

https://dspace.org/

Finding the optimal dynamic anisotropy resolution for grade estimation improvement at Driefontein Gold Mine, South Africa

Mandava, Senzeni Maggie

2016

http://hdl.handle.net/10539/20591

Mandava, Senzeni Maggie (2016) Finding the optimal dynamic anisotropy resolution for grade estimation improvement at Driefontein Gold Mine, South Africa, University of Witwatersrand, Johannesburg, <http://wiredspace.wits.ac.za/handle/10539/20591>

Downloaded from a DSpace repository. https://dspace.org/

FINDING THE OPTIMAL DYNAMIC ANISOTROPY RESOLUTION FOR GRADE ESTIMATION IMPROVEMENT AT DRIEFONTEIN GOLD MINE, SOUTH AFRICA

Senzeni Maggie Mandava

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 Mining Engineering.

February, 2016

DECLARATION

I declare that this research report is my own, unaided work. It is being submitted in partial fulfilment of requirements for the Degree of Master of Science in Mining Engineering at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other University. The information on which this dissertation is based was obtained from Sibanye after request and approval to use data from their Driefontein Gold mine for research purposes.

_____

(Signature of candidate)

_____ day of _____ _____

ABSTRACT

Mineral Resource estimation provides an assessment of the quantity, quality, shape and grade distribution of a mineralised deposit. The resource estimation process involves; the assessment of data available, creation of geological and/or grade models for the deposit, statistical and geostatistical analyses of the data, as well as determination of the appropriate grade interpolation methods. In the grade estimation process, grades are interpolated/extrapolated into a two or three – dimensional resource block model of a deposit. The process uses a search volume ellipsoid, centred on each block, to select samples used for estimation. Traditionally, a global orientated search ellipsoid is used during the estimation process. An improvement in the estimation process can be achieved if the direction and continuity of mineralisation is acknowledged by aligning the search ellipsoid accordingly. The misalignment of the search ellipsoid by just a few degrees can impact the estimation results. Representing grade continuity in undulating and folded structures can be a challenge to correct grade estimation. One solution to this problem is to apply the method of Dynamic Anisotropy in the estimation process. This method allows for the anisotropy rotation angles defining the search ellipsoid and variogram model, to directly follow the trend of the mineralisation for each cell within a block model. This research report will describe the application of Dynamic Anisotropy to a slightly undulating area which lies on a gently folded limb of a syncline at Driefontein gold mine and where Ordinary Kriging is used as the method of estimation. In addition, the optimal Dynamic Anisotropy resolution that will provide an improvement in grade estimates will be determined. This will be achieved by executing the estimation process on various block model grid sizes. The geostatistical literature research carried out for this research report highlights the importance of Dynamic Anisotropy in resource estimation. Through the application and analysis on a real-life dataset, this research report will put theories and opinions about Dynamic Anisotropy to the test.

ACKNOWLEDGEMENT

I would like to express my gratitude to and acknowledge the following people and institutions that assisted me in compiling this research:

- Leon Tolmay of Sibanye – for awarding me this great opportunity by supplying the data as well as supervising, guiding and supporting me to execute this research;
- Dr. Christina Dohm of Anglo American – for her prompt response to my last minute request to review my work. Special thanks to her for her suggestions and guidance;
- Mark Wanless, Principal Geologist and Partner at SRK Consulting – for his endless patience, guidance and support in the execution and review of this research. I am truly grateful;
- Prof R.C.A. Minnitt – for his patience in supervising the research and for assistance in administrative matters at the University of the Witwatersrand;
- Victor Simposya, Principal Resource Geologist and Partner at SRK Consulting – for giving me some valuable insight into the fundamentals of Geostatistics;
- SRK Consulting (Pty) Ltd – for awarding me the time and financial support to turn my dream into a reality and to giving this research to fruition;
- To my friends and colleagues here at SRK, Raymond and Livhu – thank you for your support, proof reading of the numerous drafts and your words of encouragement. You made the journey easier;
- And finally to my Husband, Osmond, my pillar, my strength, without whose love, support and encouragement this paper could not have been realized.

Thank you so much!

To my Children

Seth, Kayla and Garikayi

TABLE OF CONTENTS

DECLARATION **2**
ABSTRACT ..... **i**
ACKNOWLEDGEMENT ..... **ii**
LIST OF FIGURES ..... **vi**
LIST OF TABLES ..... **viii**
LIST OF SYMBOLS ..... **ix**
NOMENCLATURE **10**
1 Introduction **12**
1.1 Background 13
1.2 Study Rationale 14
1.3 Investigation Motivation 16
1.4 Study Objectives 17
1.5 General Geology 18
1.6 Witwatersrand Geological Setting and Miniralisation 19
2 Case Study **24**
2.1 Project Location 25
2.2 Topography and Climate 25
2.3 Local Geology 26
2.4 Sampling and Assaying 27
2.5 Geological Model and Interpretation 28
2.6 Resource Estimation 28
3 Methodology and Theoretical Descriptions of Estimation Processes **30**
3.1 Exploratory Data Analysis - General Process 30
3.1.1 Data Validation 30
3.1.2 Statistical Analysis 32
3.2 Exploratory Data Analysis (EDA) 38
3.2.1 Data Transformations 41
3.2.2 Data Declustering 41
3.2.3 Gaussian Anamorphosis 42
3.3 Geostatistical Theory 45
3.3.1 Stationarity 46
3.3.2 The Covariance Function or Variogram 47
3.3.3 Calculating Semi-Variograms 49
3.4 Quantitative Kriging Neighbourhood Analysis 55
3.4.1 Sample Number Optimisation 55

3.4.2 Block Discretization Number Optimisation 57
3.5 Geostatistical Estimation 58
3.5.1 Kriging 59
3.5.2 Dynamic Anisotropy Application on 30×30 m Block 61
3.6 Domaining 62
3.6.1 Statistical Analysis on Domains 64
3.6.2 Variogran Analysis on Domains 66
3.7 Optimal DA Resolution Determination 69
4 Results on 30×30 m Block Estimates 70
4.1 Validating the Estimates 70
4.1.1 Global Statistics 70
4.1.2 Visual Inspection of Plots 71
4.1.3 Swath Validation 73
4.1.4 Scatter Plots 75
5 Discussion 79
5.1 Domains for 30×30 m Blocks 79
5.2 Optimal DA Resolution 83
6 Conclusions 93
7 References 96
Appendices 108
Appendix A: Variogram Maps 109
Appendix B: Fitted Back Transformed Semi-variograms 113
Appendix C: Domain Grade Distribution Validation Plots 117
Appendix D: Sample Number Optimisation 120
Appendix E: Block Discretization Number Optimisation 122
Appendix F: Grade Distribution Validation Plots for Different Block Sizes 124
Appendix G: Scatter Plots 128

LIST OF FIGURES

Figure 1.1: South Africa's Gold producing operations and amount of gold produced in 2013 (Source: Alex Létourneau of Kitco News, *http://www.forbes.com/sites/kitconews/2014/06/20/worlds-largest-gold-producing-countries-south-africa/*, 2014). 12

Figure 1.2: Oriented search ellipsoid (Source: Copyright © 1983 – 2013 CAE Mining Corporate Limited). 15

Figure 1.3: Witwatersrand Goldfield and Location of Major Gold mines (Source: *http://www.sec.gov/Archives/edgar/data/865492/000086549203000016/ascii_gbg20f2.htm*, 2003). 20

Figure 1.4: Witwatersrand Basin depositional model (Source: Kloof-Driefontein Complex (KDC) – Technical Short Form Report, 2010). 21

Figure 2.1: Sibanye location map (Source: *https://www.sibanyegold.co.za/operations*, 2015). 24

Figure 2.2: Project Location (Kloof report "modified", unpublished, 2012). 25

Figure 2.3: Modified schematic 3D section through the Driefontein mine ore body looking north. (Source: Kloof-Driefontein Complex (KDC) – Technical Short Form Report, 2010). 26

Figure 3.1: Excerpt of chip sample excel file used. 31

Figure 3.2: Chip sample location plan colour coded on Au cm.g/t. 32

Figure 3.3: Normal distribution curve (Source: *http://civilblog.org/2014/05/11/what-is-standard-deviation-and-how-to-calculate-it-with-an-example-calculation/* , 2015). 34

Figure 3.4: Lognormal distribution curve (Source *http://www.engineeredsoftware.com/nasa/lognormal.htm*, 2015). 36

Figure 3.5: Example of negatively and positively skewed distributions (Source: *http://mvpprograms.com/help/mvpstats/distributions/SkewnessKurtosis*, 2015). 37

Figure 3.6: General forms of Kurtosis (Source: *http://mvpprograms.com/help/mvpstats/distributions/SkewnessKurtosis*, 2015). 37

Figure 3.7: Cm.g/t histogram. 40

Figure 3.8: Q-Q plot – cm.g/t. 40

Figure 3.9: Anamorphosis plot – cm.g/t. 43

Figure 3.10: Gaussian plot – cm.g/t. 44

Figure 3.11: Q-Q plot – cm.g/t. 44

Figure 3.12: Relationship between the covariance function and semi-variogram (Source: *http://inspirehep.net/record/1189977/plots*, Gentile *et al.*, 2013). 48

Figure 3.13: Omni – directional experimental semi-variogram model for cm.g/t. 50

Figure 3.14: Variogram map plot for Gaussian cm.g/t. 52

Figure 3.15: Fitted semi-variogram models for cm.g/t gaussian transformed data for directions N50 (left) and N140 (right). 52

Figure 3.16: Cm.g/t back transformed fitted semi-variogram model. 53
Figure 3.17: QKNA sample number optimisation graph 57
Figure 3.18: Graph showing sensitivity *of* $\boldsymbol{C}(\boldsymbol{V},\boldsymbol{V})$ to block discretization. 58
Figure 3.19: Orientation of mineralization plot. 61
Figure 3.20: Domaining – Zone 1, 2 and 3 63
Figure 3.21: Domaining – Zone 4 and 5 63
Figure 3.22: Cm.g/t histograms for domains 65
Figure 3.23: Fitted semi-variogram models for cm.g/t gaussian transformed data for directions N50 (left) and N140 (right) for Zone 1 66
Figure 3.24: Fitted semi-variogram models for cm.g/t gaussian transformed data for directions N60 (left) and N150 (right) for Zone 2 67
Figure 3.25: Fitted semi-variogram models for cm.g/t gaussian transformed data for directions N150 (left) and N60 (right) for Zone 3 67
Figure 3.26: Fitted semi-variogram models for cm.g/t gaussian transformed data for directions N50 (left) and N140 (right) for Zone 4 68
Figure 3.27: Fitted semi-variogram models for cm.g/t gaussian transformed data for directions N50 (left) and N140 (right) for Zone 5 68
Figure 4.1: OK estimates (above) vs sample data and DA estimates (below) vs. sample data plots. 72
Figure 4.2: Swath orientation and spacing. 73
Figure 4.3: OK estimates and DA estimates vs. sample data swath plots. 74
Figure 4.4: OK estimates (above) and DA estimates (below) vs. sample data scatter plots. 76
Figure 5.1: Scatter plots of OK estimates (above) vs sample data and DA estimates (below) vs. sample data – 2 zones 81
Figure 5.2: Scatter plots of OK estimates (above) vs sample data and DA estimates (below) vs. sample data– 3 zones 82
Figure 5.3: Slope of regression against block size. 87
Figure 5.4: Correlation coefficient against block size. 87
Figure 5.5: Diagrammatic representation showing relationship between block sizes, “resolution of Dynamic Anisotropy” and localised directions. 89
Figure 5.6: Correlation coefficient against block size (domained). 90
Figure 5.7: Slope of regression against block size (domained). 91
Figure 5.8: Estimate variances and smoothing (Source: Krige, 1996). 92

LIST OF TABLES

Table 3-1: Statistics from the sample file 38
Table 3-2: Semi-variogram model parameters for cm.g/t 54
Table 3-3: Normalised semi-variogram model parameters for cm.g/t 54
Table 3-4: Search parameters for cm.g/t 62
Table 3-5: Zonal Statistics 64
Table 3-6: Domain semi-variogram model parameters for cm.g/t 69
Table 3-7: Domain normalised semi-variogram model parameters for cm.g/t 69
Table 4-1: Global and local statistics 70
Table 5-1: Domain global and local statistics for 30 m by 30 m blocks 79
Table 5-2: Line of regression equations and correlation coefficients on domains 83
Table 5-3: Search ellipsoid orientation, search range parameters and minimum and maximum number of samples used during estimation process for each zone and block size 84
Table 5-4: Line of regression equations and correlation coefficients on blocks and domains 86

LIST OF SYMBOLS

Billion Years	Ga
Covariance	$C(0)$
Degrees	^{o}
Degrees centigrade	^{o}C
Distance vector	h
Grams per Tonne	g/t
Greater than	>
Function of x	Y(x)
Lamda	λ
Less than	<
Less than or equal to	$\leq$
Location parameter	θ
Mean	μ
Nugget Effect	C0
Number of samples	n_i
Ounce	oz
Percent	%
Target sample	$_i$
Semi-variogram	γ
Standard deviation	σ
Variogram	2γ
Variogram Sill	C1
Variance	σ^2
Weight	w_i

NOMENCLATURE

Bottom band	BB
Carbon seam	CS
Centimetre grams per tonne	cm.g/t
Channel Width	CW
Carbon Leader Reef	CLR
Coefficient of variation	CoV
Cubic metre	m^3
Dynamic Anisotropy	DA
Exploratory data analysis	EDA
Gram per tonne	g/t
Gold	Au
Hectares	ha
Inverse power distance	IPD
Inter quartile range	IQR
Kilogramme	kg
Kilometre	km
Kriging efficiency	KE
Metre	m
Metric tonnes	t
Middlevlei Reef	MVR
Million troy ounces	Moz
Million pounds	Mlb
Minimum	Min
Maximum	Max
Samples	SAMP

Selective Mining Unit	SMU
Slope of regression	SLOR
Square metre	m^2
Standard Deviation	Stddev
Standard normal (gaussian) variable	N(0,1)
Three Dimension	3D
Troy ounce	oz
Two Dimension	2D
Quantile-quantile	Q-Q
Variance	Var
Ventersdorp Contact Reef	VCR
Work place names	WPNAME

1 Introduction

South Africa is the world's fifth largest gold producer, (World Gold Council, 2015). A total of five million and twenty four thousand ounces of gold were produced in 2013 from the various operating mines throughout the country, (Létourneau, 2014) see Figure 1.1 below. The rate of deterioration in terms of gold mining output from South Africa is staggering. In 1983 the country produced 21.847 million ounces of gold, roughly 64 % of the world's output. Today it accounts for a little over 6 %.

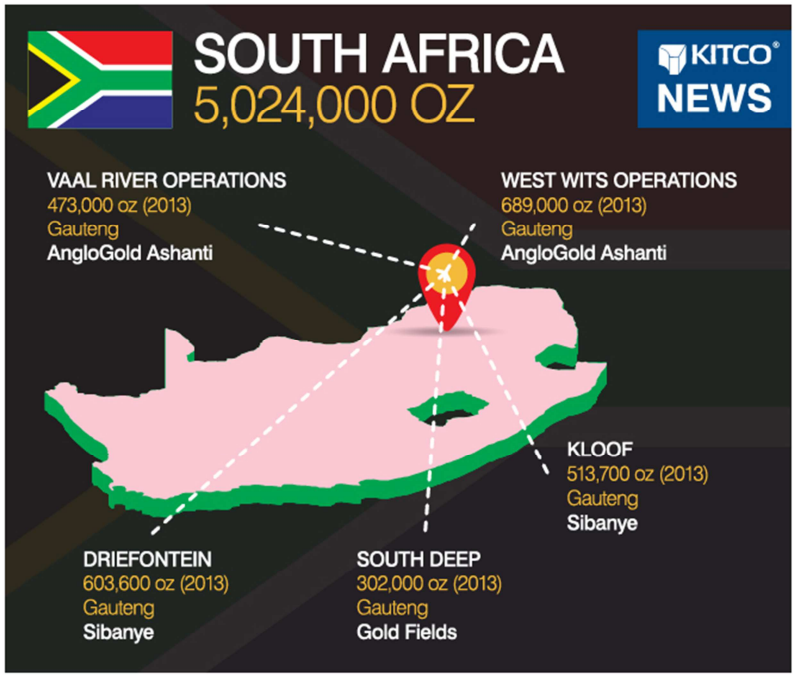


Figure 1.1: South Africa's Gold producing operations and amount of gold produced in 2013 (Source: Alex Létourneau of Kitco News, *http://www.forbes.com/sites/kitconews/2014/06/20/worlds-largest-gold-producing-countries-south-africa/*, 2014).

Over the last decade, the country has seen a steady decline year-on-year as mines are getting older, no new discoveries are being made, grades are dropping and mines are going deeper and deeper into the ground. The drop in its global ranking is not due to resource depletion (South Africa ranks close behind the world leader Australia in reserves) but rather to its high costs of production. In fact South Africa has the highest mining cash costs among all major producing regions (The Globe and Mail, 2011).

South African gold mining industry extensively utilizes geostatistical techniques to provide estimates for gold reserves and resources, and to assist with daily mine valuation and mine planning (Krige and Dowd, 2000). Geostatistical estimation is used in the mining industry to estimate mineral grades in unmined areas. The technique of linear kriging (G. Matheron, 1991) is mostly utilized, providing best linear unbiased estimates (in terms of minimum estimation variance – Armstrong, 1998) of metal grades. The theoretical development of geostatistics can be traced back to initial studies carried out by Herbert Sichel and Danie Krige in the late 1940's and early 1950's (Sichel, 1947 and Krige, 1951), as well as the pioneering work of Georges Matheron (Matheron 1963). Since then, geostatistical theory has expanded into an extensive field of research and knowledge.

1.1 Background

The geostatistical theory has its origins in the South African gold mining industry and it is documented to have evolved out of a requirement for improved prediction techniques for gold grades to justify the expense of deep-level mining (Houlding, 2000). Although the theory of regionalised variables was formalised by Matheron (1970), much of the groundwork had been performed earlier on the South Africa gold mines, principally by Krige (1951) (Braun *et al.*, 1987). Geostatistics is a widely accepted and applied method for estimating Mineral Resources. In the book, "Using Geostatistics to Estimate the Resources of a Narrow Vein Gold Deposit", Roy (2000) mentions that in general, it is believed that because of the often erratic nature of vein mineralisation, geostatistics lends itself more readily to larger, more uniform deposits.

Gold vein deposits are normally characterised by a skewed grade distribution in which a small number of higher grade samples can cause an overestimation in the surrounding low grade estimation block areas during the grade estimation process when using non-

geostatistical grade estimating methods such as arithmetic or graphical methods. The same is also true for the gold sedimentary deposits and most other metal deposits where the metals are present in trace quantities. These outliers can also cause variogram distortions such as pure nugget effect which in turn affects the grade estimate results.

When undertaking resource estimation on any deposit with narrow vein or layer-like geometry, the variable of interest (i.e. grade) is not a suitable variable for direct kriging. This is because grade of mineralised intercepts is clearly defined on varying supports (Bertoli *et al.*, 2003). The ore intercepts are very variable in width and the width is such that subcell filling is impractical or inaccurate. Due to the nature of the orebody which is thin and of variable thickness, it is most suitable to do the grade estimation in two-dimensions ("2D"). The 2D estimation approaches have significant advantages over three-dimensions, ("3D") approaches when dealing with a narrow vein or layer-like geometry (Bertoli *et al.*, 2003). Generally, the entire ore width will be mined. The main motivation for 2D estimation is that bias due to non-additivity of grade is avoided. In the 2D approach, grade can also be defined as the ratio of two other variables (thickness and gold accumulation – the product of grade by thickness potentially weighted by bulk density, cm.g/t values, i.e. gold per horizontal area), which are amenable to direct kriging (Chiles and Delfiner, 1999). Hence accumulation is kriged independently from width and the grade is derived by the ratio of the two resultant estimates.

1.2 Study Rationale

Mineral Resource estimation provides an assessment of the quantity, quality, shape and grade distribution of a mineralised deposit. The reliability of Mineral Resource estimates depends not only on the quality and number of samples, but also on the degree of continuity of the mineralisation (Blais and Carlier, 1968). The more continuous the mineralisation, the better will be the correlation between the grade of a sample and the unknown true grade of

the block represented by this sample. The grade estimation process can be summarised as follows:

- An assessment of the quantity and quality of the data available including database management and verification;
- The creation of 2D and/or 3D geological and grade models for the deposit as well as an assessment of stationarity, and possible sub-domaining of the deposit;
- Statistical and geostatistical analyses of the data and the determination of the most appropriate grade and density interpolation methods; and
- Classification and reporting according to accepted and internationally recognised guidelines.

In the grade estimation process, grades are interpolated/extrapolated into a 2D or 3D geological block model of a deposit. This process uses a search volume ellipsoid (Figure 1.2), centred on each block, to select samples used for estimation using one of the several interpolation algorithms, for example: Nearest Neighbour, Inverse Power of Distance, ("IPD"), Ordinary Kriging, ("OK"), Simple Kriging, Log Kriging, Sichel's t estimator, Ordinary Macro Kriging or Simple Macro Kriging.

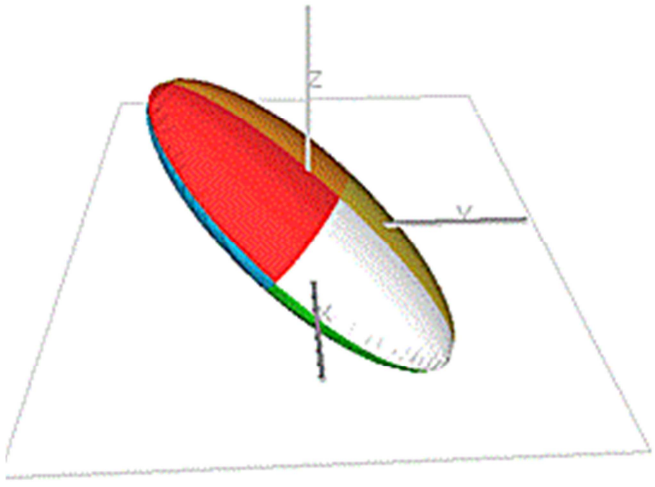


Figure 1.2: Oriented search ellipsoid (Source: Copyright © 1983 – 2013 CAE Mining Corporate Limited).

Traditionally, a global orientated search ellipsoid is used during the estimation process. The continuity of mineralisation usually varies with direction and thus the search ellipsoid should be oriented to follow the direction of mineralisation. The misalignment of the search ellipsoid by just a few degrees can impact the estimation results e.g. a block might be estimated to be below the cut-off grade when it is in fact above the cut-off grade and vice-versa.

The size, shape and orientation of the search ellipsoid are determined from variography analyses identifying the major directions of continuity of mineralisation in the deposit (Zabrusky, 2013). Variability of a variable in the sample data is dependent on both magnitude and mineralisation trend (LeMay, 1998).

Variography analyses and geostatistical estimation include directional control that allows:

- The establishment of the principal directions of anisotropy;
- The appropriate semi-variogram models to be derived; and
- The selection of adjacent samples in the estimation process to be controlled (Houlding, 2000).

The description of anisotropy and local customisation both depend heavily on a good understanding of the genesis of the data set and qualitative geology input. Spatial variability of a regionalised variable maybe isotropic: if the axes (X, Y and Z) of the ellipsoid are of equal length, or anisotropic: when the axes are of different lengths with different variabilities in orthogonal directions. The use of an anisotropic search window produces a small but consistent improvement in all estimating techniques (Bohling, 2005).

1.3 Investigation Motivation

The author has been approached by Driefontein Gold Mine to study the application of Dynamic Anisotropy, (“DA”) and then derive an optimised resolution for an area at the mine

that lies on a slightly undulating gently folded limb of a syncline. The Geostatistics and Evaluation Consultant for the mine, has supported with input and provided the dataset to be used for the investigation. The mine currently uses the OK method of estimation on a 30×30 m 2D block model to calculate their Mineral Resource estimates. Appropriately representing grade continuity in undulating/meandering and folded structures can be a challenge during the grade estimation process and one solution to this problem is to consider the application of DA during the estimation process. Therefore, the basis of this investigation is to confirm whether the application of DA in such a situation will indeed provide an improvement to the grade estimation process. In addition, the optimal DA resolution that will provide an improvement in grade estimates will also be determined.

1.4 Study Objectives

Glacken and Gray (2011), Morrison and Brown (2012), Zabrusky (2013) and in the JORC Resource report for Tulu Kapi Gold Deposit in Ethiopia produced by KEFI Minerals Plc (2014) applied DA estimation technique to calculate Mineral Resource estimates. Their studies have shown that applying DA in their estimation techniques, optimised estimation orientations thus producing better grade estimates.

This research report will describe the application of DA to an area at Driefontein Gold Mine that lies on a slightly undulating gently folded limb of a syncline and where OK is used as the method of estimation. The estimation results calculated from applying DA interpolation technique will be compared against those that are calculated using OK interpolation technique.

In addition, the optimal DA resolution that will provide an improvement in grade estimates will be determined. This will be achieved by executing the estimation process on various block model grid sizes. The estimation block size which is currently being used at

Driefontein mine is a 30×30 m grid. This study will be executed on estimation block model range of different resolutions; 30×30 m, 60×60 m, 120×120 m and 240×240 m grids.

It is good practice to continually scrutinize and test the theories of geostatistics, and to attempt to improve upon or derive new methods of dealing with spatial data. This is also the intention of this research. It is apparent how important DA is in geostatistics, and through the study of available geostatistical literature, the application and the analysis of real-life data; this research report intends putting theories and opinions about DA to the test. Such findings could be of great value to the field of gold mining geostatistics (and probably also to a wider field of geostatistics and academia).

1.5 General Geology

The Kaapvaal Craton is an ancient segment of continental crust which formed in southern Africa between about 3.7 to 2.7 Ga. Much of this continental nucleus actually formed prior to 3.1 Ga by the formation of an extensive granitoid basement and amalgamation with arc-like oceanic terranes represented by mafic/ultramafic volcanics and associated sediments (De Wit *et al*., 1992). Subsequent growth of the craton was accompanied by further continental magmatic activity, possible Cordilleran-style accretion of composite terranes along the margins of the proto-continent, and the deposition of sedimentary basins. Although the question of when the Kaapvaal continent attained its cratonic rigidity is a semantic one, it is clear that by 2.7 Ga, or shortly thereafter, this segment of crust was not subjected to further major orogenesis (Robb and Meyer, 1995).

Lithologies from five major Proterozoic basins are preserved on the Kaapvaal craton spanning a time period from 3 Ga to 1.75 Ga From oldest to youngest these are the Pongola, Witwatersrand, Ventersdorp, Transvaal and Waterberg sequences. All five appear to have followed a broadly similar evolutionary history but differences in source area, palaeo-

ecology and levels of tectonic activity during basin evolution have apparently resulted in the formation and preservation of auriferous placer deposits only in the Pongola, Witwatersrand and Transvaal basins. The Pongola and Witwatersrand deposits are the only true placer deposits, with the Transvaal deposits resulting from the remobilization of low grade protore by tectonic and thermal processes. Gold production from the Witwatersrand sequence exceeds that from the other Proterozoic sequences by almost two orders of magnitude, being 98.5 % of the total, up to 1971 (Pretorius, 1976a).

1.6 Witwatersrand Geological Setting and Miniralisation

The Witwatersrand basin extends some 350 km in length in a north-easterly direction and approximately 200 km in width northwest/southeast on the Kaapvaal Craton as shown in Figure 1.3. It is an approximately 7000 m thick terrigenous sequence. The Witwatersrand Supergroup is contained within the Witwatersrand Basin which was deposited in a fluvio-deltaic environment in the centre of the Kaapvaal Craton. It is believed to have developed as an intracontinental feature which was filled with early Proterozoic sediments and volcanics. The major goldfields are developed in downwarps between basement domes and are believed to represent fluvial fans formed at the mouth of rivers having a north-western provenance (Pretorius, 1975). These separate fan systems were developed at different entry points and are preserved as the eight distinct goldfields, each with local geological variations. These sediments were subsequently reworked by longshore currents which existed within the basin. Figure 1.3 shows the location of the Witwatersrand goldfield and of the major gold mines.

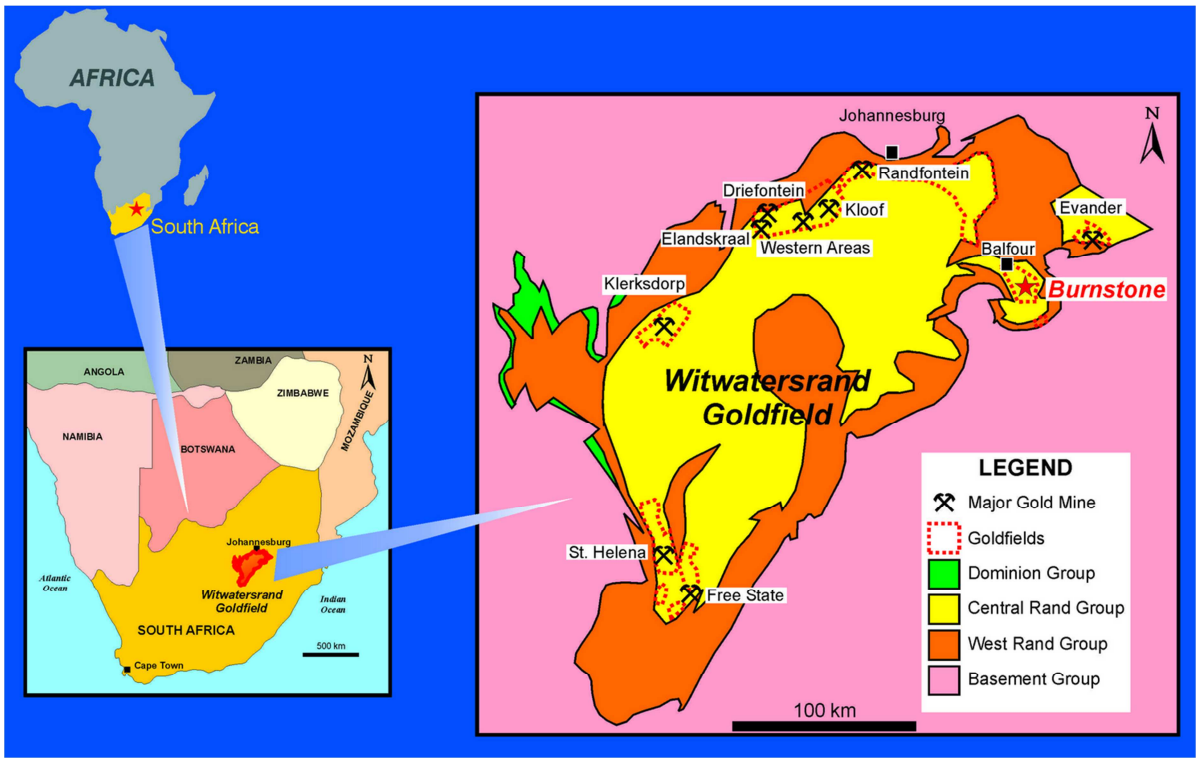


Figure 1.3: Witwatersrand Goldfield and Location of Major Gold mines (Source: *http://www.sec.gov/Archives/edgar/data/865492/000086549203000016/ascii_qbq20f2.htm*, 2003).

The Witwatersrand Supergroup was deposited somewhere in the interval 3.1 to 2.7 Ga (Armstrong *et al.*, 1986). It is divided into a lower the argillaceous and arenaceous West Rand Group and an upper Central Rand Group, the latter containing the bulk of the gold mineralisation near the base in laterally-extensive, quartz pebble conglomerate reefs. These sediments have been metamorphosed to greenschist facies (Phillips, 1987). The Central Rand Group contains a number of palaeo-placer deposits associated with angular unconformities which have parallel strikes and truncate each other towards the basin edge. The area that is mined is underlain by outliers of Karoo Supergroup shales and sandstones, followed by Pretoria Group sediments and the Chuniespoort Group dolomites. The dolomites overlie the Klipriviersberg Group volcanic rocks, which in turn cap the mineralised reef and sediments of the Central Rand Group. Figure 1.4 shows the Witwatersrand Basin depositional model.

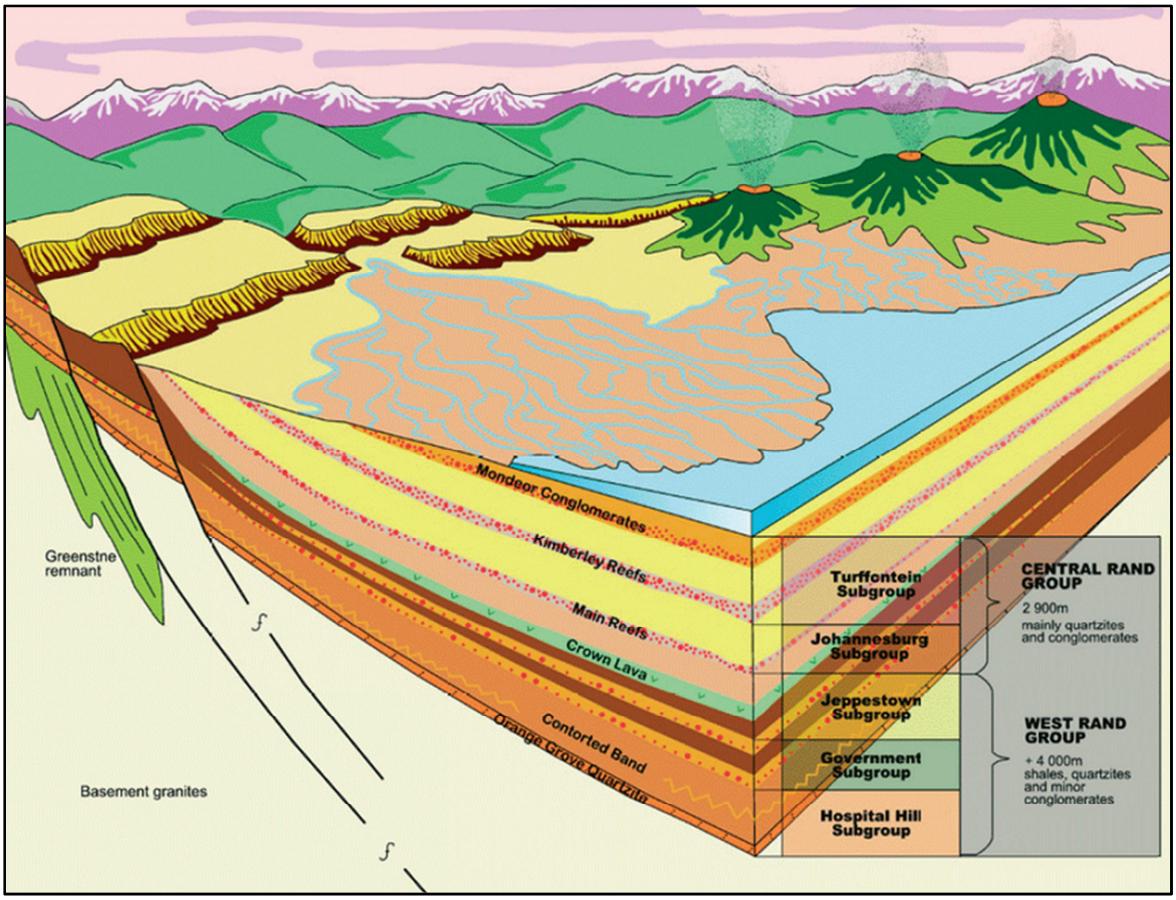


Figure 1.4: Witwatersrand Basin depositional model (Source: Kloof-Driefontein Complex (KDC) – Technical Short Form Report, 2010).

Gold mineralisation in the Witwatersrand Basin is hosted within quartz pebble conglomerate bodies, termed “reefs”, an auriferous well developed conglomerate bed, horizons developed at the base of the unconformities. The extent of unconformity is typically greatest near the basin margin and decreases toward more distal areas within the basin. Complex patterns of syn-depositional faulting have caused variations in sediment thickness within the basin. Sub-vertical to over-folded reef structures are a characteristic feature of the basin margin within certain areas. These reefs occur within the separate goldfields shown in Figure 1.3. Typically within each goldfield, there are one or sometimes two major reef units present, which may be accompanied by one or more secondary reef units. As a result of faulting and other primary controls on mineralisation, the goldfields are not laterally continuous with each other and may also be characterised by the presence or dominance of different reef units. The reefs are generally less than 2 m in thickness and are widely considered to

represent laterally extensive braided fluvial deposits or unconfined flow deposits, which formed along the flanks of alluvial fan systems that developed around the edge of what was effectively an inland sea.

The gold is mainly of detrital origin, deposited syngenetically with the conglomerates and interrelated with sedimentary features such as unconformities and fluvial channels. Although the gold generally occurs in native form and is usually associated with pyrite and carbon, most of the gold has been subsequently modified and remobilised during secondary hydrothermalism, giving rise to the most favourite depositional model currently in use, i.e. Modified Palaeoplacer Model (Gold Fields Limited, 2010). The model emphasises a control on the occurrence of ore minerals by placer-forming mechanisms, while accepting some modification by metamorphism. It is the generally accepted model for the origin of gold and uranium mineralisation of the Witwatersrand Basin. For several decades now, models using sedimentological principles were very successfully used to predict gold distribution on mine properties.

It has also been confirmed through studies done by Coward *et al.,* (1995), Beach and Smith (2007), Dankert and Hein (2010), Jolley *et al.*, (2007) and others that the basin experienced more than one episode of deformation although Dankert and Hein (2010) pointed out that little regional-scale kinematic and structural analysis has taken place across the Witwatersrand Basin despite more than 100 years of mining of gold, (Manzi *et al.*, 2013).

The origins of the gold in the Witwatersrand Basin is still a topic of ongoing debate with a strong distinction between those who believe the gold was deposited with the conglomerates at the time of sedimentation (placerists), and those hydrothermalists supporting the epigenetic origin of gold, believing that the gold was hydrothermally introduced after the sediments had been lithified (Manzi *et al.*, 2013). Despite these varied viewpoints, the most

fundamental control to the gold distribution remains the association with quartz-pebble conglomerates on intra-basinal unconformities. The reefs are laterally continuous, as a consequence of the regional nature of the erosional surfaces. Bedrock (footwall) controls have also been established, these features are interpreted to control the distribution of many of the reefs. Preferential reef development within channel systems and sedimentary features such as facies variations and channel frequency, also assist in mapping out local gold distributions. The identification and modelling of the sedimentary features within the reef units and the linkage between is a key activity for Mineral Resource estimation.

2 Case Study

Sibanye is an independent, South African mining group, which currently owns and operates four underground and surface gold operations, namely the Driefontein Operation ("Driefontein"), the Kloof Operation ("Kloof") and the Cooke Operation ("Cooke") in the West Witwatersrand region and the Beatrix Operation ("Beatrix") in the southern Free State. In addition to its mining activities, the Group owns and manages significant extraction and processing facilities at its operations, which beneficiate the gold bearing ore mined. Sibanye is the largest individual producer of gold from South Africa and is one of the world's 10 largest gold producers. In 2014, the Group produced 49,432kg (2013: 44,474kg) or 1.59Moz (2013: 1.43Moz) of gold at an All-in cost of R375,854/kg (2013: R354,376/kg) or US$1,080/oz (2013: US$1,148/oz) and invested R3.3 billion (2013: R2.9 billion) in capital at its operations (© 2015 Sibanye Gold annual reports, 2015). Figure 2.1 is a location map showing Sibanye operations.

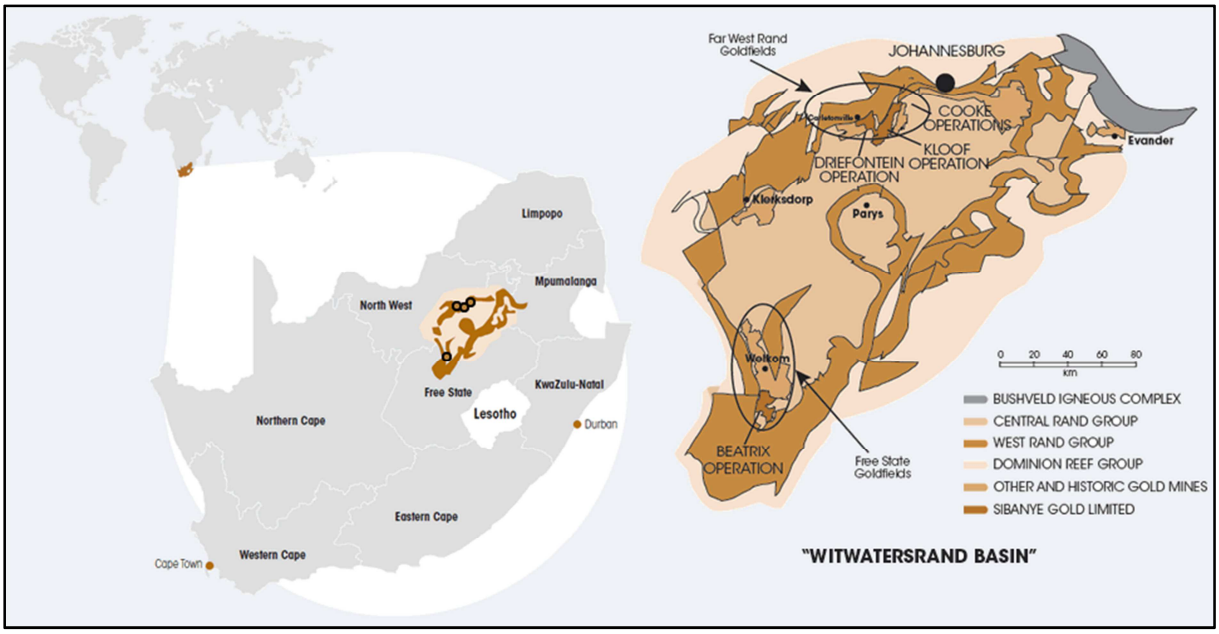


Figure 2.1: Sibanye location map (Source: *https://www.sibanyegold.co.za/operations*, 2015).

2.1 Project Location

Driefontein mine is situated between 70 and 80 km west of Johannesburg at latitude 26°24'S and longitude 27°30'E, near the town of Carletonville in the Gauteng Province of South Africa. The site is accessed via the N12 highway between Johannesburg and Potchefstroom. Geologically it is located in the West Wits Goldfield of the Witwatersrand Basin. Figure 2.2 is a map showing location details of Driefontein mine.

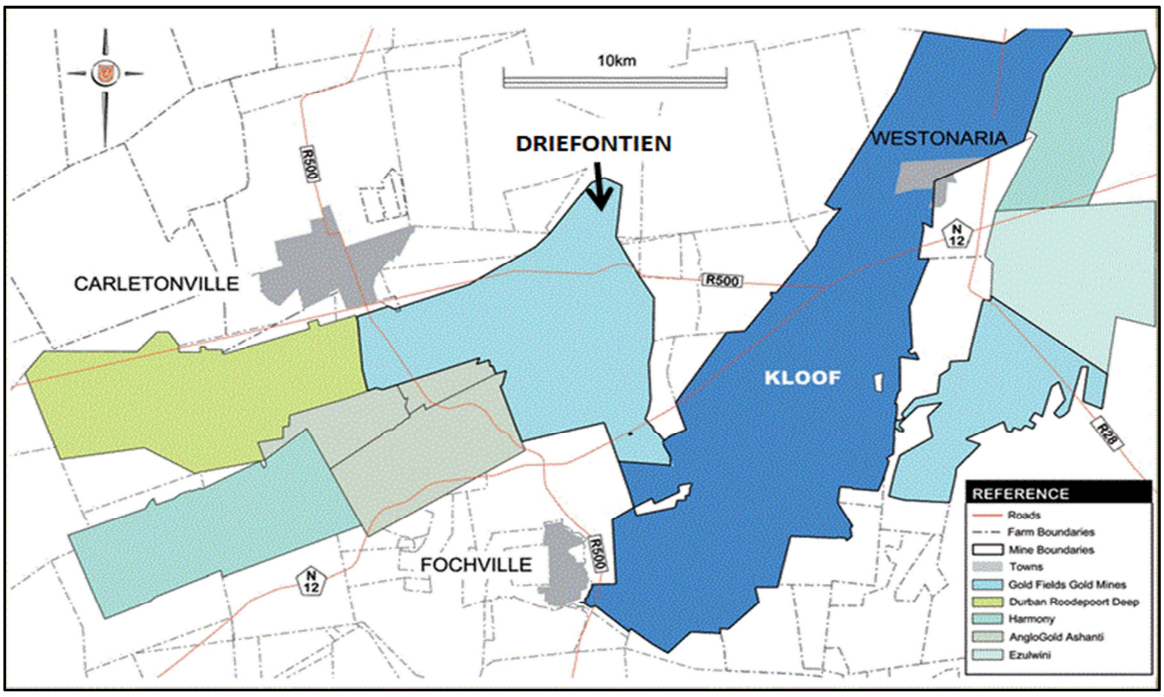


Figure 2.2: Project Location (Kloof report "modified", unpublished, 2012).

2.2 Topography and Climate

Gauteng province, borders four of the following provinces: Free State, North West, Mpumalanga and Limpopo. Topography is relatively flat and the vegetation of the area is classified as Bankenveld consisting of grassland, which is in good condition. Livestock farming is widespread in the surrounding area and no extremes of climate are experienced that affect mining operations.

2.3 Local Geology

Driefontein mine is located in the West Wits Line that forms part of the Far West Rand of the Witwatersrand Basin. This goldfield is divided into an Eastern Section and a Western Section separated by the Bank Anticline and associated faulting. The Driefontein mining right is situated in the Western Section. Driefontein Syncline dominates the structure of the Western Section. The eastern limb of the syncline dips in a south-south-westerly direction and the western limb in an east-south-easterly direction. Gold mineralisation at Driefontein mine is contained within three reef horizons: the Carbon Leader Reef ("CLR"), the Ventersdorp Contact Reef ("VCR") and the Middelvlei Reef ("MVR"), which occur at depths of between 500 m and 4 000 m. Figure 2.3 shows a schematic 3D section through the Driefontein mine ore body looking north.

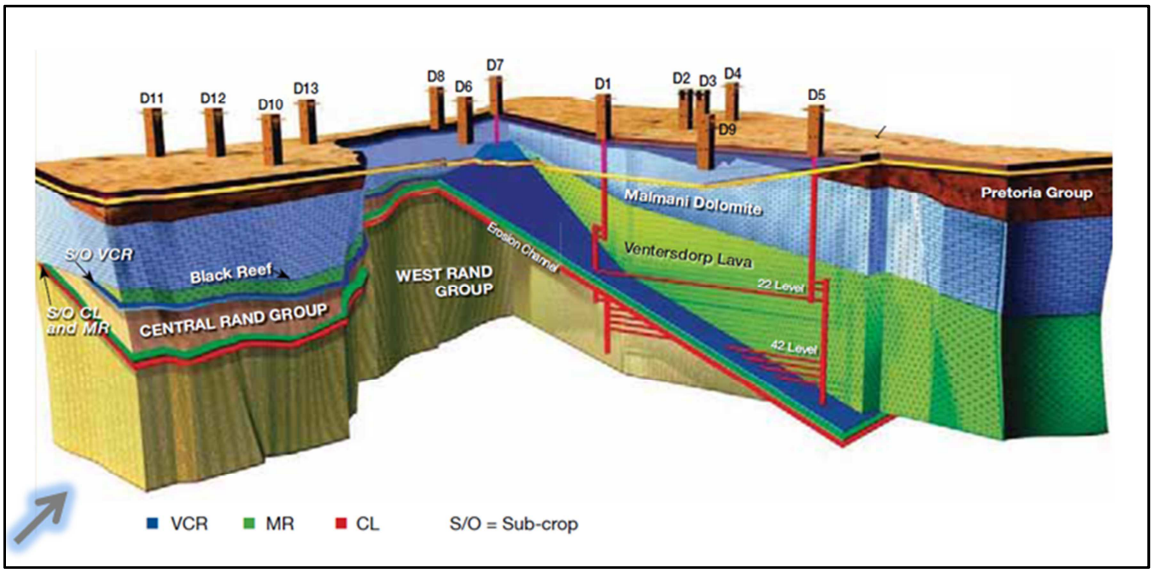


Figure 2.3: Modified schematic 3D section through the Driefontein mine ore body looking north. (Source: Kloof-Driefontein Complex (KDC) – Technical Short Form Report, 2010).

Stratigraphically, the CLR is situated 40 m to 70 m below the VCR and MVR, and is a generally high-grade reef comprising different facies. It dips to the south at approximately 25 degrees. The CLR sub-crops against the VCR in the eastern part of the mine. The west-dipping Bank Fault defines the eastern limit of both reefs. The VCR is most extensively

developed in the east, and subcrops to the west. The MVR is situated approximately 50m above the CLR. The average gold grades vary with lithofacies changes in all of the reefs.

2.4 Sampling and Assaying

Within underground workings, exposures of the reef are channel sampled. Individual channels are chipped from the stope and development working faces using a hammer and chisel and the sample chips are caught using steel pans. Driefontein mine's stope and development channel sampling interval standards are 5 m and 3 m respectively. Channels are defined perpendicular to the reef plane and each section's position is fixed by offsetting from survey pegs. The reef is segregated into lithological units and is correlated between sample sections, and individual samples of 10 – 25 cm in length are taken to reflect the internal geometry of the reef with not less than a 10 cm sample being taken on contact. The sample weight taken is of the order of 300 to 500 g. Adjacent samples spanning the hanging wall or footwall contact may be taken to increase the sample volume in the contact area, where the grades are generally the highest and most variable. The sample sections are digitised directly into the database, where the spatial validity is checked. Quality assurance and quality control ("QA/QC") procedures are done to ensure sampling protocol is maintained through planned task observations. Final submission of each sample into the Gold Fields' Integrated Resource and Reserve System ("IRRIS") database is only completed following a series of checks and approvals by the Senior Evaluators.

Driefontein mine relies on in-house assay laboratories. All underground and surface samples are dried, crushed, and pulverised and analysed using fire assay techniques. As part of QA/QC procedures at the assay laboratory, blank samples, repeat assays and inter-laboratory programmes are used to ensure that the laboratories adhere to assay standards and protocols. The laboratory is SANAS (South African National Accreditation Society) ISO17025 accredited for "Determination of Gold via Fire Assay with gravimetric finish" Method 001.

The gold assay values from the laboratories are kept in digital format in IRRIS. The long mining history and the quantity and quality of the data upon which the Mineral Resource estimates are based are sufficient to support the Mineral Resource estimates as derived. The IRRIS system is fully operational at Driefontein mine and provides an auditable trail.

2.5 Geological Model and Interpretation

Geological models are based on structural, grade and sedimentological data. The structural data are used to generate 3D models whilst the sedimentological, gold value and channel width ("CW") data are used to delineate the broadly geologically homogeneous zones called geozones. The geozones are used to constrain the statistical and geostatistical analyses that form the basis of the Mineral Resource estimation process. The geological models are validated in IRRIS, where resource blocking is also carried out. All blocks are captured according to the geological models and geozones. The resource blocks are essentially zones of geological homogeneity and stationarity (an important requirement for kriging), which take cognisance of geological structure, reef types or facies, value trends, data quantity and distribution thereof, kriging results (including kriging efficiencies and slope of regression which will be discussed in more detail in Section 3.4), infrastructure, mining method, ventilation and rock engineering.

2.6 Resource Estimation

The Mineral Resource estimation process used on Driefontein mine is based on surface and underground boreholes as well as underground channel samples and is executed in 2D. The mine is divided into various domains based on the structural and geological models which are used as the basis for the homogenous geostatistical zones or geozones. Detailed exploratory data analyses, including sample verification, histogram and cumulative frequency plots for distributional analysis, additive constant estimates, outlier checks, trend analyses, and declustering are carried out on individual geozones. The main interpolation

methodologies used at Driefontein mine are Ordinary and Simple Kriging, with Macro Kriging also being used to estimate grade in certain geozones. Sichel ‘t’ estimates and declustered averaging techniques are used for geozones that are predominantly based on relatively few surface and underground boreholes, in conjunction with peripheral mined-out stope data. Detailed variography studies are carried out on point and regularised data. Relative and traditional variograms are used for kriging purposes (on a number of regularised block sizes). Log-point variograms are used only for lognormal change of supports for grade tonnage grade curve calculations.

3 Methodology and Theoretical Descriptions of Estimation Processes

In order to obtain reliable estimates, there is a need to combine relevant and good quality data, competent geological interpretation and an appropriate grade interpolation technique. An interpretation of the available geological information and data will lead to a dependable representation of the mineralized deposit, its geometry and the analysis of the grade data will identify the grade distribution and the estimation methodology. This industry practice in resource evaluation was followed in the development of the research. An assessment of the quantity and quality of the data including statistical and geostatistical analyses of the data and the determination of the most appropriate grade estimation methods was carried out. Good quality data aspect was discussed in section 2.4 and the geological model in section 2.5. This section will concentrate on the grade analyses, variography and the estimation methodology.

3.1 Exploratory Data Analysis - General Process

Exploratory data analysis, ("EDA") is a well-established statistical tradition that provides conceptual and computational tools for discovering patterns to foster hypothesis development and refinement. It is a critical first step in analysing data in order to detect any mistakes, checking of assumptions, preliminary selection of appropriate models, determining relationships among the explanatory variables and assessing the direction and rough size of relationships between explanatory and outcome variables.

3.1.1 Data Validation

The author was supplied with a Microsoft Excel spreadsheet containing chip sample data from the CLR of Driefontein mine. An excerpt of the file is shown in Figure 3.1.

X	Y	ZONE	BHID	WPNAME	Z	SMPTYPE	SID	CMGT	CW	SW	FW	RW	VL	DATE	GORS	LEVEL	PANEL	FAULT	SAMPDIP	DATA	SYSTEM	A0	B0	FROM	TO	LENGTH	DUMMY	CHANMAX	C0
9248.8	-20751.3	12	400SE	CL40 23 0	0	1	0	664.0	106	154	0	106	0	20070532	0	0	-	-	0			330	65	0	1	1	0	210	0
9251.1	-20755.6	12	400SE	CL40 23 0	0	1	0	1745.1	93	175	0	93	0	20070532	0	0	-	-	0			330	65	0	1	1	0	210	0
9253.4	-20759.9	12	400SE	CL40 23 0	0	1	0	3515.1	110	170	0	110	0	20070532	0	0	-	-	0			330	65	0	1	1	0	210	0
9256.0	-20764.0	12	400SE	CL40 23 0	0	1	0	651.4	96	169	0	96	0	20070532	0	0	-	-	0			330	65	0	1	1	0	210	0
9258.7	-20768.0	12	400SE	CL40 23 0	0	1	0	1376.9	110	163	0	110	0	20070532	0	0	-	-	0			330	65	0	1	1	0	210	0
9260.5	-20712.2	12	400SE	CL40 23 R	0	2	0	331.3	89	0	0	89	0	20050624	0	0	-	-	0			330	60	0	1	1	0	210	0
9261.2	-20772.2	12	400SE	CL40 23 0	0	1	0	1164.4	120	150	0	120	0	20070532	0	0	-	-	0			330	65	0	1	1	0	210	0
9263.4	-20776.6	12	400SE	CL40 23 0	0	1	0	2364.7	124	175	0	124	0	20070532	0	0	-	-	0			330	65	0	1	1	0	210	0

Figure 3.1: Excerpt of chip sample excel file used.

A total of 46 951 chip samples constitute the dataset. The following data validation exercises were undertaken;

- Data checks for missing information – checks were made to ensure each of the chip samples contained all the required data types – collar, collar survey and assay information;
- Checks for duplicated chip samples and negative grades in the database;
- Check for outliers in terms of spatial location;
- Checks to determine if the samples were collected from the same domain, “zone”.

The plotting of sample locations helps to facilitate easier visual analysis in order to check for samples that may plot as outliers in terms of coordinates. They also assist in looking for and delineating trends as well as locating areas with high values and low values. Figure 3.2 shows the sample location plan colour coded in Au cm.g/t. In the southern portion of the location map, 8 samples did not have Work place names, (“WPNAME”) and were therefore removed from the database. Hence a total of 46 943 samples were used in this study.

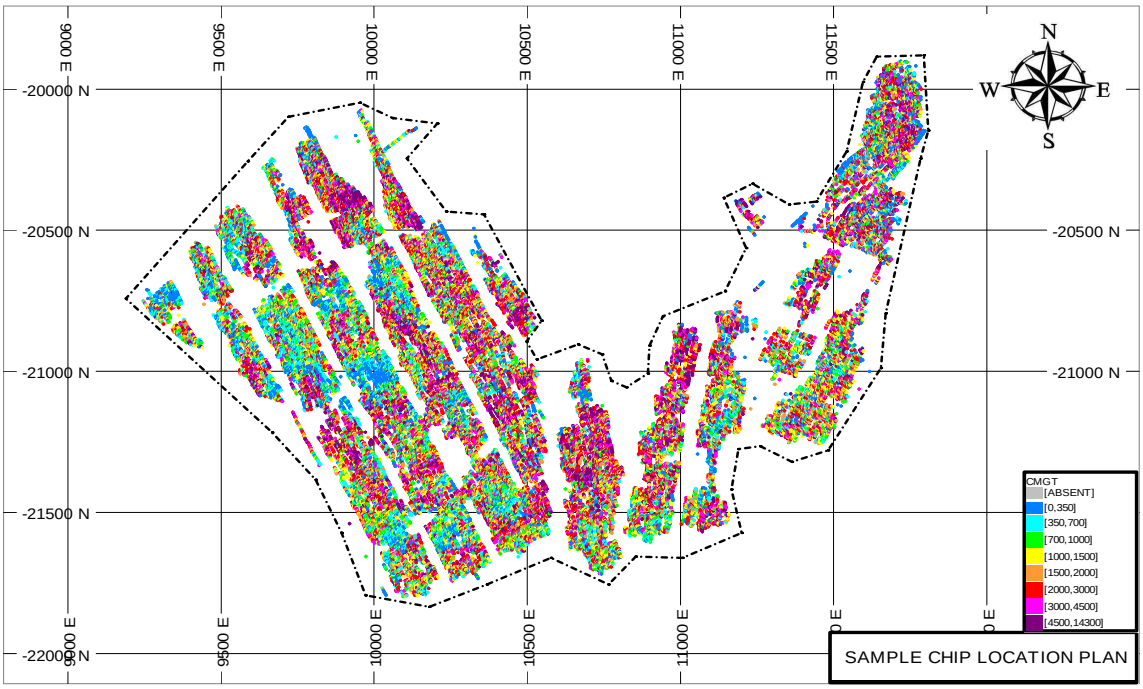


Figure 3.2: Chip sample location plan colour coded on Au cm.g/t.

3.1.2 Statistical Analysis

Any given set of samples will give rise to a distribution of values and statistics are used to describe the properties of this distribution. A wide variety of theoretical distributions have been described in statistical literature and two of these appear to be most applicable to the observed distribution of geological data. These are the normal and lognormal distributions and they have been widely applied in Mineral Resource estimation. A number of statistics, such as the mean, variance, standard deviation, normal distribution, lognormal distribution, skewness and kurtosis may be used to describe the distribution of sample values. Only some of these are of significance in Mineral Resource estimation and they will be defined before showing how they are applied.

Mean

The mean ("μ") is the most important measure of central tendency and is the average value for the distribution. It is calculated by the formula:

$$\mu = \sum_{i=1}^{n} \frac{x_i}{n}$$

where x_i comprises "n" independent measurements. The mean on its own is obviously of little value in describing a distribution. What is needed is some measure of how the values are dispersed about this mean. The most useful measures are the variance and standard deviation which are closely related to each other.

Variance and Standard Deviation

The variance ("σ^2") quantifies the average variability of a set of values about the mean of that set of values and is calculated by:

$$\sigma^2 = \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{n-1}$$

The term $(x_i - \mu)^2$ is used for two reasons. The term $(x_i - \mu)$ would sum to zero since positive and negative deviations from the mean would cancel out and in subsequent statistical use $(x_i - \mu)^2$ is more useful than the modulus of $(x_i - \mu)$ which is the unsigned deviation from the mean. The standard deviation ("σ") is merely the square root of the variance and one or the other statistic is used as required by the circumstances. The standard deviation is of great value in statistical analysis because:

1. It reflects the dispersion of values so that the variability of different sets of data may be compared in terms of the standard deviation;
2. It permits the precise interpretation of values within a distribution;
3. It, like the mean, is a member of a mathematical system which permits its use in more advanced statistical analysis (Ainslie, 1981).

Coefficient of Variation

The Coefficient of Variation ("CoV") is a statistic that is often used as an alternative to skewness to describe the shape of the distribution. It is defined as the ratio of the standard deviation to the mean. It measures variability in the data. In practice, high CoV values

(CoV>1.0) indicates the presence of some erratic high sample values that may have a significant impact on the final estimates (Isaaks and Srivastava, 1989).

The Normal Distribution

This is the most common theoretical probability distribution used in statistics. It is completely defined by its mean and standard deviation. The general formula for the probability density function of the normal distribution is:

$$f(x) = \frac{1}{\sigma\sqrt{(2\pi)}}\, e^{\{(x-\mu)^2/2\sigma^2\}}$$

Where "μ" is the location parameter and "σ" is the scale parameter. The case where μ = 0 and σ = 1 is called the standard normal distribution. This distribution has the form of a symmetrical bell-shaped curve, see Figure 3.3. An important feature of this distribution is that it is possible to calculate the probability with which the value of a sample from a normally distributed population will lie within a specified range. The equation for the standard normal distribution is:

$$f(x) = \frac{1}{\sqrt{(2\pi)}}\, e^{(\frac{-x^2}{2})}$$

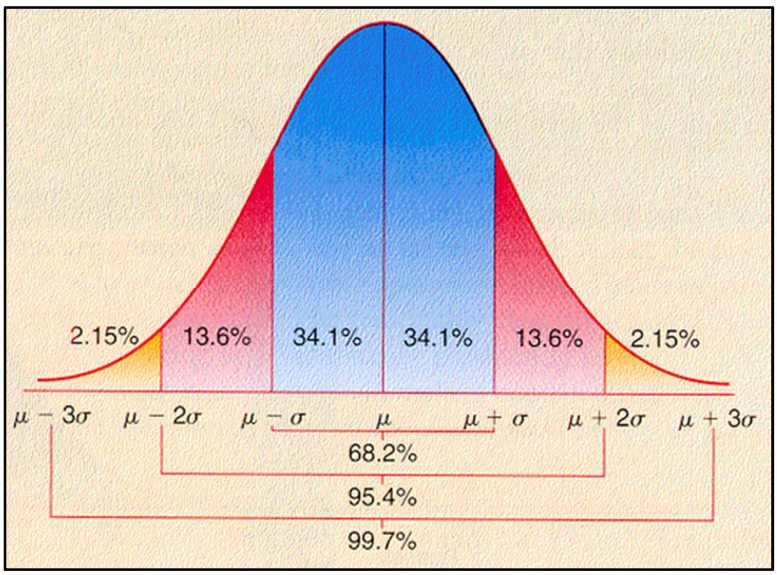


Figure 3.3: Normal distribution curve (Source: *http://civilblog.org/2014/05/11/what-is-standard-deviation-and-how-to-calculate-it-with-an-example-calculation/* , 2015).

Lognormal Distribution

A lognormal distribution has a positively skewed curve with a long drawn-out tail towards the higher values. A variable, *x* is lognormally distributed if $f(x) = ln(x)$ with "*ln*" denoting the natural logarithm. The transformation of data with this kind of distribution into natural logarithms results in a normal distribution. The general formula for the probability density function of the lognormal distribution is:

$$f(x) = \frac{1}{(x-\theta)\sigma\sqrt{(2\pi)}}\, e^{-\{\frac{(\frac{ln(x-\theta)}{m})^2}{2\sigma^2}\}} \qquad x \geq \theta\,;\ \mathrm{m}, \sigma > 0$$

Where "σ' is the shape parameter and is the standard deviation of the log of the distribution, "θ" is the location parameter and "m" is the scale parameter (and is also the median of the distribution), see Figure 3.4. The mean and variance of a lognormally distributed data cannot be calculated by using the equations given earlier because these statistics are unduly influenced by the high values in the drawn-out tail of the distribution. Hence the equation for the standard lognormal distribution is:

$$f(x) = \frac{1}{x\sigma\sqrt{(2\pi)}}\, e^{\{\frac{(\ln(x))^2}{2\sigma^2}\}} \qquad x \geq 0\,;\ \sigma > 0$$

If $x = 0$, then $f(x) = 0$. The case where θ = 0 and m = 1 is called the standard lognormal distribution. The case where θ equals zero is called the 2-parameter lognormal distribution. Krige (1960) showed that data which did not fit this distribution could be made to do so by addition of a constant (β) to each value of the data set. This is termed the three parameter lognormal distribution and data which has been transformed in this manner can be manipulated in the same way as lognormal data.

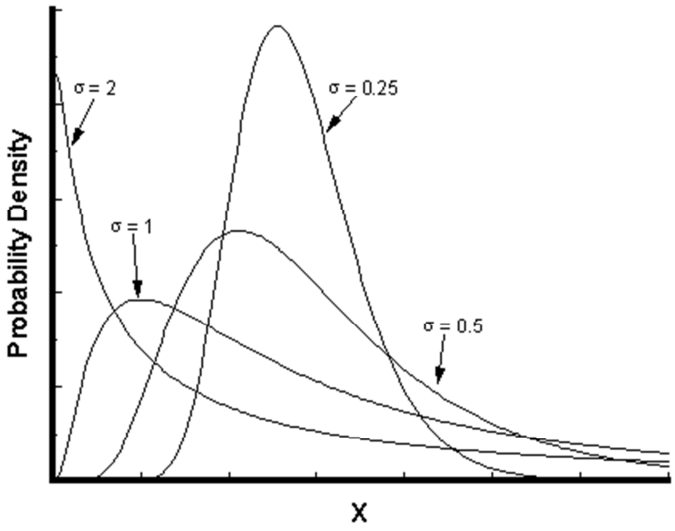


Figure 3.4: Lognormal distribution curve (Source *http://www.engineeredsoftware.com/nasa/lognormal.htm*, 2015).

Skewness and Kurtosis

A fundamental task in many statistical analyses is to characterize the location and variability of a data set. A further characterization of the data includes skewness and kurtosis. Skewness is the degree of departure from symmetry of a distribution. A distribution, or data set, is symmetric if it looks the same to the left and right of the centre point. A positively skewed distribution has a "tail" which is pulled in the positive direction. A negatively skewed distribution has a "tail" which is pulled in the negative direction. The equation for skewness is:

$$\text{Skewness} = \frac{\Sigma(x_i - \mu)^3}{n\sigma^3}$$

(Wheeler, 2004)

A diagram showing examples of a negatively and positively skewed distribution is shown in Figure 3.5.

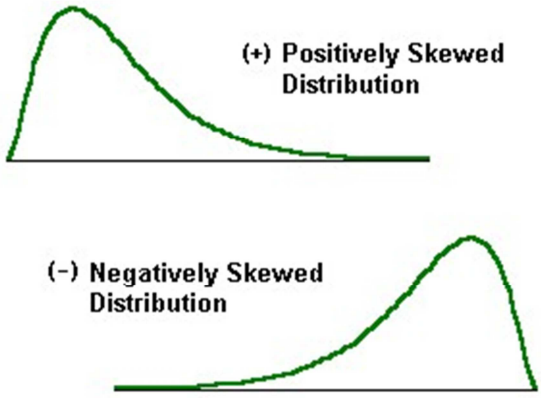


Figure 3.5: Example of negatively and positively skewed distributions (Source: *http://mvpprograms.com/help/mvpstats/distributions/SkewnessKurtosis*, 2015).

Kurtosis is a measure of whether the data are peaked or flat relative to a normal distribution. That is, data sets with high kurtosis tend to have a distinct peak near the mean, decline rather rapidly, and have heavy tails. Data sets with low kurtosis tend to have a flat top near the mean rather than a sharp peak. A uniform distribution would be the extreme case. The equation for skewness is:

$$\text{Kurtosis} = \frac{\Sigma(x_i - \mu)^4}{n\sigma^4}$$

(Wheeler, 2004)

Diagram showing general forms of Kurtosis is shown in Figure 3.6.

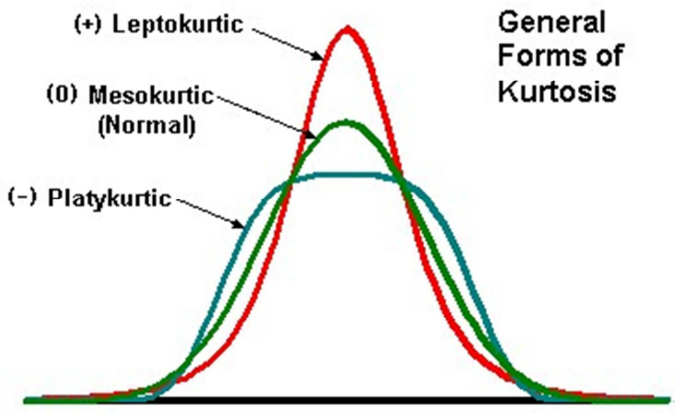


Figure 3.6: General forms of Kurtosis (Source: *http://mvpprograms.com/help/mvpstats/distributions/SkewnessKurtosis*, 2015).

The interpretation for skewness and kurtosis are as follows:

- When skewness > 0 – This is a right skewed distribution with most values concentrated on left of the mean and extreme values to the right;
- When skewness < 0 – This is left skewed distribution with most values concentrated on the right of the mean and extreme values to the left;
- When skewness = 0 – This implies that the that the mean is equal to the median and the distribution is symmetrical around the mean;
- When kurtosis > 3 - Leptokurtic distribution, sharper than a normal distribution, with values concentrated around the mean and thicker tails. This indicates that there is a high probability for extreme values;
- When kurtosis < 3 - Platykurtic distribution is flatter than a normal distribution and has a wider peak. The probability for extreme values is less than that for a normal distribution, and the values are wider spread around the mean;
- When kurtosis = 3 - Mesokurtic distribution is similar to a normal distribution.

3.2 Exploratory Data Analysis (EDA)

Preliminary analysis was performed on the entire dataset by calculation of the main statistical moments (mean, minimum value, maximum value, standard deviation, coefficient of variation, skewness and kurtosis). The results are presented in Table 3-1.

Table 3-1: Statistics from the sample file

Field	NSamp	Min	Max	Mean	Var	Stddev	Skewness	Kurtosis	CoV
Grade/cm.g/t	46 943	0.00	14 250.90	2 218.85	5 066 287.71	2 250.84	2.00	4.77	1.01
CW/cm	46 943	1.00	0.62	26.07	347.40	18.64	3.62	23.81	0.71

The following observations are made from Table 3-1:

- The mean value for cm.g/t and CW is 2218.85 cm.g/t and 26.07 cm respectively;
- The values of skewness and kurtosis for cm.g/t are 2 and 4.77 respectively. This means that this is a Leptokurtic distribution which is skewed to the right. However, Dr. Wheeler (2004) mentioned in his book, "Advanced Topics in Statistical Process

Control: The Power of Shewhart's Charts" that statistics for skewness and kurtosis simply do not provide any useful information beyond that already given by the measures of location and dispersion hence not much emphasis should be put on skewness and kurtosis values.

The dataset has CoV values of 1.01 and 0.71 for cm.g/t and CW respectively. These values are within acceptable ranges. This indicates that there are not a significant number of extreme values. The samples are closely spaced which translates to a massive quantity of values informing the distribution very well.

Histogram Plot

The purpose of a histogram is to graphically summarize the distribution of a univariate data set. The histogram graphically shows the following:

- Centre (i.e., the location) of the data;
- Spread (i.e., the scale) of the data;
- Skewness of the data;
- Presence of outliers; and
- Presence of multiple modes in the data.

A histogram of the population was constructed, see Figure 3.7. The data reflects a positively skewed distribution which is expected for a gold grade distribution. It also supports the results of the statistical analysis tabulated in Table 3-1.

Q-Q Plot

The quantile-quantile ("Q-Q") plot is a graphical technique meant to compare the distribution of a variable with a secondary distribution (theoretical distribution or distribution of an auxiliary variable). A 45-degree reference line is also plotted. If the points depart from the straight line, then the assumed distribution is called into question. The greater the departure from this reference line, the greater the evidence for the conclusion that

data set have come from populations with different distributions. A Q-Q plot of the dataset was constructed, see Figure 3.8.

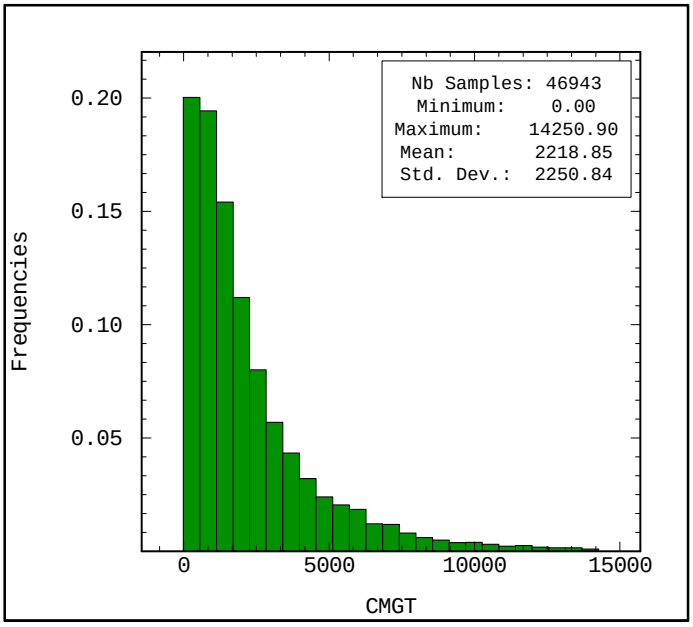


Figure 3.7: Cm.g/t histogram.

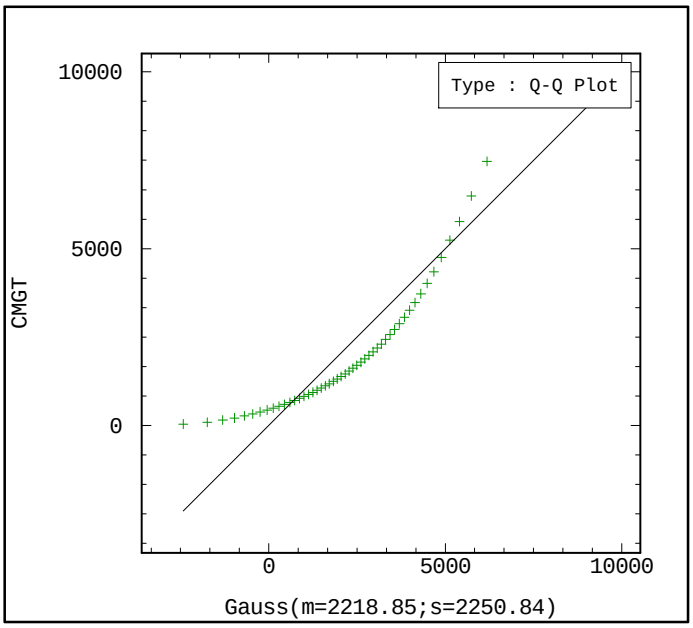


Figure 3.8: Q-Q plot – cm.g/t.

Observation from Figure 3.8 shows that points in a Q-Q plot depart from a straight line and that the distribution is skewed to the right.

Witwatersrand gold has been shown to fit a three parameter lognormal distribution. Professor Krige in 1960 extended the two parameter lognormal model to include a third parameter or additive constant (β), thus adding enormous flexibility to the model such that it would accommodate almost any distribution of gold values as well as many other minerals (Magri, 1983). Since there is an excess of low values, by the addition of a constant to all the samples the population will then be lognormal (St George, 1992). The next step was to determine the best-fit distribution by calculating the log values of the cm.g/t or consider normalising the transformation. For the purpose of this investigation, the data was transformed using the Gaussian Anamorphosis (normal score transformation).

3.2.1 Data Transformations

Transformations can be used to make the data more normally distributed and satisfy assumptions of continuous variability. Data transformations are performed before using some geostatistical methods such as uniform condition and for maps that require the normal assumption. Quantile and probability maps from Simple, Ordinary and Universal Kriging methods require the normality assumption. Frequency distributions of grades in the mining industry are often skewed and contain outliers. Depending on the amount of data available and the number of outliers, the estimation of the semi-variograms parameters may prove difficult or near impossible. In such cases, Krige and Magri have shown that if the data set is lognormally distributed then the log-transformation of the raw data can substantially improve the semi-variogram model.

3.2.2 Data Declustering

Often times the geological sampling programs are not entirely spatially random and data may have a high density of sample points near one another. This may lead to a spatial preferentiality among the data points. If data is preferentially sampled when it is spatially auto correlated, the resulting histogram from the sample may not reflect the histogram of the

population. Hence the removal of potential bias or "declustering" the data is vital. The declustering method assigns higher weights to the less densely populated sample points and lower weights to the high density areas. The process adjusts the summary statistics to be representative of the entire area of interest thereby removing potential bias, for example one area may have clustered sampled points compared to other portions (Deutsch and Journel, 1998).

Using the "Declustering" application in the Isatis software package, the sample dataset used in this study was declustered. To compute the weight, w_i attached to a target sample, i, the number of samples inside a moving window centered on this target sample were counted. The weight w_i is equal to mv/n_i where mv is the mean of all the n_i inside a window. The weight will be 1 when the number of points inside the moving window equals the mean of the n_i. This normalization allows the weights not to decrease with the number of data. Following this, a low weight will be attached to a sample which is inside a cluster while an isolated point will get the maximum weight.

3.2.3 Gaussian Anamorphosis

The Gaussian Anamorphosis (normal score transformation) was used to transform the raw variable, cm.g/t into a gaussian variable. Gaussian distributions are commonly used due to their convenient statistical properties. The new Gaussian variable was created using the empirical inversion method. The empirical inversion calculates for each raw value the attached empirical frequency and calculates the corresponding gaussian value. For mathematical reasons this function can be conveniently written as a polynomial expansion:

$$\varphi(Y) = \sum_{i=0}^{\infty} \Psi_i H_i\,(Y)$$

Where $H_i(Y)$ are called the Hermite Polynomials. In practice, this polynomial expansion is stopped to a given order. Using this technique, any function can be expressed in terms of

Hermite polynomials (30 polynomials were used in this scenario), (Baafi and Schofield, 1997).

Figure 3.9 shows the anamorphosis model which is represented with dotted lines when outside the validity bounds called the practical interval of definition of the raw variable. These bounds are the points where the model is no longer smoothly increasing. The absolute interval of definition is represented as two horizontal lines. The solid blue line represents the cumulative solution of the hermite polynomials, which informs the Gaussian Anamorphosis, and the block line represents the actual data.

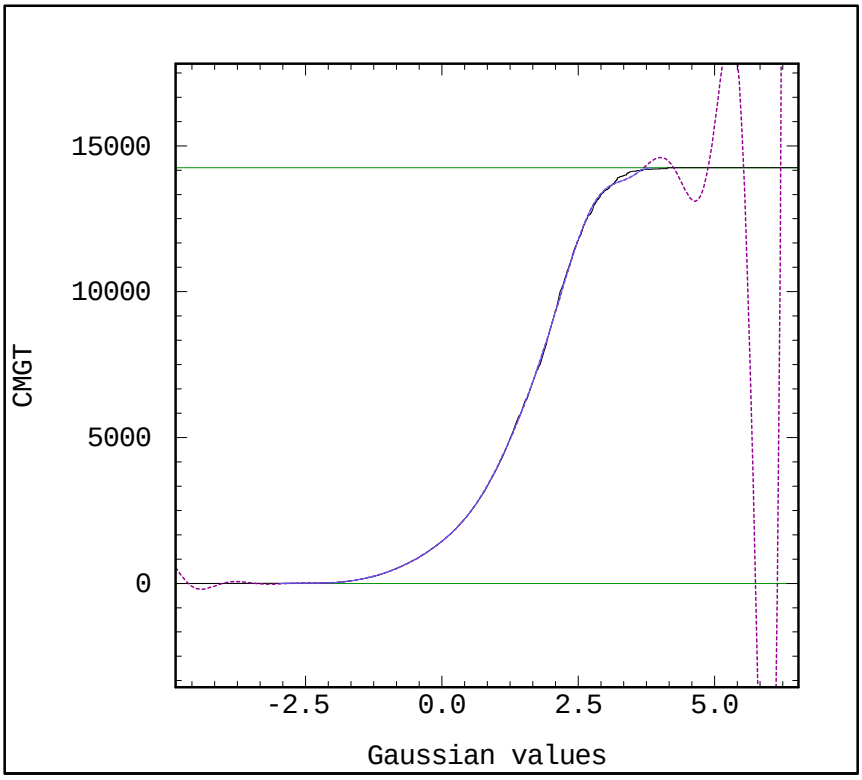


Figure 3.9: Anamorphosis plot – cm.g/t.

Histograms and Q-Q plots for the gaussian transformed dataset, Figure 3.10 and Figure 3.11 respectively were constructed for the cm.g/t values.

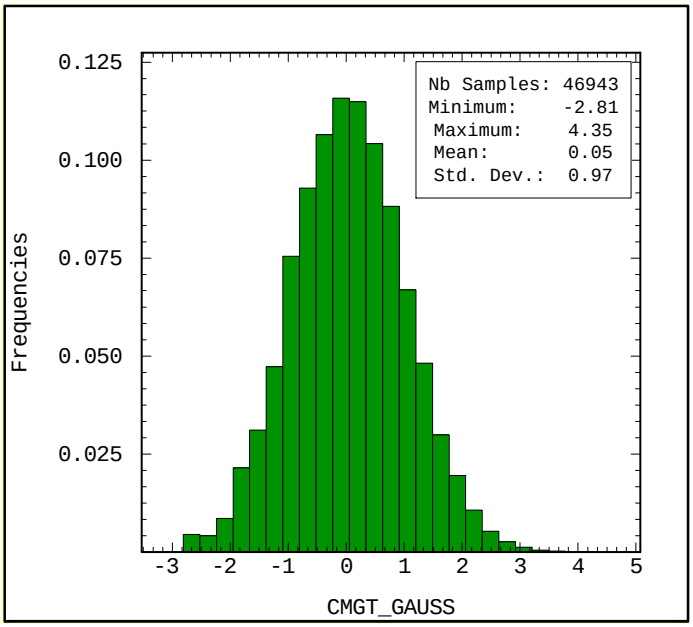


Figure 3.10: Gaussian plot – cm.g/t.

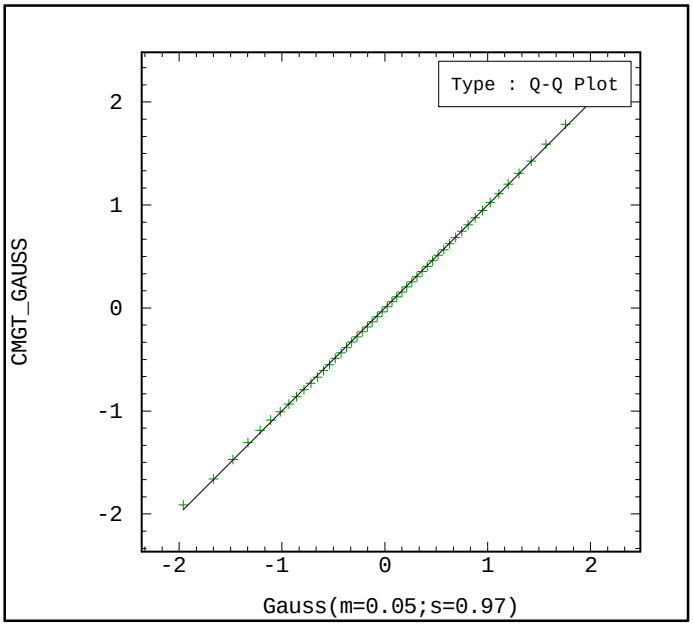


Figure 3.11: Q-Q plot – cm.g/t.

It can be shown from Figure 3.10 (histogram model shows a normal distribution) and Figure 3.11 (model follows an almost linear trend) that the Gaussian transformed model is a good approximation for the distribution.

3.3 Geostatistical Theory

Geostatistics was defined by Matheron (1962), one of the fathers of geostatistics, as "the application of the formalism of random functions to the reconnaissance and estimation of natural phenomena." Sample grade, for example, is a regionalised variable because it is distributed throughout a space. The variability of the grade throughout that space is described by a function called the ‘semi-variogram’. Through a process known as ‘kriging’, samples in space and the semi-variogram are used to estimate the mean grade of a point, area, or volume. Kriging provides the best estimate of the mean value of a regionalised variable. It provides the Best Linear Unbiased Estimator (“BLUE”) of the grade without consideration of the spatial statistics of all the estimates collectively. The focus is on local accuracy as opposed to spatial continuity (Deutsch and Journel, 1998).

During kriging, each sample is assigned a sample weight based on the distance allocation of the block being estimated. A sample further away from the centre of a block being estimated is assigned a smaller weight than a sample closer to the centre of a block. The weighted samples are then linearly combined to give the best estimate. It is the ‘BEST’ estimate because the procedure minimizes the expected error between the estimated grade and the true grade. Sample weights are calculated such that the variance of the estimate is minimised. That variance can be calculated using the sample positions and the semi-variogram function. Having the estimation variance is extremely useful because it allows the user to explore the risk or error attached to the estimate.

The basic tool of geostatistics is the semi-variograms. The geostatistical estimation of the resources of an ore deposit is highly dependent on the quality and stability of the semi-variograms of the sample grades (Champigny and Armstrong, 1990). Gold deposits commonly show a skewed grade distribution. The few high grades at the tail end of the distribution produce unstable semi-variograms that are difficult to model. If these grades are

cut to a lower grade value or removed, more stable semi-variograms are obtained. However, outlier grades may make a deposit economically viable, and their cutting or removal should be treated with caution. That is the reason why a transformation of the raw data is performed, to reduce (but not to eliminate) the effect of high grades. Semi-variograms of the transformed grades are resistant to high grade outliers and present a clear structure.

3.3.1 Stationarity

Regionalised variables are variables that are distributed throughout a space (Journel and Huijbregts, 1978). For mining application, there are two characteristics that are evident in a regionalised variable. The first characteristic is a local randomness that gives the impression of a random variable. The second is a general, structural pattern that can be represented by a function. The two can be interpreted probabilistically using random functions. A sample's grade can be considered a particular realisation of a random variable. A random function is all possible realisations of that random variable. At a point, a grade is considered a random variable. However, a pair of points that are separated by a distance are independent, but also correlated through a certain spatial structure known as the semi-variogram. Because many grade realisations would be necessary to define the random function at a certain point, certain assumptions must be made. One such assumption concerns deposit homogeneity or stationarity. Under strict stationarity, the spatial structure must be invariant under translation (Journel and Huijbregts, 1978). That means the expected mean of the random variable must be constant in any location. In practice, quasi-stationarity is assumed to exist. Under that hypothesis, the spatial structure is assumed invariant for a limited distance known as the range. The range can be rationalised as the limits of a homogeneous zone in a deposit. Variables separated beyond the range are considered uncorrelated.

3.3.2 The Covariance Function or Variogram

The empirical variogram provides a description of how the data are correlated with distance. The semi-variogram function, $\gamma(h)$, was originally defined by Matheron (1963) as half the average squared difference between points separated by distance, h. The terms semi-variogram and variogram are often used interchangeably. By definition, the semi-variogram $\gamma(h)$ is one half of the variogram $2\gamma(h)$.The semi-variogram is calculated as:

$$\gamma(h) = \frac{1}{2|N(h)} \sum_{N(h)} (z_i - z_j)^2$$

where $N(h)$ is the set of all pairwise Euclidean distances $i - h = h$, $|N(h)|$ is the number of distinct pairs in $N(h)$, and z_i and z_j are data values at spatial locations i and j, respectively. In this formulation, h represents a distance measure with magnitude only. Sometimes, it might be desirable to consider direction in addition to distance. In such cases, h will be represented as the vector **h**, having both magnitude and direction.

Considering a stationary random function Y with known mean m and variance σ^2, the mean and variance are independent of location, that is, $m(\mu) = m$ and $\sigma^2(\mu) = \sigma^2$ for all locations μ in the study area (Gringarten and Deutsch, 2001). The variogram can also be defined as:

$$2\gamma(h) = Var[Y(\mu) - Y(\mu + h)] = E\{[Y(\mu) - Y(\mu + h)]^2\} \quad (1)$$

That is, the expected squared difference between two data values separated by a distance vector **h** is the variogram.

Whilst the variogram is a measure of variability and increases as samples become more dissimilar, the covariance on the other hand is a statistical measure that is used to measure correlation (it is a measure of similarity). Covariance provides a measure of the strength of the correlation between two or more sets of random variates.

$$C(h) = E\{[Y(\mu).Y(\mu + h)] - m]\} - m^2 \quad (2)$$

By definition, the covariance at $h = 0$, $C(0)$, is the variance σ^2. The covariance $C(h)$ is 0.0 when the values h-apart are not linearly correlated. Expanding the square in Equation (1) leads to the following relation between the semi-variogram and covariance:

$$(\gamma(h) = C(0) - C(h) \; or \; C(h) = C(0) - \gamma(h)) \qquad (3)$$

This relation depends on the model decision that the mean and variance are constant and independent of location. These relations are the foundation for variogram interpretation. That is, (1) the "sill" of the variogram is the variance, which is the variogram value that corresponds to zero correlation; (2) the correlation between $Y(\mu)$ and $Y(\mu + h)$ is positive when the variogram value is less than the sill; and (3) the correlation between $Y(\mu)$ and $Y(\mu + h)$ is negative when the variogram exceeds the sill. The relationship between the covariance function and semi-variogram is best illustrated in Figure 3.12.

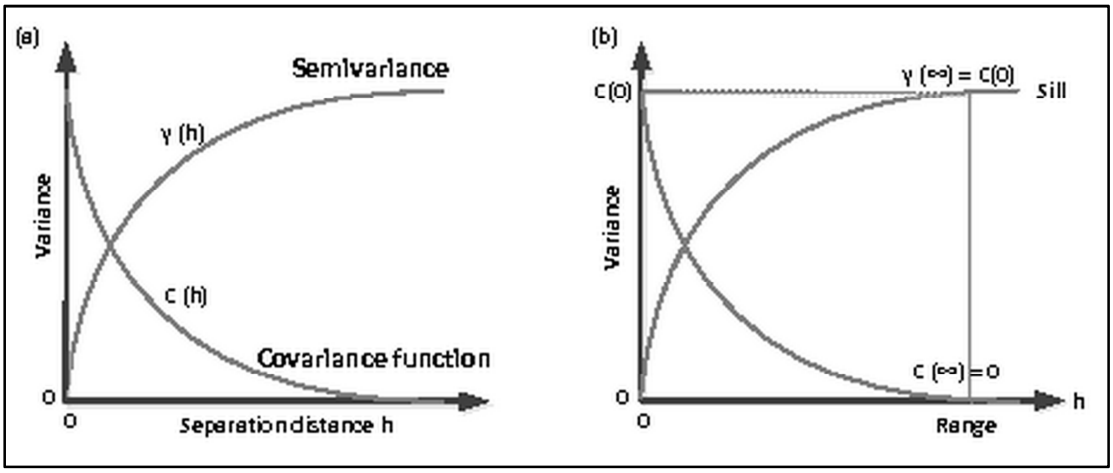


Figure 3.12: Relationship between the covariance function and semi-variogram (Source: *http://inspirehep.net/record/1189977/plots*, Gentile *et al.*, 2013).

The covariance structure can be anisotropic, meaning that there can be many different covariance functions that describe the covariance structure along different orientations. For example, variables spaced in the horizontal direction might follow a certain covariance function, while those spaced in the vertical direction might follow another.

The main goal of a variogram analysis is to construct a variogram that best estimates the autocorrelation structure of the underlying stochastic process. Most variograms are defined

through several parameters; namely, the nugget effect, sill and range. The nugget effect represents microscale variation or measurement error. It is estimated from the empirical variogram as the value of $\gamma(h)$ for $h = 0$. At a separation distance of zero, two variables have the maximum covariance C_o. The corresponding variogram value $\gamma(0) = 0$, meaning that the variance of two variables located at the same point is equal to zero. The variogram function is discontinuous in that at a very small separation distance, the variogram function equals a certain value known as the nugget value. The nugget value is caused by measurement errors and micro variability in mineralisation (Journel and Huijbregts, 1978). It is a local randomness that is similar to the 'white noise' of other phenomena. The sill represents the $\lim_{h\to\infty} \gamma(h)$ which represents the variance of the random field. The range is the distance (if any) at which data is no longer auto correlated.

3.3.3 Calculating Semi-Variograms

When a drill hole or channel sample is taken orthogonally to a horizontal bench or vein, the graded sample grade is found by integrating the grade over the intersection length and dividing by the length (Journel and Huijbregts, 1978). In the discrete case, that amounts to calculating the weighted average grade (with respect to length) over the vein thickness. The regularizing process decreases the overall grade variability. Intuitively, that makes sense. Considering a drillhole that is sampled at one metre intervals, those samples have a certain variance. In a gold deposit, the grades might appear erratic. If the samples are then composited (averaged) over five metre intervals, the variance of the composites will be less than the variance of the one metre samples. In other words, the composite grades will appear to be less erratic than the one metre samples.

Spatial variability of the dataset was assessed using semi-variograms. All semi-variograms used in this research will be modelled by spherical semi-variograms of the form:

$$\gamma(h) = \begin{cases} c_0 \left[\frac{3}{2} \frac{h}{a_0} - \frac{1}{2} \left(\frac{h}{a_0} \right)^3 \right], & for\ h \leq a_0 \\ c_0, & for\ \ h > a_0 \end{cases}$$

Semi-variogram calculation and modelling was performed on the gaussian transformed dataset using Geovariances Isatis Software. An omni-directional experimental semi-variogram for the dataset was constructed, see Figure 3.13.

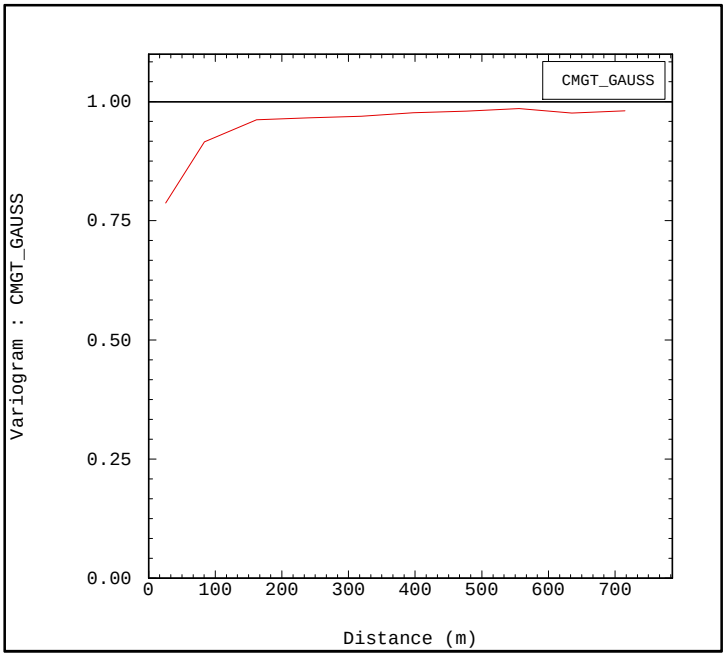


Figure 3.13: Omni – directional experimental semi-variogram model for cm.g/t.

Observation from Figure 3.13 shows that the experimental semi-variogram structure is well developed and has a long range. However, the nugget effect is high (approximately 50 % nugget to sill ratio) which is expected for a gold deposit. Using OK estimation technique with a poor semi-variogram model may produce worse estimates than the other simpler methods and is most successful when anisotropy is properly described and when the semi-variogram is locally customised (Isaaks and Srivastava, 1989). Therefore it is very important to make sure that the anisotropy is properly measured. Hence, anisotropic tests on the dataset were carried out.

A variogram map, a tool in Isatis software package, was constructed to determine whether the dataset is characterised by isotropic (omni-directional) or anisotropic (directional) behaviour. The anisotropic semi-variance surface or “variogram map” provides a visual picture of semi-variance in every compass direction. The software displays the results as a map illustrating the Gamma values in each lag for each sector by means of a colour scale. This is usually done along each of the three primary planes (XY, YZ, and XZ). However, in the case for this study, this was done along the XY planes. This therefore allows one to more easily find the appropriate principal axis for defining the anisotropic semi-variogram model. A transect in any single direction is equivalent to the semi-variogram in that direction: the surface (Z-axis) is semi-variance. The centre of the map corresponds to the origin of the semi-variogram $\gamma(h) = 0$ for every direction (Isaaks and Srivastava, 1989 and Goovaerts, 1997).

Based on the visual assessment of the map, semi-variograms from several directions were selected to look at the quality of the structure in the semi-variogram, and further assess the anisotropy. An analysis on the spatial continuity of the gaussian variable, cm.g/t in all the directions of the space in a 2D view was done to pick the possible anisotropies, see Figure 3.14 . The variogram map is displayed using a fan representation.

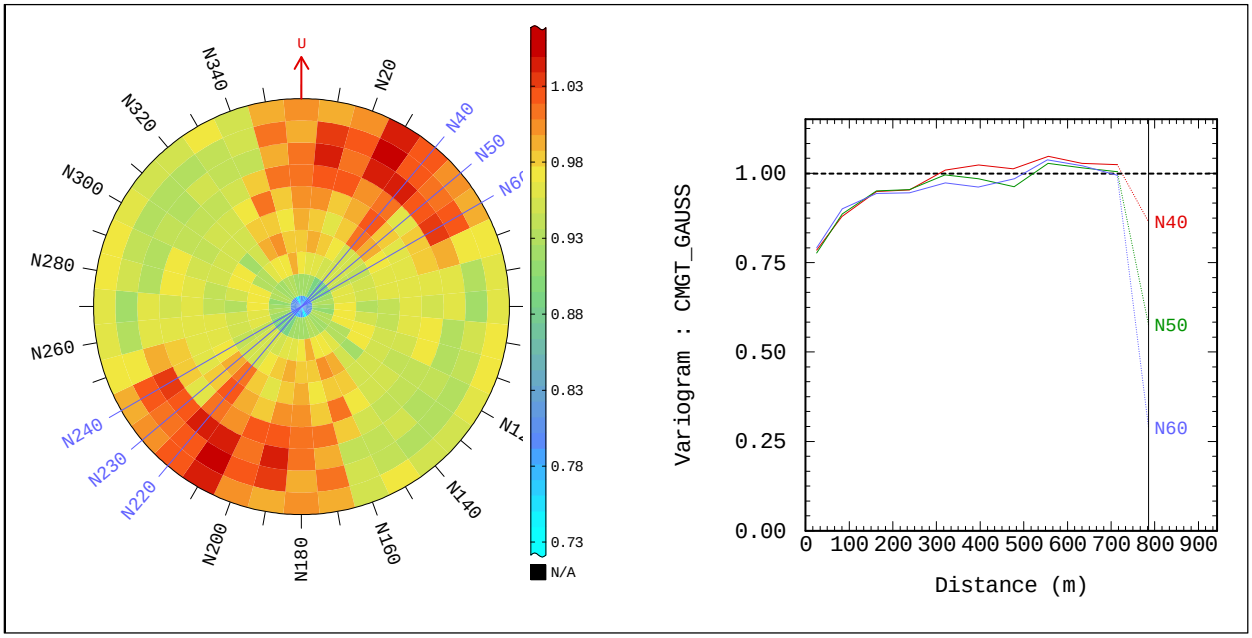


Figure 3.14: Variogram map plot for Gaussian cm.g/t.

Specific directions (N40°, N50° and N60°) highlighted from this variogram map were extracted and their corresponding experimental semi-variograms displayed. Directions N50° and the direction orthogonal to this plane, N140° were used in the calculation of the experimental directional semi-variogram model. Figure 3.15 shows the fitted semi-variogram models for cm.g/t gaussian transformed data for directions N50°, (left) and N140° (right), across and along strike respectively.

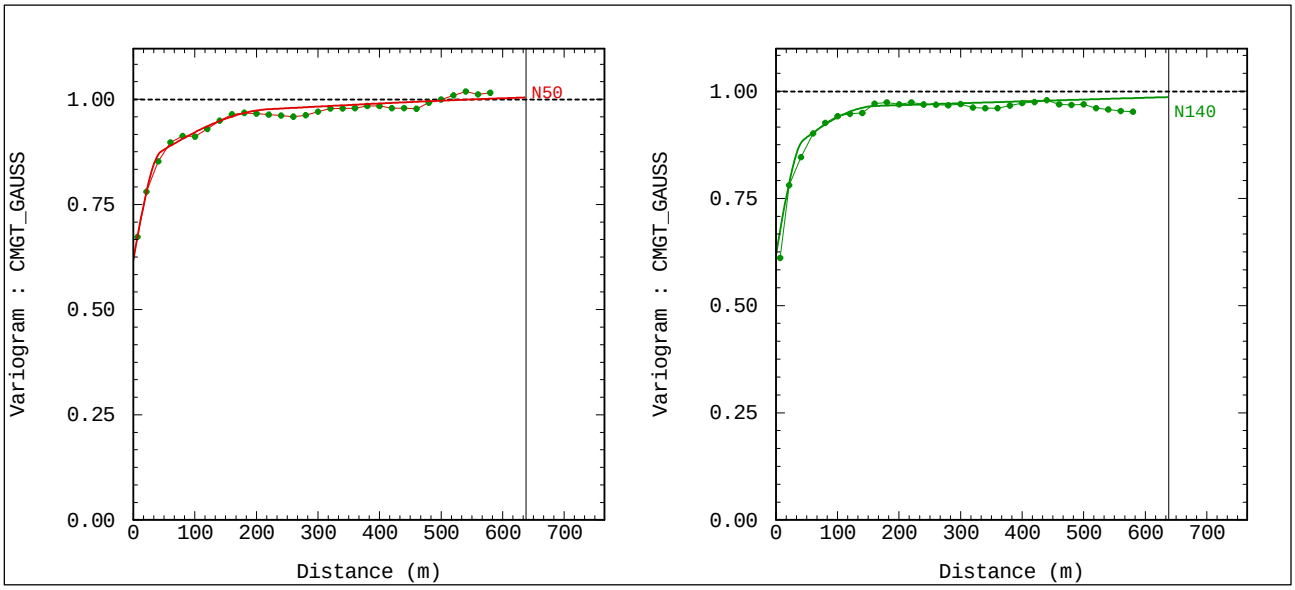


Figure 3.15: Fitted semi-variogram models for cm.g/t gaussian transformed data for directions N50 (left) and N140 (right).

Transformed estimates are rarely needed, so a back transform is used to restore the estimate to some meaningful value. Like the original transformation, the back transform does need to have an actual analytical expression; it could be a graphical procedure or any other consistent and repeatable method for restoring the transformed estimate to an estimate of the original value (Isaaks and Srivastava, 1989). The semi-variogram model (Figure 3.15) was then back transformed from the gaussian cm.g/t into a discretized semi-variogram (Figure 3.16) using the already saved anamorphosis function calculated during the transformation of the raw cm.g/t into the gaussian cm.g/t as explained in chapter 3.2.3. The raw variable (Y) and the gaussian variable (Z) are linked by the following anamorphosis function phi:

$$Z = phi\ (Y)$$

The semi-variogram value was calculated from the model for each lag and, using the anamorphosis function, the associated raw semi-variogram value was deduced and experimental semi-variograms fitted with theoretical models.

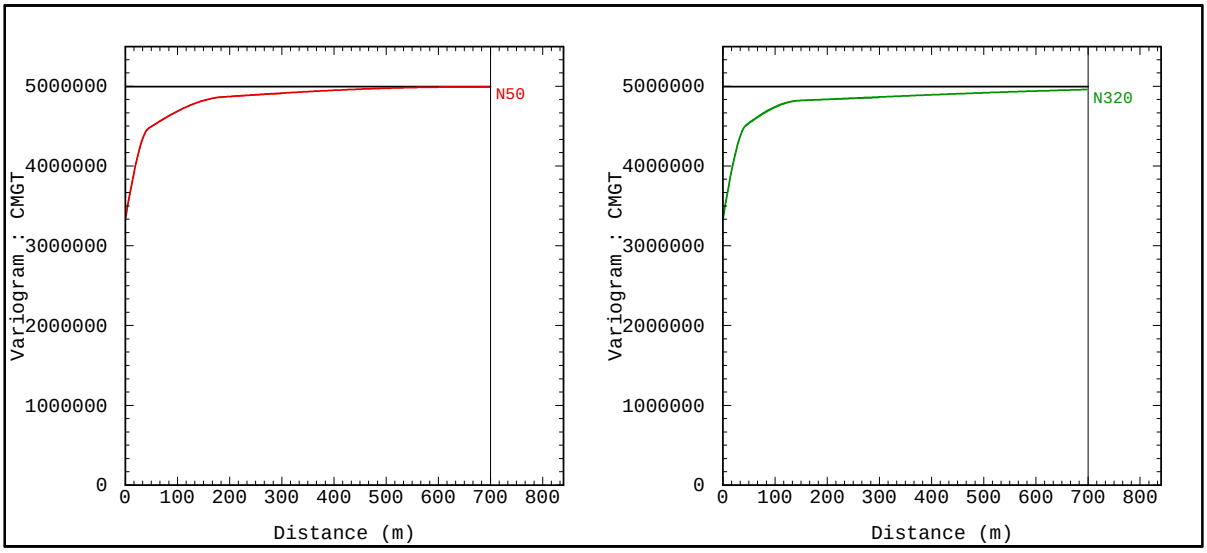


Figure 3.16: Cm.g/t back transformed fitted semi-variogram model.

The following observations were made from Figure 3.15 and Figure 3.16;

- The experimental semi-variograms are robust and well-structured and therefore easy to model;
- The models fitted onto the experimental semi-variograms are reasonably robust;

- The ranges are long indicating grade continuity for a longer distance;
- The nugget effect is rather high which indicates high spatial dependence. Therefore at small distances, the variable exhibits a rather low degree of continuity suggesting a high spatial variation at small distances. This is typical of what is expected in a coarse gold deposit;
- The semi-variogram model also shows that there is no trend in the dataset because it stabilises at the population variance. It also shows that it is extracted from the same domain i.e. the data have second order stationarity. The semi-variograms also show that there are different sills in different directions. This is common in a channelized environment.

Table 3-2 shows details of the semi-variogram model parameters for cm.g/t used for estimation.

Table 3-2: Semi-variogram model parameters for cm.g/t

				Range 1			Range 2			Range 3	
Zone	**Strike**	**Nugget**	**Sill 1**	**X**	**Y**	**Sill 2**	**X**	**Y**	**Sill 3**	**X**	**Y**
All	50	3335010	937042.5	43.99	43.99	502345.7	187.32	147.2	216331.7	650.06	1049.73

Due the high variance values recorded, the data was normalised such that the total variance equals one. Table 3-3 summarises the normalised semi-variogram parameters used in the estimation process.

Table 3-3: Normalised semi-variogram model parameters for cm.g/t

				Range 1			Range 2			Range 3	
Zone	**Strike**	**Nugget**	**Sill 1**	**X**	**Y**	**Sill 2**	**X**	**Y**	**Sill 3**	**X**	**Y**
All	50	0.66824	0.18776	43.99	43.99	0.10066	187.32	147.2	0.0433	650.06	1049.73

The modelling of the short range structures adequately reflects the continuity seen in the experimental data, and the longer ranges modelled will have little impact on the quality of the resource estimates. The semi-variograms indicate spatial continuity of 44 m along and across strike.

3.4 Quantitative Kriging Neighbourhood Analysis

As the data locations become farther away from the location where the value is unknown they may not be as useful when predicting the value at an unmeasured location. At some distance, the points will have no correlation with the prediction location and it is possible that they may even be located in an area much different than the unknown location. Therefore it is common practice to specify a search neighbourhood that limits the number and the configuration of the points that will be used in the predictions. There are two controlling mechanisms to limit the points used, namely, specifying the shape of the neighbourhood and establishing constraints on the points within and outside the shape. For estimation methods that can handle any number of nearby samples, the most common approach to choosing the samples that contribute to the estimation is to define a search neighbourhood within which all available samples will be used. The search neighbourhood is usually an ellipse centered on the point being estimated. The orientation of this ellipse is dictated by the anisotropy in the spatial continuity (Isaaks and Srivastava, 1989). Therefore, defining a search neighbourhood within which all available samples are used is very important. The ellipse will be oriented with its major axis parallel to the direction of maximum continuity.

Quantitative Kriging Neighbourhood Analysis ("QKNA") is a quantified approach to choosing the search neighbourhood in a kriging estimate (Vann *et al.*, 2003). In this study, it was used to calculate the maximum number of samples used as well as the number of discretization points used for each block size.

3.4.1 Sample Number Optimisation

A study to determine the maximum number of samples to be used in an estimate was done using QKNA exercise. In their paper titled. "Quantitative Kriging neighbourhood analysis

for the mining geologist - a description of the method with worked case examples", (Vann et al., 2003) state that, the criteria to consider when evaluating a particular kriging neighbourhood, in the order of priority used in practice are:

- Slope of regression, ("SLOR") of the sample data value vs. the estimates;
- Weight of the mean in simple kriging;
- The distribution of kriging weights (including the proportion of negative weights);
- The kriging variance ("KV");
- Percentage blocks filled;
- Mean of the estimate vs. mean of the sample data.

The slope is given in terms of the covariance and the variance of the estimated blocks by the expression:

$$\frac{Z^*}{Z} = \frac{Cov\,(Z_v\, {Z_v}^*)}{Var\,({Z_v}^*)}$$

where:

Z^* / Z is the slope of the regression

Z_v is the true block grade

${Z_v}^*$ is the estimated block grade

QKNA was run for a range of increasing sample number, see Figure 3.17. It shows a plot of the SLORs, kriging efficiencies and estimation variance done on a 30×30 m block size using a quadrant search. The equation used to compute the kriging efficiency is:

$$KE = \frac{BV - KV}{KV}$$

Where *KE* = kriging efficiency, *BV* = theoretical variance of blocks within the domains.

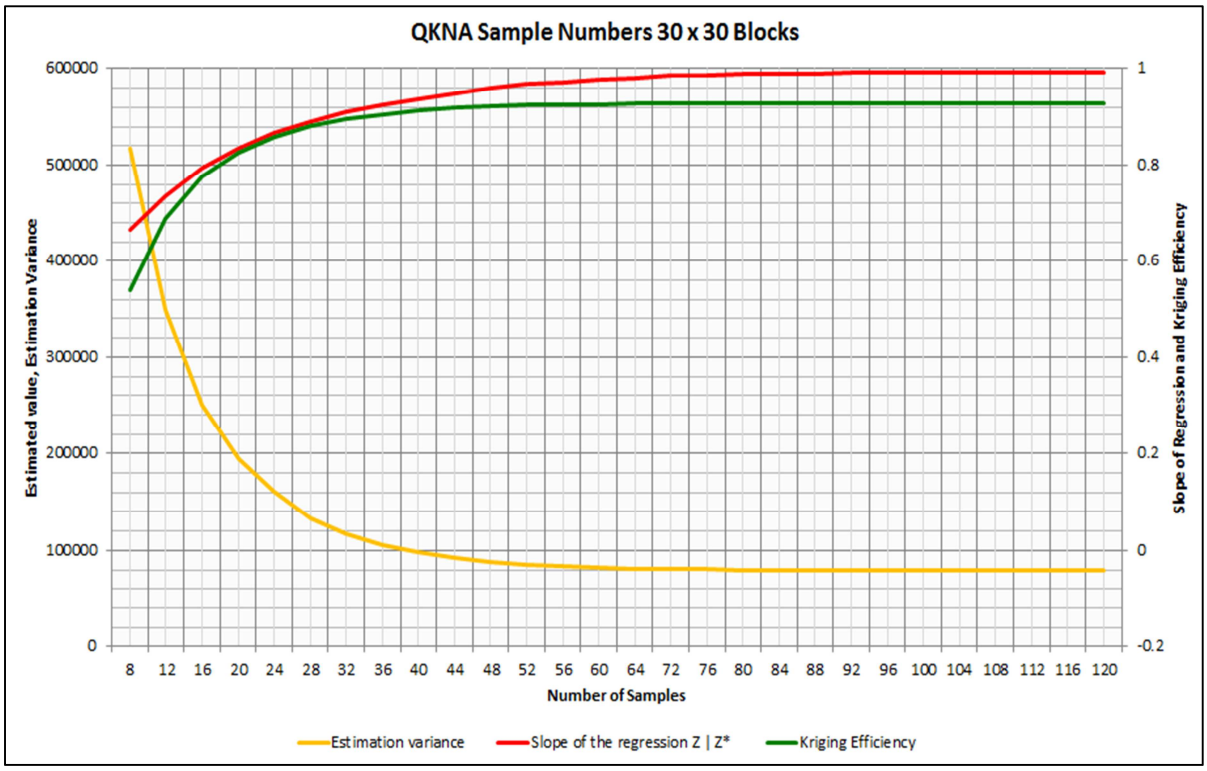


Figure 3.17: QKNA sample number optimisation graph

Ideally, the SLOR, $(\frac{Z^*}{Z})$ should be very close to 1.0 and thus implying conditional unbiasedness. In these circumstances, the true grade of a set of blocks should be approximately equal to the grade predicted by the kriged estimation (Vann *et al.*, 2003). For the purpose of this exercise, the writer chose 8 and 36 for the minimum and maximum number of samples respectively to be used for the 30×30 m block size because further increasing the number of samples will not significantly increase the SLOR and KE (the line graphs flatten) thus causing over smoothing in the resulting estimates.

3.4.2 Block Discretization Number Optimisation

Vann et al., 2003, also state that the discretization in block kriging is used to calculate the point-block average values of variogram (or covariance) functions, i.e. $\bar{\gamma}(x_i, V)$ or $\bar{C}(x_i, V)$. If the estimation consists of producing the optimal average value of a variable over a target block *(v)*, the average covariance between the data point and the target cell i.e. $\bar{C}(v, V)$, as well as the average covariance of the cell i.e. $\bar{C}(V, V)$, have to be calculated. A block is

regularly discretized according to the number of discretization points along X, Y and Z. The points are randomly and independently moved along X, Y and Z. The moving distance in a direction is smaller than the size of the block along the same direction divided by twice the number of discretizations along this direction. These moved points are used to compute $\bar{C}(v,V)$ and $\bar{C}(V,V)$. Calculations for each discretization were done several times, moving the origin of the discretization grid each time. According to Vann et al., 2003 stable results indicate that the discretization is adequate. Figure 3.18 shows a graph of the impact of discretization on the resulting range of $\bar{C}(V,V)$ values.

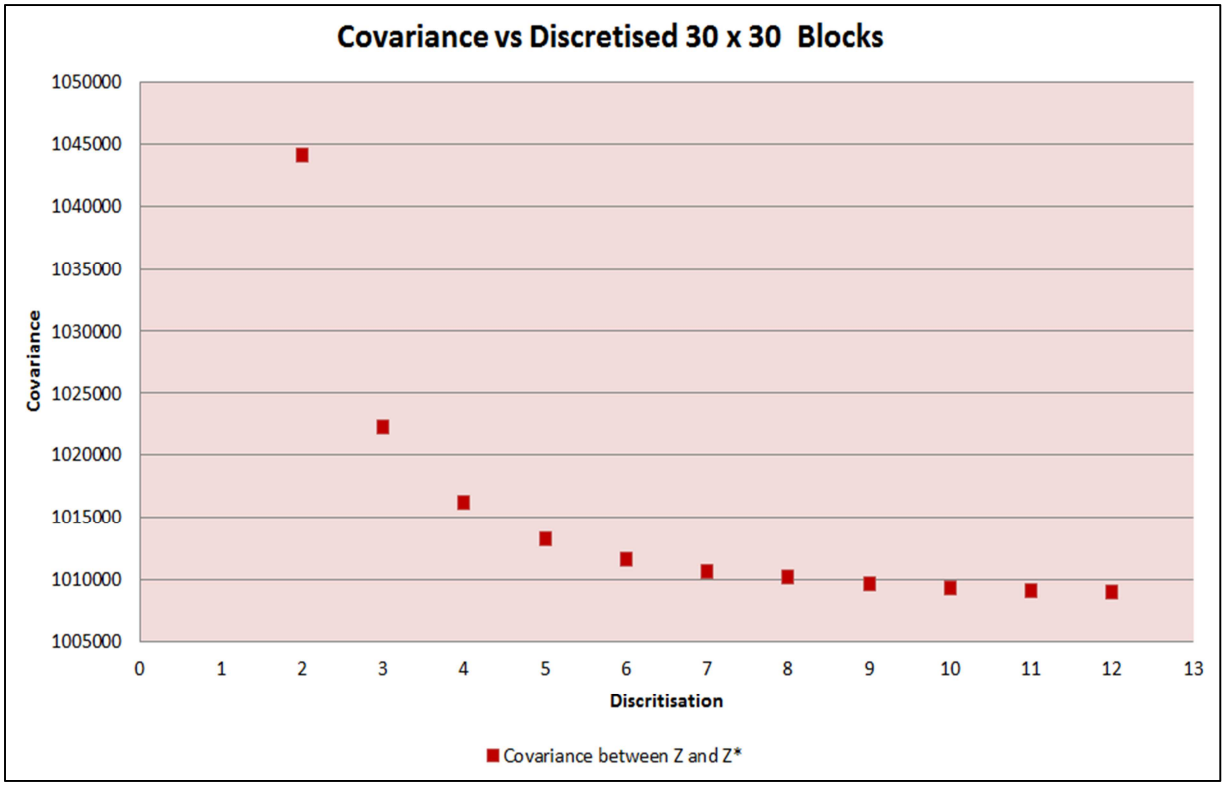


Figure 3.18: Graph showing sensitivity *of* $\bar{C}(V,V)$ to block discretization.

Observations from Figure 3.18 show that the variance for results, plotted on the Y-axis does not show significant continuing downward trend beyond about 5×5 m and 6×6 m. However for the purpose of this study a discretization value of 5×5 m was chosen.

3.5 Geostatistical Estimation

Numerous authors have discussed the fundamentals of geostatistics in detail such as Deutsch and Journel, 1998 and Isaaks and Srivastava, 1989. Geostatistical estimation can be defined

as the prediction of the value of an attribute at a location at which it is unknown from measured values of that characteristic at a number of known locations through a function defining the spatial correlation between values (Dohm, 2010). Both kriging and simulations have had much theoretical development and are well described in most geostatistical literature and are mostly used in the mining industry. Most geostatistical methods rely on the assumptions of stationarity, which is seen as the decision to pool data within a given area or domain. Making proper decisions about stationarity is very important as they are critical for the representativeness and reliability of the geostatistical tools used (Deutsch and Journel, 1998). Samples from geozones or domains of similar geology i.e. rock type, chemical characteristic, structure, grade are grouped together. Dividing a dataset into areas that are acceptable with regards to stationarity is therefore essential in geostatistics (Isaaks and Srivastava, 1989).

3.5.1 Kriging

Geostatistical interpolation ("kriging") was developed by Matheron, a French mathematician in the early 1960s based on the Master's thesis of Danie G. Krige, the pioneering plotter of distance-weighted average gold grades at the Witwatersrand reef complex in South Africa. It provides the best linear unbiased prediction for spatially dependent properties based on assumptions on covariances (Journel and Huijbregts, 1978; Vieira *et al.*, 1997). Kriging predicts the value of a function at a given point by computing a weighted average of the known values of the function in the neighbourhood of the point. It makes use of Gauss-Markov theorem to prove independence of the estimate and error and has therefore been frequently used for the spatial interpolation of mineral deposits (Sousa, 1989). However, the great variability on ore concentration in conjunction with sparse sampling may mask the spatial dependence (Journel, 1983).

OK is the form of kriging used mostly because it works under simple stationarity assumptions and does not require knowledge of the mean (Chilès and Delfiner, 2012). A kriged estimate Z^* of the grade of a given block is the linear combination of available sample grades X, which satisfies the following conditions:

i. The estimate must be unbiased, whatever the true unknown mean grade of the deposit may be. That is, on average the estimated values will equal the unknown values – "stationarity". There should be no drift or trend; and

ii. The estimation errors (Z-Z*), the estimation variance, will be as small as possible – Normal distribution.

Because of its weighed moving average, the kriging method resembles the interpolation method, but, since it determines the weights according to the spatial analysis provided by the experimental semi-variogram, it diverges from the simple linear interpolation and Inverse Square of distance methods. The kriging estimator supplies unbiased estimations, that is, on the average; there is no distortion between the estimated and actual values of a same point. Another important characteristic is the minimum variance, which means that, although there may be differences between the estimated and measured values from point to point, these differences are very low. As a linear estimator, a kriging procedure may be represented by the weighed linear combinations of measured data, or a moving average that considers a variability structure of the measured variable, expressed by the semi-variogram and by the localization of known values (Porto de Carvalho *et.al.*, 2004). In this concept, the points closest to the position that will be interpolated have a major weight in relation to the most distant ones. The kriging estimator is formulated as follows:

$$Z^*(x_0) = \sum_{i=1}^{N} \lambda_i \, Z(x_i)$$

where N is the number of measured values $Z(x_i)$ used in the prediction, and λ_i is the weight associated to each measured value $Z(x_i)$.

3.5.2 Dynamic Anisotropy Application on 30×30 m Block

Using Datamine Studio 3 software, digitized strings were used to define the orientation (the direction of best continuity) of the mineralisation trend. The string file was input to the ANISOANG process command in Studio 3 to create a "points" file. Figure 3.19 shows the points as arrows and each "arrow" is a locally interpreted "direction".

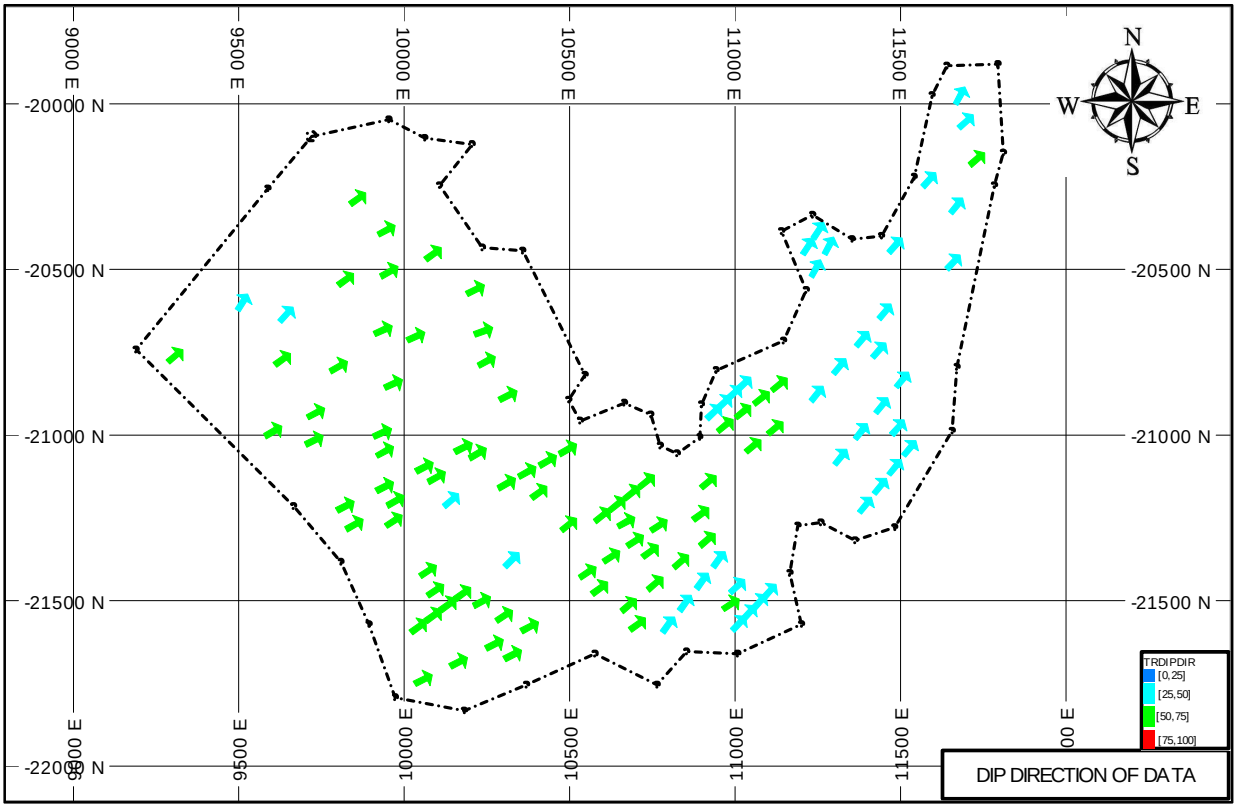


Figure 3.19: Orientation of mineralization plot.

A 30×30 m block model for the research area was generated. Using the "ESTIMA" process in Studio 3, the circular IPD method of interpolation was used to interpolate the dip direction data into the block model. The DA option in Studio 3 allows the anisotropy rotation angles for defining the search volume to be defined individually for each cell in the model. The search volume ellipse was oriented precisely and followed the trend of the mineralization. Since the estimation process was done in 2D, the dip was set at 0, and only the dip direction was defined. The final step was to interpolate the grade into the block model using the DA option in Studio 3, this time using the DA angles for the blocks instead of a set search ellipse.

During the estimation process, the space around the centre of a block being estimated was divided into four quadrants by the axial planes of the data search ellipsoid. This parameter ensured that the samples informing an estimate were relatively evenly spread within the block. At least four quadrants had to be filled before a local estimate was carried out. For all estimates, a quadrant search was carried out using a minimum and optimum number of samples per quadrant of 2 and 9 respectively with a discretization pattern of 5 (X) by 5 (Y) by 1 (Z) points used. The final search neighbourhood parameters used in the estimation process and based on the semi-variogram models in section 3.3.3 are summarised in Table 3-4.

Table 3-4: Search parameters for cm.g/t

	Search Range (m)		Number of samples	
Variable	U	V	Min	Max
cm.g/t	45	45	8	36

3.6 Domaining

Grade domaining constitutes a common practice in the mineral industry as it may increase the dispersion of the estimation errors and provoke a conditional bias in the resource estimation. Even though results from the variogram analysis discussed in section 3.3.3 indicate that the dataset was extracted from the same domain and have second order stationarity, another study was performed to find out whether domaining would improve the quality of resource estimates from the 2 methods of interpolation. Hence, the dataset was domained using the mineralisation trend as a guide to split it into the various zones. Two scenarios of domaining based on mineralisation trend were implemented:

- First scenario, the data set was split into 3 zones, see Figure 3.21; and
- Second scenario, where the dataset was split into 2 zones, see Figure 3.20.

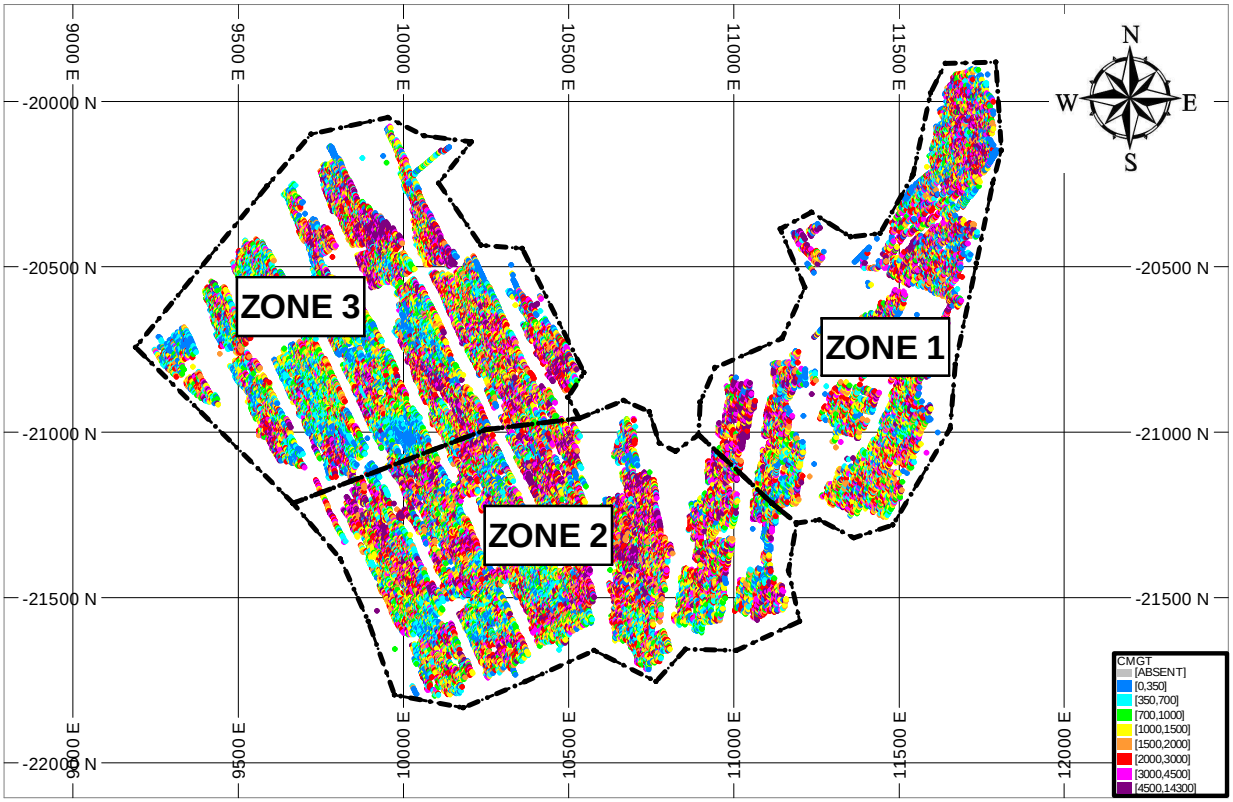


Figure 3.20: Domaining – Zone 1, 2 and 3

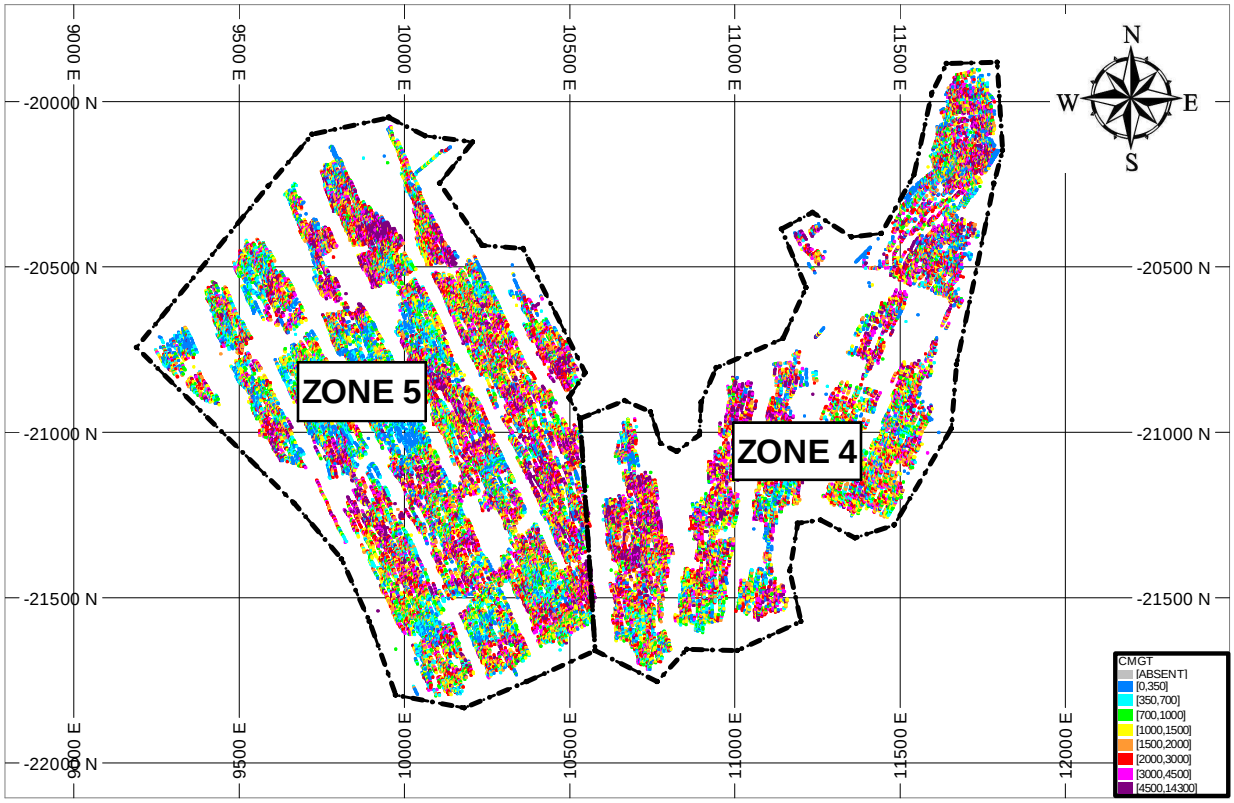


Figure 3.21: Domaining – Zone 4 and 5

3.6.1 Statistical Analysis on Domains

The statistical results for the domains are presented in Table 3-5.

Table 3-5: Zonal Statistics

ZONE	FIELD	NSamp	Min	Max	Mean	Var	Stddev	CoV
1	cm.g/t	9 512	0.00	14 222.00	2 242.64	4 844 673.93	2 201.06	0.98
	CW	9 512	2.00	210.00	19.72	318.05	17.83	0.90
2	cm.g/t	18 900	0.00	14 200.42	2 314.71	5 065 762.80	2 250.72	0.97
	CW	18 900	3.00	220.00	24.80	288.89	17.00	0.69
3	cm.g/t	18 531	0.00	14 250.90	2 108.88	5 158 822.36	2 271.30	1.08
	CW	18 531	1.00	210.00	30.63	378.98	19.47	0.64
4	cm.g/t	16 018	0.00	14 222.00	2 350.32	5 049 761.13	2 247.17	0.96
	CW	16 018	2.00	210.00	18.55	231.80	15.23	0.82
5	cm.g/t	30 925	0.00	14 250.90	2 150.76	5 061 258.36	2 249.72	1.05
	CW	30 925	1.00	220.00	29.97	362.79	19.05	0.64

The following observations are made from Table 3-5:

- Zone 2 has the highest mean value of 2314.71 for cm.g/t when compared with Zone 1 and 2 whilst Zone 3 has the highest average CW value of 30.63;
- Zone 3 has the highest CoV value of 1.08 indicating it has a higher variability when compared with the rest of the other zones.
- When comparing the CoV value of the dataset that is not domained in Table 3-1 which is 1.01 it is observed that it is lower than the values observed for zone 3, 1.08 and zone 5, 1.05. The CoV values for zones 1, 2 and 4 have reduced slightly when comparing with the un-domained dataset.

Referring to results shown in Figure 3.21, Figure 3.20 and Table 3-5 it can be deduced that domaining will not help in reducing variability in the dataset. This can be attributed to the fact that the samples are closely spaced and the inherent variability of gold values.

Histogram Plots

Histograms of the samples in the zones were constructed, see Figure 3.22. As is expected and similar to the un-domained data, the data in the various zones all reflect a positively skewed distribution.

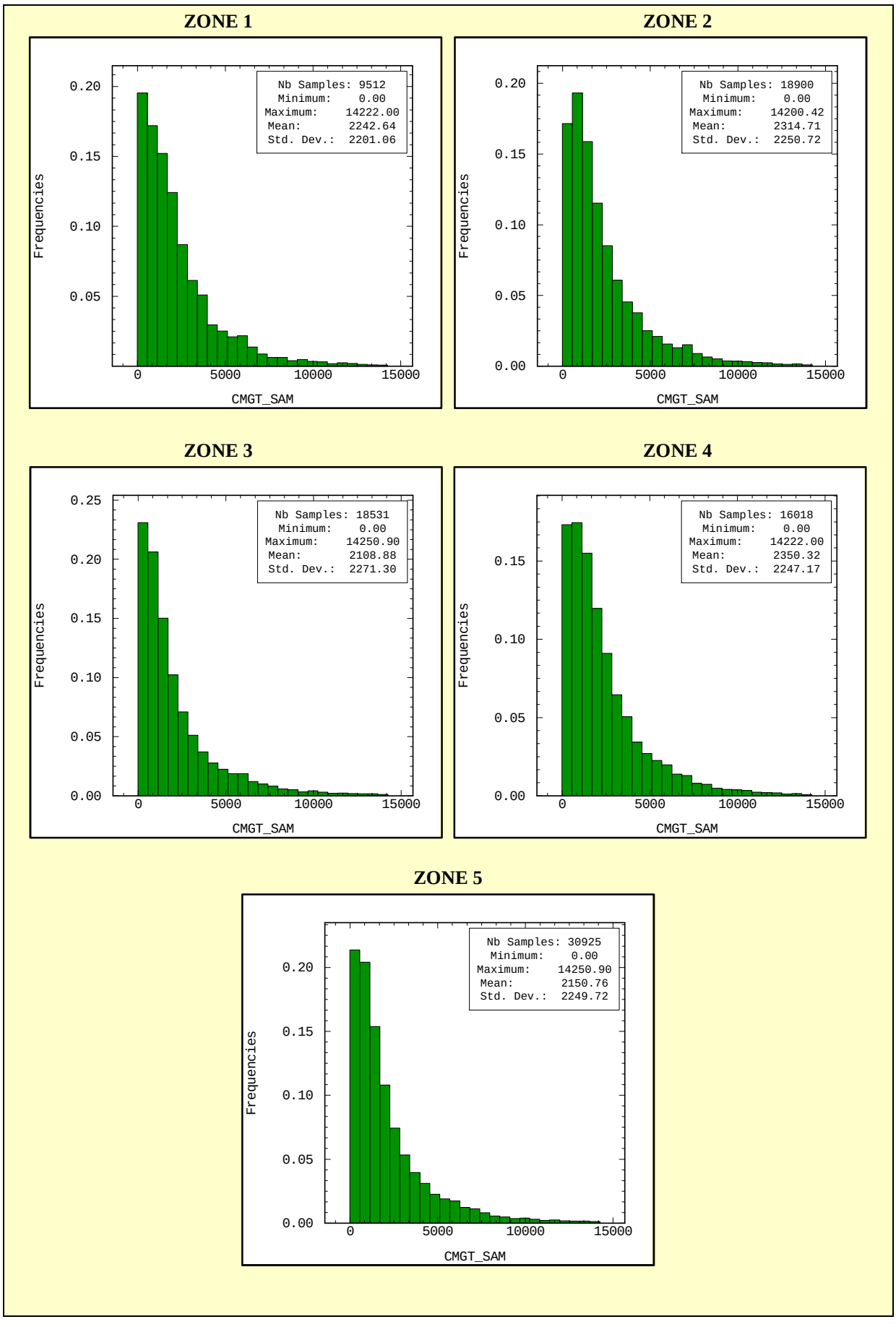


Figure 3.22: Cm.g/t histograms for domains

The data reflects a positively skewed distribution which is expected for a gold grade distribution.

3.6.2 Variogran Analysis on Domains

Data in the various domains was declustered, transformed and semi-variograms generated and modelled. Figure 3.23 to Figure 3.27 shows the modelled semi-variogram models for cm.g/t gaussian transformed data for directions N50° (left) and N140° (right). The variogram maps and the fitted semi variogram models (back transformed) are shown in Appendices A and B respectively.

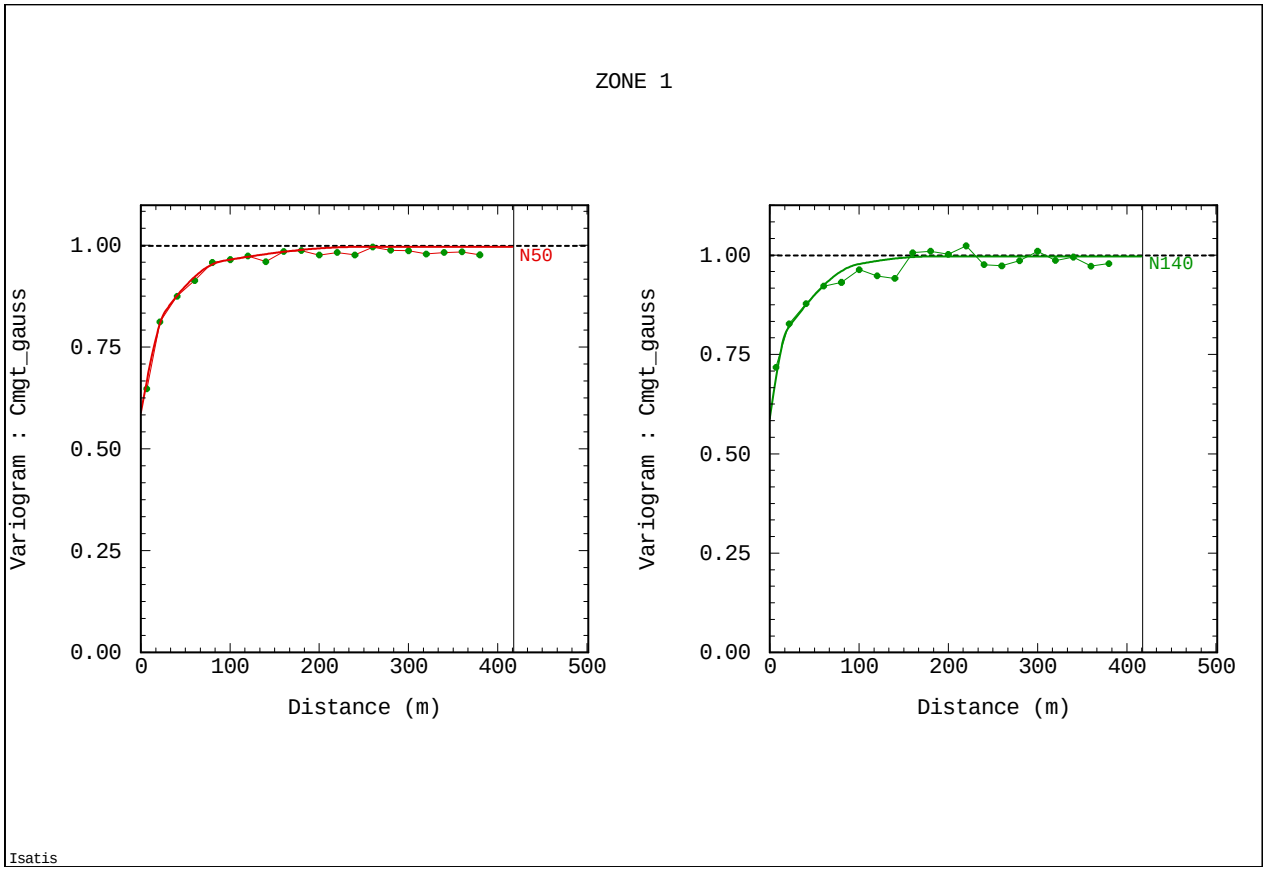


Figure 3.23: Fitted semi-variogram models for cm.g/t gaussian transformed data for directions N50 (left) and N140 (right) for Zone 1

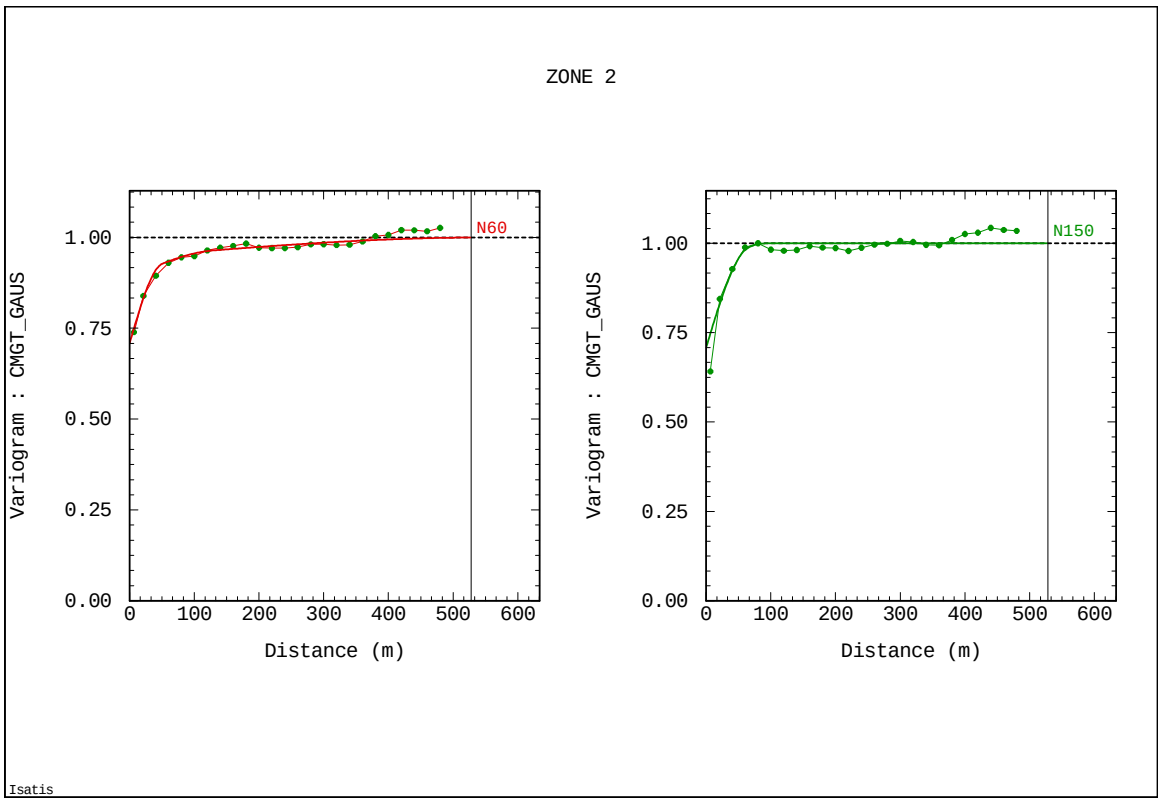


Figure 3.24: Fitted semi-variogram models for cm.g/t gaussian transformed data for directions N60 (left) and N150 (right) for Zone 2

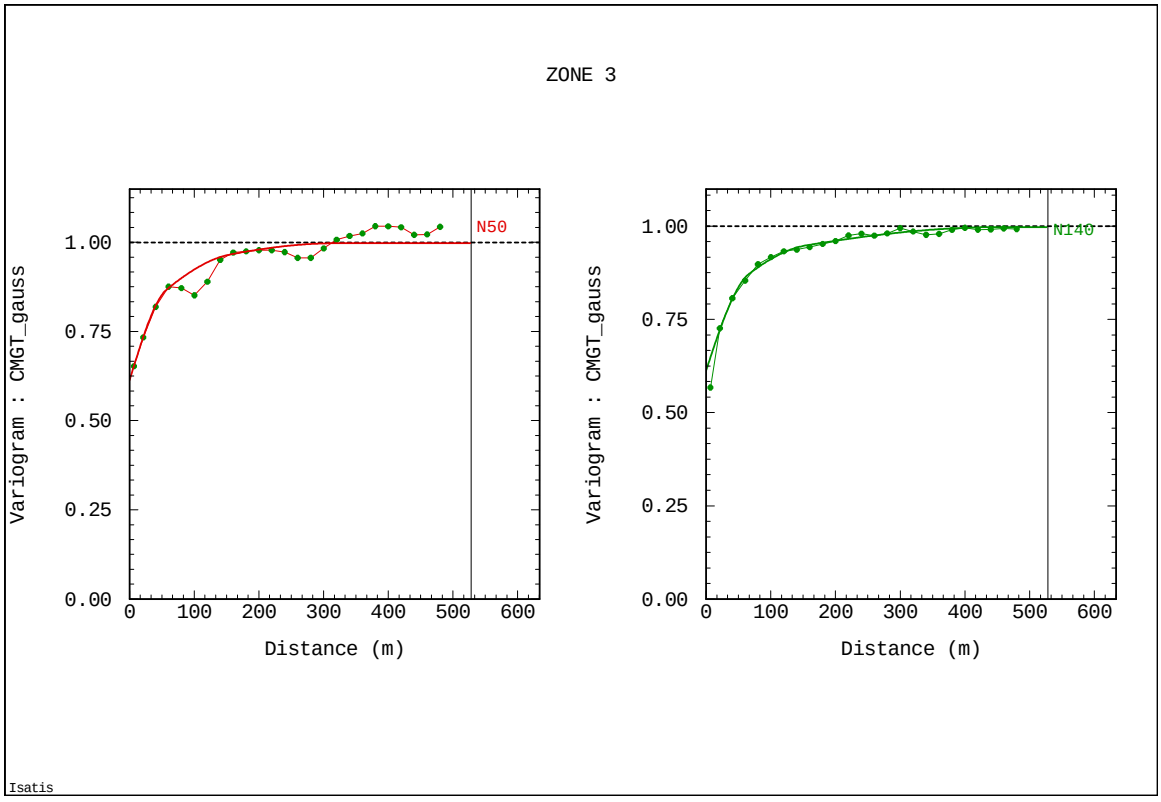


Figure 3.25: Fitted semi-variogram models for cm.g/t gaussian transformed data for directions N150 (left) and N60 (right) for Zone 3

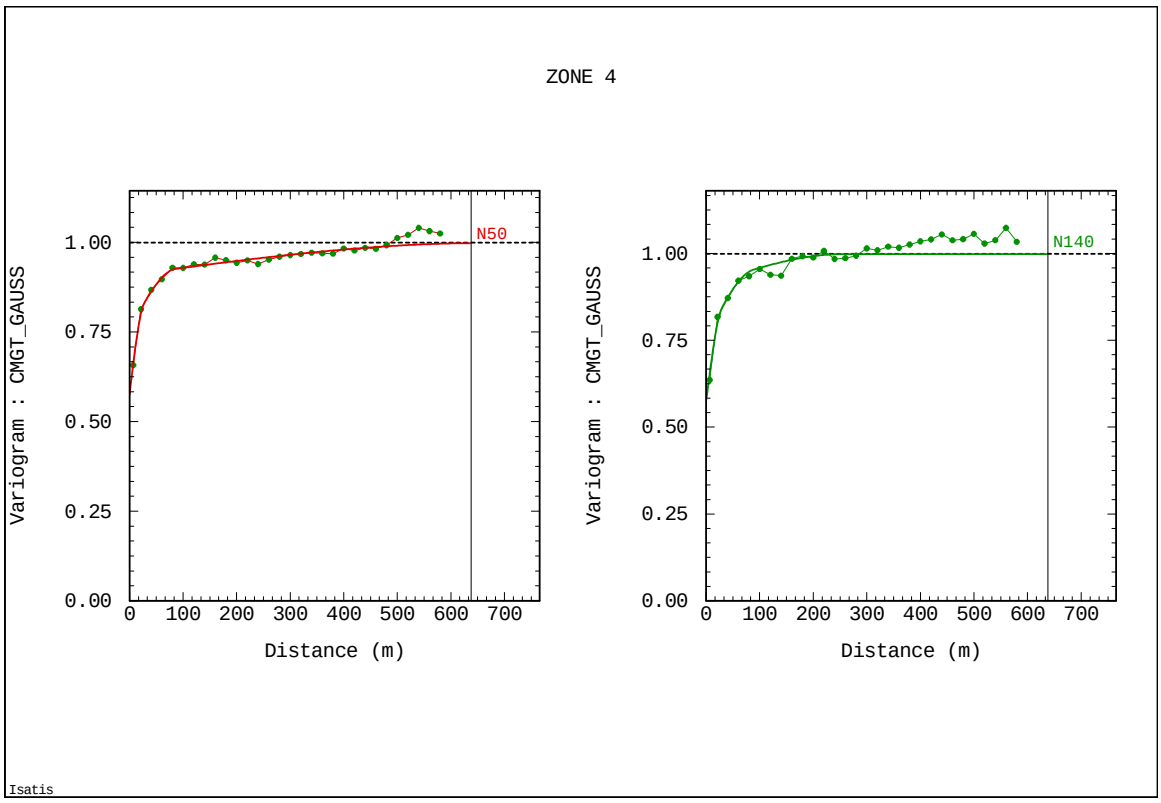


Figure 3.26: Fitted semi-variogram models for cm.g/t gaussian transformed data for directions N50 (left) and N140 (right) for Zone 4

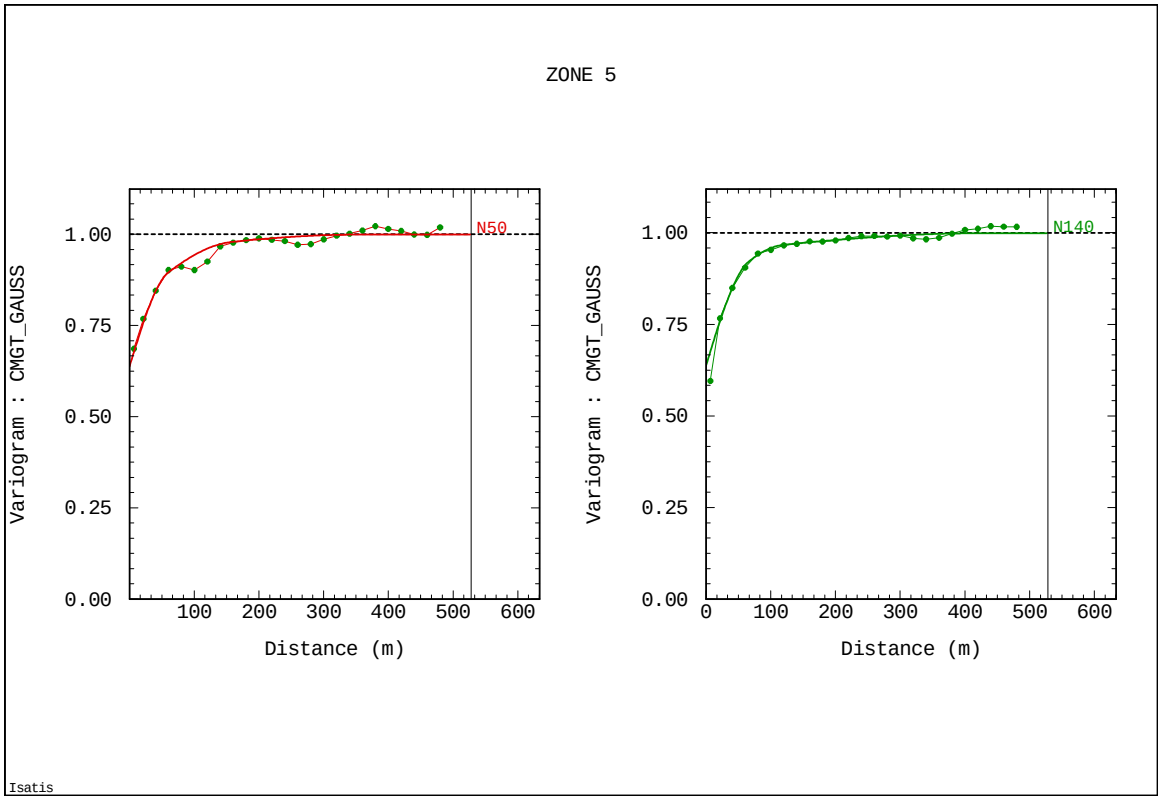


Figure 3.27: Fitted semi-variogram models for cm.g/t gaussian transformed data for directions N50 (left) and N140 (right) for Zone 5

Similar results are observed for the fitted semi-variogram models in the various domains when compared with the undomained dataset. The experimental semi-variograms for all the zones are also robust, well structured and thus easy to model. The models fitted well to the experimental semi-variograms. Generally, the ranges are long for all zones indicating that the grade continuity is very high. Table 3-6 and Table 3-7 shows details of the semi-variogram model parameters and the normalised semi-variogram parameters for cm.g/t used for estimation in all the domains.

Table 3-6: Domain semi-variogram model parameters for cm.g/t

				Ranges 1			Ranges 2			Ranges 3	
Zone	**Strike**	**Nugget**	**Sill 1**	**X**	**Y**	**Sill 2**	**X**	**Y**	**Sill 3**	**X**	**Y**
1	50	3219961	693217.6	29.45	20.45	725603.4	87.75	95.75	298901	250.69	180.69
2	60	3731880	821588.7	50.08	63.85	265428.6	185.77	82.00	143051	680.91	120.27
3	50	3274694	722830.6	58.42	63.42	492697.5	150.48	150.48	363665.3	308.79	460.79
4	50	3132192	833827.6	28.67	31.67	655931.2	91.32	91.32	358071.2	703.19	273.19
5	50	3425603	734673.7	59.96	65.96	575797.4	154.87	114.87	206123.4	385.37	485.37

Table 3-7: Domain normalised semi-variogram model parameters for cm.g/t

				Ranges 1			Ranges 2			Ranges 3	
Zone	**Strike**	**Nugget**	**Sill 1**	**X**	**Y**	**Sill 2**	**X**	**Y**	**Sill 3**	**X**	**Y**
1	50	0.65212	0.14039	29.45	20.45	0.10066	87.75	95.75	0.0433	250.69	180.69
2	60	0.75210	0.16558	50.08	63.85	0.14695	185.77	82.00	0.0605	680.91	120.27
3	50	0.67465	0.14892	58.42	63.42	0.05349	150.48	150.48	0.0288	308.79	460.79
4	50	0.62895	0.16743	28.67	31.67	0.10151	91.32	91.32	0.0749	703.19	273.19
5	50	0.69313	0.14865	59.96	65.96	0.13171	154.87	114.87	0.0719	385.37	485.37

The following can be deduced from Table 3-6: Zone 4 has the longest range in the X direction meaning that the grade continuity is quite high in comparison with the other zones.

3.7 Optimal DA Resolution Determination

In order to determine the optimal DA resolution for grade estimation improvement, the estimation process was executed on various block model range of different resolutions which include; 30×30 m, 60×60 m, 120×120 m and 240×240 m grids for both estimation techniques, DA and OK. The resulting estimates were then compared against each other; see Section 5 for results and discussion.

4 Results on 30×30 m Block Estimates

4.1 Validating the Estimates

A number of checks were conducted on the kriging results to validate the estimates. The block estimates were validated against the input composite sample "source" data using the following methods:

- Global statistics;
- Visual inspection of regularised block estimates against composites of the drillhole intersections;
- Swath plots; and
- Scatter plots of block estimates vs the average block values = "Actual".

4.1.1 Global Statistics

The estimates were validated against the source data by calculating and comparing the global statistics of the sample data, and the estimates produced by DA as well as estimates produced by OK interpolation techniques. Table 4-1 shows the global and local statistics for comparison purposes.

Table 4-1: Global and local statistics

Variable	NSamp	Min	Max	Mean	Stddev	CoV	Diff with cm.g/t-sample
cm.g/t-sample	46 943	0	14 250.90	2 218.85	2 250.84	1.01	
OK estimates	2 795	4.94	7 742.17	2 114.07	1 062.99	0.50	-4.72%
DA estimates	2 795	4.94	7 701.44	2 116.57	1 064.61	0.50	-4.61%

The following observations are made from Table 4-1:

- The block estimate average grades are slightly lower than the average grade for sample data with percentage differences of 4.72 % and 4.61 % for OK and DA estimates respectively;
- The estimates calculated using DA interpolation technique have a mean that is closer to the mean of the sample data. However, the average grades calculated by both

estimation techniques are comparable with values of 2114.07 and 2116.57 for OK and DA estimates respectively;

- The CoV values for the both estimates have halved with values of 0.50. This is consistent with literature that states that kriging smooths data due to the reduction in variability of the data.

4.1.2 Visual Inspection of Plots

In this method, the block estimates and composites are presented as a plot for comparison purposes. The sample data was declustered, a process which adjusted the full data set to give a more representative and evenly spaced set of samples. Then the average value of all samples within the regular 2D grid cell was calculated. The sample data (plotted as points) and block estimates are then plotted on the same plot superimposed on each other and colour coded with the same legend for cm.g/t, see Figure 4.1.

The visual comparison plots show that the colour coding of the sample data closely matches that of the block estimates for both estimation techniques, which indicates that grade distribution of the composites compares well with the observed distribution in the block model. However, it is difficult to determine which method of estimation is better using these plots.

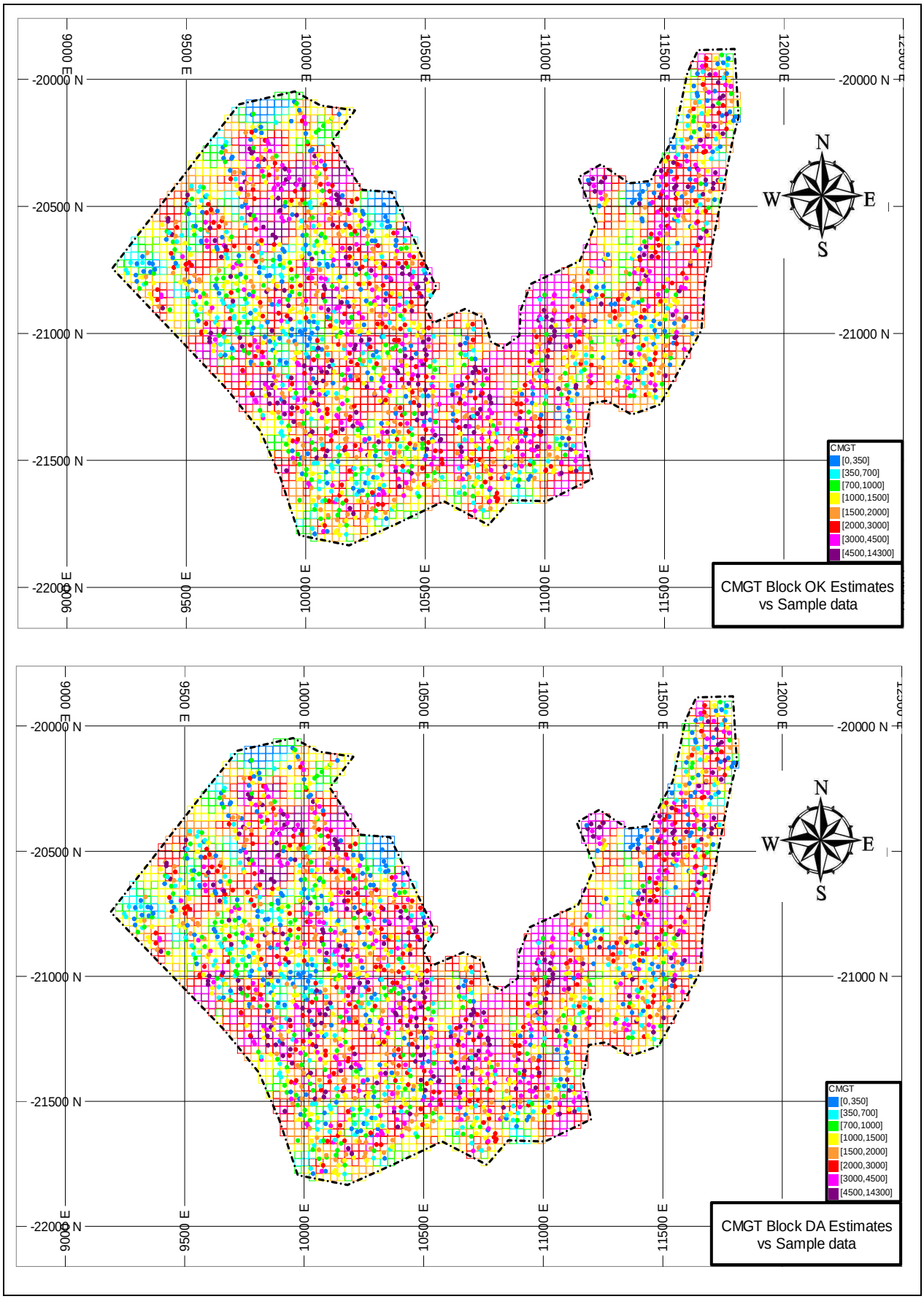


Figure 4.1: OK estimates (above) vs sample data and DA estimates (below) vs. sample data plots.

4.1.3 Swath Validation

The block model estimates were also validated against the composites by swath analyses. Swaths are perimeters generated with a spacing of 70 m, equivalent to the section lines along strike of the mineralized zone (Figure 4.2). Sample composites and block model estimates within each swath were extracted and the mean statistic compared with the expectations of close correspondence between the two sets of data. The results of the swath validation for cm.g/t sample data and block model OK and DA estimates are shown in Figure 4.3.

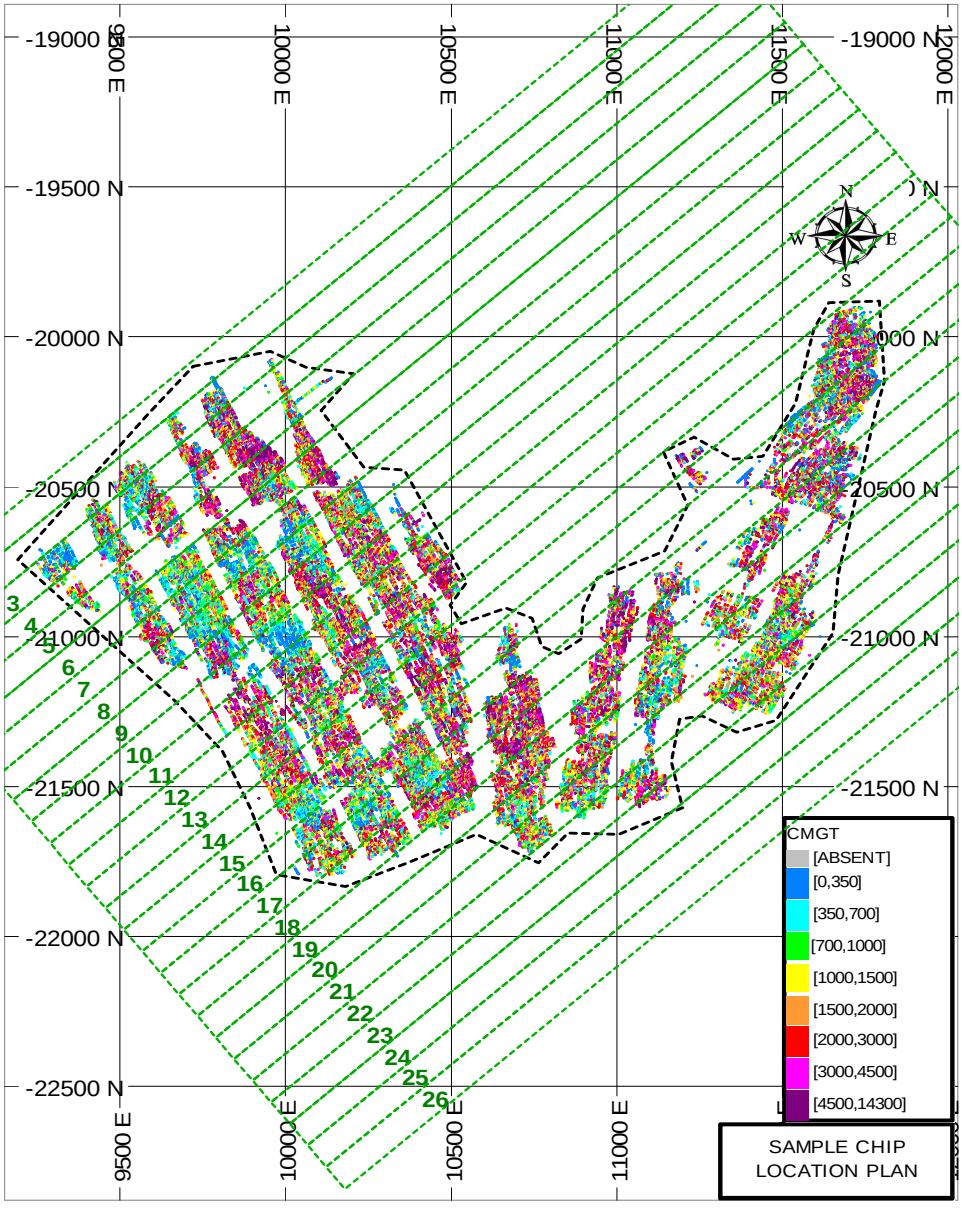


Figure 4.2: Swath orientation and spacing.

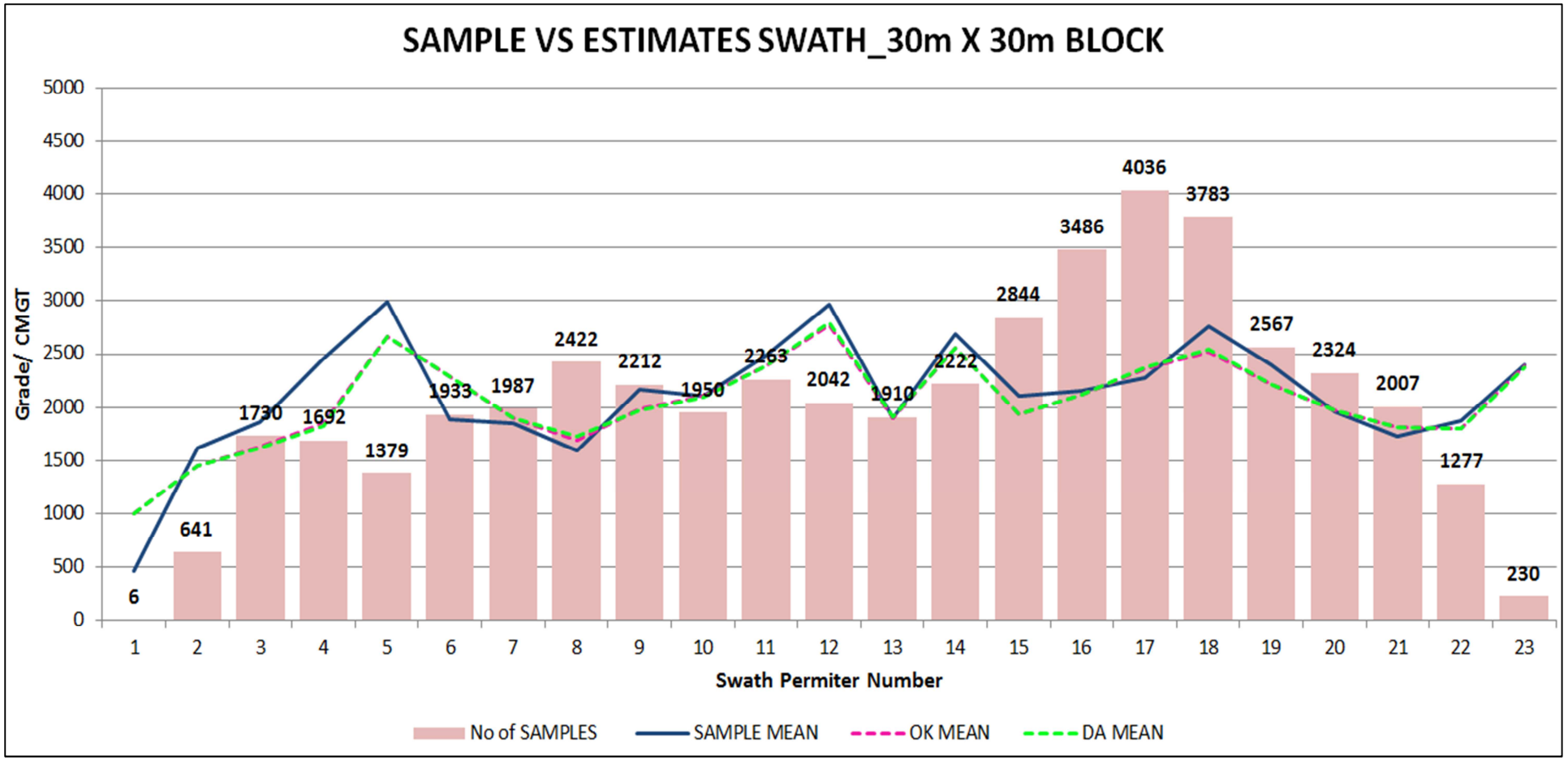


Figure 4.3: OK estimates and DA estimates vs. sample data swath plots.

Observations made from Figure 4.3 show that there is an expected degree of smoothing evident in the grade estimates for both methods of interpolation. Generally there is a relatively good correspondence between the swath mean composite grades and the respective block model estimates. The OK and DA estimates graphs tend to follow the sample mean grade values quite well. However, there is an under estimation as shown in swaths 3, 4, 5 and 6 in Figure 4.3. It is observed that there are a low number of samples in these areas. This can be explained as follows; the high value samples were not used to inform the estimates in the unsampled areas during the interpolation process because of the low range values and search orientation of the variogram model used during the process.

It can also be observed that using the swath plots to determine which one of the estimation methods is better is quite difficult.

4.1.4 Scatter Plots

The estimates calculated from both the DA and OK interpolation processes were also validated against the sample data using scatter plots. The sample data was regularised where the average value of all samples within a grid cell of 30×30 m in X and Y directions were extracted and the mean statistic compared with corresponding block model estimate with similar block IJK values. The IJK field in a block model is an index value giving the location of the parent cell for each cell or subcell. All subcells in a parent cell have the same IJK value. Scatter plots showing the relationship between the block model estimates, and sample data are shown in Figure 4.4. The strength and direction of the linear relationship between them were determined and the equation of the line that best fits the data calculated.

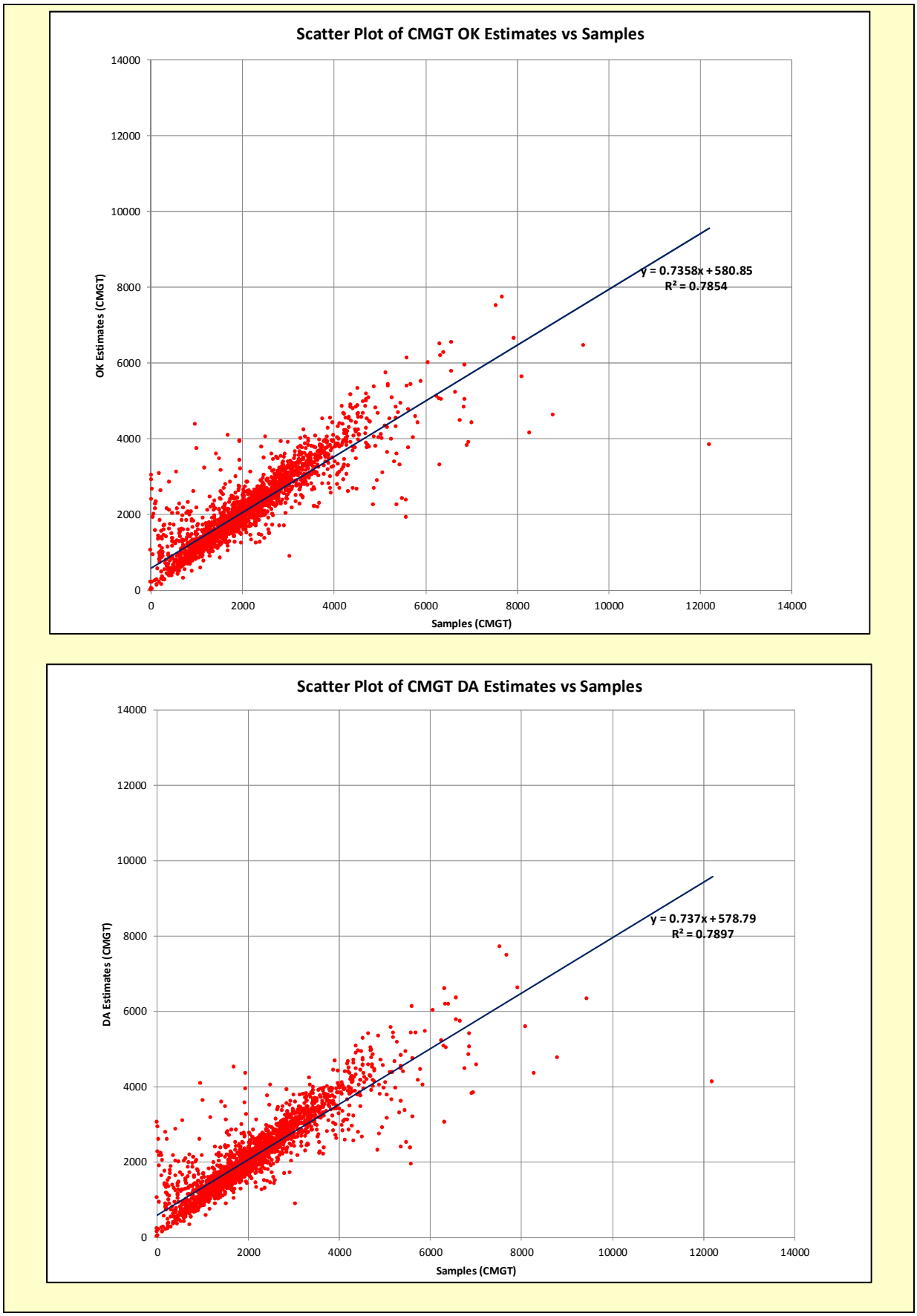


Figure 4.4: OK estimates (above) and DA estimates (below) vs. sample data scatter plots.

A linear regression fit was applied using the method of regression analysis. The author used the “linear regression” function in EXCEL software spreadsheet which provided a biased fit of the data between the sample data and estimates. The reason for this is that the software assumes one variable is independent (samples) and the other is dependent (estimates). The underlying assumption is that the independent variable is precisely known while the output is not.

Regression analysis generates an equation to describe the statistical relationship between one or more predictor variables and the response variable. The formula for the best-fitting line (or regression line) is $y = a + bx$, where x is the slope of the line and “a” is the y-intercept. The y-intercept is the value on the y-axis where the line crosses. The regression lines for the graphs in Figure 4.4 were calculated and are as follows;

$$y = 0.736x + 580.85$$ ……for sample data vs OK estimates

$$y = 0.737x + 578.79$$ ……for sample data vs DA estimates

Where x is the sample data and y is the estimate.

When x=0, y = 580.85 for OK estimates and 578.79 for DA estimates. However, the intercept is not considered in evaluating the performance of the estimator.

The slope for both methods of interpolation is positive and is 0.736 and 0.737 for OK and DA estimates respectively. This indicates that the estimates calculated from both types of interpolation processes are similar, implying that when the sample grades increase by a factor of 1, the value of the corresponding estimates increases by a factor of 0.74 and a larger constant has to be added ~ 580 cmg/t for both OK and DA estimates.

The correlation coefficient, R measures the precision, strength and direction of a linear relationship between two quantitative variables on a scatterplot. It’s a statistical measure of

how close the data are to the fitted regression line. It is also known as the coefficient of determination. It is the percentage of the response variable variation that is explained by a linear model. The value of R is always between +1 and –1. The following observations are made from Figure 4.4:

- There is a positive and strong linear relationship between the sample composites and estimates produced by both interpolation methods;
- The relationship between samples and DA estimates is much stronger with a correlation coefficient value of 0.889 in comparison to an R value of 0.886 which is recorded for OK estimates. However the percentage difference between the 2 estimation techniques is quite small, 0.003.

Therefore, this indicates that the application of DA estimation technique in 2D over the 30×30 m block model does not have a significant impact on improving the quality of estimates.

5 Discussion

5.1 Domains for 30×30 m Blocks

Estimates were also calculated separately on the various domains discussed in Section 3.6. Although the estimates were calculated inside the hard boundaries of the various domains, these estimates were validated against the source data by calculating and comparing the global statistics of the source data for the entire dataset.

Table 5-1 shows the global and local statistics for comparison purposes.

Table 5-1: Domain global and local statistics for 30 m by 30 m blocks

Zone	Variable cm.g/t	NSamp	Min	Max	Mean	CoV	Diff with Sample cm.g/t
ALL	Sample	46 943	0	14 250.90	2 218.85	1.01	
	OK estimates	2 795	4.94	7 742.17	2 114.07	0.50	-4.72%
	DA estimates	2 795	4.94	7 701.44	2 116.57	0.50	-4.61%
4 and 5	OK estimates	2 795	4.53	8 194.09	2 110.70	0.50	-4.87%
	DA estimates	2 795	4.53	8 194.05	2 113.41	0.50	-4.75%
1, 2 and 3	OK estimates	2 795	14.80	7 772.52	2 120.27	0.50	-4.44%
	DA estimates	2 795	9.58	7 501.72	2 119.53	0.50	-4.48%

The following observations are made from Table 5-1:

- The average block estimate grades calculated using both estimation methods are slightly lower than the sample data meaning that there is an overall under estimation in the blocks. However, the percentage differences are less than 5 % and so are within acceptable levels;
- Comparing results from the estimates calculated within the domains; it is observed that splitting the data set into 3 zones, gives estimates with average mean values that are closer to the average sample mean for both methods of interpolation. The value of OK estimates in the domain with 3 zones, 2120.27 cm.g/t is closer to the average of the sample values, 2218.85 cm.g/t. However, the converse is true when comparing the results for the domain that was split into 2 zones.

Domaining the dataset into the 3 separate zones and calculating the estimates within these zones seems to actually “improve” the estimates to a certain extent when just global estimates are considered. This observation postulates that even more detailed domaining would again continue to improve estimates. Referring to the global means again shows that the OK interpolation technique yields better estimates than those calculated using the DA method of interpolation because the average estimate values are closer to the sample mean. However, just comparing the global statistics does not give a clear indication which one of the interpolation methods is better than the other and it also does not give a good indication on how the estimates are spread across the entire dataset because it is just a global average.

Grade distribution plots of the block estimates superimposed on the sample plots were generated for comparison purposes; see plots in Appendix C. The visual comparison plots also show that the colour coding of the sample data closely matches that of the block estimates for both methods of interpolation.

However, it is again difficult to determine which method of interpolation is better using these plots. Therefore, the OK and DA estimates calculated in the different domains were also validated against the sample composites using scatter plots and the same criteria explained in Section 4.1.4. Scatter plots showing this relationship are shown on Figure 5.1 and Figure 5.2. The strength and direction of the linear relationship between them were determined and the equation of the line that best fits the data calculated.

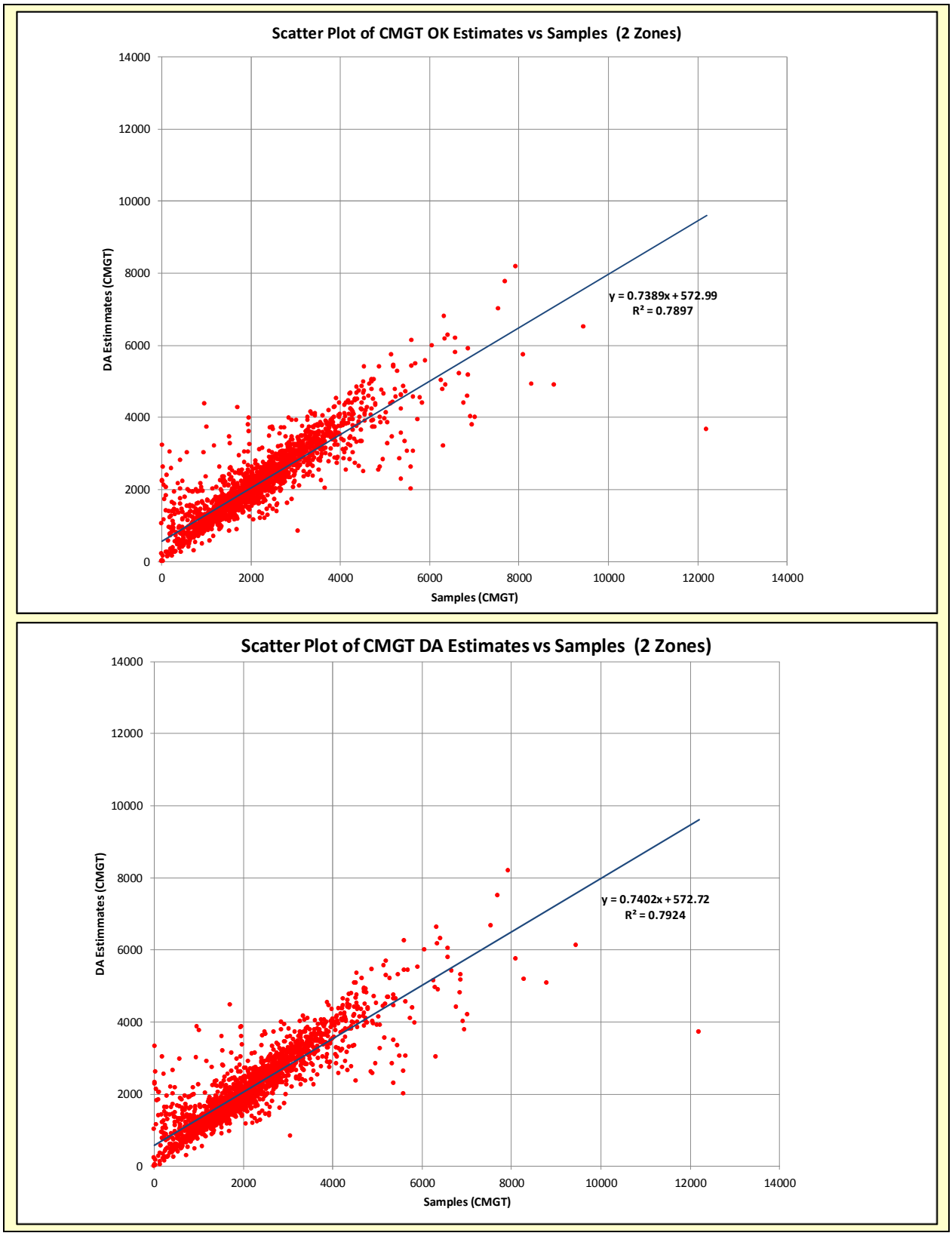


Figure 5.1: Scatter plots of OK estimates (above) vs sample data and DA estimates (below) vs. sample data – 2 zones

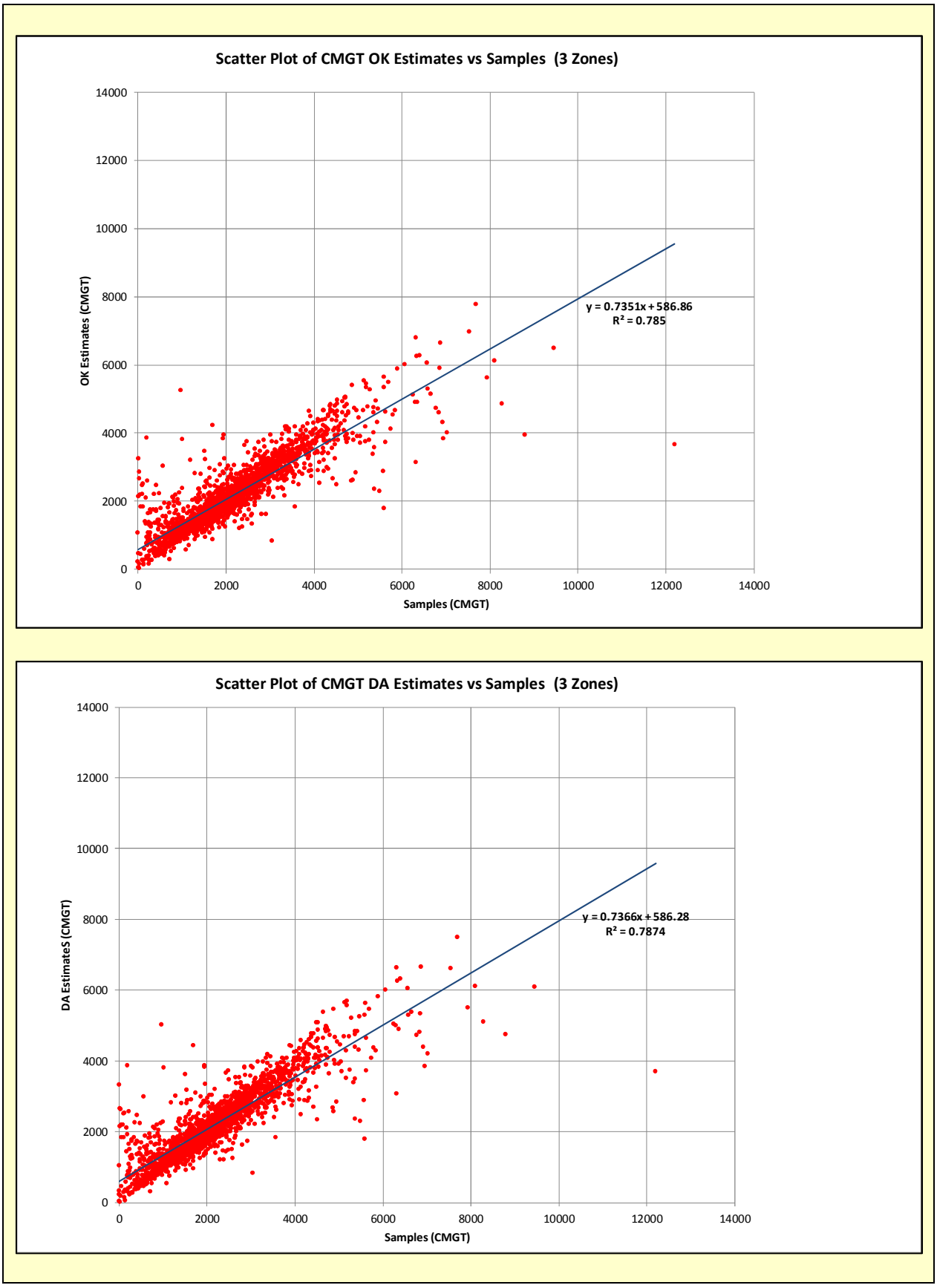


Figure 5.2: Scatter plots of OK estimates (above) vs sample data and DA estimates (below) vs. sample data– 3 zones.

Table 5-2 summarizes the linear regression equations as well as the correlation coefficients determined for the different estimation methods on the domains.

Table 5-2: Line of regression equations and correlation coefficients on domains

Block	Zones	Est Method	R^2	R	Regression equation
30 × 30	All	DA	0.790	0.889	y = 0.7370x + 578.79
30 × 30	All	OK	0.785	0.886	y = 0.7358x + 580.85
30 × 30	4 and 5	DA	0.792	0.890	y = 0.7402x + 572.72
30 × 30	4 and 5	OK	0.790	0.889	y = 0.7389x + 572.99
30 × 30	1, 2 and 3	DA	0.787	0.887	y = 0.7366x + 586.28
30 × 30	1, 2 and 3	OK	0.785	0.886	y = 0.7351x + 586.86

*where zone = All means no domaining on data set

The following is deduced from Table 5-2:

SLOR and R for both types of estimation techniques and for all the domains are all similar with a value of 0.74 and 0.89 respectively. This means that domaining the data set has no effect on improving the quality of estimates. The results therefore help in deducing that the domaining of this data set will not help in improving the quality of estimates calculated by both interpolation techniques.

5.2 Optimal DA Resolution

To determine the optimal resolution that DA will provide an improvement in grade estimates, DA and OK methods of estimation were further computed onto the following block sizes; 60×60 m, 120×120 m and 240×240 m grids and the resulting estimates compared.

First and foremost, QKNA exercises were done on the different block sizes to determine the maximum and minimum number of samples to be used during the estimation process. Sample number optimization as well as a sensitivity analysis of discretization on covariance for the same variogram and block size (block discretization number optimization) was done. These plots are shown in Appendix D and E. Observations from the discretization plots in Appendix E also indicate that the variances for results do not show significant continuing

downward trend beyond about 6×6 m for all bock sizes. Hence, a discretization value of 5×5 m was used for all block sizes. The following is deduced from Appendix D: In all the block sizes tested, there is a step change in the quality indicators assessed here (SLOR and KE). Although using a higher number of samples would result in a higher quality estimate, the estimation variance decreases, indicating more smoothing of the estimates. The minimum and maximum number of samples determined from these graphs is tabulated in Table 5-3.

For all estimates, a quadrant search was carried out with a discretization pattern of 5 (X) by 5 (Y) by 1 (Z) points used. The final search neighbourhood parameters used in the estimation process including the minimum and maximum number of samples and based on the semi-variogram models in Section 3.3.3 and 3.6.2 are summarised in Table 5-3.

Table 5-3: Search ellipsoid orientation, search range parameters and minimum and maximum number of samples used during estimation process for each zone and block size

Block size	Zone	Rotation (Degrees Clockwise)	Search Range (m)		Number of samples	
			U	V	Min	Max
60 × 60	ALL	50	45	45	8	40
	1	50	30	20	8	40
	2	50	50	65	8	40
	3	50	60	65	8	40
	4	50	30	30	8	40
	5	50	60	65	8	40
120 × 120	ALL	50	45	45	40	80
	1	50	30	20	40	80
	2	50	50	65	40	80
	3	50	60	65	40	80
	4	50	30	30	40	80
	5	50	60	65	40	80
240 × 240	ALL	50	45	45	40	100
	1	50	30	20	40	100
	2	50	50	65	40	100
	3	50	60	65	40	100
	4	50	30	30	40	100
	5	50	60	65	40	100

Interpolations on the various block sizes were then carried out using both DA and OK estimation techniques and the resultant estimates were validated. Grade distribution validation plots for different block sizes were generated for comparison purposes; see plots in Appendix F. The visual comparison plots do not clearly show the relationship between the sample data and block estimates for both types of estimation techniques.

Scatter plots to determine the kind of relationship between sample data and DA estimates as well as the relationship between sample data and OK estimates were generated, see Appendix G. The sample composites were regularised where the average value of all samples within a grid cell of the 60×60 m, 120×120 m and 240×240 m in X and Y directions were extracted and the mean statistic compared with corresponding block estimate with similar IJK value. The strength and direction of the linear relationship between them were determined and the equation of the line that best fits the data calculated and tabulated.

Table 5-4 summarizes the line of regression equations as well as the correlation coefficients determined by the different types of estimation methods on various block sizes and on the domains as well.

A graph plotting the SLOR and R against the un-domained estimation blocks for the two methods of interpolation was generated; see Figure 5.3 and Figure 5.4 respectively.

Table 5-4: Line of regression equations and correlation coefficients on blocks and domains

Block	Zones	Est Method	R^2	R	Regression equation
30 × 30	All	DA	0.790	0.889	y = 0.7370x + 578.79
30 × 30	All	Normal OK	0.785	0.886	y = 0.7358x + 580.85
30 × 30	4 and 5	DA	0.792	0.890	y = 0.7402x + 572.72
30 × 30	4 and 5	Normal OK	0.790	0.889	y = 0.7389x + 572.99
30 × 30	1, 2 and 3	DA	0.787	0.887	y = 0.7366x + 586.28
30 × 30	1, 2 and 3	Normal OK	0.785	0.886	y = 0.7351x + 586.86
60 × 60	All	DA	0.797	0.893	y = 0.8760x + 284.68
60 × 60	All	Normal OK	0.797	0.893	y = 0.8793x + 280.83
60 × 60	4 and 5	DA	0.763	0.873	y = 0.8437x + 339.09
60 × 60	4 and 5	Normal OK	0.770	0.877	y = 0.8574x + 320.45
60 × 60	1, 2 and 3	DA	0.760	0.872	y = 0.8315x + 364.41
60 × 60	1, 2 and 3	Normal OK	0.765	0.875	y = 0.8431x + 349.78
120 × 120	All	DA	0.615	0.784	y = 0.8734x + 209.99
120 × 120	All	Normal OK	0.612	0.782	y = 0.8709x + 214.42
120 × 120	4 and 5	DA	0.601	0.775	y = 0.8717x + 246.09
120 × 120	4 and 5	Normal OK	0.637	0.798	y = 0.8953x + 185.59
120 × 120	1, 2 and 3	DA	0.628	0.793	y = 0.9013x + 218.74
120 × 120	1, 2 and 3	Normal OK	0.649	0.805	y = 0.9273x + 155.46
240 × 240	All	DA	0.260	0.510	y = 0.9777x + 65.710
240 × 240	All	Normal OK	0.295	0.543	y = 1.0138x - 47.016
240 × 240	4 and 5	DA	0.413	0.642	y = 1.1577x - 320.95
240 × 240	4 and 5	Normal OK	0.413	0.642	y = 1.1577x - 320.95
240 × 240	1, 2 and 3	DA	0.307	0.554	y = 0.9169x + 181.59
240 × 240	1, 2 and 3	Normal OK	0.275	0.524	y = 0.8438x + 309.60

*where “zone = All” means no domaining on data set

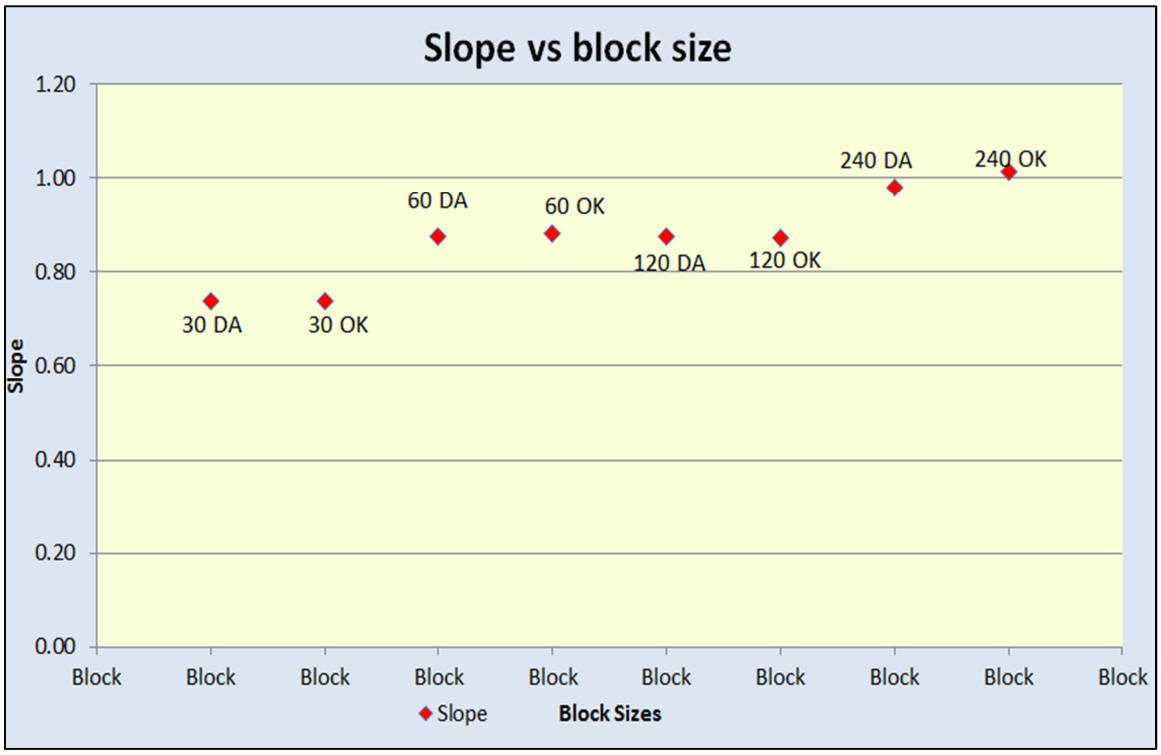


Figure 5.3: Slope of regression against block size.

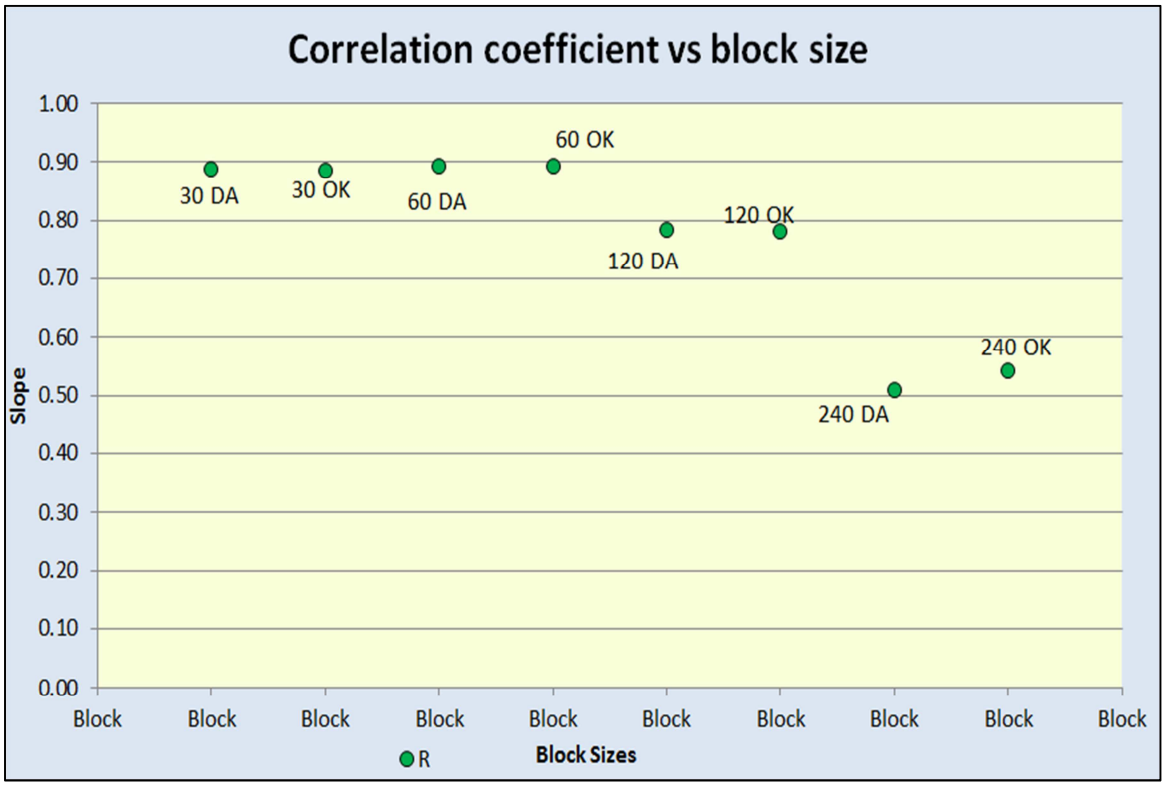


Figure 5.4: Correlation coefficient against block size.

The following can be deduced from Table 5-4, Figure 5.3 and Figure 5.4:

- SLOR line increases significantly with increasing block size implying that increasing the block size increases the quality of estimates for both interpolation methods. However this can be attributed to over smoothing due to biases. The SLOR for the 240×240 m grid block is greater than one implying that there is a conditional bias of the estimates. This is fully explained in the Figure 5.8 and sections below;
- There is a small to insignificant difference between the SLOR for both types of interpolation techniques executed on the same block model size implying that application of DA interpolation technique on the dataset supplied either has minimal to no effect on improving the quality of estimates. This can be explained by the fact that the study was executed in a 2D environment where the dip component of DA has been ignored, “the z coordinate is a constant value throughout the database”. This is a limitation for this exercise and further explains why the results between OK and DA estimation techniques are very similar;
- Generally, R reduces significantly with increasing block size meaning that the relationship between the samples and the estimates weakens as the block sizes increase with the exception of 240×240 m grid block whose R value has slightly increased for both interpolation techniques when compared with R for 120×120 m grid block. R is just a measure of precision and is not an accurate measure of how close the estimates are to the sample grades. The value of SLOR is a better measure of accuracy and is thus a better tool to use to compare the methods of interpolation.
- The traditional estimation technique, which is OK, becomes a better method of estimation with increasing block size in comparison to the DA method of estimation. This can also be explained pictorially in Figure 5.5 which shows that as the block sizes increase, the anisotropy tends toward the regional anisotropy and localized directions become misleading. In other words, as the block sizes approach the level of resolution of DA, the benefits thereof are reduced and as observed may even

become a confounding factor in estimation leading to less accurate and smoothing of estimates;

- Domaining the dataset has minimal effect on improving the quality of estimates calculated by both estimation techniques for the smaller block sizes. However, as the block sizes increase, domaining the dataset tends to improve the quality of the estimates;

These observations can be attributed from the block size being significantly larger than the sample spacing and also because all blocks are well informed. There are a lot of samples that are available for estimation within a large block and yet just a limited number of them are used to estimate the grade in the block during the interpolation process. Hence there is also limitation on no extrapolation as no samples lying outside the block being estimated are used during the interpolation process.

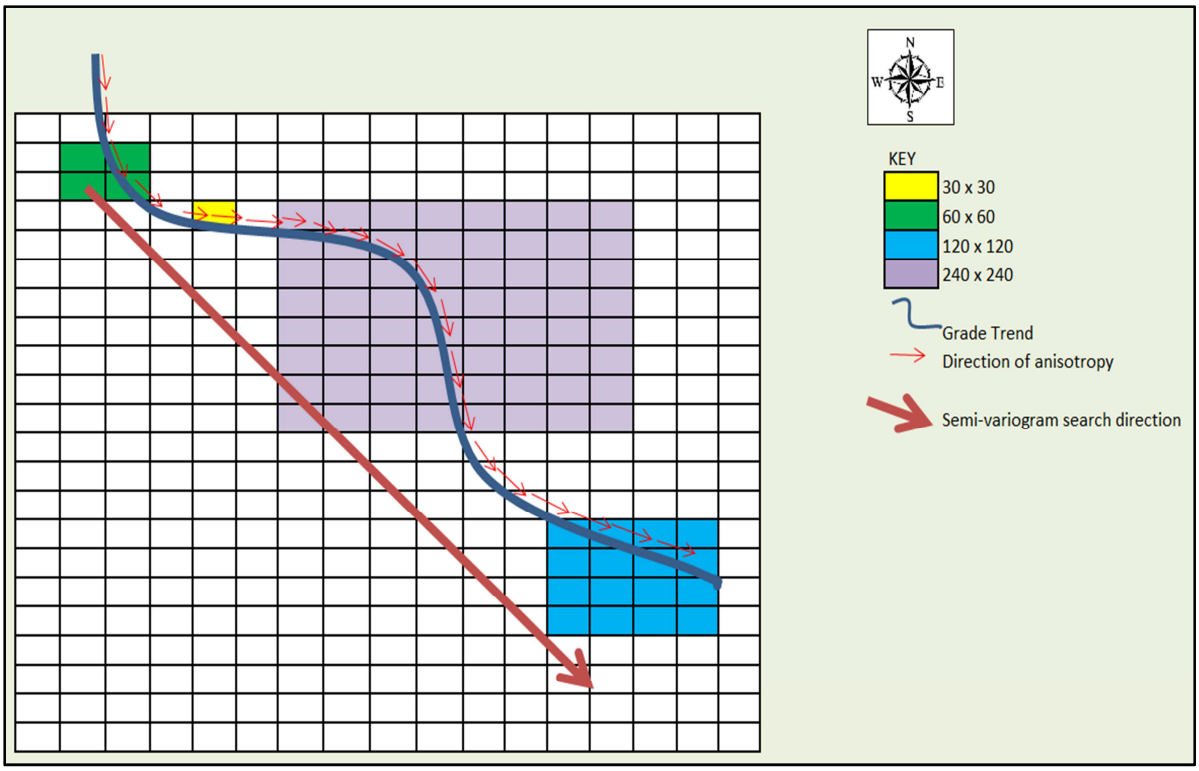


Figure 5.5: Diagrammatic representation showing relationship between block sizes, "resolution of Dynamic Anisotropy" and localised directions.

A graph plotting R against the domained estimation blocks (with increasing block size) for the two types of estimation methods from Table 5-4 was also generated; see Figure 5.6.

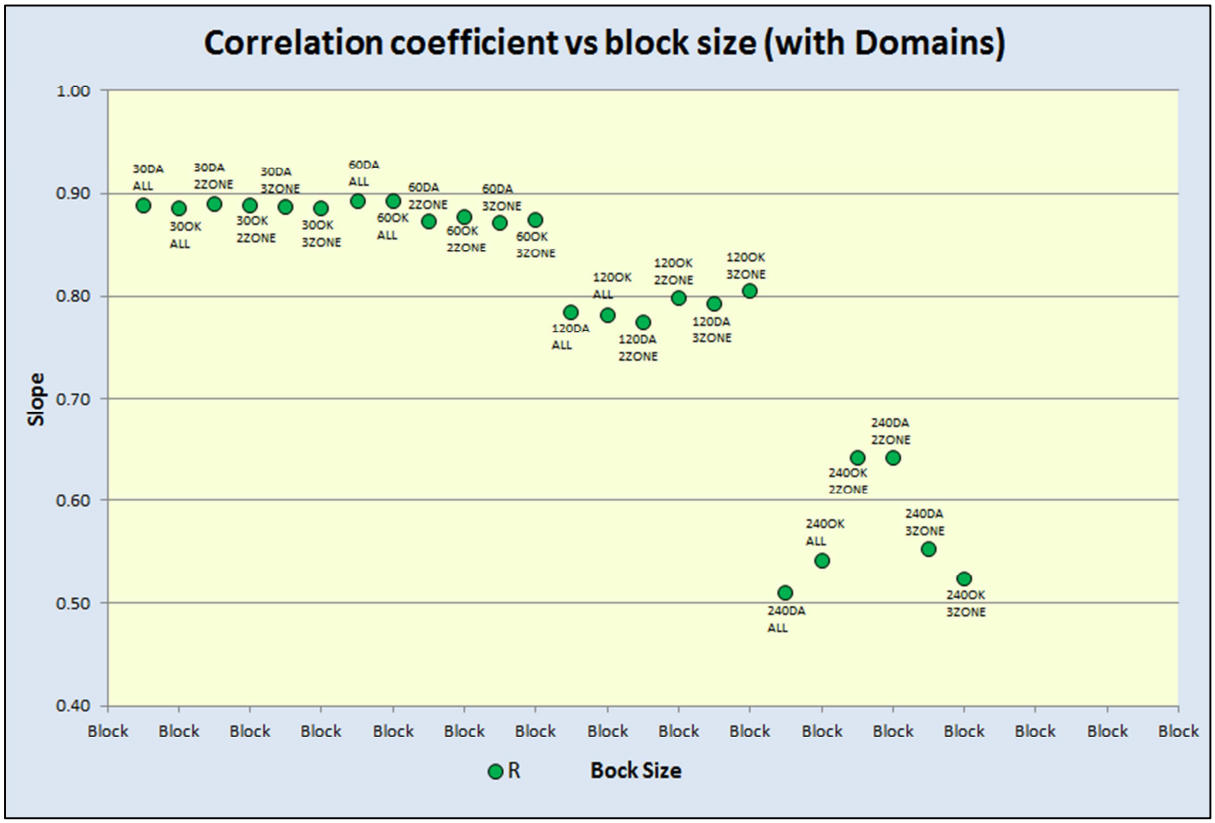


Figure 5.6: Correlation coefficient against block size (domained).

It is observed from Figure 5.6, that the value of R gradually decreases with increasing block size. This implies that the quality of estimates suffers as the size of the estimation blocks gets bigger. This can be attributed to the fact that smaller block sizes have the effect of improving the quality of the estimates because of a denser sample data grid where samples tend to share weights.

A graph plotting the SLOR against the domained estimation blocks (with increasing block size) for the two types of estimation methods was also generated; see Figure 5.7.

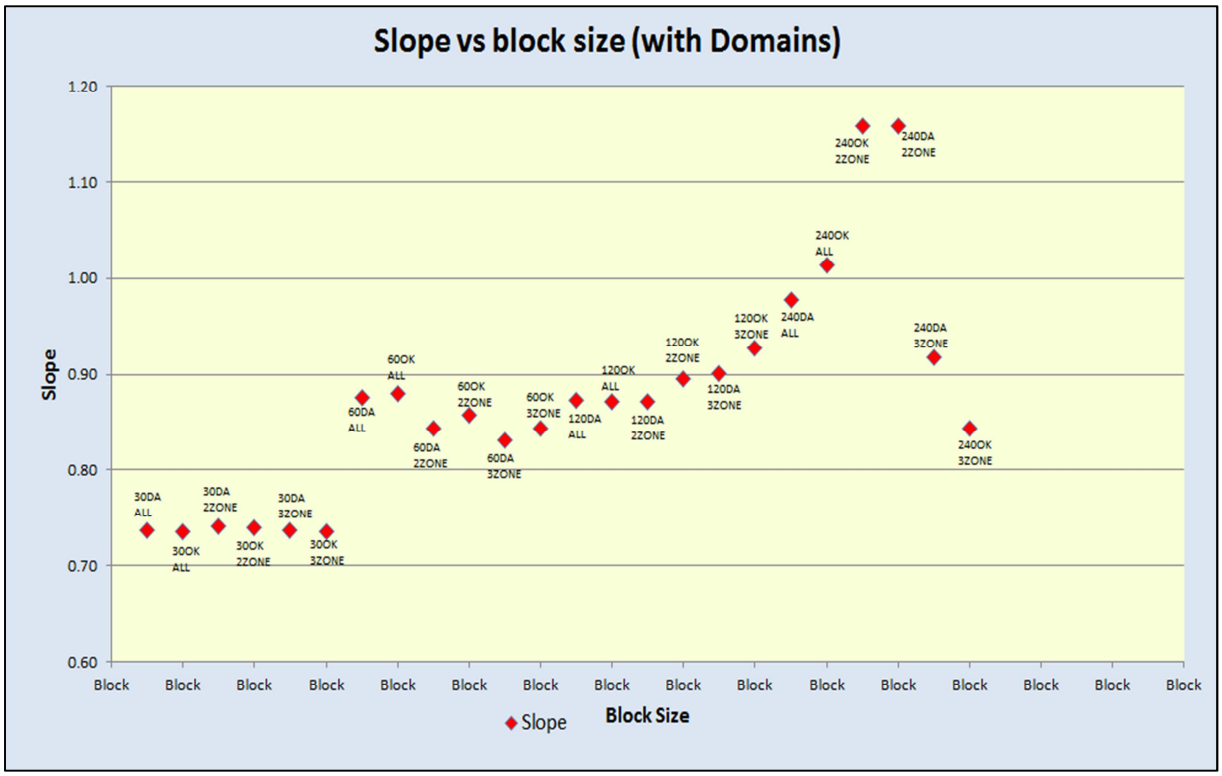


Figure 5.7: Slope of regression against block size (domained).

It is observed from Figure 5.7 that domaining the dataset has improved the quality of the estimates to a certain degree as the block sizes increase although this is not the case for the smaller 30×30 m block size where domaining had no effect on improving the estimates. However, domaining the dataset on 240×240 m block has unfavourable effects on the SLOR. It has actually decreased the quality of estimates determined from both methods of interpolation.

It is also observed from Figure 5.7 that generally, there is an upward trend in the SLOR as the size of estimation blocks increases for both types of interpolation methods which translate to an increase in the quality of estimates. However, this is not the case as the increase in SLOR is due to over smoothing due to biases as previously explained. This can be further explained referring to Figure 5.8. The definition of the SLOR of actual (x) on estimated block values (y), is:

$$\text{slope} = \text{r}\left(\frac{\sigma_x}{\sigma_y}\right)$$

If the slope is to be unity for unbiasedness and "*r*" is less than unity because estimates are never perfect, $\left(\frac{\sigma_x}{\sigma_y}\right)$ must be larger than unity, i.e. the standard deviation (or variance) of the real block values must be larger than that of the estimated values. The gap between these two variances, i.e. the smoothing effect can, therefore, only be reduced by increasing the correlation between estimates and real values, i.e. by improving the efficiency of the valuation technique or by providing more data (Krige, 1996), see Figure 5.8.

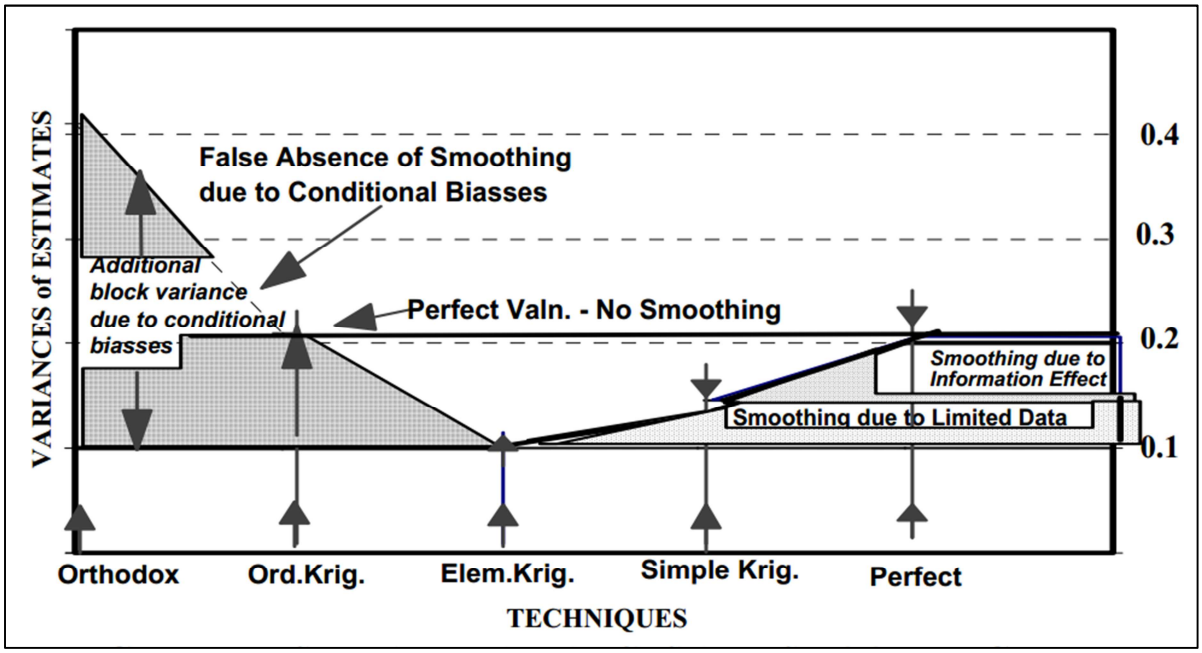


Figure 5.8: Estimate variances and smoothing (Source: Krige, 1996).

6 Conclusions

Representing grade continuity in flexured, undulating and folded structures can be a challenge to correct grade estimation. However, the advancement of technology has helped in utilizing a dynamic search ellipse for grade estimation that allows for a more accurate representation of the grade within such structures. DA is the method of estimation that is recommended to be used in such cases. This method determines the data that will be used during the estimation of a single block. It allows the search ellipsoid to change block by block throughout the model to account for the local variations by using the dip and dip direction interpolated into each cell within a block model to rotate it. DA will then follow local variations along the strike of the deposit thus leading to improved sample selection, kriging efficiency during the grade estimation process. This method therefore, provides a fast, relatively simple and mostly effective option to use for grade estimation for undulating and folded structures.

The author was approached by Driefontein Gold Mine to study the application of DA and then derive an optimised resolution for an area at the mine that lies on a slightly undulating gently folded limb of a syncline. The mine currently uses the OK method of estimation on a 30×30 m block size to calculate their Mineral Resource estimates. DA and OK interpolation techniques were applied to the data set extracted from the case study area and the resultant estimates calculated from both interpolation methods were compared. Scatter plots showing the relationship between the block model estimates and sample data were plotted. The strength and direction of the linear relationship amongst them were determined and the equation of the line that best fits the data calculated. Analysis of the results revealed that the SLOR and R for both estimation techniques were positive and almost similar. The estimates calculated from both estimation techniques correspond well with the sample data used during the estimation processes.

To determine the optimal DA resolution that will provide an improvement in grade estimates, the estimation process using DA and OK estimation techniques was executed on block model range of different resolutions; 30×30 m, 60×60 m, 120×120 m and 240×240 m grids. Analysis of the results using the equation of the line of best fit of the scatter plot revealed that the SLOR increased significantly whilst R decreased with increasing block size. This implies that although increasing the block size improved the quality of estimates for both estimation techniques, because R reduced significantly with increasing block size, the relationship between the sample data and the estimates weakens as the block sizes increases. However this can be attributed to over smoothing due to biases. The SLOR for the 240×240 m grid block is greater than one implying that there is a conditional bias of the estimates. The block sizes are significantly larger than the sample spacing and also all blocks are well informed and therefore there is also limitation on no extrapolation as no samples lying outside the block being estimated were used during the interpolation process for larger blocks.

The dataset was domained using the mineralisation trend as a guide to find out whether domaining would improve the quality of resource estimates. Analysis of the results revealed that domaining the dataset has minimal effect on improving the quality of estimates calculated by both estimation techniques for the smaller block sizes. However, as the block size increases, domaining the dataset tends to improve the quality of the estimates.

This study has shown that generally, the application of DA estimation technique during the estimation process on two dimensional 30×30 m block model generated from a slightly undulating area at Driefontein mine does not have a significant impact on improving the quality of estimates. This can be attributed to the fact that smaller block sizes have the effect of improving the quality of the estimates, probably because samples are closer to one

another and therefore tend to share weights. However, this is not the case with larger block sizes.

The results also reveal that the OK estimation technique becomes a better method of interpolation with increasing block size in comparison to the application of DA estimation technique during the estimation process. In other words, as the block sizes approach the level of resolution of DA the benefits thereof are reduced and as observed may even become a confounding factor in estimation leading to less accurate estimates. This is due to the fact that as the block sizes increase, the anisotropy tends toward the regional anisotropy and localized directions become misleading. However, the study has also shown that doubling the grid size from 30×30 m to 60×60 m improves the quality of estimates significantly.

The results of this study can also be explained by the fact that it was executed in a two-dimensional environment where the dip component of DA has been ignored, “the z coordinate is a constant value throughout the entire database”. This is a limitation for this exercise and further explains why the results between OK and DA estimation techniques are very similar.

Therefore the author recommends that for future studies, Driefontein Gold Mine considers implementing a similar exercise based on three dimensional and widely spaced data or regularising the point data into data blocks and using simple or related kriging techniques. This study was also done in a mined out area and therefore for future studies, the author recommends that further investigations be executed in an exploration area with low sample data density and test whether the DA method of interpolation will improve the quality of estimates.

7 References

Ainslie, L. C. 1981: Geology and Ore Reserve Estimation of the Witwatersrand-Type Gold Deposits with Specific Reference to the Welkom Goldfield, Dissertation submitted in partial fulfilment of the requirements for the degree of Master of Science (Mineral Exploration) at *Rhodes University*.

Armstrong, M. (1998) Basic Linear Geostatistics, Berlin: Springer-Verlag.

Baafi, E. Y., and Schofield, N. A. 1997: Geostatistics Wollongong "96 Volume 2, Quantitative Geology and Geostatistics, *Kluwer Academic*, pp. 869-870.

Beach, A., Smith, R. 2007: Structural geometry and development of the Witwatersrand Basin, South Africa. In: Ries, A.C., Butler, R.W.H., Graham, R.H. (Eds.), Deformation of the Continental Crust: The Legacy of Mike Coward: *Geological Society, London*, Special Publications, pp. 272, 533-542.

Bertoli, O. , Job, M. , Vann, J. and Dunham, S. 2003: Two-Dimensional Geostatistical Methods — Theory, Practice and a Case Study from the 1A Shoot Nickel Deposit, Leinster, Western Australia, *5th International Mining Geology Conference*, Bendigo, Vic, 17 - 19 November 2003.

Behrens, J. T. 1997: Principles and Procedures of Exploratory Data Analysis, Psychological Methods Copyright 1997 by the *American Psychological Association*, Inc. Vol. 2, No. 2, pp. 131-160.

Blackney, P. C. J and Glacken, I.M. (1998) A practitioner's implementation of indicator kriging. Submitted for Publication in *The Geostatistical Association of Australasia* "Beyond Ordinary Kriging" Seminar Perth, Western Australia.

Blais, R. A. and Carlier, P. A. 1968: Applications of Geostatistics in Ore Evaluation, *Canadian Institute of Mining and Metallurgy*, Ore Research Estimation and Grade Control. Special Volume, Montreal, Quebec, pp. 42.

Blais, R. A. and David, M. 1972: Geostatistical Ore Reserve Estimation, *South African Institute of Mining and Metallurgy*, APCOM 72. *Application of Computer Methods in the Mineral Industry: Proceedings of the Tenth International Symposium, Johannesburg, South Africa, 10-14 April 1972*, SAIMM, 1972. pp. 27-33.

Bohling, G. 2005: Introduction to Geostatistics and Variogram Analysis, C&PE 940, *Kansas Geological Survey*, pp. 19-20.

http://people.ku.edu/~gbohling/cpe940/Variograms.pdf

Braun, R. 1987: Ore Reserve Estimation at Western Deep Levels Gold Mine, 20th *International Symposium on the Application of Computers and Mathematics in the Mineral Industries, Volume 3*: Geostatistics. Johannesburg, SAIMM, pp. 233-243.

Champigny, N. and Armstrong, M. 1990: Effect of high grades on the geostatistical estimation of gold deposits, *AIME Transactions Volume 288*, SME.

Chilès, J-P and Delfiner, P. (2012) Geostatistics-Modelling Spatial Uncertainty, *John Wiley & Sons, New York*, pp. 489.

Clarke, Isobel. 1979: Practical Geostatistics *Applied Science Publishers Ltd*, London. pp. 118-123.

Coetzee, C. B. 1976: The Mineral resources of South Africa. *Department of Mines, Geological Survey, Pretoria*, pp. 462.

Coward, M. P., Spencer, R.M., Spencer, C.E. 1995: Development of the Witwatersrand Basin, South Africa. In: Coward, M.P., Ries, A.C. (Eds.), Early Precambrian Processes: *Geological Society, London*, Special Publications, 95, pp. 243-269.

Dankert, B. T., Hein, K.A.A. 2010: Evaluating the structural character and tectonic history of the Witwatersrand Basin. *Precambrian Research*, pp. 177, 1-22.

David, M. and Blais, R. A. 1972: Geostatistical Ore Reserve Estimation, *Application of Computer Methods in the Mineral Industry: Proceedings of the Tenth International Symposium*

Deutsch, C. V. 2011: Unpublished document from the Citation Course in Geostatistics, under the auspices of the *Geostatistical Association of Southern Africa, Johannesburg*.

Deutsch, C. V. and Journel, A.G. 1998: Geostatistical Software Library and User's Guide, Second Edition, Oxford University Press, *Applied Geostatistics Series*, pp. 13- 131.

De Wit, M. J., Roering, R., Hart, R. J., Armstrong, R. A., De Ronde, C. E. J., Green, R. W. E., Tredoux, M., Peberdy, E. and Hart, R. A. 1992: Formation of an Archaean continent. Nature, pp. 357: 553-562.

Driefontein Gold Mine Mineral Resource Management Team. 2012: Competent Person's Report on The Material Assets of Kloof Gold Mine, Prepared for Gold Fields Limited. pp. 16-24.

Dohm, C.E. 2010: Course Notes: Geostatistical Methods in Mineral Resource Evaluation, GDE, University of the Witwatersrand.

Dowd, P. A. 1992: Basic Geostatistics for the Mining Industry, University of Leeds

Dowd, P. A. 2000: The future for geostatistics: An Academic Perspective, In: Kleingeld, W. J., and Krige, D. G. eds, *Geostats 2000 Cape Town: Proceedings of the Sixth International Geostatistics Congress held in Cape Town, South Af*rica, vol. 1, Geostatistical Association of Southern Africa, pp. 12-25.

Engineering statistics handbook:
http://www.itl.nist.gov/div898/handbook/eda/section3/eda3669.htm

Frimmel, H E, Groves, D I, Kirk, J, Ruiz, J, Chesley, J and Minter, W E L, 2005. The formation and preservation of the Witwatersrand goldfields, the world's largest gold province, *Economic Geology 100th Anniversary Volume*, pp. 769-797 (Society of Economic Geology).

Gibson, M. A. S., Jolley, S. J., Barnicoat, A.C. 2000: Interpretation of the Western Ultra Deep Levels 3D seismic survey. The Leading Edge, pp. 19, 730–735.

Glacken, I. M. and Gray, D. B. 2011: Resource estimation at the Kylylahti copper-gold-cobalt mine, Ountokupmpu, Finland, 2011, *Fennoscandian Exploration and Mining Conference*, Optiro, Perth.
http://fem.lappi.fi/c/document_library/get_file?folderId=506958&name=DLFE-10161.pdf

http://www.goldfacts.org/en/economic_impact/countries/south_africa/)

Goovaerts, P. 1997: Geostatistics for Natural Resources Evaluation. *Oxford University Press, Applied Geostatistics Series*, pp. 98, 483.

Grant, C. L. B and Morrison, S. R. 2012: RIO Grande Cu-Au-Ag Project,
Northwest Argentina, Regulus Resources Inc., Teta Tech Wardrop, Toronto, Ontario

http://www.forbes.com/sites/kitconews/2014/06/20/worlds-largest-gold-producing-countries-south-africa/

Gringarten, E. and Deutsh, V. 2001: Variogram Interpretation and Modeling, Mathematical Geology, Vol. 33, No. 4, *International Association for Mathematical Geology*

Horscroft, F. D. M., Mossman, D. J., Reimer, T. O. and Hennigh, O. 2011: Witwatersrand Metallogenesis: The Case For (Modified) Syngenesis, Microbial Mats in Siliciclastic Depositional Systems Through Time, SEPM (Society for Sedimentary Geology), *SEPM Special Publication* No. 101, ISBN 978-1-56576-314-2, pp. 75–95.

Holding, S. W. 2000: Practical Geostatistics: Modeling and Spatial Analysis. Manual, *Springer-Verlag Berlin Heidelberg New York*, pp. 1-14.

Isaaks, E. H. R. and Srivastava, M. 1989: Applied Geostatistics, *Oxford University Press*, Inc. pp. 150-170, 186, 339-340, 529.

Jacek, M. and Monika, W-B. 2012: Variability Anisotropy of Mineral Deposits Parameters and Its Impact on Resources Estimation – A Geostatistical Approach, *Gospodarka Surowcami Mineralnymi / Mineral Resources Management, 2012,* vol. 28, issue 4, pp. 113-135.

Jolley, S. J., Stuart, G.W., Freeman, S. R., Knipe, R.J., Kershaw, D., McAllister, E., Barnicoat, A. C., Tucker, R. F. 2007: Progressive deformation of a Late Orogenic Thrust System, from duplex development to extensional reactivation and disruption: Witwatersrand Basin, South Africa. *Geological Society, London,* Special Publications, 272, pp. 543-569.

Journel, A. G. (1983) Non-parametric estimation of spatial distributions, *Mathematical Geology 15*, pp. 445-468.

Journel, A. G. and Huibregts, C. H. J. 1978: Mining Geostatistics, *Academic Press Inc. (London) Ltd*. pp. 175-183, 262-274.

Journel G. 1974: Geostatistics for Conditional Simulations of ore Bodies, *Econ, Geol 69* (5). pp. 673-687.

KEFI Minerals Plc. 2014: Upgraded JORC Resource at Tulu Kapi Gold Deposit in Ethiopia, http://www.londonstockexchange.com/exchange/news/market-news/market-news-detail.html?announcementId=11889758

Kloof-Driefontein Complex (KDC) – Technical Short Form Report, 2010

Kloof Gold Mine Mineral Resource Management Team. 2012: Competent Person's Report on The Material Assets of Kloof Gold Mine, Prepared for Gold Fields Limited. pp. 20-31.

Krige, D. G. 1976: Some basic considerations in the application of geostatistics to the valuation of ore in South African gold mines, *Journal of the South African Institute of Mining and Metallurgy, Johannesburg*.

Krige, D. G. 1981: Lognormal-de Wijsan Geostatistics for Ore Evaluation, *South African Institute of Mining and Metallurgy, Johannesburg*.

Krige, D. G. 1994: A statistical approach to some basic mine valuation problems on the Witwatersrand, *Journal of the Chemical, Metallurgical and Mining Society of South Africa*, pp. 95-111.

Krige, D. G. 1996: A basic perspective on the roles of classical statistics, data search routines, conditional biases and information and smoothing effects in ore block valuations, *Conference on Mining Geostatistics, Kruger National Park, Sept.1994, UNISA Publication 1996.*

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., *International Geostatistics Congress, Wollongong, Australia, 1996*.

Krige, D. G. 2000: Half a century of geostatistics from a South African perspective, In: Kleingeld, W. J., and Krige, D. G. eds, Geostats 2000 Cape Town: *Proceedings of the sixth International Geostatistics Congress held in Cape Town, South Africa, vol. 1, Geostatistical Association of Southern Africa*, pp. 3-11.

Koglin, N., Frimmel, H. E., Lawrie, W. E. and Brátz, H. 2010. Trace element characteristics of different pyrite types in Mesoarchean to Paleoproterozoic placer deposits, Mineral Deposita, pp. 259-280.

Landmark, G. J. 1990: Vaal Reefs Mine, South Africa, and the Role of the Geologist, Mine Geologists' Conference, 2 – 5 October 1990, Mount Isa, Queensland, Australia, *Australasian Institute of Mining and Metallurgy, Parkville, Victoria*: AUSIMM, 1990.

LeMay, Jr. N. E. 1998: Variogram Modeling and Estimation, *University of Colorado at Denver*, Master of Science in Applied Mathematics Thesis. pp. 21-24.

McCarthy, T S and Rubidge, B, 2005. . The Story of Earth & Life. A Southern African Perspective on a 4.6-Billion-Year Journey, Cape Town, Struik Publishers/Johnnic Publishing Group.

Magri, E. J. 1983: Improved Geostatistics for South African Gold Mines, *University of the Witwatersrand*, Ph.D. Thesis.

Manzi, M. S. D., Hein, K. A. A., King, N. and Durreheim, R. J. 2013: Neoarchaean tectonic history of the Witwatersrand Basin and Ventersdorp Supergroup: New constraints from high-resolution 3D seismic reflection data, School of Geosciences, *University of the Witwatersrand Johannesburg*, P Bag 3, WITS, 2050, Republic of South Africa.

Matheron, G. 1963: Principles of geostatistics, *Economic Geology*, vol. 58, pp. 1246-1266.

Matheron, G., Kliengeld, W. J. 1987: The Evolution of Geostatistics, APCOM 87. *Proceedings of the Twentieth International Symposium on the Application of Computers and Mathematics in the Mineral Industries. Volume 3: Geostatistics. Johannesburg*, SAIMM, 1987. pp. 9-12.

Miller, S. L. 1983: Geostatistical Evaluation of a Gold Ore Reserve System, *The University of South Africa*, MSc. Dissertation.

http://www.miningaustralia.com.au/news/the-world-s-ten-largest-gold-mines#comments

Minnitt, R. C. A., Hewitt, T. Lavery, B. and A Johnston, A. 2011: Mineral Resource Modelling and Reconciliation in Ultra Deep Underground Gold Mines, *Eighth International Mining Geology Conference / Queenstown, New Zealand*, 22 - 24 August 2011.

Morrison, R. S. and Brown, C. L. B. 2012: Technical Report, Rio Grande CU-AU-AG Project, Republic of Argentina Prepared for Regulus Resources Inc. dated January 19, 2012, Toronto, Ontario, pp. 84-94.

Nwaila, G. Becker, M. Ghorbani, Y. Pertersen, J. Reid, D. L. Bam, L. C. de Beer, F. C. and Franzidis, J-P. 2013: A Geometallurgical Study of the Witwatersrand Gold Ore at Carletonville, South Africa, *The Second AUSIMM International Geometallurgy Conference / Brisbane, Qld,* 30 September - 2 October 2013.

Paz-Ferreiro, P., Vázquez, E. V. and Vieira, R. S. 2010: Geostatistical analysis of a geochemical dataset, Bragantia vol.69 supl.0 Campinas.

Philips, G. N. 2013: Origin of Witwatersrand Gold – A Viable Source Region and Source Process for Gold, *World gold Conference / Brisbane, Qld*, 26 - 29 September 2013.

Pretorius, D. A. 1976a: Gold in the Proterozoic sediments of South Africa; systems, paradigms and models: in Wolf, K.H., (ed.), Handbook of stratabound and stratiform ore deposits, Vol. 7. *Elsevier Publishing Co., Amsterdam,* pp. 1-28.

Porto de Carvalho, J. R., Vieira, S. R., Magalhães, P. S. G and Vendrusculo, L. G. 2004: *Computational System for Geostatistical Analysis, Sci. Agric*. (Piracicaba, Braz.), v.61, n.1, pp. 100-107.
http://www.regulusresources.com/Portals/0/pdfs/20120119_RioGrande_43-101_Report.pdf

Robb, L. J. and Meyer, F. M. 1995: The Witwatersrand Basin, South Africa: Geological framework and mineralization processes, Ore Geology Reviews 10 (1995), Department of Geology, *University of the Witwatersrand*, Private Bag 3, Wits 2050, South Africa, pp. 67-94.

Roy, W. D. 2000: Using Geostatistics to Estimate the Resources of a Narrow Vein Gold Deposit, A Thesis Submitted to the Faculty of Engineering in Partial Fulfilment of the Requirements for the Degree of Master of Applied Science, *Department of Mining and Metallurgical Engineering, Dalhousie University – Daltech, Halifax*, Nova Scotia.

http://www.sibanyegold.co.za/

http://www.sec.gov/Archives/edgar/data/865492/000086549203000016/ascii_gbg20f2.htm

Sichel, H. S. 1947: An Experimental and Theoretical Investigation of Bias Error in Mine Sampling with Special Reference to Narrow Gold Reefs, Transactions Institute Mining Metallurgy, London, vol. 56, pp. 403-443.

Snowden, V. 1989: Estimating Recoverable Reserves – Developments in Applied Geostatistics, AIG Seminar, November, *Snowden Associates Newsletter.*

St George, J. D. 1992: Application of Geostatistics To Gold Deposits, Proceedings of the 26th Annual Conference 1992, *New Zealand Branch of the Australasian Institute of Mining and Metallurgy*.

http://www.theglobeandmail.com/report-on-business/international-news/global-exchange/south-africas-mining-industry-mired-in-disarray/article2035440/

Thurston, M. and Armstrong, M. 1987: The Application of Log Semi-variograms to the Kriging of Moderately Skew Raw Data, APCOM 87. Proceedings of the Twentieth

International Symposium on the Application of Computers and Mathematics in the Mineral Industries. Volume 3: *Geostatistics. Johannesburg*, SAIMM. pp. 53-64.

Van der Westhuizen, W. A., De Bruiyn, H., Meintjes, P.G. 1991: The Ventersdorp Supergroup: an overview. *Journal of African Earth Sciences*, pp. 13, 83–105.

Vann, J., Jackson, S and Bertolli, O. 2003: Quantitative Kriging neighbourhood analysis for the mining geologist - a description of the method with worked case examples. 5th International Mining Geology Conference, Melbourne. *The Australasian Institute of Mining and Metallurgy*, pp. 215 – 223.

Vermaakt, D. T., Chunnet, I.E., 1994. Tectono-sedimentary processes which controlled the deposition of the Ventersdorp Contact Reef within the West Wits Line. In: Anhaeusser, C.R. (Ed.), Proceedings of the 15th Congress of the Council for Mining and Metallurgical Institutions, Volume 3 (Geology): *South African Institute of Mining and Metallurgy*, pp. 117–130.

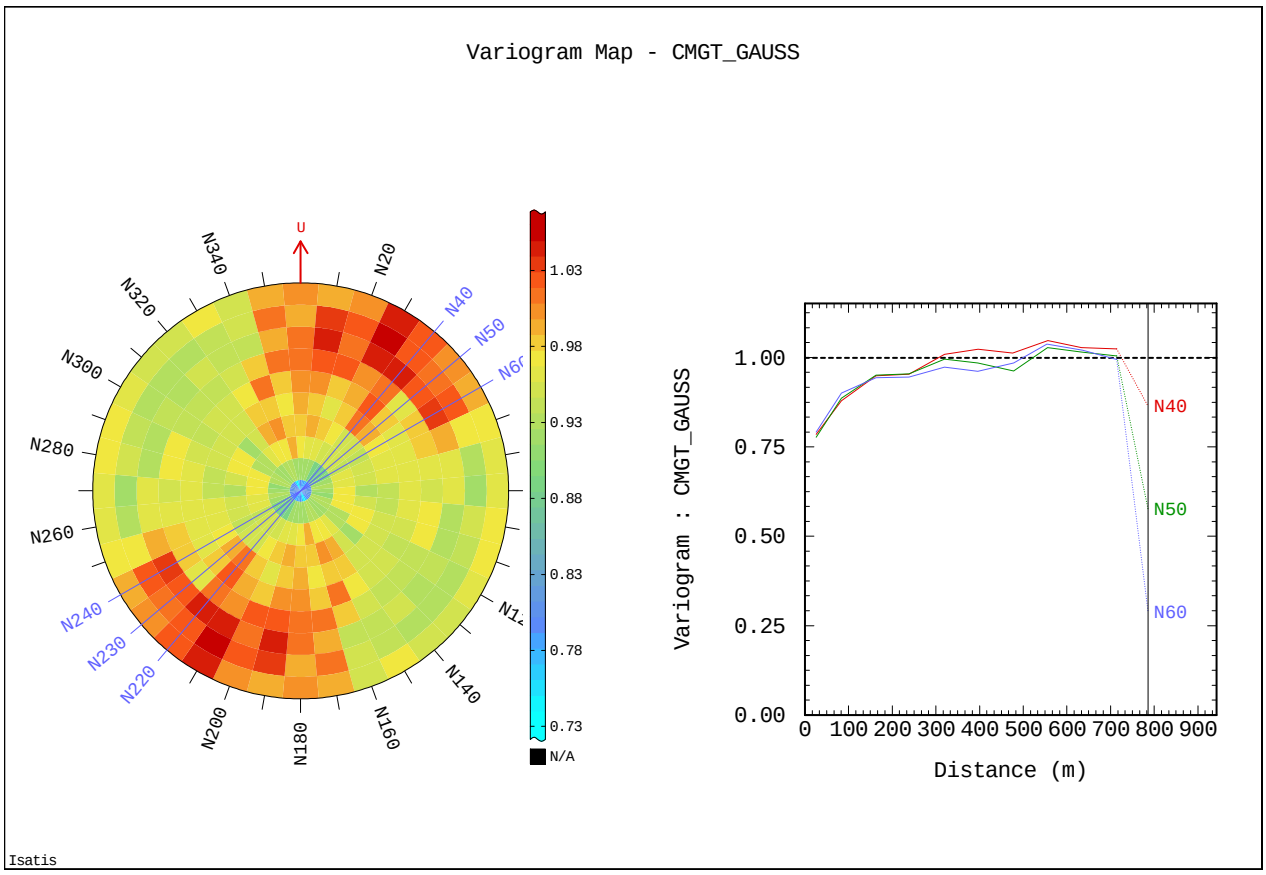
Yaohong, D. Jiang, ThreeDify's CTO, Mining Software Solutions, ThreeDify GeoModeler v1.5.0 adds support for Dynamic Anisotropy, Ottawa.
http://threedify.com/threedify-geomodeler-v1-5-0-adds-support-dynamic-anisotropy/

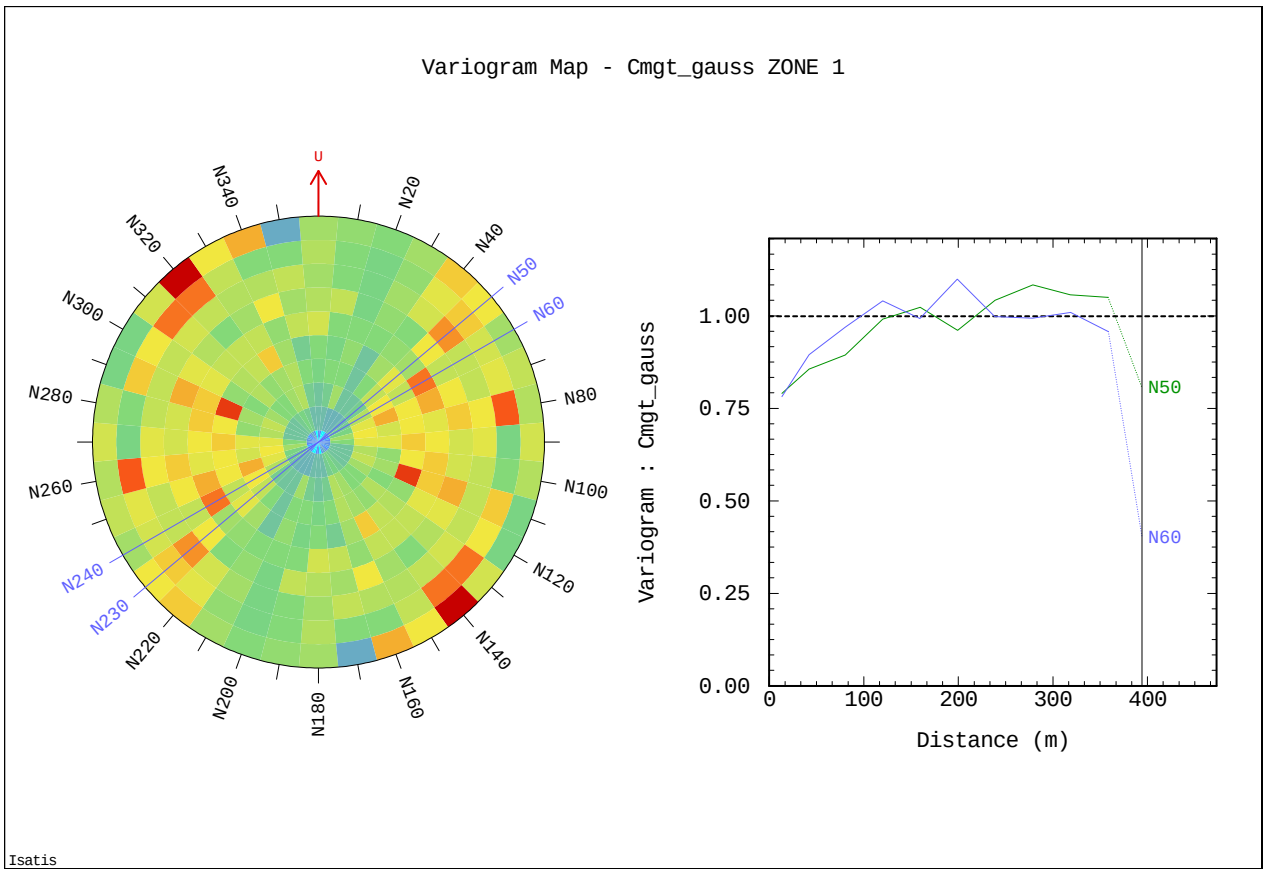
Wheeler, D. J. 2004: Advanced Topics in Statistical Process Control: The Power of Shewhart's Charts, SPC PRESS (Statistical Process Control)

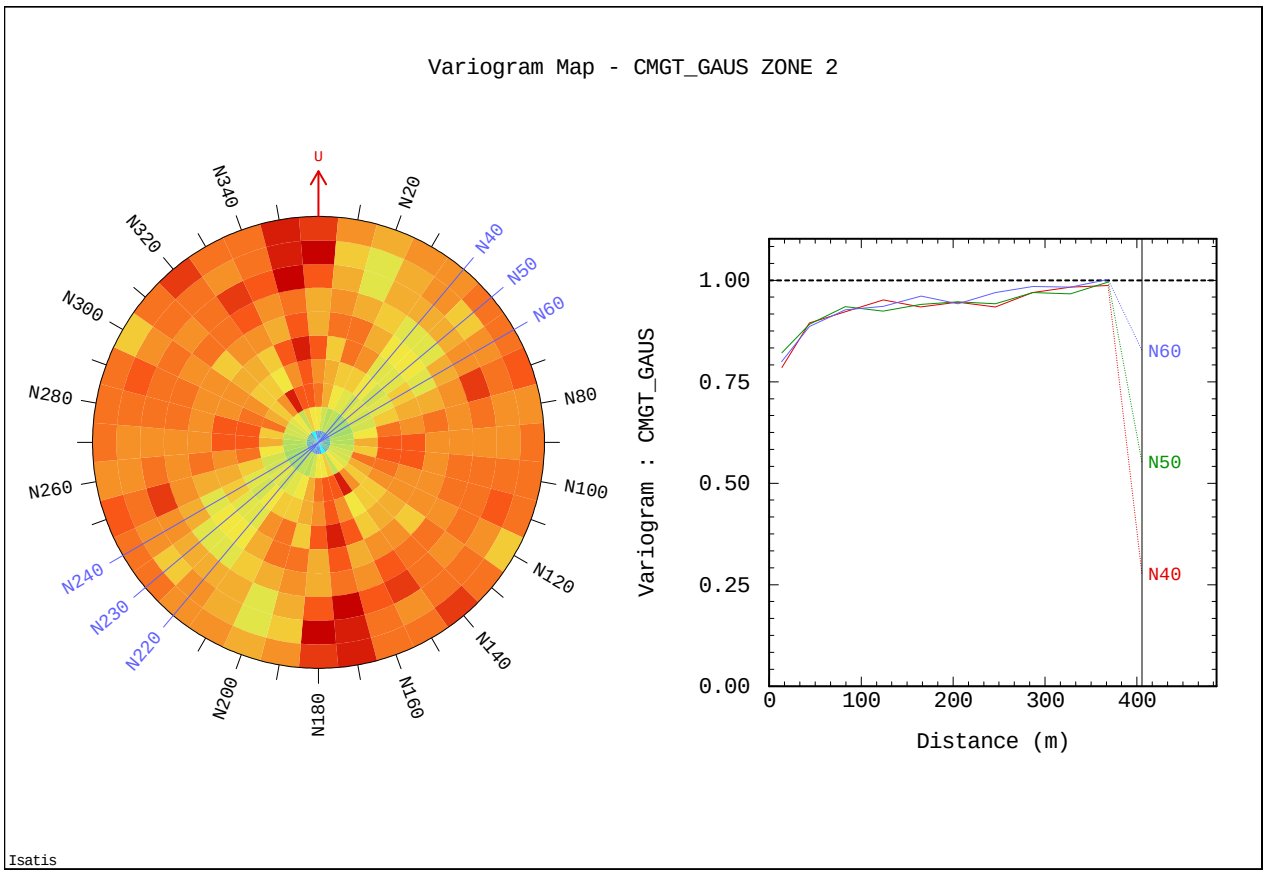
Zabrusky, K. Gustavson and Associates. 2013: Using Dynamic Anisotropy to Improve Resource Estimation, *SME Annual Meeting, Denver*, CO, Preprint 13-128.

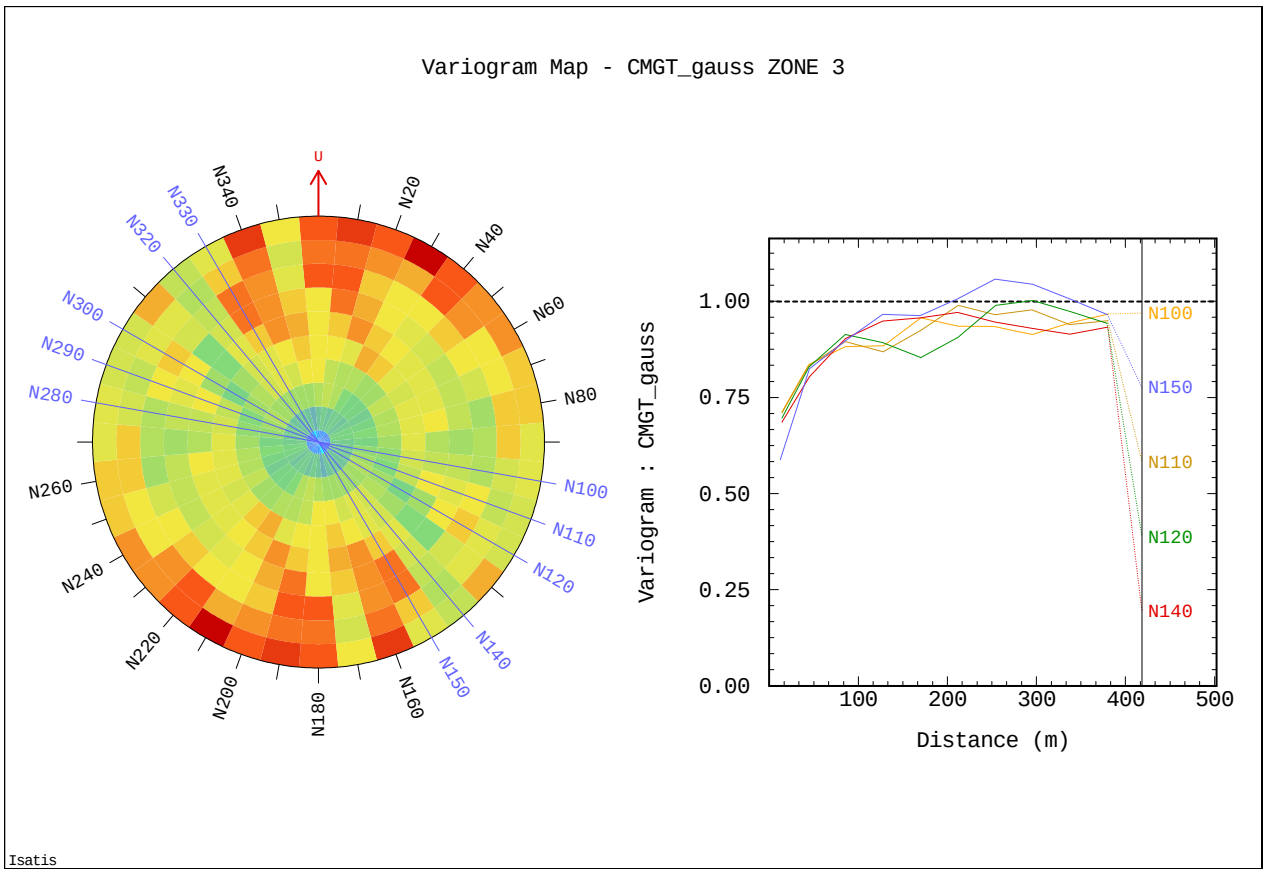
Appendices

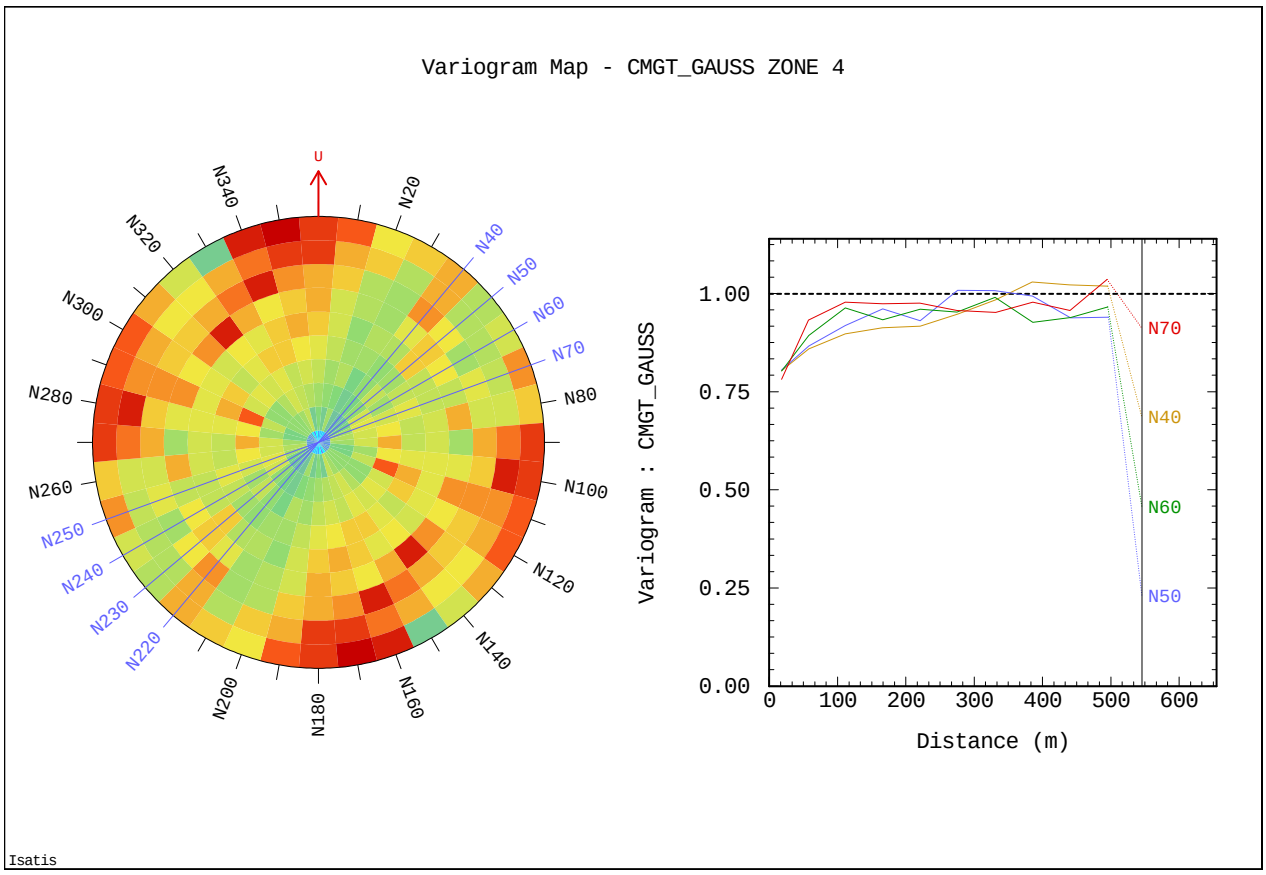
Appendix A: Variogram Maps

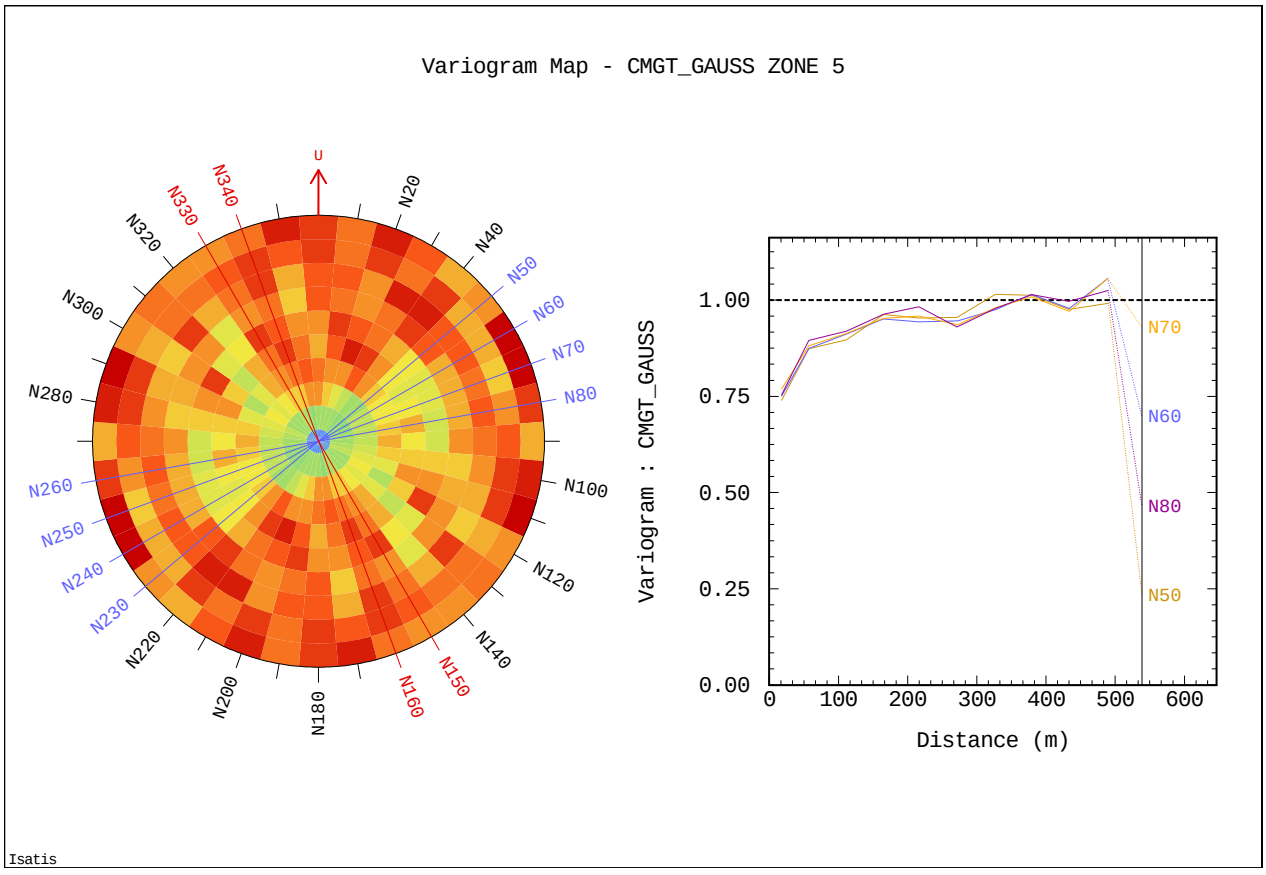

Variogram Map - CMGT_GAUSS
U
N20
N40
N50
N60
N80
N100
N120
N140
N160
N180
N200
N220
N240
N260
N280
N300
N320
N340
1.03
0.98
0.93
0.88
0.83
0.78
0.73
N/A
Variogram : CMGT_GAUSS
1.00
0.75
0.50
0.25
0.00
0 100 200 300 400 500 600 700 800 900
Distance (m)
N40
N50
N60
Isatis


Variogram Map - Cmgt_gauss ZONE 1
U
N20
N40
N50
N60
N80
N100
N120
N140
N160
N180
N200
N220
N240
N260
N280
N300
N320
N340
Variogram : Cmgt_gauss
1.00
0.75
0.50
0.25
0.00
0 100 200 300 400
Distance (m)
N50
N60
Isatis

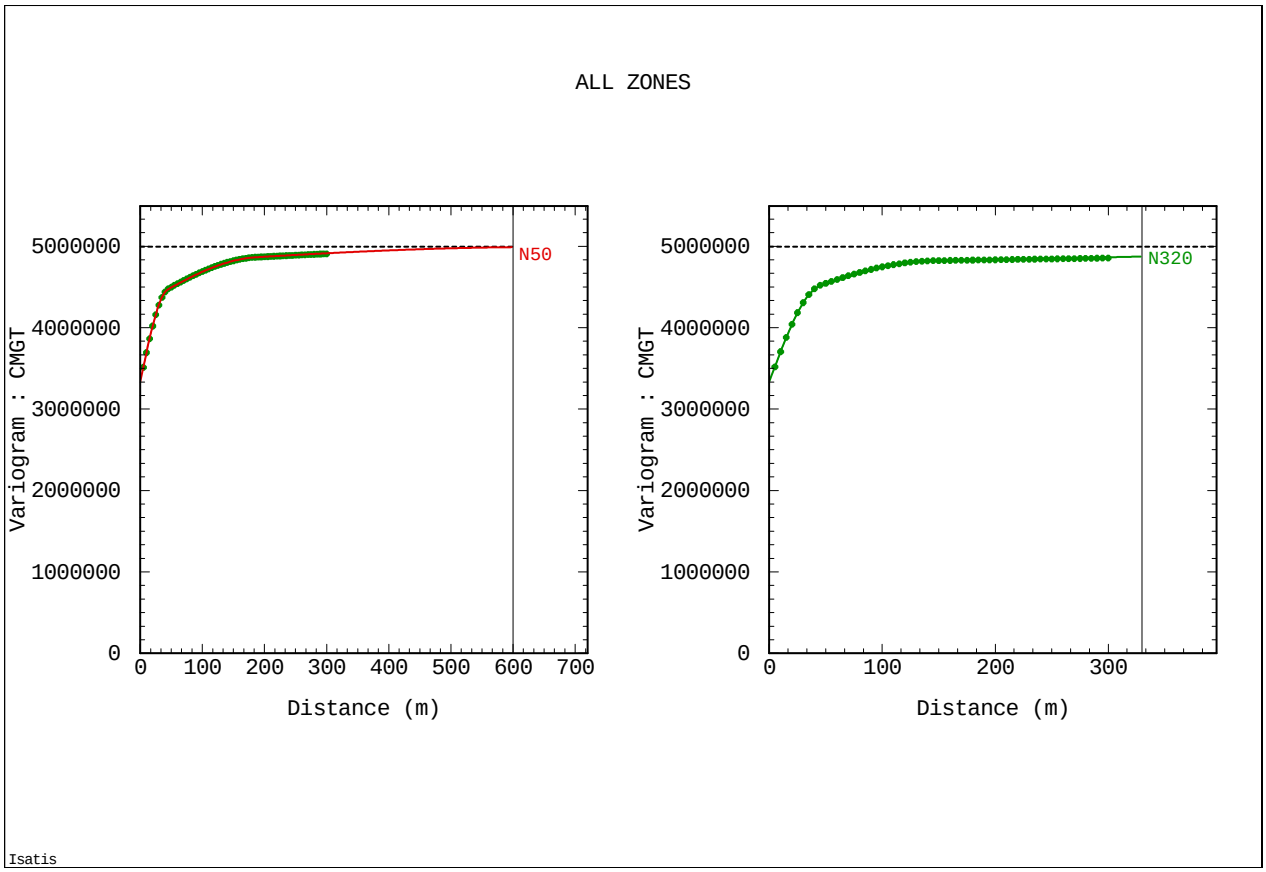

Variogram Map - CMGT_GAUS ZONE 2
U
N20
N40
N50
N60
N80
N100
N120
N140
N160
N180
N200
N220
N230
N240
N260
N280
N300
N320
N340
Variogram : CMGT_GAUS
1.00
0.75
0.50
0.25
0.00
0
100
200
300
400
Distance (m)
N60
N50
N40
Isatis

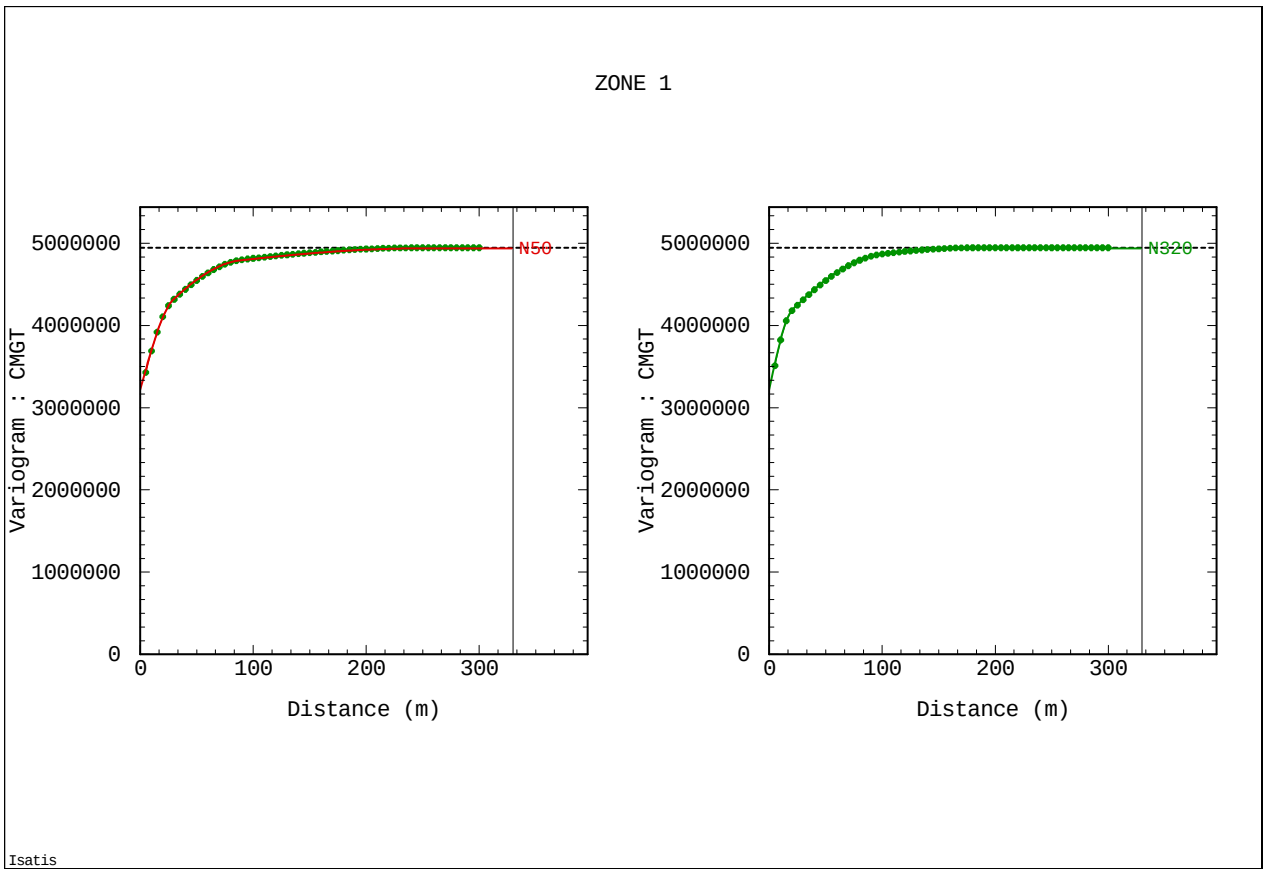

Variogram Map - CMGT_gauss ZONE 3
U
N20
N40
N60
N80
N100
N110
N120
N140
N150
N160
N180
N200
N220
N240
N260
N280
N290
N300
N320
N330
N340
Variogram : CMGT_gauss
1.00
0.75
0.50
0.25
0.00
100
200
300
400
500
Distance (m)
N100
N150
N110
N120
N140
Isatis

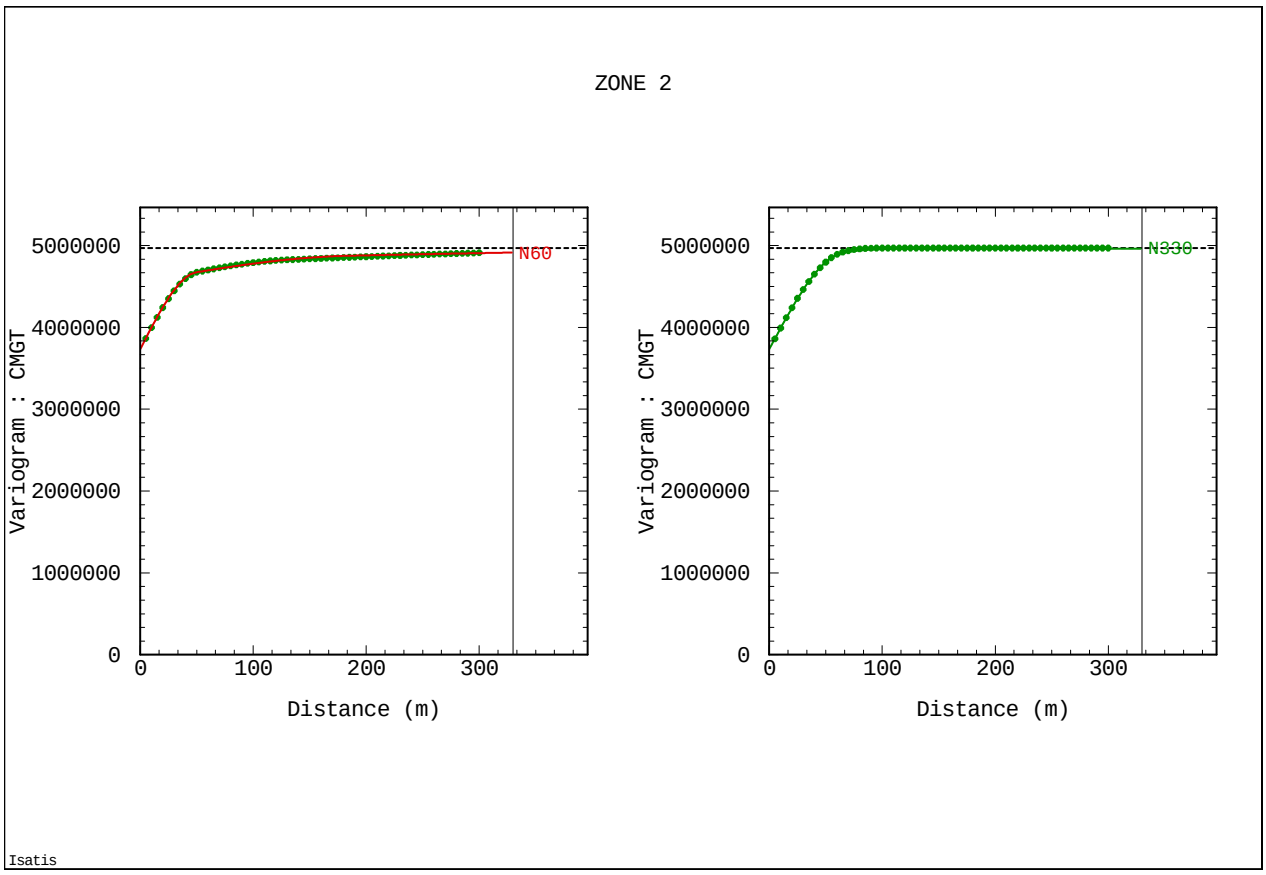

Variogram Map - CMGT_GAUSS ZONE 4
U
N20
N40
N50
N60
N70
N80
N100
N120
N140
N160
N180
N200
N220
N230
N240
N250
N260
N280
N300
N320
N340
Variogram : CMGT_GAUSS
1.00
0.75
0.50
0.25
0.00
0
100
200
300
400
500
600
Distance (m)
N70
N40
N60
N50
Isatis

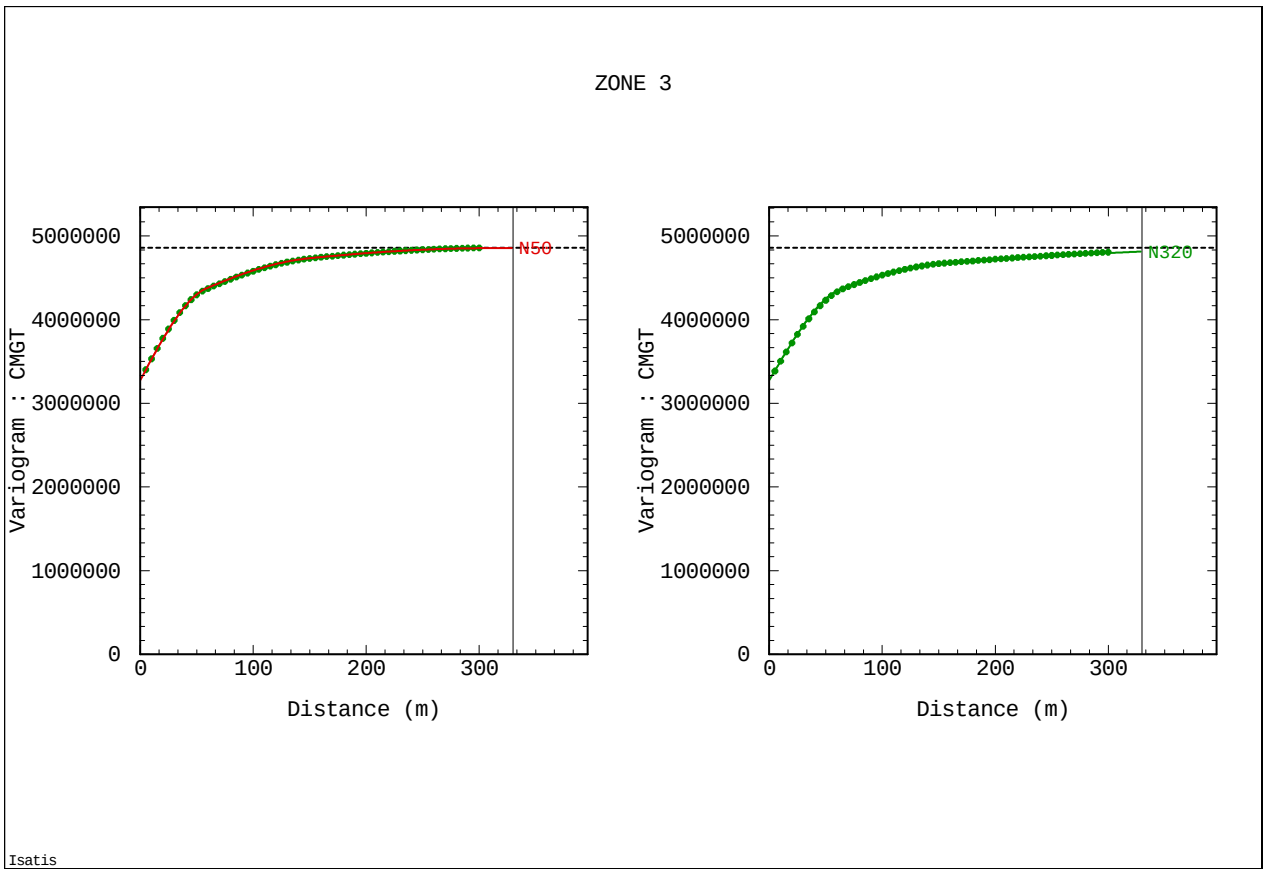

Variogram Map - CMGT_GAUSS ZONE 5
U
N20
N40
N50
N60
N70
N80
N100
N120
N140
N150
N160
N180
N200
N220
N230
N240
N250
N260
N280
N300
N320
N330
N340
Variogram : CMGT_GAUSS
1.00
0.75
0.50
0.25
0.00
0
100
200
300
400
500
600
Distance (m)
N70
N60
N80
N50
Isatis

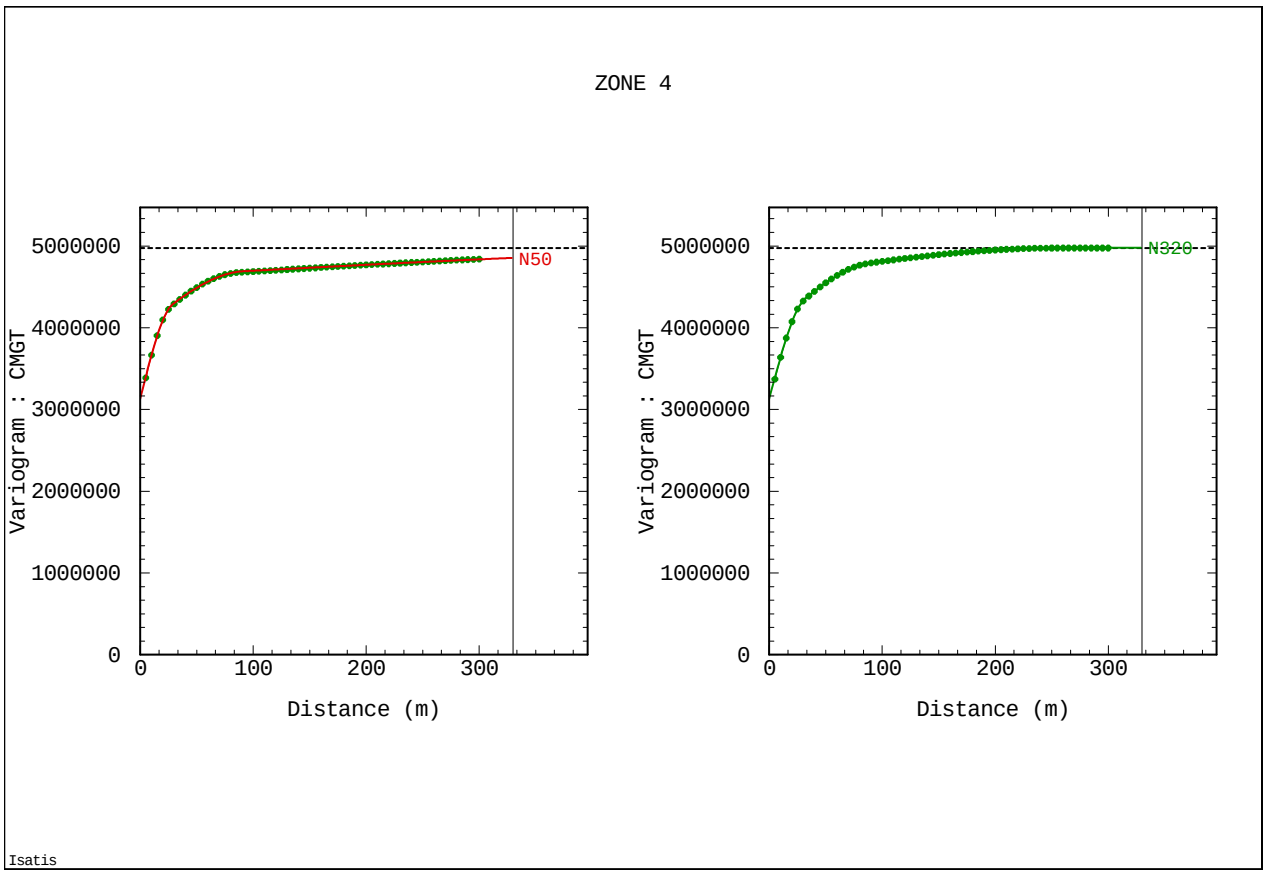
Appendix B: Fitted Back Transformed Semi-variograms

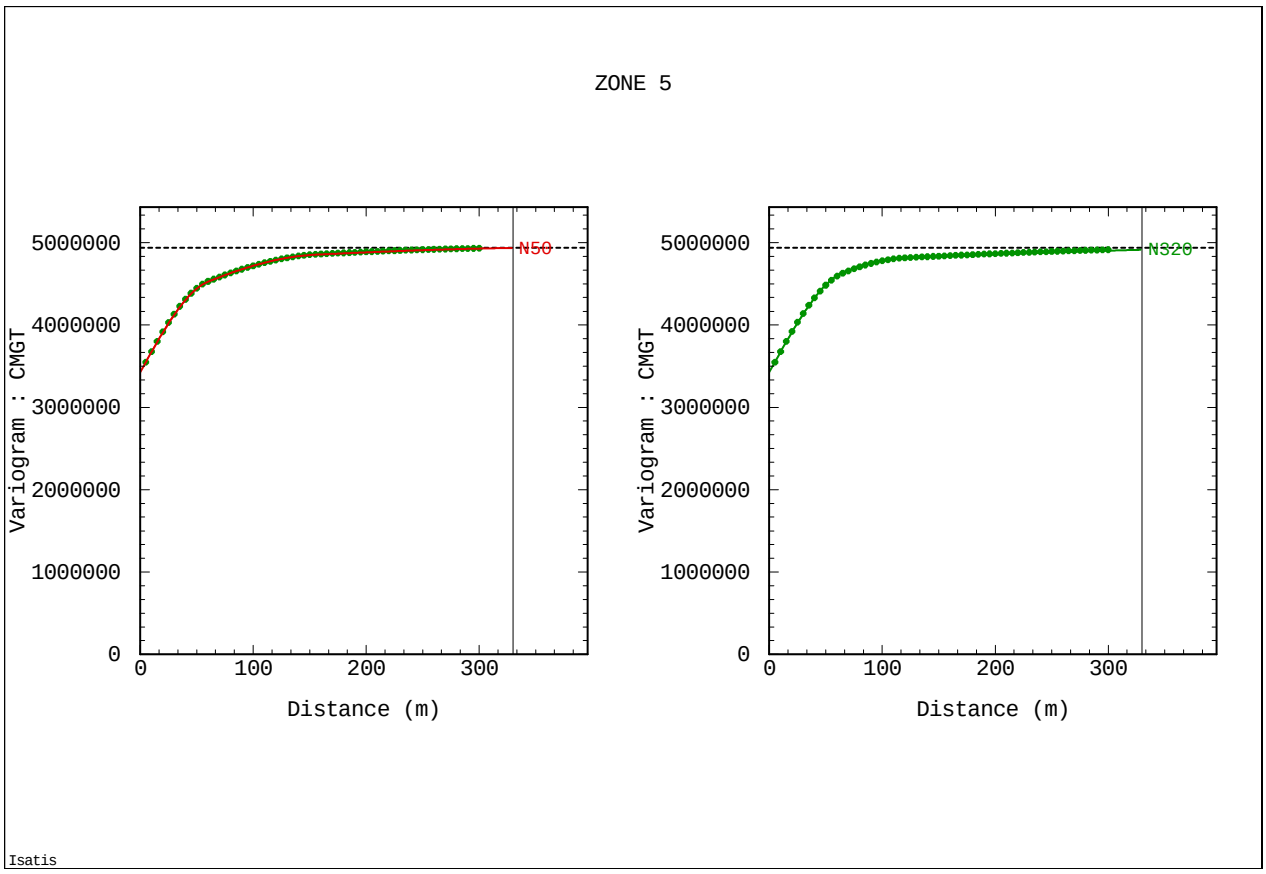

ALL ZONES
Variogram : CMGT
5000000
4000000
3000000
2000000
1000000
0
0
100
200
300
400
500
600
700
Distance (m)
N50
Variogram : CMGT
5000000
4000000
3000000
2000000
1000000
0
0
100
200
300
Distance (m)
N320
Isatis


ZONE 1
Variogram : CMGT
5000000
4000000
3000000
2000000
1000000
0
0
100
200
300
Distance (m)
N50
Variogram : CMGT
5000000
4000000
3000000
2000000
1000000
0
0
100
200
300
Distance (m)
N320
Isatis

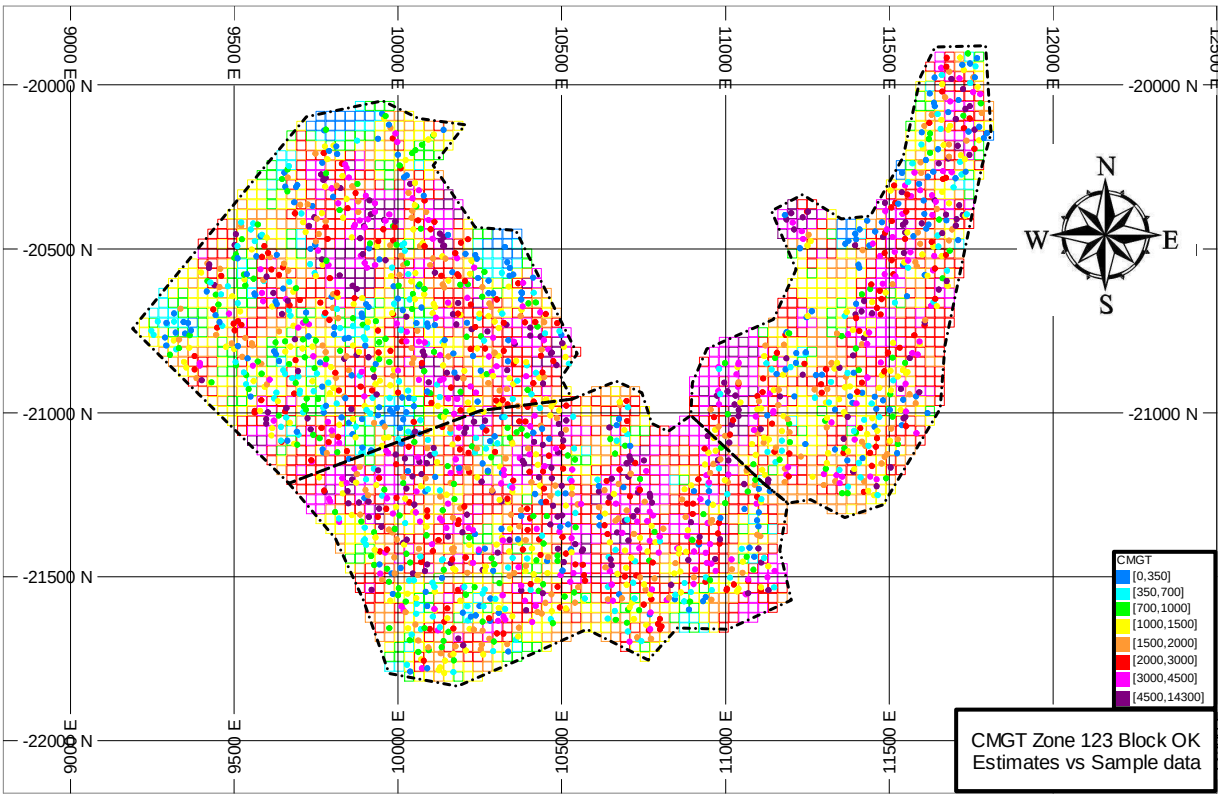

ZONE 2
Variogram : CMGT
5000000
4000000
3000000
2000000
1000000
0
0
100
200
300
Distance (m)
N60
Variogram : CMGT
5000000
4000000
3000000
2000000
1000000
0
0
100
200
300
Distance (m)
N330
Isatis

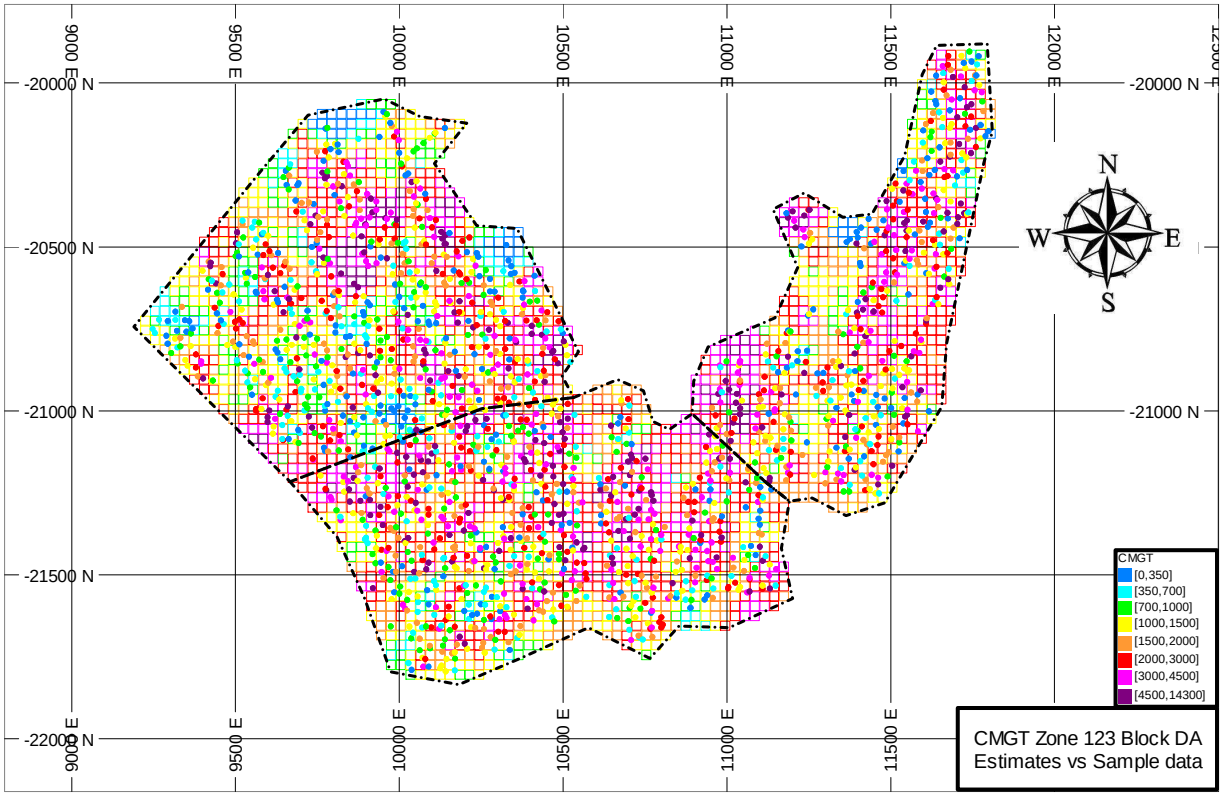

ZONE 3
Variogram : CMGT
5000000
4000000
3000000
2000000
1000000
0
0
100
200
300
Distance (m)
N50
Variogram : CMGT
5000000
4000000
3000000
2000000
1000000
0
0
100
200
300
Distance (m)
N320
Isatis

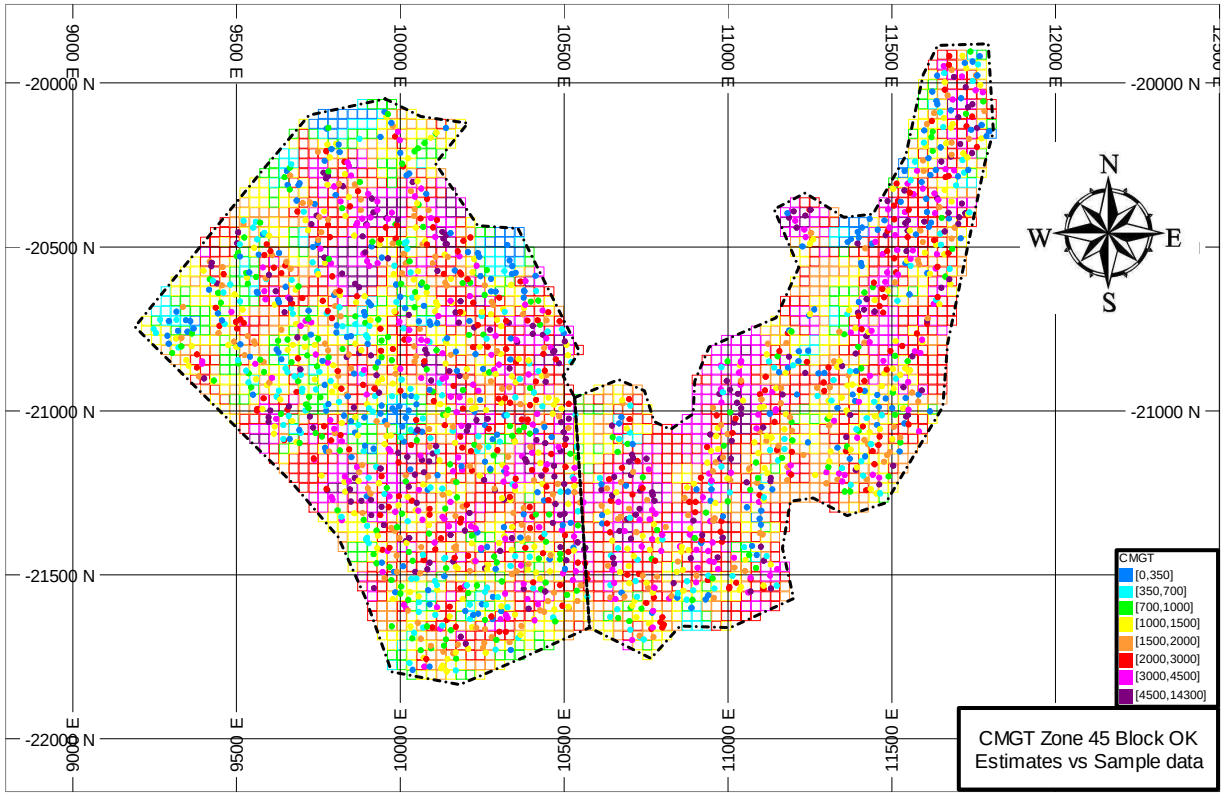

ZONE 4
Variogram : CMGT
5000000
4000000
3000000
2000000
1000000
0
0
100
200
300
Distance (m)
N50
N320
Isatis

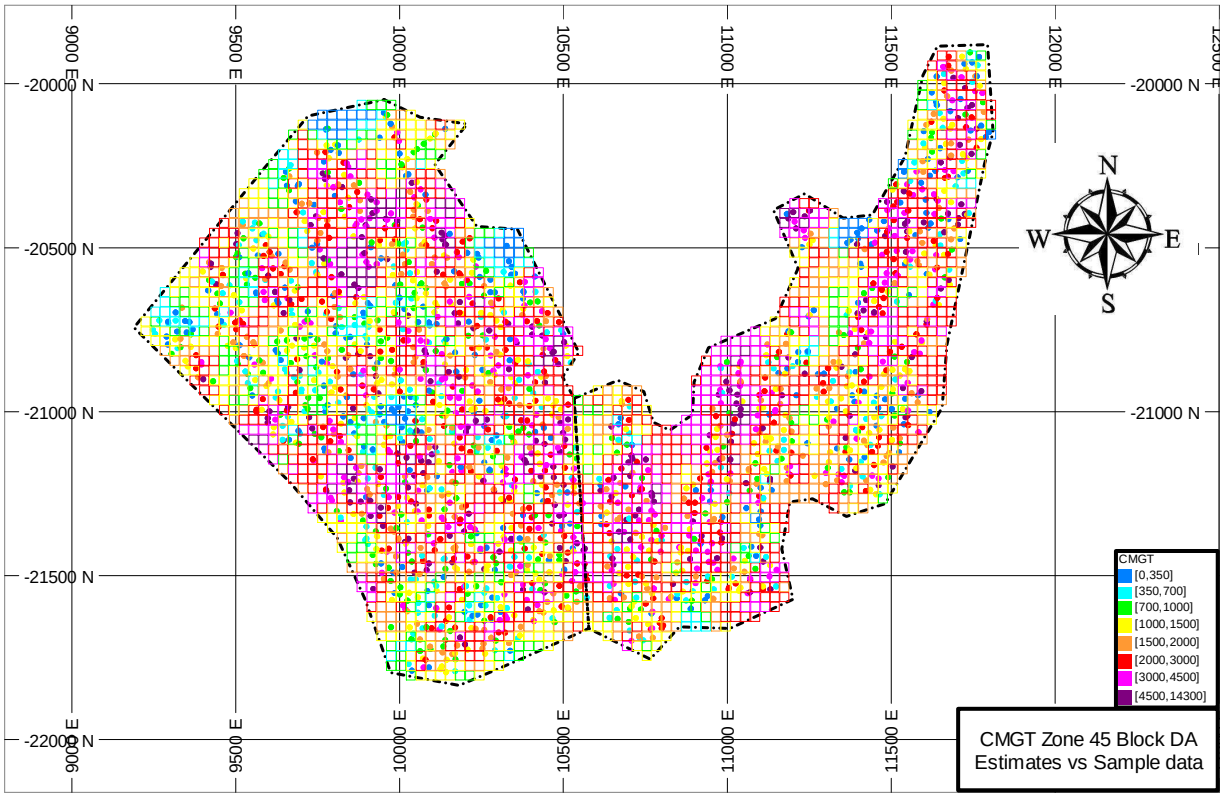

ZONE 5
Variogram : CMGT
5000000
4000000
3000000
2000000
1000000
0
0
100
200
300
Distance (m)
N50
N320
Isatis

Appendix C: Domain Grade Distribution Validation Plots

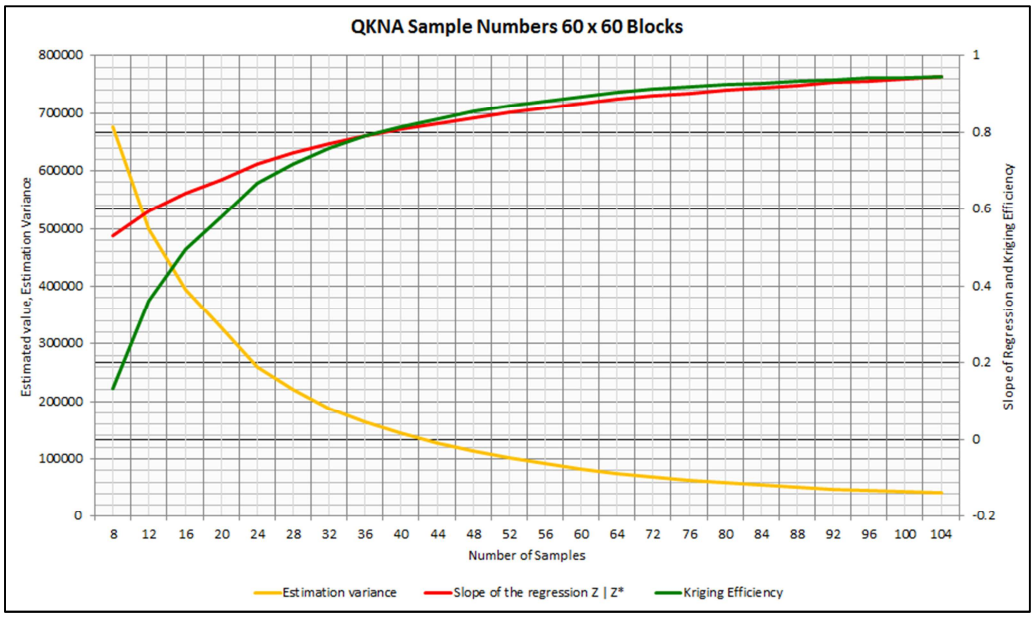

9000 E
9500 E
10000 E
10500 E
11000 E
11500 E
12000 E
-20000 N
-20500 N
-21000 N
-21500 N
-22000 N
N
W
E
S
CMGT
[0,350]
[350,700]
[700,1000]
[1000,1500]
[1500,2000]
[2000,3000]
[3000,4500]
[4500,14300]
CMGT Zone 123 Block OK
Estimates vs Sample data

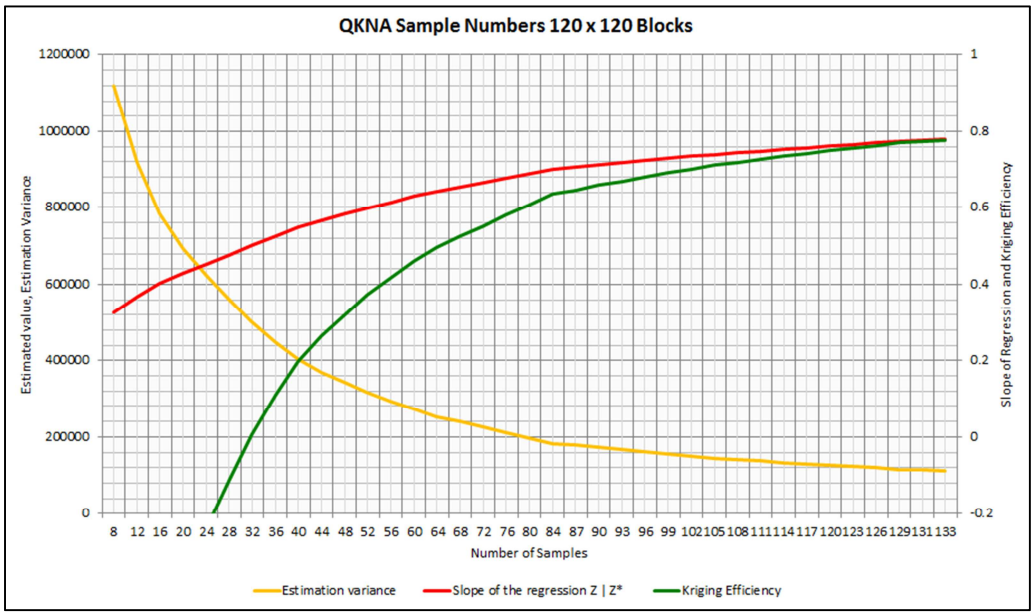

9000 E
9500 E
10000 E
10500 E
11000 E
11500 E
12000 E
-20000 N
-20500 N
-21000 N
-21500 N
-22000 N
N
W
E
S
CMGT
[0,350]
[350,700]
[700,1000]
[1000,1500]
[1500,2000]
[2000,3000]
[3000,4500]
[4500,14300]
CMGT Zone 123 Block DA
Estimates vs Sample data

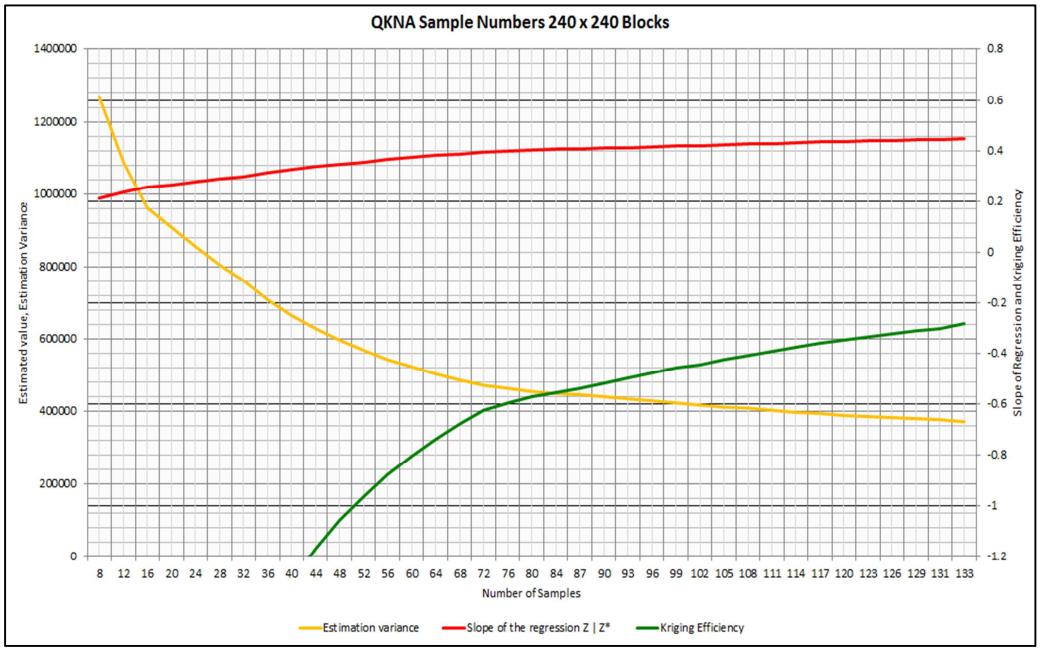

9000 E
9500 E
10000 E
10500 E
11000 E
11500 E
12000 E
12500 E
-20000 N
-20500 N
-21000 N
-21500 N
-22000 N
N
W
E
S
CMGT
[0,350]
[350,700]
[700,1000]
[1000,1500]
[1500,2000]
[2000,3000]
[3000,4500]
[4500,14300]
CMGT Zone 45 Block OK
Estimates vs Sample data


9000 E
9500 E
10000 E
10500 E
11000 E
11500 E
12000 E
12500 E
-20000 N
-20500 N
-21000 N
-21500 N
-22000 N
N
W
E
S
CMGT
[0,350]
[350,700]
[700,1000]
[1000,1500]
[1500,2000]
[2000,3000]
[3000,4500]
[4500,14300]
CMGT Zone 45 Block DA
Estimates vs Sample data

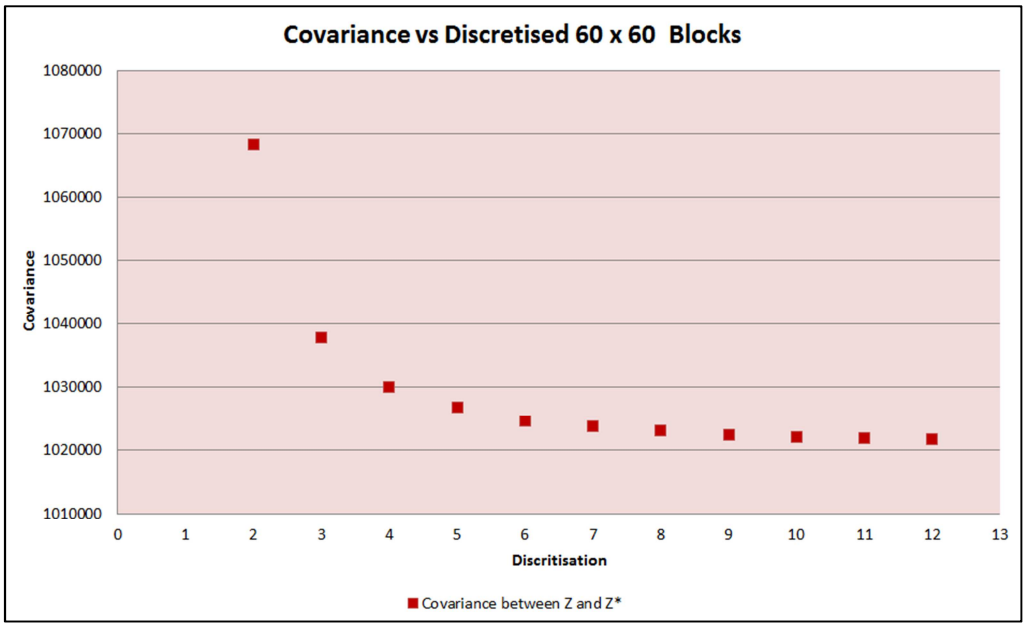
Appendix D: Sample Number Optimisation

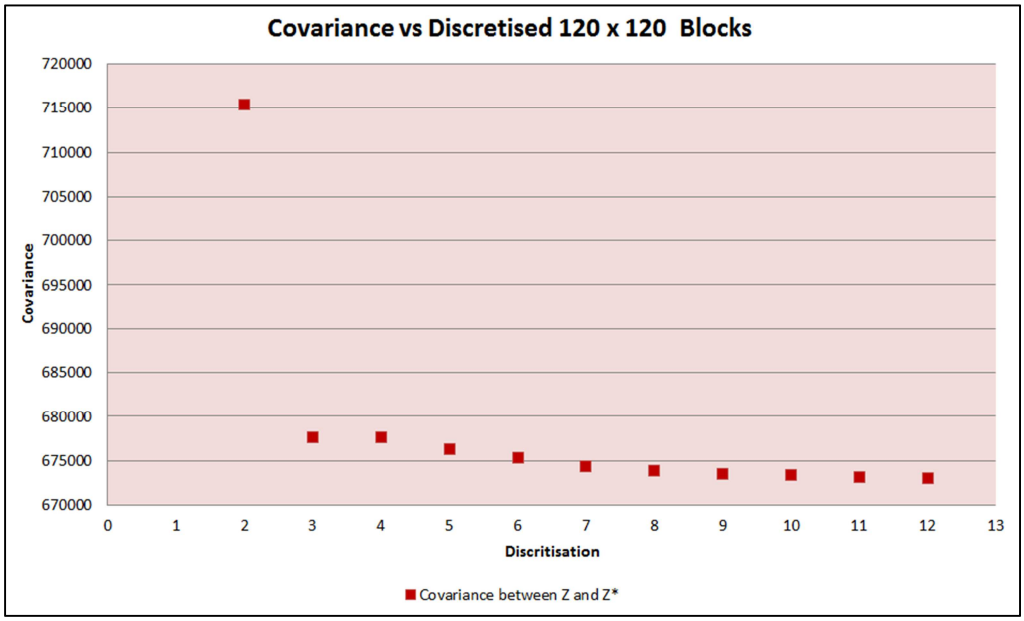

QKNA Sample Numbers 60 x 60 Blocks
Estimated value, Estimation Variance
Slope of Regression and Kriging Efficiency
Number of Samples
Estimation variance
Slope of the regression Z | Z*
Kriging Efficiency

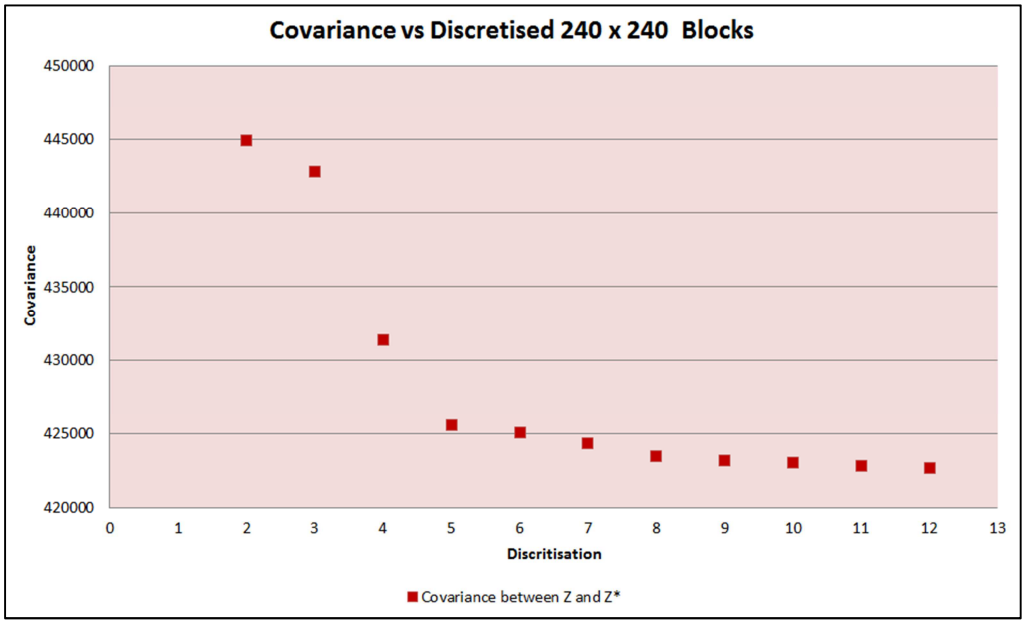

QKNA Sample Numbers 120 x 120 Blocks
Estimated value, Estimation Variance
Slope of Regression and Kriging Efficiency
Number of Samples
Estimation variance
Slope of the regression Z | Z*
Kriging Efficiency


QKNA Sample Numbers 240 x 240 Blocks
Estimated value, Estimation Variance
Slope of Regression and Kriging Efficiency
Number of Samples
Estimation variance
Slope of the regression Z | Z*
Kriging Efficiency

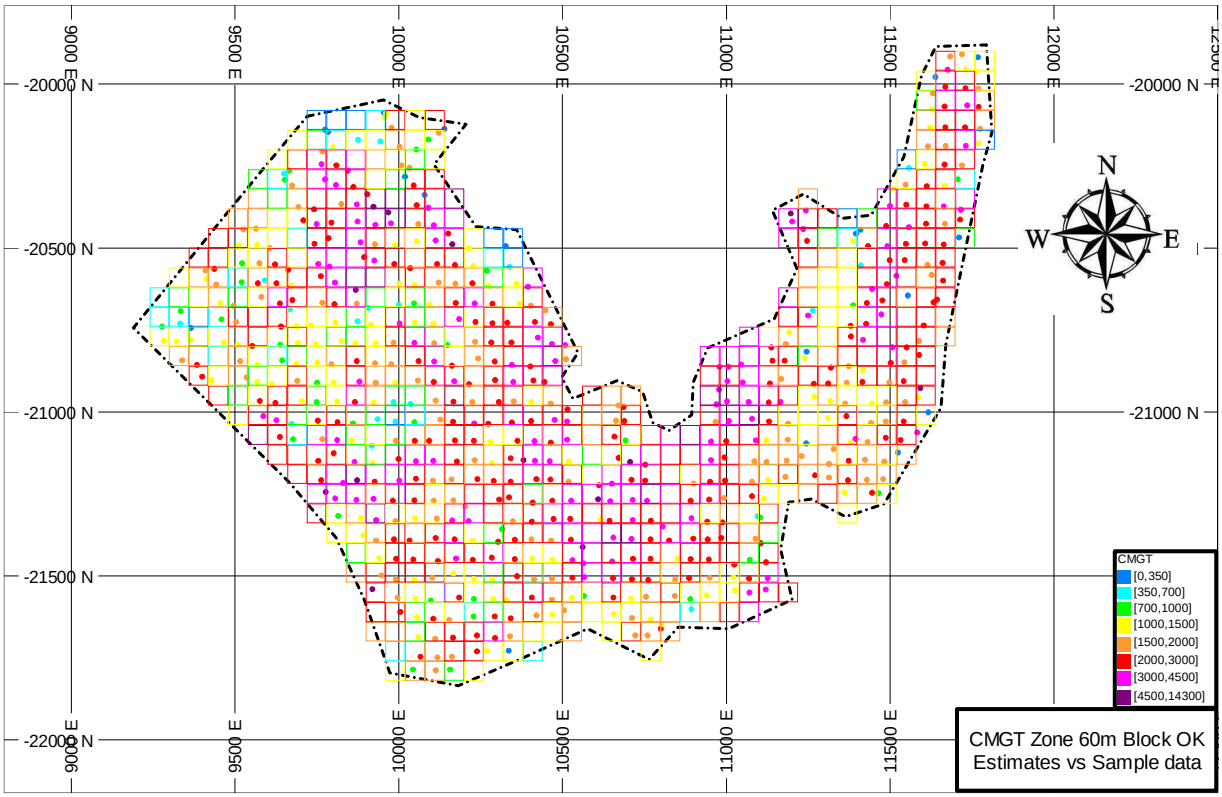
Appendix E: Block Discretization Number Optimisation

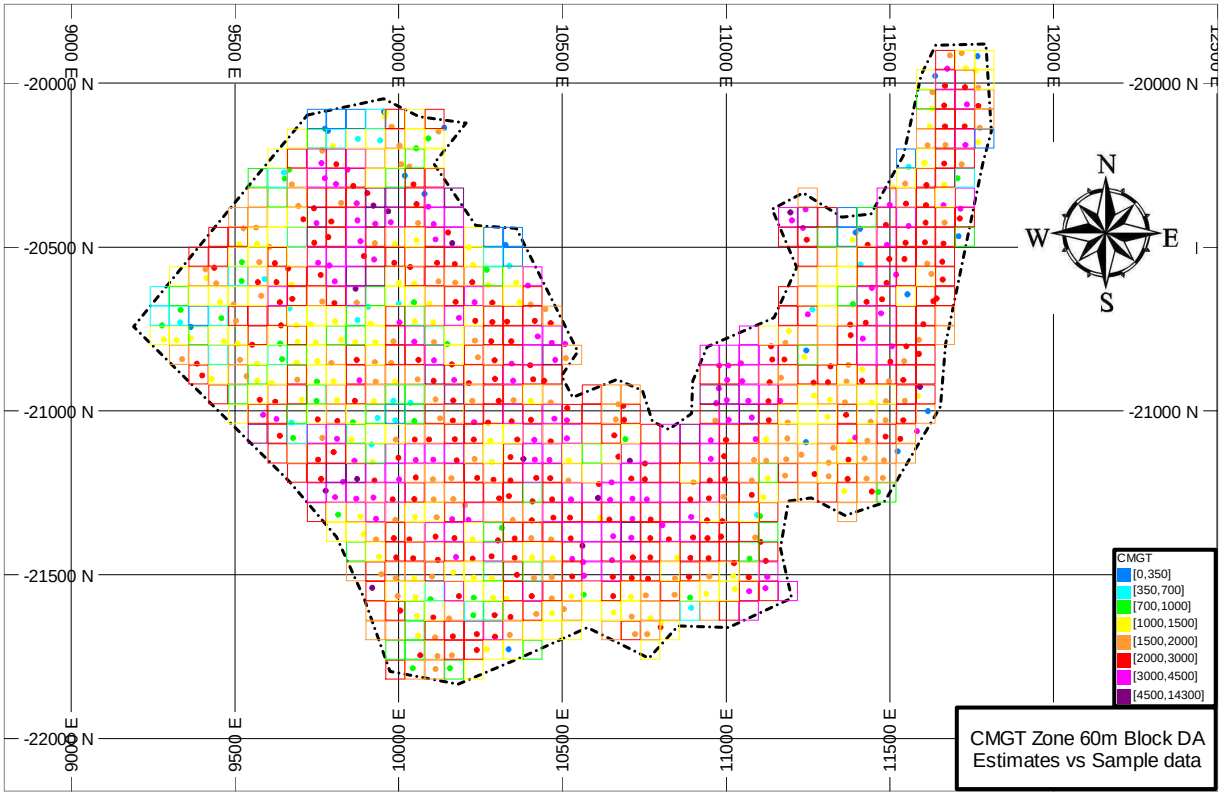

Covariance vs Discretised 60 x 60 Blocks
1080000
1070000
1060000
1050000
1040000
1030000
1020000
1010000
Covariance
0 1 2 3 4 5 6 7 8 9 10 11 12 13
Discritisation
Covariance between Z and Z*

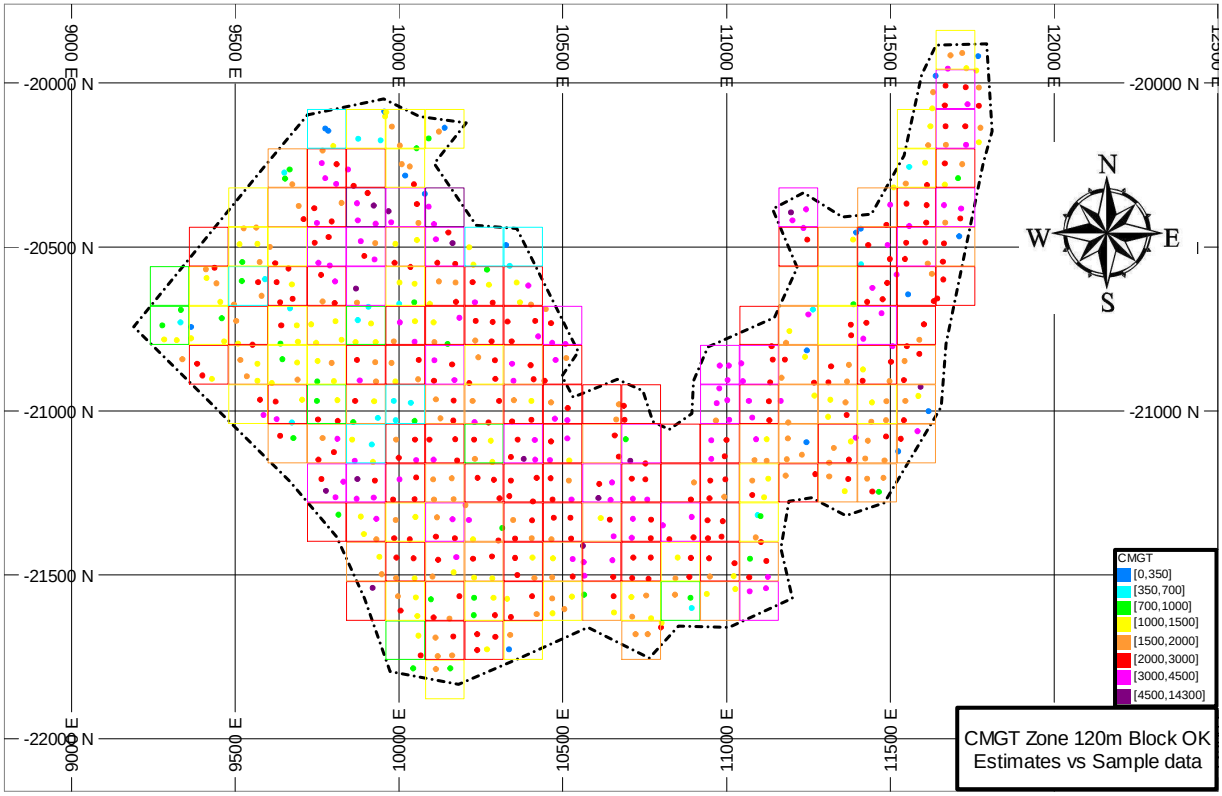

Covariance vs Discretised 120 x 120 Blocks
720000
715000
710000
705000
700000
695000
690000
685000
680000
675000
670000
Covariance
0 1 2 3 4 5 6 7 8 9 10 11 12 13
Discritisation
Covariance between Z and Z*

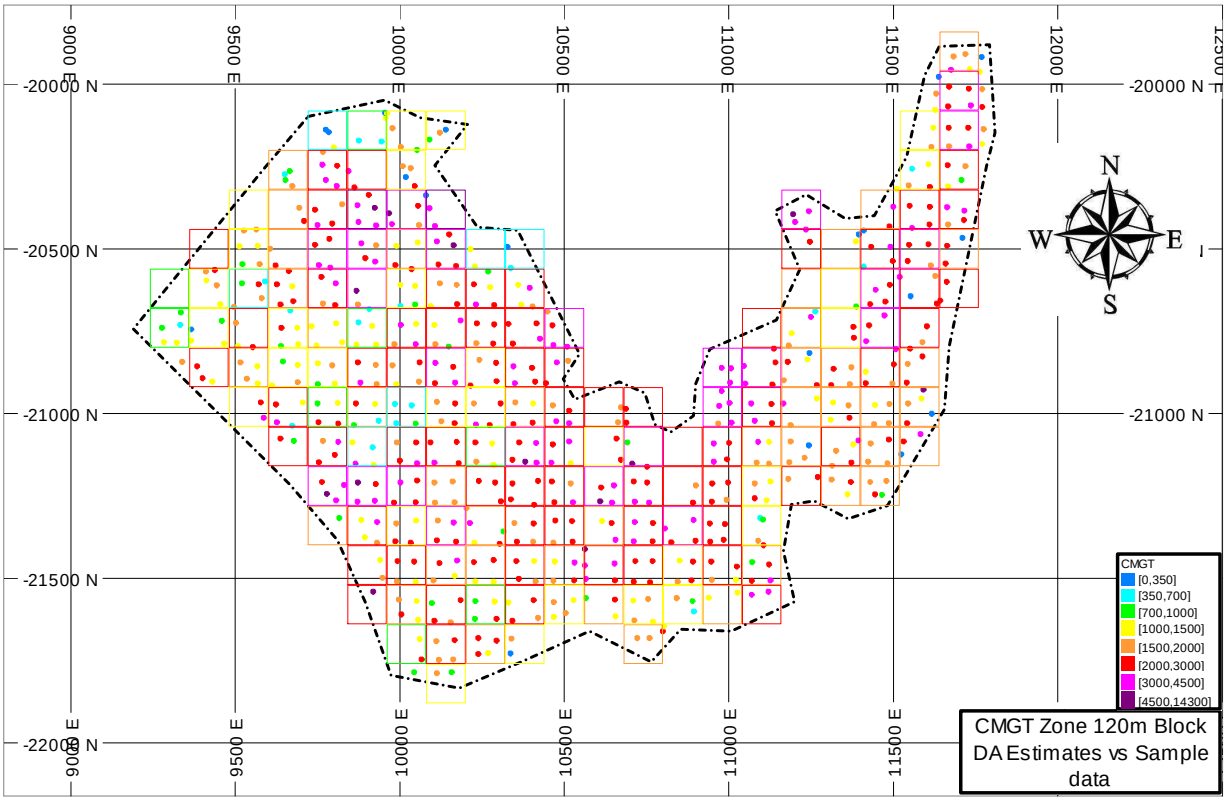

Covariance vs Discretised 240 x 240 Blocks
450000
445000
440000
435000
430000
425000
420000
Covariance
0 1 2 3 4 5 6 7 8 9 10 11 12 13
Discritisation
Covariance between Z and Z*

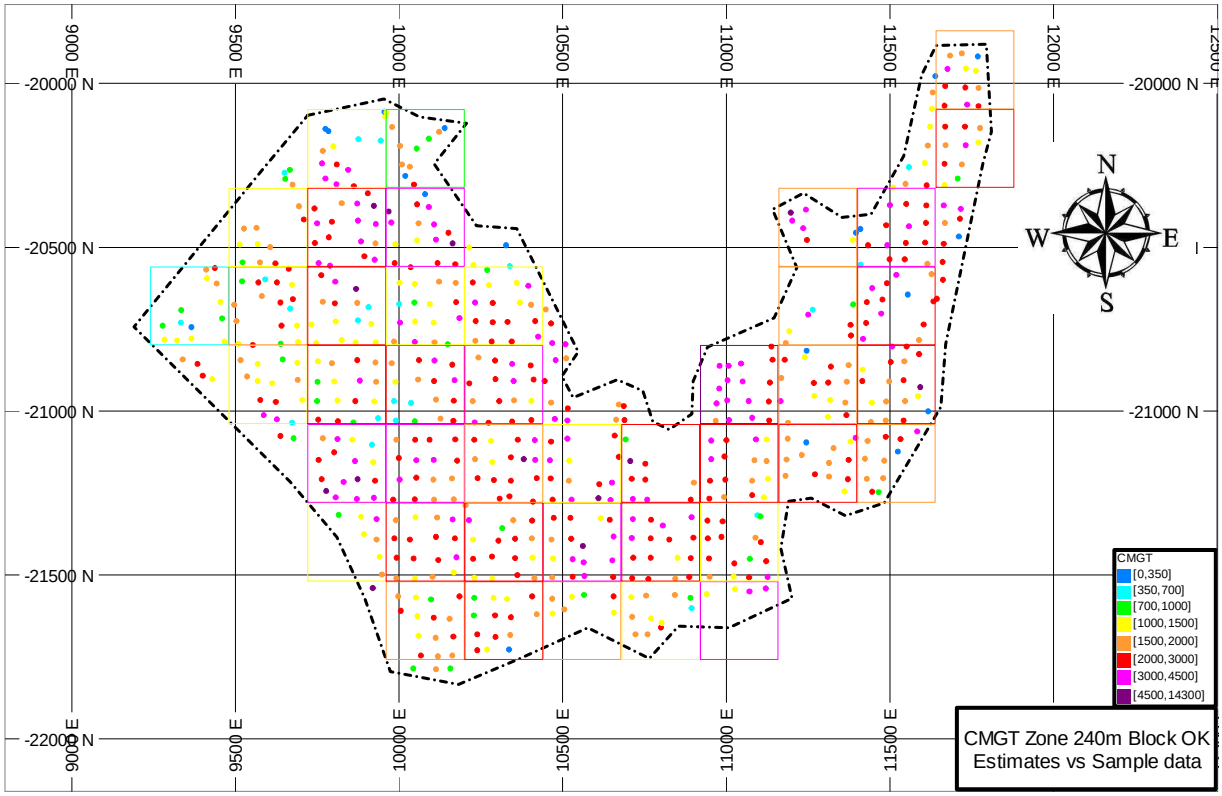
Appendix F: Grade Distribution Validation Plots for Different Block Sizes

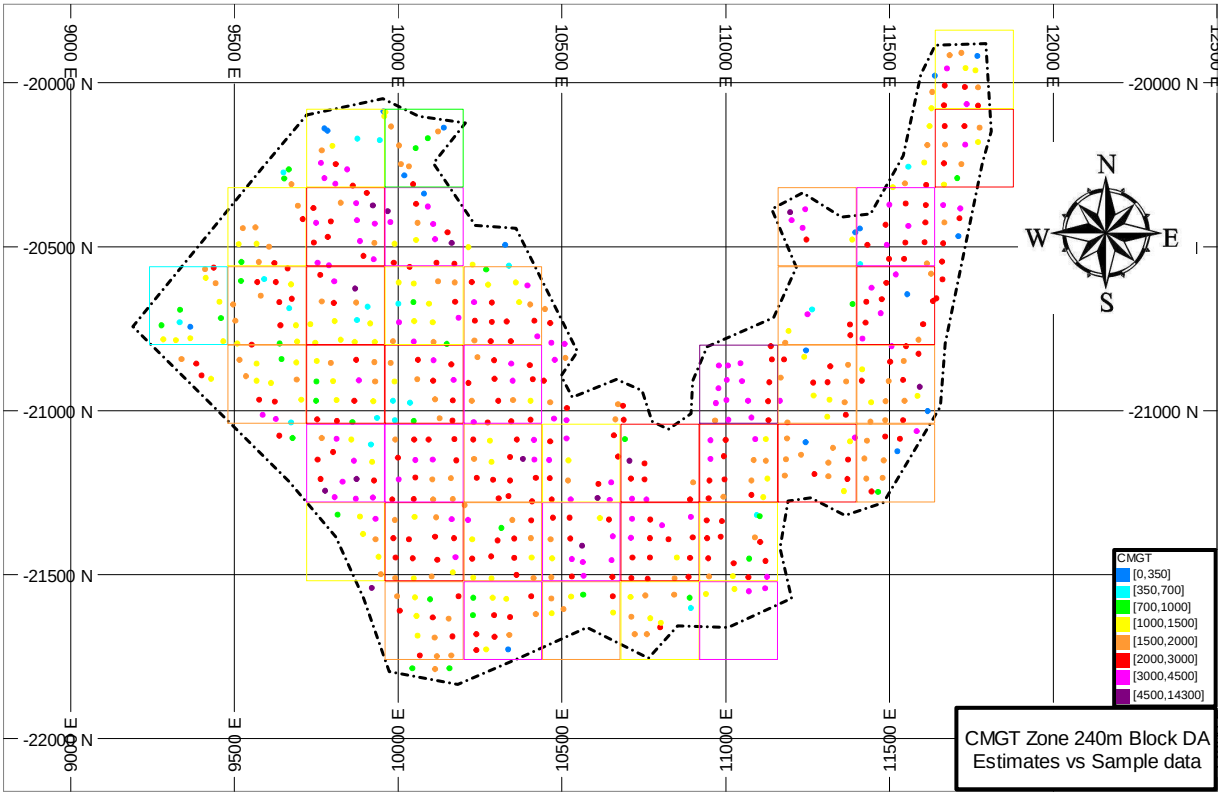

9000 E
9500 E
10000 E
10500 E
11000 E
11500 E
12000 E
12500 E
-20000 N
-20500 N
-21000 N
-21500 N
-22000 N
N
W
E
S
CMGT
[0,350]
[350,700]
[700,1000]
[1000,1500]
[1500,2000]
[2000,3000]
[3000,4500]
[4500,14300]
CMGT Zone 60m Block OK
Estimates vs Sample data


9000 E
9500 E
10000 E
10500 E
11000 E
11500 E
12000 E
12500 E
-20000 N
-20500 N
-21000 N
-21500 N
-22000 N
N
W
E
S
CMGT
[0,350]
[350,700]
[700,1000]
[1000,1500]
[1500,2000]
[2000,3000]
[3000,4500]
[4500,14300]
CMGT Zone 60m Block DA
Estimates vs Sample data

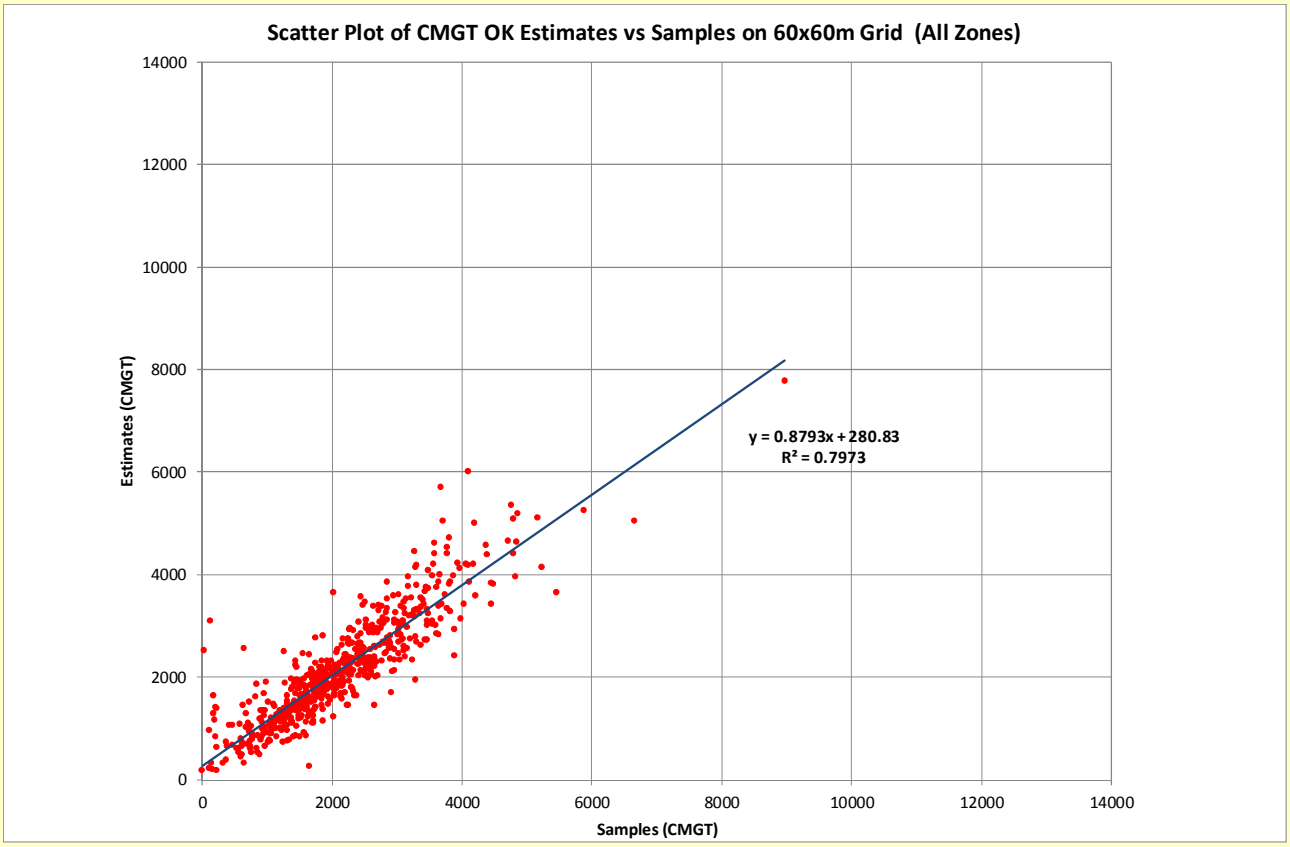

9000 E
9500 E
10000 E
10500 E
11000 E
11500 E
12000 E
-20000 N
-20500 N
-21000 N
-21500 N
-22000 N
N
W
E
S
CMGT
[0,350]
[350,700]
[700,1000]
[1000,1500]
[1500,2000]
[2000,3000]
[3000,4500]
[4500,14300]
CMGT Zone 120m Block OK Estimates vs Sample data

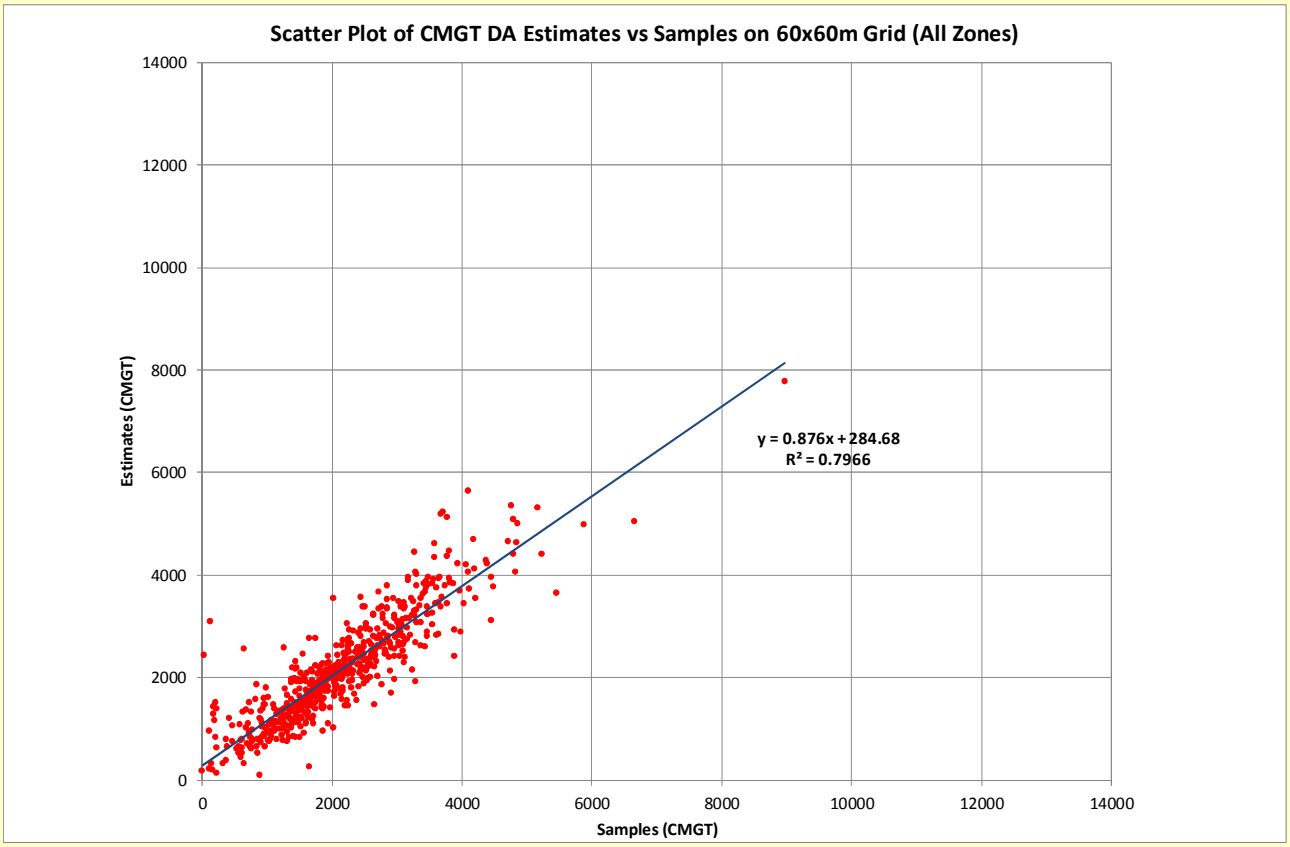

9000 E
9500 E
10000 E
10500 E
11000 E
11500 E
12000 E
-20000 N
-20500 N
-21000 N
-21500 N
-22000 N
N
W
E
S
CMGT
[0,350]
[350,700]
[700,1000]
[1000,1500]
[1500,2000]
[2000,3000]
[3000,4500]
[4500,14300]
CMGT Zone 120m Block DA Estimates vs Sample data

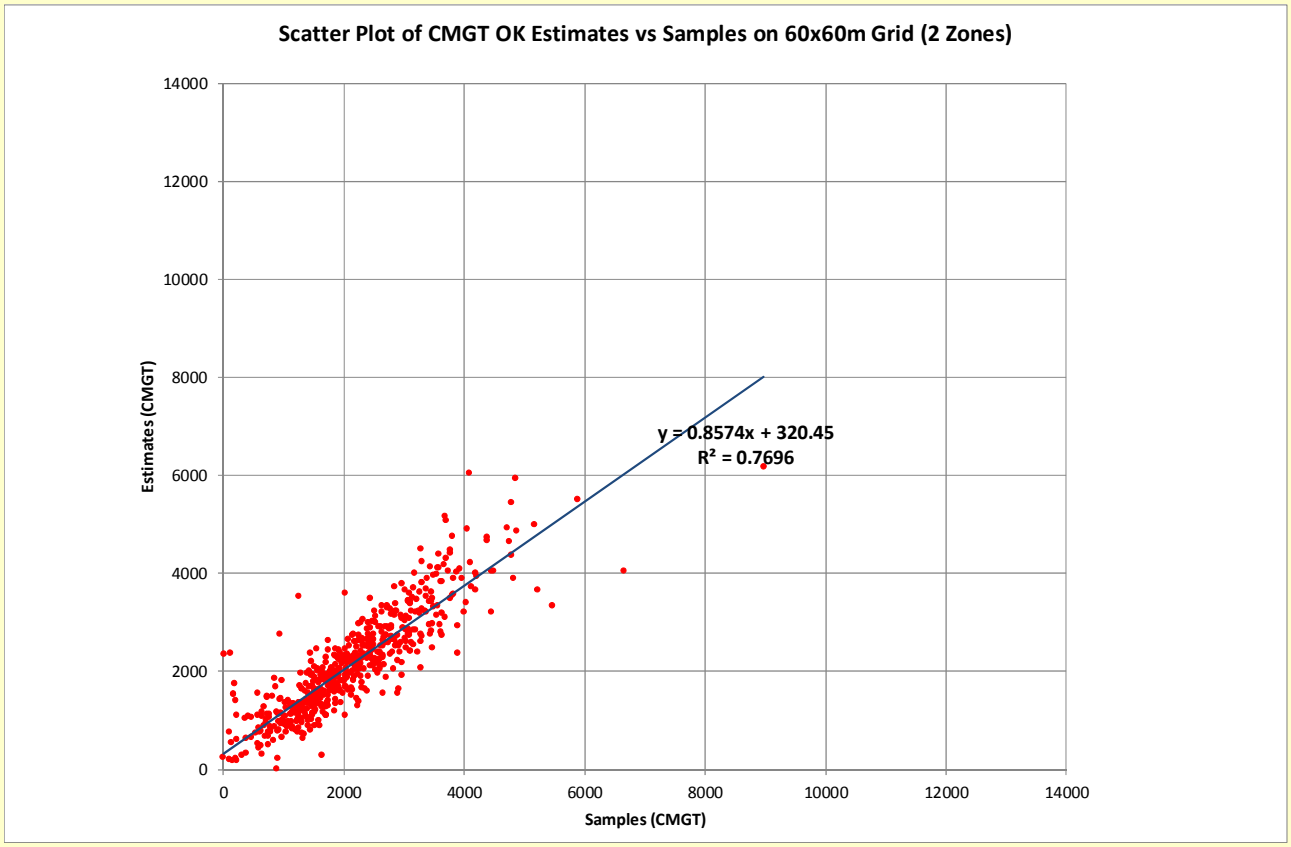

9000 E
9500 E
10000 E
10500 E
11000 E
11500 E
12000 E
12500 E
-20000 N
-20500 N
-21000 N
-21500 N
-22000 N
N
W
E
S
CMGT
[0,350]
[350,700]
[700,1000]
[1000,1500]
[1500,2000]
[2000,3000]
[3000,4500]
[4500,14300]
CMGT Zone 240m Block OK
Estimates vs Sample data

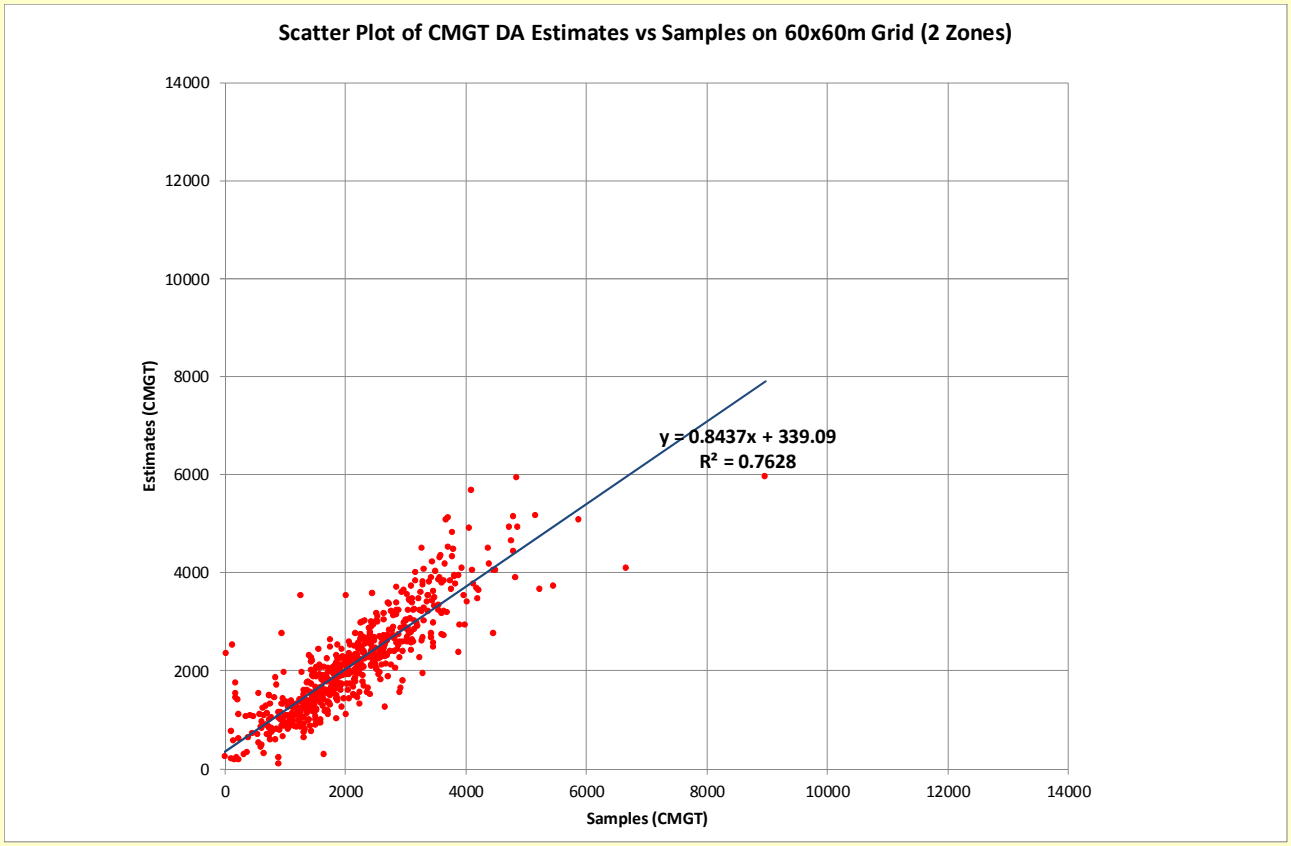

9000 E
9500 E
10000 E
10500 E
11000 E
11500 E
12000 E
12500 E
-20000 N
-20500 N
-21000 N
-21500 N
-22000 N
N
W
E
S
CMGT
[0,350]
[350,700]
[700,1000]
[1000,1500]
[1500,2000]
[2000,3000]
[3000,4500]
[4500,14300]
CMGT Zone 240m Block DA
Estimates vs Sample data

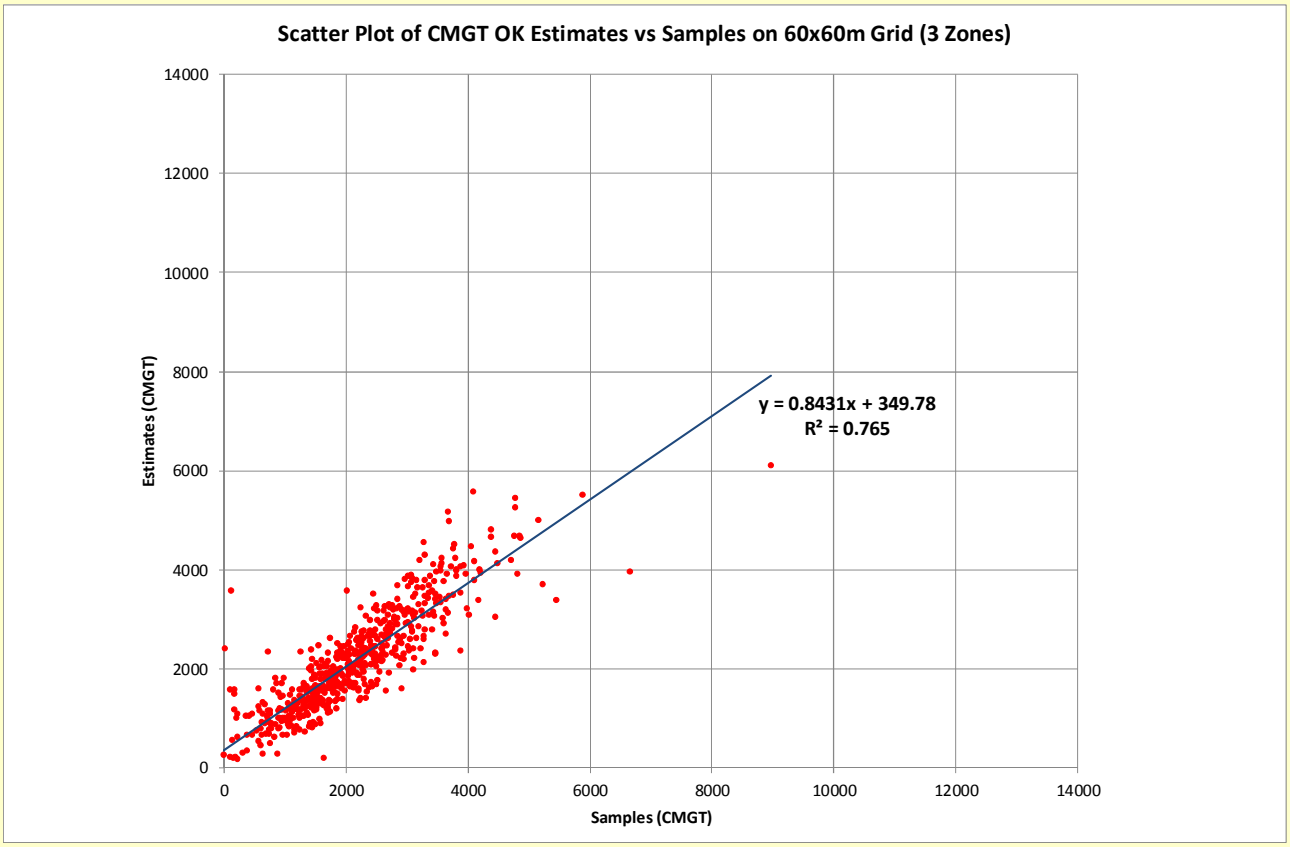
Appendix G: Scatter Plots

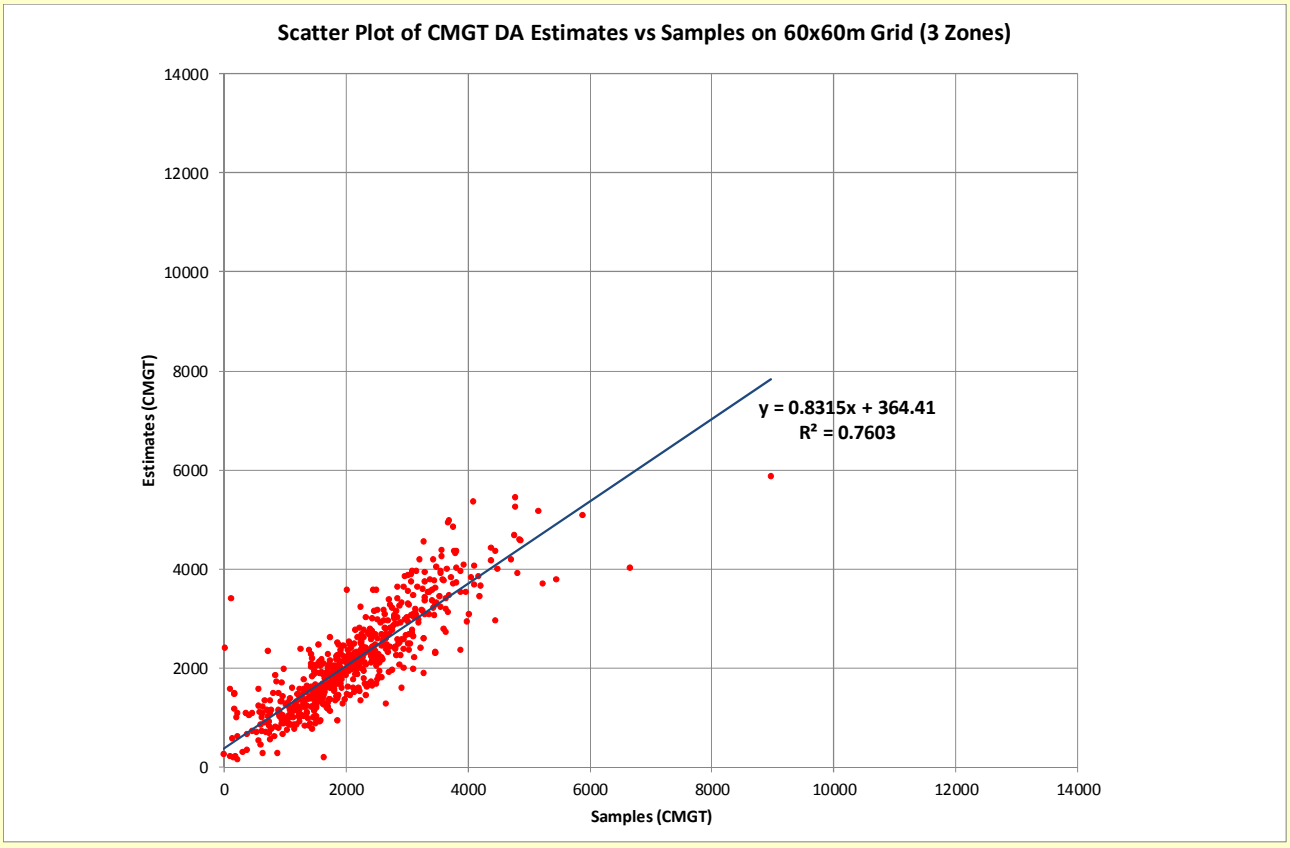

Scatter Plot of CMGT OK Estimates vs Samples on 60x60m Grid (All Zones)
Estimates (CMGT)
Samples (CMGT)
y = 0.8793x + 280.83
R² = 0.7973
0
2000
4000
6000
8000
10000
12000
14000

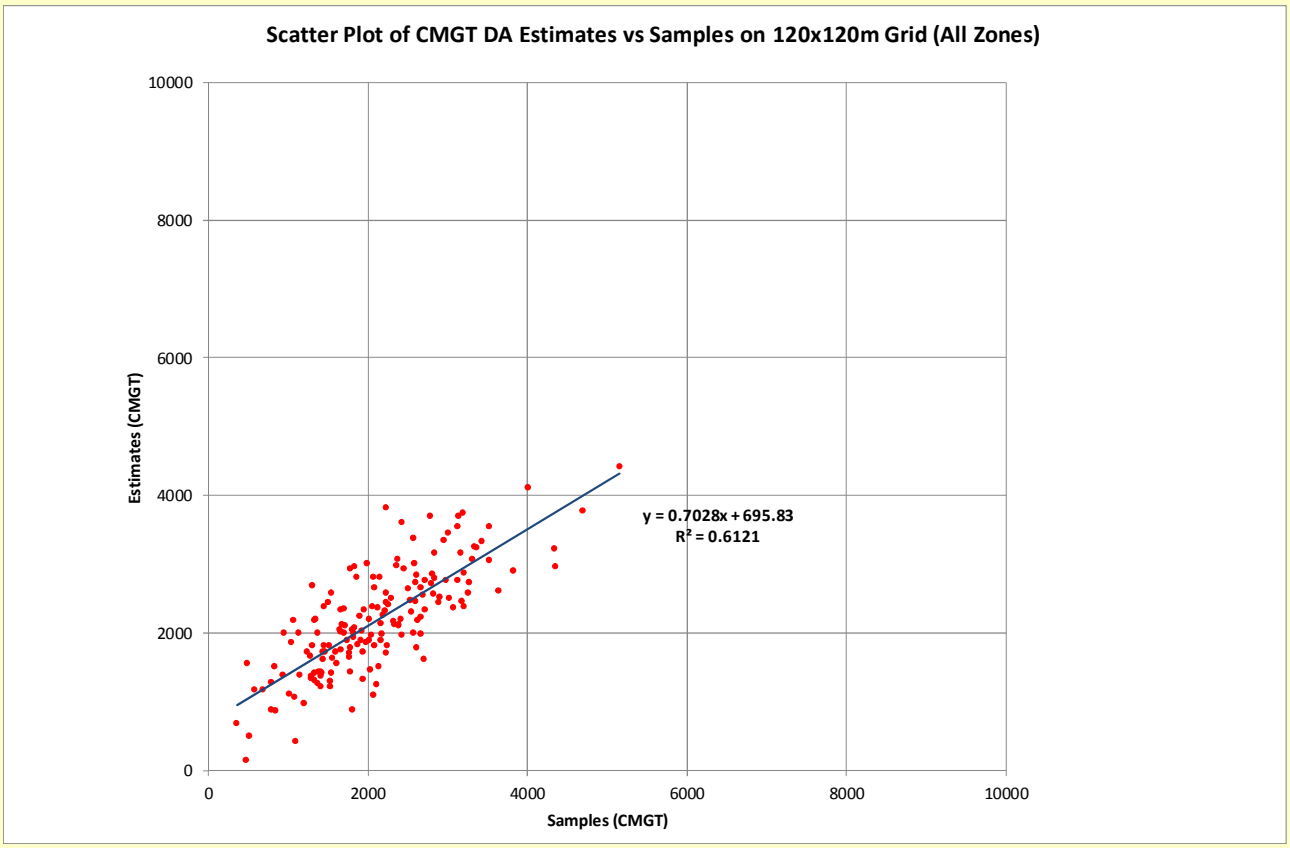

Scatter Plot of CMGT DA Estimates vs Samples on 60x60m Grid (All Zones)
Estimates (CMGT)
Samples (CMGT)
y = 0.876x + 284.68
R² = 0.7966
0
2000
4000
6000
8000
10000
12000
14000

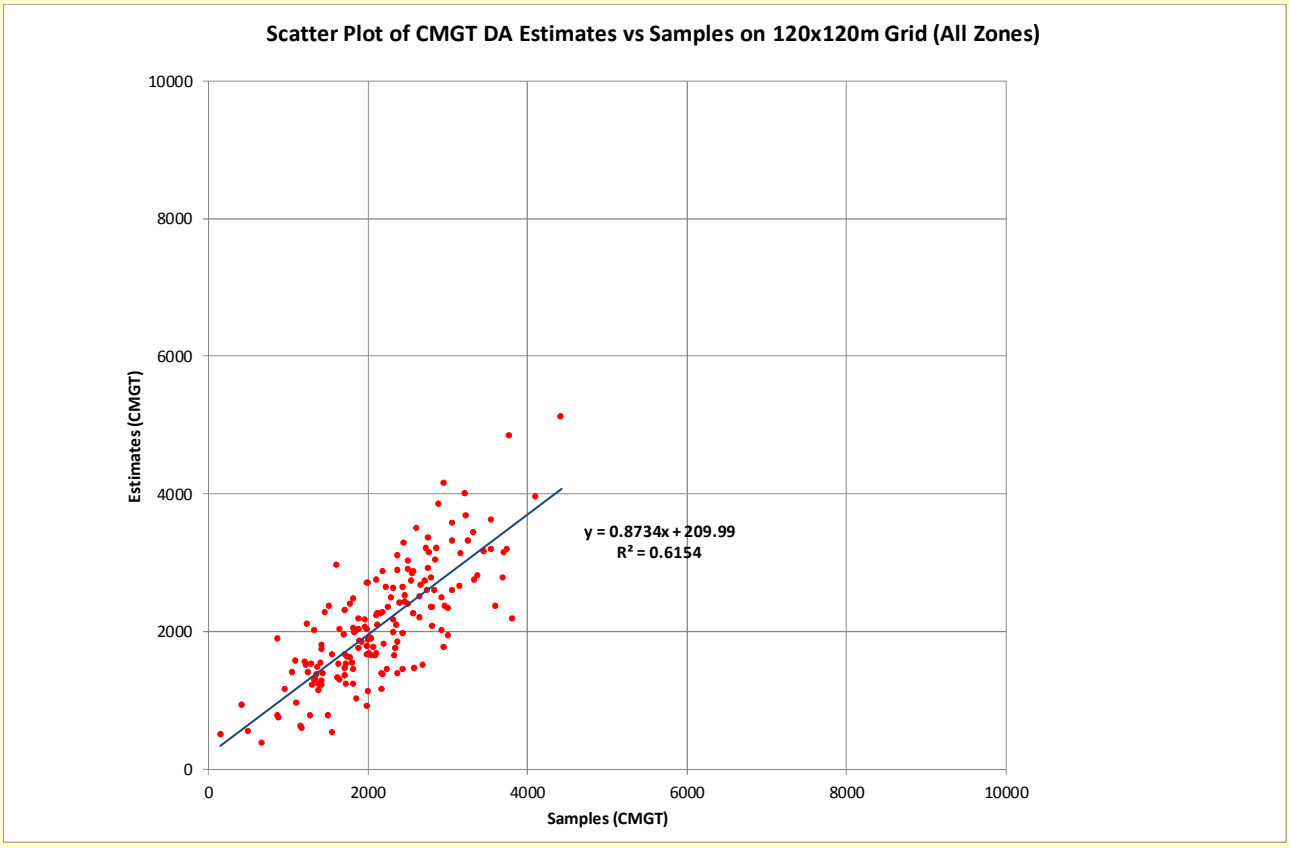

Scatter Plot of CMGT OK Estimates vs Samples on 60x60m Grid (2 Zones)
Estimates (CMGT)
Samples (CMGT)
y = 0.8574x + 320.45
R² = 0.7696
0
2000
4000
6000
8000
10000
12000
14000

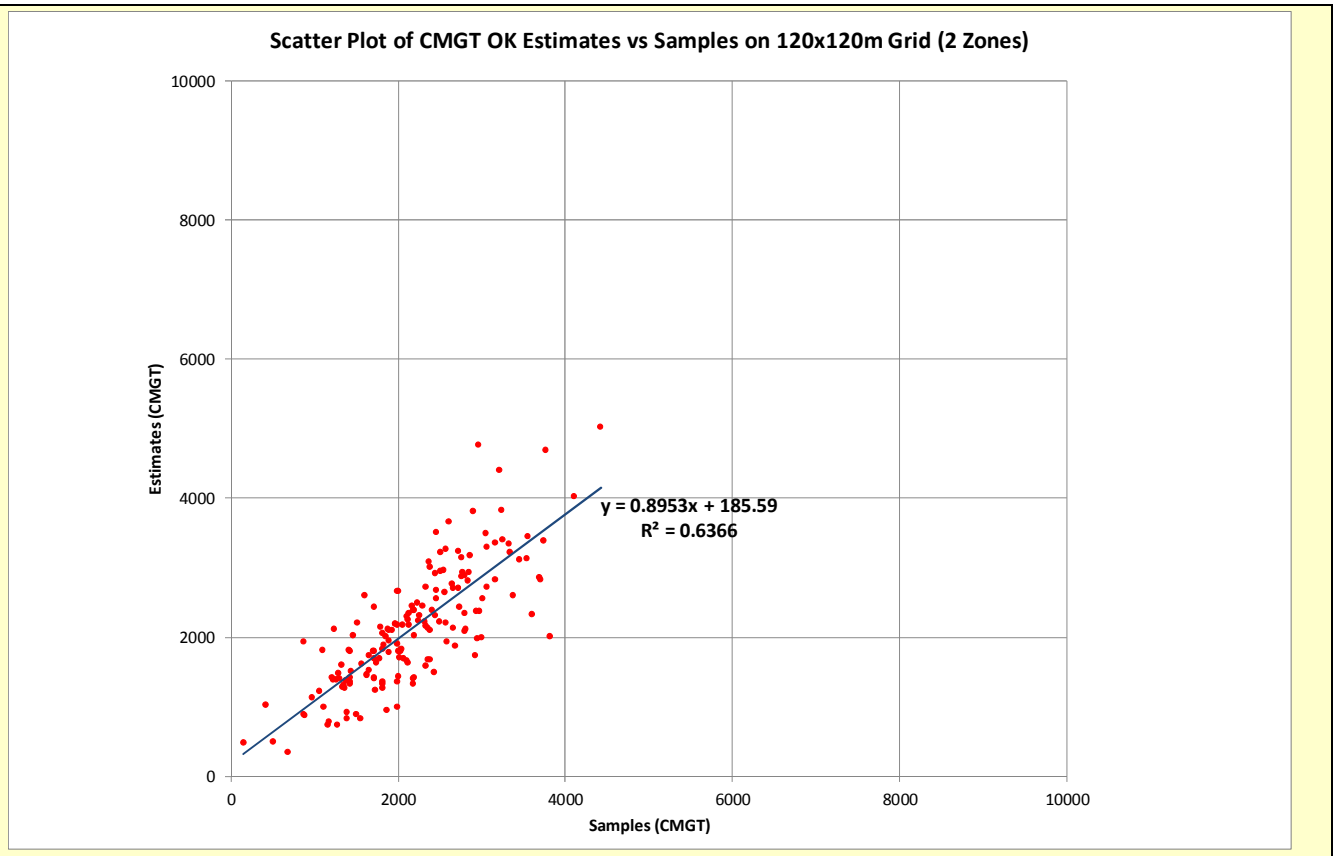

Scatter Plot of CMGT DA Estimates vs Samples on 60x60m Grid (2 Zones)
Estimates (CMGT)
Samples (CMGT)
y = 0.8437x + 339.09
R² = 0.7628
0
2000
4000
6000
8000
10000
12000
14000

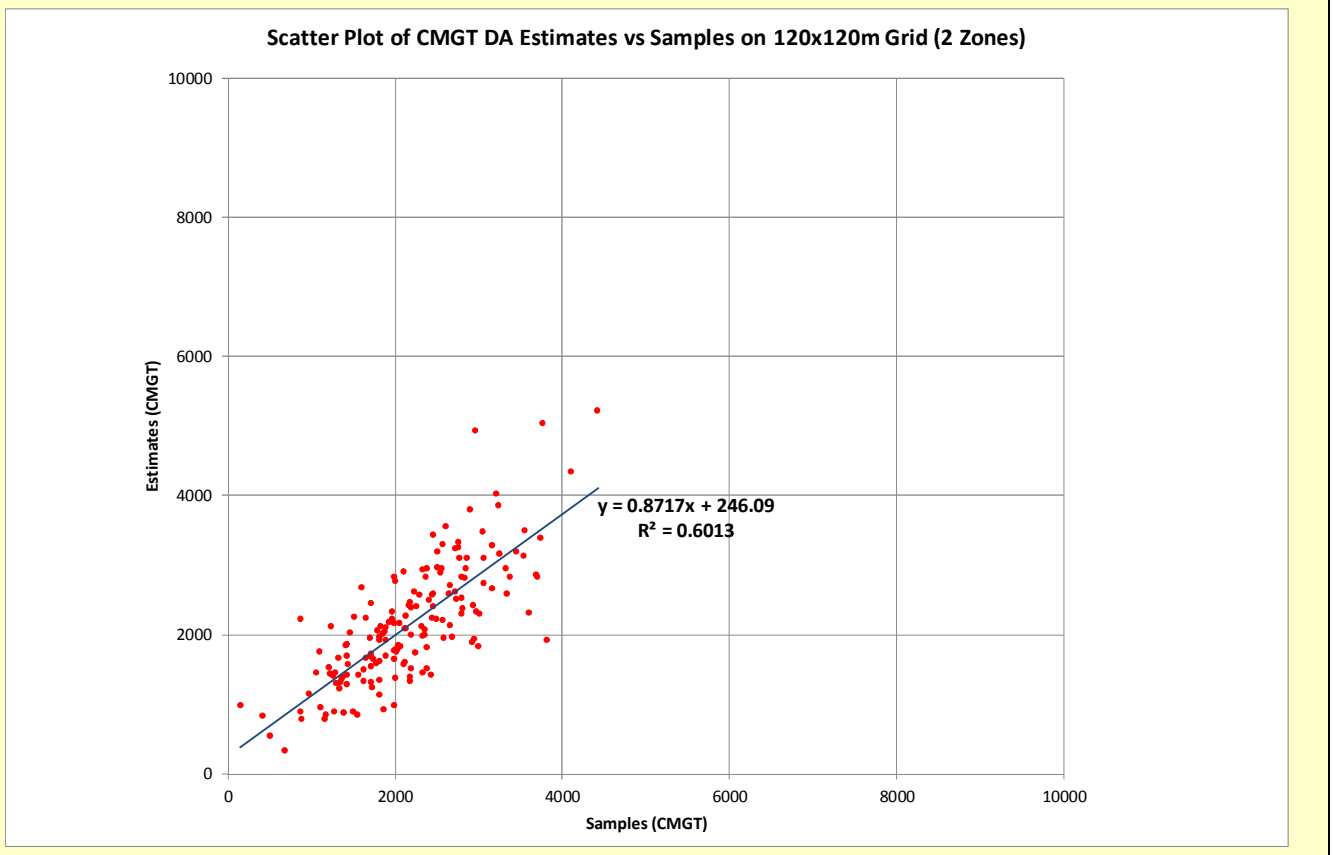

Scatter Plot of CMGT OK Estimates vs Samples on 60x60m Grid (3 Zones)
y = 0.8431x + 349.78
R² = 0.765
Estimates (CMGT)
Samples (CMGT)
0
2000
4000
6000
8000
10000
12000
14000

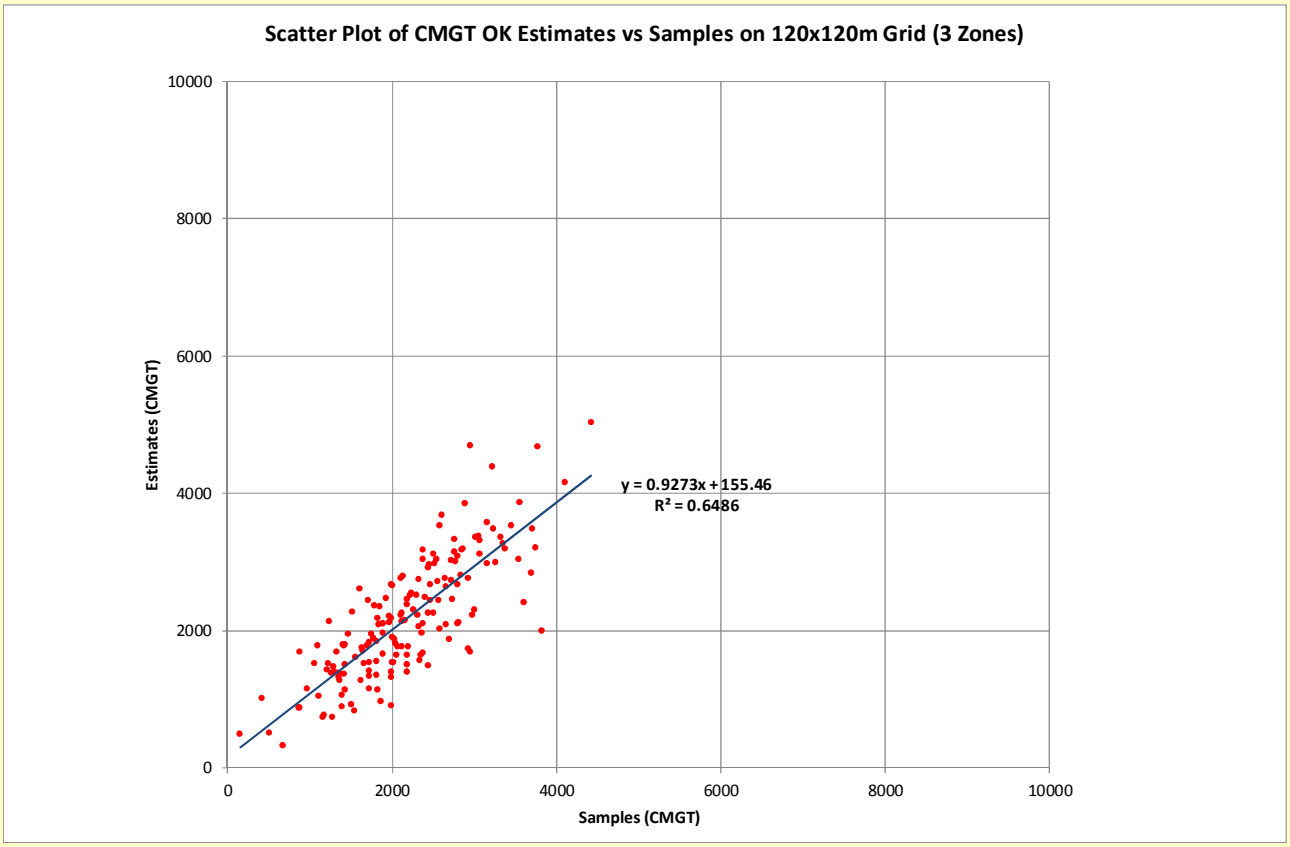

Scatter Plot of CMGT DA Estimates vs Samples on 60x60m Grid (3 Zones)
y = 0.8315x + 364.41
R² = 0.7603
Estimates (CMGT)
Samples (CMGT)
0
2000
4000
6000
8000
10000
12000
14000

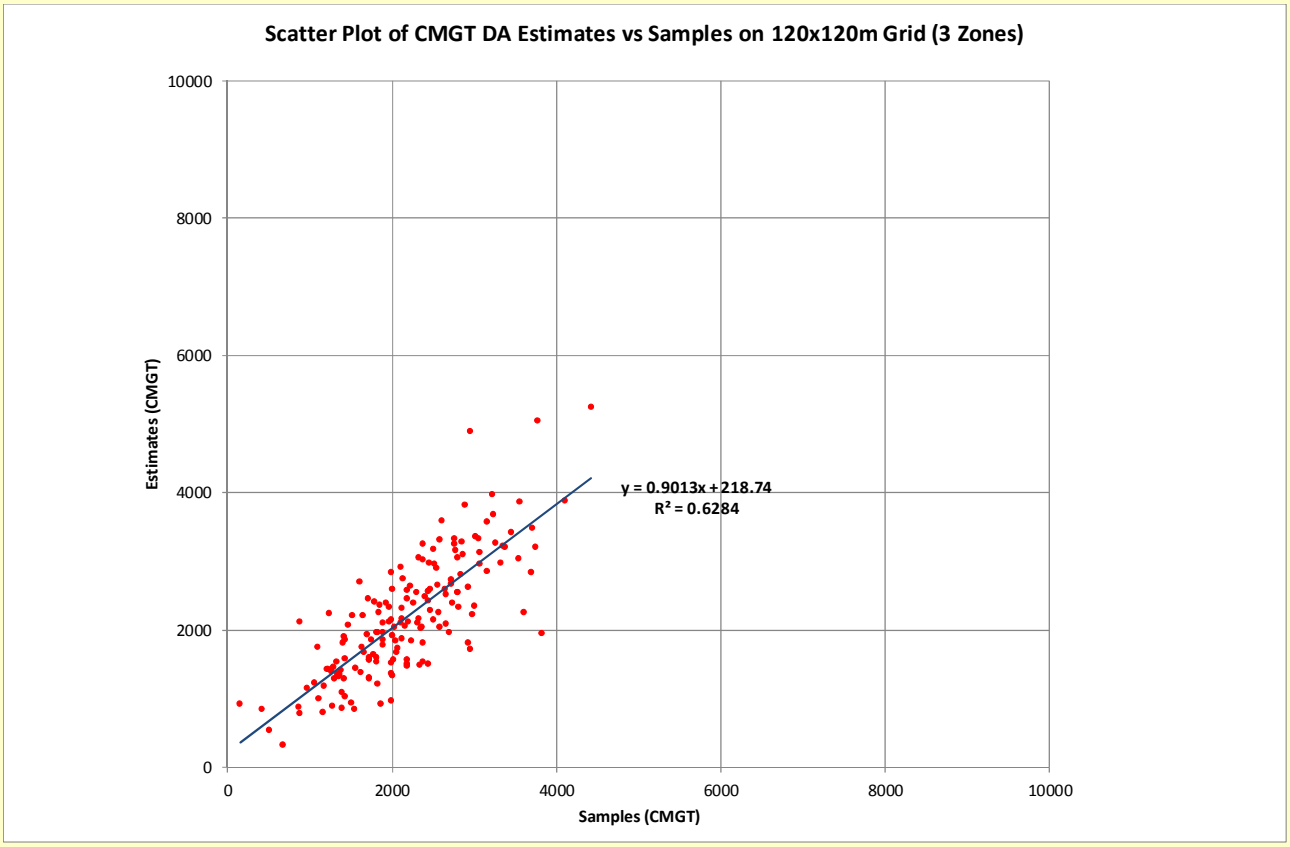

Scatter Plot of CMGT DA Estimates vs Samples on 120x120m Grid (All Zones)
Estimates (CMGT)
Samples (CMGT)
y = 0.7028x + 695.83
R² = 0.6121
0
2000
4000
6000
8000
10000

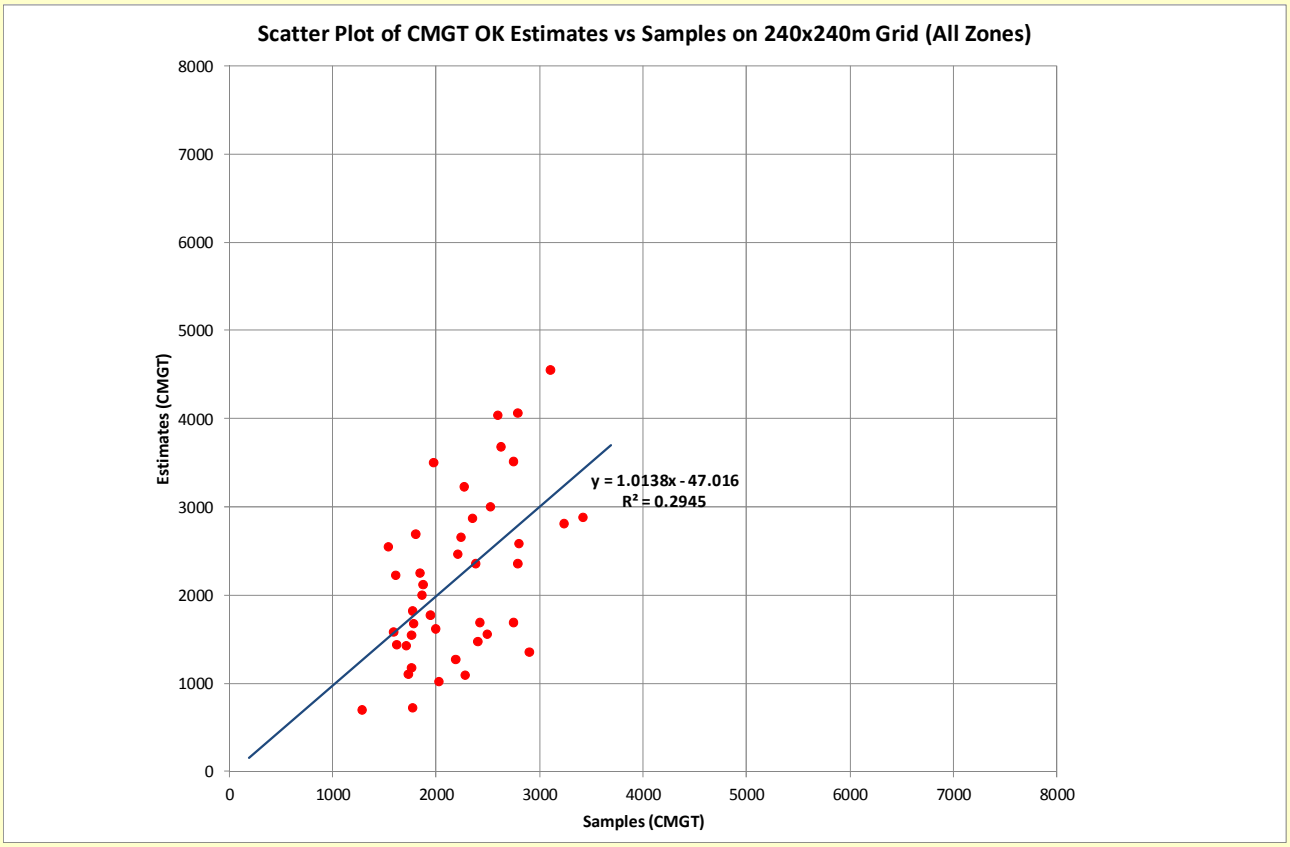

Scatter Plot of CMGT DA Estimates vs Samples on 120x120m Grid (All Zones)
Estimates (CMGT)
Samples (CMGT)
y = 0.8734x + 209.99
R² = 0.6154
0
2000
4000
6000
8000
10000

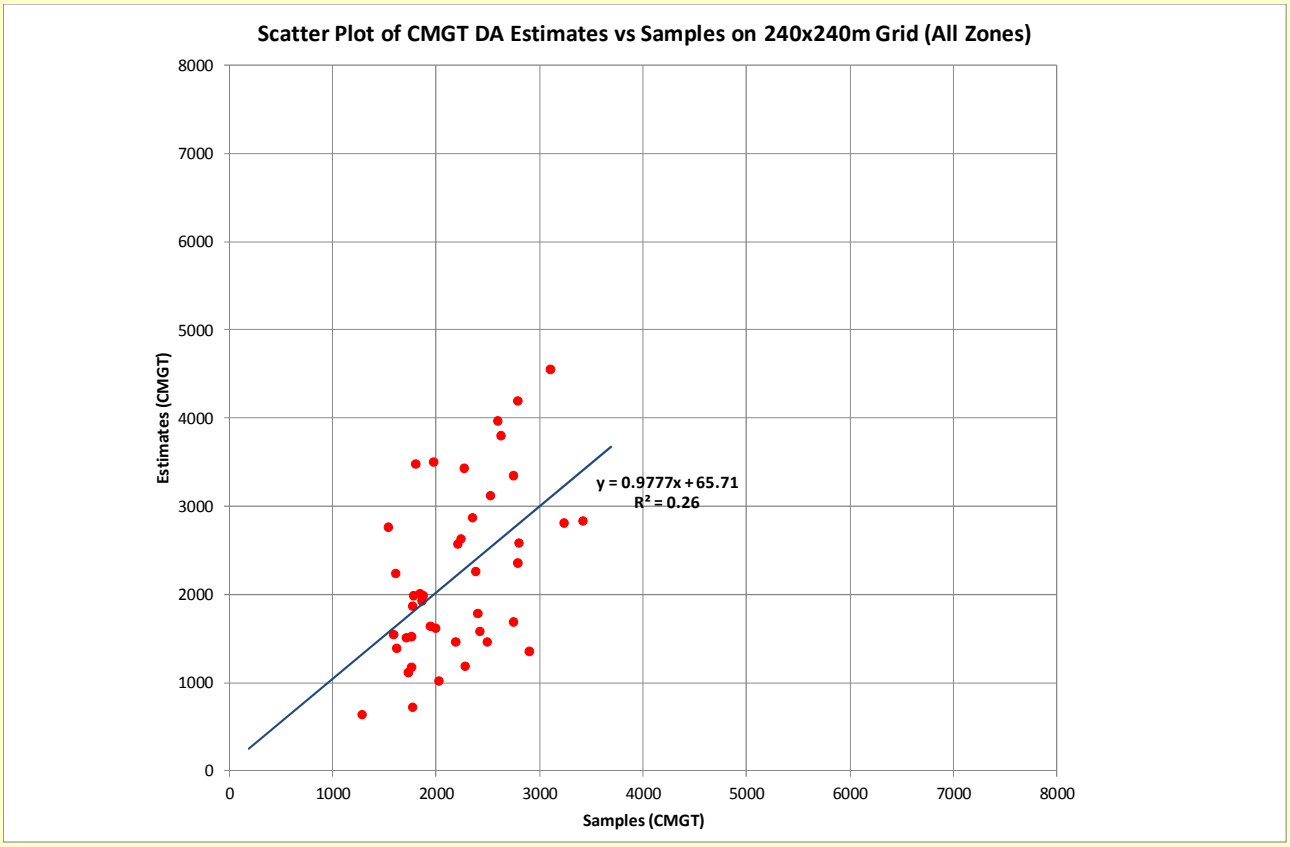

Scatter Plot of CMGT OK Estimates vs Samples on 120x120m Grid (2 Zones)
Estimates (CMGT)
Samples (CMGT)
y = 0.8953x + 185.59
R² = 0.6366
0
2000
4000
6000
8000
10000

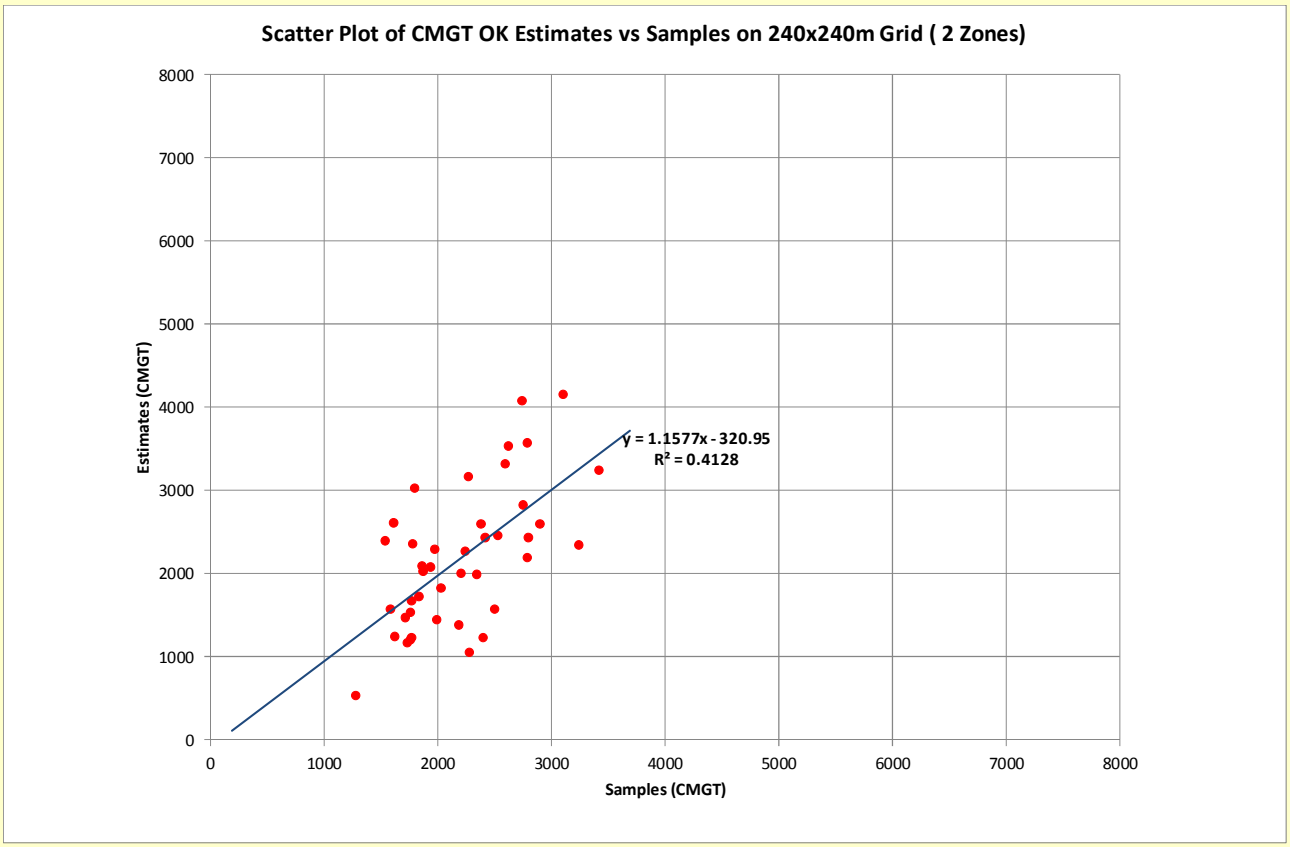

Scatter Plot of CMGT DA Estimates vs Samples on 120x120m Grid (2 Zones)
Estimates (CMGT)
Samples (CMGT)
y = 0.8717x + 246.09
R² = 0.6013
0
2000
4000
6000
8000
10000

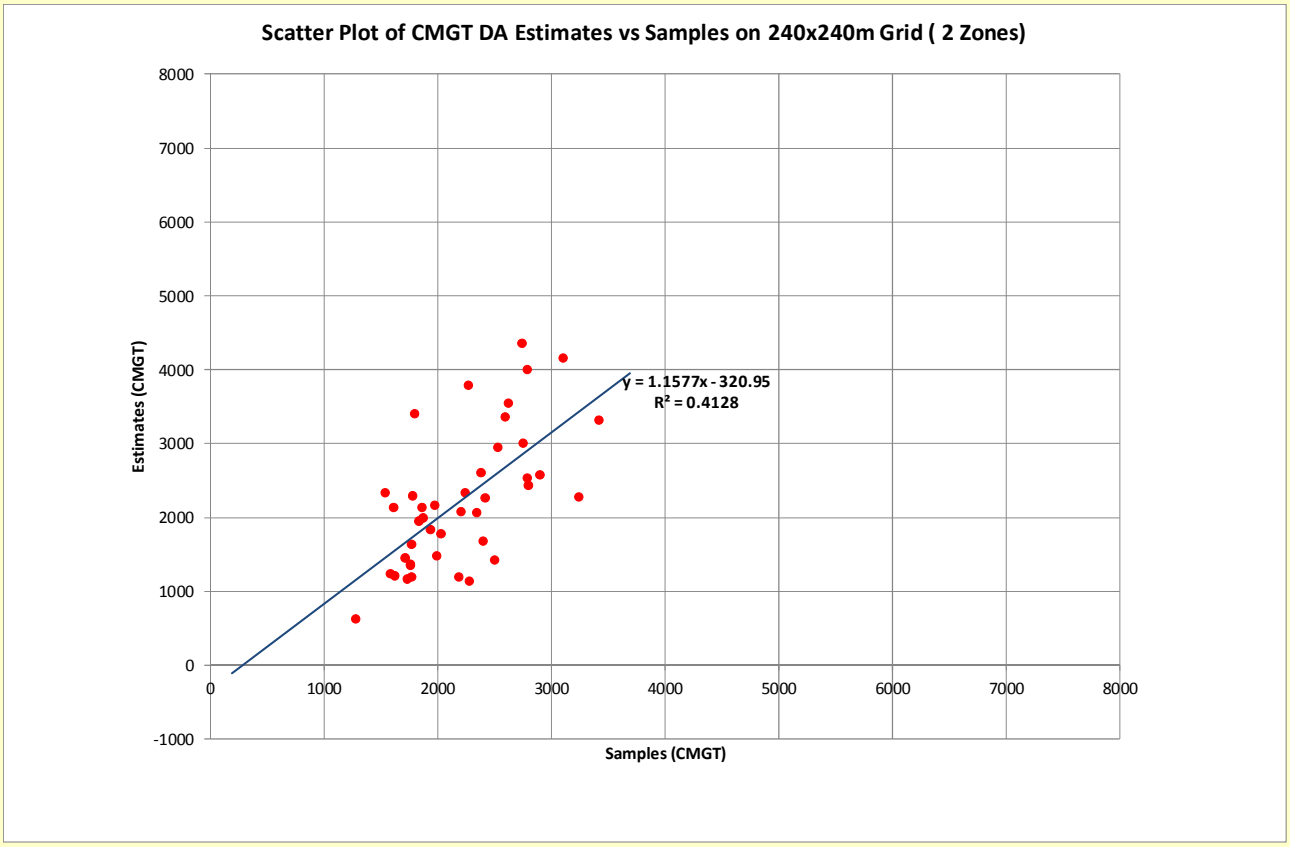

Scatter Plot of CMGT OK Estimates vs Samples on 120x120m Grid (3 Zones)
Estimates (CMGT)
Samples (CMGT)
y = 0.9273x + 155.46
R² = 0.6486
0
2000
4000
6000
8000
10000

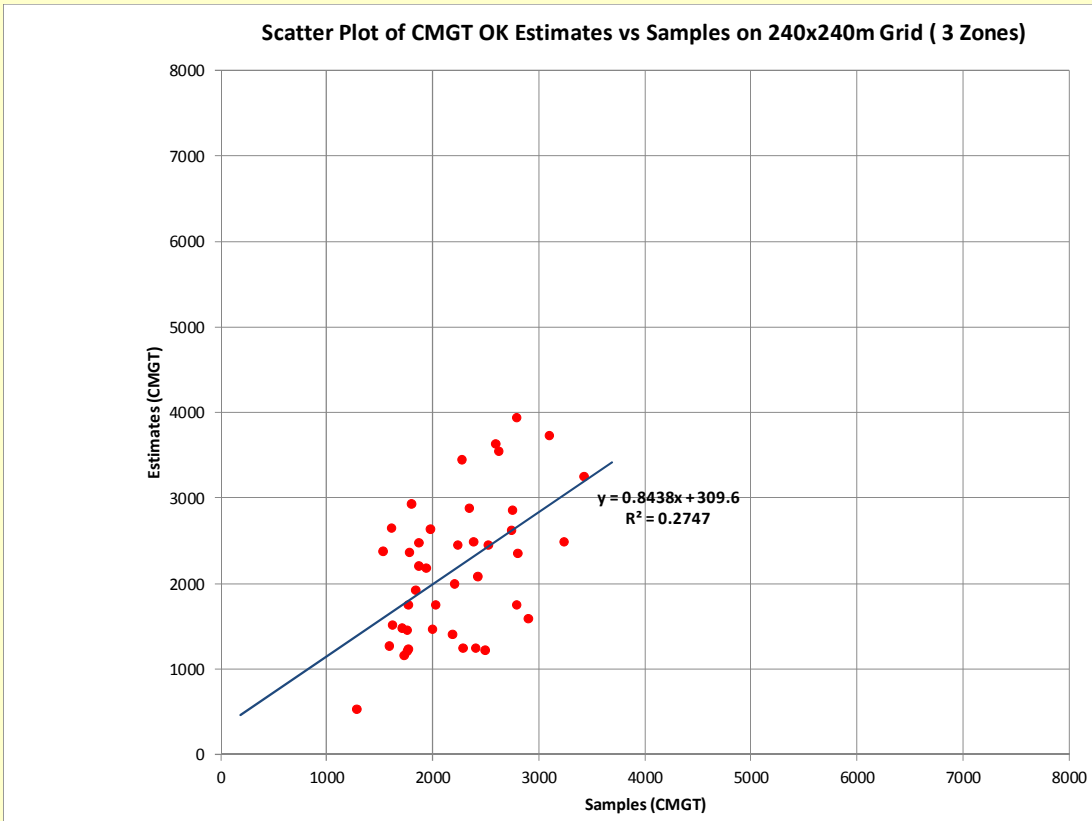

Scatter Plot of CMGT DA Estimates vs Samples on 120x120m Grid (3 Zones)
Estimates (CMGT)
Samples (CMGT)
y = 0.9013x + 218.74
R² = 0.6284
0
2000
4000
6000
8000
10000

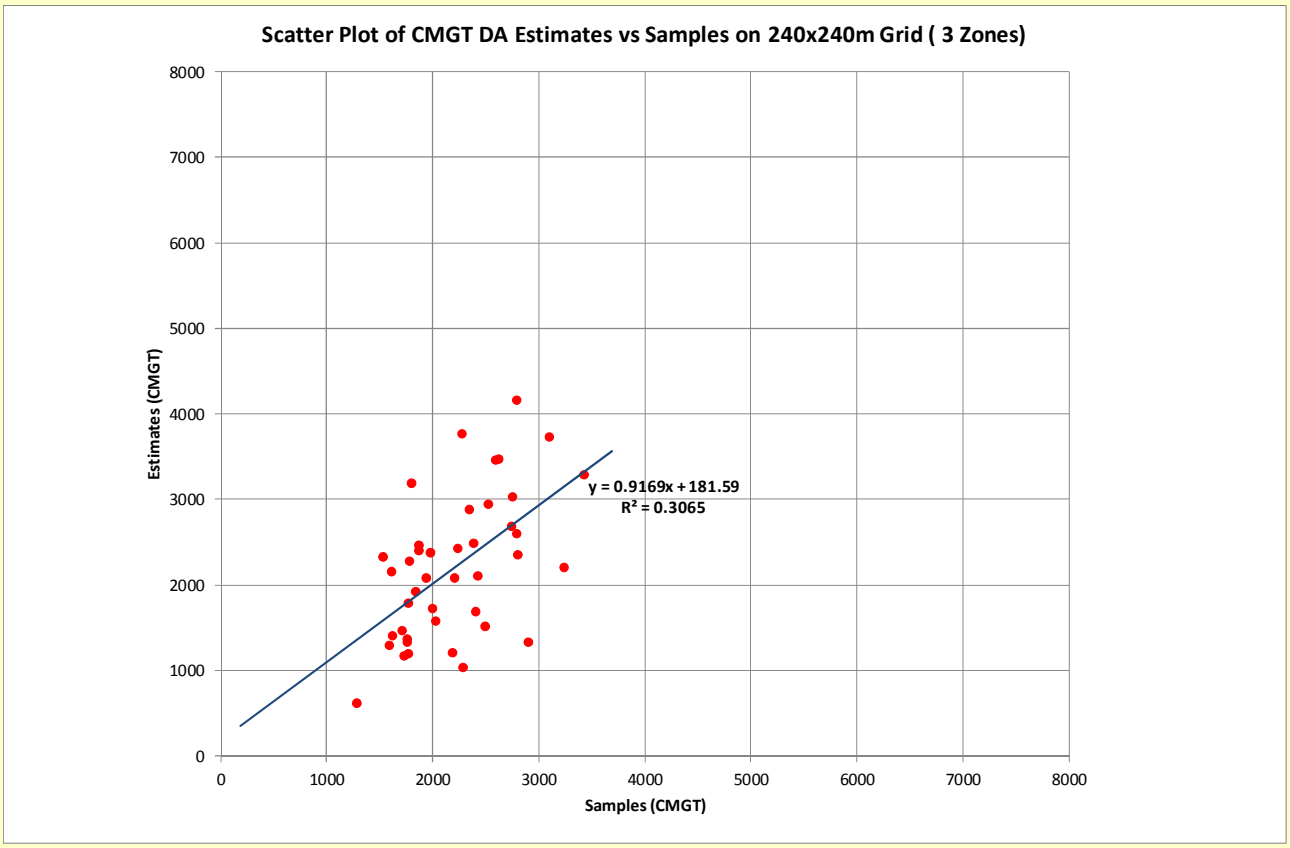

Scatter Plot of CMGT OK Estimates vs Samples on 240x240m Grid (All Zones)
Estimates (CMGT)
Samples (CMGT)
y = 1.0138x - 47.016
R² = 0.2945
0
1000
2000
3000
4000
5000
6000
7000
8000


Scatter Plot of CMGT DA Estimates vs Samples on 240x240m Grid (All Zones)
Estimates (CMGT)
Samples (CMGT)
y = 0.9777x + 65.71
R² = 0.26
0
1000
2000
3000
4000
5000
6000
7000
8000


Scatter Plot of CMGT OK Estimates vs Samples on 240x240m Grid (2 Zones)
Estimates (CMGT)
Samples (CMGT)
y = 1.1577x - 320.95
R² = 0.4128
0
1000
2000
3000
4000
5000
6000
7000
8000


Scatter Plot of CMGT DA Estimates vs Samples on 240x240m Grid (2 Zones)
Estimates (CMGT)
Samples (CMGT)
y = 1.1577x - 320.95
R² = 0.4128
-1000
0
1000
2000
3000
4000
5000
6000
7000
8000


Scatter Plot of CMGT OK Estimates vs Samples on 240x240m Grid (3 Zones)
Estimates (CMGT)
Samples (CMGT)
y = 0.8438x + 309.6
R² = 0.2747
0
1000
2000
3000
4000
5000
6000
7000
8000


Scatter Plot of CMGT DA Estimates vs Samples on 240x240m Grid (3 Zones)
Estimates (CMGT)
Samples (CMGT)
y = 0.9169x + 181.59
R² = 0.3065
0
1000
2000
3000
4000
5000
6000
7000
8000