and standard deviation is also more pronounced. With the reason for inverted relationship previously identified confirmed namely similar high mean values and a relatively low standard deviation around them.

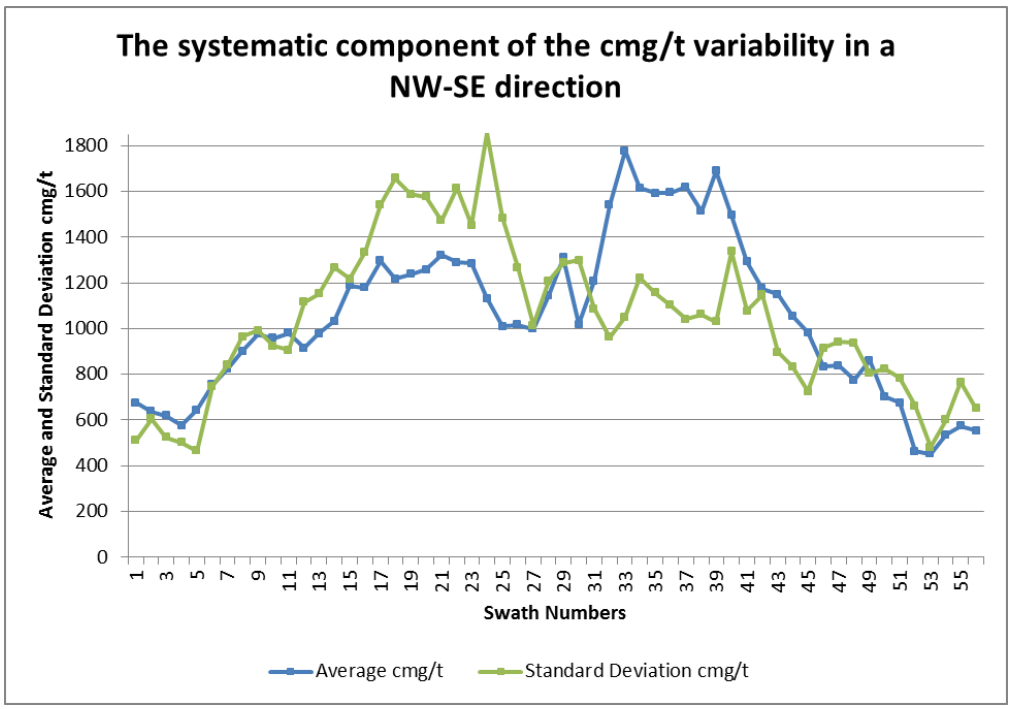


Figure 16: *Systematic component of variability in a NW/SE direction – 90m swaths*

The author wanted to know whether there is less variability in the one direction than in the other direction perpendicular to it, as this would assist with choosing the directions of the anisotropy of the variogram calculations and modelling to follow later. Swaths were taken every 500m for the NE to SW direction as shown in Figure 17 below.

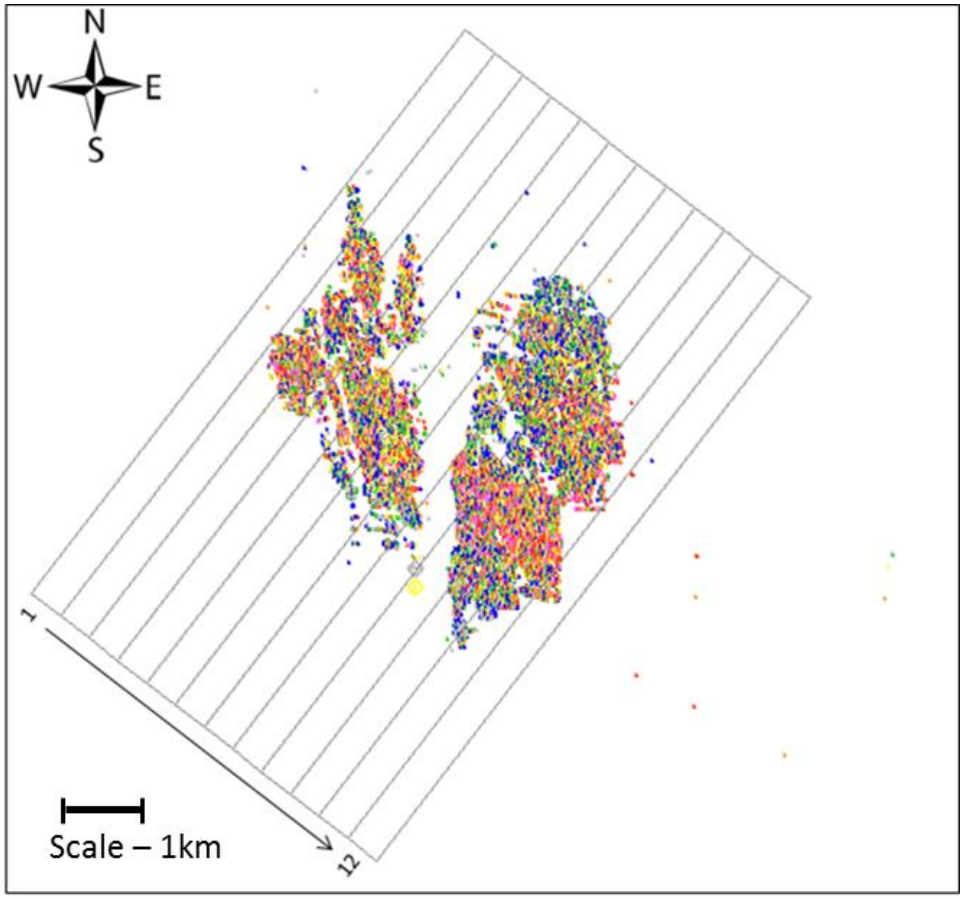


Figure 17: *cmg/t plot of the structured component of variability in a NE/SW direction – 500m swaths*

Changes in the average and standard deviations of the cmg/t values moving from Swath1 to Swath12 in the NE/SW direction are shown in Figure 18. Again, the first and last swaths (1 & 12) have significant less data and the interpretations took this into consideration.

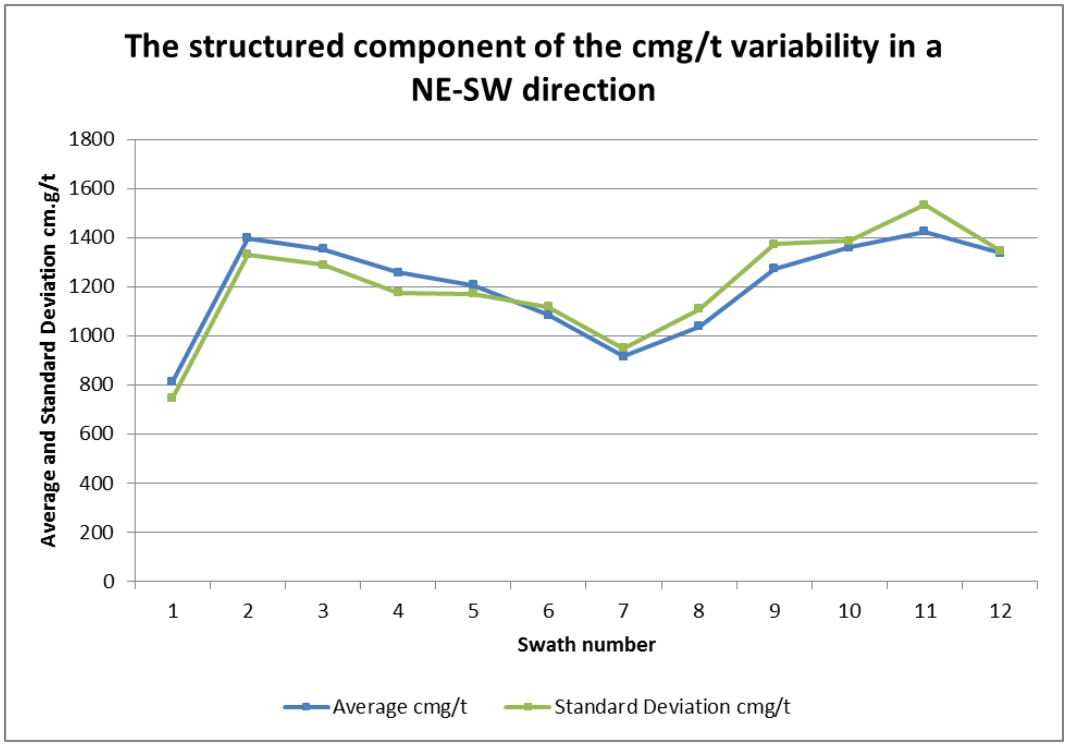


Figure 18: *Structured component of variability in a NE/SW direction – 500m swaths*

Recognising the previous observation and comparing Figures 14 & 18 one notices that the average cmg/t values in the swaths on a global scale in the NW/SE direction roughly vary between 800 and 1600cmg/t, and in the NE/SW direction between 950 and 1400cmg/t.

Swaths were also taken from the NE/SW direction on a smaller scale (every 90m) and are shown in Figure 19.

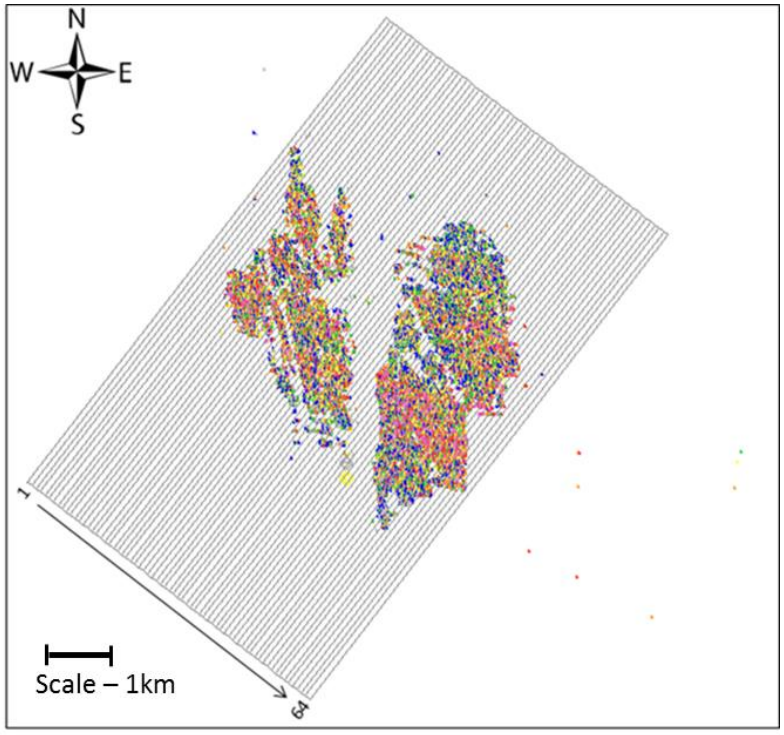


Figure 19: *cmg/t plot of the systematic component of variability in a NE/SW direction – 90m swaths*

Changes in the average cmg/t moving from swath 1 to 64 are shown in Figure 20.

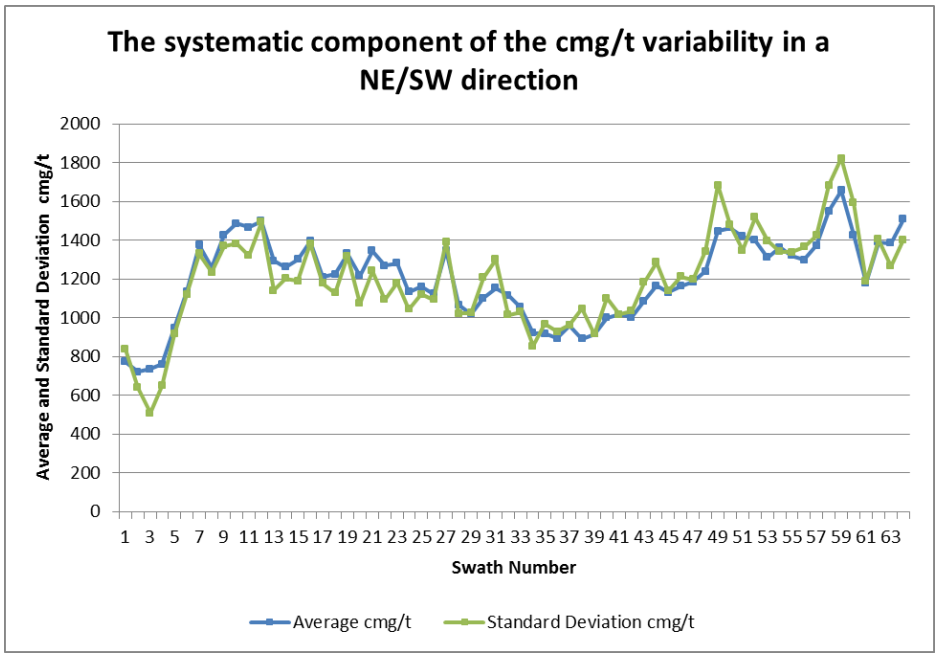


Figure 20: *Systematic component of variability in a NE/SW direction – 90m swaths*

As before those swaths at the beginning and end, Swath1 to Swath 7 and Swath 60 to Swath 64 has significantly less data and this was taken into consideration. The range of cmg/t values is wider in the NW/SE direction than in the NE/SW direction.

3.2 Facies Plan

A facies plan identifies zones that should have geological homogeneity and consistent lithology, alteration and grades. The lease area of Tshepong mine consists of 3 different facies types namely BCF, LF and MF. The mine is dominated by BCF which is, a multiple scour polymictic package. The LF on the other hand is a clean oligomictic conglomerate with a single pebble lag. The contact between the BCF and the LF is not distinct but rather an interfingering contact over a broader area.

3.2.1 Facies history

The initial facies plan of Tshepong mine was created by means of unconstrained global kriging routines to define zones with different gold distributions. With the limited information that was available in 1997 it was established that the boundary between the LF and the BCF strikes in a NW to SE direction as displayed in Figure 21.

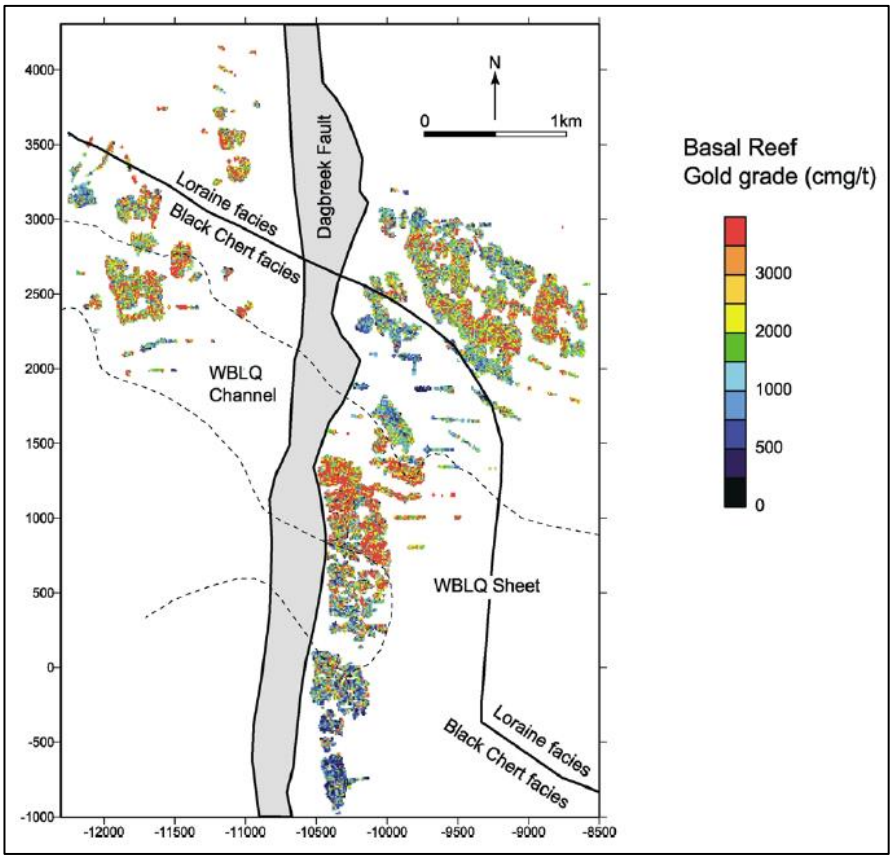


Figure 21: *Distribution of gold superimposed on the Basal reef facies type boundaries* (Jolley et al., 2004) .

These zones were extrapolated into un-mined areas in the Tshepong lease area. Many of these un-mined areas have since been mined-out and a lot of new information became available to assist with the updating of the facies plan.

3.2.2 Facies descriptions

The mine-scale stratigraphy of the Khaki Shale, Waxy Brown Leader Quartzites and the Basal reef is displayed in Figure 4. In the Orange Free State, the Basal reef is regarded as the principal carrier of gold. Directly on the UF1 quartzites lies the BCF and the LF of the Basal reef.

Black Chert Facies:

The BCF dominates the lease area. The make-up of the Black Chert facies contains up to 25% black chert clasts. The Loraine facies can also contain black chert clasts but only up to 5% of the make up (Freeman et al., 1999).

This is a small difference between the facies and can complicate the identification of a facies from boreholes but is relatively easy identified from mined exposures. Looking at boreholes it is possible that one can sometimes mistake Loraine facies for Black Chert facies.

Lorraine Facies:

The Loraine facies comprises of a single, clean, oligomictic pebble lag. In the northern parts of the mine the quartzite above the reef is cleaner and better sorted. The reef width is also thinner in the northern part of the zone.

Melkkraal Facies:

The Melkkraal facies is a small pebble conglomerate with mostly smoky quartz pebbles that makes up about 30% of the package. Siliceous chert are scattered in the matrix and makes up approximately 5% of the matrix. This is very low-grade facies with no visible carbon.

3.2.3 Updated facies plan

To ensure that geological homogeneity exists within each geozone it was important to look at how the facies corresponds with the geozones. In this research an up to date facies plan was created and compared with the existing geozones. The updating of the facies plan entailed the analysis of conglomerate descriptions of the mined-out areas, boreholes and geological structures and the updated facies plan is displayed in Figure 22.

The updated boundary between the BCF and the LF runs in a NW/SE direction and mostly falls on dykes that strike in the same direction. This boundary has the same strike as it had on previous facies plans, but the position thereof has been moved to a more accurate position. This position

has been derived from the additional information gathered during the significant amount of mining that has taken place on either side of the boundary, since the initial interpretation and extrapolation.

At a certain point the boundary turns south and is an indication that a small amount of lateral movement might have occurred along the Dagbreek fault. A small triangular area where the LF overlaps the BCF occurs in the middle of GZ5.

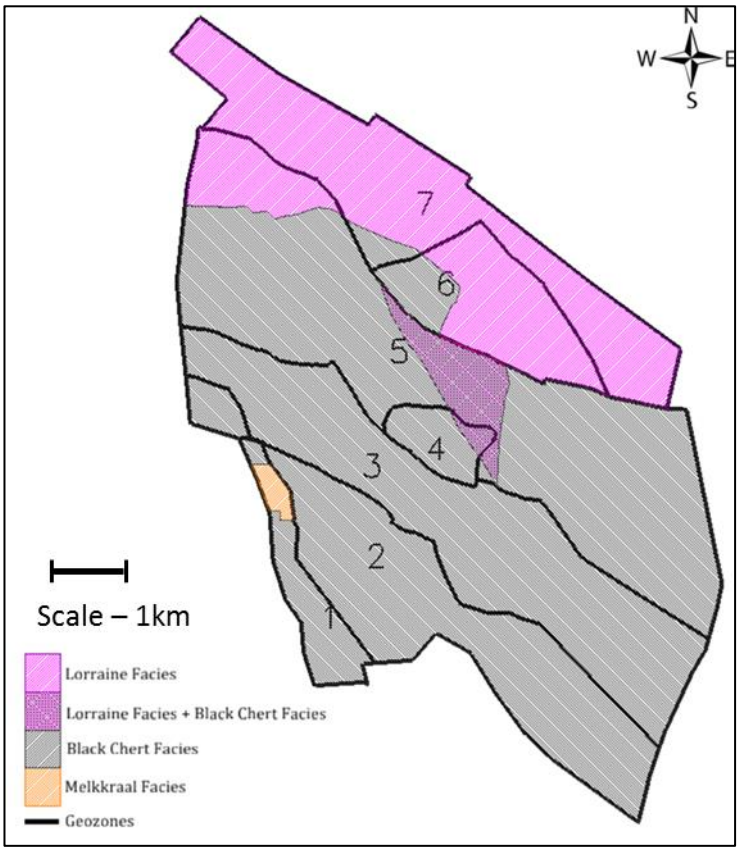


Figure 22*: Tshepong Mine 2017 Updated facies plan with the old geozones*

Comparing the updated facies plan to the geozones, highlighted areas where the geozones and the facies zones did not match. Geozones 5, 6 and 7 contains both BCF and LF within the same geozones.

The geozone delineations affects each technical service department on the mine differently. The person responsible for the Mineral Resource and Mineral Reserve blocks prefers geozones to fall within large geological

structures because if a geozone falls next to the structure, he must split the Mineral Resource and Mineral Reserve blocks on the geozone boundaries and geological structure. This results in a big increase in the number of blocks that has to be maintained.

The geostatistician prefers the geozones to represent areas of stationarity and geostatistical homogeneity. The geologist prefers geozones that are based on facies and structure. The Ore Reserve manager knows that he is not allowed to schedule mining in areas that are below cut-off, so if the geozones are based on grade, Mineral Resource and Mineral Reserve blocks that falls in low-grade geozones will not be mineable. If the geozone contains low and high-grade data, the global mean will be higher and will assist to ensure that lower grade areas are above cut-off. If one delineates low-grade areas, all the Mineral Resource and Mineral Reserve blocks in the domain will be low-grade (below cut-off) and would then be excluded from the LoM scheduling.

Using grade alone to determine domain boundaries can be risky because it could result in over estimation in the domain and under-estimation in the next domain (Glacken & Snowden, 2001).

This research will improve the confidence attached to the Tshepong Mineral Resource model. Keeping the purpose of the research in mind, the most important factors to consider for domains are:

- Geological homogeneity
- Geostatistical homogeneity
- Stationarity
- Lithology, alteration, grades
- Mineralisation
- Data within the zone must have the same characteristics.
- Trend

3.3 Geological Homogeneity

Geological homogeneity is where the mineralization, lithology, alteration and grades are similar within a geozone. There exists an area where the BCF overlaps with the LF. The BCF always occurs above the LF.

The MF only occurs in one small area of the mine, the development to the south of the MF confirms that the MF does not extend any further to the south in GZ1.

Changes in the geozones are shown in Figure 23, on the left is the new geozone delineation. The previous geozones were plotted on the updated facies plan to highlight the areas where geozones contained more than one facies (right hand side of Figure 23).

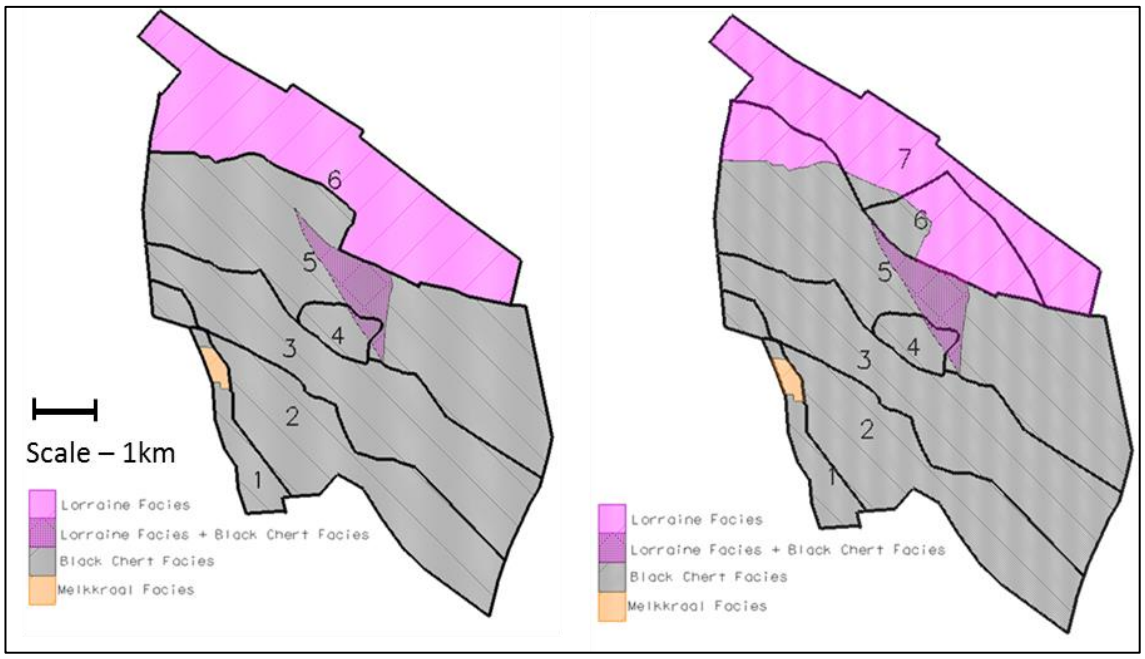


Figure 23: *New geozones with updated facies plan (left), old geozones with updated facies plan (right)*

The statistical analysis in section 3.3.1 below shows how significant the differences in the cmg/t populations of the facies are, and why the geozones were changed. The northern border of GZ5 was moved to match the border between the BCF and the LF. GZ7 only had a few LF boreholes and was merged with GZ6 to create a zone that only contains LF.

3.3.1 Statistical Analysis of each Facies

The summary statistics of the facies data (cmg/t) is presented in Table 3. By comparing the different statistical parameters, one realises how different the populations are, and this emphasises how important it is to use facies boundaries for domaining purposes.

Table 3: *Summary statistics of cmg/t per facies*

	BCF	BCF + LF	LF	MF
Mean	1325	1070	847	620
Standard Error	2.987	6.529	4.179	11.505
Median	958.65	823	641.02	457.785
Mode	523	584	500	455.98
Standard Deviation	1368.6	956.7	812.0	591.2
Sample Variance	1873138.5	915247.8	659282.6	349459.9
Coefficient of Variation	1.03	0.89	0.96	0.95
Kurtosis	66.0	30.9	45.2	15.7
Skewness	4.7	3.4	4.3	2.9
Range	62313.0	23209.0	22652.1	7376.0
Minimum	0.01	1	0.28	0.7
Maximum	62313	23210	22652.35	7376.67
Sum	278062720.6	22980220.1	31988944	1636190.51
Count	209932	21469	37753	2640

A statistical analysis had been carried out within each defined domain using the data that falls within each facie zone. The amount of data available for each facies differs, therefore to compare the facie populations with one another the relative frequencies distributions were calculated and are presented in Figure 24.

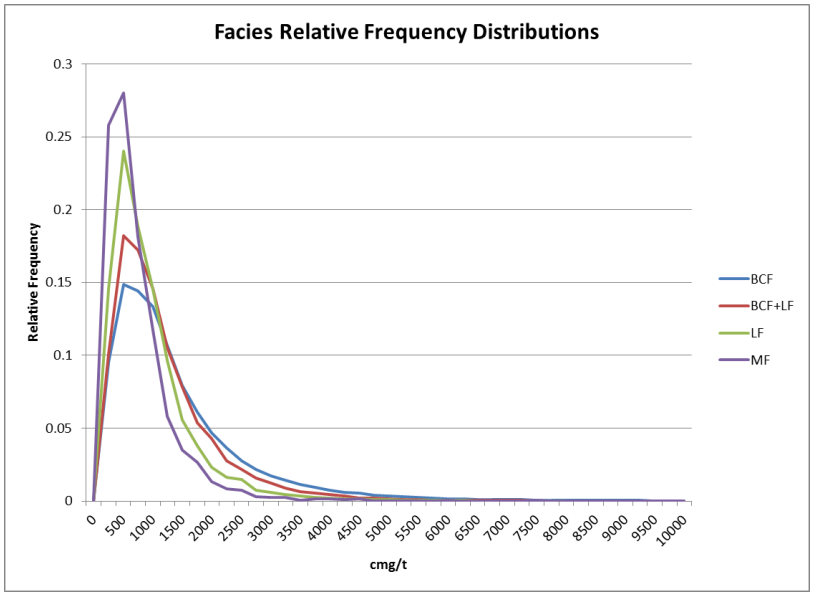


Figure 24: *Relative frequency distribution of the cmg/t values of each facies*

The shapes of the facies distributions are all positive skew, confirmed by the skewness varying from 2.9 for the MF to 4.7 for the BCF. There are however, unique characteristics in the individual facies that are described below.

The MF peaks and flattens very quickly and indicates that the chances for high values are very low.

LF takes longer to peak and the tail of the distribution stays above the MF which indicates that there are higher grades in the LF than the MF.

The peak of the area where the LF overlays the BCF is a lot lower that the peaks of the MF and the LF and the tail of the distribution takes longer to flatten. This indicates that there are generally higher grades in the overlay area than the MF and the LF.

BCF has the lowest peak and the tail of the distribution takes the longest to flatten out. Chances for high grades are the best in the BCF.

Cumulative relative frequency distributions were compared to see how the facies differs on the 75th percentile (Figure 25).

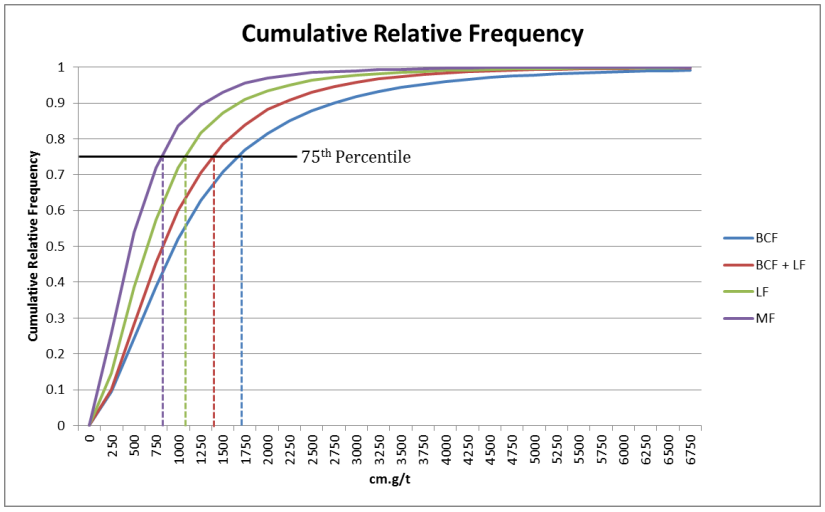


Figure 25: *Cumulative relative frequency distribution of each of the facies with the 75th percentile identified*

The graph shows that the four facies cmg/t populations are in fact very different in shape. The right (blue) distribution (BCF) on this plot indicates that it has the highest cmg/t values. The left (purple) distribution (MF) on this plot indicates that it has the lowest cmg/t values.

The MF rises steeply to about 875cmg/t, where after the rate of change decreases and the tail flattens fairly quickly. 75% of the grades of the MF are lower than 805cmg/t.

The LF has better gold grades than the MF. The LF rises and flattens off slower than the MF. 75% of the grades of the LF are lower than 1060cmg/t.

The area where the LF overlays the BCF has better gold grades than the LF. The line rises and flattens off slower than the LF. 75% of the grades in this area are lower than 1378cmg/t. The line falls between the BCF and the LF which highlights that this is a mix between the two facies. The BCF line plots below the other facies in the cumulative distribution confirming the high cmg/t values, with 75% of its gold grades below 1672cmg/t.

The means, median and standard deviations for the facies are graphically summarised in Figure 26. This comparison shows that the statistics within the facies are variable, that the distributions are positive skew and

additionally highlights the importance of considering facies in the domaining process.

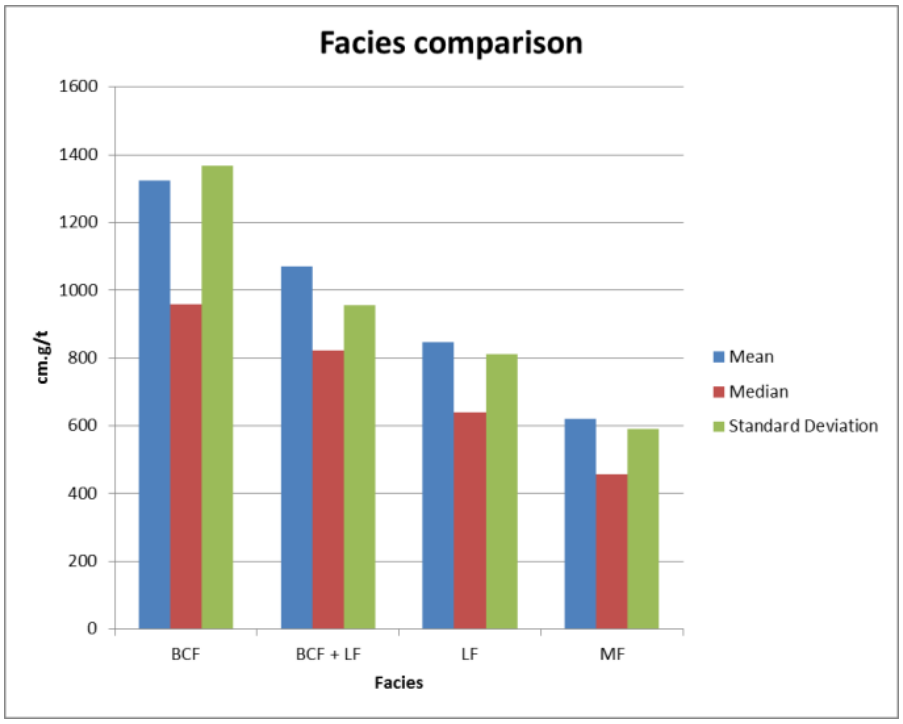


Figure 26: *Facies comparison between mean, standard deviation and median*

4 EXPLORATORY DATA ANALYSES

The statistical characteristics of a mineral deposit reveal the underlying distribution of the population (Armstrong, 1998).

In this chapter consideration is given to the statistical analyses and graphical presentation of the cmg/t values, and the interpretation of the results thereof.

The domaining that followed from the global EDA process assisted with the selection of relevant data for each geozone. A further EDA was conducted on each geozone after completing the domaining process to ensure that the domaining achieved its objective. This exercise and process increased the researcher's understanding of the characteristics of the data.

EDA was done separately for each geozone and included the following:

- Data validation through basic statistics
- Histograms
- Central tendencies, average, median and mode
- Spread, variance, standard deviation & coefficient of variation
- Shape of the distribution – skewness and kurtosis
- Cumulative relative frequencies
- Scatter grams
- Percentiles analysis

This process was used to identify outliers, abnormal values and non-homogeneous data. This analysis is time-consuming, but it is a vital step; experience has taught the researcher that it is better to spend more time on EDA than to recommence a study. An additional benefit is that a better understanding of the data and deposit is achieved in the process, enhancing ore body knowledge.

Armstrong, M.A. (1998), recommends that should more than one peak be visible on the histogram, the data must be rechecked to ensure that the data comes from a homogenous population.

In this research the histograms were examined to check for outliers and identify whether more than one mode is present in the dataset.

The data was plotted spatially to check for spurious coordinates and for abnormal values that might not have been picked up in the histograms. The data was also checked for numerical errors.

4.1 Statistical Analysis per Geozone

The goal of data analysis is to get familiarised with the sample data and to find any anomalies or outliers. Sample statistics provides information about the samples and is a stepping stone to the understanding of the population per geozone.

The summary statistics of the input data (cmg/t) is presented in Table 4.

Table 4: *Summary statistics of cmg/t per geozone*

Geozone	1	2	3	4	5	6
Mean (cmg/t)	642	974	1580	693	1232	847
Standard Error	10.6	5.6	5.0	5.9	3.9	4.2
Median (cmg/t)	480	673	1145	522	950	641
Standard Deviation (cmg/t)	596	1110	1576	699	1113	812
Sample Variance $(cmg/t)^2$	355226.0	1232523.8	2482577.0	488883.7	1238383.2	659282.6
Coefficient of Variation	0.93	1.14	1.00	1.01	0.90	0.96
Kurtosis	15.6	66.4	62.9	56.9	46.5	45.2
Skewness	2.9	5.1	4.6	5.3	4.1	4.3
Range (cmg/t)	7376	32510	62313	13065	33878	22652
Minimum (cmg/t)	0.7	1.0	0.1	1.0	0.0	0.3
Maximum (cmg/t)	7377	32511	62313	13066	33878	22652
Count	3164	39499	97407	13843	80128	37753

The relative histograms in Figure 27 depict strong positive skewness, which is confirmed by the high coefficients of variation and the means that are higher than the medians in the different geozones.

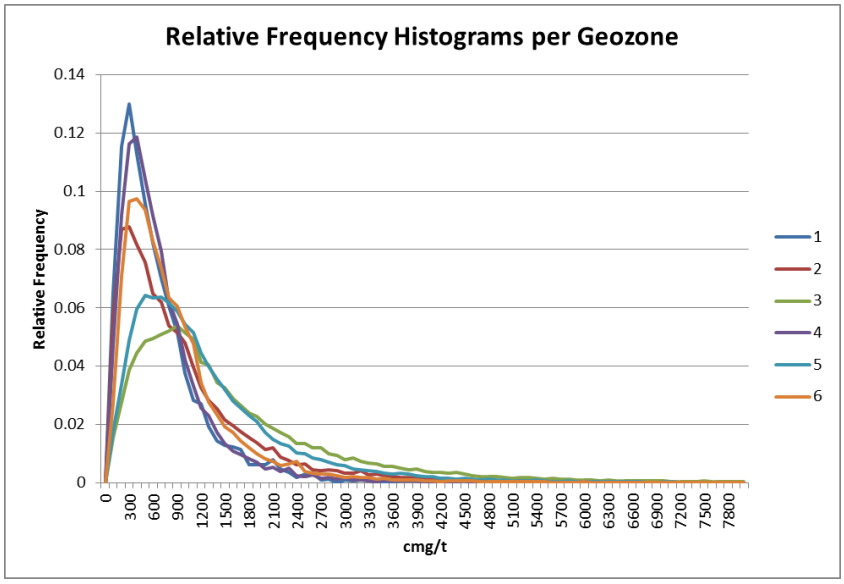


Figure 27: *Relative frequency histograms (cmg/t) per geozone*

In Figure 27 the relative histograms of GZ1 and GZ4 are remarkably similar. These two geozones are mainly low-grade zones with sporadic higher cmg/t values. The histograms of GZ3 and GZ4 are similar in shape regarding the lower end of the distributions and so are the histograms of GZ2 and GZ6. The overall conclusion is that whilst the statistical populations have similar distribution shapes, they are not the same and there are subtle differences, confirming that the domaining process was necessary and successful.

Figure 28 below compares the cumulative relative frequencies per geozone.

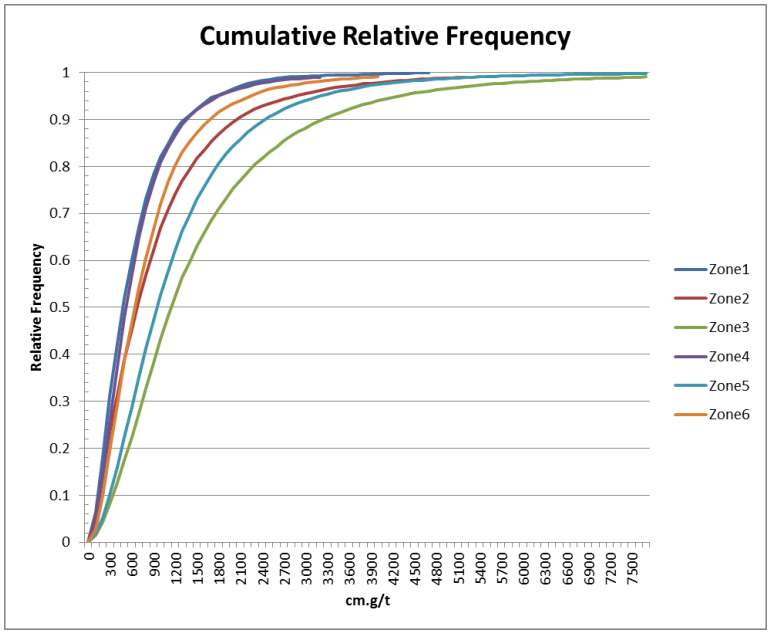


Figure 28: *Cumulative relative frequency per geozone*

GZ3 plots below the other geozones in the cumulative distribution confirming the high cmg/t values. The differences in statistical distributions of the geozones (Table 4) are further enhanced by the graph in Figure 29, and the similarity between GZ1 and GZ4 is again emphasised. Despite this similarity, the spatial position of these two geozones prevented them from being combined into one geozone.

In Figure 29 below is the statistical comparison of the cmg/t value per geozone.

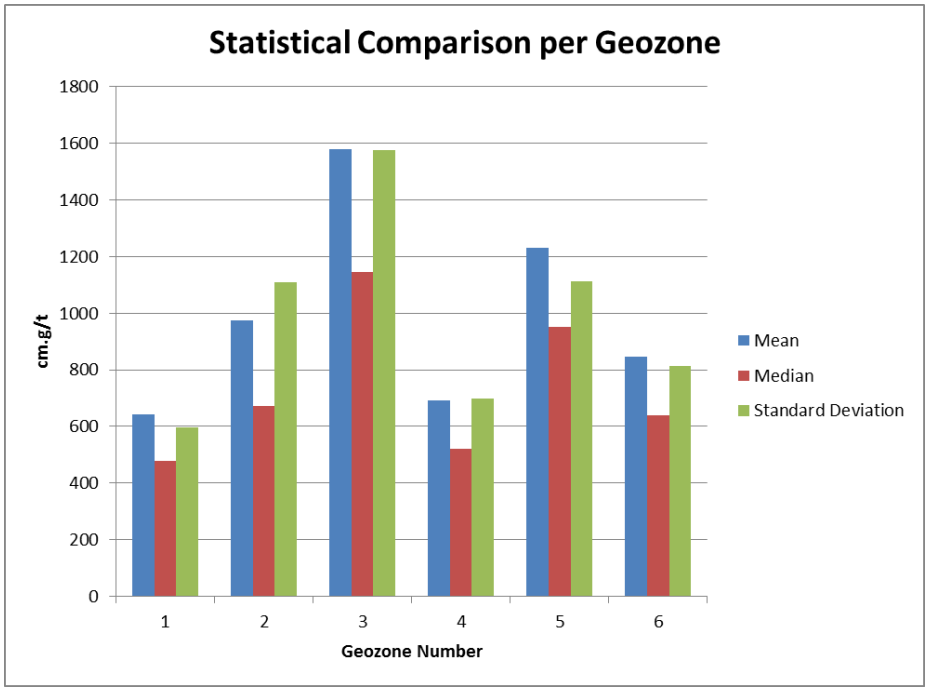


Figure 29: *Statistical comparison per geozone*

As the mean of the geozones increase, so does the standard deviation. The higher the mean, the more deviation can be expected from the mean. The higher the mean, the more erratic values can be expected. The median is less than the mean in all cases, again confirming the skew nature of the gold values. It is this positive skewness that accounts for the positive relationship between the mean and the variance that is often observed; the higher the mean the larger the variance.

If values fluctuate widely, the chances of accurate local estimations are not good. Any estimation method will benefit from data with low variability and will suffer from data with high variability (Isaaks & Srivastava, 1989).

4.2 Capping of High Grades

A capping technique that truncates data to a chosen capped grade value mitigates the influence of high values in the tail of the distribution. Capping reduces the risk of unrealistic high-grade estimates. The data was capped on the 95th percentile value per geozone and the capped values are shown in Table 5. All the values greater than the 95th percentile of a specific geozone were capped. The reduction in the sampling variance is evident and the

capped data eases the variogram modelling process and assists in the avoidance of potential over-estimation due to extreme sample values.

Table 5: *Capped value per geozone and the variance of the capped and uncapped data*

Geozone number	95th percentile value (cmg/t)	Variance of the capped data $(cmg/t)^2$	Variance of the uncapped data $(cmg/t)^2$
1	1748	211 122	360 954
2	2867	559 317	1 172 169
3	4304	1 268 967	2 537 765
4	1806	213 752	490 428
5	3206	681 501	1 257 567
6	2256	330 036	702 825

The statistical comparison of cmg/t values between the capped and the uncapped historical data is in Table 6 below:

Table 6: *A statistical comparison of cmg/t values between the capped and the uncapped historical data base files (Phakisa data is excluded in this comparison).*

	Capped data base	Uncapped data base
Mean	1216	1292
Median	947	947
Mode	4207	523
Standard Deviation	968	1312
Coefficient of Variation	0.79	1.02
Minimum	0.4	0.4
Maximum	5876	62312

The mean of the uncapped data base is higher than the capped data base due to the removal of hard-coded capping. The mode of the capped data base was

the value that was capped the most times, the mode of the uncapped data base is the current mode. The change in coefficient of variation indicates that the shape of the distribution of the sampling data changed after removing the hard-coded capped values.

5 ANALYSIS OF THE SPATIAL CONTINUITY OF THE MINERALISATION

The spatial structure of a variogram is inherent of the mineralisation in the geozone and can be defined by the type of variogram, the range and the nugget effect (Krige, 1996B).

5.1 The Variogram

The variogram is a function of the differences and the distance h, that the samples are apart. The experimental variogram is calculated from:

$$\gamma(h) = \frac{\sum(Z(x) - Z(x+h))^2}{2N_h}$$

Equation 1

Where γ(h) is the experimental variogram value at distance h, Z(x) is the sample value at location (x), Z(x+h) is the sample value at location (x+h) and h is a vector determining the distance and direction between the samples. The h value is referred to as the lag distance between the samples and N_h is the number of pairs in that direction and that distance between the samples.

The variogram is the most traditional choice when it comes to analysing spatial continuity. To analyse spatial continuity is not easy and one must be prepared to repeat steps several times (Isaaks & Srivastava, 1989).

A variogram indicates how different values become as the distance between the data points increases. It is important to calculate variograms in at least four directions to ensure the detection of anisotropy. The more directions considered the greater the possibility of identifying the direction of highest continuity on the mineralisation.

Tshepong is mining a tabular ore body and there is no option of a down-hole variogram. There is no sense in calculating the downhole variogram to get an estimation of the nugget effect however, in this research an additional point

was added at 0.08m (very close to the origin) to the experimental semi-variogram by making use of the semi-variance calculation of the $\gamma^*(0.08)$ in the above equation where Z(x) is the grade of the first sample and Z(x+h) is the grade of the check sample that is 8cm apart from it, and N_h is the number of check sample pairs. This additional point increased the confidence in the estimation of the nugget effect and is discussed in section 5.2.1.

5.1.1 Modelling of the variogram

When fitting the models, it is important to look at the nugget effect, the slope at the origin, the sills and ranges, the total sill and anisotropies (Armstrong, 1998). The sill is usually modelled to where the variogram stabilises, and fitting of the range is assessed visually (Clark, 1979).

Isotropic and anisotropic variograms:

In an isotropic variogram the range is the same in all directions. This means that there is no preferred direction of spatial continuity. If the sill of the variogram remains the same in different directions and only the ranges change, we have geometric anisotropy. If the sills are different in different directions we have zonal anisotropy which usually occurs if one direction is across a channel with high variability and the other direction is along a channel of high continuity (McKillup & Darby Dyar, 2010).

Anisotropy refers to the case where the range is longer in one specific direction than in the other usually perpendicular direction. Thus, an anisotropic variogram indicates the preferred direction of spatial continuity.

Omnidirectional variograms:

An omnidirectional variogram, on the other hand is the variogram that is calculated as the average of all the directional variograms and can be used to give a general indication of the global spatial variability in a geozone (McKillup & Darby Dyar, 2010).

It is sometimes possible to identify anomalous anisotropy purely due to the layout of the sampling grid and that, one should be aware of this possibility for creating artificial anisotropy. Those are cases where sampling is more continuous along strike for example (say every 5m), than across strike (say every 10m to 15m). The possibility of this scenario is however very unlikely for Tshepong mine since sampling layout is 5m along and 5m across strike.

5.2 The Nugget Effect – Spatial continuity near the origin of the variogram

In this research the nugget effect estimate for each variogram has been improved by adding the semi-variance calculated from the check samples at the bottom of each sample section as previously explained.

Figure 30 shows the distribution of g/t differences of the check samples for each geozone at a very small h (lag distance) of 8cm. The averages of differences between the two check samples are all close to zero, the distributions appear symmetrical around zero, but the spreads are relatively wide. This observation on spread is in line with the nuggety nature of the Wits gold mineralisation; in fact, that closed spaced variability is the where the term nugget effect comes from. From the preceding observations one can conclude that the semi-variance of the check samples can be used as an additional experimental semi-variogram value $\gamma(0.08)$ in the individual geozones. Details of this will be discussed in the section to follow.

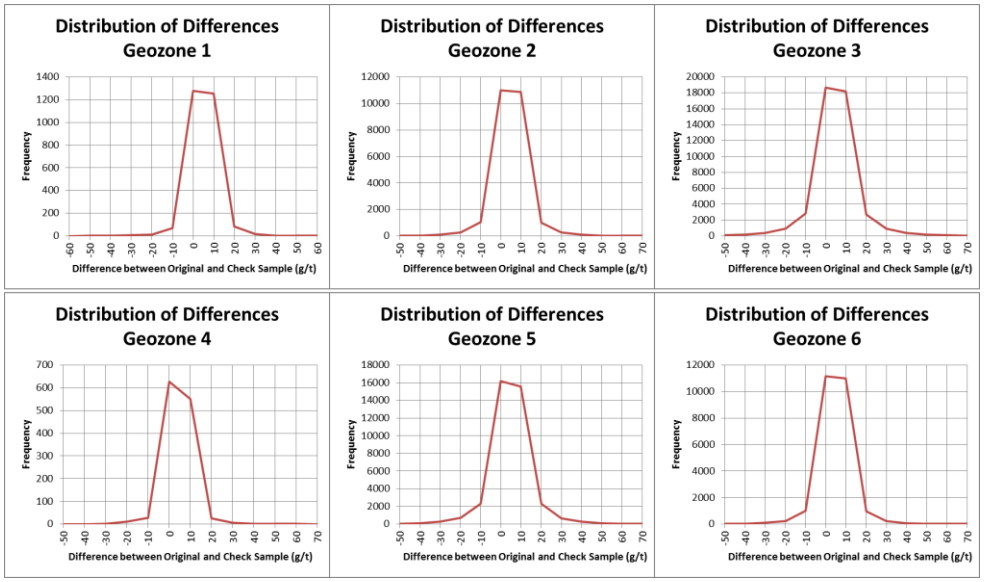


Figure 30: *Distribution of differences in grade between the check samples per geozone*

5.2.1 Nugget effect

The vertical jump from the origin of the variogram to the first variogram value at a very small distance is called the nugget effect. It is an inherent random variability that ores have at distance close to zero.

The sampling error is one contributor to the nugget effect (Pitard, 1993)

When fitting a variogram model to the experimental variogram it is crucial to take care with the estimate of the nugget effect, because this parameter has the biggest influence on the kriging results. The value of a semi-variogram ($\gamma(h)$) at h = 0 is zero, but factors like short scale variability and sampling errors causes sample values separated by only a few centimetres to be very dissimilar and causes the discontinuity at the origin of the variogram.

Improving the estimation of the nugget effect will therefore have a direct benefit on the confidence of the Mineral Resource estimation.

The nugget is usually estimated by extrapolating from the first three to four semi-variogram points at larger lags towards the variogram axis (y-axis), at the origin the lag is zero and the variogram value is zero. Figure 31 below is an example of a semi-variogram. Point A is 5m from the origin. To estimate the nugget effect, one would normally draw a straight line from point B

through point A (the point with the most number of pairs would get more weight) until the line intersects the variogram axis, and this would be used to estimate the nugget. In this research, an additional point (point X) was added to the experimental semi-variogram by using the semi-variance of the check sample data.

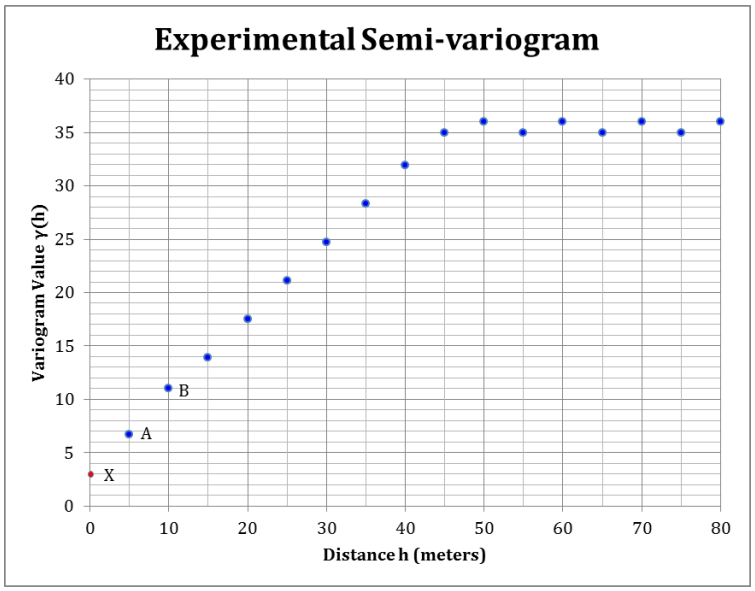


Figure 31: *Experimental semi-variogram*

During the underground collection of samples at Tshepong mine, the bottom contact of the Basal reef, two samples are chipped right next to each other (check samples) as shown in Figure 32. Both these samples are assayed and recorded in the sampling data base, and it is possible to extract these values from the database with a tailored data query. The centre points of these samples are 8cm apart and the variance of these samples will provide estimates for variogram values at a lag of 0.08m (point X in Figure 31), close to the origin of the variogram.

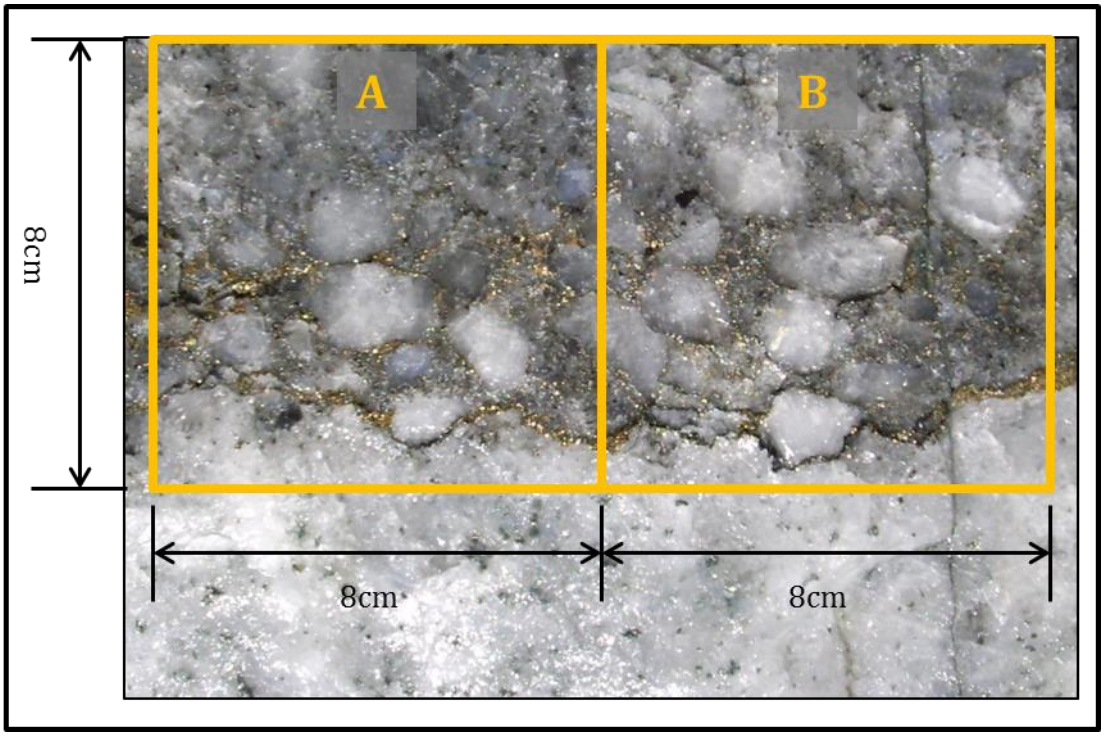


Figure 32: *Illustration of check samples*

As many check sample pairs are available, the confidence in this estimated point on the variogram will be an improvement in the quality of the nugget effect estimate.

Table 7 shows the total sample variance, the covariogram value and the standardised γ value at 0.08m of the check sample pairs in mg/t (meter grams per ton) values. The original g/t values were converted to mg/t by multiplying it by 8cm (all check samples are 8cm high) to get to cmg/t and were then divided by 100cm to get to mg/t values.

The reason for this was to reduce the magnitude of the variances in cmg/t and was done for the representation of the information in Table 7. This action has no impact on the results because a variance is unaffected by a constant value. The first theorem of a variance states that, $Var(cX) = c^2\ var(X)$, where c is a constant and X is the variable (Dohm, 2015 B).

Table 7: *Variance, covariogram, and standardised γ at 0.08m of the mg/t values of the check samples per geozone*

Geozone	C = Total sample variance (mg/t)²	C(h) = COV (original, check) = Covariogram	γ(h) = Sill - C(h) h =0.08m Variogram	No of check sample pairs	Standardised γ(0.08m)
1	53.243	31.278	21.965	2729	0.413
2	210.739	120.475	90.265	24907	0.428
3	380.261	261.663	118.598	45997	0.312
4	50.805	32.430	18.375	1256	0.362
5	507.083	342.015	165.069	38947	0.326
6	230.350	142.165	88.185	43573	0.383

Table 7 represents the information at sample scale and not at block scale. The last column of Table 7 is the recommended estimate to be used in the estimation of the nugget effect of the standardised sample variogram models.

For future updates with new data the methodology discussed above should be applied to calculate these nugget effect estimates. The scatter plots of the check sample pairs in the different geozones are shown in Figure 33 below.

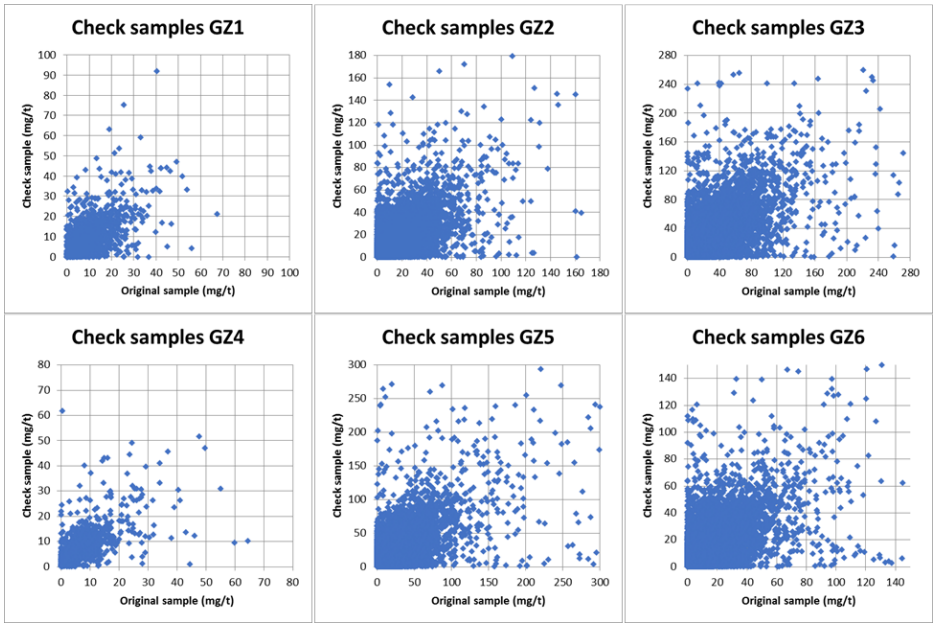


Figure 33: *Scatter plots of the original sample against the check sample (mg/t)*

5.2.2 Regularisation

At Tshepong Mine the Mineral Resources are estimated as Measured (30m x 30m) blocks, Indicated (60m x 60m) blocks and Inferred (120m x 120m)

blocks. Final Mineral Resource classifications are based on the halos that are generated from the respective block models, this is discussed in section 9.

Figure 34 below shows the regularised data that is used for the Indicated and Inferred block models.

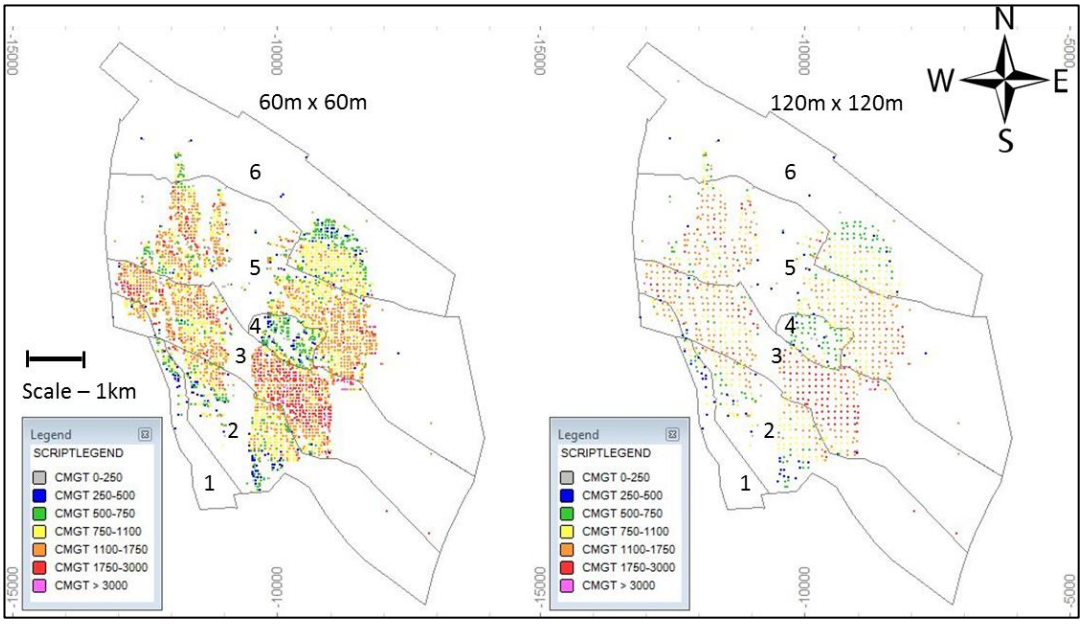


Figure 34: *Colour coded plots of the 60m x 60m (left) and 120m x 120m (right) regularised cmg/t data*

A summary of the statistics of the raw (not regularised) data that was used for the experimental semi-variograms of the 30m x 30m block models are shown in Table 4. The summary of the statistics of the regularised data that was used for the 60m x 60m and the 120m x 120m block models are shown in Table 8 below.

Table 8: *Summary statistics of the 60m x 60m and the 120m x 120m regularised data*

	60 x 60 Blocks						120m x 120m Blocks					
Geozone	1	2	3	4	5	6	1	2	3	4	5	6
Mean	631	959	1543	702	1202	813	610	936	1528	724	1204	788
Standard Error	34.4	19.1	19.1	20.7	13.3	14.2	43.9	28.9	32.7	43.7	21.8	23.4
Median	631	920	1456	646	1189	809	648	918	1433	663	1202	803
Standard Deviation	271	450	654	309	463	297	236	396	611	371	434	279
Sample Variance	73236.6	202672.6	427534.9	95356.0	213958.6	88490.7	55914.1	15664.1	373458.3	137306.2	188382.8	77971.5
Coefficient of Variation	0.43	0.47	0.42	0.44	0.38	0.37	0.39	0.42	0.40	0.51	0.36	0.35
Kurtosis	1.9	6.6	18.1	16.5	9.6	6.0	0.4	1.8	21.7	23.8	12.1	0.6
Skewness	0.9	1.5	2.4	2.6	1.4	1.1	0.6	0.6	2.8	3.8	1.6	0.0
Range	1391	3754	8726	3053	5377	2791	967	2720	6936	3053	4657	1474
Minimum	215	21	6	56	5	77	297	20	239	56	81	80
Maximum	1606	3775	8732	3108	5382	2868	1264	2740	7175	3108	4738	1554
Sum	39134	531471	1800194	155869	1462126	355322	17681	174982	534854	52117	475702	111842
Count	62	554	1167	222	1216	437	29	187	350	72	395	142

In all the geozones the change of support effect is reflected by the reduction of the CV (Coefficient of Variation) from sampling data, to the 60m x 60m blocks and then reducing further for the 120m x 120m blocks with an anomaly for GZ4 where it increased. The increase in the standard deviation and the mean at this block size and the "small" areal extent of GZ4 could account for this.

5.2.3 Variogram contour maps of the spatial variability - Varmaps

In this research, the value trend analyses reported on in section 3.1 provide an indication of the directions of most and least spatial continuity, the transport flow direction, together with the facies interpretations gave an indication of the direction of spatial continuity of the cmg/t values.

An omnidirectional variogram with a large directional tolerance is a good start to finding the direction of minimum and maximum continuity. By creating a contour map of the omnidirectional variogram one will produce a graphical representation of the continuity that can be used to identify the directions of interest (Isaaks & Srivastava, 1989).

The variogram map is a two-dimensional plot of the sample semi-variogram for all experimentally available separation vectors (Deutsch & Schnetzler, 2009).

The varmap is usually colour coded or presented by variance contours highlighting anisotropies and low and high variances. It also shows that as the distance increases the correlation between points decreases. It allows one to see general trends and the direction of continuity (Dohm, 2015 B).

The colour coded plot of the 60m x 60m and the 120m x 120m regularised data shown in Figure 34 above assisted with the interpretation of the variogram contour maps that follows.

The contour maps in Figure 35 were created for each geozone, the green lines indicate variogram values in the selected directions and the lag distances.

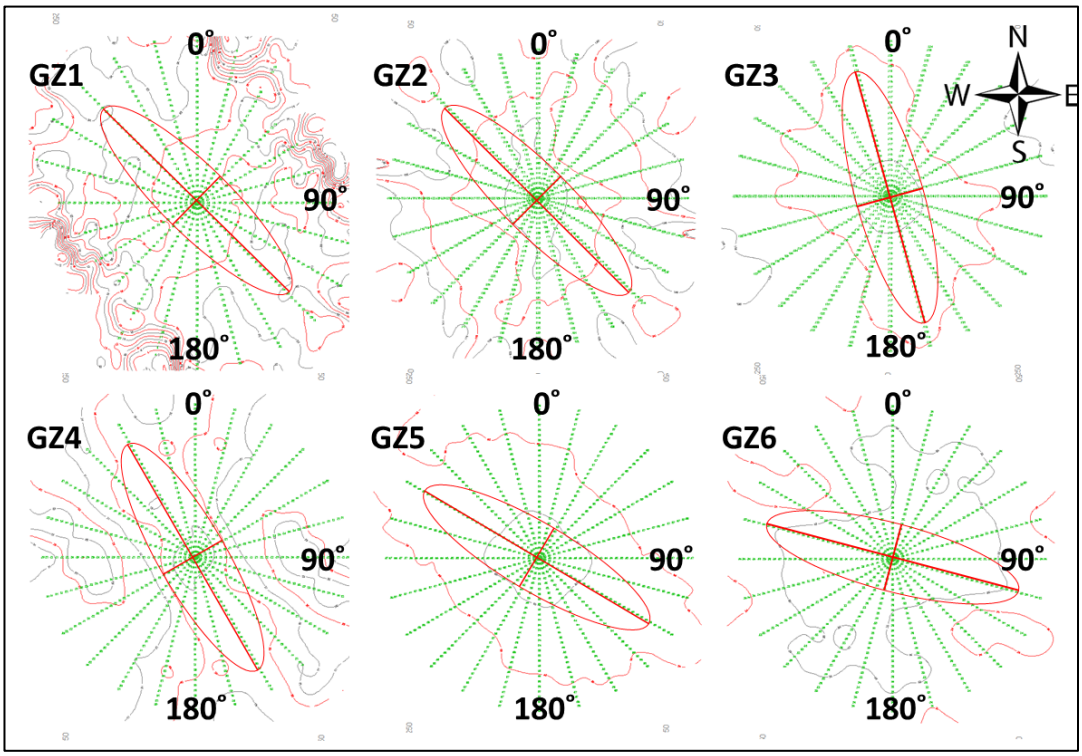


Figure 35: *Variogram contour maps based on the mg/t sampling data for each geozone*

The directions of most and least spatial continuity were identified by plotting ellipses on the experimental variograms contours in the varmaps shown in Figure 35. The orientation of the ellipses is also a function of the data configuration.

The ellipse of GZ1 has a small width compared to its length which indicates that there is a lot more continuity in 135˚ direction than in the 45˚ direction. The data in GZ1 is clustered in the North-Western corner, see Figure 6. The interpretation of the variography might be different in the future when more data becomes available as mining progresses in this geozone.

GZ2 has an ellipse where the width is in good proportion to the length and shows more continuity in the 135˚ direction than the 45˚ direction, a clear example of anisotropy.

In GZ3, the shape of the first contour is very close to a circle, which means that the continuity at close-range is more or less the same in all the directions. The second contour shows a clear ellipse with more continuity in the 165˚ direction than in the 75˚ direction. When modelling the variogram for this geozone the first structure should have the same range in all directions, and the second structure should have a longer range in the 165˚ direction than in the 75˚ direction.

GZ4 has an ellipse where the width is in good proportion to the length and shows more continuity in the 150˚ direction than in the 60˚ direction.

In GZ5, again the shape of the first contour is very close to a circle, which means that the continuity at close-range is similar in all the directions. The second contour also shows a clear ellipse with more continuity in the 120˚ direction than in the 30˚ direction. When modelling the variogram for this geozone the first structure should have the same range in all directions, and the second structure should have a longer range in the 120˚ direction tan in the 30˚ direction.

GZ6 has an ellipse where the width is in good proportion to the length and shows more continuity in the 105˚ direction than in the 15˚ direction.

On average the contour maps also indicate that the direction of continuity is in the NW/SE direction, and supports the findings of the swath analysis, transport flow direction and the facies analysis.

At Tshepong mine 60m x 60m block size is used for estimating Indicated blocks. For the estimation of blocks of this size the 5mx5m data has been regularised into 60m x 60m. The resultant variogram contour maps per geozone for this support appear in Figure 36.

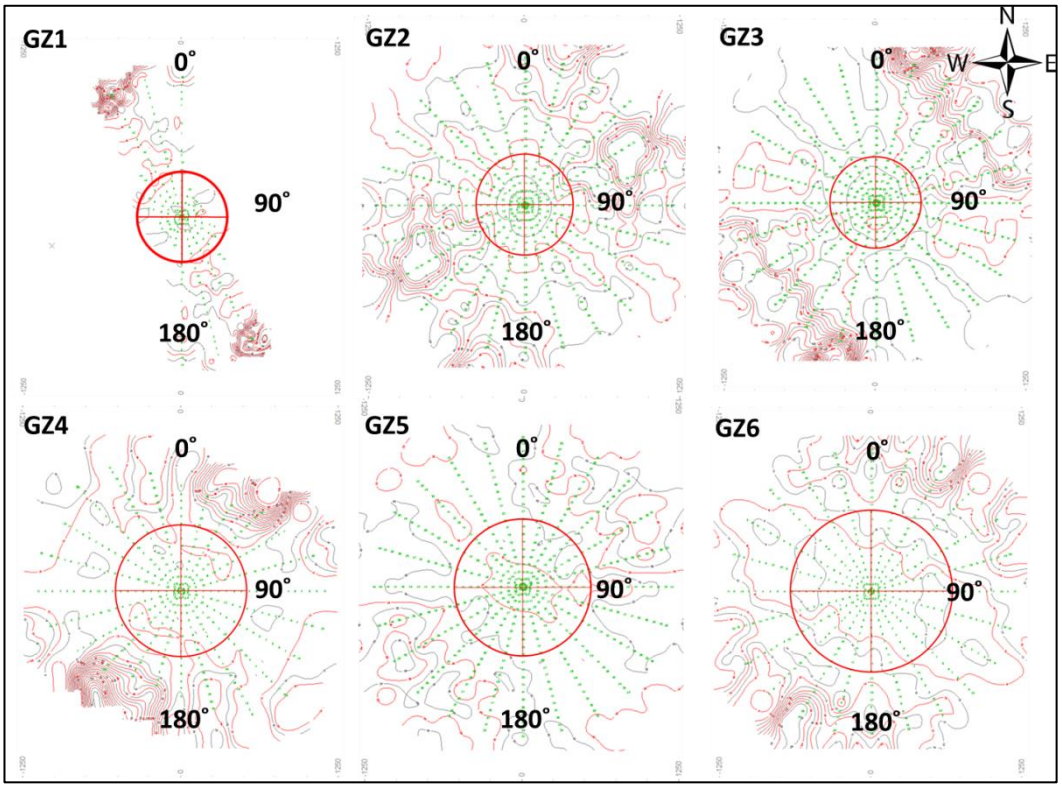


Figure 36: *Variogram contour maps based on the 60m x 60m regularised data*

Anisotropy could be identified in all of the geozones except for GZ2. While there appears some evidence of anisotropy in some of the geozones, the lack of numbers of pairs at higher lag distances, resulting from regularisation made reliable interpretation of anisotropy difficult and therefor omnidirectional variograms were modelled.

The variogram contour maps per geozone shown in Figure 37 were created from the 120m x 120m regularised blocks; the size used for estimation of Inferred blocks.

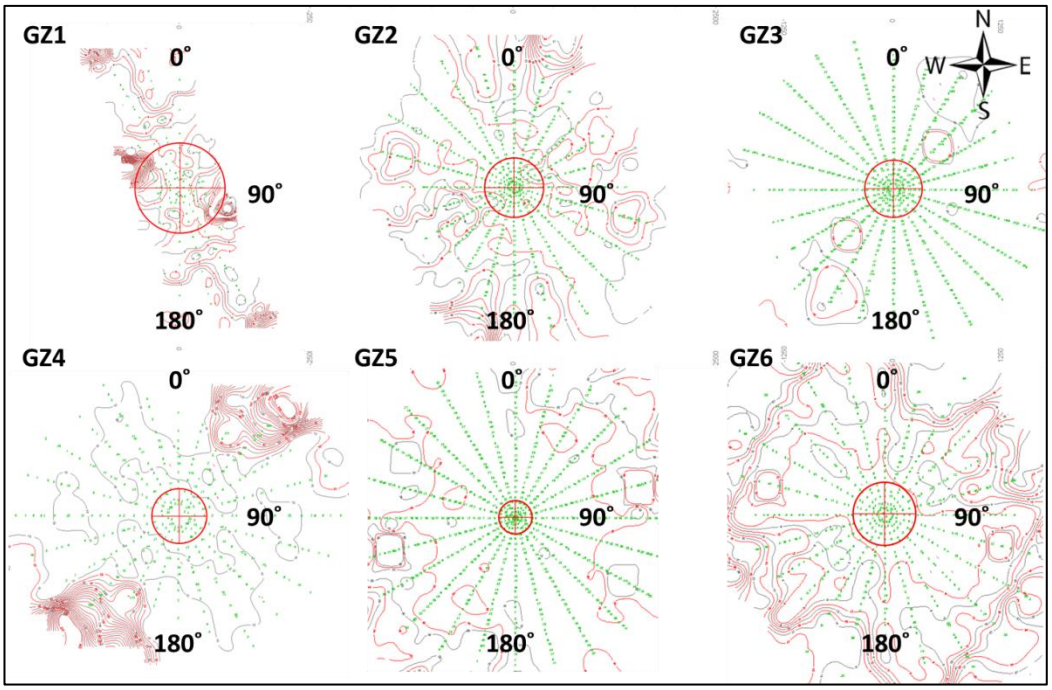


Figure 37: *Variogram contour maps of 120m x 120m regularised blocks*

Anisotropy could be identified in all of the geozones, again the anisotropies were difficult to model because the number of pairs in the directions of continuity was not sufficient to warrant anisotropic modelling. Again, due to the lack of the number of pairs resulting from regularisation, omnidirectional variograms were modelled.

5.3 Variogram Models

The variogram quantifies and measures spatial correlation between sample values in the mineral deposit. The weighting assigned to samples during kriging is related to this spatial correlation. The estimation variance is a measure of the estimation error. The optimum weighting assigned to the samples is done in such a way to ensure a minimum estimation variance, and this is known as the kriging variance.

The experimental points on the semi-variogram represents the observed variability between samples that are at specific distances apart and describes the underlying structure in terms of variability of the deposit. The variability of the grade is measured with respect to the spatial dispersion of data by

considering the variance between pairs of data points over specific separation distances.

The range of influence of the variogram model is the distance at which the variogram levels off and reaches the sill – the estimated population variance. Samples that are further away than the range are spatially uncorrelated (Clark, 1979).

When fitting a model to the experimental variogram of each geozone, the nugget effect was modelled by taking cognisance of the γ value at 0.08m, the first few experimental semi-variogram points, and the number of pairs per point.

Spherical models with nested structures were used in all the geozones. The variograms were calculated in all directions with a 22.5˚ increment between each direction starting with an azimuth of 0˚. The lag tolerance was half a lag distance. The lag distance was the same for all geozones and coincided with the data spacing. The number of lags varied depending on the range of influence. The variograms were calculated in mg/t to reduce the size of numbers during calculations. The values were scaled back to cmg/t in a later step.

In this research experimental semi-variograms were created for estimation into the 30m x 30m block models for each geozone from 5m x 5m sampling grid data. Experimental semi-variograms were created for the 60m x 60m block model for each geozone by using data that was regularised into 60m x 60m blocks, and for the 120m x 120m block model for each geozone by using data that was regularised into 120m x 120m blocks.

5.3.1 GZ1

The contour map of GZ1 in Figure 35 showed that there is more continuity at small distances in the 135˚ direction. The variograms in Figure 38 below were used for the 30m x 30m estimation block model. A small anisotropy was observed in the first structure of the variogram model. The estimated nugget

is 65% of the total sill, which higher than the standardised variance of the check samples (Table 7, GZ2) and which was modelled taking the number of pairs of the shorter lags into consideration.

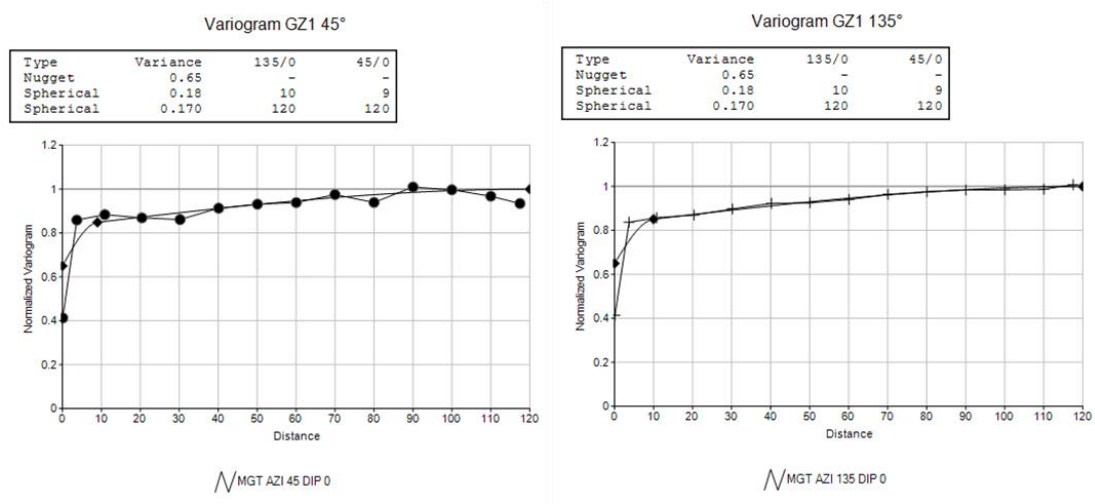


Figure 38: *Anisotropic variograms of GZ1*

The variogram contour map for the 60m x 60m block model shown in Figure 36 indicated that there is more continuity in the 0° direction. However, bearing mind the shape of GZ1, one realises that on a 60m x 60m regularised grid there would be a scarcity of data points and it does not come as a surprise that there is lack in the number of pair in these directions.

The anisotropy observed was recognised as an example of the data configuration creating an artefact in the interpretation of anisotropy at this scale of data. The fitted model is an isotropic variogram model and is shown in Figure 39 below. The variogram was not modelled beyond 360m, as there was insufficient data to assume spatial correlation beyond this point.

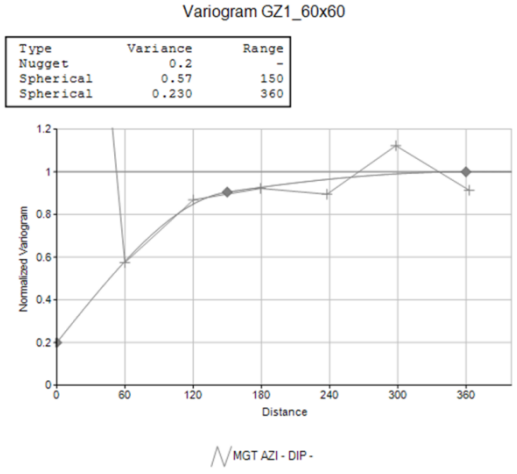


Figure 39: *Isotropic variograms of GZ1 (Indicated confidence 60m x 60m blocks)*

The variogram contour map for the 120m x 120m model for GZ1 shown in Figure 37 indicated that there is more continuity in the 150° direction than in the 60° direction. This anomaly is again, because of the lack of data on a 120m x120m regularised grid in this geozone, as well as the shape of the geozone.

An omnidirectional variogram was modelled and is shown in Figure 40 below.

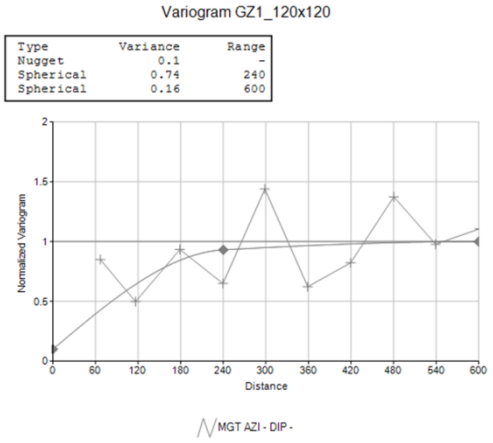


Figure 40: *Isotropic variogram of GZ1 (Inferred confidence 120m x 120m blocks)*

5.3.2 GZ2

The variogram contour map for GZ2 in Figure 35 indicated that the direction of continuity was in the 135˚ direction. The variograms in Figure 41 are for

the 30m x 30m block model. The second structure of the model had a longer range in the 135˚ direction than in the 45˚ direction. The nugget is 49% of the total sill, which higher than the standardised variance of the check samples (Table 7, GZ2) and again modelled by taking the number of pairs of the shorter lags into consideration.

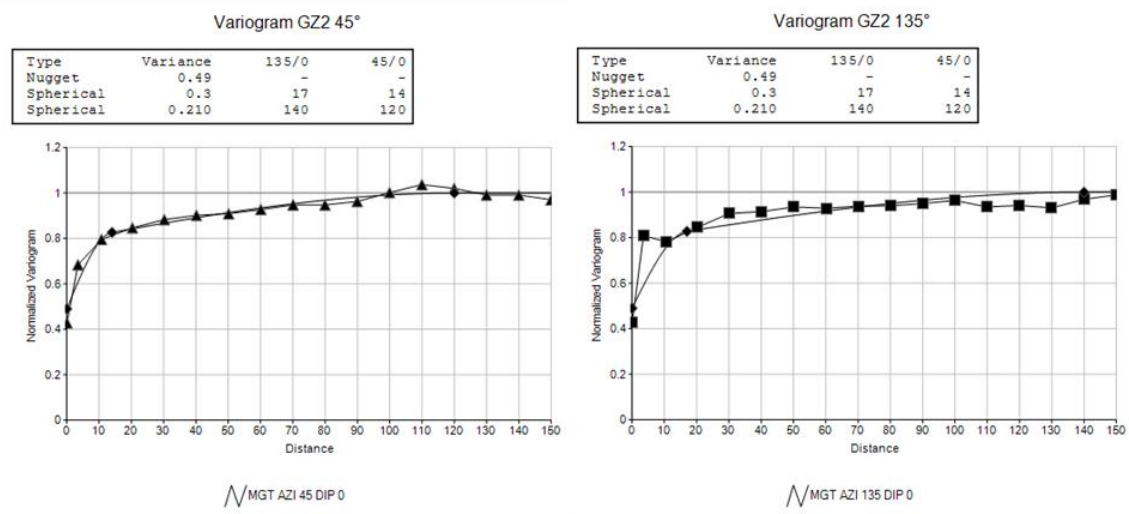


Figure 41: *Anisotropic variograms of GZ2*

The variogram contour map for the 60m x 60m model shown in Figure 36 indicated that the continuity is the same in all directions. The omnidirectional variogram for the 60m x 60m model for GZ2 is shown in Figure 42 below.

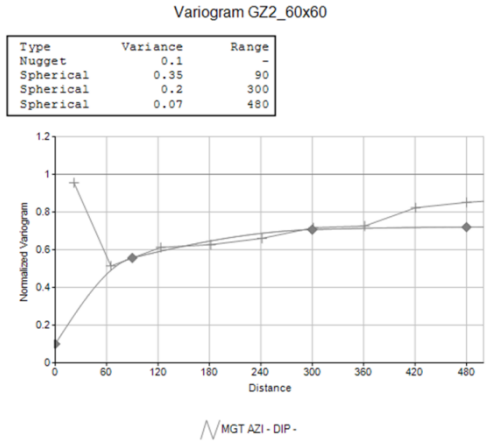


Figure 42: *Isotropic variogram of GZ2 (Indicated confidence 60m x 60m blocks)*

A model with three structures was used. A decision was made not to model beyond 480m due to trend in the geozone.

The variogram contour map for the 120m x 120m model for GZ2 shown in Figure 37 indicated that there is more continuity in the 135° direction than in the 45° direction.

However, due to the lack of pairs in these directions it is believed that this change in continuity direction might be an artefact and a single structure isotropic variogram model was modelled and is shown in Figure 43 below.

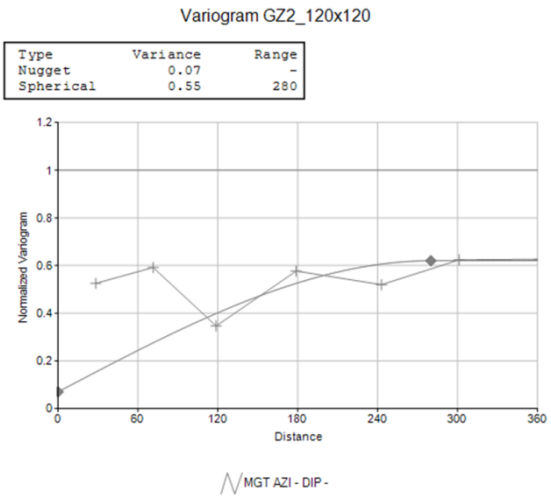


Figure 43: *Isotropic variogram of GZ2 (Inferred confidence 120m x 120m blocks)*

Trend was observed beyond 280m and it was decided that 280m will be the maximum range of the variogram.

5.3.3 GZ3

The variogram contour map of GZ3 in Figure 35 indicated that there is more continuity in the 165˚ direction. The variograms in Figure 44 are for the 30m x 30m model. The nugget is 42% of the total sill, which is slightly higher than the standardised variance of check samples (Table 7, GZ3) and it was modelled by taking the number of pairs in the shorter lags into consideration.

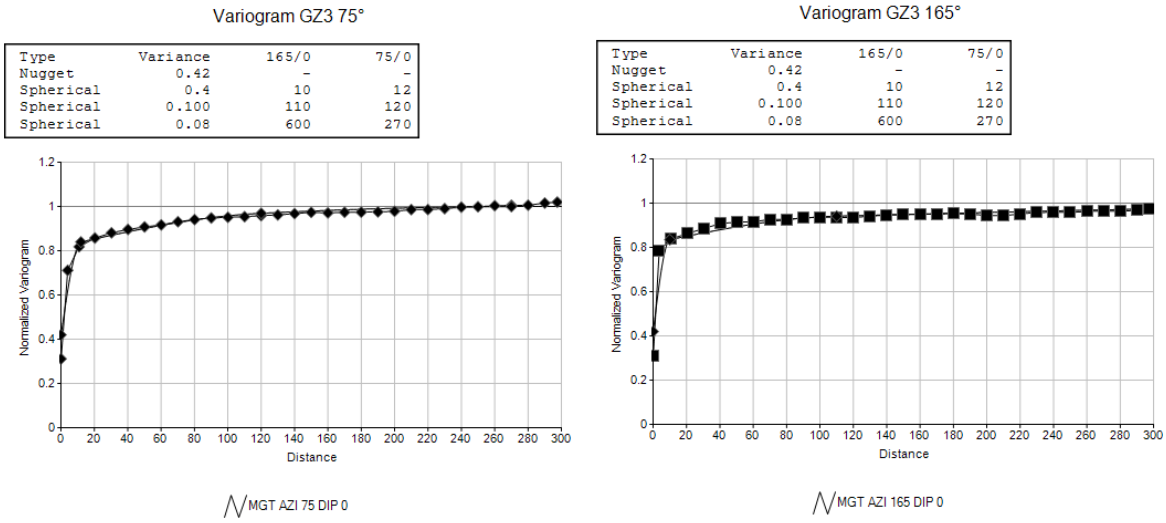


Figure 44: *Anisotropic variograms of GZ3*

The variogram contour map for the 60m x 60m model shown in Figure 36 indicated that the continuity is more in the 0° direction than in the 90° direction. The change in direction of the observed continuity is due to the shape of GZ3, on a 60m x 60m regularised grid there are less data points which results in a lack in the number of pair in these directions. The anisotropic variogram models for the 60m x 60m model for GZ3 are shown in Figure 45 below.

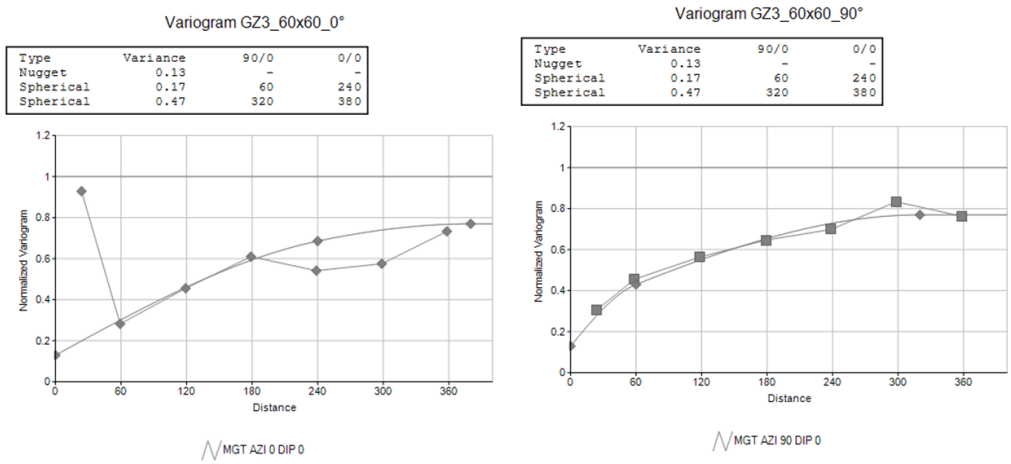


Figure 45: *Anisotropic variograms of GZ3 (Indicated confidence 60m x 60m blocks)*

Some evidence of trend was observed beyond 400m and it was decided that 400m will be the maximum range of the variograms.

The variogram contour map for the 120m x 120m model for GZ3 shown in Figure 37 indicated that long range continuity in the 135° direction and a short-range continuity in the 45° direction. Due to the lack of pairs it is believed that the change in the direction of continuity is not real but rather arises due to scarcity of 120m x 120m blocks. A single structured isotropic variogram model was fitted and is shown in Figure 46 below.

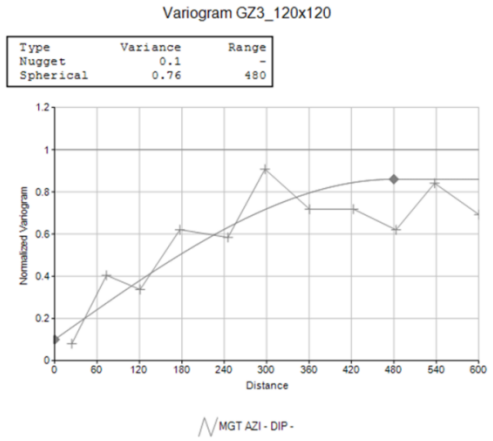


Figure 46: *Isotropic variogram of GZ3 (Inferred confidence 120m x 120m blocks)*

5.3.4 GZ4

The variogram contour map of GZ4 in Figure 35 indicated that there is more continuity in the 150˚ direction. The variograms in Figure 47 are for the 30m x 30m model. The nugget is 72% of the total sill which is much higher than the standardised variance of check samples (Table 7, GZ4). There are only a few pairs for this geozone, and the γ value at 0.08m did not have a large influence on the nugget effect.

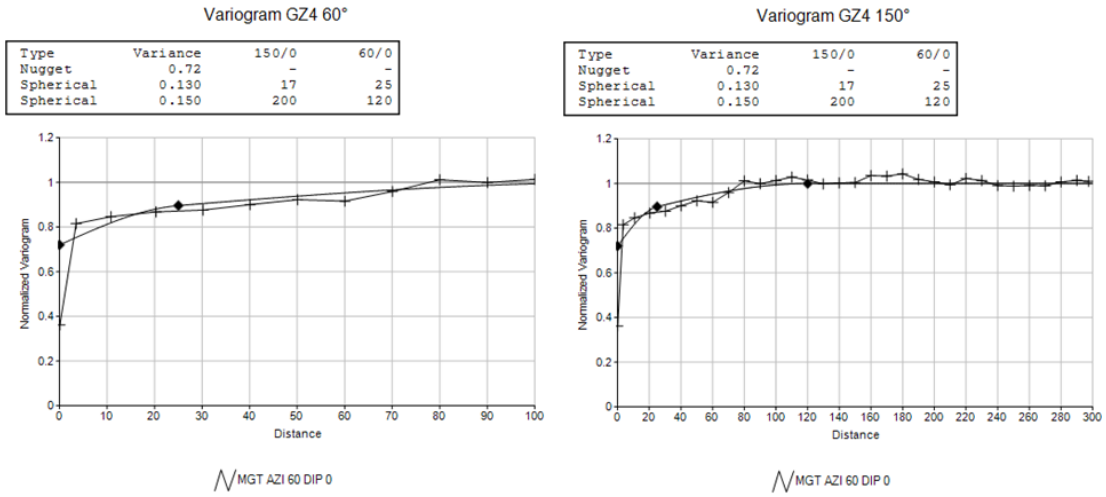


Figure 47: *Isotropic variograms of GZ4*

The variogram contour map for the 60m x 60m model shown in Figure 36 indicated that the continuity is more in the 120° direction than in the 30° direction. Due to the lack of pairs in these directions the interpreted anisotropy was ignored. The modelled omnidirectional variogram for the 60m x 60m data in GZ4 is shown in Figure 48 below.

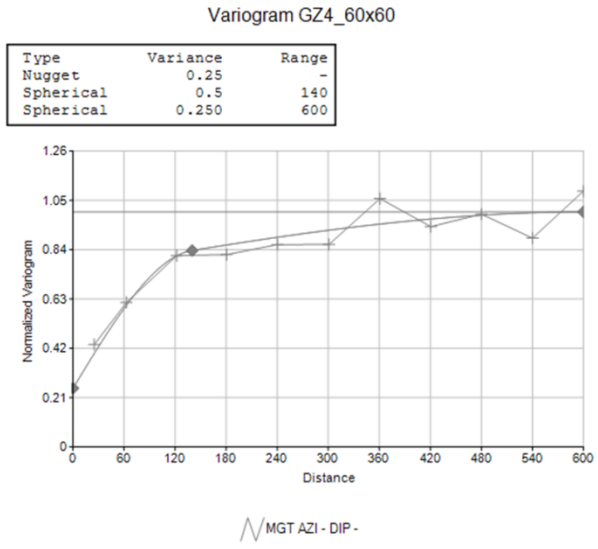


Figure 48: *Isotropic variogram of GZ4 (Indicated confidence 60m x 60m blocks)*

GZ4 does not have Inferred areas and does not need a variogram for a 120m x 120m model.

5.3.5 GZ5

The variogram contour map of GZ5 in Figure 35 showed that there is more continuity in the 120˚ direction. The variograms in Figure 49 are for the 30m x 30m model. The variograms confirmed the direction of continuity. The nugget is 65% of the total sill. This is again higher than the standardised variance of check samples (Table 7, GZ5) the standard procedure of modelling the variogram and weighing it towards the larger number of pairs in the shorter lags into consideration.

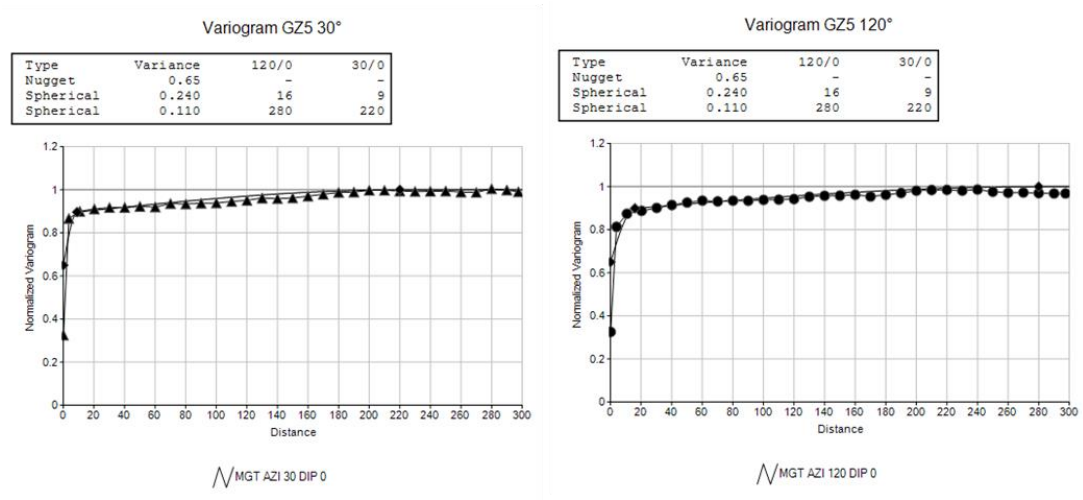


Figure 49: *Anisotropic variograms of GZ5*

The variogram contour map for the 60m x 60m model shown in Figure 36 indicated that the continuity is more in the 135° direction than in the 45° direction. Again, like in GZ3, the change in direction of the observed continuity is due to the shape of GZ5, on a 60m x 60m regularised grid there are less data points which results in a lack in the number of pairs in these directions. As for previous geozones the lack of pairs in these directions prevented the fitting of an anisotropic model. The omnidirectional variogram for the 60m x 60m model for GZ5 is shown in Figure 50 below.

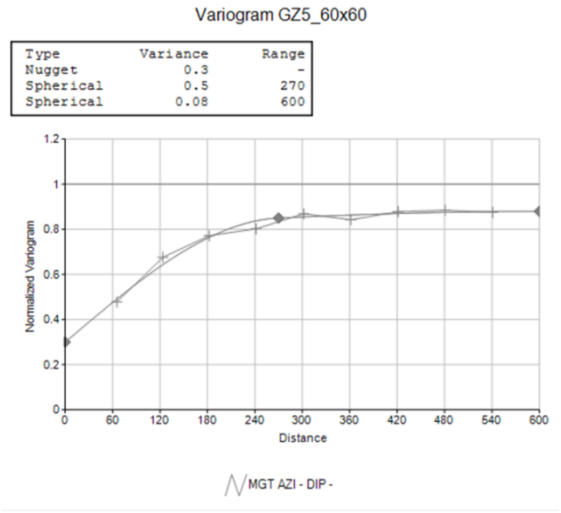


Figure 50: *Isotropic variogram of GZ5 (Indicated confidence 60m x 60m blocks)*

Trend was observed beyond 600m and it was decided to restrict the range to 600m.

The variogram contour map for the 120m x 120m model for GZ5 shown in Figure 37 indicated that there is more continuity in the 150° direction than in the 60° direction. As before, the lack of pairs in these directions, prohibited modelling of this anisotropy. The single structure isotropic variogram model for GZ5 is shown in Figure 51 below.

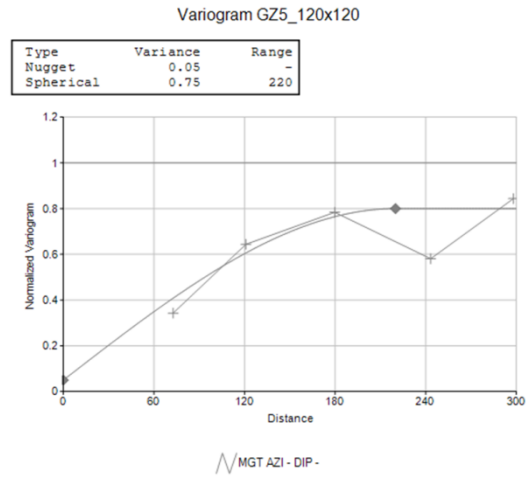


Figure 51: *Isotropic variogram of GZ5 (Inferred confidence 120m x 120m blocks)*

Trend was observed beyond 220m and the range was modelled to 220m.

5.3.6 GZ6

The anisotropy in GZ6 in Figure 35 was in a different direction to the other anisotropic geozones. The variograms in Figure 52 are for the 30m x 30m model. This can be due to the LF in GZ6 compared to the BCF in the other geozones. The nugget is 60% of the total sill. This is again higher than the standardised variance of check samples (Table 7, GZ6) and again the standard procedure of modelling the variogram by weighing it towards the larger number of pairs within the shorter lags was followed.

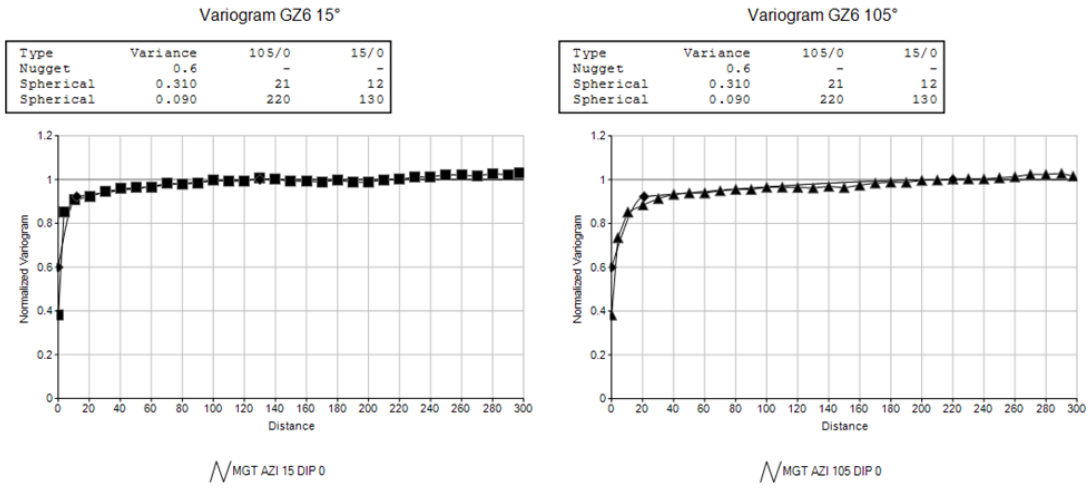


Figure 52: *Anisotropic variograms of GZ6*

The variogram contour map for the 60m x 60m model shown in Figure 36 indicated that the continuity is more in the 105° direction than in the 15° direction. Again, like in previous geozones, the change in direction of the observed continuity is due to the shape of GZ6, on a 60m x 60m regularised grid there are less data points which results in a lack in the number of pairs in these directions. Due to the lack of pairs in these directions it was not possible to create an anisotropic model. The omnidirectional variogram for the 60m x 60m model for GZ6 is shown in Figure 53 below.

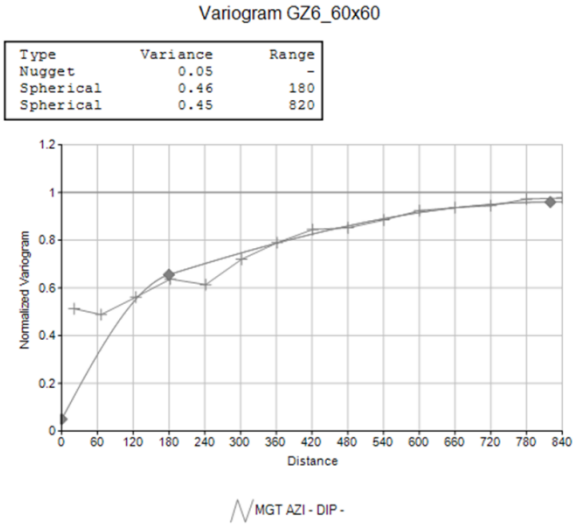


Figure 53: *Isotropic variogram of GZ6 (Indicated confidence 60m x 60m blocks)*

Trend was observed beyond 820m but it was still possible to fit a reasonable model.

The variogram contour map for the 120m x 120m model for GZ6 shown in Figure 37 indicated that there is more continuity in the 105° direction than in the 15° direction. Due to the lack of pairs in these directions it was not possible to create an anisotropic model. The single structure isotropic variogram model for GZ6 is shown in Figure 54 below.

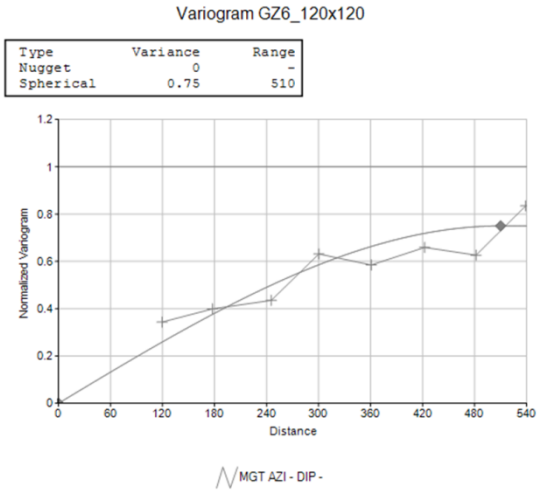


Figure 54: *Isotropic variogram of GZ6 (Inferred confidence 120m x 120m blocks)*

Due to trend the second sill was higher than the first sill and would cause poor results. It was thus decided to only model up to the first sill and to limit the range to 520m.

5.3.7 Stationarity

An important concept in geostatistics is that of stationarity which refers to a situation where a stationary random function is self-repeating in space and homogeneous. A distribution of $Z(x_1)$, $Z(x_2)$,…,$Z(x_k)$ should be the same as $Z(x_1+h)$,$Z(x_2+h)$,…,$Z(x_k+h)$ to be able to make statistical inference possible, where $Z(x_k)$ is the sample value at location (x), $Z(x_k+h)$ is the sample value at location (x+h) and h is the distance between the samples locations.

In order to achieve stationarity and homogeneity geozones must be delineated in such a way that all the data within a geozone has the same statistical and geostatistical characteristics.

The distribution of differences (Z(x)-Z(x+h)) can be measured to check for stationarity and spatial correlation.

- If this distribution is symmetrical it is an indication of stationarity
- If the mean of the differences in cmg/t is close to zero for all h values one can assume stationarity
- The spread of the difference distribution for all lag distance values give an indication of the spatial correlation in the geozone, the narrower the distribution of differences the more the spatial correlation (Isaaks & Srivastava, 1989)

Stationarity occurs when values have similar differences at similar distances apart, and this is observed in the experimental semi variograms. The data would reach a sill and beyond this lag distance it would appear as a constant variability. If an increasing variability is observed beyond a specific lag distance, then this would be indicative of a trend in the values. When this occurs, the practical solution is to only model the variogram up to that lag

distance, effectively assuming there does not exist any spatial correlation beyond that lag.

The trends observed for the 60m x 60m and 120m x 120m blocks experimental variograms, were interpreted as being to be a function of the scarcity of data at those block sizes and were therefore ignored and omnidirectional variogram models were fitted up to the lag of stable relative variability. No additional stationarity tests were carried out in this research as the evidence in the variograms shown in Figure 38, Figure 41, Figure 44, Figure 47, Figure 49 and Figure 52 for the individual geozones, was deemed sufficient to assume stationarity.

6 QUANTITATIVE KRIGING NEIGHBOURHOOD ANALYSIS

The practitioner is responsible for the selection of a kriging neighbourhood, and this can have a significant impact on the quality of the kriging estimate. As an example, a neighbourhood that is too restrictive will generate results with big conditional bias (Krige, 1994) (Krige, 1996 A) (Krige, 1996 B).

Kriging can only be a best linear unbiased estimator if a suitable kriging neighbourhood is selected. A few tests were done to quantitatively assess the kriging neighbourhood based on work by Vann et al. (2003) which forms the base of this chapter.

6.1 Search Parameters

The slope of the variogram close to the origin and the relative nugget effect have a great influence on the choice of the kriging neighbourhood.

Pure nugget effect means that there is no correlation between any two sample sections in the geozone. In a case where the nugget is high, good estimation would require that the neighbourhood gets progressively larger close to the origin of the variogram. By limiting the search neighbourhood samples would be uncorrelated to the true grade of the estimated blocks. As the nugget increase, the weights assigned to samples will become increasingly similar.

The percentage of the modelled nugget variance to the total sill for each geozone is displayed in Table 9.

Table 9: *Relative nugget per geozone at sample scale*

Geozone 1	Geozone 2	Geozone 3	Geozone 4	Geozone 5	Geozone 6
65%	49%	42%	72%	65%	60%

The search ranges were optimised keeping in mind the relative nugget and that the slope of the variogram model close to the origin has a considerable influence on estimation.

6.2 Number of Samples

The number of samples used for estimations were optimised for each geozone. Averaging takes place if the number of samples used for estimation is too much and the variance of the estimate increases if the number of samples is too small. The higher the nugget as a percentage of the sill, the more samples are required for estimation. The number of samples used for estimation can be changed in the search parameter file, where the minimum and maximum number of samples is set.

To create a graph that indicates what the effect of the number of samples has on the variance, blocks of 30m x 30m in each geozone were estimated with models that were created by using different numbers of samples. The blocks were created right next to mined-out areas that were sampled on regular intervals. For each geozone, 15 models were created, the first model by using one sample and the last model using 15 samples. The blocks were estimated 15 times using the 15 different models, and the estimation variance of each model was recorded next to the number of samples that was used for the estimation in each geozone.

The estimation variance decreased as the number of samples used for estimation increased. The variance was plotted against the number of samples to see how many samples were necessary for estimation (Figure 55).

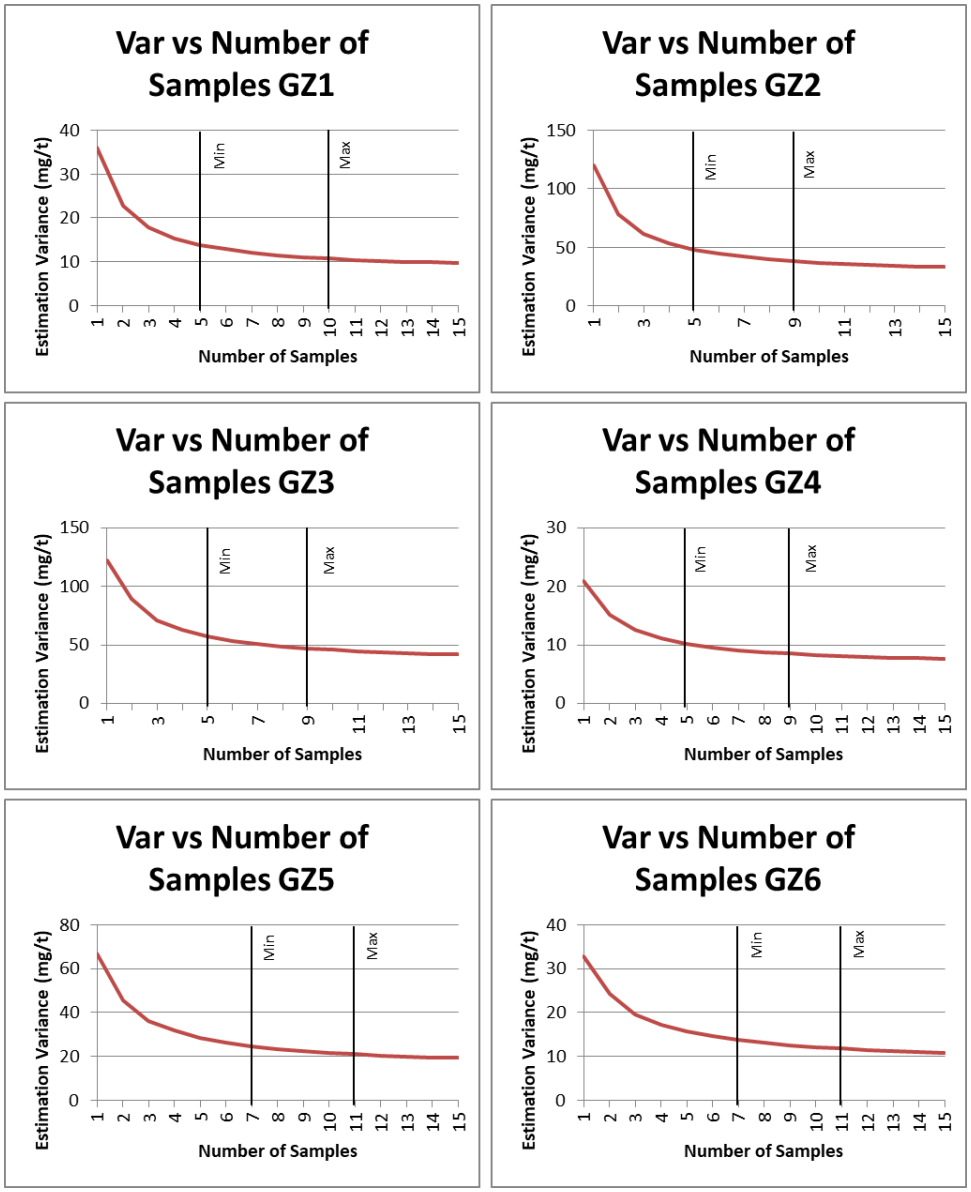


Figure 55: *Estimation variance vs. number of samples for 30m x 30m blocks in each geozone*

The curve of the estimation variance flattens off as the number of samples increases. It reaches a point where increasing the number of samples does not have a big effect on reducing the variance anymore, and this is the point where the variance stabilises, the number of samples at the point where the variance stabilises can then be used for the minimum number of samples in the search parameter file. The point where the variance does not decrease significantly by increasing the number of samples was chosen as the maximum number of samples to be used for the estimation.

To find the minimum number of samples one must find an acceptable balance between estimation variance and number of samples. Four of the variance curves in Figure 55 improved drastically by increasing the number of samples from 1 to 5 (geozone 1 to 4), and the other two curves (geozone 5 and 6) by increasing from 1 to 7. From 5 samples onwards, the variance still decreases but not as much as it decreased with fewer samples. A summary of the data is shown in Table 10 below.

Table 10: *Estimation variance of mg/t values vs. number of samples used to estimate*

Number of samples	1	2	3	4	5	6	7	8	9	10
Estimation variance GZ1 (mg/t)²	36.00	22.82	17.83	15.38	13.79	12.86	12.01	11.50	11.07	10.68
Estimation variance GZ2 (mg/t)²	119.57	77.81	61.25	53.26	47.94	44.35	41.74	39.88	38.24	36.84
Estimation variance GZ3 (mg/t)²	122.04	89.44	71.19	62.75	57.35	53.58	50.74	48.63	47.01	45.71
Estimation variance GZ4 (mg/t)²	21.00	15.19	12.51	11.12	10.17	9.52	9.05	8.74	8.50	8.31
Estimation variance GZ5 (mg/t)²	67.12	45.62	36.32	31.70	28.28	26.27	24.63	23.38	22.37	21.63
Estimation variance GZ6 (mg/t)²	32.86	24.36	19.50	17.27	15.63	14.56	13.74	13.07	12.58	12.11

For geozones 1 to 4 where the kriging variance curves became stable at 5 samples, 5 samples were chosen as a minimum number of samples, and for geozones 5 and 6, 7 samples were used as a minimum number of samples.

6.3 Discretisation

To decide on how many discretisation points to use, models were created with different amounts of discretisation points in units of compatible lengths to the block dimensions.

The slope of regression improved as the number of discretisation points increased. The number of discretisation points was increased several times until results stabilised. The models that were created by using fewer discretisation points could not generate Mineral Resource models with a good coverage, and many of the estimated blocks had a slope of regression of less than 0.7.

It was found that the optimum discretisation for the Measured block model was 6m x 6m for the 30m x 30m blocks. The discretisation also ties in with the number of and position of the samples in a 30m x 30m block.

This was sufficient to generate a Mineral Resource model with a good coverage of blocks that had a slope of regression higher than 0.7.

The Indicated 60m x 60m blocks used 4 and the Inferred 120m x 120m blocks used 8 discretisation points.

6.4 Conditional Bias

Conditional bias is related to the fact that the sample volume is much less than the actual extracted volume. Kriging has a smoothing effect that causes higher actual values to be estimated lower and lower actual values to be estimated higher. This happens because the estimates are smoothed towards the mean. A case of conditional unbiasedness is where the regression curve is linear and the slope equal to one (Rivoirard, 1987).

When considering the estimated value and the true value the slope of regression can be used to diagnose conditional biases. Conditional bias has a slope of regression of less than one. The ideal slope of regression is where the actual grade is equal to the estimated grade (Deutsch, 2007). However, in practice there is never a perfect correlation.

The regression of the true values given the estimate is an indication of the conditional bias (Figure 56).

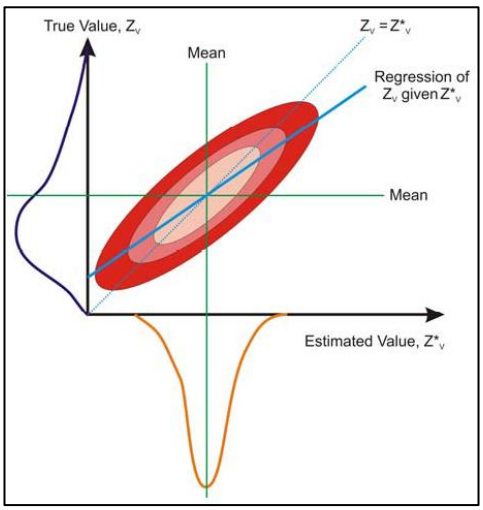


Figure 56: *The True value (Z_v) plotted against the Estimated value (Z^*_v) (Deutsch, 2007)*

The true value (Z_v) of blocks will always remain unknown, and the theoretical value of the conditional slope of regression is derived from the covariogram. For ordinary kriging the slope of regression can be calculated from the kriging weights and the covariogram function is shown in Equation 2 below (Deutsch, 2007).

$$b = \frac{Cov(Z_V, Z_V^*)}{\sigma^2_{Z_V^*}} = \frac{\sum_{i=1}^{n}(\lambda_i \bar{C}(v_i, V)}{\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j C(v_i, v_j)}$$

Equation 2

In practice however, a comparison of the empirical conditional bias and the theoretical conditional bias as calculated from the first part of the slope of regression formula above is possible, when the "true" block value is estimated from the mined-out blocks.

In this research the "true" block value was obtained from mined-out areas by calculating the average of the sample values in the 30m x 30m, for blocks that contained at least 4 samples in the block. The input data was declustered to the same block size as the block model with the same origin. The function "IJKGEN" was used to assign numbers to each declustered block to correspond with the block model in space. This allowed for the comparison between each declustered block (true value) to each model block (estimated value) that plots on exactly the same positions in space. Each geozone were done separately because the estimation parameters are different for each geozone.

The plot of the "true" block values (y-axis) and the estimated block values (x-axis) is very informative. The conditional covariance cloud, conditioned by the kriged block values, provides a graphical representation of the empirical conditional bias and the slope of the fitted trend line give the empirical value of the slope of regression. These conditional covariance clouds are shown in Figure 57.

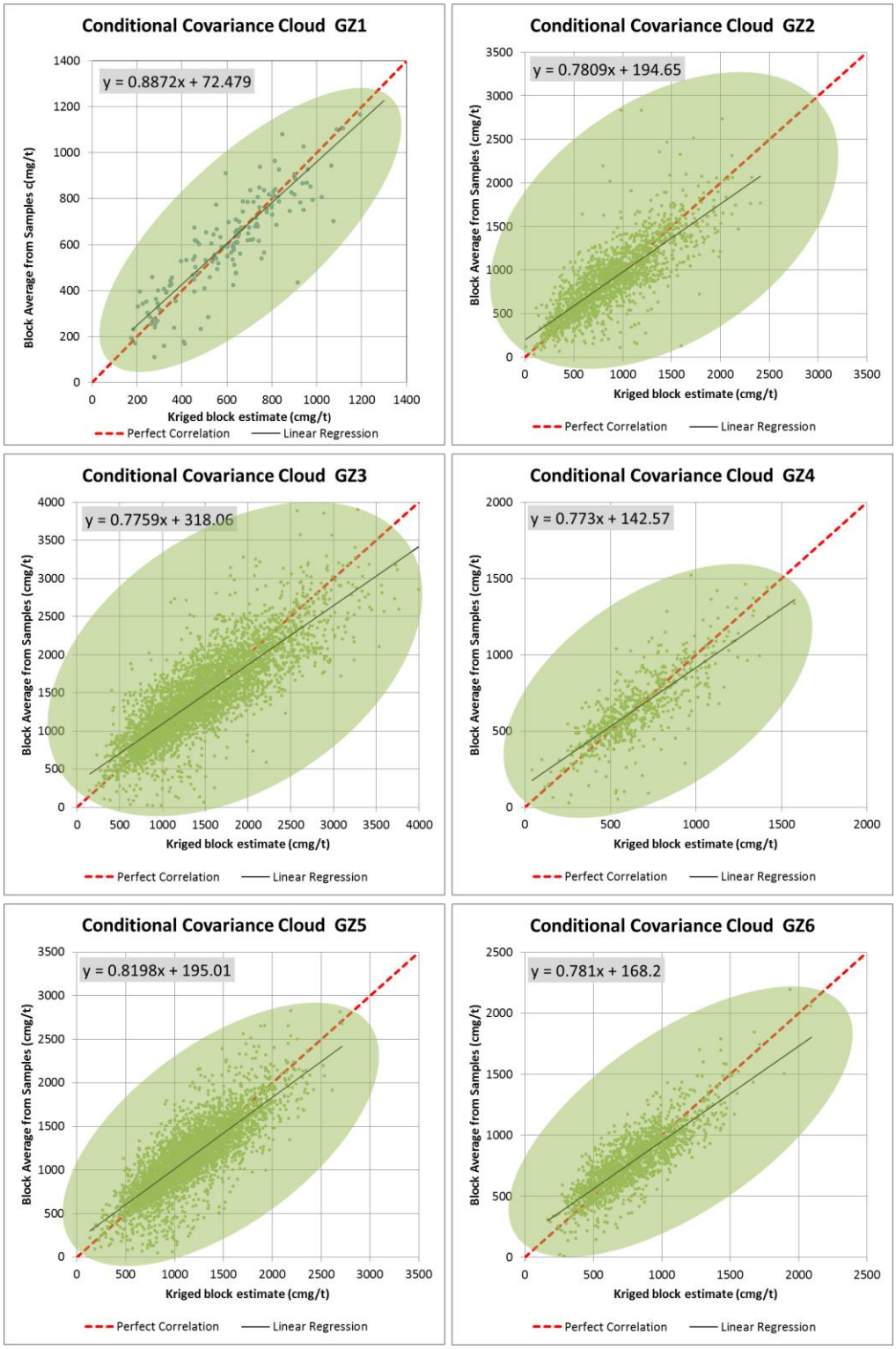


Figure 57: *Conditional covariance clouds of the Kriged block estimate values versus the Block average values from samples for block sizes of 30m x 30m, highlighting conditional bias*

This exercise is purely indicative and the slopes of the fitted trend lines for each geozone can be compared with the average slope of regression calculated from the estimated kriged block data (Table 11). In the kriging process the slope of regression and kriging efficiency are theoretically calculated for each estimated block by using Equation 2 defined previously and Equation 3 defined below:

- The covariogram model, which is derived from the semi-variogram model (this relationship is shown below)
- The sample locations, as observed in the ore body and captured in the data file
- The size and position of the test block relative to the sample positions (these are user defined)
- The number of kriging equations, that are a function of the minimum and maximum number of samples (also user defined)
- The kriging weights come from solving the kriging equations

The graphical and the theoretical relationships between the covariance and the semi variance under stationarity are given in below (Figure 58) (Dohm, 2015 B).

$C(h) = C - \gamma(h)$ where C = Variogram Sill and $\gamma(h)$ = semi-variogram value at h.

$C(0) = \sigma^2$ = population variance = $C_0+C_1+C_2$ = sum of the nugget effect and the semi-variogram sills.

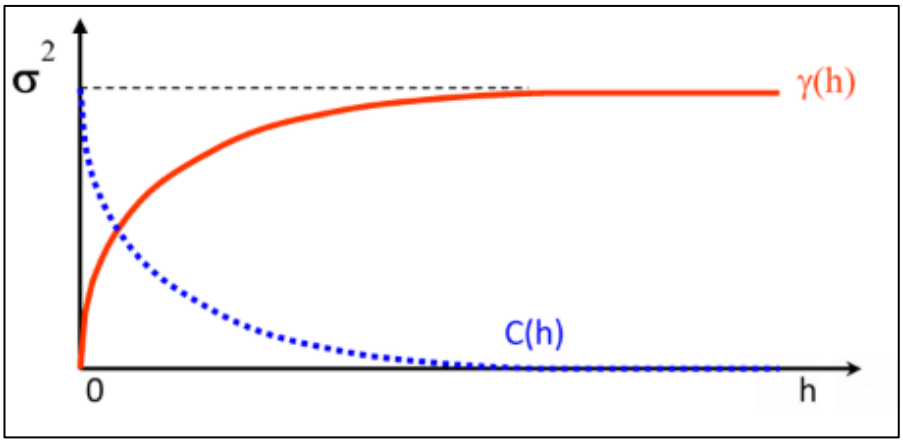


Figure 58: *The theoretical relationships between the covariance and the semi variance (Dohm, 2015 B)*

The covariance functions for sample support (v) and block support (V) are given below:

- $C(v, v) = \{ C_0+C_1+C_2 \} - \gamma(v, v) = \sigma^2 - \gamma(v, v)$ (covariance relationship at samples support v)
- $C(V, V) = \{ C_0+C_1+C_2 \} - \gamma(V, V) = \sigma^2 - \gamma(V, V)$ (covariance relationship at blocks support V)
- $C(v, V) = \{ C_0+C_1+C_2 \} - \gamma(v, V) = \sigma^2 - \gamma(v, V)$ (covariance relationship of sample v and block V)

Kriging Efficiency in terms of the covariogram is shown in Equation 3 below:

$$\text{Kriging Efficiency} = \text{KE} = \frac{\bar{C}(V,V) - \sigma_K^2}{\bar{C}(V,V)}$$

Equation 3

Where σ_K^2 is the kriging variance and in terms of the covariogram is shown in Equation 4 below:

$$\sigma_K^2 = \bar{C}(V,V) - \sum_{i=1}^{n} \lambda_i \bar{C}(v_i,V) + \lambda$$

Equation 4

It is important to routinely check the extent of conditional biases of ore block valuations, because block valuations that are subject to conditional biases leads to higher error variances and lower efficiencies. The factors that affect these biases for example range, nugget, data patterns, number of samples, block size and extrapolation distance can be analysed (Krige, 1996).

After a block has been mined out it is important to do a follow-up validation of the acceptability and quality of estimates by using the data that became available in the block. Never reduce the smoothing effect of kriging at the expense of increasing the conditional biases (Krige, 1996).

It is clear that conditional bias (the smoothing of the block estimates as opposed to the true values) is present in all the geozones, with acceptable slopes that are less than but close to one. Table 11 shows the correlation coefficient and the block factor between the actual and estimated values per geozone.

Table 11: *Correlation coefficient and block factor between actual value and estimated value of the 30m x 30m blocks*

Geozone number	Correlation Coefficient	Slope of Regression	Block factor
1	0.81	0.89	99.5%
2	0.80	0.78	100.0%
3	0.82	0.78	100.1%
4	0.80	0.77	101.0%
5	0.81	0.82	101.1%
6	0.83	0.78	100.2%

A perfect correlation coefficient between actual and estimate would have a value of one. The correlation coefficients in all the geozones are above 0.8 which shows a very good correlation between the true block values and the estimated block values. The slopes are the empirical slopes that are shown in Figure 57.

The block factor is a value expressed as a percentage between the value called for in the estimate and the value accounted for in the actual. All of the block factors are close to 100%.

The QKNA process followed in this research assisted in the reduction of conditional bias and the improvement of regression of the estimated blocks.

7 MINERAL RESOURCE ESTIMATION

Some assumptions had to be made about the data to ensure validity of the inferences that follow from using the data. The first assumption that was made relates to the precision of the sample assays where if you took exactly the same sample and submit it to exactly the same analysis the result should be within a certain tolerance. The results of the duplicate samples of Tshepong mine are within the required limits and validate the assumption that the sample assays are precise (Section 2.5 and Appendix B).

A second assumption was made regarding the accuracy of the sample assays where it was assumed that the sample assay values are in the proximity of the true value at that location. The results of the CRM samples of Tshepong mine are within the required limits and validate the assumption that the sample assays are accurate (Section 2.5 and Appendix B).

A third assumption was made that the samples are collected from a continuous and homogeneous population of all the possible samples and that the measured variable at the sample locations also exists at every un-sampled location that falls within the domain. There should also be no sudden changes of any characteristics and that estimations will be made within known geological constraints. The facies and domain analysis carried out allowed for this assumption to be valid (Section 3.3.1).

A forth assumption was made that all the values of un-sampled locations are related to the values at sampled locations. This spatial relationship was observed in the variography analysis and there is comfort in this essential assumption to spatial estimation methods (Section 5.3).

Datamine® Studio 3 was used for the creation of the Mineral Resource block models. Through the process of kriging of the observed sample values the block model is filled. Ordinary Kriging was used for estimates for 30m x 30m blocks, which at Tshepong mine are classified as Measured blocks. These blocks have the highest density of data and are close to current mining faces. Simple macro kriging was used for the estimation of Indicated (60m x 60m

blocks), less data and further away from current mining, and for the estimation of Inferred (120m x 120m blocks) with even fewer information and furthest away from current mining positions. The Mineral Resource confidence as a function of estimation methodology and estimation block size is standard practice at Tshepong mine. The ore body is tabular, and the block model is in two dimensions.

7.1 Block Model

Different block sizes have different variances, as blocks get bigger the variance between blocks decreases and vice versa. This is known as the volume variance or change of support effect. The variability in the underlying data is much bigger than the variance of a kriging estimate. The global mean of the deposit will always remain the same, only the block size can be changed, not the amount of metal in the deposit. Kriged estimates are globally unbiased.

Blocks that are too small will result in a very high variance and blocks that are too large will result in over-smoothing. There are not many constraints on Smallest Mining Unit (SMU) due to the conventional mining technique used on Tshepong mine. The face length of panels varies, ledging can have face lengths that vary between 10 and 15 metres and the face length of stope panels can vary between 15 and 40 metres on average. Block sizes that were chosen in the past, to cater for a good balance between variability and smoothing of estimates to overcome the volume variance, were honoured in this research.

At Tshepong blocks that are targeted for:

Measured confidence have a block size of 30m x 30m.

Indicated confidence have a block size of 60m x 60m.

Inferred confidence have a block size of 120m x 120m.

The proto-model has been defined on blocks of (X, Y, Z) = (120, 120, 1). The blocks were sub-celled into the relevant sizes to cater for Measured (30, 30, 1), Indicated (60, 60, 1) and Inferred (proto model) Mineral Resource confidence categories.

The parameters of the proto-model are shown in Table 12.

Table 12: *Parameters of the proto-model*

Model Parameter	Value
X origin	-17000
Y origin	-8500
Z origin	-0.5
Cell size in the X direction	120
Cell size in the Y direction	120
Cell size in the Z direction	1
Number of Cells in the X direction	150
Number of Cell s in the Y direction	150
Number of Cell s in the Z direction	1

Any area within the geozones that were not covered by the 120m x 120m estimate received a global mean estimate as summarised in Table 13.

After the kriging estimation of Measured, Indicated and Inferred models, the models were added together to produce one model with estimates and their associated Mineral Resource confidence categories.

Overall the model provides reliable estimates of grade on local scale and the averages of block estimates are in line with the input sample data. The final kriged model annotated on cmg/t is displayed in Figure 59.

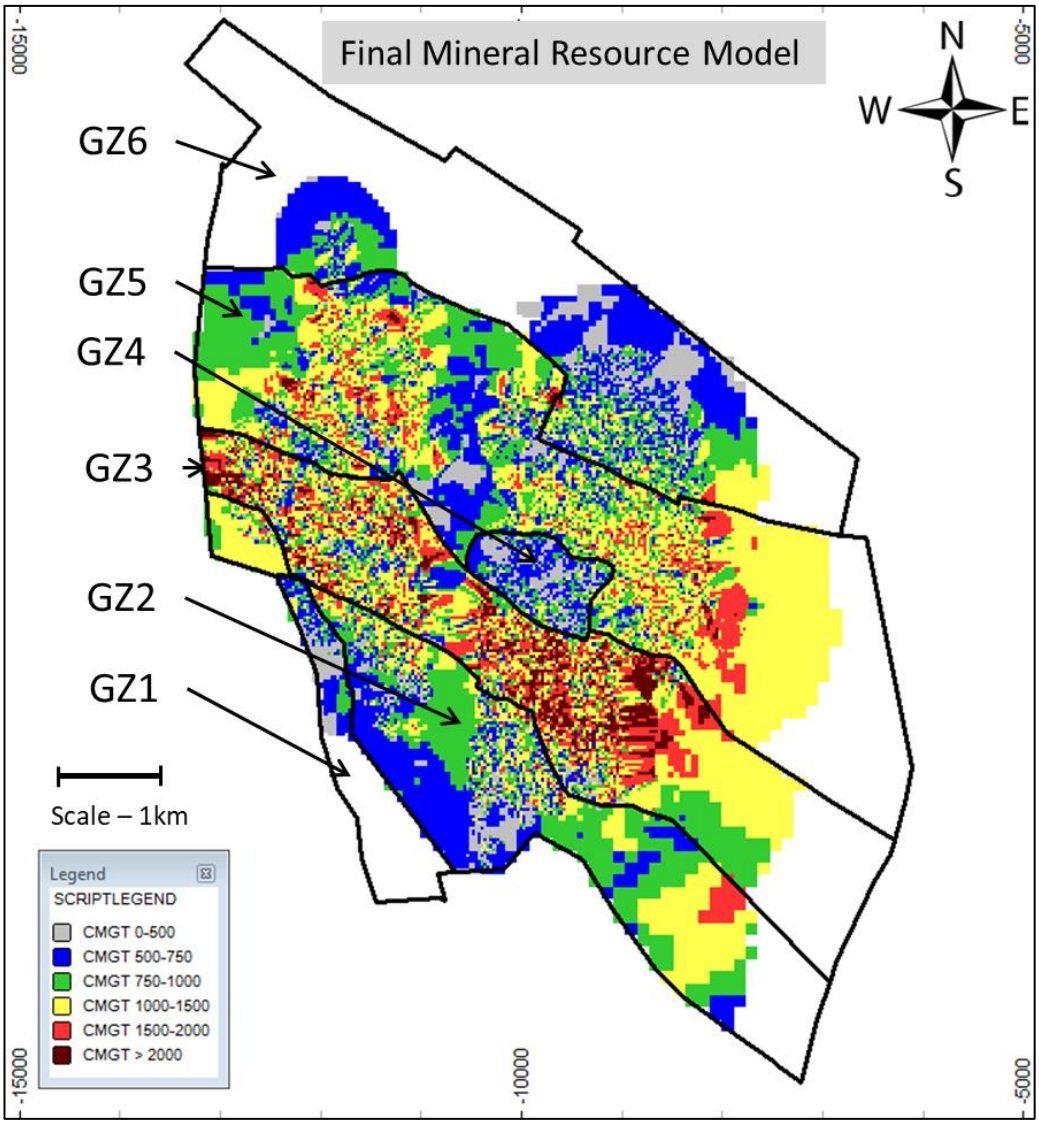


Figure 59: *Final Mineral Resource model (cmg/t)*

7.2 Global Means

A global mean value is a conservative average value per geozone. Mines tend to mostly mine higher grade areas, and leave low-grade areas, so an arithmetic mean is not the best estimate for a global mean. To calculate the declustered global mean per geozone the data was regularised into blocks starting at a block size of 30m x30m, increasing each side by 10m up to blocks of 200m x200m.

Each of these blocks received the average value of the samples that falls in the block, and the blocks receive a weight according to the number of

samples that falls in the block. The average of the regularised blocks was then taken as the average value for the respective block sizes. The average values of the blocks were then plotted against the corresponding block sizes (Figure 60).

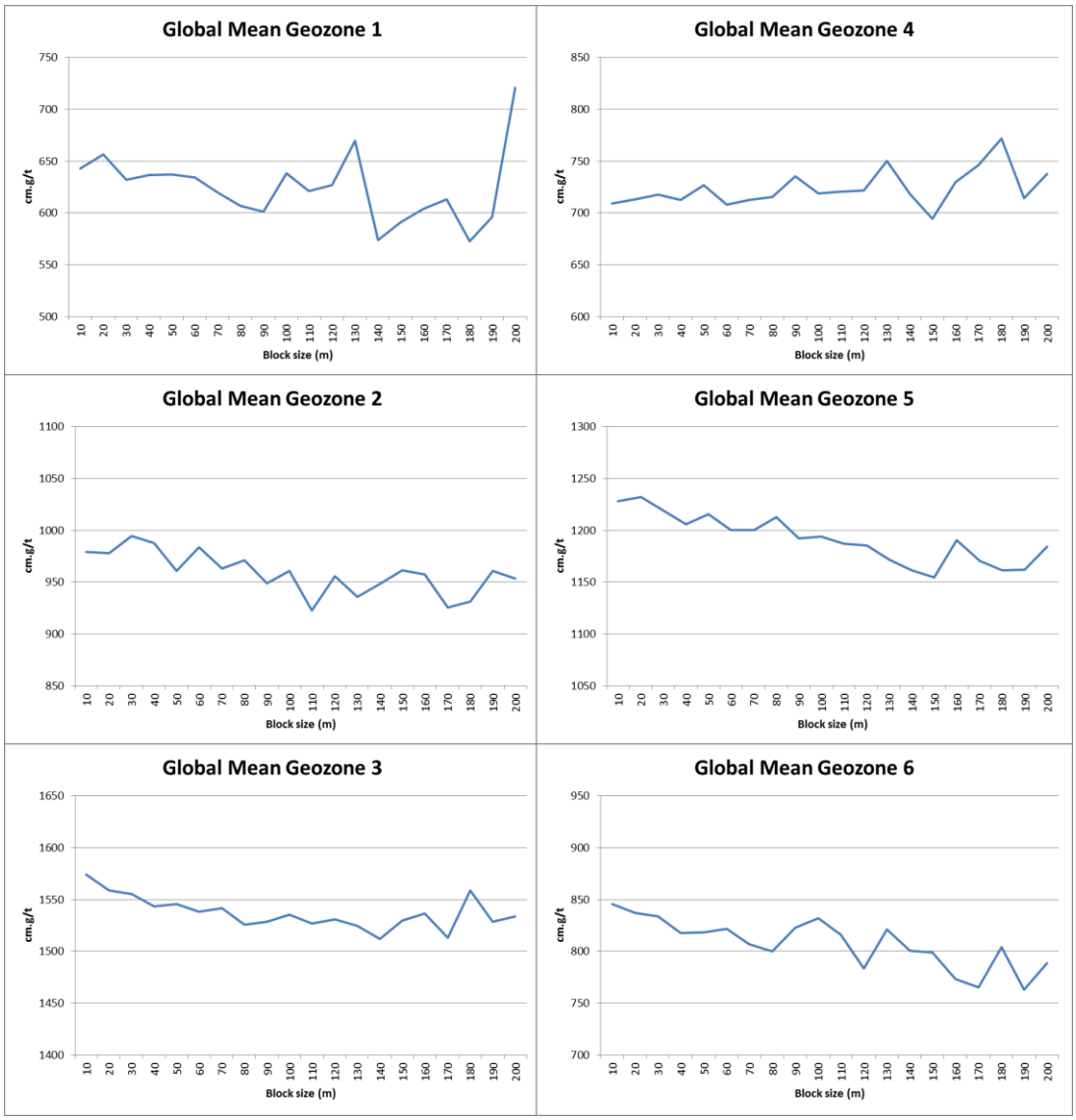


Figure 60: *Global mean per geozone per square block size*

The mean values of the blocks vary slightly as the block sizes change due to the weight assigned to the blocks according to the number of samples within each block. To give a conservative global value for the geozone the lowest declustered average value of the data was used. A comparison between the global mean values of the old model and the refined model generated in this

research, are displayed per geozone in Table 13 below. The global mean values per geozone of the refined model are all higher than that of the old model and highlights the impact of the uncapped historical data file.

Table 13: *Comparison between global mean values (cmg/t) of the old model and the refined model per geozone*

Geozone number	Global mean value of the refined model (cmg/t)	Global mean value of the old model (cmg/t)
1	572	570
2	923	862
3	1512	1390
4	694	645
5	1154	1089
6	765	774

7.3 Simple Macro Kriging

Simple macro kriging is used for the estimation of blocks in the Indicated and Inferred Mineral Resource categories (standard practice on Tshepong mine). Simple macro kriging uses larger support sizes and requires a "MKNUG" field to be added to the declustered data file. The number of samples in each block is not the same and MKNUG shows how the variability of blocks differs as the number of samples within blocks changes. The Indicated area is covered with 60m x 60m blocks and the input data for the simple macro kriging consists of data that is at block support and data that is at sample support. The block support data are located at the centre of the 60m x 60m regularised blocks and the value at these positions is the average of the data from the mined-out area that falls within a block of this size. The point support data are the relevant boreholes in the geozone at composite support. To distinguish between block support data the MKNUG field is introduced in the input file to appropriately assign weights to the borehole data and the block data. The calculation process of this nugget variance is explained below.

The kriging matrix includes the MKNUG field and solving the kriging equations ensures the weight for the samples at block support and samples at composite support are optimally and correctly assigned. A higher variability (MKNUG) will lead to smaller weights assigned to samples.

The MKNUG calculation is done for each declustered 60m x 60m block in each geozone. The calculation is done by calculating the total variance for all the 60m x 60m blocks in a geozone that contains one or more samples (this will be the highest variance and close to the sample variance in the geozone). The total variance is then calculated for 60m x 60m blocks that contain 2 or more samples (all the blocks that contained only one sample are excluded in this step, which means that the total variance in this instance will be lower than the total variance in the previous calculation). This process is then repeated for blocks that contain 3 or more samples, up to blocks that contain 50 or more samples, each geozone is considered separately. The total variance decreases every time that blocks with fewer samples are removed from the calculation. The total variances are then plotted against the minimum number of samples per block as in Figure 61 below.

In this research, the researcher came across some interesting findings regarding the MKNUG process, for example there were a number of cases where the variance did not reduce as the minimum number of samples increased, an assumption that was made when this methodology was introduced.

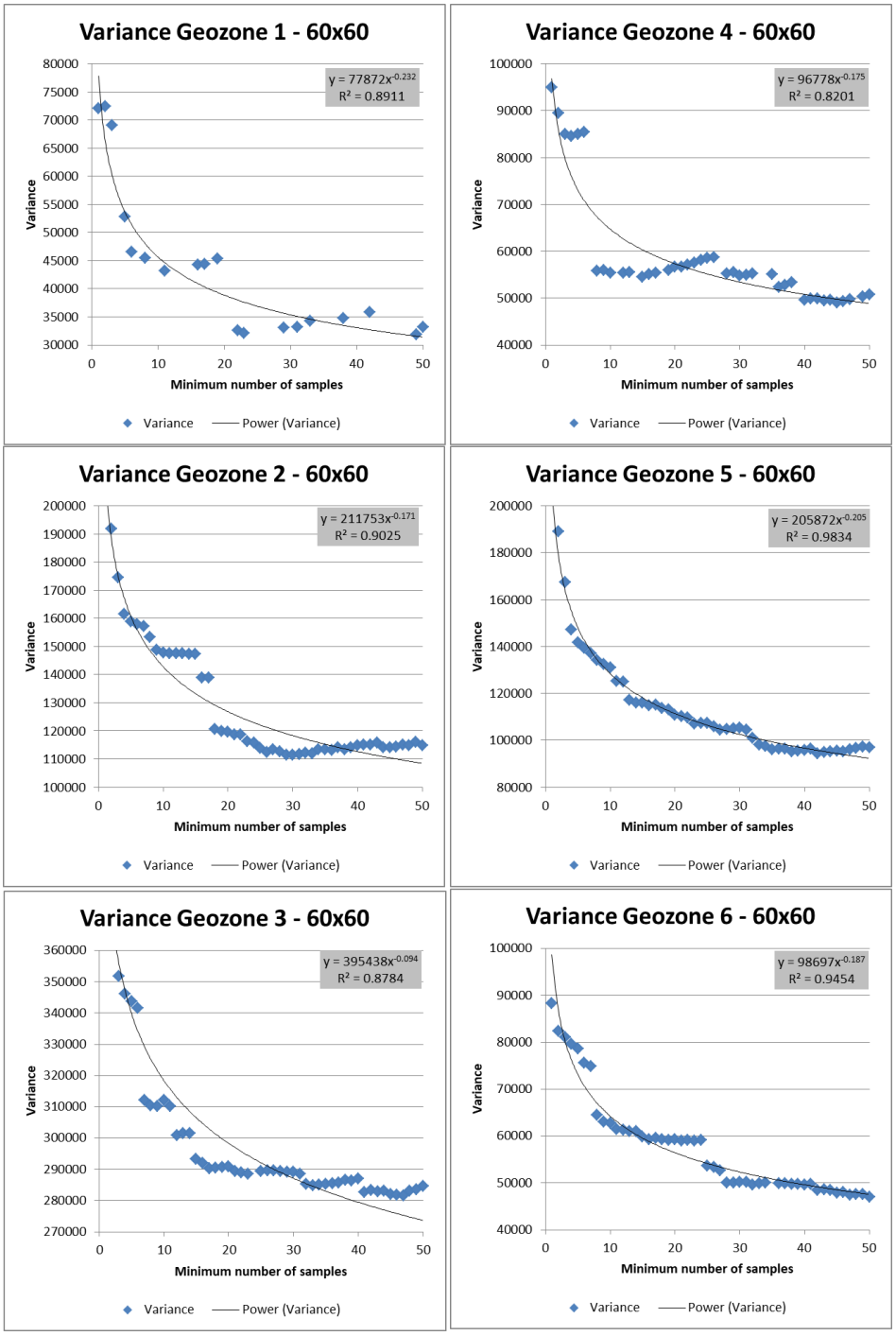


Figure 61: *Variance versus minimum number of samples for the MKNUG calculation for 60m x 60m blocks*

The equation of the curve for each geozone was then used for the calculation of the MKNUG field that was added to the declustered data file. The same was done for the Inferred model at a 120m x 120m scale and the results appear in Figure 62 below.

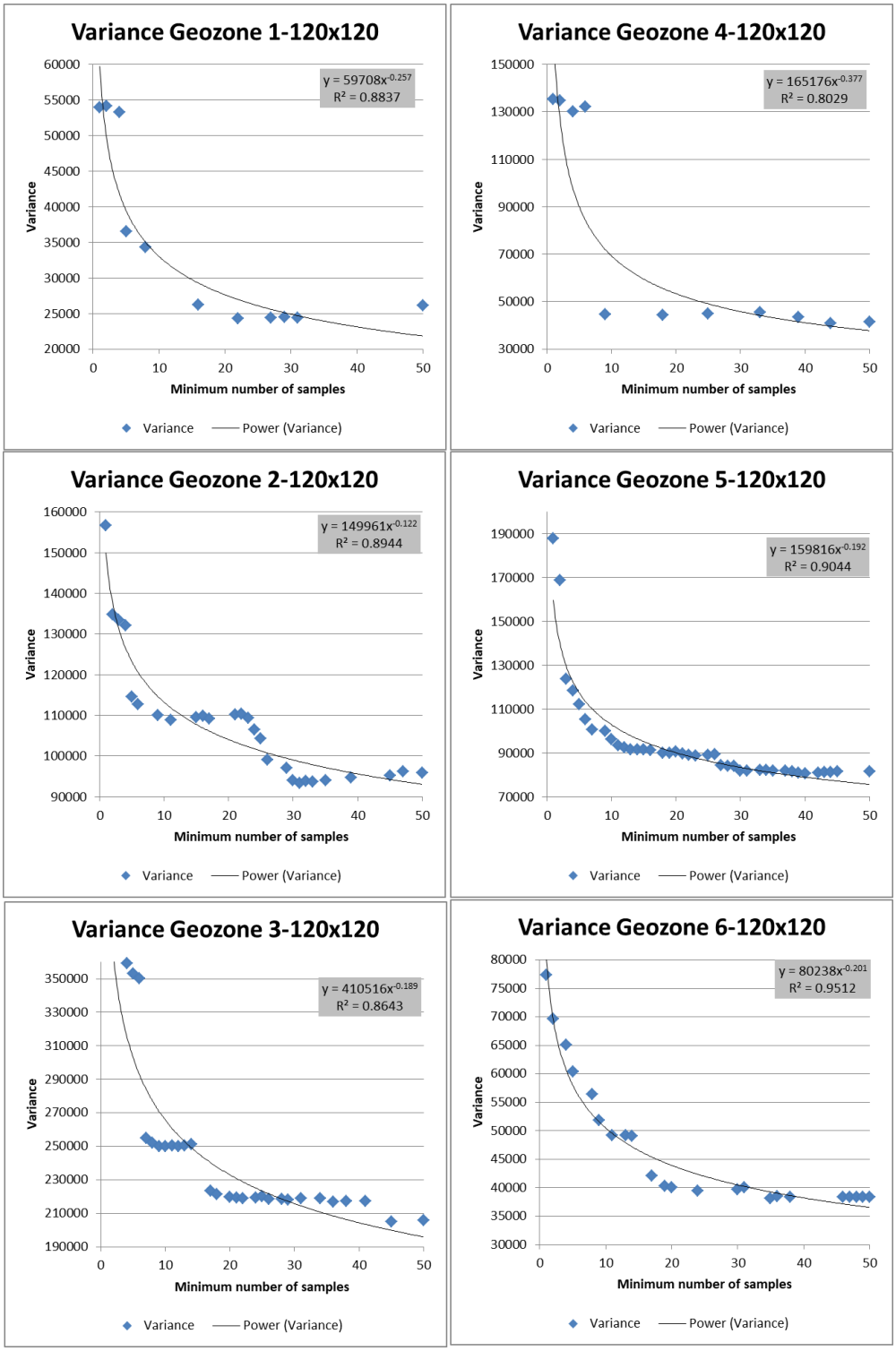


Figure 62: *Variance versus minimum number of samples for the MKNUG calculation for 120m x 120m blocks*

8 MODEL VALIDATIONS

The Mineral Resource model was checked by using several techniques based on work by Glacken and Snowden (2001):

- Jack-knifing test to compare estimates with all data to estimates with removed data (Figure 63)
- Histograms of the input data were compared with the histograms of the block estimates for each geozone
- The mean grades of the input data were compared to the mean grades of the block estimates
- Graphical plots of the input data and the block model with the same legends were created and checked for similarity
- Block factors between the input data and the block estimates were calculated (Table 11)
- Measures of conditional bias empirical and from the block model – Slope of Regression and Kriging Efficiency

8.1 Jack-knifing

To test the validity of the variogram model fitted to the experimental variogram the methodology of jack-knifing (cross-validation) can be followed. This consists of eliminating data temporarily and then kriging the area of the removed data with the remaining samples.

Blocks with a size of 2500m^2 were created and placed at random locations throughout the geozones. If blocks are small, fewer data will be deleted, and if blocks are big, more data will be deleted. If blocks are too small, the small amount of deleted data will not have a big influence on the estimated values. The block size in this exercise was chosen to be the same as the area that one mining crew takes to mine in about 6 months. This provides a good indication of how accurately the estimations area for 6 months' worth of mining from current sampling information.

The blocks were placed on mined-out locations that were covered by sampling on regular intervals. The data in these blocks were deleted and the Mineral Resource block model was re-created estimating values for these blocks. The valuation results of the blocks that were valuated with the deleted data were then compared to the valuation results of the normal Mineral Resource block model with all the data included.

The blocks were spaced far apart to ensure that deleted data of one block does not affect the estimation of a neighbouring block. This meant that the block model with the deleted data had to be re-created several times after blocks were moved to new locations, a very tedious exercise. The comparison of the results with and without the data per geozone is displayed in Figure 63.

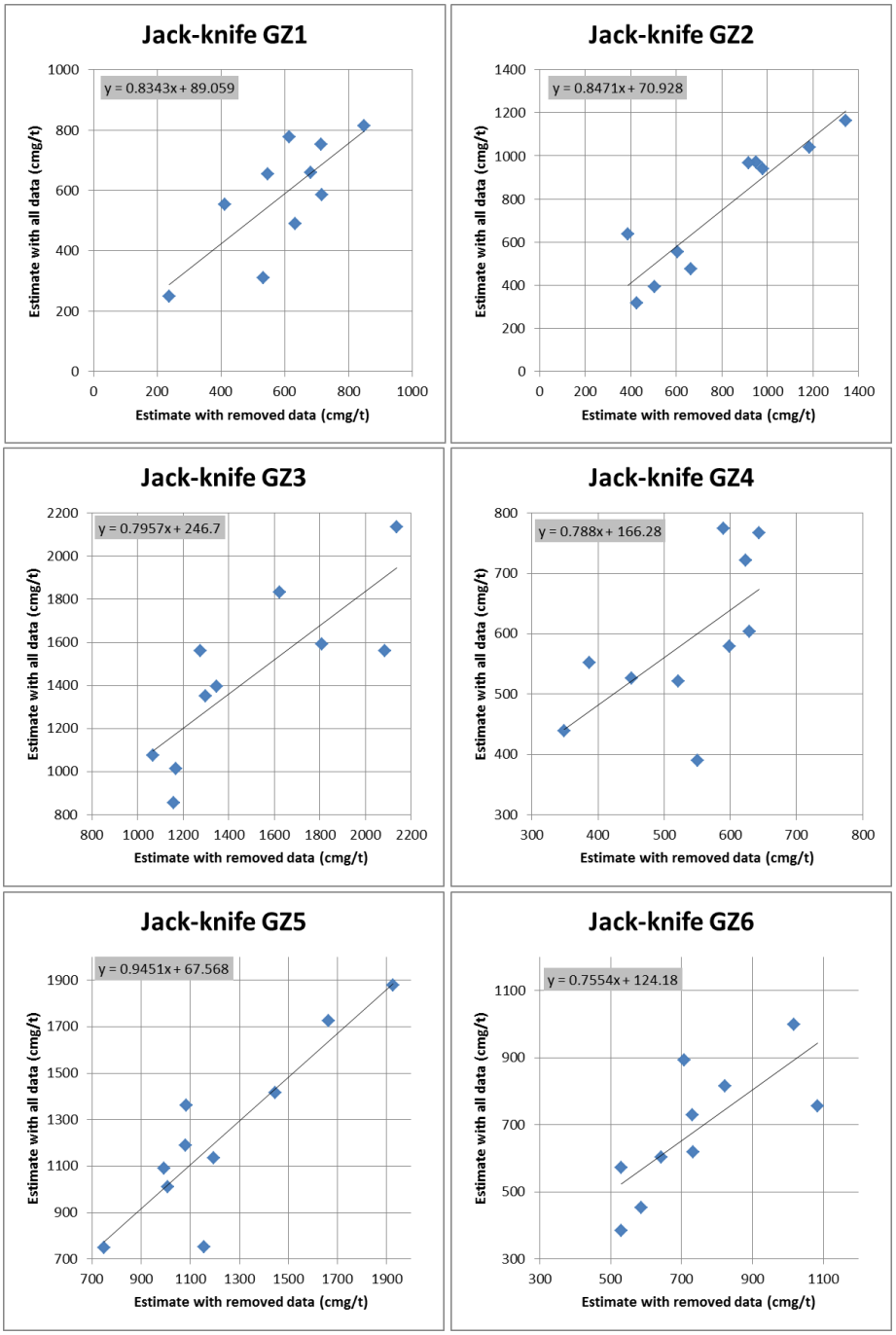


Figure 63: *Results of 60 Jack-knife blocks comparing the estimate with all data present against the estimate with the deleted data*

The results of the jack knifing exercise are summarised in Table 14. The differences between the estimates with data in the blocks and those estimated by jack knifing in the geozones are within 9%; with the two low-grade geozones GZ4 and GZ6 accounting for the greatest negative and positive % differences respectively.

Table 14: *Jack-knife results per geozone*

GZ	Estimate (all data) (cmg/t)	Jack Knife Estimate) (cmg/t)	% Difference	Correlation Coefficient	Slope of Regression
1	584	593	+1.6%	0.76	0.83
2	745	796	+6.8%	0.91	0.85
3	1437	1496	+4.1%	0.81	0.80
4	587	534	-9.0%	0.63	0.79
5	1230	1230	0.0%	0.89	0.95
6	682	739	+8.3%	0.75	0.76

The correlation coefficient is above 0.75 except for the low-grade geozone, GZ4. The regression slopes in all the geozones are above 0.75. The overall results are considered acceptable.

8.2 Distribution Comparison

The cmg/t distributions of the models were compared to that of the input data. Due to a degree of smoothing and the change of support effect that are inherent in kriging the distribution of the estimated blocks is less skewed than the distribution of the input data for the variable of interest.

The change of support effect, is due to the fact that samples have a higher variability than blocks, as the volume increases the variance decreases. The spike at the end of the tail of the input data distributions represents the 5% of the values that were capped.

The cmg/t distribution comparison of GZ1 is shown in Figure 64 below. The distribution of the model is less skewed than that of the capped input data. The spike in the tail of the distribution of the capped data is not present in the distribution of the block model estimates. This observation could indicate that the high values that were capped do not occur in clusters in GZ1 but are scattered throughout the geozone and few blocks are estimated around the capped value. The 30m x 30m block estimate of GZ1 shows some evidence of

bimodality which was also observed in Figure 59 where the annotated block estimates were mainly in two of the colours of the specified legend for cmg/t values.

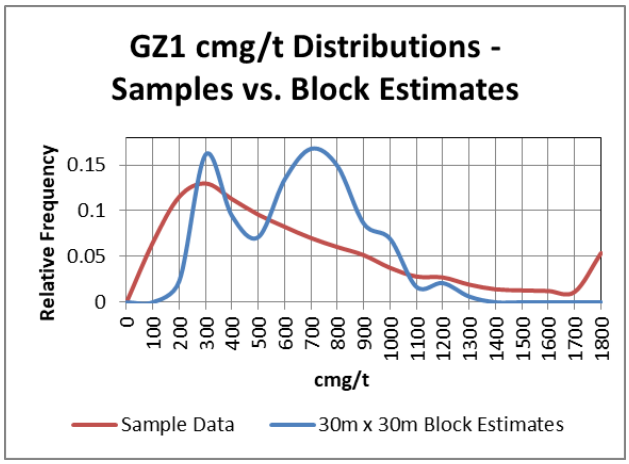


Figure 64: *GZ1 cmg/t distribution comparison of the 30m x 30m block estimates and the sample data*

The cmg/t distribution comparison of GZ2 is shown in Figure 65 below. The same observation that was made in GZ1 can be made for GZ2 where the 5% values that were capped are scattered throughout the geozone. The change of support effect is more pronounced.

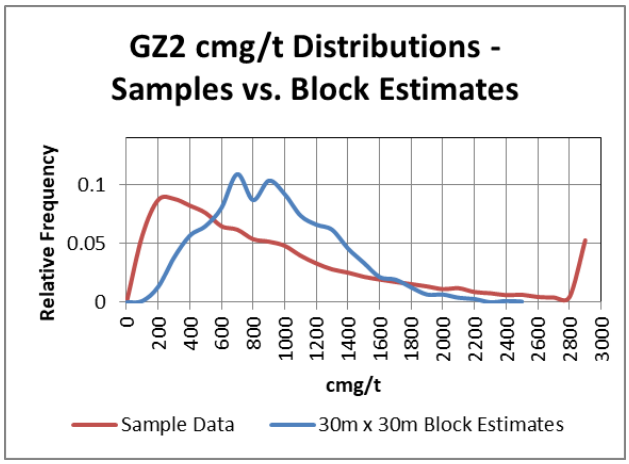


Figure 65: *GZ2 cmg/t distribution comparison of the 30m x 30m block estimates and the sample data*

The cmg/t distribution comparison of GZ3 is shown in Figure 66 below. For GZ3 there is no spike in the tail of the distribution of estimates. It has a long tail and the change of support effect is as visible as in GZ1 and GZ2. This

shows that some of the estimated blocks for the geozone were as high as the capped value and that the high-grade sampling values might occur in clusters.

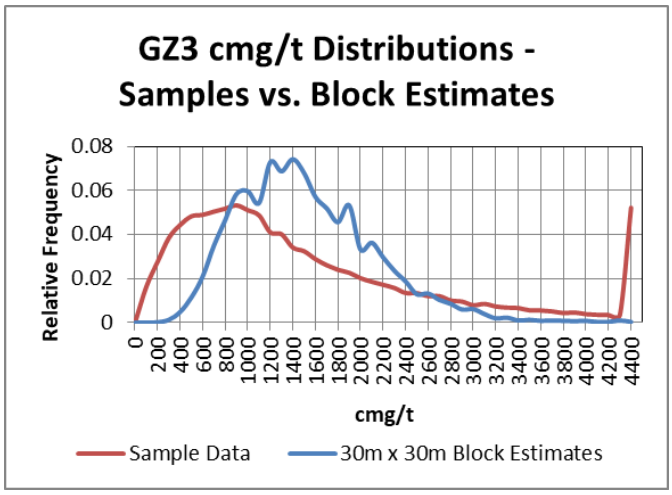


Figure 66: *GZ3 cmg/t distribution comparison of the 30m x 30m block estimates and the sample data*

The cmg/t distribution comparison of GZ4 is shown in Figure 67 below. The same observations that were made for GZ1 and GZ2 can be made for GZ4.

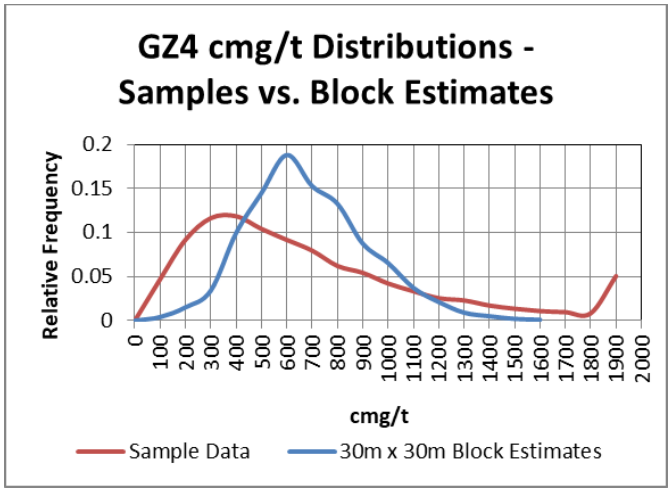


Figure 67: *GZ4 cmg/t distribution comparison of the 30m x 30m block estimates and the sample data*

The cmg/t distribution comparison of GZ5 is shown in Figure 68 below. The spike in the tail of the capped data distribution indicates that these capped values are also scattered throughout the geozone.

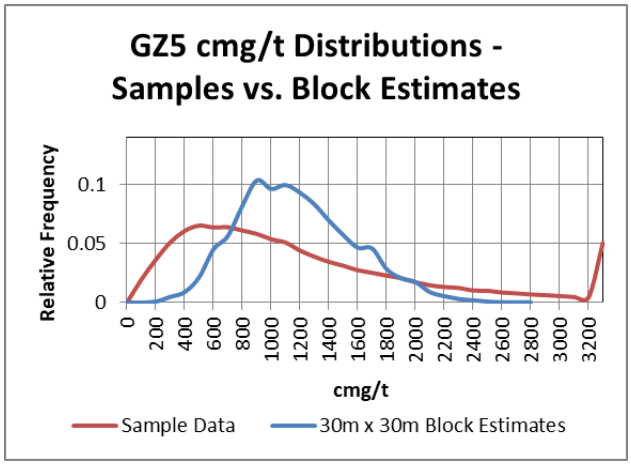


Figure 68: *GZ5 cmg/t distribution comparison of the 30m x 30m block estimates and the sample data*

The cmg/t distribution comparison of GZ6 is shown in Figure 69 below. For GZ6, the same observation was made in GZ5 where these high values are scattered throughout the geozone.

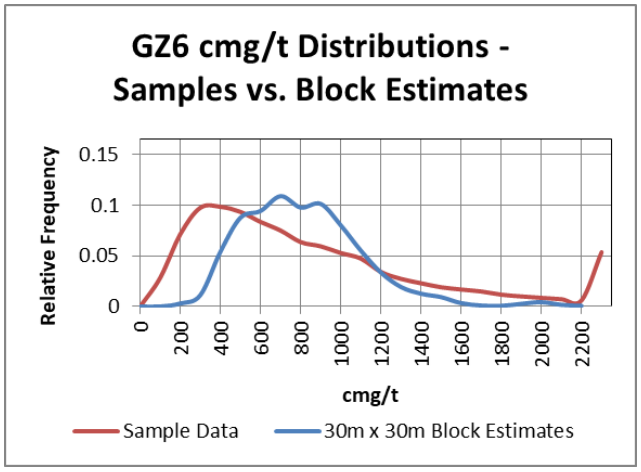


Figure 69: *GZ6 cmg/t distribution comparison of the 30m x 30m block model estimates and the input data*

8.3 Mean Grade Comparison

The mean grades of the input data per geozone were compared to the mean grades of the model. It is important that the mean grade of the model is similar to the mean grade of the input data. If for instance the mean grade of the model is significantly higher than the mean grade of the input data, the model is estimating grades that are not going to be achieved when mining. If the means differs too much, the model is not representative of the data and the parameters and neighbourhood of the model must be revisited (Glacken & Snowden, 2001).

The mean cmg/t comparison in Figure 70 below shows that the average value of the model is slightly less than the average value of the input data for each geozone.

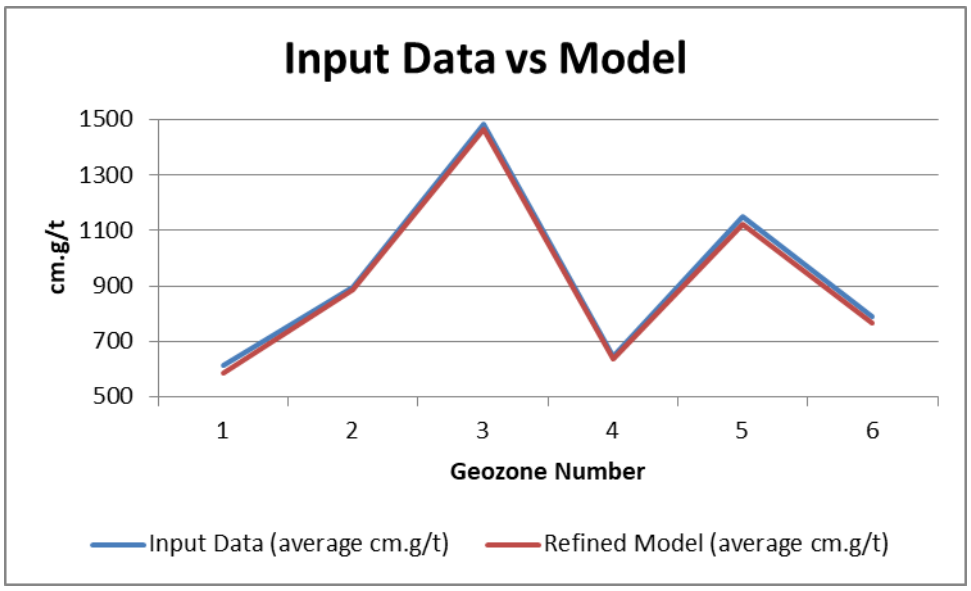


Figure 70: *Average cmg/t value of the input data compared to the average value of the model per geozone for the 30m x 30m blocks*

The measures of conditional bias for the 30m x 30m block model are shown in Table 15 below.

Table 15: *Conditional bias measures for 30m x 30m blocks*

Geozone	Average Slope of Regression 30m x 30m blocks	Empirical Slope of regression Estimated as the Regression Slope of the "actual" block value given the estimated Block value (Figure 57)	Average Kriging Efficiency
1	0.81	0.89	0.73
2	0.80	0.78	0.83
3	0.82	0.78	0.76
4	0.80	0.77	0.80
5	0.81	0.82	0.81
6	0.83	0.78	0.81

The kriging efficiencies are high in all the geozones and confirms a low kriging variance. The kriging efficiency of a perfect estimation is 1, however, in practice there is never a perfect correlation (Deutsch et al., 2014).

9 MINERAL RESOURCE CLASSIFICATION

The classification of Mineral Resource and Mineral Reserve blocks is done based on the degree of geoscientific knowledge and confidence. As the level of confidence increases the classification of Mineral Resources can change from Inferred to Indicated, and from Indicated to Measured.

The generated model was annotated according to block size to display the Mineral Resource categories. The 30m x 30m blocks are red, the 60m x 60m blocks are green and the 120m x 120m blocks are yellow (left portion of Figure 71).

Blocks within the Measured Resource confidence are 30m x 30m blocks and these have a slope of regression of more than 0.7. All the 30m x 30m blocks, that had a slope of regression less than 0.7 were downgraded to the Indicated category by replacing them with 60m x 60m block estimates. This is done to comply with the Harmony standard which states that any estimation block that has a slope of regression less than 0.7 should be “cut away” and excluded from the Measured model.

Once the model was annotated on the Mineral Resource category legend, halos strings were created. The halos are used as a guide for the classification of Mineral Resources according to the requirements of the Harmony standard. The Measured confidence halo string falls on the edge of the 30m x 30m model blocks. The Indicated halo string falls on the edge of the 60m x 60m model blocks.

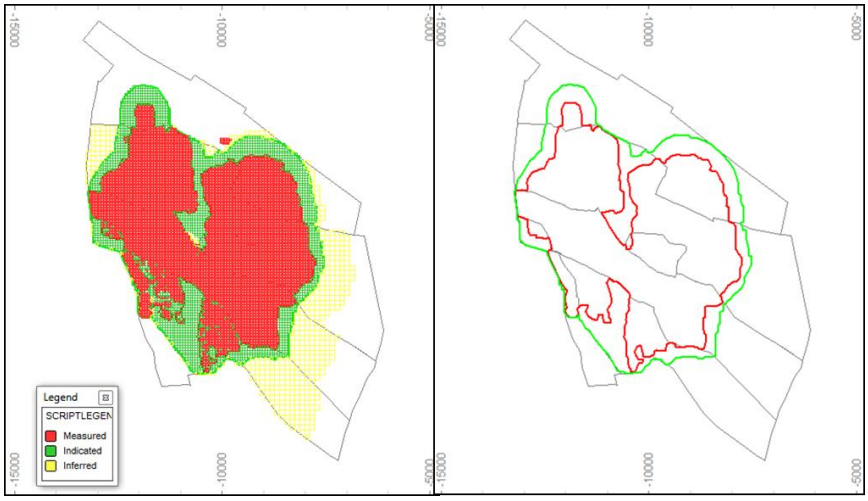


Figure 71: *Halos for Mineral Resource classification and the Mineral Resource block model annotated on Mineral Resource category (Left). The halo strings (Right)*

The halos are purely based on geostatistical parameters, so the confidence and understanding of geological factors may override the halos in some areas. These areas are investigated by a competent team and changes to halos are signed off by the shaft Geologist, the Ore Reserve Manager and the Geostatistician.

The difference between the Mineral Resource estimates of the old model and the refined model are shown in Table 16 below. The comparison was done between the total block models at zero cut-off and was split into the confidence limit categories (Measured, Indicated and Inferred).

Table 16: *Comparison between the total block models (refined model and old model)*

Confidence Category	Measured		Indicated		Inferred	
Mineral Resource Model	Refined Model	Old Model	Refined Model	Old Model	Refined Model	Old Model
Mean (cmg/t)	1129	1086	844	935	1037	995
Standard Error	3.8	2.8	6.4	5.5	13.5	17.0
Median (cmg/t)	1041	1025	773	936	1038	1053
Standard Deviation	533	453	315	232	306	278
Sample Variance (cmg/t)²	283 713	205 445	99 137	53 906	93 522	77 487
Coefficient of Variation	0.47	0.42	0.37	0.25	0.29	0.28
Kurtosis	1.9	1.5	2.2	-0.3	-0.2	-0.6
Skewness	1.0	0.9	1.2	0.2	0.1	-0.4
Range (cmg/t)	4329	3840	2779	1600	1975	1425
Minimum (cmg/t)	8	52	187	416	338	313
Maximum (cmg/t)	4337	3892	2966	2016	2313	1738
Number of blocks	19954	25604	2399	1797	516	268

The refined kriging neighbourhood caused the Measured block model to have an increase and the Indicated block model to have a reduction in value. The value of the refined Inferred block model is higher than that of the old model because many Inferred blocks receive a global mean value, and as previously stated, the global means of the refined model is higher than that of the old model (Table 13).

10 RECONCILIATION BETWEEN THE OLD AND THE NEW MODEL

The Mineral Resource model of 2016 was compared to the new refined model. The comparison in Figure 72 shows the difference in the average estimated cmg/t between the two models for each geozone.

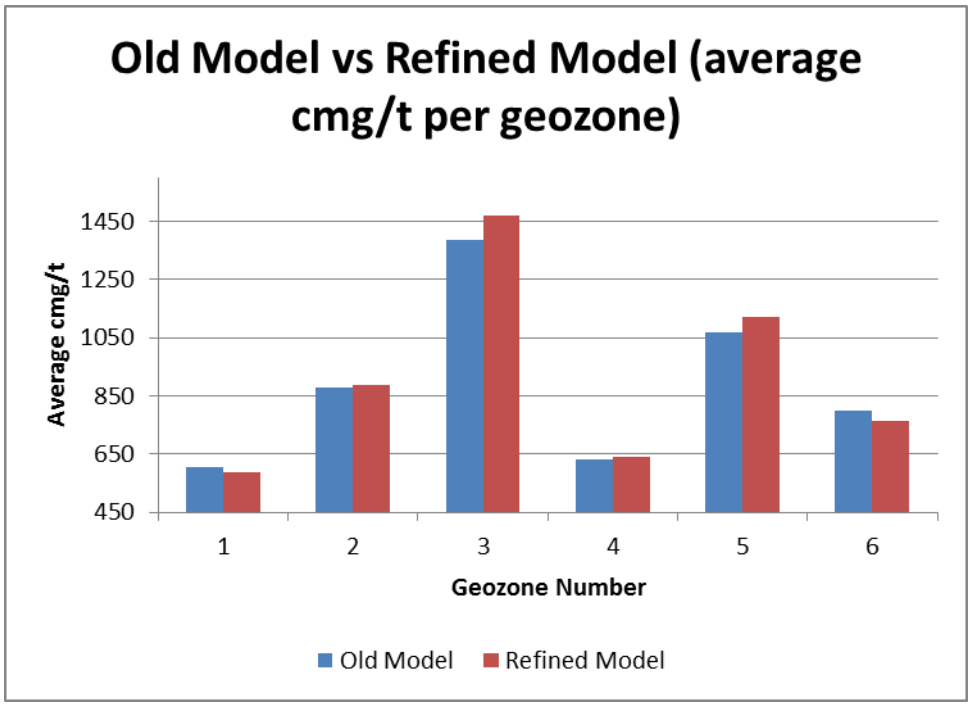


Figure 72: *Reconciliation of cmg/t values between the refined and the old model*

The estimation of the value for GZ1 dropped with 3.3%. GZ1 is mainly covered with Inferred blocks that received a global mean value, and the drop in the global mean value influenced the value of the geozone negatively.

The zonal mean cmg/t value of GZ2 increased slightly by 0.7%. GZ2 contains a large amount of historical data that was capped on two different values. Using the uncapped historical data caused a difference in the global mean value.

The estimated mean cmg/t value of GZ3 increased by 5.8%. This geozone contained a lot of data of Phakisa mine that was removed from the data base. The historical data in this geozone was capped on three different values. Using the correct and uncapped data caused the change of global mean value.

GZ4 had a small change in zonal mean cmg/t value with an increase of 0.8%. This is the smallest geozone and does not contain any Inferred areas. The small amount of historical data did not have a negligible effect on the mean cmg/t value.

The estimated mean cmg/t value of GZ5 increased by 5.3%. GZ5 is a large geozone that contains a lot of historical data and Inferred areas. A small amount of data in GZ5 became a part of GZ6. The higher the global mean cmg/t value and the uncapped historical data caused the increase in value.

The new GZ6 is a combination of the old GZ6 and GZ7. A small portion of GZ5 also became a part of GZ6. The estimated value of GZ6 dropped by 4.6% since low-grade data from the northern part of GZ5 and the low-grade borehole data of GZ7 was included in the new GZ6.

A Q-Q plot comparing the cmg/t values of the old and the refined model (30m x 30m) is shown in Figure 73 below.

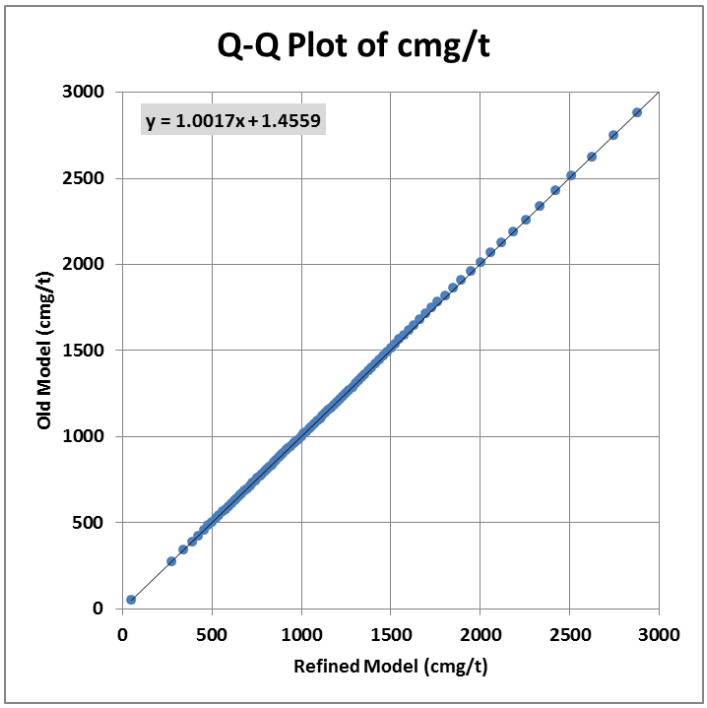


Figure 73: *A Q-Q plot of cmg/t values between the old and the refined model (30m x 30m)*

The Q-Q plot shows that the cmg/t points only deviate slightly from a straight line, these small deviations indicates that the old and the refined model do not have any major differences in cmg/t values.

A grade tonnage comparison was done between the estimations of the refined and the old model and is shown in Figure 74 below.

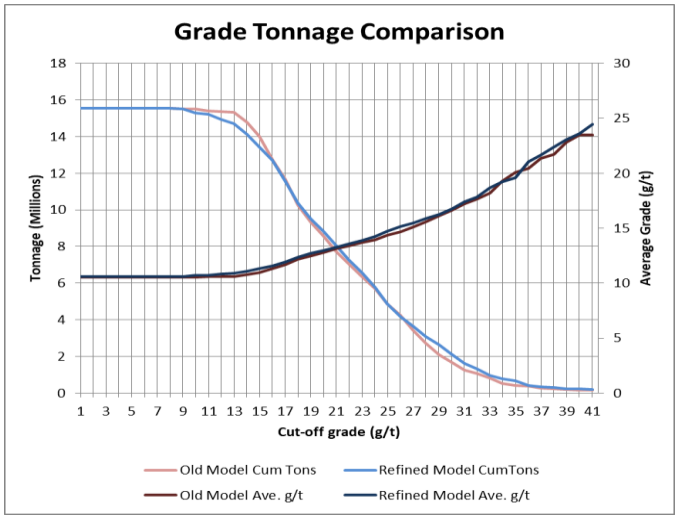


Figure 74: *Grade tonnage comparison between the estimations (Measured blocks) of the refined and the old model*

The refined model shows an increase of 0.06g/t in grade and 989kg of gold when considering a 0g/t cut-off grade. The refined model also shows an increase of 0.34g/t in grade and 849kg of gold when considering a 6.5g/t cut-off grade. The summary of the comparison is shown in Table 17 below.

Table 17: *Comparison of grade, tonnage and gold content between the refined and the old block model (Measured blocks).*

		Refined Model	Old Model
Resource estimate @ Cut-off of 0g/t	Average Grade above Cut-off (g/t)	10.59	10.53
	Metric tonnes above Cut-off	15 544 109	15 544 109
	Content above Cut-off (kg)	164 632	163 643
Resource estimate @ Cut-off of 6.5g/t	Average Grade above Cut-off (g/t)	11.10	10.77
	Metric tonnes above Cut-off	14 119 592	14 783 079
	Content above Cut-off (kg)	62 552	61 703

The improved estimate of the Mineral Resource shows an increase in gold content and higher average grade, all due to a clean database, kriging neighbourhood selection and improved domaining.

11 CONCLUSION AND RECOMMENDATIONS

The research objectives were met by improving the quality of the input data, the accuracy of the geological domains, improvement of the nugget effect modelling and the kriging neighbourhood parameters which ultimately produced the refined model with improved Mineral Resource estimation and within the classification categories for Tshepong mine.

Sometimes not all the data gathered is used because it was gathered for another purpose e.g. check sampling, this data that can however, be exploited to improve the quality of the Mineral Resource model. If the data is available, time should be made for analysing all data and uncover the underlying features and incorporate it to current thinking and methods.

Below is a summary of the important actions, findings and conclusions:

Validation and Changes to the input data:

Several changes were made to the input data to include only relevant data. The sampling data of Phakisa mine was removed from the Tshepong data base due to the incompatible support size of the samples of the two mines.

The historical data file was corrected by replacing the hard-coded capped values with the original uncapped assay values. The uncapped values have higher cmg/t values, and this had an impact on local and global mean estimations in some areas of the mine.

Using the uncapped data base had affected the local estimations in the older areas of the mine that are currently blocked as pillars in the Mineral Resource.

Facies plan:

The updating process of the facies plan was based on the information of conglomerate descriptions of mined-out areas and new bore holes. The resulting new facies plan had an impact on the geozone data changes and assisted with the grouping of data with similar characteristics.

Changes to geozone boundaries:

Changes to the geozone boundaries were made in the northern parts of the mine. These changes were made based on the facies plan that confirmed areas of geological homogeneity. Statistical analyses of the new geozones confirmed stationarity within each geozone, and that every geozone has its own population characteristics and type of grade distribution. The geozone changes had a positive impact on the consistency of the data within each geozone and improved the confidence in the Mineral Resource estimations.

The updating of the facies plan led to changes of the geozone boundaries which had an impact on the 2017 Mineral Resource and Mineral Reserve blocks. The blocks that were in GZ6 before the geozone changes became part of GZ5 after the geozone changes. These blocks were mostly below the cut-off value which excluded them from the LOM scheduling. These blocks are now situated in a higher grade geozone and will now be included in the next LOM plan.

Nugget effect estimation from an additional point on the variogram:

The semi-variance analysis of the check sample data provided additional experimental variogram information at 0.08m that was used in enhancing of the modelling of the nugget effect of the variograms for each geozone. This information provided a point closer to the origin of each variogram and provided more confidence in the estimation of the nugget effect. The nugget effect has the most impact on the kriging estimation process, where weights are assigned to samples according to their proximity to the block being estimated as well as their spatial relationships with the other samples being considered in the estimation of the block. The estimation process has been improved by the more accurate nugget effect estimation.

Analysis of spatial continuity:

Some of the geozones were estimated by using anisotropic variograms. The selection of short and long-range directions was based on the results of the

spatial continuity analysis (swath analysis and variogram contour maps). The variogram contour maps were used in conjunction with the experimental semi-variograms to choose the most appropriate direction of continuity.

Optimisation of kriging neighbourhood parameters:

The QKNA assisted with the selection of the best kriging neighbourhood parameters such as search volumes, number of samples, discretisation, which minimised conditional bias and improved estimation variance.

Improvement of the Mineral Resource estimate:

The classification of Mineral Resource blocks is dependent on the geological model and the ranges of the variograms of the 30m x 30m, 60m x 60m and 120m x 120m block models. The geological model was improved by correcting the input data, updating of the facies plan, and the adjustment of the geozone boundaries. The variograms were improved by a better estimation of the nugget effect, the analysis of spatial continuity, and a better selection of a kriging neighbourhood through the QKNA process.

The way forward:

For Tshepong mine to benefit from this research, the following should form part of the annual geostatistical Mineral Resource estimation updates:

- The uncapped historical data file must be used instead of the capped file
- Separate models must be created for Tshepong and Phakisa mine due to the incompatibility of samples; the models represent distinctly different parts of the reef
- The facies plan must be updated before considering any geozone changes, and if any geozone changes are made, stationarity within the zones must be confirmed by statistical analysis
- The check sample data must be analysed before updating the variogram models to ensure the best estimation of the nugget

- Swath plots and variogram contour map analysis should form part of the analysis of spatial continuity
- QKNA techniques must be applied during the optimisation process of the kriging neighbourhood
- Jack-knifing should form part of the model validation process
- The Mineral Resource classification methodology should be improved
- Care should be taken with the determination of the nugget variance equations to be used for input into the MKNUG in the macro kriging process. The equation of the line is influenced by each variance point on the graph, and only points that does have a corresponding minimum number of samples should be plotted.
- A question still remains namely why does the standard procedure for 120m x 120m blocks also use as maximum only 50 samples in the block to capture the variability at this block size.

REFERENCES

Armstrong, M., 1998. *Basic Linear Geostatistics*. Illustrated ed. Springer Science & Business Media.

Clark, I., 1979. *Practical Geostatistics*. Alloa: Geostokos Limited.

Deutsch, C.V., 2007. *The slope of regression for kriging estimators*. University of Alberta.

Deutsch, C. & Schnetzler, E., 2009. *Statios*. [Online] Available at: http://www.statios.com/help/varmap.html [Accessed 30 June 2016].

Deutsch, J.L., Szymanski, J. & Deutsch, C.V., 2014. Checks and measures of performance for kriging estimates. *The Jounal of The South African Institute of Mining and Metallurgy*, 114(March 2014), pp.223-29.

Dohm, C.E., 1995. *Improvement of Ore Evaluation Through the Identification of Homogeneous Areas Using Geological, Statistical and Geostatistical Analyses.* PhD Thesis. Johannesburg: University of te Witwatersrand.

Dohm, C.E., 2015 A. *MINN 7007 - Statistical Valuation of Ore Reserves*. Course lecture notes. Johannesburg: School of Mining Engineering University of Witwatersrand.

Dohm, C.E., 2015 B. *MINN 7006 - Geostatistical Methods in Mineral Resource Evaluation*. Course lecture notes. Johannesburg: School of Mining Engineering University of Witwatersrand.

du Toit, R., 2015. *Practical Implementation of Statistical and Geostatistical Techniques on Tshepong Mine's Geozone 5*. Johannesburg: University of the Witwatersrand.

Emery, X. & González, K.E., 2007. Probabilistic Modelling of Litological Domains and its Application to Resource Evaluation. *The Jounal of The South African Institude of Mining and Metallurgy*, 107, pp.803-09.

Emery, X. & Ortiz, J.M., 2005. Estimation of Mineral Resources Using Grade Domains: Critical Analysis and a Suggested Methodology. *The Jounal of The South African Institude of Mining and Metallurgy*, 105, pp.247-55.

Freeman, S.R. et al., 1999. *Tshepong Mineralisation Model*. Research. Rock Deformation Research Group.

Glacken, I.M. & Snowden, D.V., 2001. Mineral Resource Estimation. *The Australasian Institude of Mining and Metallurgy*, pp.189-98.

Hengl, T., 2009. *A Practical Guide to Geostatistical Mapping*. Amsterdam: University of Amsterdam.

Isaaks, E.H. & Srivastava, R.M., 1989. *An Introduction to Applied Geostatistics*. New York: Oxford University Press.

Jolley, S.J., Knipe, R.J. & Henderson, I., 2004. Structural Controls on Witwatersrand Gold Mineralisation. *Journal of Structural Geology*, 26, pp.1063-86.

Krige, D.G., 1994. An analysis of some essential basic tenets of geostatistics not always practiced in ore valuations. In *Proceedings Regional APCOM*. Slovenia, 1994. Computer Applications and Operations Research in the Minerals Industries.

Krige, D.G., 1996. A Practical Analysis of the Effects of Spatial Structure and of Data Available and Accessed, on Conditional Biases in Ordinary Kriging. In *Proceedings Fifth International Geostatistical Congress, Geostatistics Wollongong*., 1996.

Krige, D.G., 1996A. A basic perspective on the roles of classical statistics, data search routines, conditional biases and information and smoothing effects in ore block valuations. In *Proceedings Conference on Mining Geostatistics*. Kruger National Park, South Africa, 1996A. Geostatistical Association of South Africa.

Krige, D.G., 1996B. A practical analysis of the efects of spatial structure and of data available and accessed, on conditional biases in ordinary kriging. In Baafi, Y.E. & Schofield, N.A., eds. *Proceedings Fifth International Geostatistical Congress*., 1996B. Geostatistics Wollongong.

McKillup, S. & Darby Dyar, M., 2010. *Geostatistics Explained*. New York: Cambridge University Press.

Ortiz, J.M. & Emery, X., 2006. Geostatistical Estimation of Mineral Resources with Soft Geological Boundaries: a Comparative Study. *The Jounal of The South African Institude of Mining and Metallurgy*, 106(8), pp.577-84.

Pitard, F.F., 1993. *Pierry Gy's sampling theory and sampling practice.* Colorado: CRC Press.

Ridge, J.D., 2013. *Annotated Bibliographies of Mineral Deposits in Africa, Asia (Exclusive of the USSR) and Australia*. Revised ed. Elsevier.

Rivoirard, J., 1987. Two Key Parameters When Choosing the Kriging Neighborhood. *Mathematical Geology*, 19(8), pp.851-56.

Russel and Associates, 2014. [Online] Harmony Available at: https://www.harmony.co.za/investors/reporting/annual-reports [Accessed August 2015].

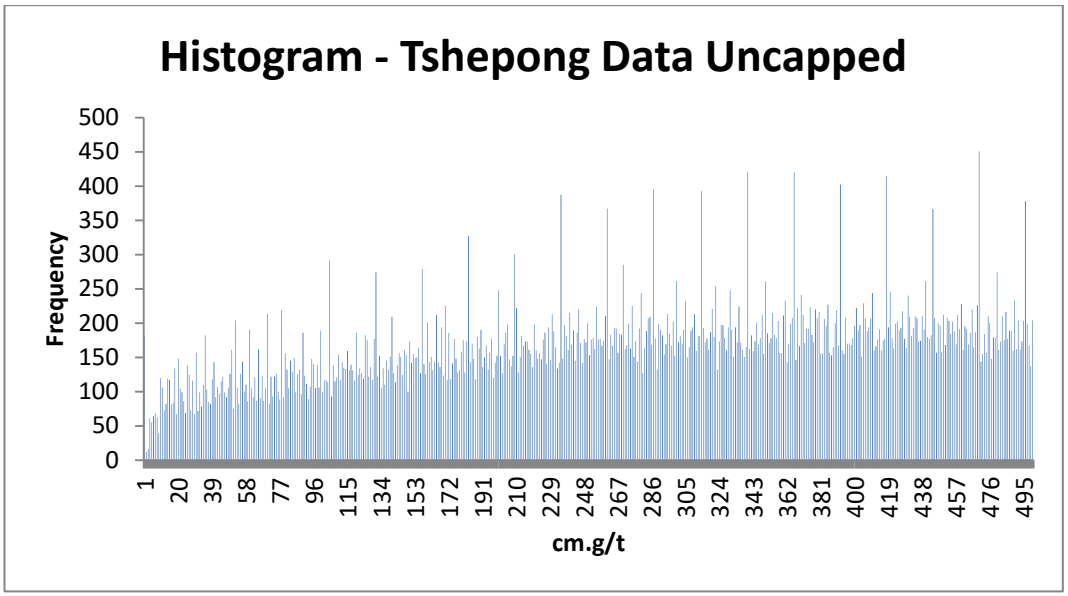
Superior Mining International Corporation, 2011. [Online] Available at: http://superiormining.com/ resources/maps/120123-MangalisaNeighbourOps.jpg [Accessed 2015 August 2015].

Tankard, A. et al., 1982. *Crustal Evolution of Southern Africa*. New York: Springer-Verlag.

Vann, J., Jackson, S. & Betroli, O., 2003. Quantitative Kriging Neighbourood Analysis for the Mining Geologist. In *Fifth International Mining Geology Conference*. Bendigo, 2003.

APPENDIX A HISTOGRAM OF HISTORICAL DATA

Figure A-1: Histogram of the first 500cmg/t of the historical data at a 1cmg/t bin size.



APPENDIX B PROGRESSIVE QAQC GRAPHS

Figure B-1: Scatterplot of AMIS0100 (expected value is 44.97g/t)

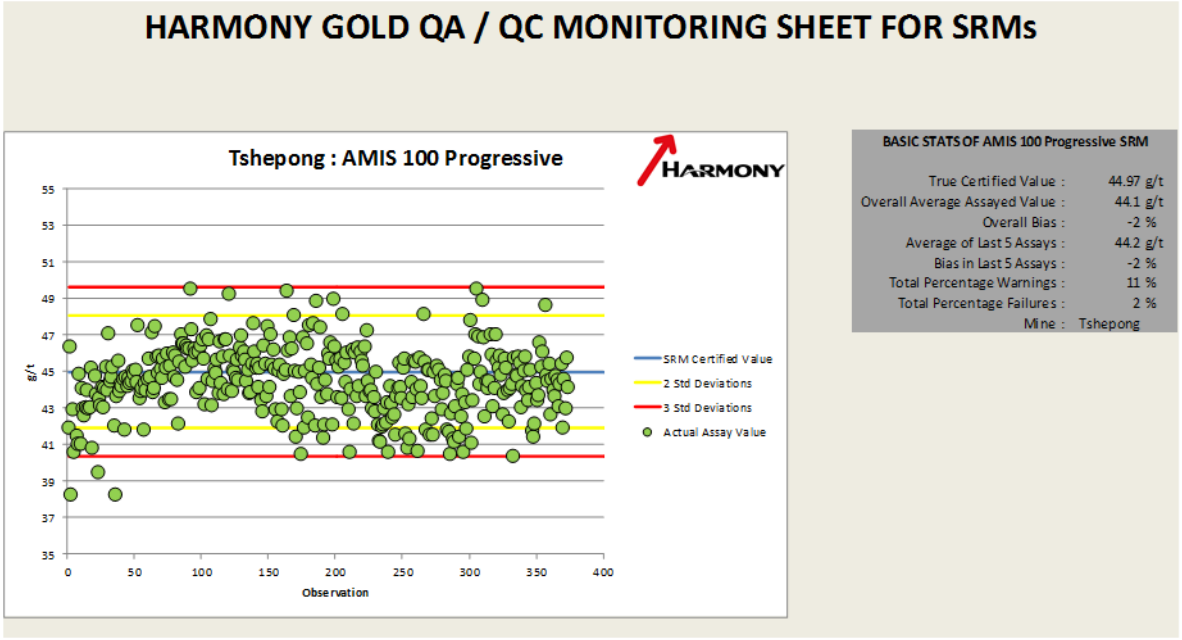


Figure B-2: Scatterplot of AMIS0103 (expected value is 4.73g/t)

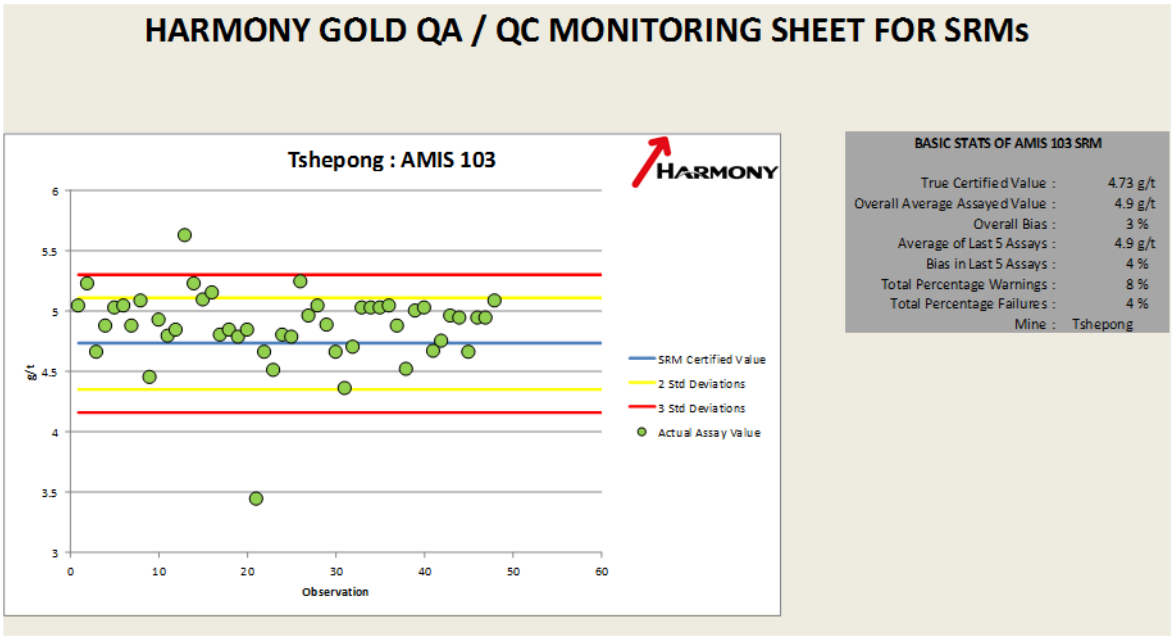


Figure B-3: Scatterplot of AMIS0108 (expected value is 0.063.g/t)

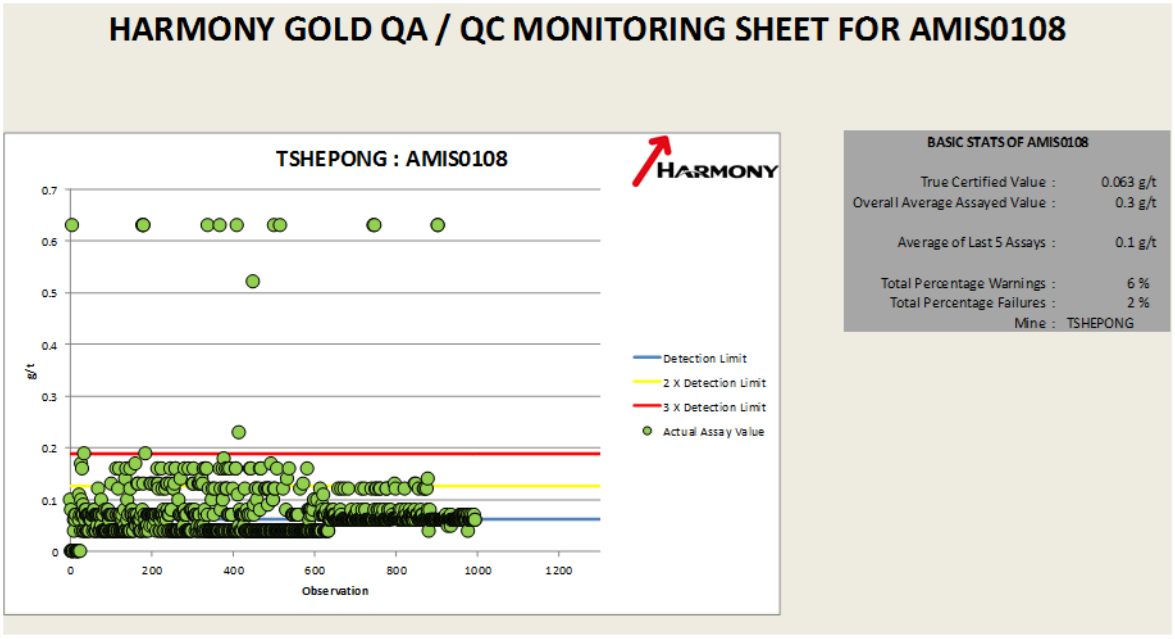


Figure B-4: Scatterplot of AMIS0111 (expected value is 20.48g/t)

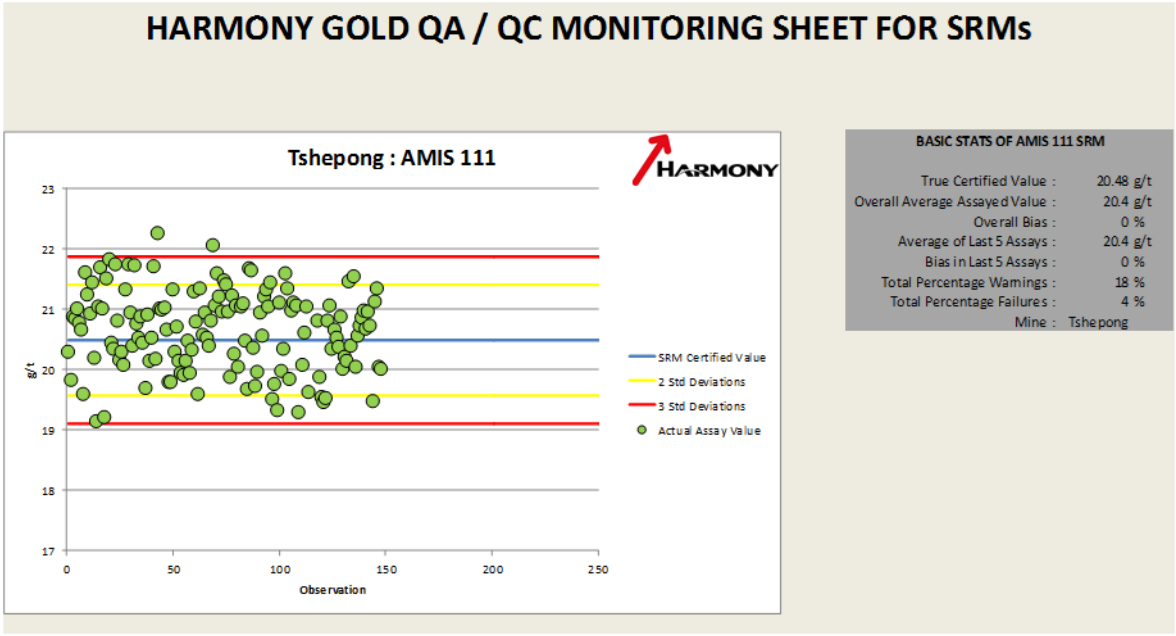


Figure B-5: Scatterplot of AMIS0112 (expected value is 12.27g/t)

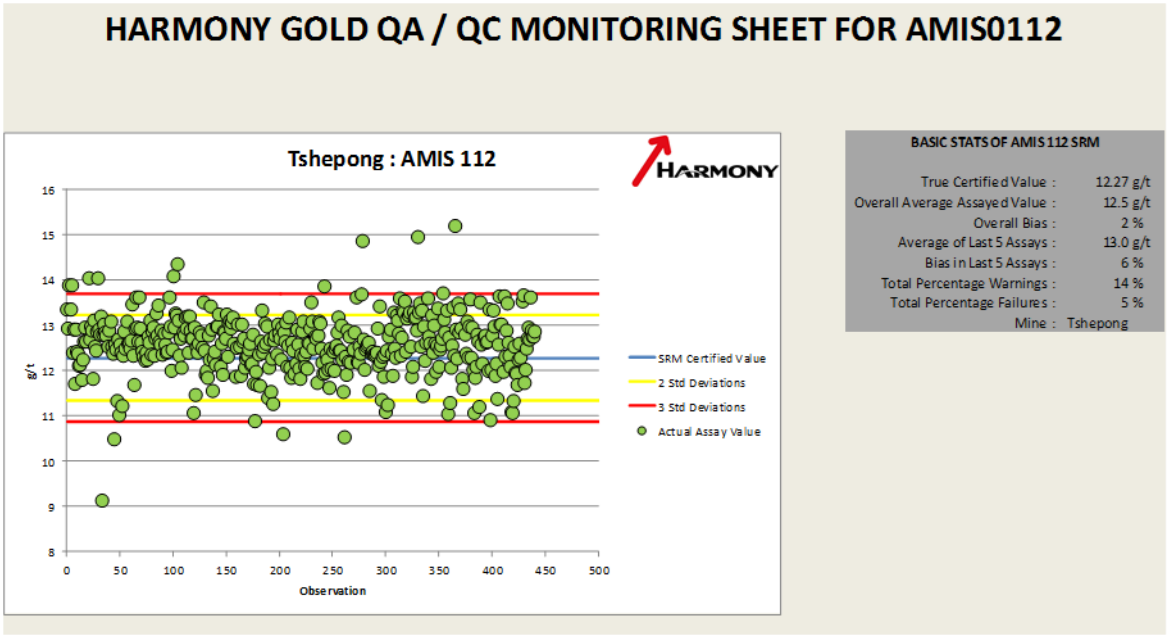


Figure B-6: Scatterplot of AMIS0133 (expected value is 310.00g/t)

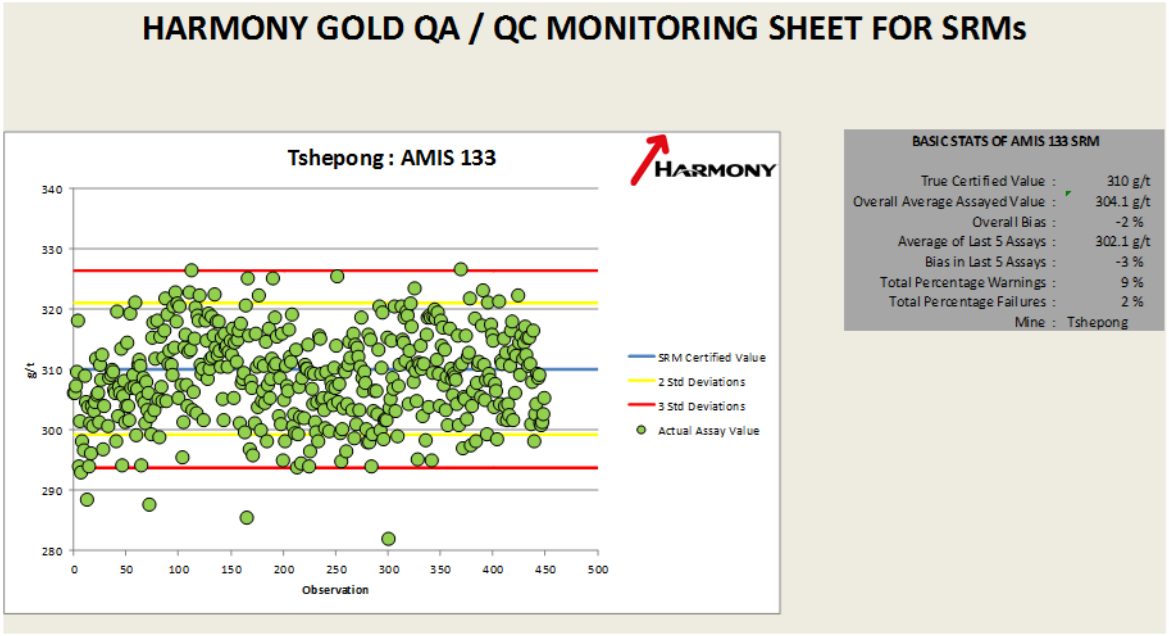


Figure B-7: Scatterplot of AMIS0184 (expected value is 14.28g/t)

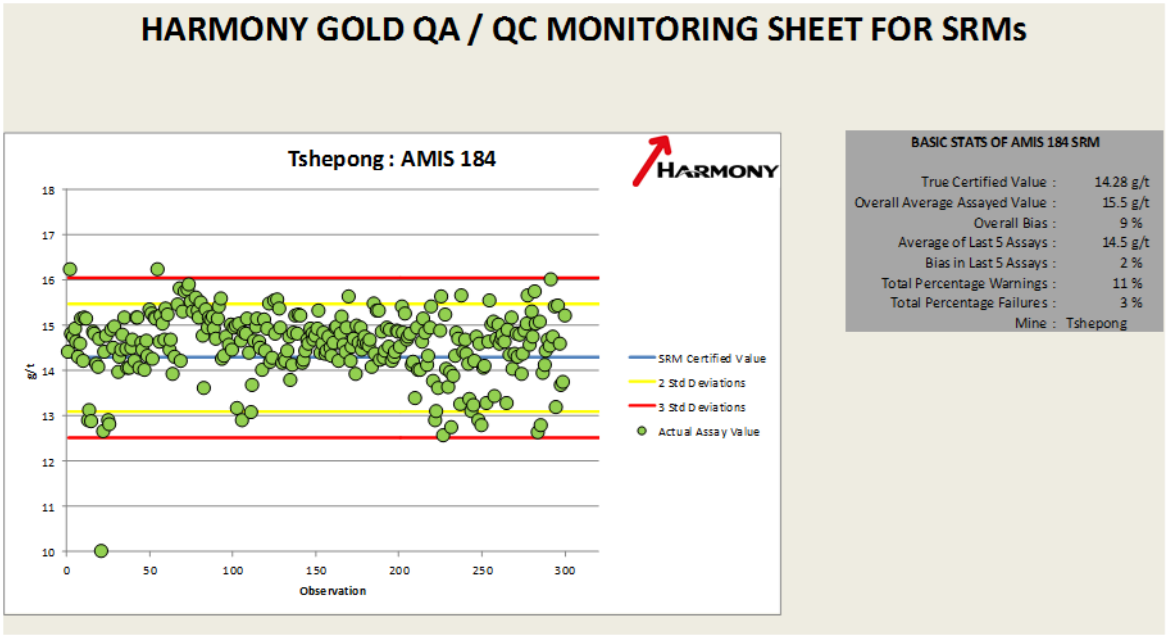


Figure B-8: Scatterplot of AMIS0245 (expected value is 88.42g/t)

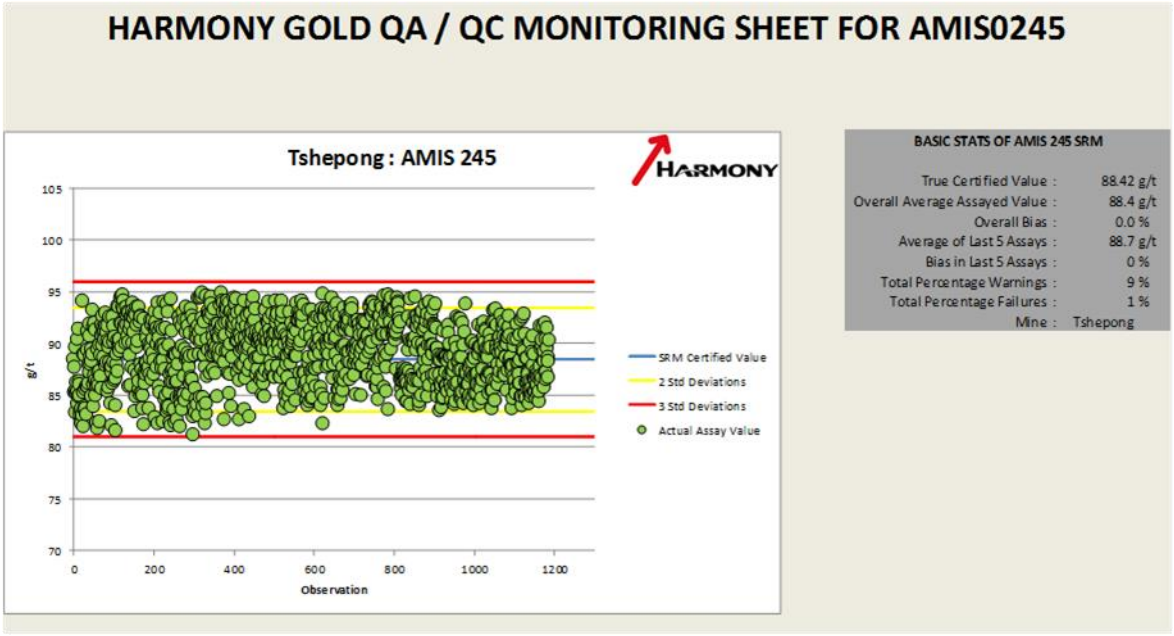


Figure B-9: Scatterplot of AMIS0303 (expected value is 8.78g/t)

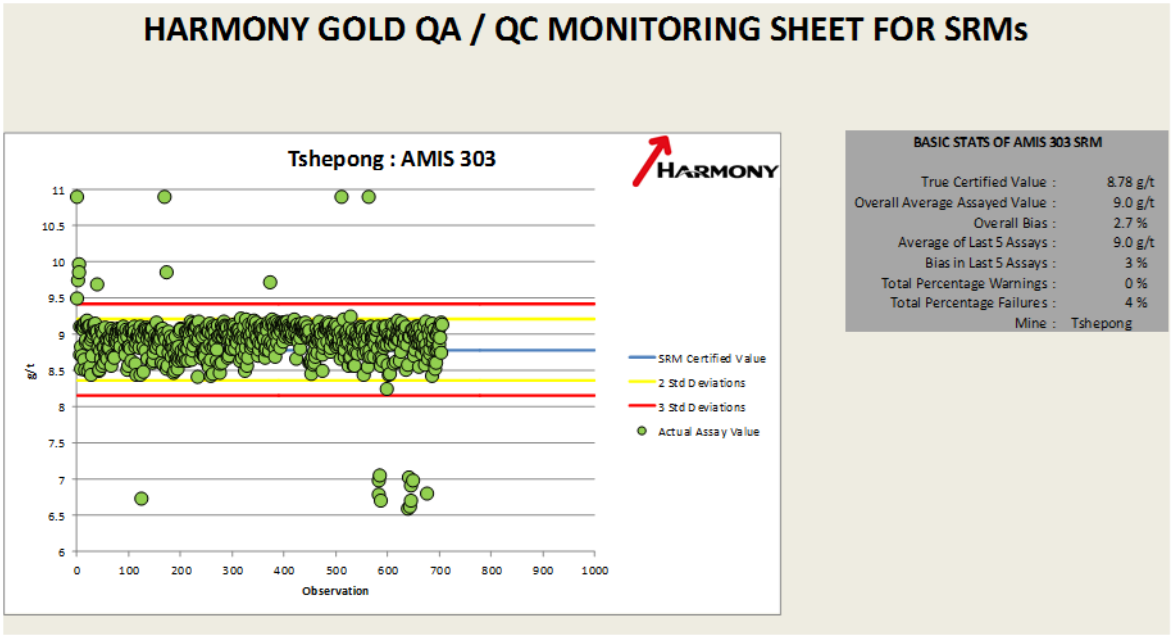


Figure B-10: Scatterplot of AMIS0369 (expected value is 26.36g/t)

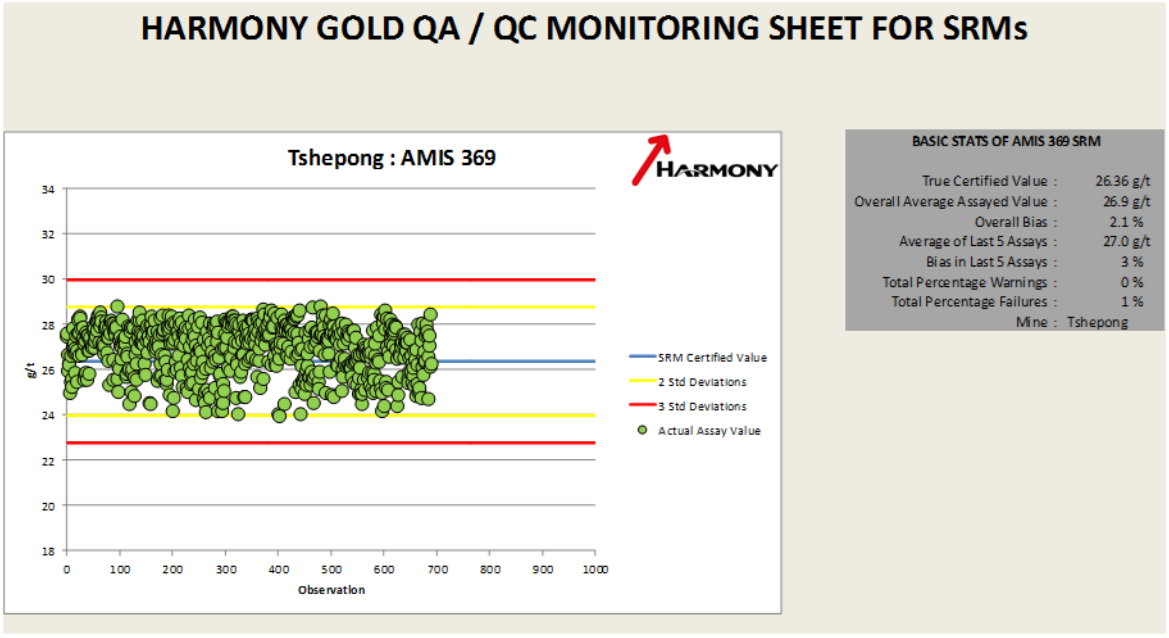


Figure B-11: Scatterplot of Course Blank 19mm (expected value is <0.189g/t)

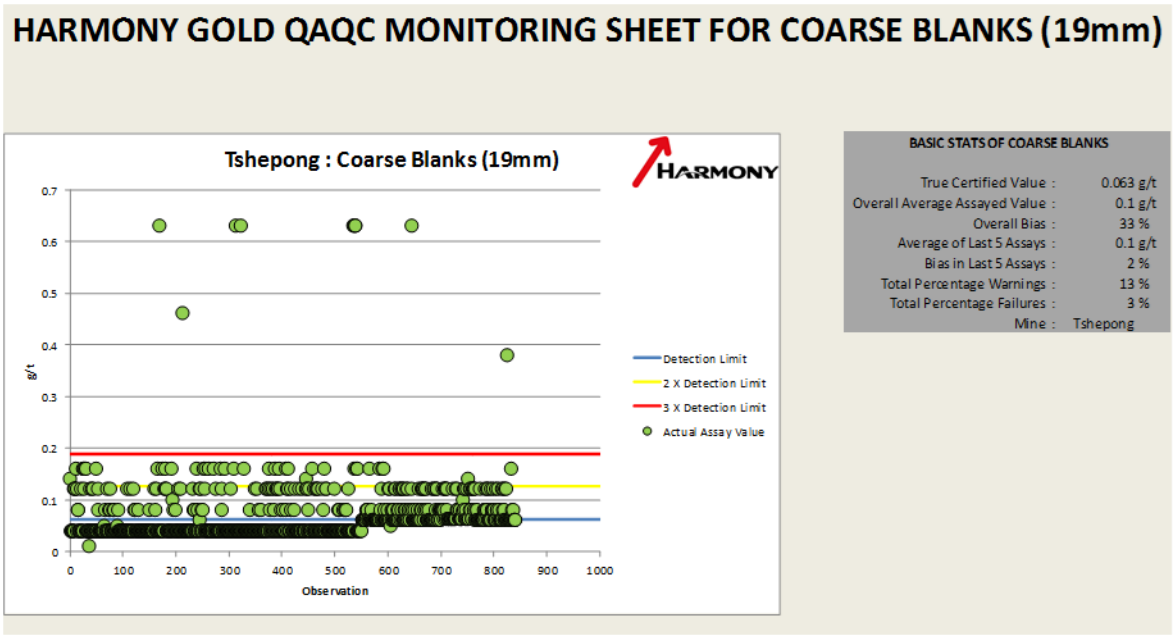


Figure B-12: Scatterplot of Course Blank 4mm (expected value is <0.189g/t)

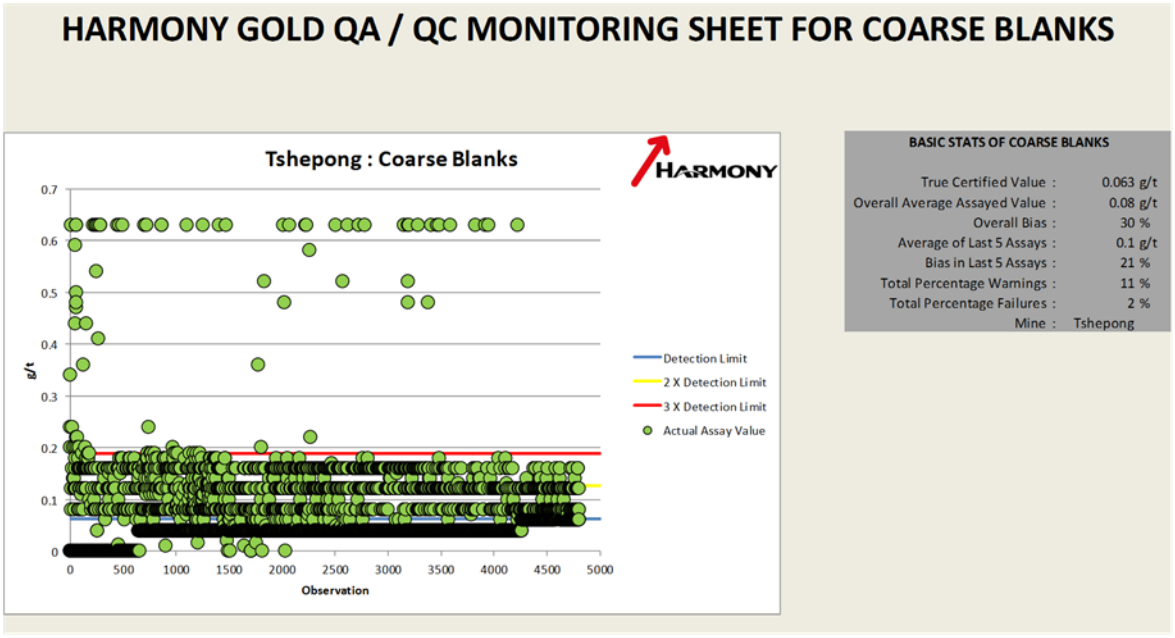


Figure B-13: Scatterplot of Duplicate Samples (expected 80% of duplicate samples to fall within 20% absolute relative difference)