Federation University Australia

Mine Evaluation Optimisation

PhD candidate: Igor Grigoryev

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy
in the School of Science, Engineering and Information Technology

December 2, 2019

Declaration of Authorship

I, Igor Grigoryev, declare that this thesis titled, "Mine Evaluation Optimisation" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at Federation University Australia.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at Federation University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: I. Grigoryev I. Grigoryev

Date: December 2, 2019

Abstract

The definition of a mineral resource during exploration is a fundamental part of lease evaluation, which establishes the fair market value of the entire asset being explored in the open market. Since exact prediction of grades between sampled points is not currently possible by conventional methods, an exact agreement between predicted and actual grades will nearly always contain some error. These errors affect the evaluation of resources so impacting on characterisation of risks, financial projections and decisions about whether it is necessary to carry on with the further phases or not.

The knowledge about minerals below the surface, even when it is based upon extensive geophysical analysis and drilling, is often too fragmentary to indicate with assurance where to drill, how deep to drill and what can be expected. Thus, the exploration team knows only the density of the rock and the grade along the core.

The purpose of this study is to improve the process of resource evaluation in the exploration stage by increasing prediction accuracy and making an alternative assessment about the spatial characteristics of gold mineralisation. There is significant industrial interest in finding alternatives which may speed up the drilling phase, identify anomalies, worthwhile targets and help in establishing fair market value.

Recent developments in nonconvex optimisation and high-dimensional statistics have led to the idea that some engineering problems such as predicting gold variability at the exploration stage can be solved with the application of clusterwise linear and penalised maximum likelihood regression techniques.

This thesis attempts to solve the distribution of the mineralisation in the underlying geology using clusterwise linear regression and convex Least Absolute Shrinkage and Selection Operator (LASSO) techniques. The two presented optimisation techniques compute predictive solutions within a domain using physical data provided directly from drillholes. The decision-support techniques attempt a useful compromise between the traditional and recently introduced methods in optimisation and regression analysis that are developed to improve exploration targeting and to predict the gold occurrences at previously unsampled locations.

Acknowledgements

I would like to acknowledge the special help received from many people during the course. First and foremost, special mention goes to my supervisors, Michael Tuck and Adil Bagirov. My PhD has been an amazing experience and I thank them wholeheartedly, not only for their academic support and constructive critique. I am grateful to the School of Mathematical and Physical Science of University of Technology Sydney for collaboration in research. My thanks are extended to Dr. E. Hajilarov for his assistance in various points of the thesis. I would also like to thank R. Schodde from MinEx Consulting for permission to use his long-term study, which have helped to bring direct targeting strategy into better focus. I thank Pr. A. Kruger from the Centre for Informatics and Applied Optimisation for his great support in my studies. I am grateful to the external reviewers for extensive comments on an earlier version of this thesis.

Igor Grigoryev
School of Science, Engineering and Information Technology
Mt. Helen December 2, 2019

Contents

List of Figures

List of Tables

Listings

Notations

$\mathbb{A}$	data set
A, B	matrices
β	vector of coefficients
$\beta*$	oracle
β_0, β_1	logistic regression coefficients
E	mathematical expectation
$\mathbb{R}$	real number
$\mathbb{R}^n$	n-dimensional Euclidean space
f	objective function
$\mathbb{N}$	set of natural numbers
ℓ_1	LASSO
ℓ_2	ridge regression
λ	ridge penalty
x, y, z	vectors
x_i	i-th component of vector x
$\|x\|$	norm of x in $\mathbb{R}^n$, $\left(x^T x\right)^{\frac{1}{2}}$
$H\left(p, \alpha\right)$	hyperplane
S, U	sets
$\nabla f\left(x\right)$	gradient of function f at x
$\xi \in \partial\left(x\right) f$	subgradient of function f at x
$\tilde{f}_k\left(x\right)$	piecewise quadratic model of function f at x
$\mathbf{X}$	$N \times p$ matrix of observed predictors, where $\mathbf{X_j}, j = 1, ..., k$ are columns of $\mathbf{X}$
$\gamma(h)$	semivariogram value at separation h
$Z(x)$	random variable at place x
ANN	artificial neural networks
CAD	computer-aided design and drafting (CADD)
CLR	Clusterwise Linear Regression
DGM	Discrete Gradient Method
i.i.d.	independent identically distributed
GIS	geographic information system
`lambda`	the sequence. Default= 0, `glmnet` chooses its own sequence
LASSO	least absolute shrinkage and selection operator
LOO	leave-one-out cross-validation
MSE	mean squared error
HDS	high-dimensional statistics
`nfolds`	number of folds, by default= 10
ppm	one part per million =1 g/t; 3ppm=0.088 Oz/t
RC	regression coefficient
RSS	residual sum of squares
RMSE	regression mean-squared error (measure of performance)
TS1, TS2	test sets (TS1=Test set 1, TS2=Test set 2)
UML	Unsupervised Mashine Learning
WLSE	weighted least squares estimation

Chapter 1

Introduction

This thesis is devoted to one of the most intriguing mysteries for humankind since ancient times - the prediction of mineral occurrence in the crust of the earth. Predicting the total amount of worthwhile elements at depth beneath barren rock cover and the manner, in which they are distributed has been one of the greatest challenges during all the history of mining.

Any mining project begins with the exploration stage, which at the start of mine development includes an approximation of the economic returns expected. Despite exploration not guaranteeing that mining will occur in the area being explored, without exploration, there would be no new resources for future mining. The *Mining Act* 1992 regards exploration and the mineral extraction as part of the overall mining process. In Victoria, mineral exploration and mining activities are regulated under the *The Mineral Resources (Sustainable Development) Act* 1990 (Vic)

A key point in the evaluation remains the proper description and design of a conceptual model, upon which mine design and net present value projections can be built. The evaluation of the economic potential in the early stages of exploration is very important because it is used to determine the sale price of a project, set the relative contributions for each of the partners in any Joint Venture and confirm that the benefits exceed the cost of the exploration.

The geological settings of some deposits, such as iron ore, are often more clearly understood and relatively easy to evaluate (Selley et al., 2005; Geoscience Australia). However, the distribution of gold in more complex geological structures, such as a high-nugget, vein-type deposits is more difficult to evaluate (JORC 2012, s.11). The results obtained often give rise to divergent opinions and scientific debates.

The history of civilisation has been in part shaped by the discovery and retrieval of gold. Gold exploration is risky and a marginally economic endeavour, which takes time, persistence, innovation, intuition and significant investment. However, nothing is more rewarding in the mining industry than discovering a big gold-bearing formation.

Australia still holds a strong competitive position in gold mining and the economics of exploration are usually positive. However, surveys such as that conducted by Mackenzie & Doggett (1992) have shown that the return from mine development and production in some cases is only $1 Million greater than the associated cost of finding and delineating an economic deposit.

A recent survey by Schodde (2017a) shows that the average discovery cost is A$70/oz and is rising by A$10/oz per decade and most recent discoveries have been of small size and low value. World-class gold resources, known as Tier 1 deposits are rare and only 2-3 are found each year. None were found in Australia in the last decade.

Despite most Australian gold exploration projects are not worth much at the early stages, exploration requires significant funds from investors, who are ready to take a

risk. However, foreign investors are reluctant to put a high value on gold exploration projects, because they have limited access to the data necessary to realise the exploration potential.

Mineral deposits are commonly overvalued and rarely undervalued. The benchmarking studies carried out by MineEx Consulting from 2004 to 2017 reports that the Australian valuers tend to over-value projects by 40% and the factors most associated with incorrect evaluation are gold market price fluctuations and incorrect ore reserve estimations.

The price behaviour of gold usually differs from most other metals. The price of gold does not usually change significantly. However, inaccurate drill-core database interpretation presents challenges to accurately predicting the profit of gold mining operations over the long-term. The history of mining activity shows that more mines fail because of ore reserve errors than any other reason. To avoid problems associated with financial expectations, it is very important to predict as precisely as possible the continuity of the mineralisation and the amount of dilution within the explored areas.

Until the 1930s gold discovery was more a matter of luck than scientific prediction (Bradley, 1925; Richert, 1929). There was no way of confirming the true values of the area away from the measured points. Interpolation was a type of spatial prediction.

Articles published in the 1930s and 1940s (Richert, 1929; Solderberg, 1930) raised a number of questions about the applicability of interpolation for predicting ore reserves. Attempts at rigorous statistical testing were not generally attempted until the 1950s. From year to year, gold exploration was becoming more and more expensive as new deposits became harder to find (R. Schodde in Schodde, 2017a reports the increase of the average discovery cost from A$10/oz in 1980 to A$70/oz in 2017). Many studies and technical papers reveal the deficiencies in understanding the variability. This affected the capacity to accurately predict how the minerals were distributed.

This situation changed substantially, thanks to geostatistics introduced by Matheron (1963) and refined by David (1977). Since then, a number of probabilistic and algorithmic approaches play an important role in applications used for evaluation.

With less than complete spatial information, the industry is often forced to make assumptions as to what material is in the ground and how it will perform when extracted and processed. In the past, some mining companies used unconventional methods in the process of prediction, including human sensing and intuition. In 1980, Uri Geller, famous for his psychokinetic demonstrations, predicted the existence of large coal deposit in South Africa near the Zimbabwe border. In 1886 Geller successfully predicted locations of gold and diamonds in previously unexplored areas of the Solomon Islands.

1.1 Exploratory Drilling

A generalised approach to modelling a deposit has three distinct components.

1. The geometry of the geological units that formed and host the ore body.
2. The attribute characterisation in terms of assays of all commodities to be mined.
3. The value model in terms of economic extraction.

The key element of this approach is the ability to predict all these components prior to actual mine development. This makes the prediction of gold distribution fundamental to the financial success of the entire project.

However, deposit geometric properties and physical characteristics are never completely known, but are projected from the exploration drilling (Code of Practice 2012). The

results of the exploration drilling program are the most useful data source for preliminary resource evaluation and provide the basic input for defining recoverable reserves.

Diamond drilling performed from the surface or underground provides explorers with an incomplete information of the underlying mineral content, the presence or absence of the gold. Drilling results often reveal hidden details and change the whole direction of the project. In this thesis, the data taken from cores is used as a basic input for data processing, representation of drillhole traces, statistical visualisation and 3-D modelling.

The main objectives of exploratory core-drilling are

1. Obtaining core-samples for testing, gathering geophysical, geochemical information.
2. Finding anomalies and determining the geometry of the mineralisation.
3. Creating a model of the deposit.

This process is described in Pan et al. (1993) and Knight et al. (2007) as ore delineation.

Drilling results delineate exploration targets and lead to at an Inferred Mineral Resource in conjunction with JORC 2012. This drilling data contains hidden information that is used to guide additional exploration drilling activities.

Each stage of exploration is designed to get to the next decision point, that is, whether or not to continue exploration, based on results of the previous stage. As delineation information accumulates, the lateral and vertical limits of the resource are determined and the grades to mineralisation are attributed. At the Pre-Feasibility phase (JORC 2012, s 29), Mineral Resources become Ore Reserves.

Another important aspect is very high cost of exploration drilling.

Exploration for minerals remains a risky endeavour because of its great uncertainties (Selley et al., 2005). The cost of drilling can be very high, but the payoff from a success can be great. The cost of a too tight drilling grid could affect investment decisions. Costs of drilling, in areas of varying geographic remoteness, can range between A$100 to A$200 per metre (the cost is rarely disclosed, but can be found in some ASX Announcements). Considering that it would be too expensive to obtain data over all the areas of the exploration lease, the number of drillholes is usually kept to a minimum.

Annels & Dominy (2003) demonstrate that reliable evaluation depends on the number and quality of samples taken during drilling. Insufficient patterns of drill holes over the exploration lease can also affect evaluation. Indeed, from an engineering viewpoint, errors in data interpretation begin with sampling, continue through data input, data manipulation and the incorrect interpretation of the results.

Taking a very limited amount of geologic data and extrapolating the interpreted data to cover unknown areas is a real problem in making accurate predictions. Some data is always lost during core recovery. This is due to equipment failure, drilling bit deviation and improper core handling and recording. Lost, badly recovered or badly recorded core data always affects the accuracy of prediction and can potentially lead to the wrong classification of entire deposit.

1.2 What is forecasted?

What changes are expected for the prediction of gold resources? Professor T. McCuaig from the Centre for Exploration Targeting suggests that the industry will be forced to undergo a dramatic shift away from exploiting large, ever lower grade resources to

seek high quality discoveries. This shift will require innovations in understanding the underlying geology and its largest scale footprints.

Professor Richard Schodde from MineEx Consulting in his long-term outlook for gold 2017–2057 (Schodde, 2017b) points out that to maintain long-term production at 2017 levels, the industry needs to double its current discovery performance. It has been forecasted that by 2032, half of all Australia's gold production will come from deposits yet to be discovered. In forty years time almost all of Australia's future gold production will come from exploration successes.

To predict and discover economical gold deposits, the industry needs to develop techniques that can provide greater precision and in less time. The banking community requires a better understanding of the dynamic nature of the brownfields search space–the potential of the near-mine region.

The latest stock market news updates and business newspapers report that speeding up the exploration process by one year will result in a 0.2–0.3 Million ounces of gold per year increase in Australia's mine production in the medium term. Another established fact (Schodde, 2017b) is that the effectiveness of exploration is declining. Thus, the in-situ value of gold discovered per exploration dollar spent is currently only A$11 compared to A$23 in 2000 and A$57 in 1980-1990.

Recent sensitivity studies carried out by Centre for Exploration Targeting and MinEX Consulting showed that each additional dollar spent on exploration generates an extra A$11.40 in revenue. For the gold industry to maintain production at current levels in the longer term, it will either need to

1. Double the amount spent on exploration, or
2. Double its discovery performance

To improve the performance multiple approaches to prospectivity and analysis will be required to understand the geological uncertainty.

To do this, extra research is required into the development of new detection methods and software packages. These are the tools that employ automation and the power of the computer to improve the accuracy of prediction. It is also recommended that traditional evaluation tools be viewed critically, as alternatives may exist, which in some cases are better suited to achieving the required objectives.

1.3 Methods for analysing the evaluation information

A number of the studies have recognised the complexity of the spatial prediction problem at hand and the difficulty of solving the problem with established systems. The prediction of values at unmeasured points are currently made from the regionalised variable theory (Matheron, 1963) and based on the concept of a random function. Many researchers have pointed at a number of disadvantages of that theory (Philip & Watson, 1986; Armstrong & Champigny, 1989; Rossi & Deutsch, 2014). In some cases, these may generate errors in the assumptions (in section 2.4.1) of the geometry of a deposit.

Any existing method used in resource evaluation has drawbacks to a greater or lesser extent. Geostatistics is not an exception and some techniques have been criticised for a number of disadvantages. Journel & Huijbregts (1978) and Rossi & Deutsch (2014) have specified several assumptions, which in some cases are not the best options in spatial applications.

Studies by Cox & Miller (1965), Shurtz (1985, 1994), Myers (1989), Merks (1992), Henley & Watson (1998) and Chilès & Delfiner (2012) consider the assumptions of continuity, the smoothness and differentiability of the regionalised functions of the conventional tools such as kriging, which has the highest sensitivity to the number of variables and the quality of samples involved in computation. The authors discuss disadvantages of assumptions and argue that some methods are incapable of providing a measure of the magnitude of the errors. Some authors admit that the estimate of grade is misleading unless it provides not only an average value but also a probability for the volume considered.

Both business and science agree that the errors in evaluation can never be eliminated due to the heterogeneity of rock mass and a limited number of available drilling samples. Geological uncertainty often defies attempts to describe the distribution of mineralisation correctly. Therefore, the inaccuracies and risks that arise from uncertainty must be also well understood and the resulting consequences carefully managed.

1.4 Prediction as an optimisation problem

Some statisticians (e.g. Galit, 2010) describe a predictive model as a method that produces predictions, regardless of its underlying approach: parametric, nonparametric, data mining, statistical models, etc.. In many cases it is possible to formulate the prediction problem as an optimisation problem (Hahens & Doveton 1991; Kinnicutt et al., 1994). Optimisation methods can also be applicable to solving some problems related to the prediction of gold in unexplored areas.

The application of optimisation approaches for resource definition is not a new topic in the literature. A number of techniques that are capable of predicting the existence of mineral occurrences can be found in Clarici et al. (1993), Yama & Lineberry (1999), Melkumyan & Ramos (2011) and Rodrigues et al. (2015). Collectively, all these studies outline the critical role of optimisation in predicting spatial variation below the surface and provide important insights into finding an optimal solution to the problem.

There is a large volume of mathematical studies (e.g. Converse, 1970; Polyak, 1987; Fletcher, 1987; Rockafellar, 1994) describing the role of optimisation techniques in recovering hidden patterns in data. While the traditional optimization based predictive models applied to mine evaluation are usually convex, many important problems, such as optimization models of the distribution of gold mineralisation and consequently, evaluation of exploration/ mine potential involve nonconvex functions.

1.5 The Goals of The Thesis

A number of publicly available online ASX reports suggest that evaluation and investment in gold mining remain the realm of the investor able to evaluate the relative probability of a project. A number of studies referenced in this chapter identify the need for accuracy and precision of evaluation in early stages of exploration, which is the key to decide whether it is necessary to carry on the further phases or not.

There has been considerable debate in the literature on the applicability of some conventional techniques when applied to complex gold deposits. The spatial and temporal distribution of gold deposits is extremely heterogeneous (Groves et al., 2005).

Pragmatically, the main problems with the evaluation of such deposits stem from the lack of spatial continuity, highly erratic distribution of minerals, limited drilling data and quality of sampling.

It is fair to suggest that some extra research is required to develop alternative techniques that are less sensitive to heterogeneity and can make more accurate prediction of the existence of anomalies and the deposit potential. The purpose and long-term goal of this thesis is improving the accuracy of mine evaluation by increasing prediction accuracy and facilitating alternative prediction of gold mineralisation with the application of modern methodologies.

1.6 Summary of Chapter

Resource estimates are not precise calculations, being dependent on the interpretation of limited information and the available sampling results.

There is a significant industrial interest in finding alternatives which satisfy the industry requirements (section 1.2), may speed up the drilling phase, identify worthwhile targets and promising footprints, helping in establishing the fair market value of a gold deposit in the open market.

Apart from those listed above, two more important aspects have to be addressed.

At the exploration phase, the knowledge about the gold occurrence, even when it is based upon extensive geophysical analysis, is often too fragmentary to indicate with assurance where to drill, how deep to drill and what can be expected when you drill. The exploration team knows only the density of the rock and the grade along the core.

An important constraint on mine evaluation is the data visualisation, in particular, the availability of high-resolution predictive maps. The disadvantage of modern 3-D commercial software packages with integrated statistical tools is the lack of transparency. The underlying source codes by which resource estimation is completed are not accessible and the manuals have limited information. As a result, in some cases it is hard to understand what is being done just based on the available instructions.

Two alternative and conceptually distinct methods are developed through this thesis:

Method 1 is a machine learning technique, more specifically, Clusterwise Linear Regression (CLR). This model helps to achieve relatively high accuracy of predicting local instances of gold in unexplored areas and beyond the exploration lease. Each predicted structure has its own set of x, y, z coordinates and an attribute, such as gold grade in Au g/t. The predicted points can suggest to the exploration team the presence or absence or extent of the gold in lease.

Method 2 refers to predictive mapping with the application of high dimensional statistics. The method is based on the Least Absolute Shrinkage and Selection Operator (LASSO) methodology and is used to generate maps of unexplored areas using a g/m^3 metrics. The method allows the automation of high-dimensional analysis, provides the minimisation of generalisation errors in dimension reduction while maximising predictive performance. Method 2 is a platform for creating further, more complex techniques.

The two proposed methods are viewed as alternatives to the traditional techniques used in the preparation of the Exploration Target Statement and Public Reporting of Exploration Results. It will be demonstrated that in a general sense, that the proposed methods in some cases are less affected by the limitations existing in some conventional methods.

Organisation of the Thesis

This thesis is concerned with developing two regression methods, which can improve prediction of the probability of mineral occurrence below the surface, which, in turn, leads to decisions on the amount of funds to be deployed in exploration and mining. There is arguably a more important need to research, design predictive methods to ensure that the developed methods be in applied to solving real-world tasks. The structure of this thesis is as follows:

Following this introduction, chapter 2 provides a general introduction to basic mining and investment terminology, discusses common techniques used in evaluation, surveys their advantages and disadvantages within the context of engineering evaluation as well as general description of the commonly used spatial techniques. The factors that affect the performance of these techniques are discussed. Finally, the role of machine learning in solving spatial variation is reviewed.

Chapter 3 is the most creative part of the thesis. The problem of predicting hidden data is reformulated into a nonconvex optimisation problem. The design of the overall configuration required for recovering hidden data is presented and the implementation of the algorithms is discussed. The chapter ends with the presentation of a complete design of the direct detection predictive system based on machine learning techniques.

Chapter 4 presents real-world geological data with a complex structure and having a large number of internal boundaries. A data preparation process to convert raw drilling data into a mathematical format is discussed. A series of tests are conducted to study and compare the behaviour of predictive sets obtained during multiple program runs. 3-D renders with mathematical content are used as a way to deliver locally optimal solutions within visualised contexts.

Chapter 5 looks at the research problem from a different view and presents an alternative, convex predictive method based on ℓ_1 penalties empowered by `glmnet` R-package, which sets the regression coefficients to exactly zero. The method selects predictors, shrinks their coefficients toward zero relative to the least-squares estimates. A number of predictive maps generated by the method are introduced.

Important findings are drawn and future directions for research are discussed in the final Chapter 6 of the thesis.

The following important matters are presented in greater details in the Appendices:

Appendix A: data transformation from raw to output code.
Appendix B: UML system code.
Appendix C: LASSO `R` code.
Appendix D: Matrices of Kernel densities for predictive prototypes.
Appendix E: Output format of prediction

Chapter 2

Literature Review

The economic viability of a project depends on reliable exploration targeting, which requires the prediction of the presence or the absence or the extent of mineralisation in a deposit. This chapter provides a general introduction to basic exploration, mining and investment terminology, discusses common techniques used in evaluation and surveys, their advantages and disadvantages within the context of engineering evaluation. The chapter identifies from all the information which could be considered what is actually necessary to develop predictive methods that can help improve exploration targeting and predict the probability of gold occurrence, which lead to more accurate evaluation of a deposit.

2.1 Terminology and Definitions

Australia has one of the world's largest share of economic gold reserves and is also a large contributor to the global gold trade (A Minerals Council of Australia). Since gold was discovered in the 1850s, gold mining has been a key contributor to Australia's wealth and prosperity. In 2017, gold in Australia produced 9.7 Moz of gold, generated $A16 billion in revenue and employed 28,000 people at 76 gold operations (Lucas, 2018)

Australian mining legislation differs substantially from the mining laws of other countries. Each of the States and Territories has its own legislation governing the exploration and extraction of minerals.

In Queensland, for example, depending on the project, resource permits and authorities for exploration and mining can granted under the *Environmental Protection Act* 1994 (Qld) and *Mineral Resources (Sustainable Development) (Mineral Industries) Interim Regulations* 2018 once a resource proponent has completed an *Environmental Impact Statement*.

Exploration is the process of searching for deposits of minerals in the ground. The purpose of exploration is to locate areas where mineral resources may be present, to establish the quality and quantity of those resources and to investigate the viability of extracting the resource. In the Northern Territory, exploration is regarded as the amount of work necessary for the discovery and assessment of the potential value of minerals in the title area (*Mineral Titles Act* 2018 (NT)).

According to Australia Minerals, one of the first-order challenges for the industry is being able to reliably predict at the deposit-scale the depth to prospective basement and the character of the overlying cover.

In NSW, before exploring for minerals, an explorer must first obtain an Exploration Licence under the *Mining Act* 1992, pt. 3. An exploration Lease is usually granted by the State for a period of five years and gives the licence holder the exclusive right to:

- explore minerals within a designated area for the evaluation of the potential for mining in the area.

Exploration licences do not permit mining, nor do they guarantee that a mining lease will be granted.

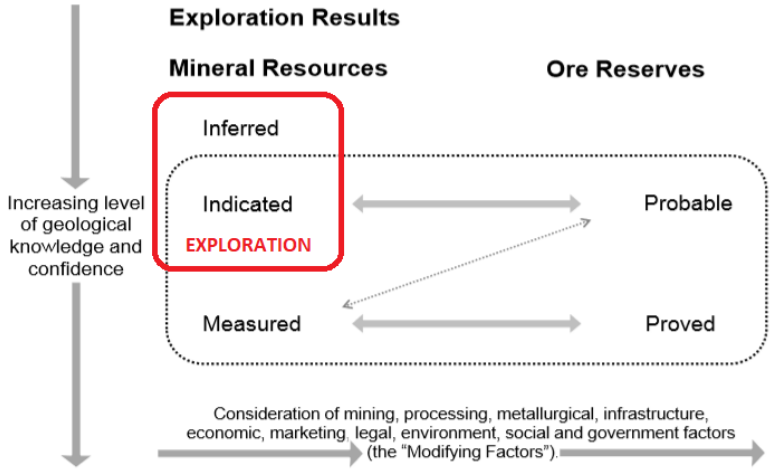

Figure 2.1. Exploration defines Inferred or Indicated Resources, in some cases Resources may be upgraded as to Measured. Adapted from JORC 2012

It is necessary to consider the four essential stages for the development of mines in Australia:

1. The initial exploration stage: a range of methods are used, and predictions made prior to validation by drilling.
2. The further detailed exploration stage: exploratory drilling defines mineral resources (see Figure 2.1) with reasonable prospects for being developed into standalone mines or neighbouring mines. This thesis focuses on the predictions made in this stage.
3. The mining stage (not covered in this thesis), which begins with Pre-Feasibility and Feasibility studies.
4. Opening a surface or underground mine and exploiting the raw material. Mineralisation may not be classified as a reserve unless the determination has been made that the mineralisation could be economically and legally produced or extracted at the time the reserve determination is made.

A Mineral Resource (see JORC 2012, s. 20) is a concentration of minerals some of which are valuable, but most are not. The amount of valuable minerals that a Mineral Resource contains is stated as % or a grade which is the average amount of grams of gold per ton.

Figure 2.1 shows the relationship between Exploration Results, Mineral Resources and Ore Reserves. Exploration deals with Resources, which are sub-divided into three categories based on the level of confidence in the results of testing (see JORC 2012, ss. 21–23)

1. Inferred Resources – the lowest level of confidence, but the data provides enough information to imply the results.
2. Indicated Resources – high level of confidence but not as high as with Measured Resources. The data provides enough information to assume the results.
3. Measured Resources – the highest level of confidence and the data provides enough information to confirm the results.

Only Indicated and Measured Resources can be upgraded to Reserves. To upgrade Resources to Reserves a Competent Person (JORC 2012, s. 4) either a Pre-Feasibility or Feasibility study.

2.1.1 Exploration: investor's perspective

There are a large number of private and publicly listed mining companies and individuals who are actively looking to acquire or invest in exploration projects. At any stage, exploration projects can be bought, sold, optioned and joint ventured on the basis of their perceived potential for the existence of an economic mineral deposit. From an Australian Securities Exchange (ASX) viewpoint, the value of an exploration property is therefore based on this exploration potential. Those seeking information about an Australian investment opportunity in a resource project may find the following available:

Exploration Target Statement: in the early stages of exploration the quantity of data available is generally not sufficient to allow any reasonable estimates of Resources. In such cases, exploration results could include information on outcrop sampling, assays of drill hole intersections, geochemical results and geophysical survey results, used in the interests of investors (JORC 2012, s 18). Importantly, these results do not form part of a declaration of Mineral Resources.

JORC 2012, s 17 sets out the detailed requirements that must be satisfied when reporting an Exploration Target. Among other things, it requires that any statement referring to potential quantity and grade of the target be expressed as ranges and must include:

a. A detailed explanation of the basis for the statement, including a specific description of the level of exploration activity already completed, and

b. A clarification statement within the same paragraph as the first reference of the Exploration Target in the Public Report, stating that the potential quantity and grade is conceptual in nature, that there has been insufficient exploration to estimate a Mineral Resource and that it is uncertain if further exploration will result in the estimation of a Mineral Resource.

JORC 2012, s 17 states that ranges of tonnages and grades must be represented only as approximations. For example: from 0.3 Moz to 0.6 Moz of gold, or 1.1 Moz–1.8 Moz.

Under the ASX disclosure standards, if the Exploration Target Statement includes information relating to ranges of tonnages and grades, these must also be represented as approximations and supported by a description of the process used to determine the grade and the tonnage ranges. In Banking terms this is a project that does not have a reported JORC, 2012 Resource, or a Bank Feasibility Study in progress or completed, and it is not in production.

Despite the fact that the Exploration Target Statement is not regarded as completed exploration[1], there are many investors who consider this stage is highly prospective. As for a lot for sale in an auction, such projects are usually denoted as an Exploration Investment Opportunity.

From an engineering perspective, early drilling results may delineate the Exploration Target and can lead to an Inferred Mineral Resource. The drilling data contains hidden information that can and should be used to inform and optimise additional drilling and sampling activities.

[1] ASX Announcements, Tables 1. and 2. available at https://www.asx.com.au/

Public Reporting of Exploration Results: JORC 2012, s 18 allows this stage to include data that might be of use to investors, but which does not form part of a declaration of Mineral Resources or Ore Reserves. Reports must include relevant information such as exploration context, type and method of sampling, relevant sample intervals and locations, distribution, dimensions and the relative location of all relevant assay data, methods of analysis, and data aggregation methods. If true widths of mineralisation are not reported, an appropriate qualification must be included in the Public Report. Where assay and analytical results are reported, they must be reported using one of the following methods, selected as the most appropriate by the Competent Person:

1. either by listing all results, along with sample intervals. Examples of this approach can be found in Reporting of Exploration Results, s. 2)
2. or by reporting average grades of mineralised zones, indicating clearly how the grades were calculated.

More information on Mineral Exploration reporting can be found at WAMEX.

JORC 2012, s 19, states that Public Reports of Exploration Results must contain sufficient information to allow a considered and balanced judgement of their significance. Reports must include relevant information such as exploration context, type and method of sampling, relevant sample intervals and locations, distribution, dimensions and relative location of all relevant assay data, methods of analysis, data aggregation methods ... on any of the other criteria (JORC 2012, p.26, Table 1) that are material to an assessment.

Many public reports using method 1 can be found in open sources. Reports providing details following method 2. are rare.

A question an investor may ask is *"how much drilling do we have to do before we can reasonably expect a defined Indicated Resource?"*.

JORC 2012 s 19, states that to predict geological continuity, maximum likelihood-based statistical confidence intervals are developed. Interval estimates for the % of drillholes having values of the variable of interest above a benchmark value are set. A level of confidence that the interval contains the % of drillholes having values of the variable of interest above the benchmark are also set for the expected grades in low-and high-grade areas. Modelling of grade continuity assists in determining the amount of drilling needed to update Inferred to Indicated Resource.

One of the principles governing the operation and application of the JORC 2012 is Materiality. This principle requires that a Public Report of Exploration Results contains all the relevant information that investors would reasonably require, and reasonably expect to find in the report, for the purpose of making a reasoned and balanced judgement regarding the Exploration Results being reported.

At any stage of gold exploration, investors are usually unhappy with the amount of information provided. There is never enough drilling information. Apart from the information that is publicly announced, experienced investors are always looking for external information that can increase confidence in the investment.

At this stage, the starting bid and fair open market value of an exploration project is hard to determine. The Banking community evaluates a project as a Possible, Probable or Speculative buy.

Reporting of Mineral Resources: JORC 2012, s 20 presumes that the location, quantity, grade, continuity and other geological characteristics of a Mineral Resource are known, estimated or interpreted from specific geological evidence and knowledge, including sampling. Mineral Resources satisfy the requirement that there are reasonable prospects for eventual economic extraction, regardless of the Inferred, Indicated and Measured

categories. From an investor point of view such a project is a project that has a reported JORC Resource, a completed Bank Feasibility Study and is in production.

Exploration can attract investment at all stages. Before auction, every investor is informed that inferred mineral resources (see Figure 2.1) have a great amount of uncertainty as to their existence and great uncertainty as to their economic and legal feasibility. Therefore. it can not be assumed that all or any part of an inferred mineral resource will ever be upgraded to a higher category such as indicated or measured. Prospective buyers are cautioned not to assume that mineral Resource will ever be converted into Reserves.

From 1 December 2014, JORC 2012, s 29, requires Ore Reserves to be defined by studies at Pre-Feasibility or Feasibility level as appropriate. *The Mineral Resources (Sustainable Development) Act* 1990 (Vic) provides that subject to a mineral resource being identified, it is expected that the holder of an exploration licence will work towards preparation of a mineralisation report and ultimately establishing the mineral resource to at least an inferred standard (see Figure 2.1) within the meaning of JORC 2012.

However, depending on compexity of a project, the estimated inferred mineral resources may not form the basis of Pre-feasibility and Feasibility mining studies, except in rare cases. Investors are usually cautioned not to assume that all or any part of an inferred mineral resource exists or is economically mineable.

2.1.2 Exploration categories

Exploration companies are often broken down into three categories. It is essential to consider these categories that are usually used by the investment and banking communities. More information is available online from Productivity Commission Inquiry Report.

a. Greenfields exploration – relies on the predictive power of ore genesis models to find mineral deposits in previously unexplored areas or in areas where they are not already known to exist.

b. Grassroots exploration – when an explorer has a conceptual idea about where a deposit might be and spends money to see if the mineralisation is really there. Grassroot projects are the riskiest projects in the mining business.

c. Brownfields exploration – explorers look for deposits near or adjacent to an already operating mine. The possibility that existing data might be used makes the risk in brownfield exploration considerably lower than in greenfield exploration. Because the facilities for mining and processing the ore are often already built and paid for, the additional marginal capital cost for processing any new ore is very low.

After having delineated a deposit with drilling, the next step is opening a surface or underground mine and extracting the rock mass by means of the special mining techniques and procedures of mineral processing.

2.1.3 Information disclosure and validation of exploration results

Despite gold is not a critical commodity (Geoscience Australa), detailed information on the verification of the gold distribution is extremely hard to find in open sources. Mining companies provide drilling data and the employed methods recommended in Table 1 of JORC 2012, p.26, but rarely share step-by-step calculations explaining how exactly predictions were made, how uncertainty was managed, how the 3-D wireframes

and vertices of the mineralised domains were digitised and how mineral resource models were calibrated.

In the 1970s, members of some institutions[2] and banks have played an important role in encouraging exploration and mining companies to disclose more meaningful information about ore resources and the applied methods of estimation.

One reason for this is that the literature and technical notes on gold prediction is read and company announcements well monitored by investors and interested persons. As previously mentioned, at the early stages investors have limited access to the data necessary to take an investment decision. But not all. The reality is that some investors informed better than others.

Any leak or reference or inaccurate statement (pessimistic or over-optimistic) may lead to a rise or fall in the trading price of stocks and securities and provoke a query from the ASX.

In some cases, this may also give rise to insider trading allegations where the information was not disclosed to the broader market. Every gold company has some “security" rules – no leaks. However, if requested, mining companies usually provide the actual drilling data and some extracts from official production notes.

Grassroots exploration companies tend to operate privately and competitively, and they are also not eager to reveal the methods they employ to find anomalies or hit worthwhile targets, or how well or badly the employed methods have worked out. Since the average cost of gold discovery is currently A$70/oz (Schodde, 2017a), exploration is financed entirely by investors and once their cash runs out, they will have to rise finance or stop. As a result, the average delay between discovery and mine development in Australia is 13 years.

Another reason is that most of the large gold deposits located close to the surface have already been discovered. There have been a decreasing number of World-Class Tier-1 discoveries globally over the past 20 years. They are extremely rare (reported by Centre for Exploration Targeting and MinEX Consulting).

The practical validity of the reserve estimation technique usually lies in a comparison of estimates with the reconciliation phase which is a comparison of estimated tonnage and grades with real measurements. A French independent software vendor Geovariances defines the term reconciliation as: “checking the validity and robustness of any resource estimate is a mandatory step to ensure models remain current and useful”. The company states that avenues for validation and reconciliation are

1. geostatistical inspection of resource estimates, benchmarking against data inputs.
2. using past and current production data to improve the predictive performance of models.

In Parker (2012) the term reconciliation is defined as a process that allows determination of the ability of a mine to produce the tonnage and grade that were estimated in the ore reserve. Refer to H.M. Parker, for precious metals mine a good annual reconciliation between mine and mill is generally acceptable as 10%.

RSC Mining and Mineral Exploration Services from Perth offers reconciliation of resource model with mine production, and mine production with mineral processing information, which provides key information on the quality of the entire system.

Importantly, the exact reconciliation of Indicated and Inferred Resources (see Figure 2.1) is hardly possible. Detailed studies and reports on geological reconciliation in later

[2] https://www.finsia.com/

stages (versions by Parker and RSC) are rarely published and even when published, too few details are provided to enable a realistic evaluation of the prediction. Detailed comparisons of geostatistical predictions (suggested by Geovariances) with reality are also rarely covered in the literature.

Ideally, to verify prediction, the gold ore extracted from a particular block needs to be closely monitored, a record made of every pass of a truck with continuous checking the grade with at least an x-ray fluorescence analysis and other on-site techniques. Then checking stock piles, conveyors, crushers and ball mills–all against the amount of gold predicted and produced by smelter. This way, however, high cost and technical complexity make exact reconciliation of gold the hardest task.

2.2 Background and Previous Work

The fundamental starting point for mine evaluation is to determine the characteristics of ores to be developed and mined. To obtain these, successful exploration is required to realise the enhanced economic returns offered at the start of mine development. Schodde (2017b) lists reasons why new methods to drive the next wave of discoveries should be found:

1. The decreasing efficiency of finding new deposits.
2. The average cost of a new discovery has increased to $70 per once.
3. A shift of investment from major to junior exploration companies.
4. A reduced rate of discovery of world-class gold deposits Tier 1.
5. A shift of investment from greenfields to brownfields exploration.
6. A reduced investment efficiency for the discovery of economic deposits.

2.2.1 Historic definitions of prediction

As early as 1870 the research into the existence of gold in the ground was kept confidential by fearful authorities. It was believed that gold fever could potentially cause anarchy in the small fledgling British colonies. Several scattered references were made in the mining engineering literature to the need for statistics for predicting gold-bearing ores from chemical assays. Hale (1881), for example, made a suggestion that the gain from the application of statistics in the gold mining and milling business would be large.

In the 1900s the literature mentions a variety of predictive practices such as averaging depth and area (Bradley, 1925), cross-sections (Richert, 1929) and grouping of assay values in blocks (Solderberg, 1930). The results of these practices were usually employed in the control of mining operations, accounting purposes, figuring depletion, calculating depreciation and development charges.

Although mathematical principles were involved, the evaluation of mines was not considered an exact science and wide experience was usually required in the selection of methods and the interpretation of results in order to arrive at near-correct conclusions. The accuracy of the final results, it was believed, would depend to a large extent upon estimator's assumptions, his experience and the soundness of his judgement.

Bradley (1925) and Solderberg (1930) noticed that these differences in judgements can lead to serious problems in the initial evaluation of a deposit when the geologic resource estimates are interpreted to represent ore reserves but later found to be misleading and overestimated. From a modern science perspective, these methods were not mathematically

correct in some instances, but most of them were used in the past and it was assumed that the applied methods gave satisfactory results.

The purchase of a prospect was often dependant on the question of ore extension because it was seldom that the ore was fully developed and the purchase price was almost invariably based on prediction of the distribution. However, there was no way to measure the reliability of the estimation. Solderberg (1930) mentions that a number of modifying factors were often applied to make corrections when the estimate went wrong.

The first published studies on the use of statistical techniques for prediction of ores appeared about 1920s. Wattermeyer (1919) discussed the application of the theory of probability in predicting grades. The benefits of the involvement of statistics into reserve calculation were considered by Harrison & Hanover (1929). The authors suggested the application of graphical extrapolation and interpolation of ores by use of assay plans, projections and cross-sections. A method of statistical grouping and weighting of assays for predictive purpose was proposed in Jakson & Knaebel (1934).

In a number of technical documents in the 1970s, for example in Dowding (1976a), the term geospatial predictive modelling is defined as analytical production of 3-D graphics by using data from drillholes.

A considerable amount of modern literature has been published on spatial prediction. According to Noble (2011) the traditional approach to mine evaluation is based on collection of data, *predicting* spatial properties and physical characteristics of the mineral occurrence, modelling the size, shape and grade of the ore body.

A number of published studies on spatial analysis and GIS systems suggests that most applications are based on interpolation with synonymous with *optimality predicting* in space using observations taken at known nearby locations. The objective of more recent predictive techniques (Henley, 2001; Hengl, 2007; Gaetan & Guyon, 2010) is to assess the *likelihood* of the distribution.

Another term found in the literature is *geospatial predictive modelling* (GPM), which is rooted in the principle that the occurrences of the events being modelled are limited in distribution. Hengl (2007) describes GPM as a process for analysing events through a geospatial filter in order to make statements of the likelihood for event occurrence.

A more recent addition to spatial modelling, the term *optimal prediction* usually refers to ore reserve estimation, the aim of which is a numerical or graphical model (Rossi & Deutsch, 2014) that accurately predicts the grades and tonnage of ores to be extracted.

Some statistical literature on spatial exploratory data analysis, for example, Bivand et al. (2013) link prediction to geostatistical data that could be in principle measured anywhere, but typically comes as measurements at limited a number of observation locations. In a study by Pimpler (2017) on the development of spatial statistical tools and ArcGIS systems, the prediction is viewed as modelling relationships among data variables associated with geographic features.

2.2.2 What level of accuracy of resource estimation is accepted?

JORC 2012, s 25 states that mineral resource estimates are not precise calculations, being dependent on the interpretation of limited information on the location, shape and continuity of the occurrence and on the available sampling results. A crucial question can be asked then: *"What is an acceptable level of resource estimation accuracy?"*

Australian legislation does not mandate or recommends the level of accuracy for estimation. However, the ASX disclosure of key assumptions underpinning mineral

resources and reserves estimates suggests the following levels of estimation accuracy:

- Scoping study (greenfields and early grassroots exploration): a level of accuracy of +/- 30 to 40%.
- Preliminary feasibility study (grassroots exploration and early stages of development): a level of accuracy of +/- 20 to 25%.
- Feasibility study (pre-development and development stage): a level of accuracy of +/- 10 to 15%

2.3 Classification of Methods for Spatial Prediction

To avoid the difference of opinion, the provided below notations and equation are referenced and presented in the forms as they are in the original sources. A good review of thirty-eight spatial interpolation methods can be found in Li & Heap (2008). The methods most frequently cited in the literature are summarised in Table 2.1.

Predicting spatial continuity is generally solved by two methods: interpolation and extrapolation. Interpolation is the mathematical function that gives the exact value for every known control point, so that the surface it defines benefits all the known data. Interpolation performs the task of prediction by estimating the values at unsampled areas using data from observations within the same region. According to Rendu (1994), since there are no ways of confirming the true values of the field away from the control points, interpolation is viewed a type of spatial prediction.

Statistical methods used for resource estimation (in grassroots and brownfields stages) are generally divided into the traditional geometric methods that are done manually on plans or sections and interpolation methods (Sinclair & Blackwell, 2004).

A formal discussion of the state of interpolation requires some notations. Webster & Oliver (2007) admit that nearly all currently known interpolation methods can be represented as weighted averages of sampled data and usually share the same formula:

$$\hat{z}_a\left(x_0\right) = \sum_{i=1}^{n} \lambda_i z(x_i). \tag{2.1}$$

Here $\hat{z}_a$ is the estimated value of the primary variable at the point of interest x_0, z is the observed value at the sampled point x_i, λ_i is the weight assigned to the sample point and n represents the number of sampled points used for the estimation.

One of the most widely used in GIS applications technique is inverse distance , which described in Hengl (2007) as a mechanical method. The term mechanical means that a user is prompted to accept the default parameters suggested by a GIS package.

Inverse distance is a method that fits only the continuous model of spatial variation. In some situations, inverse distance can perform as good as the statistical models. To determining the weights, the inverse distances from all points to the new point are used:

$$\lambda_i\left(S_0\right) = \frac{\frac{1}{d^\beta\left(S_0,S_i\right)}}{\sum_{i=0}^{n} \frac{1}{d^\beta\left(S_0,S_i\right)}}; \quad \beta > 1. \tag{2.2}$$

Here $d^\beta\left(S_0, S_i\right)$ is the distance from the new point to a measured point, β is a coefficient to adjust weights, λ_i is the weight for neighbour i. The higher the β, the less influence posed on distant points.

A method of predicting the values of a variable at points outside the region covered by existing observations is called extrapolation. Burrough & McDonnell (1998) regard this method as part of interpolation. JORC 2012, Table 1, s 3 states that the maximum distance of extrapolation from data points should be defined.

A strong relationship between two variables can predict one variable if the other is known. Isaaks & Srivastava (1989, p.33) admit that the simplest type of prediction is linear regression.

The literature reveals a great number of interpolation techniques which are divided into two main types: deterministic and stochastic:

- Deterministic interpolation creates surfaces from measured points, based on either the extent of similarity or the degree of smoothing. It predicts a value that is identical to the measured value at a sampled location and is known as an exact interpolator.
- Stochastic methods incorporate the concept of randomness and provide both estimations (deterministic part) and associated errors (Li & Heap, 2008).

The third mentioned in the literature type is called *combined methods* which are capable of using both primary and secondary information.

Interpolation methods are generally classified as local and global approaches:

- Local methods in geostatistics predict the value of an unknown point based on the values of neighbourhood samples (Hoerl & Kennard, 1970). Local methods usually operate within a small area around the point being estimated and usually capture the short-range variation (Webster & Oliver, 2007). However, more correct definition of the term local refers more to the fact that only local information is used, as opposed to global.
- Global methods (Isaaks & Srivastava, 1989; Burrough & McDonnell, 1998) calculate entire dataset to generate prediction for a particular point. In Li & Heap (2008) global methods use all available data in the region of interest to derive the estimation and capture the general trends.

As previously mentioned, the traditional approach to mine evaluation is based on the collection of data, predicting of spatial properties of the mineral occurrence and modelling the ore body. Cressie (1989b) states that the data collected from different spatial locations needs a spatial model that indicates dependence between measurements at different locations.

Other definitions of predictive geostatistics can be found in David (1977), Webster & Oliver (1992), Hengl (2007), Chilès & Delfiner (2012) and Bivand et al. (2013) where the prediction is derived from the spatial structure called the variogram:

$$\gamma(x) = \frac{1}{2} E\left[(Z(x+h) - Z(x))^2\right] \tag{2.3}$$

Here E is the expectation (prediction), $x \in R^2$ is a position vector and $h \in R^2$ is a separation vector. Equation (2.3) is only defined if $Z(x)$ is an intrinsic random function, which suggests that the increment $Z(x+h) - Z(z)$ must be the second-order stationarity. More information on stationarity can be found in Journel & Huijbregts (1978), Isaaks & Srivastava (1989) and Diggle & Ribeiro (2007).

Good illustration of the application of variography to a real-world data can be found in a study by Shi et al. (2000).

Table 2.1. The spatial interpolation methods (Li & Heap, 2008).

Deterministic methods	Geostatistical Univariate methods
Nearest neighbours (Clark & Evans, 1954)	Simple kriging
Natural neighbours distance (Sibson, 1981)	Ordinary kriging
Inverse distance weighting (Shepard, 1968)	Block kriging
Regression models (Hardy, 1971)	Factorial kriging (Goovaerts, 1997)
Trend analysis (Oldham & Sutherland, 1955)	Dual kriging (Goovaerts, 1997)
Splines (Webster and Oliver, 2001)	Indicator kriging (Journel, 1983)
Thin plate splines (Wahba & Wendelberger, 1980)	Disjunctive kriging
Classification (Burrough and McDonnell, 1998)	(Armstrong and Matheron,1986)
Regression tree (Breiman et al., 1984)	Universal kriging (Matheron, 1969).
Kernel smoothing (Hoerl and Kennard, 1970)	Model-based kriging
Geostatistical multivariate methods	**Combined methods**
Universal kriging	Classification combined with other methods
Probability kriging	Trend surface analysis combined with kriging
Kriging with an external drift	Lapse rate combined with kriging
Simple cokriging	linear mixed model
Ordinary cokriging	Regression trees with kriging
Indicator kriging	Residual maximum
Colocated cokriging (Goovaerts, 1997)	Likelihood-empirical best predictor
Simple kriging with locally varying means	Gradient plus inverse distance squared
Multivariate factorial kriging	Regression kriging

2.3.1 Steps of resource evaluation

A question might be raised: *"is there any particular method by which the target must be identified or the evaluation of Mineral Resources must be done?"*

No. Any mandatory method of evaluation is not specified. However, the ASX Listing Rules have the force of legislation being enforced under the *Corporations Act* 2001. They are regarded as statutory regulations. That change came after the ASX became a publicly traded company and questions were asked about it regulating itself. JORC (2012) is part of the Listing Rules and has the force of legislation. As legislation, it only recommends, rather than mandates. In practice, evaluation is being done by more than three methods.

JORC 2012, s 19 states that reports must include the relevant information listed in Table 1.p.26 that is material to an assessment.

JORC 2012, s 20 provides that the geological evidence and knowledge required for the estimation of Mineral Resources. Public Reports must include sampling data of a type, and spacings, appropriate to the geological, chemical, physical, and mineralogical complexity of the mineral occurrence for all classifications [of Mineral Resources].

A number of recommended steps can be found in the geostatistical literature. Isaaks & Srivastava (1989) recommend steps of evaluation, where the first objective is to determine the location, volumes and grade of the mineralised zones. Next, the authors recommend steps for mining stage to determine how the tonnage, grade and the economic feasibility of the project will vary with the selectivity of the mining method chosen. To do this, it is recommended accommodating the following aspects of three-dimensionality to a level of precision which is compatible with an acceptable degree of financial risk:

i. An ability to define and represent geological boundaries with precision.
ii. A precise method of determining the volumes with algorithms for estimation of grade within these volumes.

Webster & Oliver (2007) outline an alternative, a two-stage concept of predicting the value of an unknown function $z(x)$ which can be made in two steps:

i. a modelling step based on measured data.
ii. a prediction step based on prediction model.

Hengl (2007, p.10) defines the steps of modern 3-D spatial prediction modelling as the process which predicts values of a sampled variable over the whole area of interest. It is assumed that the input samples are representative, a variable at some new location s_0 can be derived using a spatial prediction model:

$$\hat{z}\left(S_0\right) = E\left\{Z\left|z\left(s_i\right),\ q_k\left(s_0\right),\ \gamma\left(h\right),\ s \in \mathbb{A}\right.\right\}. \tag{2.4}$$

Where $z(s_i)$ is the input point dataset, $\gamma(h)$ is the covariance model defining the spatial autocorrelation structure, $q_k(s_0)$ is the list of deterministic predictors, also known as covariates or explanatory variables, which needs to be available at any location within $\mathbb{A}$.

According to the book of Hengl (2007) a spatial prediction model comprises a list of procedures to generate predictions of value of interest given the calibration data and spatial domain of interest. This concept is involved in many mapping, geo-coding and spatial statistical packages available in the CRAN `R` library.

JORC 2012, Table 1, s 2, Data Aggregation Methods, suggests:

> ...in reporting Exploration Results, weighting averaging techniques, maximum and/or minimum grade transitions (cut-off high grades) and cut-off grades are usually material and should be stated.

JORC 2012, Table 1, s 3, Estimation and Reporting of Mineral Resources, it is suggested that any explanation include:

> ...the nature and appropriateness of the estimation technique(s) applied and key assumptions, including treatment of extreme grade values, domaining, interpolation parameters and maximum distance of extrapolation from data points. If a computer assisted estimation method was chosen include a description of computer software and parameters used.

2.4 Geostatistical prediction: literature review

The thesis does not engage with geostatistics and does not provide a comprehensive review of geostatistical tools. However, central assumptions in geostatistics will be addressed.

The origins of geostatistics can be identified in Kolmogorov's search for a method of optimal interpolation in the 1930s. The prefix [*geo*] comes from geology. Geostatistics (also known as spatial statistics) has its origins in mining and was firstly mentioned in the study of Sichel (1947) on the application of the lognormal distribution in a gold mine. This study was followed by the unpublished thesis of D.Krige (cited as Krige, 1951) on the application of regression analysis between sampling and blocks, which set the stage for linear geostatistics.

In 1951, a mining engineer D. Krige (Krige, 1951) observed that samples taken close to each other are more likely to have similar values than if taken farther apart. This observation is the foundation on which geostatistics characterises values are defined in 3D space. According to Cressie (1989a), Krige had predicted areal gold concentrations in a South African mine based upon large amounts of data which exhibited strong positive

correlation. Krige showed that the over-and-under evaluation of blocks, can be explained by statistical theory.

Matheron (1963) introduced a probabilistic interpretation to *regionalised variables theory* (RVT) that led to the development of the emergence of geostatistics as an ore reserve estimation technique in early 1960s in France and this spread worldwide. His theory of regionalised variables states that any variable, which is related to its position (i.e. exhibits spatial correlation) and support or volume in space, is called a regionalised variable (more information on RVT can be found in Sinclair & Blackwell, 2004, p.10). The most remarkable contribution of Matheron was the construction a theoretical framework for modern geostatistics that rests upon the random function model.

Since then, many geospatial models based on that theory utilise an understanding of the spatial relations of sample values within a mineral body. In fact, almost all variables encountered in geospatial sciences can be regarded as regionalised variables. Most regionalised variables in reserve estimation display two aspects:

i. Random, which consists of highly irregular and unpredictable variations.
ii. Structured, which reflects spatial characteristics of the regionalised phenomena.

David (1977) outlines the purposes of the RVT:

i. to express the spatial properties of regionalised phenomena in adequate form.
ii. to solve the problems of estimating regionalised variables from sample data.

Some tools based on RVT, for example kriging, has a number of advantages over conventional spatial prediction techniques. Cressie (1989b) reports that kriging accurately explains the process of developing a model for the correlation structure and use this proposed model to predict responses at unsampled locations.

Predicted values are based on the proportion of total sample variability accounted by random noise reports (Isaaks & Srivastava, 1989). That means the noisier the sample set and the less the individual samples represent their immediate vicinity, the more they are smoothed and the greater the associated uncertainty. This phenomenon is called smoothness.

Another feature of the RVT-based techniques discussed in Isaaks & Srivastava (1989) is that the weighting coefficients assigned to a sample are lowered to the degree that its information is duplicated by nearby samples with little variability. This effect is called declustering which helps mitigate the effects of the density of variable samples.

A number of researchers have reported that an integral part of geostatistical estimation is that it predicts not only a value but also a measure of uncertainty associated with the value. On a global scale, geostatistics has been successfully applied to ferrous, non-ferrous metals and precious metals while in some countries its application has been limited to base metals and coal. A good introduction to geostatistics can be found in Sinclair & Blackwell (2004, p.181).

Modern geostatistical techniques override the limitations of numerical methods by providing estimates together with a minimum error variance (Matheron & Kleingeld, 1987). The methods based on kriging utilise an understanding of the inter-relations of sample values for quantifying two important geological concepts:

i. The inherent characteristics of the deposit.
ii. A change in the continuity of interdependence of sample values according to the trend of mineralisation is a range of interdependence of sample values.

Interpolation in spatial statistics usually is synonymous with a spatial prediction method *kriging*, which Gandin (1974)[3], called optimal interpolation.

Kriging is described by David (1977) as an *optimisation technique* of grade estimation. Ord (1983) defines kriging as a method of statistical interpolation for random spatial processes, which presents predictors that are linear in the observations. Cressie (1989a) equates kriging with spatial optimal linear prediction, where the unknown random-process mean is estimated with the best linear unbiased estimator.

According to Matheron (1963):

> *...kriging consists [of predicting]... the grade of a panel by computing the weighted average of available samples...the suitable weights* a_1 *are determined by* $\sum a_i = 1$*...*

Danie G. Krige in Krige (1976) provides another definition:

> *...the multiple regression procedure for arriving at the best linear unbiased [predictor] or best linear weighted moving average [predictor] of the ore grade of an ore block [of any size] by assigning an optimum set of weights to all the available and relevant data inside and outside the ore block...*

Cressie (1989a) looks at various versions of kriging as an assumption about the mean function change. Based on these quantifications, kriging performs estimation with a minimum variance and an error of estimation both in local and global scales. None of these properties is taken into account in the classical methods. Technical papers also indicate that kriging techniques (different variants of kriging estimators have been developed) mark a major advance in Resource and Reserve evaluation.

Good illustration case study and the application of modern structural analysis can be found in a study by Tolosana-Delgado et al. (2019),

2.4.1 Assumptions

The main difference between classical statistics and geostatistics is the assumption of spatial dependency. In this section, the greatest strength of interpolation - assumptions are briefly discussed.

It has to be noted that there have been some disagreements between researches on the definition of the discussed below assumptions – for different authors at different times assumptions mean different matter.

Isaaks & Srivastava (1989, p.438) observe that interpolation necessarily involves a number of assumptions about how the distribution behaves at points where it has not been directly estimated.

Assumptions allow for the spatial interpolation methods to be formulated and to create a surface that is intended to best represent empirical reality (Figure 2.2). Loosely speaking, there are two assumptions:

1. the attribute data are continuous over space.
2. the attribute is spatially dependent, indicating the values closer together are more likely to be similar than the values farther apart.

The assumption of stationarity is one of the most important concepts of statistical prediction and the requirements for some interpolation methods. Figure 2.2 shows a

[3](n.d.) L. Gandin and S. Kagan; in Journal of Meteorology, Lenindrad, 1974

hypothetical 2-D distribution patterns of a metal, say, "M". The surface is obtained from kriging, in which the assumption is made that the statistical properties are the same throughout the shown area.

Azpurua & Ramos (2010) state that spatial interpolation assumes the attribute data are continuous over space. This allows the estimation of the attribute at any location within the data boundary. Another assumption discussed by Azpurua & Ramos (2010) is that the attribute is spatially dependent, indicating the values closer together are more likely to be similar than the values farther apart.

Assume a model build by interpolation in the form of a surface (Figure 2.2) with predicted grades (say, metal "M") between the sampled points.

The model in Figure 2.2 suggests that gradational character with a pronounced effect on smoothing exists between sampled points. The obtained surface predicts grade at an unsampled "red" point as 1.6 M@gram per ton, which may or may not reflect the reality. This effect is probably the major complaint made by some non-statistically oriented people. From an end-user perspective, the contours indicate just trends. Hekmat et al. (2013) observe that this approach to modelling results in excessive smoothing of the data which masks the orebody's true grade variability.

Interpolation creates optimal predicted surfaces and delivers a measure of confidence of how likely these predictions are true. For example, statistical procedures behind the contouring indicate that not much evidence is available to suggest that the existence of grades, say -3 or +7 M@g/t are likely to happen. Despite the model in Figure 2.2 is being smooth, the assumptions can make the degree of smoothness controllable. This feature is used in many commercial GIS packages, which automatically provide predictions in elegant visualised form.

From an optimisation viewpoint, the problem of predicting a value at or near "red point" is solvable. However, in this scenario, continuous surface as shown in Figure 2.2 may not be created.

Note: hypothetically, a global solution can be found by a published in 2018 on CRAN `MlBayesOpt` Bayesian R–package, which can tune hyperplanes. However, the author of this thesis failed to find firm practical evidence of the claimed.

The most frequently mentioned and cited assumption is the stationarity. It simply means that the mean and variance of values do not depend on location. Hengl (2007) defines stationarity as a property of a variable to have similar statistical properties within the whole area of interest. Oliver & Webster (2015) admit that stationarity underpins the practicality of geostatistics, and it is an assumption that enables data to be treated as though it has the same degree of variation over a region of interest.

Strong stationarity in the interpretations of Cox & Miller (1965, p.277): $Z(x)$ is stationary if for any finite number n of points $x_1, .., x_n$ and any h, the joint distribution of $Z(x_1), .., Z(x_n)$ is the same as the joint distribution of $Z(x_1 + h), ..., Z(x_n + h)$, where $Z(x)$ is a random function defined in 1-,2- or 3-space and x is a point in space, not just the finite coordinate. This form of stationarity does not imply the existence of means, variances, or covariances.

In Myers (1989) strong stationarity implies that the mean, variance and all other distribution parameters are everywhere the same. From a geostatistical point of view, this assumption is too strict and hard to be verified, so it is usually weakened.

Second-order stationarity is implied by strong stationarity. In Webster & Oliver (2007, p.52) "weak" is defined in terms of the covariance function. It is assumed that the covariance between two points is the same for a given distance and direction, regardless

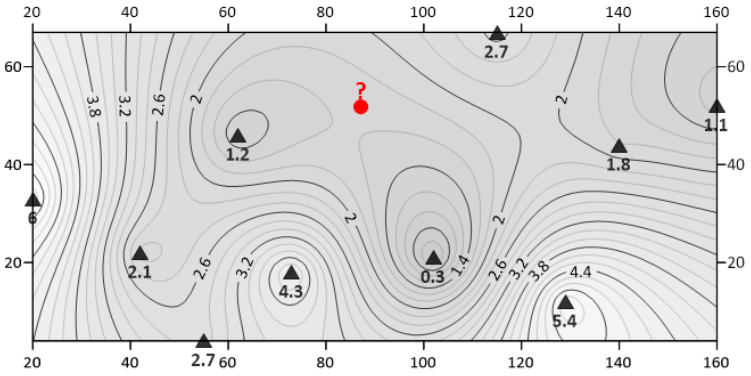

Figure 2.2. Example of a surface obtained from kriging. Triangles are the measured points, the contour lines are at interval 0.2, grid referenced in every 20 metres. What is the value at the red point?

of which two points are chosen. It was believed that this assumption may lead to a more general geostatistical analysis based on the variogram as a description of the variation.

David (1977) admits that the *second-order* stationarity is based on a condition that the expected value E of the regionalised variable $z(x)$ is the same all over the area of interest

$$E\left(z\left(x\right)\right) = m, \tag{2.5}$$

where E is the expectation. The spatial covariance of $z(x)$ is the same all over the field of interest. Then, the covariance becomes

$$E\left(\left(Z\left(x\right) - m\right)\left(Z\left(x+h\right) - m\right)\right) = K\left(x,\, x+h\right) = K\left(h\right). \tag{2.6}$$

Here h is the separation in space and the variance of the random function $z(x)$ is

$$\operatorname{var}\left(Z\left(x\right)\right) = E\left(\left(Z\left(x\right) - m\right)^2\right) = K\left(0\right) \tag{2.7}$$

and $z(x)$ and $Z(x+h)$ are the values of the random variable z at places x and $x+h$, and E denotes the expectation.

In Oliver & Webster (2015) the covariance of $z(x)$ depends only on h which is the separation between samples in both distance and direction. Hence, $z(x)$ is a function of h. Therefore, the process has a covariance only if variance of $z(x)$ is finite. However, in reality a finite variance may not exist.

For some situations when the variations of grade rather than"just" grade is considered, a finite variance may exist. For example, when the increments of the function

$$Z\left(x\right) - Z\left(x+h\right) \tag{2.8}$$

are considered, one can make the following assumption:

$$E\left(Z\left(x+h\right) - Z\left(x\right)\right) = 0 \tag{2.9}$$

$$\operatorname{var}\left(Z\left(x+h\right) - Z\left(x\right)\right) = 2\gamma\left(h\right) \tag{2.10}$$

The form (2.10) is the *variogram* which is based on differences and provided by (2.9) holds locally. In the rewritten form of (2.10) provided in David (1977):

$$\operatorname{var}\left(Z\left(x+h\right) - Z\left(x\right)\right) = E\left(Z\left(x+h\right)\right) - Z\left(x\right) - E(Z\left(x+h\right) - Z\left(x\right))^2. \tag{2.11}$$

If assumption of expected value $E\left(Z\left(x+h\right)-Z\left(x\right)\right)=0$, then

$$\text{var}\left(Z\left(x+h\right)-Z\left(x\right)\right)=\ E(Z\left(x+h\right)-Z\left(x\right))^2. \tag{2.12}$$

Equations (2.10) and (2.3) are the intuitive definitions of the variogram.

In some cases, the geostatistitians use a weaker definition of stationarity. If a random function $z(x)$ is a second-order stationary it is *intrinsic assumption*. Another word, it is the assumption that the variance of the difference is the same between any two points that are at the same distance and direction apart no matter which two chosen points. In this case the variogram becomes

$$\text{var}\left(Z\left(x+h\right)\right)+\text{var}\left(Z\left(x\right)\right)-2\text{cov}\left(Z\left(x+h\right),\ Z\left(x\right)\right). \tag{2.13}$$

The (2.13) is a more accurate formulation of an intrinsic assumption that can be applied to reserve estimation. More information on the two assumptions can be found in David (1977, p.94) and Oliver & Webster (2015, Ch.3). Second-order and intrinsic stationarity are assumptions that allow predictions and assess uncertainty in the predictions to be made.

Oliver & Webster (2015, Ch.2) discuss a case *quasi-stationarity*, which limits stationarity to local areas, i.e. with sufficient data the assumptions can be applied locally.

2.4.2 To Smooth or Not to Smooth? Criticism

Unstructured orebodies should be smoothed. Structured orebodies should only been smoothed only if there is not significant difference between the grades. An interesting discussion on smoothness and nonsmoothness of drilling data can be found in the report by Dr. Spero Carras (FAusIMM) in Carras (1990).

Criticism towards some interpolation techniques have been from financial institutions, securities analysts and stock market experts. From an end–user's viewpoint the model in Figure 2.2 does not provide enough information to make investment decision.

For example, concerns about the accuracy of geostatistical prediction has been addressed in several annual reports issued by the stock exchange companies. First small-scale investigation conducted by Miskelly (1982) discusses a number of unsuccessful examples of estimates. Miskelly (1982, p.13) points his critique towards at estimators: *"...some of the estimators appear to become fascinated by figures and to lose sight of their meaning".*

Some mining practitioners argued that geostatistics encounters the most difficulty when applied to vein-type gold deposits. For example, Clow (1991) wrote:

"...this type of treatment [geostatistics] is very poorly understood by the average mining engineer and geologist and, as a result, the output is frequently misapplied".

The recent case related to the temporarily close one of the biggest gold mines in Australia can give rise to opinions that the above statements are incorrect.

There have been complaints on inability if geostatistics to reproduce the surface exactly, the inability to account for the physical processes that created the underlying spatial distribution, for the approach to extreme values treatment and the absence of a single workflow that can be applied to every spatial problem.

Some online forums contain discussions with examples of over-estimations of mineral resources, many of which emerged from the inability of mines to match ore estimates either in terms of grade or expected tonnage. However, such form of criticism is not

constructive and do not pose alternative, positive outcomes.

Although there have been major advances in the development of spatial tools over the past decades, there remain definite areas for improvement.

A variety of constructive criticism have been levelled at geostatistical methods by some members of the geostatistical community for a number of disadvantages.

The irony is that criticism has mostly been aimed at the greatest strength of interpolation – assumptions (section 2.4.1). Discussions can be read in Journel & Huijbregts (1978), Shurtz (1985), Philip & Watson (1986), Knoll (1989), Armstrong (1994), Sinclair & Blackwell (2004), Rossi & Deutsch (2014). It was admitted that some methods are applied without sufficient regard to the accuracy of the available sampling grids, which generally provide detailed information only along the drillhole leading to incomplete information of parameters of continuity (see Figure 2.2).

Sarkar et al. (1988) argued that the conventional methods of estimation, in practice, do not provide any objective way of measuring the reliability of the estimates. The principal complaint made by Merks (1992) was that kriging tended to inflate expectations for the continuity of mineralisation between measured data points.

Philip & Watson (1986) and Shurtz (1994) criticise the semivariogram definition, the validity of kriging as the Best Linear Unbiased Estimation, the treatment of the extreme values and the significance of the estimation variance. Concern was expressed that all parameters of a semivariogram model are often determined subjectively (Remark 1).

Remark 1. The term subjectively also relates to technical jargon, which widely used among the 3-D modellers. This relates to a phenomenon known as "blown bubbles", or an automated process of building a 3-D solid, which sometimes manifested by some software developers as a significant advantage, which is not entirely correct. Because of the existence of several geomechanical, structural and hydrological constraints, the gold-bearing ores are extracted partially, even if they have already been included into the reserves, that is it, these constraints make the automatic conversion into an accurate solid(s) impracticable. Even if such subjective solids are being created, they are manually calibrated. In such cases, the wireframes are being rebuilt manually, string by string. Hence, the vertices, nodes and edges are re-positioned.

It has to be noted that because of the existence of numerous geomechanical, structural and hydrological constraints, the gold-bearing ores are extracted partially, despite the inaccessible areas included into reserves. These constraints make the automated creation of accurate solid(s) impossible. Even when such solids created automatically, they are being calibrated daily. The wireframes are being re-built manually string by string, and vertices and edges are re-positioned. However, this topic is usually avoided by CAD-based GIS software presenters during presentations.

Many critiques have been towards poor interpretability of the regression coefficients and the R^2.

Kriging has been criticised for its dependency on the fitness of variograms, which sometimes difficult to correct for smoothing of the estimates. Also, criticism has been towards the cost of computation and complexity related to interpretation of results.

Sinclair & Blackwell (2004) admit high cost of estimations and the requirement both highly trained personnel as well as substantial computing capability and demands close attention to data quality and geologic interpretation. Loquin & Dubois (2010) describe a case of computational burden of kriging. The authors admitd the poor interpretability of its influence coefficients that depend on the fitness of variograms, which is some cases

are difficult to correct for smoothing of the estimates.

Falivene et al. (2010) argued that in some cases traditional geostatistical methods were incapable of fulfilling the requirements to accuracy because many deposits demonstrate high degree of complexity. The complexity may be due to foldings, irregularities, variation in grade or combinations of these.

Some researchers admit that the smoothness in the interpolated models is related to the number of times it is mean-square differentiable. For instance, Chilés & Delfiner (2012) found that the more times it is differentiable, the greater is the degree of smoothness in the obtained models, i.e. the surfaces provided by kriging are often over-smoothed.

Adamu & Brandon (2014) admit that some geostatistical techniques consider mainly spatial correlation in the datasets, but they violate the aspect of observing independence of attributes and randomness in data distribution.

The shortcomings of stationarity have been examined and discussed in Goovaerts (1997) and Webster & Oliver (2007) and criticised on many instances by Shurtz (1985), Myers (1989); Armstrong & Champigny (1989), Sinclair & Blackwell (2002), Rossi & Deutsch (2014), Oliver & Webster (2015) and many others for a property of a variable to have similar statistical properties within the whole area of interest.

Particular criticism has been mostly towards *second-order* stationarity which, in a critic opinion, makes some methods perform poorly when the data involves discontinuities or nonlinear trends, i.e. when the same semivariogram applied over the entire area is assumed to be a function of distance. Some authors believe that this is an assumption about stationarity in the field. Instead of assuming that the variance is everywhere the same, it is assumed that the variance depends solely on the distance.

The controversy over the strict and weak stationary assumptions has raged for many years. Henley (2001) viewed stationarity as an unlikely property of a model suitable for fitting data, which clearly vary in both expected value and variance from one place to another. Henley (2001) states that it is impossible in principle to test a dataset for stationarity, then the entire body of geostatistical methods built upon any stationarity assumptions must be rejected as unscientific. Finally, he suggests seeking an alternative model, which does not require any such assumptions.

Criticism has also been towards the degree of continuity of the regionalised function from the behaviour of the semivariogram near the origin. The behaviour of continuity gives information on the interval of distances for which the spatial dependence is high. The theoretical reason involves the concept of:

i. Mean-square continuity.
ii. Smoothness.
iii. In some cases the differentiability of the regionalised function.

It is appropriate that basic definitions of these terms are provided in this section. The following proposition characterises quadratic mean continuity:

a *second-order* stationary process at point s denotes

$$\lim_{h\to 0} E\left(\left(Z\left(s+h\right)-Z\left(s\right)\right)^2\right)=0 \tag{2.14}$$

which implies the following condition:

$$\lim_{h\to 0} 2V\left(Z\left(s\right)\right)-2C\left(h\right)=\lim_{h\to 0} 2\left(C\left(0\right)-2C\left(h\right)\right)=\lim_{h\to 0}\gamma\left(h\right)=0, \tag{2.15}$$

where h in (2.15) is the distance and V is the support of the regionalised function. Condition (2.15) indicates that unless $C\left(h\right)\rightarrow C\left(0\right)$, or $\gamma\left(h\right)\rightarrow 0$ as $h\rightarrow 0$, the

regionalised function cannot be continuous at location s.

The advantages and disadvantages of *smoothness* and *differentiability* and other properties were discussed by Diggle & Ribeiro (2007).

In gold mining, the process of prediction with geostatistical tools is complicated by heterogeneity introduced by the underlying geology, the erratic distribution and the lack of reliable samples. In spite of these criticisms, the geostatistical approach is currently considered to promote a more comprehensive understanding of the spatial phenomenon.

2.4.3 A link between prediction and reality

Matheron & Kleingeld (1987) admit that every ore body is unique and sometimes it is only possible after mine closure to determine exactly what the ore body contained. A number of problems related to the comparison of predicted and encountered data and the validity of predictions were discussed in Wahlstrom (1964), Robinson & Lee (1967), Dowding (1976b). Collectively, these studies report that apart from general geological information, little success had been noted in the locational prediction of worthwhile targets in the investigated mines.

Parker (2012) discusses situations when a small estimation error can cause a mine to miss its cash flow targets. To understand and avoid sources of errors, it was suggested to gather reconciliation data at various stages of the mining process. The author admits that one of the main sources of error that should be considered is inaccuracy in the estimation of resources and reserves.

The distribution of grades within an ore body is of mixed character, being partly structured and partly random (Matheron & Kleingeld, 1987) and it is never so chaotic as to preclude all forms of forecasting characteristics of the deposit which is never regular enough to allow the use of deterministic prediction techniques.

In regard to the evaluation of gold deposit, a particular concern is the acceptance of limited and poor input data. The question raised by David (1977) and Isaaks & Srivastava (1989) was whether it was possible to reconstruct the probability distribution from limited data before actual development commences? Finally, the following was proposed:

i. The domain to be sampled must be defined.
ii. The properties of the material in the domain also must also be defined.

A study by King et al. (1982) reports that statistics should not be involved in reserve estimation until all other factors such as geological continuity and contacts, lost cores, sampling and assay errors have been identified, examined and assessed. Rendu (1994) points out the statistics realises the need for a link between true geology and statistics that is manifested at each step of a statistical study. However, Sinclair & Blackwell (2004) argue that link is often complex and not well understood and connection between real geology and geostatistics is often tenuous.

A very important aspect is that for the greenfields and grassroots projects accurate comparisons between the statistical predictions with reality are technically impossible. Rossi & Deutsch (2014) admit that an accurate model is one that reproduces well the actual tonnages and grades mined. This check can only be performed in mine development stage or if the mine is operating.

Studies by Sinclair & Blackwell (2004) and Carranza (2011) discuss a number of attempts undertaken to bridge the gap between "true" geology and statistical predictions, which is the quality of information to produce adequate estimates of worthwhile targets

in the early exploration. The authors report that the amount of information, from which reliable geological models were created was so low that any statistical inferences were unstable.

The quality of sampling has also received considerable critical attention in surveys conducted in Sinclair & Vallee (1993), Pistolski & Sinclair (1998) who point out that all available geological information should be used more to improve the reliability of ore estimates.

In return, Gy (1989) and Pitard (1993) refer to the theory of sampling which provides an insight into the causes of errors that occur during assaying. In practice, the sampling errors can never be eliminated due core loss and limited number of accessible samples. Noble (2011) points out that there should be no prediction across a discontinuity with a limited amount of information. A consistent sampling error may result in either underestimation or overestimation.

JORC, 2012, s.18 suggests a solution: ... where assay and analytical results are reported, they must be reported using methods, selected as the most appropriate by the Competent Person by ... listing all results, along with sample intervals.

Results of obtained from the sample intervals method can be met in some Public Reports and Technical Paper in the form of 2D cross-sections. The method is intuitively understandable and widely used by both exploration practitioners and the investment community.

Mine development gives the most reliable assessment of ore distribution. Another question raised in the literature (e.g. David, 1977; Matheron & Kleingeld, 1987; Isaaks & Srivastava, 1989) was that whether it was possible to reconstruct the probability distribution from a limited drilling data before actual development commences? Finally, the following was proposed:

i. The domain to be sampled must be defined.
ii. The properties of the material in the domain also must also be defined.
iii. The outset of a survey the dimensions of the units is their size, shape and orientation of the samples.

Numerous examples of accurate predictions are known. For example, a study by Chilès & Delfiner (2012, p.220) checks the accuracy of the geostatistical estimates and revise them to obtain a model “reality–prediction" for the continuation of a tunnel at Cault Clay. The observations in the service tunnel were found in good agreement with the geostatistical model.

The problem of incorporating geology into geostatistics in the literature has been accurate modelling with respect to geological controls of mineralisation, i.e. determination of lateral limits that closely reflect reality. Most researches agree that only mine development and geological reconciliation give the most reliable assessment of the difference between prediction and realistic distribution.

As previously mentioned, detailed information on a comparison of predicted characteristics with actual measurements is hard to find in open sources. At the exploration stage, accurate reconciliation is not technically possible.

The major concern of a practitioner is that nearly all interpolation methods are fundamentally averaging techniques and they produce results that are smoother than reality. As a consequence, some engineers are sceptical of these methods. Logically, the better the quality of the input data the better the results that will be achieved by an analytic process.

However, all currently known methods encounter the most difficulty when applied to gold deposits - drilling data is never enough.

2.5 Evaluation and Valuation

In exploration context, the first step of evaluation is the process of defining and reporting Resources. The second step of evaluation is financial reporting for the exploration for and evaluation of mineral resources.

Clause 6. of Australian Accounting Standard AASB 6 2015[4], states that exploration and evaluation assets shall be measured at cost.

Clause 18. of AASB 6 2016 states that exploration and evaluation assets shall be assessed for impairment when facts and circumstances suggest that the carrying amount of an exploration and evaluation asset may exceed its recoverable amount.

Wrong resource definition and consequently, reporting is subject to investigation.

Next stage is valuation, which is performed on a project to determine its financial value considering VALMIN 2015[5] recommendations. It has to be noted that evaluation is not valuation. The requirements to valuation of exploration and mining projects are outlined in VALMIN Code 2015, which does not constitute legal advice.

During exploration, there is usually an increase in stock price as investors speculate, based on drilling or other sampling results, whether the company has found anything. As the company defines resources and releases further results, investors usually become interested in the stock.

The Legislation does not specify a method by which resource estimation must be done. The JORC, 2012 recommends its adoption as a minimum standard for Public reporting. Exploration companies encouraged to provide information in their Public Reports which is as comprehensive as possible.

The exploration and mining sectors are private, invested by individuals, banks, big SuperFunds and mining companies. The outcome of the evaluation is valuation of a project in monetary terms and determining its Fair Market value.

2.5.1 The outcome of evaluation: project's fair market value

Fair market value (FMV) is the price that exploration project would sell for on the open market. Micon International, define the term fair market value as the price which is established in a free and open market by transactions between a willing and informed buyer and a willing and informed seller, both of whom are acting without compulsion and at arm's length. Xstract Mining Consultants state that the fair market value often implied by the market capitalisation of the holding company can vary significantly on a daily basis.

In banking terms, an exploration property is these on which an economically viable mineral deposit has not been demonstrated to exist. The real value of an exploration property lies in its potential for the existence and discovery of economically viable mineral deposit.

[4] Available at https://www.aasb.gov.au
[5] Available at http://www.valmin.org

Another aspect is that the open market is often driven by "emotion" and speculative information during the exploration phase. According to MinEx Consulting, Australian valuers tend to overvalue grassroots exploration projects in average by 40%.

The time taken to make a discovery and turn it into a mine is, on average, from 10 to 15 years. During this period of time, the exploration activity and updated geological data are closely monitored.

In 2018, more shares of Mineral Resources have been bought than sold by. The bidding statistics, provided by some online platforms suggests that experienced bidders are better informed than other bidders. Numerous open-sources, for example Australian Securities & Investments Commission, WallSt[6], report that leaks, inaccurate statements or references lead to a rise or fall in the trading price of shares (ASIC 2019). Former mining executive sentenced to serve 9 months for insider trading, *Media release*, Monday 11).

2.5.2 Methods used for mine evaluation

Journel & Huijbregts (1978) wrote that a need of the application of interpolation to evaluation was historically motivated by the following:

i. Estimating the amount of worthwhile metals in an orebody.
ii. Selectively mine the orebody based on a set of observations at known locations.

These days, at the greengrass and grassroots stages, interpolation is used to establish the shape, positioning of exploration target and approximate tonnage. At the mining stage, interpolation is used to accurately determine the boundaries of a deposit, grades and the tonnage.

A number of unique properties of spatial interpolation were discusses in Section 2.4.1. The goal of interpolation is to create a surface that is intended to best represent empirical reality. From a GIS-user viewpoint, the unique feature of most interpolation methods is that their smoothing properties make it possible to use it in contour mapping and cross-sectioning with isolines. This process is described in Webster & Oliver (1992), Hengl (2007) and Bivand et al. (2013) as minimising estimation errors, calculating optimal sampling distances, mapping and estimation of the size of the recoverable resource.

A number of numerical techniques and procedures summarised in Table 2.1 such as polygon and splines (Ahlberg et al., 1969, Baxter, 2004), triangulation, nearest neighbours distance and inverse distance weighted (David, 1977; Annels, 2004) have been used in the evaluation.

From 1 December 2014, clause 29 of the JORC, 2012 require Ore Reserves to be defined "at Pre-feasibility or Feasibility level as appropriate that include the application of Modifying Factors".

In the *Pre-feasibility* phase, a number of traditional tools such as histograms, Q-Q, P-P, cross-plots are used in analysing general statistics per domain. Variograms are usually used for understanding the geometrical parameters of an ore body. The assumed local confidence intervals are used for resource classification and reports.

In the *Feasibility* study phase, several interpolation techniques such as Inverse distance weighted interpolation (IDW) and kriging help in building a 3-D model and determine the volume/ tonnage, economic cut-off grade, i.e. evaluate recoverable reserves more precisely. Infill drilling updates are used to improve understanding of spatial uncertainties.

[6] A commercial graphical investment analysis https://simplywall.st/

In the *Production* stage kriging based on updates assists with detailed evaluation of grade to optimise mine planning. The grades obtained from stopes and stockpiles are compared with the previously predicted grades.

The publicly available Exploration Reports in the form of Table.1, JORC 2012 show that IDW and kriging have been the most used methods for the evaluation. Some previous studies such as that conducted by Li & Heap (2008) found that kriging methods, as a predictive tool performs better than other methods, with only a few exceptions.

2.6 Machine Learning approaches in mining engineering

Recent developments in optimisation and machine learning (ML) have led to the idea that many mining problems such as the efficiency of equipment, scheduling and spatial variability can be solved with the power of machine learning. Many recently reported experiments have established the ability of machines to learn and automatically process data to identify potentially anomalous regions.

Artificial Intelligence (AI) is the established name for the field, but the term artificial intelligence is a source of much confusion because artificial intelligence may be interpreted as the opposite of real intelligence (Poole & Mackworth, 2017).

Ever since computers were invented, people have wondered whether they might be made to learn. It is not yet how to make computers learn nearly as well as people learn, stated in Mitchell (1997). With the increasing amount of data being collected universally, machine learning (ML) systems are becoming more popular and are increasingly using data mining methods to detect patterns of anomalies. An anomaly in ML is an observation or a pattern of observations that does not conform to the expected normal behaviour of the data. It is believed the adoption of AI and ML as an exploration tool is inevitable. Two questions to be posed: *What amount of value to geological and exploration models adds the involvement of AI? Is there a false reliance on the models being produced by AI?*

Exploration is inherently risky, but, considering some economically disastrous resource estimates, are the claims that ML aid in mitigating risk valid?

Recent experiments described in Harrington (2011) have established the ability of ML to automatically identify spatial clusters, form prototypes, identify anomalous regions, recognise correspondence patterns in spatial variations, interpolate and extrapolate over multi-dimensional relationships and map their spatial variation.

A further unique feature of the ML is the ability to relate these discoveries to other geologic, geochemical and lithological interpretive objects. The most frequently mentioned in the mining literature method of ML is the artificial neural networks (ANN) technique, which has an information processing structure consisting of relatively simple processing elements, similar to the neuro-cells in the brain. Just like the brain, the ANN learns from training repeatedly on a set of data.

In supervised learning, the neural network shows both input data and the desired output data. After each trial, the ANN compare its own output with the correct output, corrects any deficiencies, and tries again, iterating until output error reaches an acceptable level. Analysis of the development of AI techniques allow another question to be posed:

"can artificial intelligence produce a model from which new targets, trends and new groups of clusters can be identified, pointing a rig operator towards locations where gold occurrences might be uncovered?"

2.6.1 ML-based techniques in mining engineering

This section reviews published sources on the application of machine learning (ML) and other artificial intelligence (AI) techniques to the problem of exploration targeting.

There is growing excitement that AI and ML will aid in exploration and help us uncover new deposits. Interest in ML has been growing steadily, and many mining companies are aware of the potential impact AI can have on the results of exploration and extraction.

In early stages of mining, a sufficiently large drillhole dataset is not usually available for accurate evaluation and setting fair value in the open market. However, this limited data cannot be disregarded because at all stages of evaluation a “go” or “no-go” decisions are required to be taken. It was believed that AI systems, would become an integral part of the evaluation in the investment decisions for explorations projects.

The late 1980s through 1990s witnessed rapid growth a number of geocomputational developments. Ways unlock the predictive mapping capability of ML were investigated by Harris (1989) and McCammon (1989). Trends in ML for resource mapping were surveyed in an extensive study by Bonham-Carter et al. (1998).

Hagens & Doveton (1991) explored an application of the cerebellar model articulate controller to a two-dimensional surface interpolation problem with encouraging results. The authors showed that the extension of the neural model to fuzzy logic permits multi dimensional and general-purpose capabilities that are not possible with lower-dimensional conventional mathematical techniques. The need to examine Al-based techniques in dealing with qualitative information in geostatistics is discussed in Dimitrakopoulos (1993).

An approach to modelling of spatial data using NN was described in Clarici et al. (1993). Wu & Zhou (1993) reported encouraging results within these limitations for solving generic ore-reserve estimations by NN. With the proposed technique, higher precision of generated interpolations, classifications and extrapolations were possible together with faster learning.

Kinnicutt (1994) describes a system called NOMAD which could be used for 3-D stratigraphic characterisation of a deposit to create ground profiles from borehole data. The authors found that this could be done by combining geostatistical and knowledge-based approaches, i.e. the statistical interpretation is combined with subjective data entered by the user. The use of ANN for classification of mineral deposits was considered by Singer & Koude (1997). Yama & Lineberry (1999) presented an ANN model for learning the spatial continuity for predicting values for given coordinates. The authors studied the "trainability" of ANN for normally distributed geologic data. The predictive power of ANN was compared to kriging on the same sub-regions. The authors found that the trained networks performed satisfactory and the developed model was capable of learning from a training set and predicting the unseen validation set. A set of examples was used to tune the parameters of the selected classifier.

Much of ML use in the described methods fell under supervised learning technique. Collectively, all these studies outlined enormous potential of ML and explored what ML can and cannot do.

Researchers believe that ML is valuable for extracting patterns from large and complex datasets. Investigations into spatial analysis with ML has especially been active over the last decade and a large number of applications to model geological data for the purpose of predicting the presence of commodities have been proposed.

Several solutions to *probabilistic modelling* with the application of ML techniques such as functional approximation, weight of evidence, parameter fitting can be found in the study of Barnett & Williams (2006).

As machine learning gets deployed in decision-making situations the ability to estimate error become essential. Skabar (2007) reports results of applying Bayesian learning techniques to the production of maps representing gold mineralisation potential over the Castlemaine region of Victoria. The application of Bayesian approach shown that optimisation of parameters such as the weight decay regularisation coefficient can be performed using training data alone, avoiding the noise introduced through split-sample validation.

Carranza (2011) discusses a number of knowledge-driven, bivariate and multivariate data-driven computation of spatial associations between known deposit-type locations and spatial evidence. The study identified the existence more than 150 journal articles and conference papers related to the application of ML to spatial modelling. An answer "which of the predictive techniques is the most efficient, therefore has the best chance of leading to ore deposit discovery?" is looked for.

Melkumyan & Ramos (2011) took a machine learning view to the problem of resource evaluation. This study describes the use of a non-parametric Bayesian method, leading to parameter estimation for solving the problem of the quantification of the uncertainty of measurements obtained from multiple sources. The squared exponential covariance function was applied to constructing a geological model using chemical assays of drill chips taken from exploration drillholes.

It was shown that through Bayesian learning, the parameters of the investigated data were obtained by optimising a parameter, which the authors called marginal likelihood. Another contribution of this study is that the data fusion mechanism continuously updated data as more drilling data become available. This approach allowed a comparison of different statistical models using different covariance functions.

A document search through Scopus revealed more than 600 modelling efforts published between 2012 and 2018 on using Bayesian learning. This type of learning was applied in GIS-related phenomena such as defining important variable relationships. Importantly, Bayesian spatial modelling does not require a Gaussian spatial process and is more flexible in generalised linear modelling.

An automated geostatistical interpolation method called Empirical Bayesian kriging was explored by Krivoruchko & Gribov (2014). It has to be noted that the reviewed GIS methods based on Boayesian learning, were fast and provided smoothed surfaces.

Zuo et al. (2011) described results of the application of support vector machine (SVM), a supervised learning algorithm to mineral prospectivity mapping. Despite the obtained with R results indicated the usefulness of SVM as a tool, it was found that weights for each (WofE) prediction approach provided higher predictive accuracy.

A study by Rodriguez et al. (2015) draws attention to a number of crucial issues related to the application of different ML classifiers such as classification trees, ANN, SVM and random forest. The study analyses statistical significance of the differences between the performance of these methods and sensitivity to data set size reduction and noise.

Probably one of the most interesting approaches to targeting exploration efforts was proposed by Joly et al. (2012). This study stands out as a topic gaining our interest, because this study was devoted to gold prediction for efficient drill targeting, filters out human bias and prioritise areas for hitting promising targets. The objective of this study

was to minimise human bias to make the decision on whether a target makes sense or is real. A number of measures were employed to evaluate the spatial associations between known deposits and predictor maps to establish weights for each predictor layer. Then these layers were combined in a predictive map using WofE approach. Unfortunately, source codes have not been provided. Abedi et al. (2013) proposed a method of reducing the cost of exploratory drilling by the application of a fuzzy knowledge-driven method called multi-criteria decision making (MCDM) technique. Ducart et al. (2016) describe the application of unsupervised learning to multi-source classification for mapping iron oxides.

A recent study by Desharnais et al. (2017) discusses the particular challenges facing the application of ML to the problem of exploration targeting. A potential workflow that can be applied to most of ML techniques to exploration targeting such as a node–by–node ore/waste classifier is proposed. It is admitted that proper validation, verification and the application of basic geological principles into computation will help limit spurious results resulting in drill targets are optimised for discovery.

The literature suggests that careful selection of the appropriate ML approach is needed to ensure the best possible results. Exploration crew feel or can estimate the distance beyond which statistical estimates are unreasonable. The ML, in contrast, may automatically point at a single gold anomaly located far beyond the exploration property ignoring faults and lithological boundaries.

The reviewed studies have exclusively been focused on supervised ML and computing a *global solution*, failing to address unsupervised and reinforced ML, which in some cases can provide results with a higher degree of precision. Since cloud SQL, storage, real-time iteration and import of geological datasets directly into multi-platform compilation system were not available, reproducing research results can be challenging even for IT-oriented researchers.

Another observation is that the visualisation, the essential foundation of any good exploration program, may appear, is undervalued in some studies, although all authors use 2-D images as an important visual component of their studies.

2.6.2 Other ML applications in the mineral industry

The late 1990s and early 2000s represented a period of a significantly increased number of publications on the application of ML to solving spatial problems. This subsection reviews ML techniques which have been developed for solving various important problems such as site characterisation, classification of rocks, development of underground openings, ground support and estimation of earthworks.

Many studies have been published on the application of ML in rock mechanics. Dershowitz & Einstein (1989) studied and discussed opportunities for improvement in rock mechanics practices such as rock fracture flows, rock wedge stability and numerical modelling through the use of ML. Other applications of ML to geologic roof classification and longwall stability prediction were surveyed by Lee & Sterling (1992), Cardon & Hoogstraten (1995) and Zhang & Bhattacharyya (1995).

A significant step towards the use of ML to manage ground conditions was made by Millar & Hudson 1994) who outlined the theory of AI behind the use of ANN in monitoring rock mechanics performance. An interesting approach to blast damage is introduced in Yu & Vongpaisa (1996) with special reference to mining operations for assessing damage by incorporating the vibration level, rock properties, site characteristics and the effects of ground support systems.

A number of AI-based systems have been proposed for geomechanical applications by Rehak et al. (1985). A tool for analysing failures and slope stability was introduced in Grivas & Reagan (1988). Ghosh et al. (1987) proposed an AI application for deciding on spacing for supporting coal mine roofs. Zang et al. (1991) applied ANN to the estimation of coal mine support. In a study by Harvey & Fotopoulos (2016), the performance of naïve Bayes, k-nearest neighbour, random forest and SVM were compared to assess geological types of rock. Random forest was found as the best performing approach.

Some ML systems found their practical application in the development of tunnelling in coal mines. Tajdar (2006) reported that ANN were suitable for predicting the ultimate pile bearing capacity in identified area of the texture of soil. The authors showed that the proposed ANN had a better operation in comparison with commonly used methods.

There is a considerable amount of literature on the application of ML in mineral processing and high-dimensional automation. A ML–based assessment of the effectiveness of mineral processing was studied in Hodouin et al. (1989) where trained ANN were used as adaptive neural controllers for real-time control of processing plants. Several fundamental studies of the use of ML for tuning machines and equipment were provided by Hales & Ynchausti (1992). The authors propose an application of ANN for learning complex non-linear relationships involved in SAG mill.

Gouws & Aldrich (1996) introduced an AI-based back-propagation algorithms to exploit information from digital images of the froth phase of flotation plants. The study explores solutions to the identification of control decisions necessary to maintain optimal operation of mineral processing.

Tolwinski & Underwood (1992) propose an algorithm for the determination of a production schedule for an open pit mine that satisfies all principal physical constraints required by a realistic mine design. The authors proposed a sequential optimisation model that describes evolution of an open pit mine over time in a natural way. The authors found that the proposed model accommodated most of the constraints that a real extraction schedule must satisfy.

In our view, some of the described techniques for solving spatial variability have been too computationally demanding, time-consuming and in some instances were not supported by validation. It appears that the research to date has tended to focus on supervised learning. Although many of the described systems were simple prototypes, some systems were progressing beyond the developmental prototype phase. Review of the literature found the most frequently used ML techniques have been ANN, Fuzzy systems, Support Vector Machines and Bayesian Methods. WofE approach was a preferable method for geologically oriented authors. Regularised and Shrinkage, Kernel and Gradient Descent methods have not been considered.

Review of the published method showed that in all ML applications, cross-validation was used as an explanatory tool. To assess the optimal value of the different parameters, the predictions were derived from all possible parameter combinations and were evaluated using the Mean Square Error (and other measures of performance) using a cross-validation procedure. The best model was the one with the lowest MSE.

There is a significant commercial interest in developing AI tools that find minerals to extract, speed up the exploration and identify the most realistic economical scenario for the project. A study by Desharnais et al. (2017) on the future of exploration targeting using ML concludes that although ML is a relatively new technique, many exploration companies are excited about the prospect.

Search through Scopus and Web of Science revealed a large number of studies on managing geological heterogeneity with unsupervised learning. For instance, Cracknell

and de Caritat (2017) examined the application of UML to highlight areas that potentially host previously unrecognised Au mineralisation.

2.7 A Review of R–Packages

Where new tools, which can help improve the efficiency of discovery of new deposits should be looked for?

For over 20 years, R has had an increasing number of contributed packages for handling and analysing spatial data. Since this thesis investigates the application of regression methods to solving high-dimensional problems, the implementation of the R-language for statistical computing and visualisation is seen as an essential component required for the design of a spatial method, based on penalised regression.

First, spatial modelling with R is supported directly by the contributed add-on packages. The user often meets a problem: when R loads a new package, errors can appear due to conflicts between packages, that is, several packages might have functions named similarly.

R–default conflict management system usually gives the most recently loaded package precedence, but not in all cases. R-environment produces a warning if that new package contains any functions that are already present in the system. But, again, in some cases, it is hard to detect conflicts, particularly when introduced by an update to an existing package.

Therefore, another raised question was "how to find out which packages are potentially conflicting?"

A review of spatial analytical tools and methods of handling spatial data through CRAN website and open-source libraries revealed that the coverage of R-packages ranges from standard techniques to new developments. A few R for spatial interpolation packages are available for download from CRAN: `gstat`, `sgeostat`, `fields` and `deldir`. Several interesting statistical R-packages and solutions to spatial prediction and visualisation have been found in Bivand et al. (2013).

Bivand et al. (2013) recommends `meuse.grid` interpolator, packages `idw` and `spatstat` for Inverse Distance Weighted. A function `lm` is suggested for spatial prediction with linear regression. A function `predict` is recommended as the provider of confidence intervals for a given confidence level. The authors recommend considering a multi-variable `gstat.cv` function. A study by Paciorek (2008) suggests functions `mgcv` and `SemiPar` to fit penalised likelihood smoothing models rather than doing kriging.

Two interesting mapping packages `sp` and `sf` appeared on CRAN at the end of 2016. However, they are under very active development and a little number of examples of successful applications are available.

Since this thesis deal with large multivariate data set containing a number of variables superior to the number of samples, attention has been paid to the packages based on *penalised regression*.

At the beginning of 2016, CRAN published a penalised regression-based clustering package `prclust`, in which unsupervised clustering is performed through penalised regression with grouping pursuit. The built-in cross-validation provides an approximately unbiased estimate of the prediction error.

The `glmnet` package was suggested by the R Studio Community and Machine Learning libraries. The packege is described as efficient for solving complex high-dimensional

problems. The package is described in details in Friedman et al. (2008), Tibshirani et al. (2010), Simon et al. (2011) and presented as fast algorithms for estimation of generalised linear models with convex penalties. This package is for making a variety of predictions from the fitted models and fitting multi-response linear regression, is a package that fits a generalised linear model via penalised maximum likelihood. By default, `10-fold` cross-validation is used to find the best model among the competing models.

Loosely speaking, `glmnet` is the R package that fits generalised linear models penalising the maximum likelihood with both the LASSO, ridge and also the mixture of the two penalties (the elastic net). To determine the minimum the `glmnet` uses cyclical coordinate descent.

CRAN suggests function `mvtnorm` to compute multivariate probabilities, quantiles, random deviates and densities as an efficient add-on to `glmnet`. Another potentially promising package was `matrix` which supports many classes of matrices, including symmetric, triangular, diagonal, both dense and sparse matrices and with pattern.

No precedents of a conflict between the selected packages have been discussed in the literature and online sources.

2.8 Data Analysis

The literature indicates that today, hardly any decision is made without the help of some sort of data. However, mistakes are made. For this reason, the art of extracting nontrivial information from data is bound to take major importance. This is the purpose of data analysis.

This is focused on three particular aspects of data analysis: classification, clustering, and pre-processing.

Classification (usually supervised learning): this problem is one of the most widely studied in the field of data analysis. It may be necessary to be able to guess a characteristic using the features. A dataset is separated into several groups, called classes, according to one of its features. The goal of supervised classification is to elaborate an algorithm to assign a new observation to one or several of the classes. A supervised learning algorithm analyses the training data and produces an inferred function, which can be used for mapping new examples. An optimal scenario will allow for the algorithm to correctly determine the class labels for unseen instances. More information, examples and source codes can be found in Harvey & Fotopoulos (2016) and Pimpler (2017)

Clustering (unsupervised learning): the problem here is less definite one than the supervised classification problem. It consists of finding “clusters", that is to group the records by similarity. The goal of clustering is to find homogeneous subgroups within the data. Intuitively, patterns within a valid cluster are more similar to each other that they are to a pattern belonging to a different cluster. Another word, similarity between observations is defined using some inter observation distance measures including Euclidean and correlation-based distance measures.

Peter Williams (see Williams, 2002, sec.4.1.) corresponds unsupervised learning to a form of probability density estimation. He described the application of the generative topographic mapping method to a high dimensional data visualisation where observations were modelled as noisy expressions of the state of underlying latent variables.

Other examples can be found in Kiaei et al. (2015) on reservoir characterisation, Ducart et al. (2016) on mapping iron oxides with UML and in Cracknell & de Caritat

(2017) on catchment-based gold prospectivity analysis.

Pre-processing. Prior to applying a clustering algorithm, it is necessary to perform pre-processing. When the characteristic can be numerically encoded (for example, distances between samples or gold grades in g/ton), different units in which the data is entered might induce different behaviours of the clustering algorithms.

It is now sensible to discuss "clustering" technique as the appropriate type of data analysis for solving such a complex prediction problem as exploration targeting.

The term cluster analysis is about discovering groups in data and it is the generic name for a variety of procedures that can be used in ML to create classification (Everitt et al., 2011). Cluster analysis finds clusters in the data such that observations are as "similar" as possible within clusters, and as "dissimilar" as it could be between clusters. In more recent literature, for example in Henning (2016) the process of clustering is described as the unsupervised classification of patterns in data.

The last three decades have witnessed a huge growth in developments in clustering. Thus, the number of published applications on this type of analysis in all scientific fields has doubled approximately once every four years, and thus the rate of growth is much faster than of even the most rapidly grown disciplines.

The literature suggests two reasons for that rapid growth:

1. the development of data mining and
2. solving clustering as a scientific problem with the AI tools.

A survey on clustering found that *cluster analysis* is usually employed in empirical sciences for the summarisation of datasets into groups of similar objects, with the purpose of facilitating the interpretation and further analysis of the data. More recent studies point out that cluster analysis is of particular importance in the exploratory investigation of highly complex datasets such as drilling assays.

Importantly, cluster analysis is usually used in situations where clustering information is not observed on the data points and one wants to get this information from the data.

In this thesis, cluster analysis is a very important component of the designed UML system for clustering homogeneous observations from drilling assays.

Another frequently cited technique is the clusterwise regression analysis described by Späth (1979) and DeSabro (1989), which is used for studying the relationship between a dependent variable and a set of explanatory variables which have observations on a sample of objects. The concept is that if the samples come from different populations, the variable indexing the populations also affect the dependent variables. It was found that the regression should be performed on individual populations separately through the corresponding sub-samples observed, or by including the population effect in the model to make valid or more reliable statistical inference.

Method 1, developed and tested in Chapters 3 and 4 is focused on a specific type of the clusterwise regression analysis. It is a method that iteratively clusters data into clusters according to the available regression pattern and then updates the regression in each cluster simultaneously until some specific stopping criterion is attained.

Conventional regression techniques are usually (often successfully) applied to homogeneous observations. However, the underlying geology and the mineral content are examples where the observations are not homogeneous. In this way, another method, known as clusterwise linear regression is considered for overcoming the heterogeneity problem in regression analysis.

2.9 Summary of Chapter

Australian laws regard exploration and then mineral extraction as part of the overall mining process. Mine evaluation has remained to the present time one of the most discussed in the technical literature topics and the one, which many experts have sought consensus for many years.

A survey of published sources found that the primary goal of spatial analysis is often viewed as prediction of the spatial process at locations, where the process is not observed. Statistical methods interpolate spatially referenced data and allow the prediction of grades for arbitrary points in the area of interest. From these predictions, spatial models are created. However, opinions regarding the accuracy of some techniques appear to be debatable.

Spatial prediction or spatial interpolation aims at predicting values of the target variable over the whole area of interest and results in statistical plots or maps. The issue is that geological, litholigical and associated physical properties of ores are not distributed isotropically. This principle is not addressed well in some interpolation algorithms. In some cases, is forms basis for complaints about the accuracy of statistical prediction. Statistical methods can be very efficient on large datasets of base metals, massive sulphides, potash, etc., but they may not provide an accurate prediction of the localised or erratic gold occurrences.

Australian legislation does not outline the method by which evaluation must be done. However, the business literature suggests that ores and geological constraints have to be assessed by more than one computational method. There has been a significant commercial interest in developing new techniques that increase chances of discovery, speed up the exploration, minimise drilling cost and identify the most realistic economical scenario for a project.

Search through Scopus, WoS bibliometric tools revealed an emerging trend: solving real-world problems with artificial intelligence is one of the fastest growing fields of academic research. For example, the research into the application of unsupervised learning to real-world problems has grown in 2015-2018 almost twice as fast as research overall each year between 2010 and 2015.

However, despite a certain rise in the involvement of AI in solving mining problems, there is the lack of fundamental studies on predicting spatial distribution of the gold, optimisation of drill targeting and problems related to deposit evaluation with ML in early exploration stages. ML approaches are among most promising to be applied to the prediction of exploration targets.

The review of studies stored online in Scopus and WoS platforms revealed that there were no precedents anywhere in the literature on solving gold prediction with NSO. Search through forums and discussions related to solving nonconvex problems held on the ResearchGate platform, have not revealed clear precedents on this matter.

It is revealed that spatial variation of the gold can be formulated as a nonsmooth nonconvex optimisation problem. Another revealed problem relates to the visualisation of the predicted results generated by AI. Highly detailed visualisation of the underlying mineral content at the grassroots/ brownfield stages is important for the identification of exploration targets and delineation of resources to confirm that the benefits exceed the cost of the exploration.

The remoteness of exploration sites demands the application of free AI-based software applications that can provide exploration teams with effective solutions as to manage

the exploration data at remote areas and how drills are targeted.

For the identification of gold exploration target zones and, consequently more efficient evaluation of a resource, the problem of extracting hidden information from a drill-core assay and forecasting the behaviour of unseen structures can be tackled by two predictive methods:

1. Reformulating problem as a regression problem using nonsmooth optimisation approach.
   Goal: development of machine learning (clusterwise linear regression) model for predicting x, y, z coordinates and grades of potentially worthwhile exploration targets. The method is designed for greenfield and brownfield exploration as well as less-explored regions.
2. Reformulating problem into a high-dimensional convex optimisation problem, in which the penalisation of parameters minimises generalisation errors in dimension reduction.
   Goal: development of a method, based on LASSO penalisation. The method is designed for the application in any exploration and mining stage.

Chapter 3

Methodology

This chapter is motivated by a need for a method that provide a view of the hidden characteristics of gold mineralisation by extracting information from drilling data.

3.1 Nonconvex Nonsmooth Optimisation

Progress in the machine learning technologies is enabling highly precise data-driven predictions that allow one to make decisions based on these predictions.

The field of optimisation is very broad and has many applications. In general, optimisation is the search for an optimal valid solution with respect to a given objective function. The literature suggests that optimisation is making possible the analysis of high dimensional geological datasets.

Optimisation problems can be classified as *convex* or *nonconvex*. In a simplified form, the two types of optimisation and their objectives are shown in Figure 3.1.

Convex optimisation means that the objective function and a feasible set are convex. Therefore the search area is convex. In a strict convex case, there exists exactly one minimum, moreover it is located inside the search area.

Nonsmooth nonconvex optimisation (NSO) is viewed as one of the most difficult tasks in optimisation. It refers to the general problem of minimising or maximising functions that are typically not differentiable at their minimisers or maximisers. A nonconvex problem is usually much harder to solve.

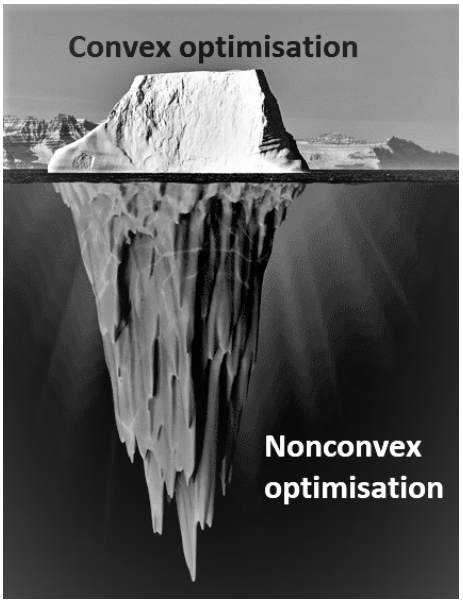

Figure 3.1. The objectives of convex and nonconvex optimisation.

In optimisation, each variable takes a numerical value and the list of variables can consequently be written as a vector $x \in A \subset \mathbb{R}^n$ where n is the number of variables. The vector x is also called a feasible solution.

The quality of a given set of variables is measured with a function q: $A \to \mathbb{R}$. The higher the value of the function q the better the quality of the parameters. The goal of optimisation is to maximise the quality.

The function q is also sometimes called *fitness function*. Alternatively, it is possible to define a cost function c: $A \to \mathbb{R}$. In this case, the goal of optimisation is to minimise the cost. Both problems are equivalent, as the cost can be defined as the opposite to the fitness: $c = -q$.

By convention, most mathematical optimisation procedures are designed to minimise. A very general formulation of the optimisation problems is then

$$\text{minimise} \quad f(x) \quad \text{subject to} \quad x \in C. \tag{3.1}$$

Here f is the objective function and C is the feasible set.

An optimisation problem may be categorised according to the form of its *objective function* f that is the function for which an optimal value is required. These functions are identified as either *linear* or *nonlinear*.

In some cases, it is practicable to locate a local minimum, and this may not be a global minimiser as shown in Figure 3.2.

Refer to the information provided by Geoscience Australia, most Australian deposits tend to occur in one, or often more than two isolated deposits surrounded by small deposits which are less predictable. Since the spatial and temporal distribution of gold deposits is extremely heterogeneous, the problem of predicting the existence such type of deposits requires the application of nonconvex functions. However, in this scenario, problem may have a large number of locally optimal solutions (Figure 3.2 and Figure 3.3)

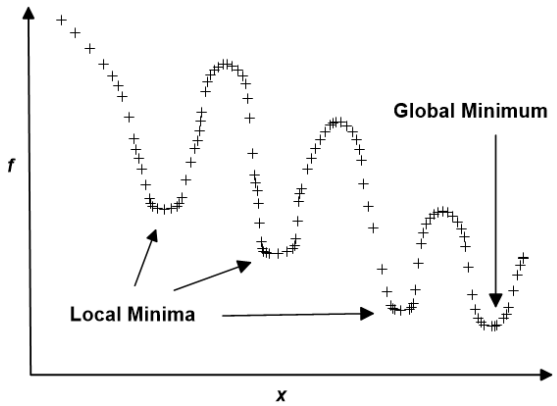

Figure 3.2. Nonconvex problem may have a large number of locally optimal solutions.

The three main challenges using optimisation models for prediction of distribution of gold deposits can be drawn from the above:

i. Prediction of the presence or absence of mineralisation in particular areas of lease.
ii. Increase the efficiency of exploration targeting,
iii. Lower the cost of exploratory drilling.

3.1.1 Precedents - what to expect?

The question is “what one can expect from the application of nonconvexity to solving geological predictive problems"?

One serious limitation of the research is that there are no clear precedents on the application of nonconvex nonsmooth optimisation methods to extracting information from core-drilling datasets anywhere in the literature.

To sum, the expectations and potential problems that may occur are as follows:

- Real-world problems can be modeled as nonconvex optimisation problems, but these problems become at least NP-hard.
- There is not a general algorithm to solve problems efficiently in all cases. In most cases approximate solutions may be sought.
- The involvement of a process that generates a sequence of approximate solutions may be required.
- NSO problems may have multiple locally optimal solutions. One gets into locally optimal solutions because of an algorithm that specifically searches for locally optimal solutions (see Figure 3.3).
- For nonconvex problems, the main drawback is the need for good initialisation (program restarts). Domain-specific knowledge to design these initialisations might be required.
- The concern is that the effect of one hidden variable cannot be expressed as a linear combination of effects of other variables.
- The concern is that the classical statistical theory may not provide a basis for the whole range of nonconvex optimisation methods.
- Predictions made by NSO models might occur outside the range of the data.
- Visualising high-dimensional solutions (when the number of predicted samples exceeds the number of samples in actual data) is not an easy task. Local solutions then, might provide an end-user with knowledge about the geometry of a nonconvex object.
- Statistically, higher variance in the data provided by nonconvexity may be expected.

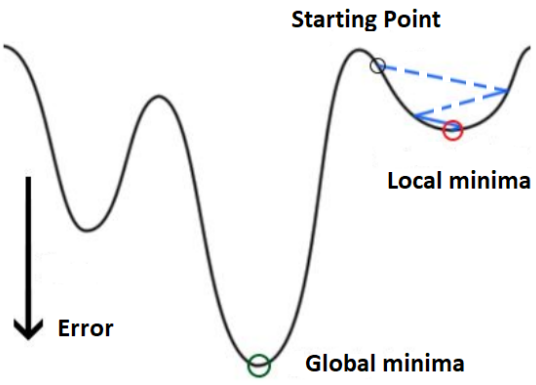

Figure 3.3. Schematic example of a nonconvex function with starting point, local minima, saddles and the error.

3.1.2 Potential problems with data visualisation

Visual representation of multivariate data, in fact, can provide insights into the structure of the underlying data. The usefulness of visualisation arises from the power of the human visual system in detecting unseen patterns. May appear, the weakest part present in some papers on solving spatial problems is in their visualisation part, where 2-D diagrams graphs and plots the level of interaction with the end-user is often minimal.

Here is another challenge: to help people from the industry comprehend the behaviour of algorithms intuitively, *how the outcomes of NSO algorithms be visualised?* Some believe that there are no other ways to visualise higher dimensional constraints or feasible sets as in 2D. Conventional visualisation methods may fail to capture the "sharpness" and "isolation" (isolated high-grade samples, other than a group of clusters) of minimisers. In the following sections, an OpenGL-based visualisation is provided to give the reader a clear view of the behaviour and properties of algorithms, including the positioning of hyperplanes (see Figure 3.6) and visual presentation of algorithm performance.

3.1.3 Cluster analysis

The term cluster analysis (CA) is about discovering groups in data. CA is usually used in situations where clustering information is not observed on the data points and one wants to get this information from the data. It is the generic name for a variety of procedures that can be used in ML to create classification (Everitt et al., 2011). These procedures form clusters of highly similar entities.

Henning (2016) defines the objective of CA as to divide a set of data points into subsets or clusters such that observations within one cluster are more similar to each other than to observations in different clusters.

In ML research, clustering is a method of unsupervised learning. This is the assignment of a set of observations into clusters (some prefer term subsets) so that observations in the same cluster are similar in some sense.

3.1.4 Regression analysis

Another frequently cited technique is the regression analysis (RA), which is used for studying the relationship between a dependent variable and a set of explanatory variables which have observations on a sample of objects. The concept of RA is that if the samples come from different populations, the variable indexing the populations also affect the dependent variables. RA is viewed as a form of predictive modelling which investigates the relationship between a dependent and independent variable, called *predictor*. This technique is often used in mineral processing for exploring relationship between the variables.

It is essential to mention that almost all known spatial techniques are based on RA in a greater or lesser extent. There are various kinds of regression techniques available to make predictions. For instance, linear regression is one of the most widely known modelling methods. For `R` environment the `lm` function is the most downloaded from CRAN. For finding the probability of event, a logistic regression is used.

In cases, when the power of independent variable is >1, a polynomial regression is worth to consider. Other methods, such as Least Absolute Shrinkage and Selection Operator (LASSO) making predictions by penalising the absolute size of the regression

coefficients. In some situations, in which independent variables are highly correlated, ridge regression can be used.

3.1.5 Clusterwise regression

This thesis is mostly focused on a specific type of regression analysis which can be referred to as clusterwise regression (CR), which is probably the most interesting technique which has great potential.

The CR is a method that iteratively clusters data into clusters according to the available regression pattern and then updates the regression in each cluster simultaneously until an equilibrium is attained.

CR is referred to estimating the class-specific hyperplanes underlying the data that randomly come from a population consisting of distinct classes. Hyperplanes (see Remark 4) do not necessarily pass through the origin in the 3D dimension (Figure 3.6).

Cluster analysis problems are divided into two classes:

1. A cluster is considered as a subject of the data points, which can be modelled adequately by a distribution from class of cluster reference distributions (CRD). These distributions have been selected to learn and imitate geological heterogeneity with respect to the data analysis problem. Therefore, CRD are often unimodal. If the class of CRD is parametric, then one is interested in the classification of the data points and parameter estimation within each cluster.
2. A cluster is considered as an area of high-density of the distribution of the whole data set. No distributional assumption is made for the single clusters.

Traditional regression techniques are usually applied to homogeneous observations. However, the underlying geology is an example where the observations are not homogeneous. In this way, another method, known as clusterwise linear regression (CLR), which simultaneously identify subgroups and associated regression functions is considered for overcoming the heterogeneity problem in regression analysis.

3.2 Method: Clusterwise Linear Regression

If in the clusterwise regression the regression functions are linear then it is called clusterwise linear regression (CLR).

There are two possible objectives when using CLR:

1. Identifying clusters of subjects that differ with respect to the set of predictors having explanatory power for the response variable.
2. Providing good predictive capability for the response variable.

To increase chance of the solvability of problem, the objective function is a combination of objectives 1. and 2.

To predict values at unknown areas we propose to apply the clusterwise linear regression (CLR) method described in Bagirov et al. (2014) for designing an algorithm to tackle the problem of predicting values between the sampled and recorded points. In this way the requirement for stationarity (section 2.4.1) can be avoided. The algorithm is able to predict probable locations and grades of gold using data directly from exploration drillholes.

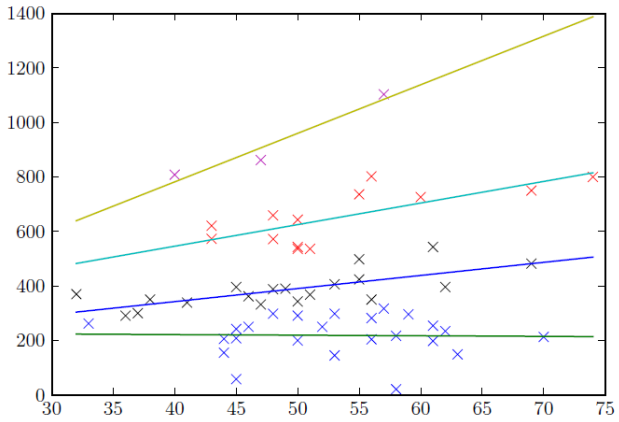

Figure 3.4. Four linear functions are used to approximate the data.

The CLR method combines unsupervised learning (or clustering) and regression techniques and finds simultaneously optimal partitions of data and regression coefficients within clusters to minimise the overall fit. Given a data

$$\mathbb{A} = \left\{(a^1, b_1), \ldots, (a^m, b_m)\right\} \tag{3.2}$$

where $a^i \in \mathbb{R}^n$ stands for an input and $b_i \in \mathbb{R}$ for its output, the aim of CLR is to divide the data into k clusters and to find regression coefficients (x^j, y_j) for each cluster $j = 1, \ldots, k$.

For regression coefficients (x^j, y_j) and a data point $(a, b) \in \mathbb{A}$ the regression error is defined as follows:

$$E_{ab}(x^j, y_j) = \left(\sum_{p=1}^{n} x_p^j a_p + y_j - b\right)^2 \tag{3.3}$$

A data point is associated with the cluster whose regression error at this point is smallest. This means that the regression error $E_{ab}(x, y)$ for the point $(a, b) \in \mathbb{A}$ is given as:

$$E_{ab}(x, y) = \min_{j=1,\ldots,k} E_{ab}(x^j, y_j) \tag{3.4}$$

Note that the regression error for a given point $(a, b) \in \mathbb{A}$ is defined as a minimum over finite number of regression errors calculated for each cluster. Then, the overall fit of the spatial predictive function is formulated as follows:

$$f_k(x, y) = \sum_{(a,b)\in\mathbb{A}} E_{ab}(x, y) \tag{3.5}$$

where $x = (x^1, \ldots, x^k) \in \mathbb{R}^{nk}$ and $y = (y_1, \ldots, y_k) \in \mathbb{R}^k$. The function f_k is called *the k-th CLR function*. For $k > 1$ it is a nonsmooth nonconvex function.

The Cluserwise Linear Regression problem is formulated as

$$\begin{cases} \textit{minimise} & f_k(x, y) \\ \textit{subject to} & x \in \mathbb{R}^{nk} \\ & y \in \mathbb{R}^k \end{cases} \tag{3.6}$$

It is a *nonconvex nonsmooth optimisation* problem. The number of clusters k is not known *a priori* and this number should be specified before solving (3.6). The number

of variables in (3.6) is $(n+1)k$ and does not depend on m, which is the number of gold samples.

Remark 2. Be noted that in the clusterwise linear regression (see (3.4)) more than one linear function is used to approximate the data. Therefore, the regression error for any point $(a, b) \in \mathbb{A}$ in (3.5) should be defined as a minimum of regression errors over all clusters. This means that one has to apply the minimum operation to define the regression error and therefore "min" operation in the definition of the regression error is not redundant. The total regression error is given as the function f_k and defined as a sum of all regression errors.

Since the problem (3.6) is nonconvex, it may have a large number of locally optimal solutions. Therefore, the success of the local search algorithm strongly depends on the choice of initial solutions (or starting poins). To find initial solutions, an algorithm introduced in Bagirov et al. (2013) is adopted. This algorithm uses the so-called *auxiliary CLR problem*. To address the nonconvexity of the problem (3.6) an incremental algorithm (IA) described in Bagirov et al. (2013) is applied.

At each iteration the Incremental Algorithm uses the auxiliary CLR problem to find a set of *initial* (starting) points. The algorithm starts with one linear function and gradually adds one linear function at each iteration. Thus, the algorithm computes sets of predictive solutions directly from drilling data.

The reason of selecting the CLR as the preferable prediction technique is that it approximates the whole drill-core assay. The use of many initial points as well as the incremental approach in the algorithm allows one to find global or nearly global solutions to the CLR problem (3.6).

3.3 Computing predictive solutions

The function, defined in (3.5), is nonsmooth and nonconvex when $k \geq 2$. The problem in (3.6) has a special structure called piecewise partial separability. This special structure is used to modify the discrete gradient method (DGM), introduced in Bagirov (2008), to solve the problem (3.6).

The piecewise partial separability allows one to reduce the number of function evaluations in the DGM. An incremental approach is used to generate starting points for the DGM. In this approach linear functions found at the k-th iteration are used to find the starting solution for the $(k+1)$-th iteration. The following remarks apply.

Remark 3. DGM is a derivative-free method for nonsmooth optimisation, which does not require the calculation of subgradients of the objective function. The DGM computes only values of the objective function and approximates its subgradients using these values. More information on the DGM can be found in Karmitsa et al. (2012) and Bagirov et al. (2014).

Remark 4. A hyperplane is a plane, which is a flat in the two-dimensional space that extends infinitely far.

3.4 Solving CLR: Incremental Algorithm

Computational learning theory can be defined as the mathematical study of efficient learning by machines. The demand for efficiency is one of the primary characteristics distinguishing computational learning theory from other techniques of spatial inference and pattern recognition in data.

The maintenance and processing of exploration data face the following challenges

1. The exploration drilling data is not usually large.
2. It is subject to many small changes (split parts of same core may provide different information)
3. The data has to be kept updated with new parameters.

Then, the question can be asked on *how do you avoid recomputing whole volume after adding new samples?*

The problem that affects non-incrementalised schemes is efficiency, which is a vital factor in the quality of output. Incremental algorithms control a sequence of input and find a sequence of solutions that build incrementally while adapting to the changes in the input. To improve efficiency the optimisation, the process of unsupervised machine learning is divided into many small incremental steps.

As previously discussed (section 3.1), the global optimisation problem may have a large number of solutions among which only global or near-global solutions are of interest.

However, some conventional global techniques cannot be directly applied to solve this problem when its size is large. From the other hand, local search algorithms can reach only locally optimal solutions, the quality of which depends on *starting points* (initial solutions) (Figure 3.3).

For efficient *local* search, it is crucial to develop a procedure for finding those good starting points from the actual drilling data. Therefore it is crucial to use a procedure for finding these starting points. This can be done, in particular, by applying incremental approach for solving general CLR problem. The algorithm incrementally finds clusters in data and approximate them using one linear function. A pseudo-algorithmic code of the Incremental Algorithm is presented in Table 3.1.

3.5 Implementation of Algorithms

The incremental algorithm gradually computes regression functions starting from one function to solve the CLR problem. Steps 1–5 in Table 3.1 avoid recomputing the whole volume of data after every small change. When the assay is expanded with new parameters and new gold interceptions are added, the algorithm saves time by only recomputing existing output. At the first iteration of the incremental algorithm, the whole data set is considered as one cluster, and one linear regression function for the data set is found. Step 2 is for computation of the next linear regression function. Then, it generates initial solutions (starting points). Finally, it builds a set of predictive solutions withing (and outside) domain.

The algorithm computes the updated data significantly faster than computing new output data from scratch.

Importantly, since the exploration drilling network grows inwards incrementally as new data gained, the algorithm learns the number of clusters and converts new input

Table 3.1. A pseudo-code of the Incremental Algorithm for finding initial solutions directly from drill hole database

Step 1	(Initialisation): Select parameters. Compute the linear regression function $(x_1, y^1) \in \mathbb{R}^n \times \mathbb{R}$ of the whole set $\mathbb{A}$. Set $l := 1$
Step 2	(Computation of the next linear regression function): set $l := l + 1$. Let $(x^1, y_1, \cdots, x^{l-1}, y_{l-1})$ be the solution to the $(l-1)$-CLR problem, find a set of solutions $\bar{A}$ to the l-th auxiliary CLR problem.
Step 3	(Refinement of all linear regression functions): for each $(\bar{u}, \bar{v}) \in \bar{A}$ select $(x^1, y_1, \cdots, x^{l-1}, y_{l-1}, \bar{u}, \bar{v})$ as an initial solution, compute their respective clusters, apply the DGM and compute a set of initial solutions $(\bar{x}^1, \bar{y}_1, \cdots, \bar{x}^l, \bar{y}_l)$.
Step 4	(Computation and recovering of hidden spatial partitions from drilling data): choose any $(\hat{x}^1, \hat{y}_1, \cdots, \hat{x}^l, \hat{y}_l) = \text{argmin } \{f_l(\bar{x}^1, \bar{y}_1, \cdots, \bar{x}^l, \bar{y}_l) : (\bar{x}^1, \bar{y}_1, \cdots, \bar{x}^l, \bar{y}_l)\}$.
Step 5	(Stopping criterion): if $l = k$, then stop. Otherwise go to Step 2.

partitions into non-stationary deterministic locally optimal solutions. Algorithm 3.1 was run in `Fortran` 95 and compiled using g95 compiler.

Figure 3.6 shows how the Incremental Algorithm learns drilling data through randomly generated and added one-by-one hyperplanes (Remark 4). Hyperplanes are generated automatically and positioned by the algorithm without humans involvement. Figure 3.6 shows such hyperplanes where the actual drilling traces shown in white. Importantly, each DGM run results in re-positioning of hyperplanes.

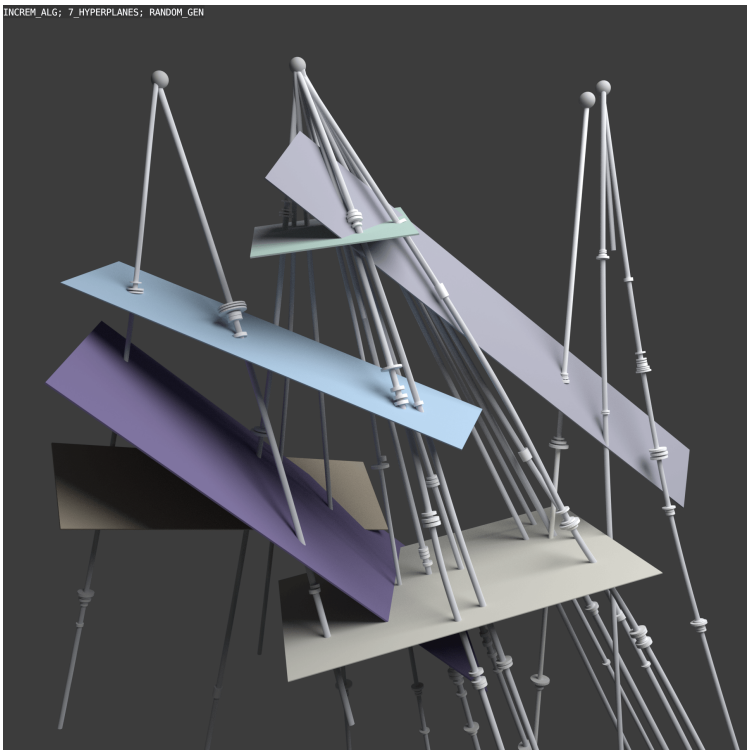

Figure 3.5. The Incremental Algorithm (3.1) learns drilling data through randomly generated and added one-by-one hyperplanes (4). Actual drilling traces are shown in white.

3.5.1 Prediction Algorithm

In order to make a prediction the following method is used: let $\ell > 1$ be a user-defined integer parameter. For any new observation (with only input data) one finds l number

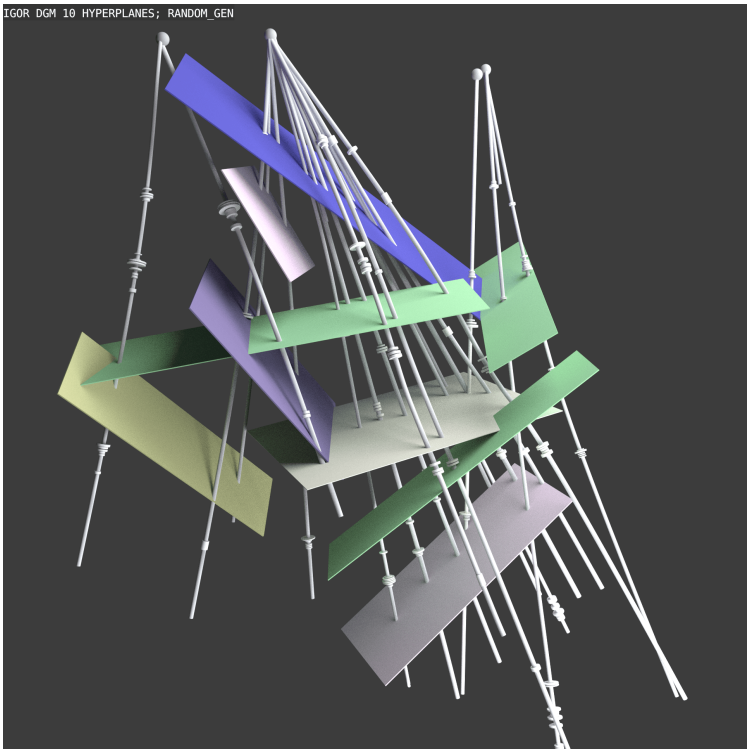

Figure 3.6. Figure shows trilling traces and 10 randomly generated hyperplanes. When the data is expanded with new information and new assayed interceptions are added, the algorithm saves time by only recomputing those outputs, which depend on the changed data.

of closest points from the training set using only input attributes. Then these points are used to determine the weights of each linear function and the prediction is made using these weights. That said, the process of prediction is one that is entirely AI controlled, subject to available drilling data.

3.6 Validation: finding the number of hyperplanes

The number of hyperplanes (linear functions) approximating the data is not known *a priori*. In order to determine the number of hyperplanes, the validation step is used. This means that the training set is further divided into two subsets: the training set containing 80% of all data used for training and 20% of the training set used for validation.

Using the new training set the IA is applied to compute hyperplanes approximating the training data and a different number of hyperplanes are used for this purpose. For each number of hyperplanes, a model is formulated to predict data from the validation set.

The flowchart of the Incremental Algorithm with the validation step is given in Figure 3.7, where the Root Mean Square Error (RMSE) is used as the measure of performance.

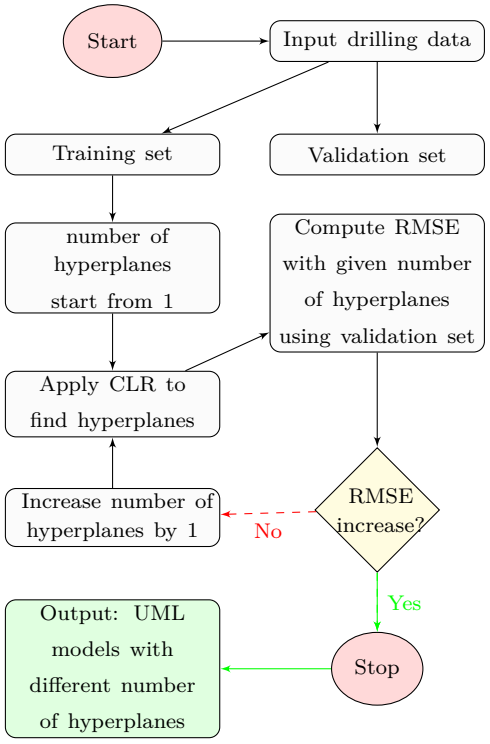

Figure 3.7. The logical structure of the adapted to solving spatial variability problems UML system where the regression mean-squared error (RMSE) is the measure of performance. Schematic representation of the matrix-free prediction procedure is broken down into subroutines according to the phases of the computation

Chapter 4

Method 1: Clusterwise Linear Regression

Results presented in this chapter have been published in the following paper: I. Grigoryev, A. Bagirov and M. Tuck "Prediction of gold-bearing localised occurrences from limited exploration data", International Journal of Computational Science and Engineering (IJCSE), Inderscience.

Detailed comparisons of predictions with realistic distribution of gold are rarely met in publicly available documents. The reason is that any references to or inaccurate statements on gold content during exploration may lead to a rapid “speculative” rise or fall in the trading price of stocks. The literature provides little evidence to suggest that the practised conventional methods used in the evaluation and direct targeting can guarantee that the predictions made reflect reality. Hence, there is no one “correct” prediction.

Machine learning, programming, pattern analysis and automatic detection systems are increasingly used to solve engineering, industrial, mathematical, analytical, medical and financial problems. Modern ML techniques emulate the human learning experience, becoming intuitive as data is sequentially processed.

Some ML techniques have the capacity to function better than human consciousness and provide an end-user with more realistic scenarios. Some extra review of research into solving nonconvex problems (e.g. Yerlikaya-Özkurt et al., 2016, Akteke-Öztürk et al., 2017; Çelic et al., 2017; Patel et al., 2018; Kara et al., 2019) indicates that numerous opportunities exist for development of new optimisation techniques, often from other fields, that may significantly benefit the efficiency of the grassroots/ brownfields exploration and improve the knowledge of the underlying mineral content.

This chapter is focused on the evaluation of the predictive performance of the CLR on an inclined deposit called "Snd". The designed UML system 3.7 is applied to solving locally optimal instances derived from the available drilling data. Since the author is under a consent agreement that is approved by the operator, for reasons of confidentiality the project and its location is not disclosed.

4.1 Convergent and Divergent Problem Solving

Guilford et al. (1956) and Campbell & Fiske (1959) state that there are two different concepts of problem solving – convergent and divergent.

Convergent means everything is coming into one area with only one solution and one outcome. There is one convergent way to solve the problem correctly.

Consider the following example of using the convergent concept for solving a problem: if one processes the same spatial data x, y, z with the same type of IDW or kriging (say simple or point kriging) using the same software – 100 consequent program runs will result in 100 of the same graphical outcome.

One of the main advantages of this concept is that the convergent methods are usually based on extensive research, often previous case studies, practical experience and tests. There are many statistical tools that can be used for data description, as part of data description and analysis of the results.

Divergent means there are many solutions to a problem and there are many ways to achieve the resulted outcome. For example, 100 UML (3.7) consequent restarts of the loop (runs) will result in 100 different graphical and numerical outcomes. Importantly, for the divergent methods a step-by-step guide to be considered or used as the basis does not necessarily exist.

Heale & Forbes (2013) state that converging results aim to increase the validity through verification and divergent findings (contradictory) can lead to new and better explanations for the phenomenon under investigation.

As a method, CLR is very well researched, described and its performance is tested against other regression techniques. But there is one nuance: in relation to the prediction of the distribution, clear precedents (section 3.1.1) on the application of the CLR to predicting gold (as well as other minerals) have not been found. In other words, the way that prediction with CLR has been done in the past shows that no correct method has been established.

The applied in this thesis CLR is purely divergent method, which is opposed to IDW or kriging. The methods are conceptually different and hardly comparable. Therefore, standard statistical tools that can be used for the description, comparison, analysis and visualisation of nonconvex output may not work.

4.2 Multiple solutions? What's next?

Since the prediction is reformulated into a nonsmooth nonconvex problem, the outcome may have a large number of locally optimal solutions. Another problem that needs to be considered is visualisation – the manner of the representation of these solutions, in which two inherent contradictions exist:

1. Combined with actual measurements the outcome should be explicitly viewed as the input for further point-cloud processing. However, without assumptions (section 2.4.1), the representation of locally optimal solutions through visual imagery in conventional style (Figure 2.2) is hardly possible.
2. The accurate reconstruction of the geometry of 3-D objects and the determination of their volumetric parameters from a limited number of solutions is problematic. The amount of data obtained from several program runs will not provide the minimal density required to form a 3-D solid.

The method for resolving these contradictions is increasing the number of solutions to a million. This resolves the paradox and provides cloud density from 1 pts/m^3 (low-resolution 369.3K vertices) to $\geq$5 pts/m^3 (high-resolution 3M vertices). Multiple UML runs will allow:

1. Creation of a raw cloud of points (for data viewing).
2. Accurate convergence into a point-cloud processing software (e.g. HighRES).
3. Further export into a CAD system such as SolidEDGE and SpaceClaim.

More specifically, this type of visualisation allows draping of a surface (similar to laying a piece of cloth over an object) over a point cloud in 4K resolution. Then, the surface is deformed to take the shape of the object along the drape direction.

To obtain volumetric parameters of an object, a raw cloud of points needs to be converted into CAD, which builds a mesh/ wireframe or meshless model. Finally, the location, orientation and volumes of an object become accurately calculated.

Correct data representation requires complex (and expensive) commercial point-cloud processing packages. Another aspect is that the development of such software requires significant engineering efforts to design, program, test and fix bugs. This research has not been funded and these packages were not accessible. As a consequence, the author cannot demonstrate the capabilities of Method 1 in its full potential.

As a solution, to represent the prediction and explain *what is this? where is that? what is going on in?*, a number of techniques from OpenGL to conventional statistics are employed to render and analyse the output. It is hoped that some of the tests conducted through this chapter have the capacity to describe, compare and categorise the output.

4.3 Organisation of Chapter

It was revealed (section 3.1.1) that there were no clear precedents on extracting locally optimal instances from exploration data as well as the 3-D representation of these instances anywhere in the literature.

An optimisator needs a visualisation tool so that the results can be viewed and clearly understood. The author is precisely aware of that it an extremely challenging task to describe locally optimal solutions visually, spatially and statistically to help the reader comprehend the internal structure of nonconvex objects and to provide all the answers in one place.

The author is aware of that an attempt to create a solid with a relatively small amount of predicted data may yield erroneous results or lead to dead-end. An attempt to reflect isolated high-grade samples, other than a cluster (or group of clusters) with traditional CAD-base systems may lead to biases in modelling. The main features of the applied method are outlined in the following sections through graphical and R-statistical explanation.

The key steps are outlined as follows:

1. Present and describe the actual data, identify high-priority regions (4.4, 4.4.1)
2. Of high importance is the answer *"what are the effective software of integration the output into real-world exploration endeavours?"* Select an accessible and location-based type of 3-D visualisation to represent the generated data coherently (4.5).
3. Prepare data for conversion into DGM solver, run program (4.6, 4.7, Appendix A).
4. Render the results through the selected type of visualisation (4.8):
    a. Visualise segments in 3-D (4.8.2).
    b. Visualise sub-segments in 3-D (4.8.4).
5. Present results by comparison, composition, distribution and relationship:
    a. Statistical features of actual and predicted data (4.8.6, 4.8.5).
    b. Compare test sets (4.9).
    c. Compare prototypes obtained from a one test set, but generated with a different number of hyperplanes.
6. Explore the effect from incrementally increasing number of hyperplanes (4.10).

7. Explore the similarity between the actual data and prediction. Explain the effect of model mimicry (if present) (4.12).
8. Draw conclusions (4.14).

In case if the effect from adding extra hyperplanes to the initially set number is insignificant, is is essential to discuss a cloud of points as the right approach to the representation of output.

4.4 Experimental Data Presentation: Geological Settings

The presented real-world exploration assay contains information on more than one element as a source of potential revenue. This applies to ferrous and non-ferrous metals such as *Cu, Ba, Co, Mg, Mn, Pb*, a number of critical technical elements and gold *Au*.

The data was obtained from the surface by "fan" diamond drilling practice (Figure 4.1), which in many cases produced multiple, but small clusters of gold close to the drilling trajectories. Thirty-eight drillholes have been drilled. To save the cost of drilling, controlled deviation of the borehole path from a parent hole was applied. The hole size has been of 96mm and the bit size used was 95.6mm. The recovered cylindrical cores have been HQ size =63.5mm from 0.3 to 1.7 metres long. Four drillholes were abandoned due to risk of flooding. Therefore, they have been excluded from the process of computation.

The assay renders very complex geometry, associated with structural deformations and forms extremely discontinuous shapes that are difficult to model and display in high resolution. No accurate engineering assumptions regarding the mineral resources were made at the time of drilling program due to the complexity of the drilling geometry and the lack of reliable data between too widely spaced drillholes and in many parts of the domain.

The existence of gold within the area, provided an opportunity for a study of the performance of the UML (Figure 3.7) on a part (70m×180m ×530m depth) of the exploration lease.

Please be noted that lithology, stratigraphy, alterations and lithochemistry have not been involved in computation because these require the application of more complex algorithms and AI techniques.

4.4.1 Testing the actual data with Discrete Fourier Transform

It is unclear whether the deposit contains sufficient gold to justify further development. The data is not enough to evaluate the exploration potential and make an investment decision. However, based on the available information, the distribution profiles ”depth from–depth to” can be reproduced with Discrete Fourier Transformation method (DFT), where the power density estimation is made by mean-square amplitude.

Although the objective of the DFT is not to locate high-grade targets, it can suggest the most promising horizontal sections of a domain. In this section, the reduction of a 3-D problem to a 2-D problem via a Fourier transform is used as a method for presenting the variation with vertical boundaries, in which the recorded gold grades are decomposed into spectrum frequencies over a continuous range.

The results of DFT shown in Figures 4.4–4.8 illustrate the correlation between Welch mean-square amplitude (WMA) and the frequency of changes per one metre of depth.

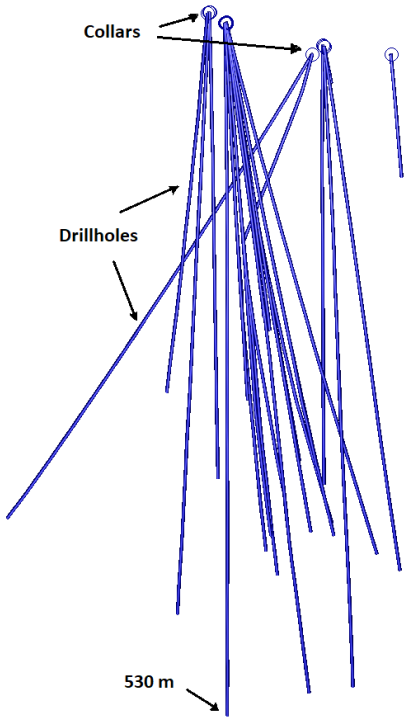

Figure 4.1. A side view of actual data with irregularly gridded fan drilling pattern in low resolution.

The frequency of changes is ranging from the surface to mRL500 (the deepest level) and indicates significant redistribution of grades with depth change.

The shapes of the curves shown in Figures 4.4–4.8 suggest that gold may occur as relatively dense dissemination throughout the host lithology from mRL200 to mRL400. The DFT indicates that insignificant concentrations from 0.0 to 2.0 Au g/t may occur between mRL0.00 and mRL200 (Figure 4.4) and between mRL400–mRL500 (Figure 4.7).

The effect of "white noise" in Figures 4.5–4.6 is equal intensity at different frequencies, providing it a constant spectral density. “White noise” is an indirect indication of the presence of potentially significant concentrations.

Pragmatically, the shapes of the curves in Figure 4.5 and Figure 4.6, indicate that the highest concentration, ranging from 2.00 to 355.3 Au g/t may lie between mRL240 and mRL340 (Figure 4.8). However, this concentration does not demonstrate feather increase with depth. It also suggests good potential for horizontal extent of Au outside lease.

The DFT suggests that approximately 50% of gold is confined to one stratigraphic interval from mRL250 to mRL350, which is being the middle portion of the domain. This interval is the target stratigraphic area, represented by relatively abrasive and unconsolidated formations. A rig operator may expect slower penetration rates than the upper and underlying levels.

Despite information is still incomplete and imprecise, it lets the exploration crew know the depth, at which targets may be located and the number rods required to intercept the targeted intervals. From an investor’s point of view, the shape of high spectral density on Figure 4.8 suggests that the mineralisation has the potential for extension beyond the exploration area, and the fees and charges for the adjacent lease might be 50% higher.

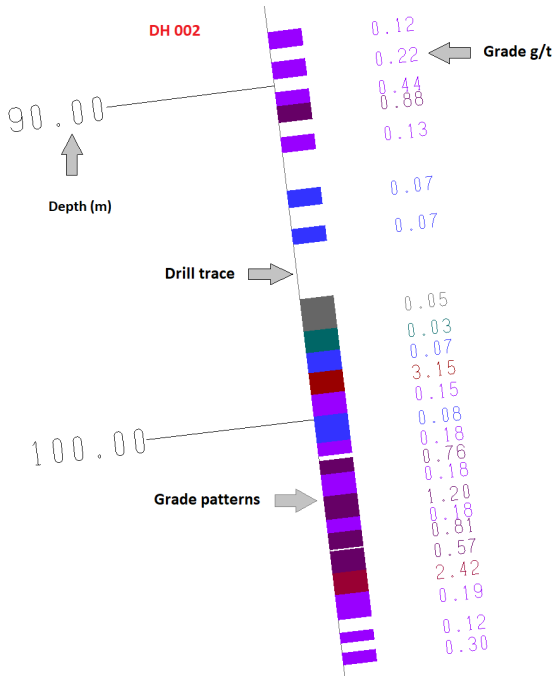

Figure 4.2. Method of visualisation 1 showing trace of drillhole DH002, depth markers, cores and the recorded grades in Au g/t. (see Method 1, Sec 2.1.1, p.12). In CAD, the intercept is expanded in Cartesian distortion view, allowing to dip in and out to see more details as needed.

From an investor's perspective, this project

1. Little evidence is available that this project has advanced enough to a stage when mineral resource estimates are determined.
2. The project has not advanced to a stage when mineral processing and metallurgical testing are appropriate.

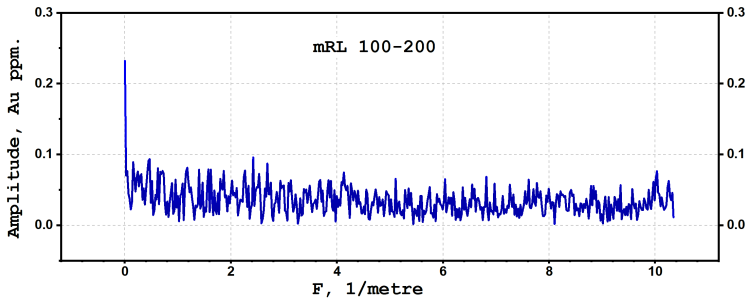

Figure 4.4. DFT mRL100–200 in the form of variation versus frequency per 1/m of depth. Minor, but equal intensities with low spectral density suggest the presence of insignificant, erratically distributed Au occurrences.

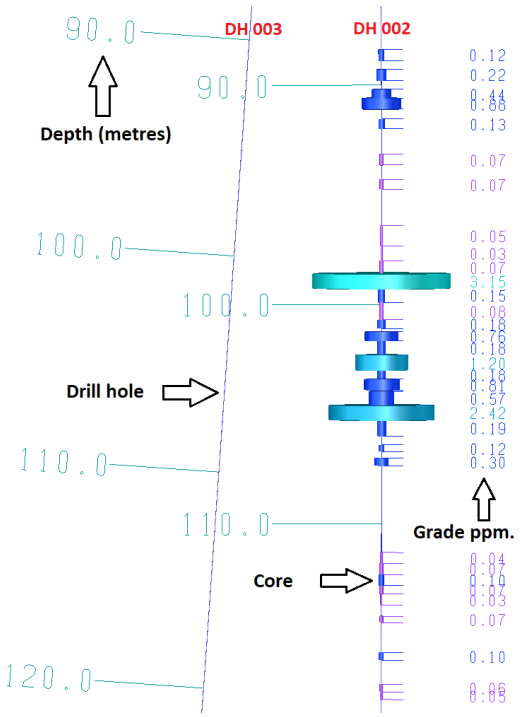

Figure 4.3. Method of visualisation 2 showing traces DH002 and DH003, depth, cores and grades in Au g/t. Gold grades along the drillhole displayed in cylinder style with variable width. Cores are in the range 0.6–1.0m in length. This method allows one to zoom into the details in low resolution

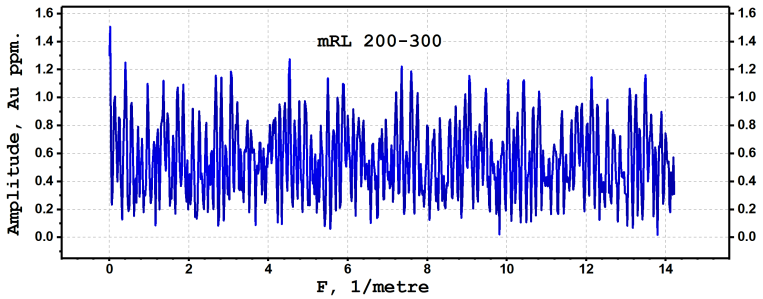

Figure 4.5. A rapid increase of Au concentrations between mRL200–300.

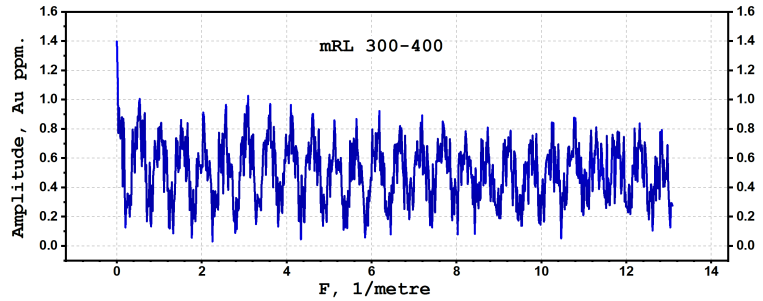

Figure 4.6. Spectrum indicates the presence of highly localised gold occurrences between mRL300–400. White noise works as an "indicator" that mirror the variation in distribution and the probable presence of gold in the form of dust and fines. Noise reduction at the right side denotes a lover variation.

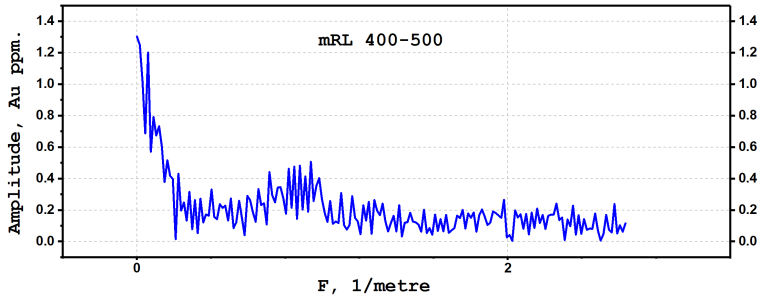

Figure 4.7. DFT mRL400-500. The Au concentrations rapidly decrease from depth ≈350 metres. A relatively Frequency per 1 metre relaxation curve suggests that the very fast decaying components are associated with smaller Au grades.

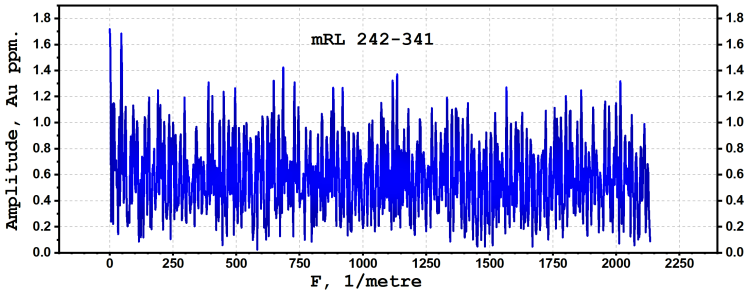

Figure 4.8. Noisy spectrum of stratigraphic layer mRL242–341 indicates high Au variation and the presence of high grades which can be associated with the epicentre of the mineralisation.

4.5 Methods of Visualisation: Actual Exploration Data

Using the data, coordinates and grades in Au g/t were assigned to 4667 *Au* recorded samples. The input variables used have been: x, y, z coordinates, the position of drilling collars on surface, the recorded length from/ to (metres), the true depth (metres), ∡azimuth° of the trajectories and ∡dip°. Please note that the drillholes in Figure 4.1 are inclined at oblique angles to intersect the veins and to maximise the information from the drill cores.

A large number of CAD/GIS software packages have been widely adopted by the industry owing to their effectiveness in abstracting out the information and understanding drilling information clearly. Generally, all these applications offer the same concept: viewing the data at different scales to see the big picture as well as to zoom local areas of interest.

It is also should be noted that accurately capture of the subsurface exploration logs and the deviation of each drillhole as one static high-resolution image is hard to visualise. The multidimensional data (as solids) is also very hard to visualise in a non-interactive way and depict the whole domain as a one static image. For example, the visualisation of the drilling grid shown on Figure 4.1 is not well suited for an accurate rendering of the geometry of traces, multiple segments and grades on a one static image (this relates to Figures/images in this thesis). Low-resolution of multiple solids also pose limitations on the accuracy of presenting the exploration pattern.

Basically, most modern packages offer two methods of drilling data visualisation

1. Method 1: the intercepts is shown in the form of grade patterns along drilling traces (Figure 4.2). This method is used for reporting (see section 2.1.1).
2. Method 2: a section of drillhole DH–002 assay grades shown in Figure 4.3 in the solid trace style, coloured and sized according to grade. This method is built-in option in many GIS applications.

Figure 4.3 has been created by method 2. It renders a fragment of a drillhole N-002 located between 85 and 120 metres below the surface. Grades of gold are shown in ppm (1 part per million = 1 gram of gold / 1 ton of rock). The gold intersections are displayed in cylinder style with variable width. The adjacent drillhole DH-003 shown no grades because it was abandoned (risk of flooding).

Method 2 is more flexible in terms of depicting drilling data using drillhole traces, grades and intervals. This type of visualisation is far from ideal: changing the visual style to show elements behind other elements in the view is required. However, this representation style is more preferable to produce visually appealing results and will be used to visualise pseudo-arrays (prototypes) that encompass the generated solutions. For example, Figure 4.3 renders a fragment of a drillhole N-002 located between 85 and 120 metres. Grades of gold are shown in ppm (1 part per million = 1 gram of gold/1 ton of rock mass).

Note that since nonconvex prototyping forms the backbone of successful visualisation, it will be focused on the application of modules that are written in Python. To do this, the visualisation (as large programming task) is broken into small, but more manageable packages available from open-source Python libraries. Prototyping by methods 1 and 2 (Figures 4.2 and 4.3) is problematic.

4.6 Predicted Data Preparation for Conversion into DGM

4567 valid samples were used for computation. The lowest gold value in data =0.01 Au ppm (0.01 Au g/t). Almost all zero values were removed by the operator to maintain continuity. The highest registered value is 355.3 Au ppm.

To facilitate the proposed UML system (Figure 3.7), the actual n-dimensional data has to be transformed into an appropriate for processing format. To do this, the following parameters were extracted from the drilling data: x, y = Cartesian coordinates, z = drilling collar elevation (metres), the true depth (m) of each drillhole, dip° and azimuth° of each straight segment of a drillhole, gold grade of each extracted core in g/ton.

Note: the reality is that drilling trajectories are not straight, many sections have a shape of “dog’s leg". If a drill bit rotates normally on a central axis, the direction of advance does not change. However, a drill hole can be unpredictable in the path that it takes. The unintentional departure of the drill bit from a planned borehole trajectory usually occurs due to change in rock density, the type of bit used, the weight on the bit, type of rods, mechanical stabilisers and the set-up of the rig. The recorded assay parameters indicate that all drillholes were not straight.

The data was extracted from the available (actual) `.accdb` file and converted into three `.txt` formats.

– inputdata1.txt: positioning of “collars” in Cartesian system, x, y and z–drilling rig set up/ collar elevation (m).

– inputdata2.txt: “survey” data with information on depth (m) of each core, true depth of each drillhole (m), azimuth° and dips° of each cores along drillholes.

– inputdata3.txt: "assay" data: the depth from/to of each side of core (m) and gold grades in ppm.

Next, the .txt files were exported into a general-purpose cross-platform Code::Blocks. For compilation into `fortran 95` format (Appendix A) a loop transformation was used.

Once compiled, the data is processed with the derivative-free discrete gradient solver (DGM) for solving nonconvex nonsmooth minimisation problems using the *training* set. The source code of the DGM for computing locally optimal solutions and sub-gradients can be found in Appendix B. Numerical experiments have been conducted on Intel Core i7-6700 CPU 4CHz and 16GB of RAM.

Two stopping criteria strategies were used:

1. The number of function evaluations is restricted to 106.
2. The program is stopped if it cannot decrease the value of the objective function in 1000 successive iterations.

4.7 Program Runs

Four program runs (DGM, Appendix B) with the involvement of a different number of hyperplanes produced four sets of locally optimal predictive solutions withing (and outside) the exploration domain:

- (Test Set 1) 1730 samples → Prototype 5 (5 hyperplanes) 1730 samples.
- (Test Set 2) 1066 samples → Prototype 7 (7 hyperplanes) 1066 samples.
- (Test Set 2) 1066 samples → Prototype 10 (10 hyperplanes) 1066 samples.
- (Test Set 1) 1730 samples → Prototype 10A (10 hyperplanes) 1739 samples.

For three-dimensional modelling of the predicted data, it was necessary to determine Cartesian coordinates x, y, z of each predicted sub-segment (virtual core) in three-dimensional space. The output format in .txt is as follows

– The elevation of drilling collars (m).

– The maximum depth of the lowest segment (m).

– Azimuth° (inclination) of each segment.

– Depth "from" (m), the depth of the upper side of each sub-segment along trace.

– Depth "to" (m), the lower depth of each sub-segment.

– Cartesian coordinates x, y, z of the upper and lower end of each sub-segment along trace.

– The length (m) of each sub-segment (virtual core).

– Grade of gold in Au g/t along each sub-segment.

4.8 Visualisation of 3-D Nonconvex Data

The standard built-in tools of commercial CAD-based applications used in the evaluation, make the visualisation of nonconvex data difficult to comprehend.

The data that an optimiser intuitively feels, required first to understand what the underlying comprehension issues might be. In relation to this study, the complexity of NSO visualisation depends on the level of representation of predicted spatial attributes

and their inter-relationships to 3-D visualisation comprehension skills of an end-user. Therefore, it is essential to discuss some aspects of 3-D visualisation comprehension.

The major concern was the inability of an engineer to associate a nonconvex nature of the predicted structures with their interior. In the author's view, the spatial attributes put into 3-D nonconvex context require a better understanding of the inter-connecting relationships. To incorporate static images into the optimisation context, it was required to select an effective strategy of representation of locally optimal solutions in 3-D.

To gain insight into the data, a number of renders with mathematical content were synthesised to demonstrate complex optimisation problems within visualised contexts. Nonconvex sets are called prototypes, or P. The number after capital letter (i.e. P5) denotes the number of hyperplanes involved into the computation.

It had been intended to display the actual data and prediction within in a one static image (in this thesis). To do this, the effect of "disproportionality" was applied to all 3-D realisations. More specifically, the radius of each <tracepath> segment has been programmatically increased by $\times 10$ times without the application of artificial extrusion.

To view nonconvex segments and sub-segments, the procedure of visualisation was done in three steps:

1. Obtain the shapes, orientation and locations of locally optimal groups of segments within the domain.
2. Iteration of nonconvex segments with theoretical target.
3. Explore sub-segment structure. To do this, a `subdivisor` modifying function is applied to avoid multiple distortions.

Figure 4.9 renders step 1 – the positioning and orientation of four sets (prototypes) of locally optimal solutions. The prototypes P5, P7, P10 and 10A were generated from four runs and initialised at different starting points. Unfortunately, the appearance of the model in Figure 4.9 is not to be somewhat complete satisfactory, because the goal to depict all four prototypes in one single render in high resolution was not achieved.

4.8.1 3-D output

The problem of solving locally optimal instances was clearly framed for the machine as to

1. recognise, learn and mimic the behaviour of the actual data
2. predict the existence of gold occurrences in the unexplored areas.

It was expected that each predicted structure would imitate structural composition of the actual data and produce sub-data such as latitude, longitude, length of predicted cores (m), from/to (m), ∡dip°, ∡azimuths° and maximal true depth (m).

Contrary to expectations, the UML took a different approach to the situation and optimised data in an entirely different way. The UML has interpreted the tasks of recognition, learning and search on its own, without human involvement and consequently, selected different approach to solving problem. More specifically, instead of strictly following the instructions set by the source code (Appendix B) to solve the problem, the UML generated its own sub-algorithm, i.e. it essentially re-programmed itself.

Instead of imitating the geometry and path of the actual drilling trajectories, locally optimal solutions were synthesised in spread-out across domain multiple unequal structures, whose long axis radiated from the actual drilling collars located on the surface.

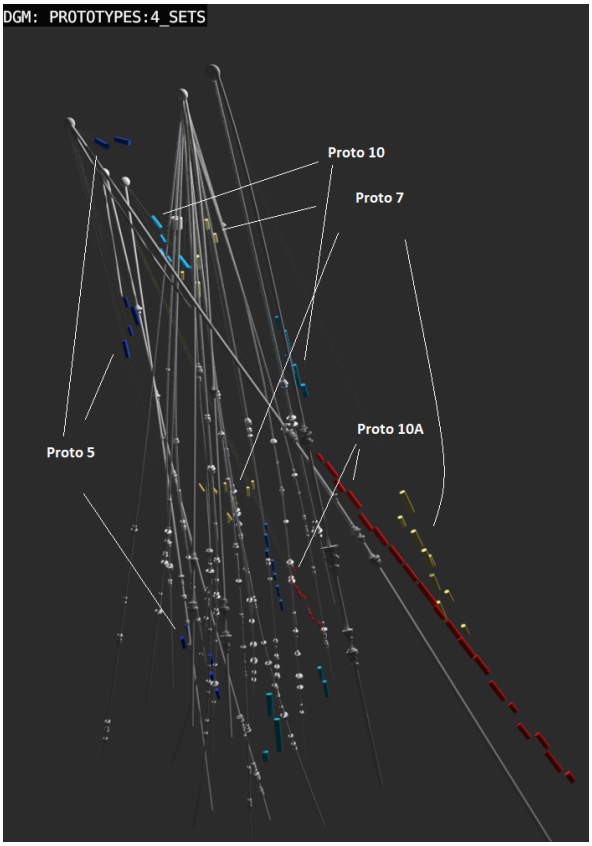

Figure 4.9. Side view of the actual drilling and predicted data looking east, with a vertical angular offset, shows positioning and orientation of 4× locally optimal sets: P5 (blue), P7 (yellow), p10 (cyan) and P10A (red).

4.8.2 Visualisation of segments

There is a graphical limitation that affects rendering time - a large number of objects (>24fps). To avoid system lags, a conflict between the multiple packages and decrease rendering time, each segment shown on Figures 4.10–4.13 was digitised as a solid. Each segment consists a number of sub-segments (virtual cores), which will be shown and discussed in section 4.8.4.

An advantage of the applied style was that it made available practically unlimited perspectives from which to view the generated segments. The model on Figure 4.9 may be rotated, panned and zoomed in countless ways, it also accommodates some important details in one image.

The generated group of segments are located away from the actual drillholes and point at zones with the presence (and absence) of gold. The trajectories in Figures 4.10 and Figure 4.11 tend to radiate from a single point on the surface (see collars in Figure 4.10) winding down in a continuous and gradually widening curve to form helix-like shapes. The angular deviation has been observed as from 0.0061 to 0.0366 `rad`.

4.8.3 Iteration of segments with theoretical target

Figure 4.12 is a 3-D side view of a theoretical target digitised as a textured solid and virtually intercepted by the actual traces and nonconvex segments. Figure displays colour-coded groups of the segments, which are the parts of 4×prototypes obtained from 4× program restarts. Theoretical targets shown in Figures 4.12–4.13 are complex

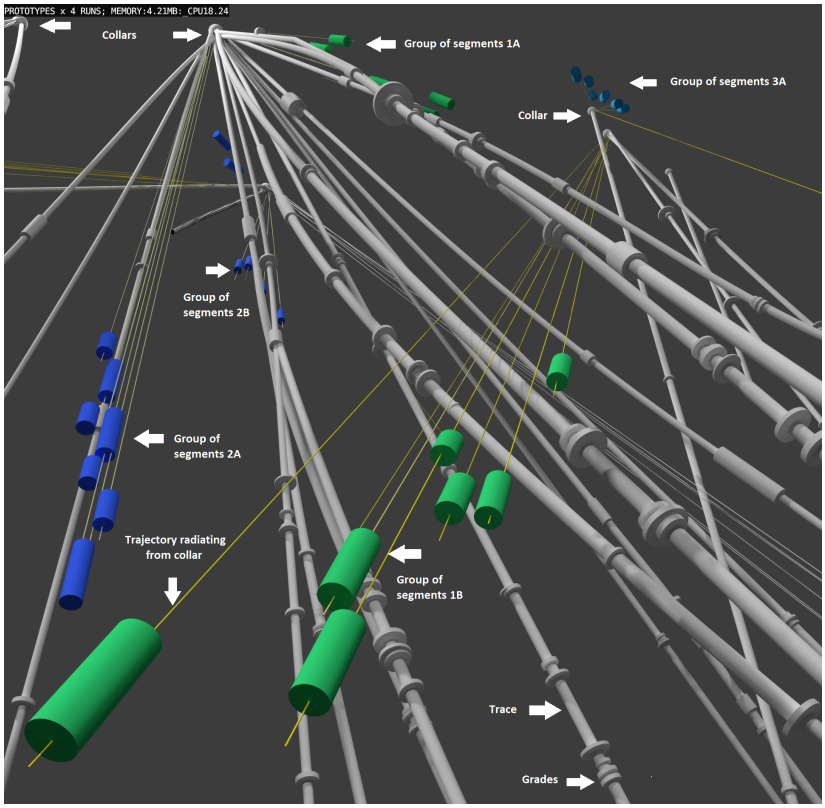

Figure 4.10. A view from below showing three groups of segments obtained from 3× program (DGM loop) re-runs. Each segment is 0.6–16.0 metres in length and contains information on Cartesian x, y, z coordinates and grades. Actual drilling traces and attributed grades shown in white. To make small details visible, segments and grades are captured by zooming.

geological structures located ≈290 metres below the surface. The assay suggests the presence of a number of voids within domain, which can pose a serious problem to ore extraction, due to the possibility of a collapse of caving. That means, some parts of the deposit cannot be mined due to high cost of support. Both actual and nonconvex sets show some signs of potentially economic mineralisation within the deposit. However, the data obtained from several program runs is not complete enough to be converted into a CAD solid.

4.8.4 Visualisation of sub-segments

To visualise nonconvex sub-segments shown on Figures 4.15–4.16, a `subdivision` function was applied. The images render sub-segments as associated gold content in a segment. The attributed grades are set as from 0.5 to 20.0 Au g/ton. One can see that each segment is represented by both high, ultra-low and 0.0 grades.

Sub-segments imitate the behaviour of the actual cores in terms of length (metres) and composition (grade). The synthesised grades point towards the locations where the gold occurrence may or may not exist. The sub-segments are designated by UML as predicted gold samples. One may notice Figure 4.16 a multiple empty “gaps", or zero grades in between the sub-segments.

Knowing Cartesian coordinate x, y, z of each predicted sub-segment, a model such as shown in Figure 4.14 can be created (as shown on Fig 4.3) with a standard CAD tool. Figure 4.14 shows the structural composition of the longest part of P10A, which is

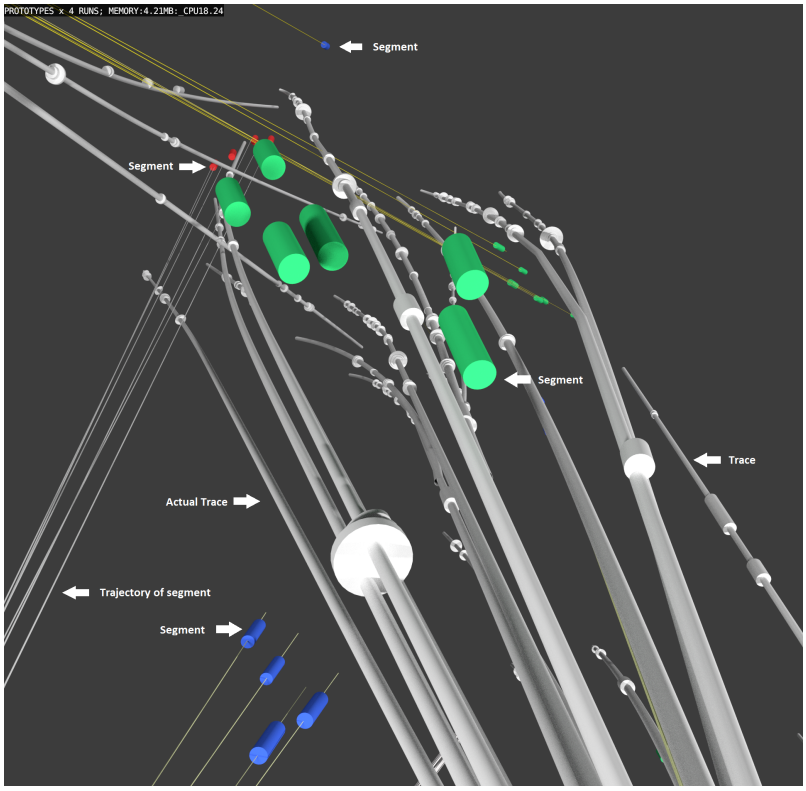

Figure 4.11. A view from above illustrating the positioning, orientation and geometry of the predicted groups of segments. Actual drilling traces and attributed grades are shown in white.

311.5m in length. The part consists of 18 segments split into a number of sub-segments ranging between 0.6–1.0 metre in length. UML points at the likelihood of the presence of barren rock along the path and probable high-grade localised occurrence (308.24 Au g/t shown in red). Gold grades in the colour-coded sub-segments are shown in the solid-trace style sized according to Au grade.

4.8.5 Statistical features of the actual and predicted data

The project is relatively small grassroots/brownfield endeavour. Deep drill holes were developed on nominal 70m×180m×530 partings. The available data is not complete enough to make an investment decision. Therefore, there is a risk that benefits will not exceed the associated cost of drilling and delineating the deposit. Basic statistics presented in Table 4.1 summarises main statistical features of the actual and predicted data. It has to be noted that all 0.0 Aug/t values have been deleted from the actual data by the operator. The lower limit of detection was set by the operator as 0.01 Au g/t.

Table 4.2 presents the outliers (extreme values). However, all the recorded in the assay high values, including the extreme once, are considered to have high importance for making predictions by optimisation. Therefore, these high values have not been removed during data preparation for conversion into the DGM solver.

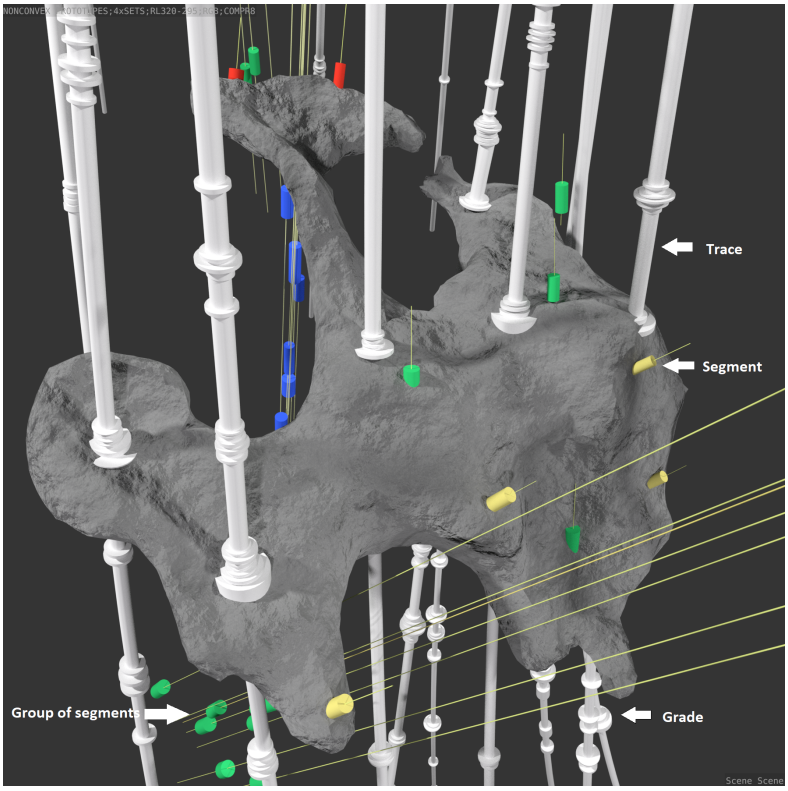

Figure 4.12. A side view showing theoretical exploration target, predicted segments, groups of segments and actual traces. Each segment shown in a cylinder style. Solid obtained from IDW (actual and predicted data combined in a one set): textural fractals =1.6

Table 4.1. Statistical representation of the actual and predicted data.

Parameter	Actual set	P5	P7	P10	P10a
N total	4667.00	1730.00	1066.00	1066.00	1739.00
Mean	1.045	1.604	1.005	1.057	1.524
Standard Deviation	8.982	10.579	10.225	10.238	10.118
SE of mean	0.131	0.254	0.313	0.314	0.243
Lower 95% CI of Mean	0.787	1.105	0.390	0.442	1.048
Upper 95% CI of Mean	1.303	2.103	1.619	1.673	2.000
Variance	80.678	111.908	104.555	104.807	102.370
Sum	4877.020	2775.355	1070.812	1127.120	2649.920
Skewness	27.200	21.537	27.243	27.127	22.629
Kurtosis	875.266	552.878	796.648	792.084	612.317
Uncorrected Sum Squares	381539.195	197940.798	112427.014	112810.888	181957.117
Corrected Sum Squares	376442.704	193488.431	111351.367	111619.144	177919.121
Coefficient Variation	8.595	6.594	10.179	9.682	6.640
Mean abs Deviation	1.444	2.196	1.411	1.454	2.138
SD times 2	17.964	21.157	20.450	20.475	20.236
SD times 3	26.946	31.736	30.676	30.713	30.353
Mode	0.020	0.375	0.000	0.000	0.230
1st Quartile (Q1)	0.040	0.357	0.187	0.105	0.180
Median	0.120	0.433	0.236	0.137	0.230
3rd Quartile (Q3)	0.530	0.503	0.287	0.163	0.280
Maximum	355.300	308.976	309.720	309.702	308.240
Index of Max	4667.000	946.000	282.000	282.000	955.000
Interquar Range (Q3 - Q1)	0.490	0.147	0.100	0.059	0.100
Range (Max - Min)	355.290	308.989	309.720	309.702	308.240
Median Absolute Deviation	0.100	0.074	0.050	0.033	0.050
Robust Coeff of Variation	1.236	0.253	0.314	0.351	0.322
P90	1.560	0.591	1.177	1.286	2.330
P95	2.600	5.646	2.647	3.032	3.800
P99	12.800	29.283	9.939	11.465	23.240

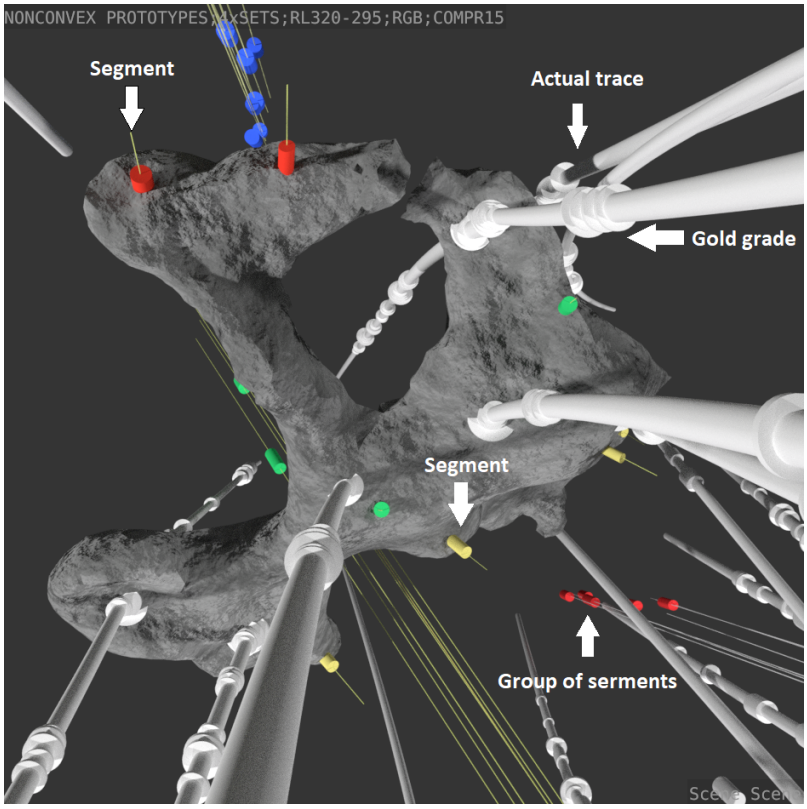

Figure 4.13. A top view showing nonconvex segments (colour-coded), groups of segments, actual drilling traces and grades, textural fractals =1.8

4.8.6 Actual Drill Core Data

A number of extreme and ultra-low values were identified as statistical outliers (Table 4.2). The data in Table 4.1 shows that grades in the actual set range from 0.01 to 355.3 Au g/t, and the extreme high grades are presented by a few samples (Table 4.2; Figure 4.23).

The bean plot in Figure 4.25 shows the existence of a number of high grades ranging from 0.01 to 355.3 Au g/t. (Berkman, 2001), (Selley, Cocks, & Plimer, 2005) suggest if the mineralisation is hosted by quartz-sulphide veining, the middle and high grades can be discovered much further from its location than previously thought.

The histogram in Figure 4.17 shows that most bins are concentrated between 0.01 and 1.7 Au g/t. Table 4.1 and Figure 4.17 indicate the prevalence of low grades (<0.50 Au g/t Au) and the presence of significant grades (between 0.50 and 1.7 Au g/t).

Box plot shown in Figure 4.24 is scaled to range 0.0–32.0 Au g/t. The distribution, the range of Standard Deviation and mean =1.045 clearly indicate that low-grade gold is dominant a occurrence, which is not necessarily localised. Data suggests that gold is highly disseminated and may exist in the form of dust, grains and small fines everywhere within the explored domain.

In practice, gold grades less than 0.5 Au g/t are usually not considered "material" as exploration drill targets due to their low grade. However, the intersections >0.5 Au g/t (see Figures 4.24 and 4.17) with up to 1 metre of internal barren rock are considered as significant and therefore, reported. From an engineering perspective, the presence of the potentially worthwhile interceptions is encouraging. Significant assays >5.0 Au g/t are usually reported separately as contained within the broader lower grade intervals.

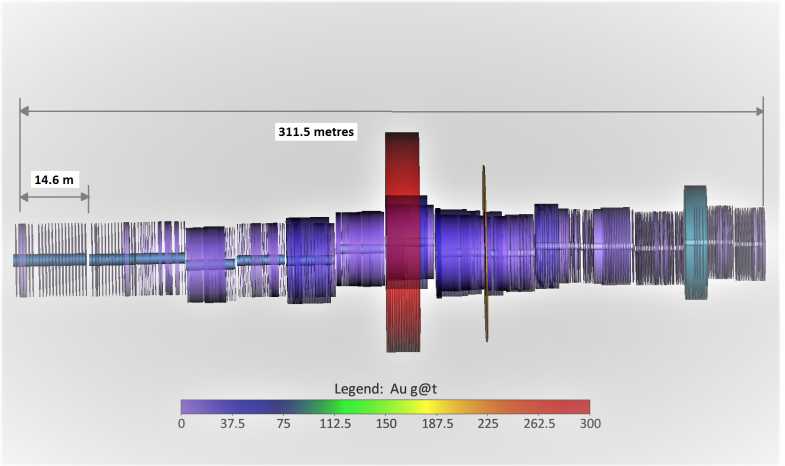

Figure 4.14. Figure shows the composition of the longest fragment of prototype 10A. Total length = 311.5m, 18 segments split into a number of sub-segments ranging between 0.3m–1.0m in length, visualised by method 1 (Sec 2.1.1). UML points at the likelihood of the presence of barren rock along the path and probable high-grade localised occurrences between RL340 and RL380.

However, due to a limited number of high grades >50 Au g/t in the actual assay, possible extensions, directions and the plunge of mineralisation are remain unclear. Consequently, the exploration target area may not be reliably identified at this stage. The amount of available drilling and predicted data is not enough and can yield probable erroneous gold distribution. More drilling is required. This project needs to be advanced to a stage when mineral reserve estimates are determined. The data is not complete enough to form the basis for ASX release.

It has to be noted that this project can be sold, bought, joint ventured at any time as is (change owner), without reporting. However, in this scenario, a new owner will have to invest in extra exploration. Once more reliable samples >2.0 Au g/t obtained, and when sufficient knowledge on the thickness of probable intersections are improved, an estimate of the averaged thickness of mineralisation can be assumed.

JORC 2012, s. 17 recommends that for grassroots projects, results must be represented only as approximations. To do this, for example, weighted average technique (Hazen, 1967, p.162), which is an average where each value has a specific weight or frequency assigned to it, can be used to determine the grade of the anomalous intervals when these intervals less than 1.0 metre have been sampled.

The capture of stable grades >8 Au g/t along 5 metres core may significantly increase sell price of this project in the open market. Obviously, more drilling is required to evaluate the exploration potential, and the exploration crew is motivated to identify higher grade zones (footprints), minimising the amount of drilling. In this case, a point cloud created by Method 1 helps the crew in achieving these goals.

4.9 Comparing test sets TS1 and TS2

The goal of this section is to visualise two test sets so that they are distinguishable and can be graphically compared. To overcome some graphical limitations, to allow nonlinear data manipulation to convert the fitted regression model into graphics, two

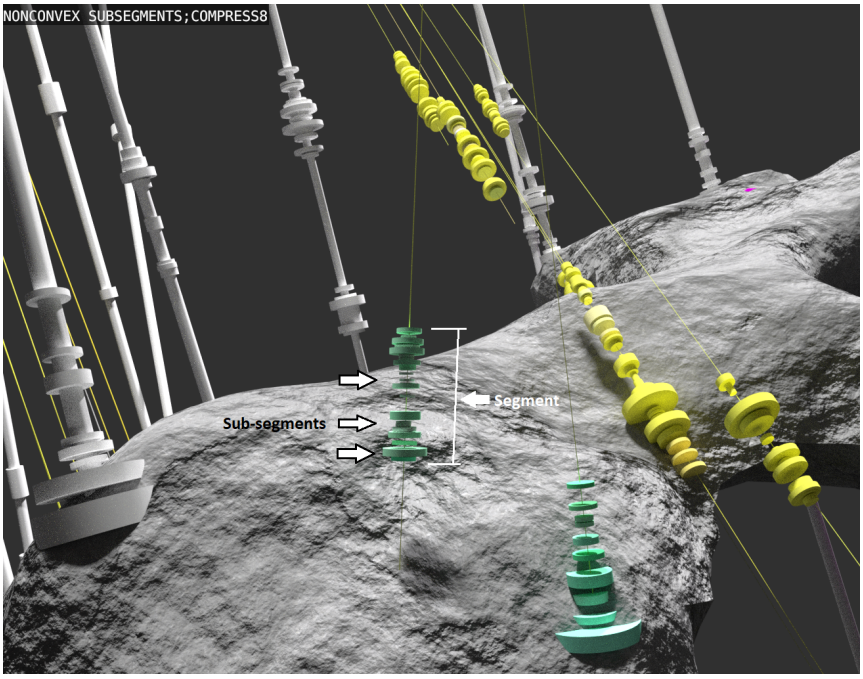

Figure 4.15. An oblique view showing nonconvex sub-segments (colour-coded), segments, groups of segments. Grade attributes range from 0.5 to 20.0 Au g/ton and shown in the solid-trace style sized according to grade. Solid obtained from IDW, textural fractals applied =1.8

extra packages `tdyr` and `dplyr` were applied. Computations in Listings 4.7–4.12 approximate the errors that make the relationship between TS1 and TS1 values relationships.

The representation of TS1 and TS2 features summarised in Table 4.3 is essential to proceed.

Bean plot in Figure 4.18 compares two distributions side-by-side. Table 4.3 highlights the highest values in TS1: 83.2; 45,7; 27.2; 22.6 and 18.9 Au g/t. The highest values in TS2 are 27.2; 22.6 and 13.5 Au g/t. The similarity is that both TS contain 0.0 values. The histograms in Figure 4.19 and Figure 4.20 indicate that ultra-low grade beans dominate the others.

Not much evidence available to suggest that TS1 and TS2 are similar.

The equivalence test (equal variances not assumed) in Listings 4.1 suggests that means in TS1 and TS2 are not equivalent. However, the result of t-test indicate that at 0.05 level, the difference between TS1 and TS2 is not significantly within [0.0;0,1].

Side-by-side TS1 and TS shown in Figures 4.21–4.22 are polar Kernel density plots, in which the densities of mean distributions are presented as spatial patterns. The compared patterns clearly indicate that the compared sets are distinct in terms of the shape and positioning of KDE patterns of mean density distribution. The shapes and positioning of the patterns give ground to suggest that TS1 and TS2 are different. Please be noted the application of KDE will be discussed in section 4.10.2.

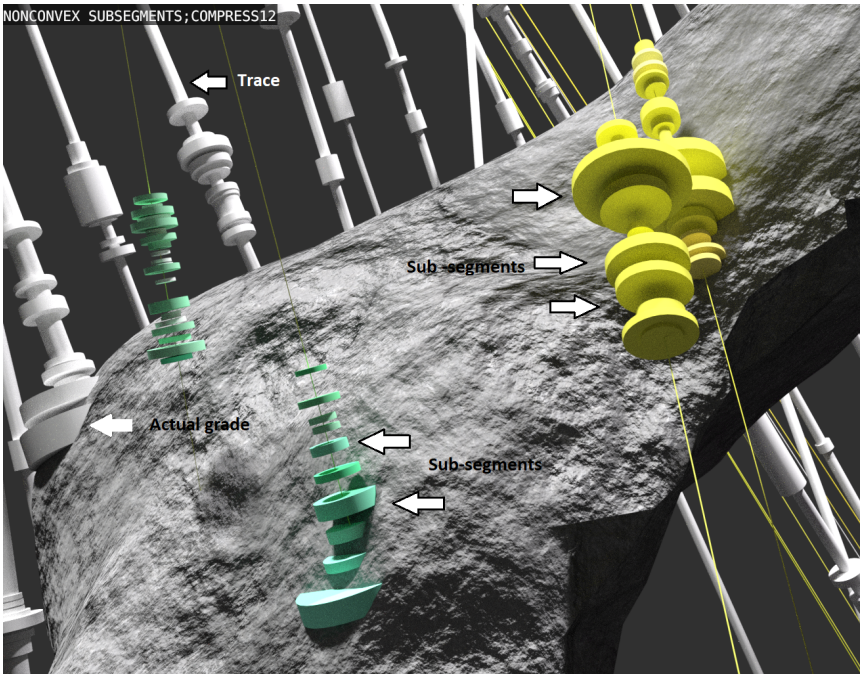

Figure 4.16. An oblique view showing iteration of sub-segments with theoretical target. Grade attributes range from 0.5 to 20.0 Au g/t. Higher grades refine the shape of target. Each program restart generates only one prototype, i.e. a group of segments (yellow, red, etc.)

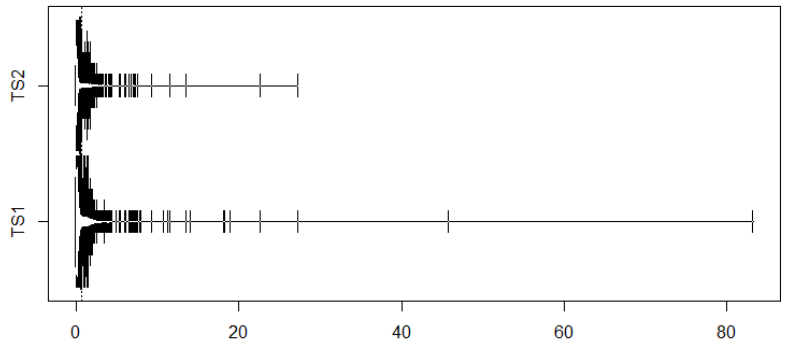

Figure 4.18. Side-by-side comparison of bean plots for TS1 vs TS2, scale 0.0–90.0 Au g/t

Table 4.2. Identified extreme values in the actual set, P5, P7, P10 and P10A. Note: 0.0 values were removed from the actual assay by the operator to keep continuity.

		Index	Value
Actual set	Highest	2367	355.300
Actual set	Highest	3263	289.000
Actual set	Highest	965	210.500
Actual set	Highest	1664	176.800
Actual set	Highest	2063	148.900
Actual set	Lowest	4957	0.01000
Prototype 5	Highest	946	308.975
Prototype 5	Highest	519	232.022
Prototype 5	Highest	663	121.418
Prototype 5	Highest	1446	65.4085
Prototype 5	Highest	934	43.2139
Prototype 5	Lowest	1433	0.00000
Prototype 5	Lowest	769	0.01920
Prototype 5	Lowest	772	0.04570
Prototype 5	Lowest	773	0.05460
Prototype 7	Highest	282	309.719
Prototype 7	Highest	782	115.909
Prototype 7	Highest	608	14.6854
Prototype 7	Highest	251	14.6854
Prototype 7	Highest	92	13.5109
Prototype 7	Lowest	314	0.00000
Prototype 10	Highest	282	309.702
Prototype 10	Highest	782	115.914
Prototype 10	Highest	608	13.9418
Prototype 10	Highest	251	13.9418
Prototype 10	Highest	92	13.4627
Prototype 10	Lowest	377	0.00000
Prototype 10A	Highest	955	308.240
Prototype 10A	Highest	528	218.590
Prototype 10A	Highest	1455	86.4600
Prototype 10A	Highest	672	72.1900
Prototype 10A	Highest	1274	42.0600
Prototype 10A	Lowest	1050	0.00000

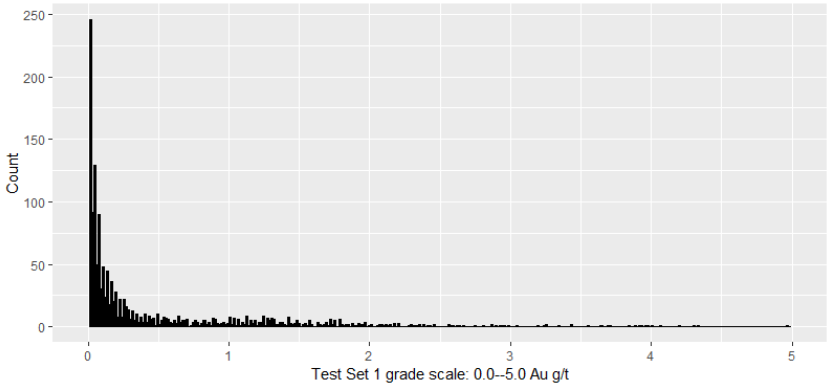

Figure 4.19. Histogram of TS1, scale 0.0–5.0 Au g/t

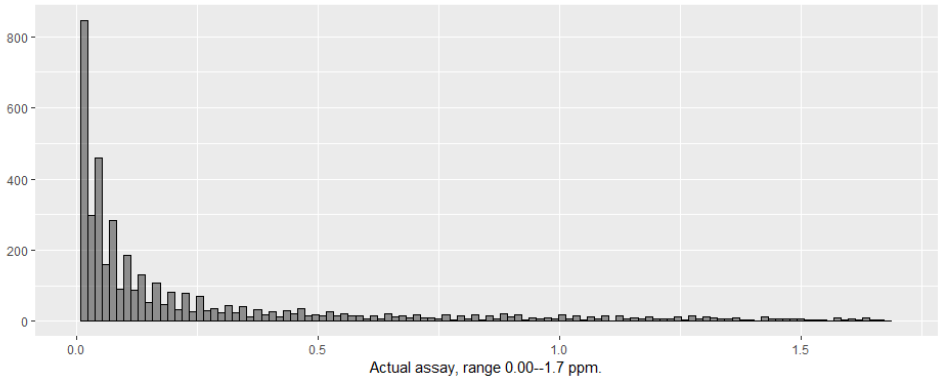

Figure 4.17. Histogram showing the distribution of grades in the actual assay, n =4667, scale 0.01–1.80 Au g/t

Table 4.3. Comparison of statistical properties and outliers for TS1 and TS2

	TS1	TS2			Index	Value
N total	1730	1066	TS1	Highest	129	83.20
Mean	0.659	0.572	TS1	Highest	352	45.70
Standard Deviation	2.782	1.545	TS1	Highest	1625	27.20
SE of mean	0.067	0.047	TS1	Highest	1070	22.60
Lower 95% CI of Mean	0.528	0.479	TS1	Highest	590	18.90
Upper 95% CI of Mean	0.791	0.665	TS1	Lowest	376	0.00
Variance	7.742	2.387	TS1	Lowest	1209	0.00
Sum	1140.610	609.800	TS1	Lowest	876	0.00
Skewness	19.187	9.473	TS1	Lowest	382	0.00
Kurtosis	497.495	130.691	TS1	Lowest	18	0.00
Coefficient of Variation	4.220	2.701			Index	Value
Mean absolute Deviation	0.851	0.720	TS2	Highest	961	27.20
SD times 2	5.565	3.090	TS2	Highest	406	22.60
SD times 3	8.347	4.635	TS2	Highest	774	13.50
Sum of Weights	1730.000	1066.000	TS2	Highest	285	11.50
Minimum	0.000	0.000	TS2	Highest	942	9.29
1st Quartile (Q1)	0.020	0.020	TS2	Lowest	691	0.00
Median	0.090	0.090	TS2	Lowest	690	0.00
3rd Quartile (Q3)	0.540	0.540	TS2	Lowest	689	0.00
Maximum	83.200	27.200	TS2	Lowest	688	0.00
Interquart. Range (Q3 - Q1)	0.520	0.520	TS2	Lowest	692	0.00
Range (Max - Min)	83.200	27.200				
P90	1.530	1.630				
P95	2.210	2.170				
P99	7.420	6.700				

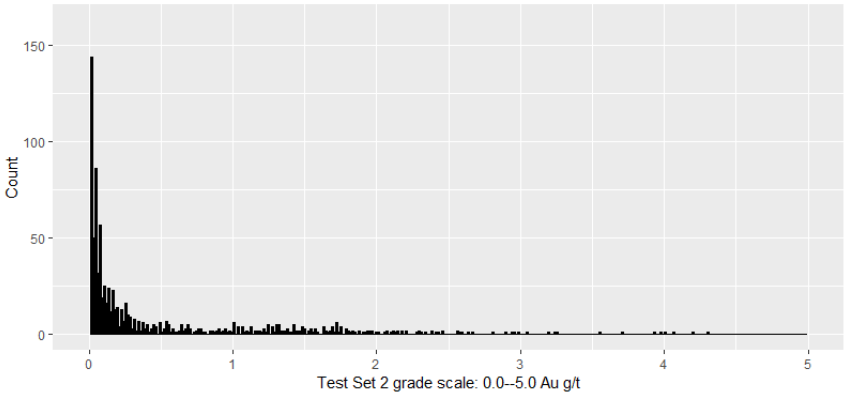

Figure 4.20. Histogram of TS2, scale 0.0–5.0 Au g/t

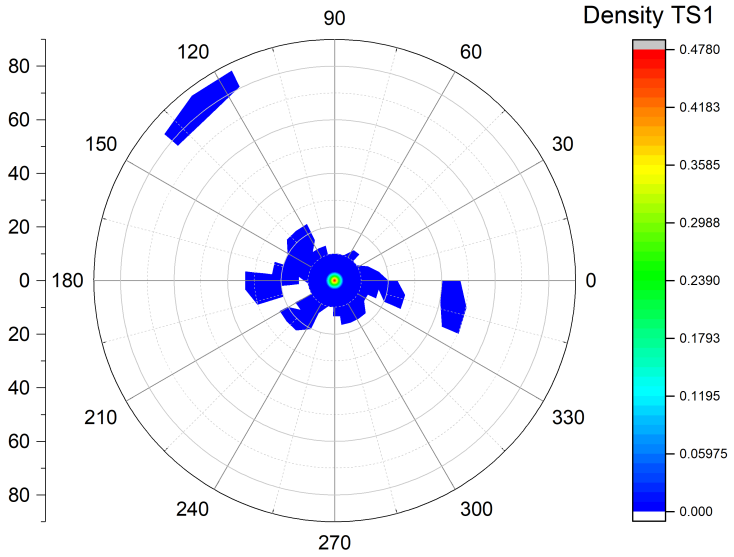

Figure 4.21. KDE of Test Set 1, n =1730. Negative sectors shown in white.

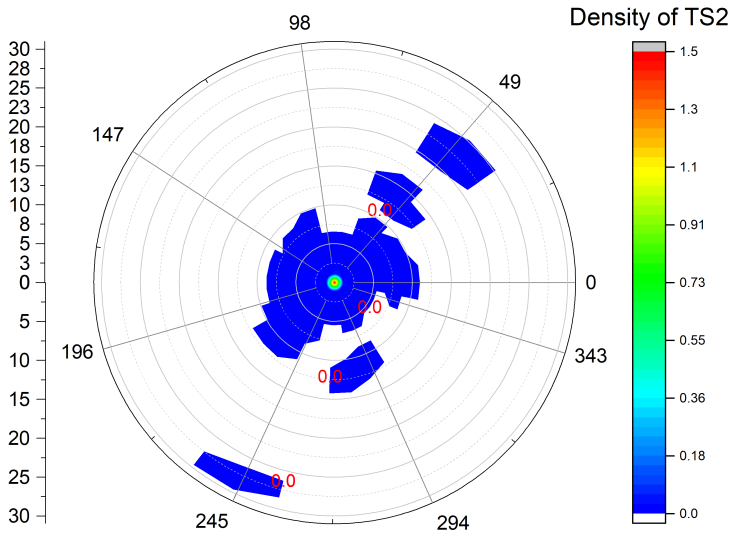

Figure 4.22. KDE of Test Set 2, n = 1066. Near-zero density sectors are labelled as 0.0

4.10 Comparison of Nonconvex Prototypes

In the present and following sections, the results of optimisation are explored from an optimisation viewpoint rather than statistical. In fact, the generated locally optimal solutions, which have not occurred naturally, should be considered as artificial objects that mimic the behaviour of the actual assay.

Listing 4.1. Equivalence test TS1 vs TS2

```
# Show statistics
          N        Mean      SD         SEM
"TS1"     1729     0.61157   1.94981    0.04689
"TS2"     1066     0.57205   1.54513    0.04732
************************************************************
# Compute equivalence analysis
# Output
     Difference  SE        Low Lim     Upp Lim  Low Bound CI  Upp Bound CI Conf Lev
     0.08727     0.08194   0.00        0.10      -0.04756      0.2221       0.95
     At 0.05 level, mean of "TS1" is not equivalent with mean of "TS2
************************************************************
#Compute t-test statistics, 0 < difference < 0.1
     Null Hypothesis     DF      t statistic  Prob
     Difference <= 0     2767    1.06495      0.1435
     Difference >= 0.1   2767    -0.15538     0.43827
Equal variances not assumed
Alternative hypothesis: 0 <Difference< 0.1.
At 0.05 level, the difference is not significantly within (0. 0,1)
```

In the UML, running the same algorithms several times yields different results per initialisation. The problem (3.6) is nonsmooth and nonconvex, therefore it also has a large number of locally optimal solutions, which are obtained by performing local search with one initialisation. That means after ten program restarts one can obtain different locally optimal solutions each time. In other words, one can not see the same result ten times.

Each DGM restart generates values for new samples, i.e. creates a new unique set of locally optimal solutions, which can not be repeated or pragmatically reproduced. Therefore, the consistency of statistical results across parameters after multiple program initialisations is not guaranteed.

However, in some cases the range, variance, standard deviation and other parameters of the generated structures may look similar statistically, but they always dissimilar enough in their shapes and geometries– millions distinct spatial combinations are possible. The only repeated feature is that each initialisation generates a set solutions of $\approx$33% of the actual dataset in size.

May appear, there is no general approach to describing local solutions by classical statistics. Formulating prediction in terms of convex optimisation is easier to solve, easier to analyse and easier to describe statistically. For example, despite the fact that in many cases global optimisation problems can be solved exactly, they still hard to be described by the classical statistical tools. It appears, the classical statistical theory does not stay behind some nonconvex nonsmooth optimisation methods.

R. Tibshirani (Tibshirani, 1996; James et al. 1996; Tibshirani et al., 2010) admitted that from a statistical viewpoint, nonconvexity does typically mean higher variance. A comparison of the variances provided in Table 4.1 as 80.678 for the actual set and 111.908, 104.555, 104,807 and 102.370 for P5, P7, P10 and 10A indicates that statement made by Tibshirani is not incorrect.

One come across a situation where the output models built from one testing set may have similar features, characteristics, shapes, etc. Sets P5 and P10A were build from Test Set 1, whereas P7 and P10 from Test Set 2. Hypothetically, pairs P5–P10A and P7–P10 may illustrate similar characteristics and behaviour. The visualised of prototypes shown in Figures 4.11–4.16 shows that the shape of each synthesised multi-object structure is much more complex and variable than in the actual dataset.

Considering that the generated prototypes are hard to describe statistically, so that

for pragmatic reasons, a number of naïve statistical approaches are used in this and the following sections to explore outputs. For example, statistical representation summarised in Table 4.1 provides some information about the mean, SD, variance and the character of the distributions.

The side-by-side boxplots shown in Figure 4.23 and Figure 4.24 display variation in samples and graphically compare the actual and predicted sets through their quartiles. From an optimisation viewpoint, apart from near-similar segments inside the rectangle in Figure 4.24 showing means and the positioning of the distribution line, the boxplots indicate that pairs P5–P10A and P7–P10 may be similar. But the feature that makes these pairs similar is unknown and has to be investigated. Therefore the following sections will look answers to questions

1. How can the obtained locally optimal solutions be described and compared?
2. How to detect the amount of difference and similarities between the models built from a one test set?
3. The objects built from one test set, but with a different number of hyperplanes (e.g. P7 and 10A) may (or may not) have similar features, characteristics and geometries. What is the effect of incrementally adding extra hyperplanes?

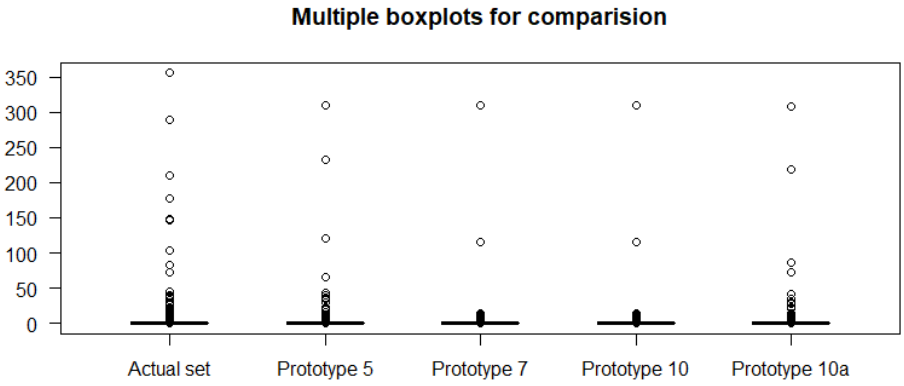

Figure 4.23. Figure showing five boxplots side-by-side for comparison. Scale 0.0–360.0 Au g/t

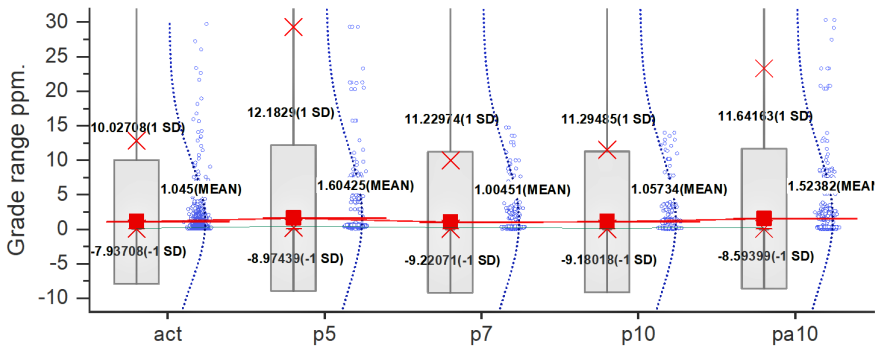

Figure 4.24. Side-by-side boxplots summarise the distribution of data stratified in five groups. The categorical factors: mean value curves–shown in red, the distribution lines–shown in blue, SD range min–max marked as red crosses, scale 0.0–32.00 Au g/t.

The application of proper statistical tools may answer the asked questions and result in detecting the difference as well as similarities between the nonconvex sets.

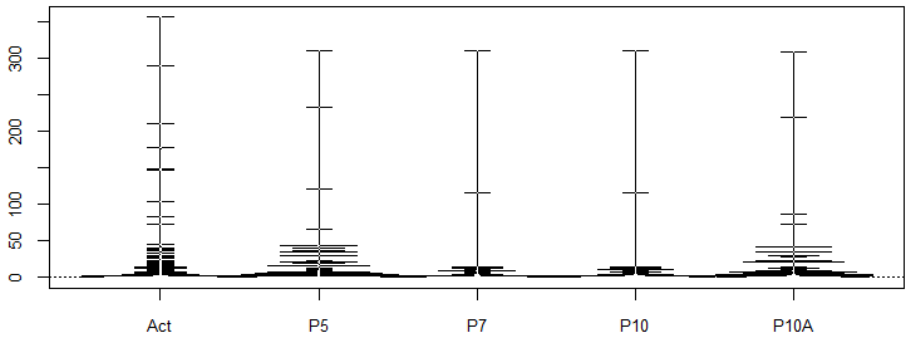

Figure 4.25. Comparison of multiple bean plots for nonconvex sets. Pair P5–P10A and pair P7–P10 appear as near-similar. Scale 0.0–360.0 Au g/t

4.10.1 P5 versus P10A (built from Test Set 1)

The structure of P5 was originally formed by (3.1) 5 hyperplanes. Set P10A has been a modification of P5, to which 5 extra hyperplanes were added incrementally (one-by-one). In this section, the comparison of results is graphically analysed ("before and after") to see if there has been an effect, i.e. an improvement in the objective function value.

The comparison of the distributions in the bean and multiple box plots (Figure 4.25 and Figure 4.23) indicates that prototypes P5 and P10A are semantically similar and display near-similar means: 1.604 vs. 1.524 (see Table 4.1). The highest predicted values (Table 4.2) have been recorded as 308.975 ppm (Au g/t) for P5 and 308.240 ppm for P10A. Results from AIC and f-tests in Listings 4.4 suggests that P5 and P10A are not statistically different. However, we still not yet have enough evidence to suggest that the compared sets are similar.

The CI in Figure 4.37 is not completely within the equivalence limits. One can not claim that the mean of the P5 is equivalent to mean of P10A.

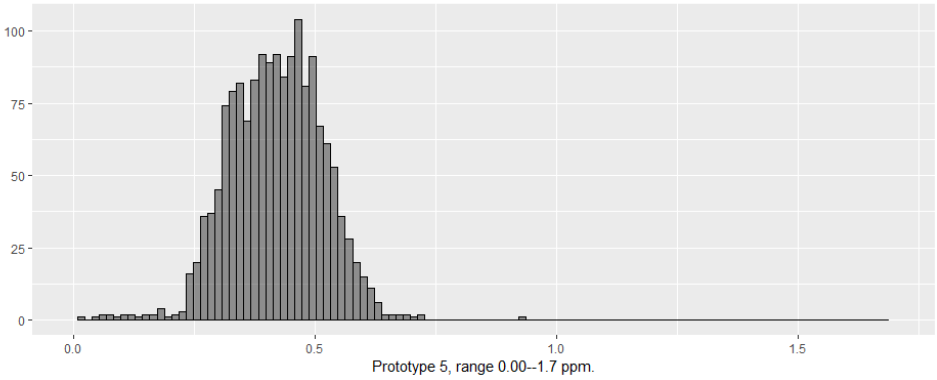

Figure 4.26. Histogram showing the distribution of grades in Prototype 5, n =1730, scale 0.00–1.8 Au g/t

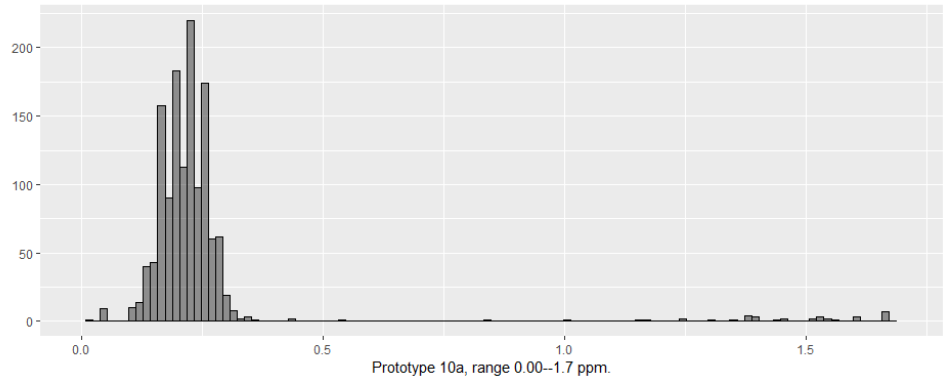

Figure 4.27. Histogram showing the distribution of grades in Prototype 10A, n =1739, scale 0.0–1.8 Au g/t

4.10.2 What is the effect from increasing the number of hyperplanes?

Which statistical feature makes the datasets similar? What could be the reason of the similarity between the sets obtained from a one test set?

The side-by-side boxplots on Figure 4.24 suggest that the answer may lie in the distribution of densities of means.

Then, an interesting possibility would be to visualise and explore the effect of adding extra hyperplanes to the existing ones, taking into account the distribution of densities of the means.

To do this and to visualise the difference between P5 and P10A, polar Kernel density estimator (KDE) (Figure 4.28 and Figure 4.29) was applied.

There are three known ways by which KDE can be generated and visualised: x, y Cartesian, Polar and Ternary. In our case, polar KDE is a useful tool for comparing spatial patterns of mean density distribution in each set.

Like the histogram, the KDE encodes the density of observations on one axis with height along the other axis. The output is a map showing the distribution of the concentrations of means in each set in the form of spatial patterns (geometrical comparison) and matrices (mathematical comparison, Appendix D). Unfortunately, the KDE theory has not made it into commonly available statistical software.

This approach allows comparison of visual effects that occurs when new hyperplanes are added to the excising number of hyperplanes. More information on the KDE, weightings and scaling can be found in studies by Uria-Tellaetxe & Carslaw (2014) and Grange et al. (2016).

For graphical analysis and model comparison a type of KDE, termed as polar estimator is used. The peculiarity of the KDE is that the mean concentrations are plotted as patterns. Bright contours around the main clustered area on the KDE shows where to highest densities are located. This type of visualisation may substantially improve understanding of the similarity/ dissimilarity between the compared sets. The matrix of the distribution of densities on the means is presented in Appendix D.

For convenience, zero value patterns and labels have been removed from both radial and angular scaling. Please note that the method does not allow the data to be weighted exactly. The principal aim of involving the Kernel density estimator is as a graphical analysis and comparison of the distribution of density patterns in data rather than for quantitative purpose.

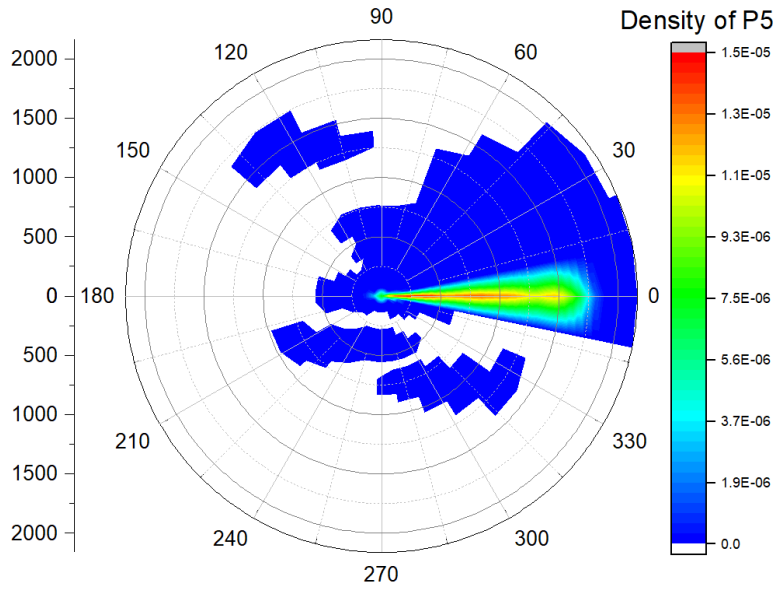

Figure 4.28. P5, n =1730, 5 hyperplanes. KDE shows the distribution of spatial patterns of the probable concentrations of the densities of means.The upper limit for data is set to 310. Red shading is statistically significant gradients indicate higher density values. Radial scale = +/- 2162. The greatest densities located between 0.0 and 1.5. Spatial patterns are semantically similar to P10A.

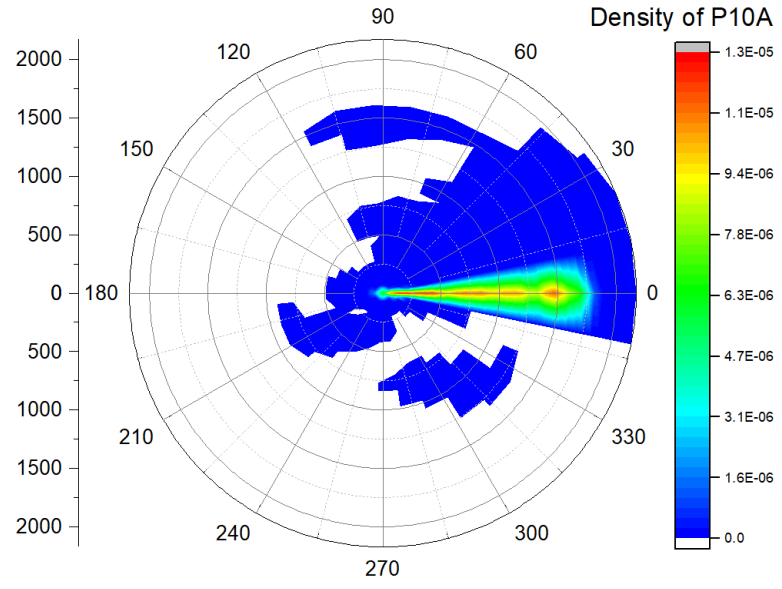

Figure 4.29. P10A, n = 1739. KDE shows the effect from adding 5 extra hyperplanes to existing 5 (cf. P5 and P10A). Radial scale = +/- 2173. The greatest densities located between 0.0 and 1.5. Spatial patterns (geometrical comparison) are semantically similar to P5.

The Kernel density plot of P5 shown in Figure 4.28 highlights the highest isolated concentrations of low grades between 0.0 and 1.8 ppm, which occur in one direction, particularly from the centre to the easterly direction 0°, reaching ≈ 1600. This single, continuous and dominant pattern in the eastern direction has high probability. The plot shows areas of positive, low and higher grade concentrations to the north–west. Comparing Figure 4.28 and Figure 4.29 one can notice that the shapes of patterns, in overall, changed insignificantly. There are some minor changes in the densities of patterns. Each sample of observed data is smeared out into a density function that represents other samples that might have also happened.

On Figures 4.28 and 4.29 some interesting phenomena can be seen - the distribution of densities in the compared plots has changed significantly. Why does this happen?

The boxes shown in Figure 4.24 indicate that there are some changes in the SD range and the shapes of distributions for P5 and P10A. Bins of P5 in Figure 4.26 are concentrated mostly between grades 0.25 and 0.7 Au g/t. Adding 5 extra hyperplanes resulted in the re-distribution of bins towards higher grades as shown in Figure 4.27. For example, compared to Figure 4.28. one may observe (Figure 4.29) some reduction of bins in range 0.25–0.4 and the appearance of new bins between 1.0–1.7 Au g/t.

Despite P5 and P10A describe the same matter and built from one test set (TS1), they are not equivalent. Histograms in Figures 4.26 and 4.27 provide a quick visualisation of the difference in distributions. Data in the both histograms are right-skewed and have more than one small localised concentrations.

One can see that bin width in both P5 and P10A are too small, show too many individual grades. However, the bins generally allow the underlying frequency distribution of the data to be seen. The equivalence test presented in Listings 4.3 indicates the presence of the difference between means. The t-test (Listings 4.3) statistics suggests that at 0.05 level the difference is not significant within (0, 0.1). Figure 4.37 shows 95% Confidence Intervals and the the difference range between the means. If the computation repeated, these parameters would not significantly change.

However, sets P5 and P10A do not demonstrate similarity in terms of bin size, positioning and the distribution over the full range of values.

Incrementally adding 1,2...5 and more hyperplanes to existing 5 gives a new dataset with re-generated of SD, variance, etc. This an attempt to re-shape the distribution and spatial positioning of the instances. However, there is little evidence available that this type of data manipulation significantly improves the accuracy of prediction.

Benefit from increasing the number of hyperplanes is the possibility to obtain a significantly lower RMSE, i.e. to obtain more refined model. In our case, 5 hyperplanes are enough to form a basis for creating millions of spatial combinations, which will be near-similar in terms of the distributions of the means.

Figure 4.30 and Listings 4.2 compare P5 and P10A, generated using the same testing set TS1. The plot shows a random pattern, 1001 rows and there is a correlation between the results. Figure 4.31 compares full range of data, the distances are zoomed from the regression line. 1740 rows involved in comparison with `sd=1.8`.

Listing 4.2. Comparison of P5 and P10A

```
# compute model from 1001 rows, reduce over-plotting
set.seed(1001)
  dataset <- data.frame(x=rnorm(1001),y=rnorm(1001),
    z=rnorm(1001))
      dataset$p5 <- with(dataset,
        rnorm(1001,mean=x+2*y+z,sd=0.7))
         m <- lm(p5~x+y+z,dataset)
       plot(predict(m),dataset$p5,
       xlab="p5",ylab="p10a", pch = 1, font.lab=1, font.sub=1, cex.lab=1)
      minor.tick(nx=10, ny=5)
        axis(side = 1, at = c(-9,-7,-5,-3,-1,0,1,3,5,7,9))
        axis(side = 2, at = c(-7,-5,-2,0,2,7))
    abline(a=0,b=1)
```

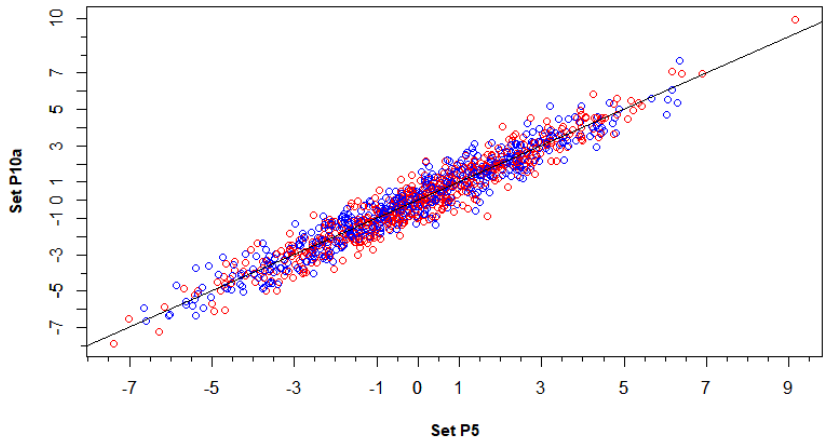

Figure 4.30. Versus Fit comparing P5 (blue) with p10A (red). Sets produced from TS1, options `set.seed(1001)` and `sd=0.7` are applied.

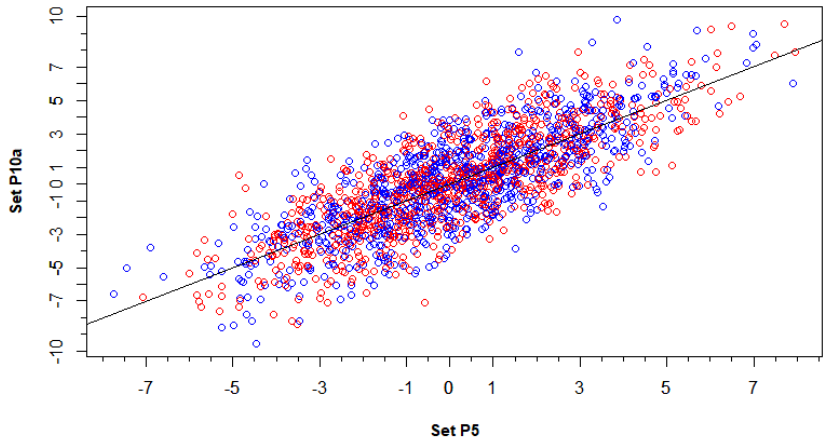

Figure 4.31. Sets P5 (blue) with p10A (red) with options `set.seed(1740)`. The distances are zoomed on from the regression line with `sd=1.8`.

Listing 4.3. Equivalence t-test P5 : P10A (built from TS1)

```
# Compute equivalence for two samples P5 : P10A
> ttest = t.test(p5~p10a)
> names(ttest)
[1] set rule "Difference = mean1 - mean2"
[2] "Alt hypothesis: lower limit < Difference < Upper limit"
# Compute statistics
# Output
         N        Mean      SD          SEM
"p5"     1730     1.60425   10.57864    0.25434
"pa10"   1739     1.52382   10.11781    0.24263
# Compute equivalence
# Output
Differ    SE      Low Lim Upp Lim Low CI     Upper     CI
0.08043   0.3515  0.00    0.10    -0.49789   0.65876   0.95
Equal variances not assumed
At 0.05 level mean p5 in not equivalent with mean p10A
# Compute t-test statistics
# Output
Null Hypothesis      DF      t_statistic   Prob
Difference <= 0      3458    0.22883       0.40951
Difference >= 0.1    3458    -0.05567      0.4778
Equal variances not assumed
Alt hypothesis 0 < Difference < 0.1
At 0.05 level the Difference is not significantly within (0, 0.1)
```

Listing 4.4. Comparison P5 and P10A (both from TS1) with Akaike and f-test

```
# Create fit parameters for P5 and P10A
# Output
                          p5            pa10
Slope        Value        47.09509      44.15915
Slope        Std_Error    7.79104       7.38808
Intercept    Value        0             0
Intercept    Std_Error    0             0
# Create fit statistics , Akaike and f-test
# Output
                          p5            pa10
Number of Points          1730          1739
Degrees of Freedom        1729          1738
Residual Sum_of_Sqr       193844.2479   178292.2304
Pearson's r               0             0
Adj. R-Square             -0.00184      -0.0021
*********************************************
# Compute Akaike test, where Akaike weight is value
# divided by the sum of these values across two models.
# Output
             AIC           Akaike Weight
Same         16223.01766   0.72364
Different    16224.94281   0.27636
According to Akaike weights, datasets are same
*********************************************
# Compute f_test
     F         Numer.DF    Denom.DF    Prob > F
     0.07481   1           3467        0.78448
Datasets are not statistically different
```

Listing 4.5. Equivalence test P7 : P10 (built from TS2)

```
# Compute equivalence for two samples P5 : P10A
> ttest = t.test(p7~p10)
> names(ttest)
[1] "Difference = mean1 - mean2"
[2] "Alt hypothesis: lower limit < Difference < Upper limit"
# Compute statistics
# Output
          N        Mean      SD          SEM
"p7"      1066     1.00451   10.22523    0.31318
"p10"     1066     1.05734   10.23751    0.31356
# Compute equivalence
# Output
Differ    SE       Low Lim   Upp Lim    Low CI     Upper     CI
-0.05282  0.44317  0         0.1        -0.78209   0.67645   0.95
Equal variances not assumed
At 0.05 level mean p7 in not equivalent with mean p10
# Compute t-test statistics
# Output
Null Hypothesis      DF      t_statistic   Prob
Difference <= 0      3458    -0.11919      0.54743
Difference >= 0.1    3458    -0.34484      0.36513
Equal variances not assumed
Alt hypothesis 0 < Difference < 0.1
At 0.05 level the Difference is not significantly within (0,0.1)
```

4.10.3 P7 versus P10 (from Test Set 2)

Next, sets P7 and P10 are compared. The structure of P7 was originally created by 7 hyperplanes, which were appointed by CLR automatically, as the most optimal variant. P10 is a modification, to which three more hyperplanes were added one-by-one. The distribution in P7 is described in Tables 4.1–4.2, Figures 4.32– 4.35 and Listings 4.5–4.6. Data in th histogram (Figure 4.32) shows that the major concentration occurs around value 0.25.

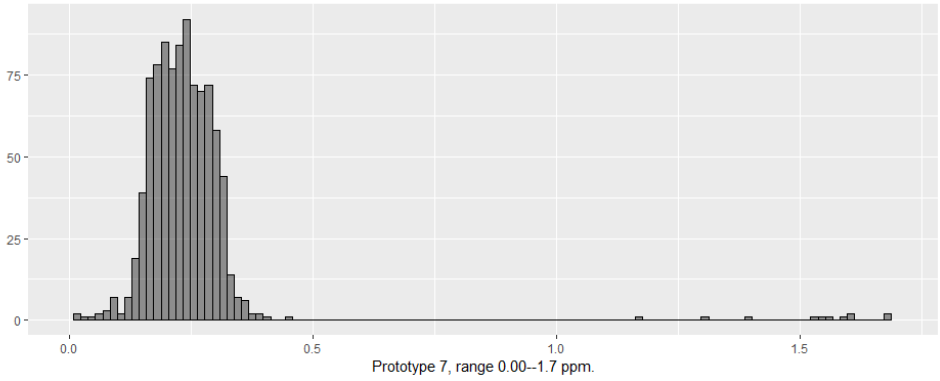

Figure 4.32. Histogram showing the distribution of grades in Prototype 7, n =1066, scale 0.0–1.8 Au g/t

The KDE density patterns in Figure 4.35 and Figure 4.36 are near similar. The highest isolated concentrations of low grades lie between 0.0 and 1.8 Au g/t scale, which occur from the centre to the easterly direction 0°, reaching ≈900. This continuous and dominant pattern in the eastern direction has the highest probability. The KDE also shows areas of low concentrations to the north–west. There some minor differences to northerly direction 90°–150° between 600 and 900. The matrix of the distribution of densities on the means is presented in Appendix D.

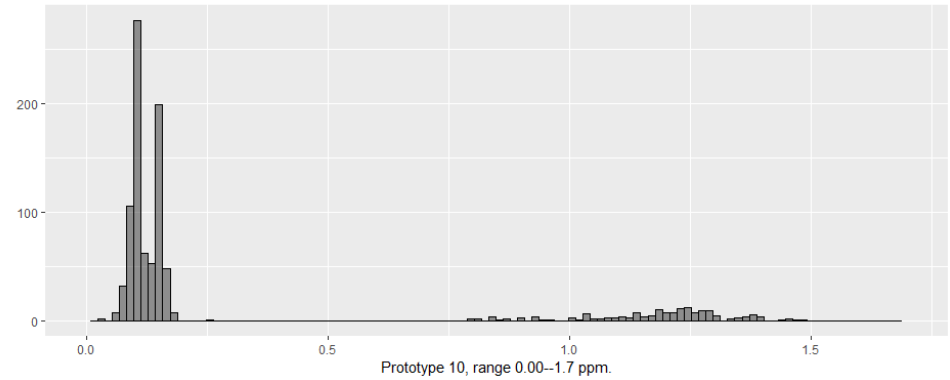

Figure 4.33. Histogram showing the distribution of grades in Prototype 10, n =1066, scale 0.0–1.8 Au g/t

The equivalence test in Listings 4.5 indicates the existence of a difference between means. However, t-test statistics suggest that at 0.05 level the difference is not significant within (0, 0.1). Figure 4.38 shows the difference range between the means of P7 and P10. It is not entirely correct to say that there is a 95% chance that the calculated confidence interval contains the true mean. Pragmatically, the CI in Figure 4.38 is not within the equivalence limits. One can not claim that the mean of the P5 is equivalent to the mean of P10A.

From KDE in Figure 4.36 along, there is no obvious indication that adding 3× extra hyperplanes significantly changed the existing mean density patterns. By contrast, the histogram in Figure 4.33 clearly indicates the re-distribution of bins towards higher grades ranging from 1.0–1.5 Au g/t.

Figure 4.34 compares full range of data, the distances are zoomed from the regression line. 1080 rows involved into comparison with option `sd=1.8`.

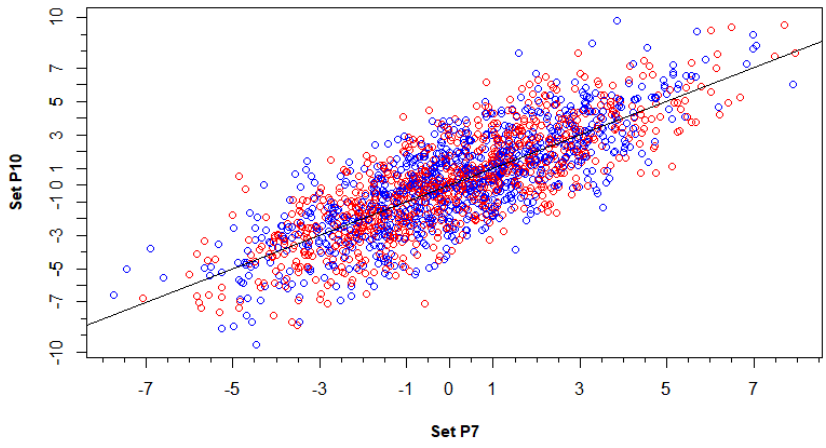

Figure 4.34. Sets P7 (blue) with p10 (red) with options `set.seed(1080)`. The distances are zoomed on from the regression line with `sd=1.8`.

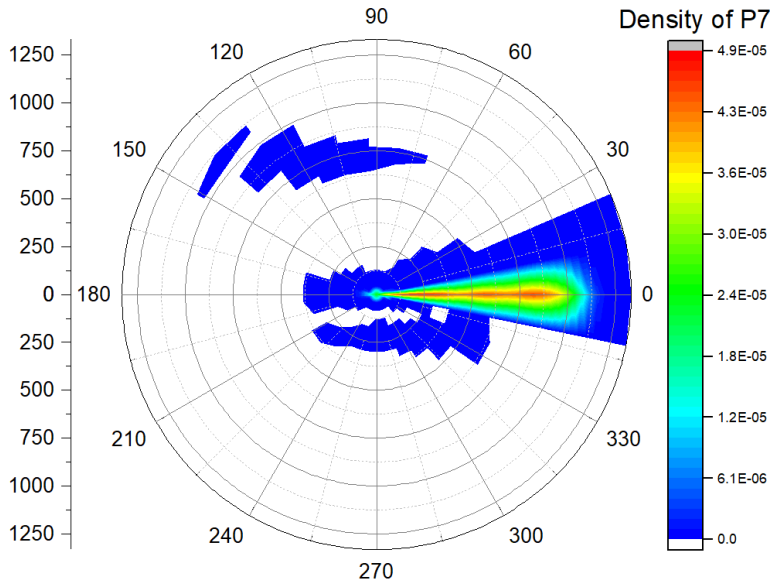

Figure 4.35. P7, n =1066, 7 hyperplanes. KDE shows the distribution of spatial patterns of the probable concentrations of the densities of means. Radial scale = +/- 1335. The greatest densities located between 0.0 and 1.2. Spatial patterns (geometrical comparison) are semantically similar to P10.

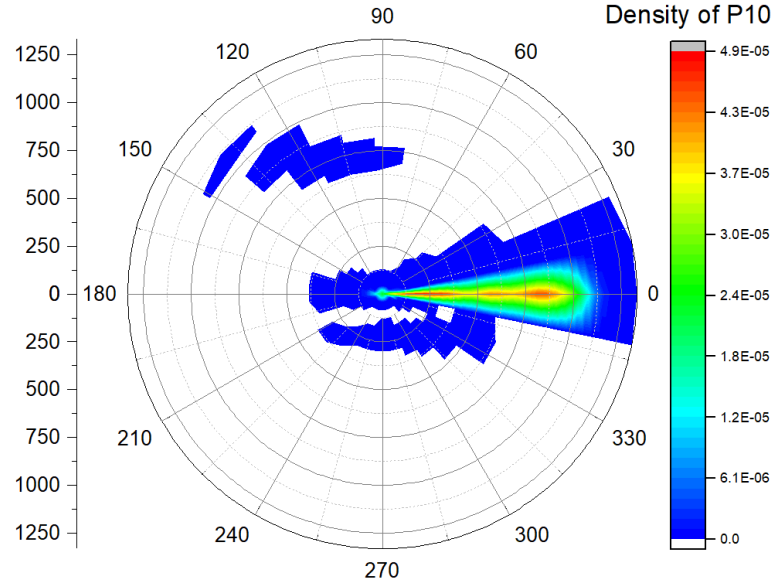

Figure 4.36. P10,n =1066. KDE shows the effect of adding extra 3 hyperplanes to existing 7 (cf. P7 and P10A). Radial scale = +/- 1332. Spatial patterns (geometrical comparison) are semantically similar to P7.are semantically similar to P7.

Listing 4.6. Comparison P7 and P10 with Akaike and f-tests (from TS2)

```
# Create fit parameters for P7 and P10
# Output
                          p7            p10
Slope       Value         41.55894      44.10214
Slope       Std Error     15.96999      15.99101
Intercept   Value         0             0
Intercept   Std Error     0             0
# Compute fit statistics
# Output
                        p7              p10
Number of Points        1066            1066
Degrees of Freedom      1065            1065
Residual Sum of Sqr     111716.63861    112010.91006
Pearson's r             0               0
Adj. R-Square           -0.00328        -0.00351
************************************************
# Compute Akaike test
AIC           Akaike        Weight
Same          9924.99433    0.72981
Different     9926.98165    0.27019
According to Akaike weights, datasets are same
************************************************
# Compute f-test
     F            Numer.DF     Denom.DF     Prob>F
     0.01266 1    2130         0.91041
Datasets are not statistically different
```

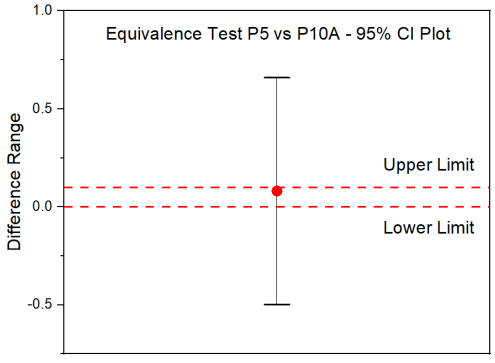

Figure 4.37. CI for P5 : P10A

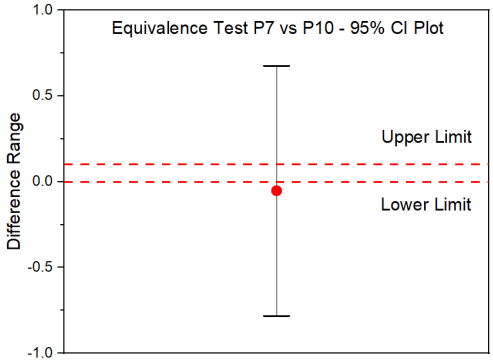

Figure 4.38. CI for P7 : P10

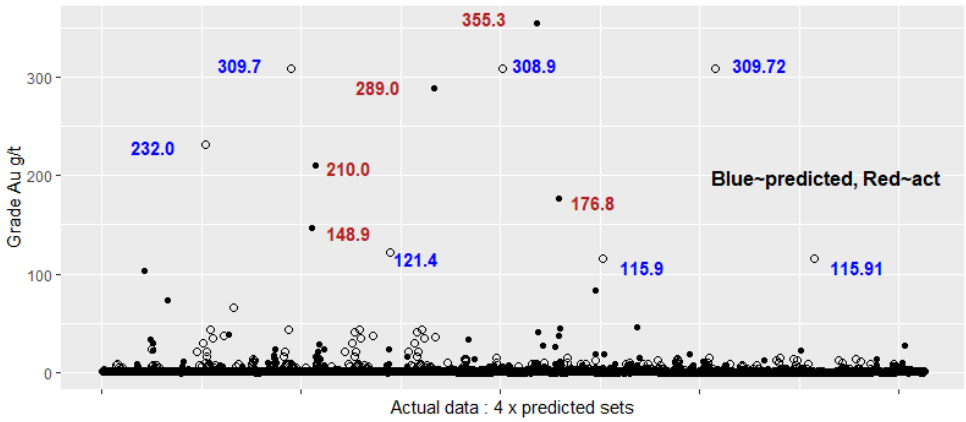

Figure 4.39. Plot compares 4× nonconvex sets combined with actual data in a one single set. High values marked in blue (actual) and red (predicted).

The goal of traditional interpolation is to create a surface that is intended to best represent empirical reality. In NSO, the data are not continuous over space. The absence of assumptions and limited number of predicted data do not allow the obtained sets to built a surface or solid.

4.11 Nonconvex Prototypes versus Test Sets

In this section, essentially the same procedure was applied to each nonconvex set to explore interesting properties of the generated models, such as how well these models have been trained and the amount of difference between the testing data and locally optimal solutions.

4.11.1 P5 versus TS1

Figures 4.43 and 4.44 show how values in P5 and Test Set 1 are related. In Figure 4.43 three horizontal areas with adjusted by `geom.point(abs)` absolute size of residuals appear

1. the upper area > 50 Au g/t is represented by four high-grade samples as high as 308, 232, 121 and 65 Au g/t marked in red.
2. the middle section is represented by a number of samples ranging 5.0–50.0 Au g/t.
3. the low area shows that the most frequent samples in P5 and TS1 occur in range 0.0–5.0 Au g/t.

As seen from Figure 4.43, the samples in the lover area have dominant sequence and the shapes of the distributions roughly form horizontal bands around the zero line. The data in Table 4.1 and Table 4.3 indicate the variance for P5 = 111.9 and for TS1 = 7.74. Table 4.3 highlights the highest values in TS1 as 83.2; 45,7; 27.2; 22.6 and 18.9 Au g/t. Therefore, the lower and middle ranges are interesting to compare for the periodicity of the samples and shapes of the distribution.

The used amount of difference between P5 and TS1 are residuals (Listings 4.7). `ggplot2` shifts the two distributions along the x-axis, preserving the shapes and relative distances in P5 and TS1. As seen from data in Figure 4.44, the models P5 and TS1 are not significantly different: P5 is following the shape of the distribution of TS1.

Listing 4.7. R–comparison P5 and TS1

```
# Compare  P5 and TS 1
# Compute residuals for P5
# Output
    Min       1Q     Median       3Q      Max              ***
 -1.632   -1.261   -1.178   -1.098   307.360              ***
# Compute coefficients for P5
# Output
              Estimate    Std.Err    t-value   Pr(>|t|)
(Intercept)    1.62338    0.26145    6.209     6.66e-10 ***
dataset TS1   -0.02900    0.09146   -0.317     0.75101  ***
*************************************************************
Residual standard error: 10.58 on 1728 degrees of freedom
Multiple R-squared: 5.822e-05, Adj. R-squared: -0.0005205
F-statistic: 0.1006 on 1 and 1728 DF, p-value: 0.7511
```

According to the data in Table 4.1, the highest values in P5 (308.87; 232.02; 121,4, etc.,) are similar to the values in the actual set (cf. Table 4.1 and Table 4.2). The question then arise *why the high values are similar?*. There is a valid and sufficient reason that explains such effect: the task of the UML system has been to recognise and mimic the behaviour of the distribution recorded in the actual assay via training and test sets.

The Incremental Algorithm (3.1) recreates all available approximations and finally, P5 (and other sets) becomes a replica of the actual drill-core assay. The blue points in P5 (Figure 4.44) follow the contour of TS1 showing the effect of "model mimicry", i.e. the ability of a model to account for the data synthesised by another data (TS1). The plot shows how much the prediction deviated from the TS1 values and how the predicted values are close to the TS1 values.

4.11.1.1 Comparison P5 versus TS1 with function `mfrow=c(2,2)`

Multiple plots shown in Figure 4.40 present a method to interpret residual terms, and determine whether there might be problems with the model.

The Residuals versus Leverage plots identifies the influential data points in the model. As is seen from plot, the influential values (52, 27, 59) located outside the red dashed Cook's distance line in the lower part. Statistics suggests that removing them would likely noticeably alter the regression results.

The Scale-Location plot shows whether the residuals are spread equally along the predictor range, i.e. homoscedastic. From a statistical point of view, the line on this plot has be horizontal with randomly spread points. The red line in the plot starts off near-horizontal at the beginning of the range, insignificantly slopes down, and then flattens again from 0 to 6. The line goes down because the residuals for P5 values are more close. Sector 0–6 data generally has uniform variance.

The normal Q-Q plot determines if the variables is normally distributed by plotting quantiles from TS1 against a P5 distribution. As is seen, the data is normally distributed and plotted in a generally straight line.

Lastly, the Residuals versus Fitted plot tests the assumptions of whether the relationship between the variables is linear, and weather there is equal variance along the regression line. Ideally, this plot should be relatively shapeless without clear patterns in the data, with no obvious outliers, and be generally symmetrically distributed around the zero line without particularly large residuals. The plot is relatively shapeless. However, there may be a number of potential outliers.

Listing 4.8. R–statistical comparison P5 versus TS1

```
# Compute two distributions
set.seed(1001)
     dataset <- data.frame(x=rnorm(300),y=rnorm(300),
         z=rnorm(300))
       dataset$p5 <- with(dataset,
             rnorm(300,mean=x+2*y+z,sd=0.3))
       m <- lm(p5~x+y+z,dataset)
          plot(predict(m),dataset$p5,
              xlab="Predicted",ylab="TS1")
                 abline(a=0,b=1)
                   theme_bw()
                par(mfrow = c(2, 2))
             plot(m)
```

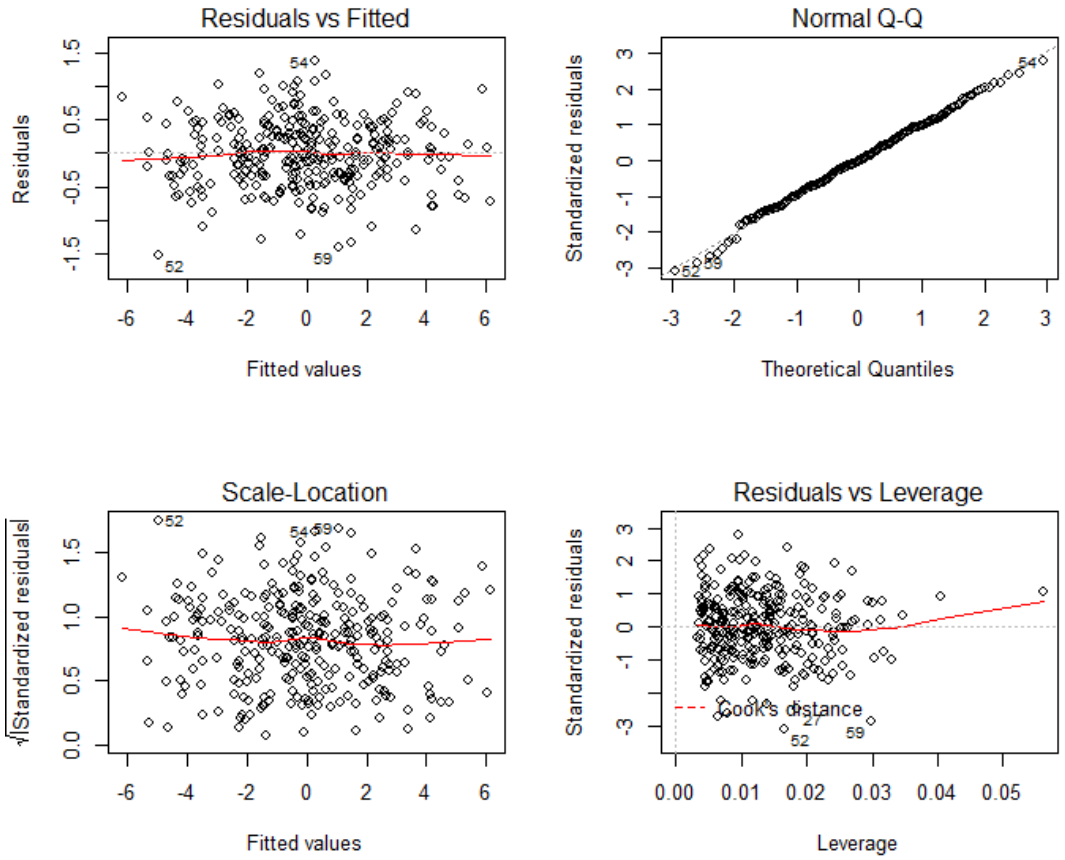

Figure 4.40. Comparison of P5 and TS1 with option `mfrow=c(2,2)`

4.11.2 P7 versus TS2

Figure 4.45 compares the distributions in TS2 and P7 in a range from 0.0 to 340.0 Au g/t, where the highest residuals marked as red circles. In the structure with adjusted absolute residuals (`geom.point(abs)`), four areas appear:

1. An upper area >100 Au g/t is presented by just two high-grade isolated samples: 309.7 and 115.9 Au g/t.

Listing 4.9. R–comparison P7 and TS2

```
# Compare  P7 and TS2
# Compute residual for P7
# Output
     Min      1Q      Median   3Q       Max              ***
  -1.084   -0.815   -0.768   -0.716   308.719            ***
# Compute coefficients for P7
# Output
              Estimate   Std.Err   t-value   Pr(>|t|)
(Intercept)   0.99707    0.33413   2.9841    0.00291   ***
dataset ts2   0.01301    0.20288   0.0644    0.94887   ***
Significant codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
Residual std error: 10.23 on 1064 degrees of freedom
Multiple R-squared: 3.867e-06,  Adjusted R-squared: -0.000936
F-statistic: 0.004115 on 1 and 1064 DF,  p-value:  0.9489
```

2. The upper limit of a middle section 5.0–15 Au g/t is represented by three highest samples: 14.6; 14.6 and 13.5 Au g/t,
3. A lower structure ranges from 0.0 to 5.0 Au g/t.

Figure 4.45 shows that the lower structure, in particular, the range 0.0–2.0, has the dominant sequence and limited periodicity of the samples (Figure 4.46). Graphical comparison of the parametrics in Listings 4.9 indicates that P7 attempts to follow the contour of TS2. One can see that the distributions are semantically similar, describe similar matter and represent similar concept.

However, a comparison of histograms in Figure 4.32 and Figure 4.20 indicates that the distributions may be different because bins and the number of values fell into each interval are completely different.

4.11.2.1 Statistical comparison P7 and TS2 with `mfrow=c(2,2)`

Statistical comparison of P7 and TS2 is presented by (Listings 4.10) is shown on Figure 4.41.

The Residuals vs. Leverage plots identified several influential values located outside (312, 327) the red dashed Cook's distance line. The Scale-Location plot shows are spread near-equally along the predictor range. The red line is flat and horizontal. This model can be viewed as homoscedastic. The data is sector (-5)–(+5) generally has uniform variance.

The Q-Q plot determines if the variables are normally distributed by plotting quantiles from P7 against a distribution TS2. As is seen, the data is normally distributed and plotted in a straight line.

The residuals vs. fitted plot on Figure 4.41 tests the assumptions of whether the relationship between the variables is linear and the whether there is equal variance along the regression line.

It can be seen that the plot has a shape with clear patterns in the data, with the presence of outliers. However, the model shows symmetrically distributed values around the zero line. Large residuals (66, 118 and 778) are present. The plot has shape and suggests that there may be a number of potential outliers.

Listing 4.10. Comparison for P7 and TS2 with mflow

```
# Compare sets, apply mflow
set.seed(999)
       dataset <- data.frame(x=rnorm(999),y=rnorm(999),
                             z=rnorm(999))
       dataset$p7 <- with(dataset,
                          rnorm(999,mean=x+2*y+z,sd=0.5))
       m <- lm(p7~x+y+z,dataset)
       plot(predict(m),dataset$p7,
            xlab="Predicted",ylab="TS2", pch = 1)
       abline(a=0,b=1)
       theme_bw()
       par(mfrow = c(2, 2))
       plot(m)
```

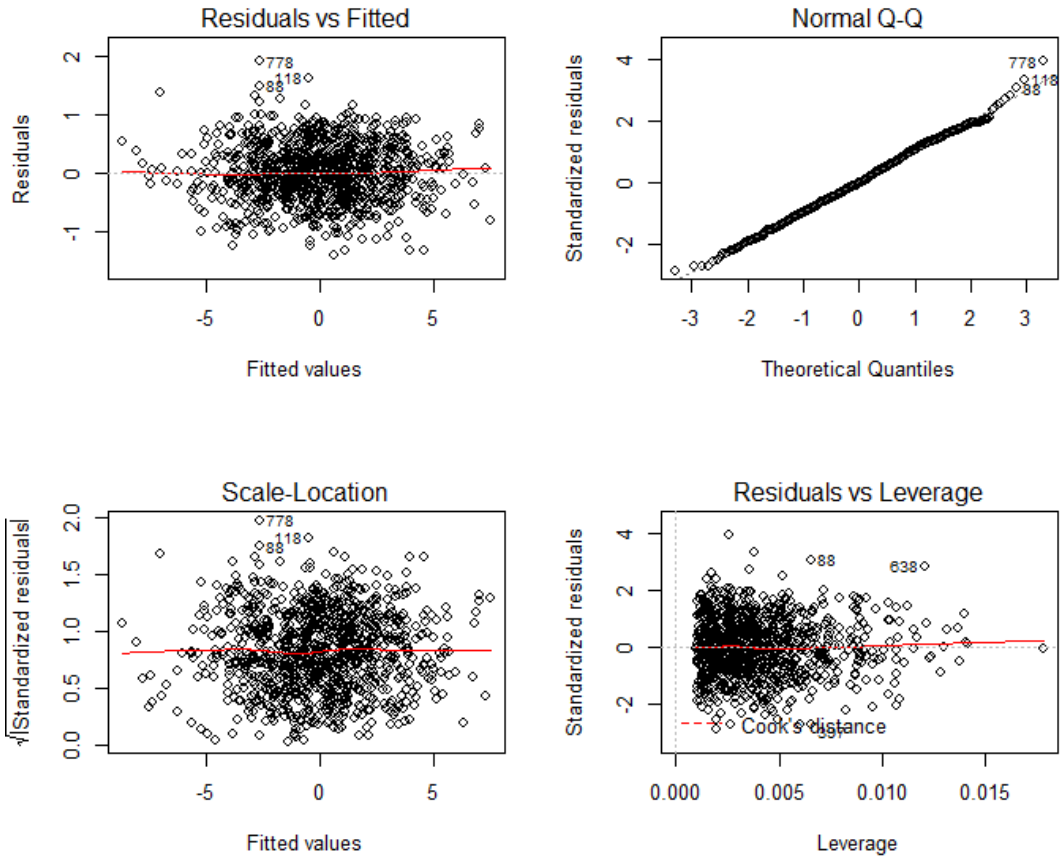

Figure 4.41. Comparison of P7 and TS2 with function `mfrow=c(2,2)`

4.11.3 P10 versus TS2

Figure 4.47 and Figure 4.48 show how values in P10 and TS2 are related. In Figure 4.48 three distinguishable areas appear

1. An upper area >100.0 is presented by just two high-grade isolated samples: 309.702 and 115.914 Au g/t.
2. A middle section 10.0–15.0 Au g/t is represented by three highest samples: a cluster of samples 13.4–13.9 Au g/t.

Listing 4.11. R–comparison parameters for P10 and TS2

```
# Compare  P10 and TS2
# Compare residuals for P10
# Output
  Min       1Q       Median     3Q       Max
 -1.059    -0.954    -0.921     -0.892    308.644
# Compute coefficients for P10
# Output
                Estimate   Std. Err   t value  Pr(>|t|)
(Intercept)    1.059074    0.334532    3.166    0.00159  ***
dataset ts2   -0.003038    0.203123   -0.015    0.98807  ***
Significant codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
#
Residual standard error: 10.24 on 1064 degrees of freedom
Multiple r-squared: 2.103e-07, Adjusted R-squared: -0.0009396
F-statistic: 0.0002238 on 1 and 1064 DF, p-value: 0.9881
```

Listing 4.12. R–comparison parameters for P10A and TS1

```
# Compare  P10A and TS1
# Compute residuals for P10A:
     Min       1Q     Median      3Q       Max
  -1.542    -1.353   -1.310    -1.250   306.698
# Compute coefficients for P10A:
                  Estimate   Std.Err    t-value    Pr(>|t|)
(Intercept)      1.54210     0.25070    6.1511    9.54e-10  ***
dataset$ts1     -0.01619     0.08770   -0.1854    0.854201
#
Significant codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
Residual standard error: 10.15 on 1728 degrees of freedom
Multiple R-squared: 1.973e-05,  Adjusted R-squared: -0.000559
F-statistic: 0.0341 on 1 and 1728 DF, p-value: 0.8535
```

3. A lower area, which shows that the most frequent samples in the sets tend to occur in range 0.0–5.0 Au g/t. The distributions in P10 and TS2 lie on one line and the amount by which P10 differs from TS2 is not significant.

P10 and P7 built from one test set - TS2. In terms of the presence of similar statistical extreme values (cf. Table 4.2), P10 is near similar to P7. A side-by-side comparison of P7 and P10 (c.f. Figures 4.46–4.48 indicates that the models are graphically comparable, they are of the same type, generated on the basis of the same representation concepts, describe similar matters and show relatedness to the TS2 and the actual data.

As appear, the parameters in Listings 4.9 and Listings 4.11 are not significantly different. It can be seen from Listings 4.11 that P10 SE=0.203 ≈ P7 SD=0.202 and parameters Pr>|t| P10 = 0.98 is not significantly different from Pr>|t|=0.94 in P7.

4.11.4 P10A versus TS1

P10A is constructed from testing set TS1. Figures 4.47–4.48 show how values in P10A and TS1 are also related. In Figure 4.48 two areas appear

1 An upper area >100.0 is presented by just two high-grade isolated samples: 308.240 and 218.59 Au g/t (see Table 4.2).

2 A middle section 40.0–90.0 Au g/t is represented by three highest values: 86.46; 72.19 and 42.06 Au g/t and

3 a low area, in which the most frequent samples in both sets occur in range 0.0–2.0 Au g/t. This range lie on one line.

Listing 4.13. R–statistical comparison for P10A and TS1

```
# Compare distributions
set.seed(1000)
  dataset <- data.frame(x=rnorm(100),y=rnorm(100),
                        z=rnorm(100))
      dataset$p10a <- with(dataset,
                      rnorm(100,mean=x+2*y+z,sd=0.1))
  m <- lm(p10a~x+y+z,dataset)
      plot(predict(m),dataset$p10a,
           xlab="predicted",ylab="TS1")
            abline(a=0,b=1)
             theme_bw()
              par(mfrow = c(2, 2))
               plot(m)
```

In terms of the presence of similar statistical extreme values (cf. Table 4.2), P10A is near similar to P5. A side-by-side comparison of P10A and P5 (Figures 4.43–4.49) suggests that the models graphically comparable, they are of the same type, conceptually comparable, describe similar matters and show relatedness to TS1.

In contrast to expectations, not much evidence available to suggest that adding 5 extra hyperplanes have significantly changed the shape of the distribution or enhanced P10A (cf. Figure 4.48 and Figure 4.44.

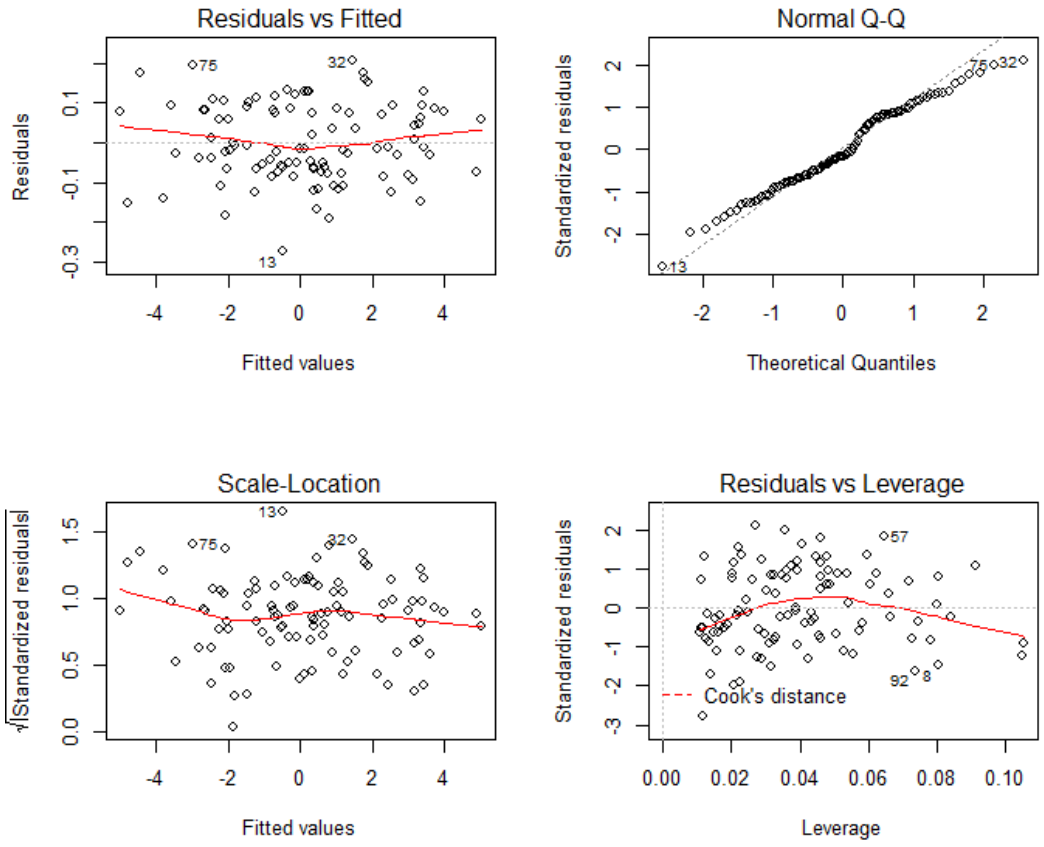

Figure 4.42. Stacked statistical plots comparing P10A and TS1.

4.12 Preliminary results

Essentially the same procedures have been applied to the all datasets.

It was found that that the distributions in P5 and P10A are graphically comparable, semantically similar, they are of the same type, conceptually comparable, describe similar matter and show relatedness to TS1 and the actual dataset.

As to P7 and P10, they are also graphically comparable, semantically similar and they are of the same type. The property that makes the compared sets similar is the distribution of the densities of means.

However, the compared sets are different in terms of the location of the generated structures, their orientation and shape. Another feature that unites the output is tha presence of co-called model mimicry. The concept of model mimicry first proposed in 1878 by Fritz Mueller in Mueller (1878). More recent interpretation made by Wagenmakers et al. (2004) states:

If model A (assume prototype) is able to closely fit the data B (assume test set) generated by model C (assume training set) formed by model D (assume actual data), model A is able to behave like model D. As appears, the UML can mimic and reproduce the properties of other objects (actual dataset).

Plot in Figure 4.44 P5 against the fitted values TS1 does not provide enough evidence of the existence of heteroscedasticity in the errors, i.e. when the variance of the residuals may not be constant.

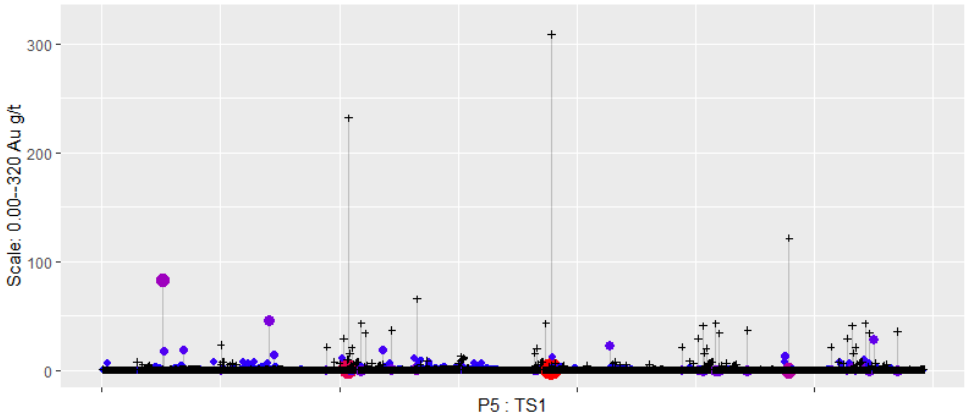

Figure 4.43. P5 versus TS1, scale: 0.0–320.0 Au g/t. Higher `abs` values marked in red.

Next, the range of Au grade for P10A and set limits y is changed again with f=`<ylim>` and the gradient of TS1 distribution is set "lightgrey". It may appear that the horizontal line (zoomed grey) work well for all parts of the data, for other sections of the fitted values. In this example, the variances for the first quarter of the data, are similar to variances for all fitted values.

Figures 4.50 and 4.44 demonstrate how well artificial logic mimics the geometry and the distribution of TS1. Figures illustrate the ability of the synthesised prototypes being “open” to a “message” from the actual data over their Test set which is acting as “token”. However, a full discussion of ”model mimicry” lies beyond the scope of this study.

Data from Listings 4.7–4.12 and Figures 4.50–4.44 indicates that the amount of error between P5/TS1 and P10A/TS1 is small. Together, P5 and P10A closely fit TS1 which is only 1/3 size “clone” of the actual drill-core assay.

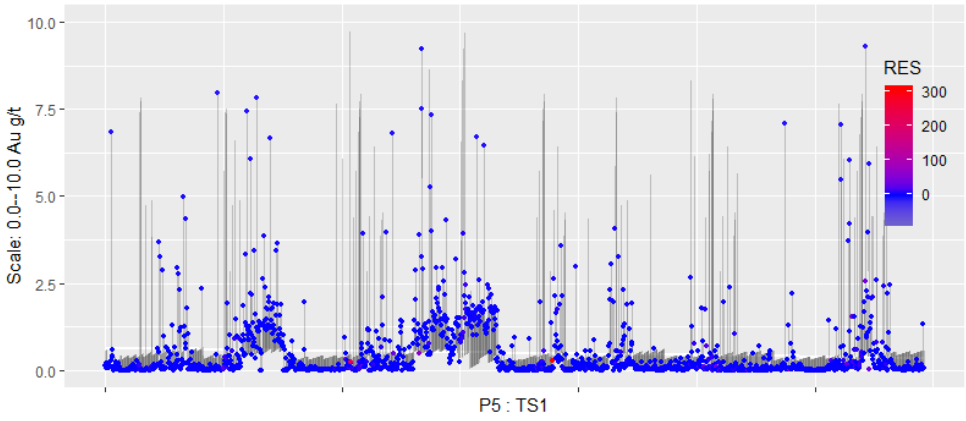

Figure 4.44. P5 (blue) versus TS1. Scale: 0.0–10.0 Au g/t

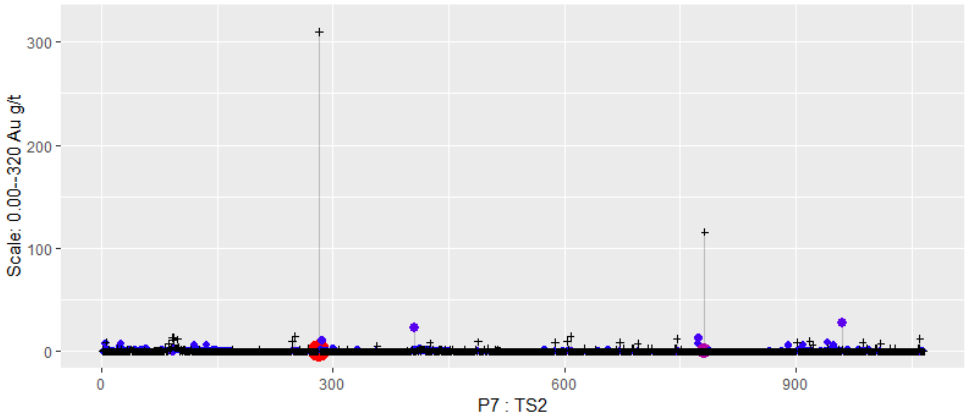

Figure 4.45. P7 (blue) versus TS2, full-range scale: 0.0–350 Au g/t. Higher `abs` values marked in red.

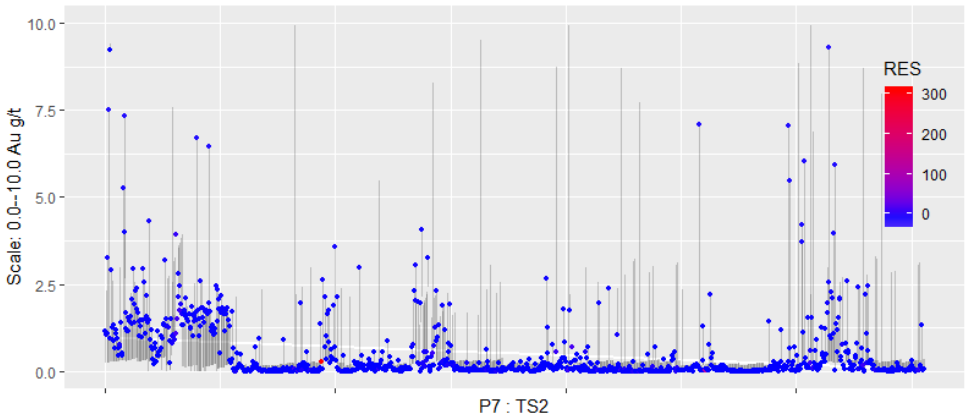

Figure 4.46. P7 (blue) versus TS2, scale: 0.0–10.0 Au g/t

This effect is known as *model mimicry* and is described in Wagenmakers et al. (2004). This example shows that P5 and P10A data is slightly biased.

The root of this problem is that the actual assay has been used twice

1. once to determine the best fitting parameters for *training* set, and
2. secondly, to assess the reasonableness of the model by the *test* sets.

To compare the amount of error in Prototypes 7 (7hp) and 10 (hp) generated from TS2 is shown in Figures 4.46 and 4.48.

To conclude, the correctness of the evidence that Prototype yields data over Test set is affected by the extent to which that training set can mimic the actual data. However, it is important to realise that models P5–P10A (Listings 4.7; 4.12) and P7–P10 (Listings 4.9; 4.11) are not generic models because they imitate the behaviour of whole actual data as well as some specific features over TS1. P5, P7 and P10 have been constructed from the data absorbed by 5,7 and 10 hyperplanes.

To obtain model P10A the topological position of hyperplanes was moderately regenerated. It may appear that the difference between the prototypes and the corresponding test set is not significant. Visual and statistical inspection of the distributions have shown that prototypes and their corresponding test sets have been graphically comparable (c.f. Figure 4.50–4.44 and Figures 4.46–4.48).

This raises new question: *if two distributions have identical fit parameters, similar shapes describe similar matters and contain contextually/semantically similar elements, would the geometries and shapes of 3-D structures be identical?*

The properties and spatial positioning of prototypes, the groups of segments and the internal structure of individual segments shown in sections 4.8.2–4.8.4 clearly indicate that the compared structures are different. However, the comparison of the densities of the means reveals a remarkable similarity between prototypes obtained from one testing data.

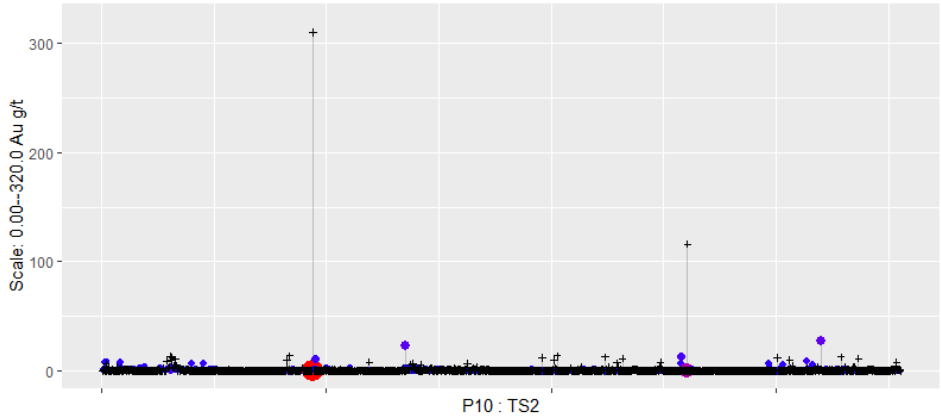

Figure 4.47. P10 versus TS2, full range scale: 0.0–320 Au g/t. Higher `abs` predicted values marked in red.

4.13 Point Cloud as the Approach to Visualisation

The increasing availability of commercial software applications for LiDAR processing point clouds and remote sensing has provided opportunities to generate high-resolution 3-D models of mineral deposits at low cost.

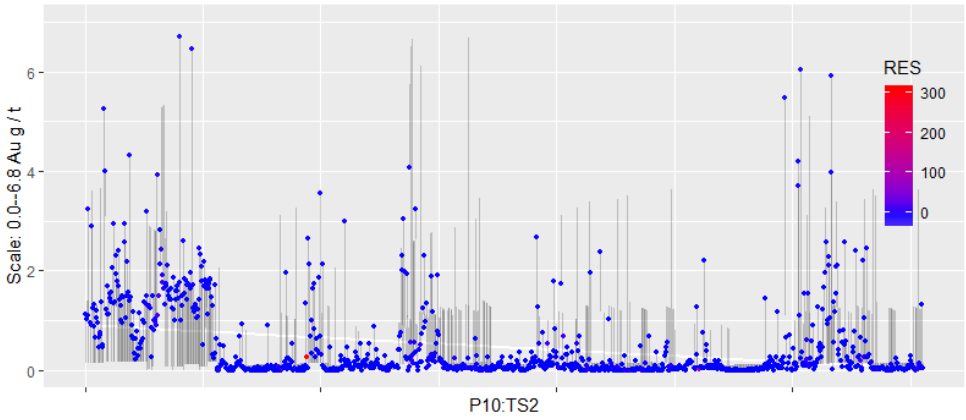

Figure 4.48. P10 (blue) versus TS2, scale: 0.0–6.8 Au g/t

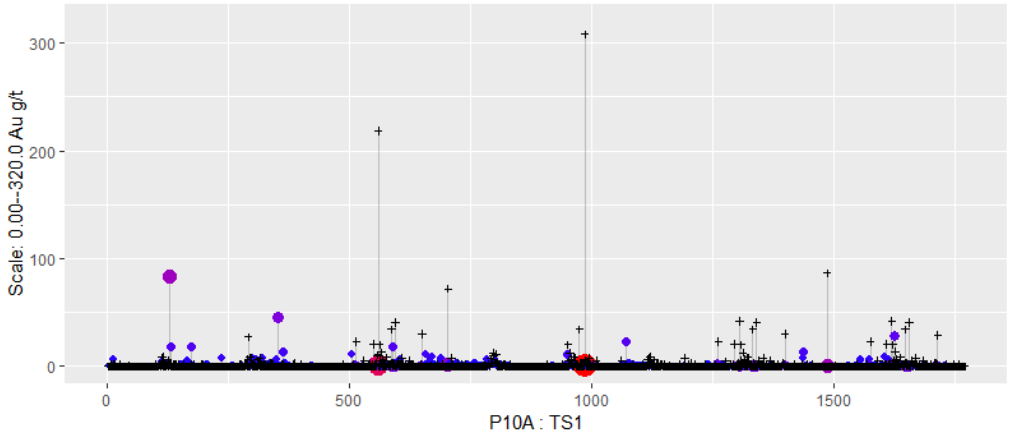

Figure 4.49. P10A : TS1, full range scale: 0.0–320.0 Au g/t. Higher `abs` values marked in red

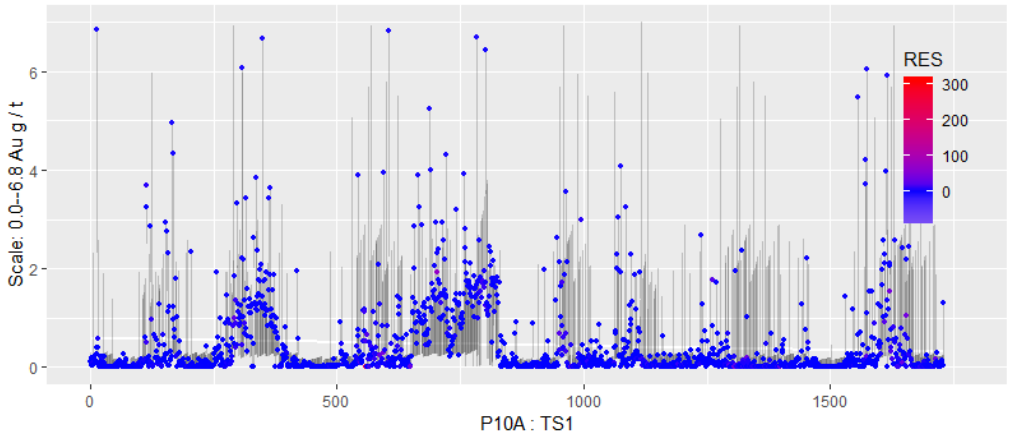

Figure 4.50. P10A (blue) versus TS1, scale: 0.0–6.8 Au g/t

A cloud in conventional LiDAR packages is a large collection of points which creates and renders objects in 3-D. In the same way, a point cloud in GIS applications (for 3-D mapping) is a large number of data points that exist in three dimensions. Technically, a point cloud in GIS is a database containing points in the 3-D coordinate system $(x, y, z +$ attributes). Point cloud imaging seeks to reveal structures hidden by the underlying rock of high density and makes it possible to identify abnormalities. The fusion of CLR and point cloud processing is an efficient way to the representation of nonconvex information, which for many technical reasons cannot be efficiently represented by standard CAD-based tools. The advantages of such approach are as follows

1. Visualising point clouds at full resolution allows modelling an orebody at the highest level of detail with a quality that outperforms many modelling software and is comparable to professional hand-made models.
2. Allows the representations of the interior of an orebody, physical interventions (drilling), as well as the distribution of grades (any mineral).
3. In case of a nonconvex set, point clouds allows the representation of volumetric data, as is done in modern medical imaging - creating visual representations of the interior of an orebody for spatial analysis and physical intervention (drilling), as well as visual representation of the distribution of the recorded grades.
4. This type of visualisation is a semi-transparent raster, which supports effective recovery of geological (and structural) features in both smooth and rough surfaces. Instead of a flat frame with an array of pixels a planner can get a 3-D "container" filled with points and the mineral content (grade, chemical attribute, etc.) may have a unique colour.
5. Remove irrelevant clusters or points.
6. The displayed results are freely switchable between mesh models for efficiency and point cloud for accuracy not achievable in the "mining" software packages.

From the typical workflow perspective, the most important thing is that point cloud is a very accurate digital record of a disseminated orebody. The result is saved in the form of a very large number of points that cover multiple surfaces of clusters, without the need to combine them in a one single structure.

4.13.1 The density of the points

It this thesis, the term density is used to describe graphical resolution of a dataset. Generally, this means the distance from a point to point. Less dense point clouds are obviously much quicker to capture.

Before creating a model or processing a point cloud, it may be necessary to understand the density of the points within the dataset. Point cloud density is an indicator of the resolution of the data: higher density means more information (high resolution) while lower density means less information (low resolution). Modern point cloud processing packages are capable of making a 3-D reconstruction of any environment by capturing thousands of aligned points. 3-D models presented in 4K resolution as clouds, consist of many millions of points, and this is far too many. In terms of density, point clouds can be divided into three groups

1. Sparse point clouds (0.5-1 pts/m^2). Clouds with such low point densities are normally collected for large scale digital height models. Subsets of these point clouds (either based on return number or classification) are used to create digital terrain models (DTM).

2. Medium-density point clouds (2-5 pts/m^2)
3. High-density point clouds (5-10 pts/m^2)

Very often a compromise between point density and cost of acquisition, these data sets are suitable for satisfying most usages. They might not be dense enough for 3-D modelling but can certainly be used to be post-processed and analysed to derive spatial objects. However, using small number of points (below 0.5 pts/m^2) to create a solid, may yield erroneous results, reflect isolated high-grade samples, other than group of clusters, leading to biases in modelling. That is why multiple re-runs of the DGM loop are highly recommended to obtain the number of predictive sets required to obtain at least a sparse cloud (0.5-1 pts/m^2). The CLR is forming the basis for generation of such sets.

4.13.2 Data conversion into 3-D mesh models

Recent collaborative development of MIDAS Information Technology Co. and Samsung allow the elimination of inefficiencies and creating mesh-free models from a cloud of points. However, meshing remains a requirement that is used in some engineering applications.

A mesh is part of a model, which is a collection of vertices, edges and faces. Assume that a face is an ordered collection of vertices connected by edges that complete a loop. It is used to describe a flat surface on the mesh. The idea is to convert point cloud data into intelligent 3D mesh models.

Assume that each predicted point has its own set of x, y, z coordinates and additional attributes such as grades. In a 3-D Cartesian coordinates system, a point can be identified by three coordinates that, taken together, correlate to a precise point in space relative to a point of origin. Then, X,Y,Z axes extend in two directions. After that, the coordinates identify the distance of the point from the intersection of the axes (0) and the direction of divergence, expressed as $+$ or $-$.

There are several techniques and software applications for converting raw point clouds to 3-D objects. Input files are raw x, y, z coordinates in .csv or .txt formats. For example, modern LIDAR widgets also process standard ASCII x, y, z files.

In general, all packages recommend a three-step process

1. Use a 3-D laser scanner to scan the area's measurement.
2. Import the obtained point cloud into point-cloud modelling software. The software lets one visualise and model the point cloud, which at that stage looks rather like a pixelised, digital version of an area of interest sits.
3. Export the point cloud from the software and import it into a CAD system.

4.13.3 The hand-made approach to orebody wireframing

Creating a wireframe (if required) is more complex process. It is critical to understand that the wireframe of orebody cannot be in any shape automatically translated into resource or reserve CAD model. In practice, there are enough reasons to create (digitise) and calibrate the wireframes manually:

1. All gold mines are being mined partially. In our case, the reefs and other structures are mostly tilted to the north east at angle from $\measuredangle 20°$ to $\measuredangle > 60$ dip°. This means when mining, the steeper the dip, the more localised the subsidence is likely to be. Faults, voids and groundwater levels also can give rise to fault steps. During

mining, steeply dipping stopes will tend to remain open and may pose a thread to operation. For production safety, highly unstable zones in some cases are being excluded from the reserve base.

2. Due to the complexity of the relationships between grades and barren rock zones, the wireframe(s) are constructed manually from manually drawn and joined parallel strings.
3. Some hand-made 3-D models of narrow-vein systems can achieve arbitrary levels of detail and fidelity.

As to point clouds, they cannot be automatically translated into a solid. Along with increasing computation power, most algorithms are still only tools shortening time for point cloud modelling. The reality is that drilling data → CAD or cloud → CAD conversion cannot be a fully automatic process as well. The point cloud interpretation process is still needed for human involvement and managing point clouds manually. However, solving this problem is a matter of time.

The density and the size of a cloud are subject to modification. Nearly all existing software packages offer to change the display settings for point clouds to simplify the display and improve performance. It is also possible to apply colour stylisations to visually represent the point cloud data.

However, manual adjustments are still labour intensive, expensive and time consuming for large mapping projects. For example, in LiDAR applications, program performance and visual noise is still manually managed by increasing or decreasing the number of visible points and point size.

4.14 Conclusion

The actual assay (4.4) showed only variability of grades from hole to hole, such that, there are many unexplored areas with potentially worthwhile targets left between the drillholes. The main goal of Chapter 4 has been the design of a machine learning system, which is capable of predicting the mineral content in these unexplored areas.

Through a survey of the literature, it has been found that the clusterwise linear regression is one of the most promising methods to be utilised. Therefore, it has been decided to explore the applicability of CLR in the hope of answering a question: *"is it possible to apply the CLR to the prediction of the existence/absence of exploration targets from the available (limited) drilling data?"*

The literature illustrated a number of advantages brought by CLR and suggested that a reformulation of the objective function could provide good forecasting outcome.

Then, the problem of predicting the gold has been reformulated into a clusterwise linear regression problem and solved as a nonconvex nonsmooth optimisation problem. A framework of a decision support predictive methodology has been designed. An execution code, by which solutions to highly constrained optimisation problem be solved reliably and efficiently has been written (Appendix B).

The system has been tuned up for solving complex geological and topological constraints as well as nonlinear problems related to the presence of heterogeneity in data. Formally, it has been attempted to identify potentially gold-bearing localised occurrences missed during exploration. Next, the developed unsupervised learning system has been applied for predicting the gold distribution over a real-world area of 6766.2 m^3 using 5334 samples, of which 4669 were found valid and were involved in the computation.

After that, to improve the accuracy of prediction, the local search has been performed with the discrete gradient method (DGM) which learned the data through class-related hyperplanes and computed four sets of locally optimal solutions, using minimal regression mean-squared error (RMSE). The system avoided some limitations of the traditional methods and constructed consistent approximations in the form of multi-object structures.

Finally, the applied UML system has brought data to optimality on its own without humans intervention and solved locally optimal instances in the actual data to optimality. Running the program several times yielded different results per each iteration.

The multiple spatial realisations permitted intrinsic heterogeneity to be transferred to the final deterministic samples at the locations generated by the proposed method. The results pointed at the likelihood of the presence in the data numerous clusters of gold which have never been previously identified. It was found that the system performed satisfactorily during all conducted runs.

4.14.1 Comparison of prototypes

To facilitate visual comparison, for histograms were plotted in the same boxplot chart. The comparison of the distributions with multiple boxpots and a beanplot revealed semantic similarity of P5 to P10A and P7 to P10. The applied statistics revealed that the difference between major parameters in the compared sets have not been significantly great.

To explore the frequency distributions of low grades (0.0–1.8 Au g/t), the prototypes were declustered with histograms with equal bin sizes. A side-by-side comparison of four histograms revealed a significant difference between the sets in terms of shapes, x-axes, the distribution of bins and contrast in the skewness of the distributions.

It was found that sets P5:P10A and P7:P10 describe same matter and they are graphically comparable. Moreover, datasets P5, P7, P10, P10A illustrate significant similarity and relatedness to the actual data. These connections between theory and the applied methods give a reason to believe the applied method was on the right track: the applied clustering provided similar and comparable notions.

A comparison of the distributions revealed significant similarities between the sets based of their structural attributes. The predicted distribution illustrated significant relatedness to the testing data and actual drill-core assay. The features that make the generated nonconvex set different were: the spatial positioning, geometry and orientation.

The results obtained from AIC (eg. Listings 4.4, lines 57–63; Listings 4.6, lines 112–116) were not entirely convincing. In a recent study by Velasco & González-Salazar (2019) a strong correlation between AIC values and estimation accuracy from validation metrics was not established.

The applied `mfrow=c(2,2)` R-function to comparison did not provide enough evidence to suggest that the compared sets were significantly different.

4.14.2 The effect from adding hyperplanes

The applied polar Kernel density estimator (KDE) established that the patterns and matrices (see Appendix D) of the probable density of means in the prototypes obtained from one test set have been near-similar. The highest isolated concentrations of low grades in P5:P10A and P7:P10 have been between 0.0–1.8 in one direction. The regions

where any points in the samples have not been observed had very similar low and ultra-low density patterns.

A comparison of the patterns and matrices have not revealed significant differences between the sets based of their structural attributes, such as the densities of the means. The compared distributions in P5:P10A → Test set 1 and P7:P10 → Test Set 2 illustrate significant relatedness to the actual drill-core assay.

The effect of adding extra hyperplanes has not been significant and has not changed the patterns of probable densities of the means. Statistics suggested that the highest and the lowest values in P5:P10A and P7:10A are near similar. This is a good sign. The exploration team can utilise any set, no matter how many hyperplanes were added to the preset number. Theoretically, one of the predictive models produced from a one Test Set, may mathematically provide better prediction, but the amount of that advantage will be newer known.

To sum, it was found that the increase of the number of hyperplanes does not lead to any significant increase of the accuracy of prediction. Another word, when the number of linear regression functions is getting larger the incremental algorithm does not necessarily produce significantly more accurate solutions.

Importantly, the KDE suggested a significant structural difference between Test sets 1 and 2

4.14.3 Prediction

The actual assay suggests that the deposit is polymetallic and contains a combination of different metals (copper, magnesium, nickel, manganese), and some critical technical elements such as cadmium. Importantly, it is technically possible to recover and separate most of these metals from this deposit.

The distribution of gold is compounded by the presence of highly disseminated low and low–middle grades in the form of dust and small grains, which may make mineralisation economic. There are some signs of multiple, from low to middle-grade concentrations may occur within the strike extensive N and N-E branch.

According to Berkmam (2001), Table 4.4.2, the average predicted diameter of gold spheres (grains) may range from 0.08 to 0.127mm. That means, at the production phase, the low–middle grade ore treatment will probably involve gravity concentration–crushing. The particles about 80-110 μm will probably be liberated by fine grinding. However, to process the low-grade material on site, the dissolving of gold in a weak solution of sodium cyanide may be required (an important cost actor).

A possible existence of locally closely-spaced low-middle grade occurrence, potentially clustered fine-grained samples and a number of closely spaced groups of middle-grade coarse particles (0.12 to 0.5mm) in the N and N-E parts of domain.

Ninety four gold-bearing structures, ranging from 1.0m to 15.0m in length have been predicted. The average expected grade of the low-grade structures ≈+0.7 Au g/t. The operator may expect numerous scattered low-middle grade occurrence ≈+1.3 across the domain.

The shortest predicted interval =0.6m with grade +0.5 Au g/t was generated outside the search area (lease) at depth 560m. Six distinct high-priority exploration targets have been predicted mostly in the NN-E part between mRL210 and mRL360. One of the potentially worthwhile targets may occur in the northern end of the domain.

The predicted data suggests that multiple intersections ranging from 3.5 to 15.0m in length generally grade between 1.27 and 4.85 Au g/t in ∡azimuths 214–320° at depth from 210 to 330 metres. Numerous insignificant interceptions between 1.0 and +1.3 Au g/t may exist in the eastern direction. The interceptions listed below suggest the length of segments (structures), the average grade and the depth in metres.

- Segment 11: 14.0m @ 2.34 Au g/t from 285.0m
- Segment 11: 14.7m @ 1.28 Au g/t from 300.0m
- Segment 23: 14.9m @ 4.39 Au g/t from 285.3m, likelihood of multiple occurrence
- Segment 24: 14.0m @ 1.44 Au g/t from 300.2m
- Segment 25: 13.9m @ 1.27 Au g/t from 315.0
- Segment 28: 14.0m @ 0.79 Au g/t from 360.0m, likelihood of coarse gold
- Segment 38: 13.5m @ 4.73 Au g/t from 285.0m
- Segment 39: 14.5m @ 2.44 Au g/t from 300.0m
- Segment 42: 14.0m @ 2.76 Au g/t from 345.0m, likelihood of occurrence >50 g/t
- Segment 48: 20.0m @ 4.85 Au g/t from 430.0m, likelihood of occurrence >10 g/t
- Segment 60: 14.0m @ 4.91 Au g/t from 286.0m, likelihood of occurrence >20 g/t
- Segment 62: 3.50m @ 2.11 Au g/t from 315.9m
- Segment 69: 14.0m @ 2.63 Au g/t from 285.0m
- Segment 78: 14.9m @ 3.58 Au g/t from 270.1m
- Segment 79: 14.5m @ 4.59 Au g/t from 285.0m
- Segment 85: 14.9m @ 3.23 Au g/t from 210.0m, likelihood of occurrence >50 g/t
- Segment 89: 15.0m @ 3.53 Au g/t from 270.0m, likelihood of occurrence >30 g/t
- Segment 90: 14.5m @ 4.50 Au g/t from 285.0m

Two zones of potentially stable mineralisation in N and N-E parts of domain are suggested be the focus of the future drill campaign to commence the process of making discoveries. However, a very limited number of data, do not let low density to form a cloud of points. Accurate delineation of the two mineralisation zones without extra probes is problematic.

The most important information is that the exploration of adjacent areas can be economic.

4.14.4 Limitations

The solutions from UML provide the exploration crew with useful information on projected grades in areas of the domain with previously unknown characteristics. However, the limited number of predicted samples obtained from four program runs do not let exhibiting trends in the average gold grade, as well as the fluctuations of grades across the entire domain.

Pragmatically, prediction by optimisation should determine (or at least suggest) the shapes of solids. More specifically, the graphical output must be an accurate representation of 3-D scalar data (not necessarily a mesh-based), which is volume. Importantly, the amount of actual and predicted data must be large enough for the comprehensible representation of the distribution and conversion that distribution into a 3-D solid.

The roots of this limitation lie in the absence of geostatistical assumptions. On the one hand, unlike interpolation, the applied type of optimisation assumes the attribute data are not continuous over space. Therefore, the UML does not allow the estimation of the attribute at any selected location within the exploration boundary.

On the other hand, the attributes in NSO are spatially independent and the values closer together are not necessarily similar than the values farther apart.

As a consequence, these features make the UML unable to automatically create a 3-D solid and provide that solid with volumetric parameters.

Finally, it was not investigated how potential end-users may comprehend the used in chapter 4 3-D visualisation and what graphical properties make it more difficult for end-users to understand the insights it is attempting to convey.

4.14.5 Solutions to volume visualisation

Volume visualisation remains one of the most research topics in applied visualisation, such as magnetic resonance tomography (MRT) in which the volumetric information is provided by virtual geometric solids.

In chapter 4 a number of renders with mathematical content were synthesised to demonstrate complex optimisation problems within visualised contexts. The presented stage-performed 3-D mapping has been abstracting locally optimal solutions into a visualisation, consisting of primitives such as lines and cylinders. Each generated point has been attributed by grade in Au g/t.

To make a near-photorealistic environment and show locally nonconvex solutions, a 3-D viewpoint has been created with the involvement of the OpenGL interface. However, the used type of visualisation with a low point density is not new and does not let practical triangular meshes to be created and a complete 3-D solid to be constructed.

Since the UML provided adequate information about locations and grades of predicted points, it would be a straightforward step to extend Method 1 to a modern 3-D situation - a cloud of points.

To make mine evaluation effective, it is proposed to generate million(s) points to form a point cloud of middle or high density. Multiple DGM runs can generate millions of data points in space. In this case, groups of predicted points become a cloud of points. Converted into CAD, a high-density the cloud forms multiple semi-transparent, an MRT image-like objects.

In this scenario, the need of a wireframe in forming external an external surfaces is questionable. Depending on task, such point cloud can be viewed and explored in both the OpenGL view and in the DXT in orthogonal projection. The shape, dip and strike of such optimised structures is presented as a collection of points that represent a "transparent" 3-D shape of the underlying structures defined by a selected coordinates system. In a 3-D Cartesian coordinates system, a point is identified by three coordinates x, y, z and grade.

4.14.5.1 Pessimistic solution

While the locally optimal solutions do not provide volumetric information, these can assist in the identification of the orientation, depth and orientation of a hidden exploration target. In the pessimistic scenario, the number of optimisation runs and consequently, the number of generated locally optimal solutions remains limited and such dataset becomes a non-uniform 3-D set without normals. That means the transformation of optimisation results into triangular meshes is hardly possible. In this case, the information on the predicted points with x, y, z Cartesian coordinates and the attributed grades allows to

1. Identify high priority high-grade targets and their locations.
2. Explore the potential of the zones adjacent to lease.
3. The identified high-grade zones can be used for future exploration efforts.
4. Evaluate the benefits of increasing the area of lease.

4.14.5.2 Pragmatic solution

A point of cloud becomes manageable. Increasing the number of locally optimal solutions to a million allows their convergence into a point-cloud processing software (e.g. HighRES integrated point cloud processor) and further export into CAD systems such as SolidEDGE or SpaceClaim. This approach allows to drape a surface over a point cloud or wrap a mesh around a point cloud.

1. Test high priority high-grade targets and their locations.
2. Predict zones of potentially worthwhile intersections.
3. Explore the potential of the zones adjacent to lease.
4. Orient trajectories during drilling.
5. The identified high-grade zones will be used for future exploration efforts.
6. Evaluate the benefits of increasing the area of lease.

4.14.5.3 Optimistic solution

Fog, consisting of a millions points is exported into a point cloud post- processing software, which provides accurate conversion into a transparent geological view. This approach allows to

1. Identify high priority and high-grade targets and their locations.
2. Explore the potential of the zones adjacent to lease.
3. Orient trajectories during drilling.
4. The identified high-grade zones will be used for future exploration efforts.
5. Evaluate the benefits of increasing the area of lease.
6. Reconstruct a number of continuous scalar functions which can be evaluated at any arbitrary location.
7. Determine the shapes and volumes of multiple solids.
8. Exclude potentially problematic areas from the reserves.

Chapter 5

Method 2: Penalised Regression - LASSO

Method 2 was developed in collaboration with the School of Mathematical and Physical Sciences, University of Technology Sydney. Results obtained in this chapter were presented on 4 April 2019 at the 3rd Australasian Commodity Markets Conference, Sydney and submitted for publication in the paper: I.Grigoryev, J.Hinz and A.Novikov "Commodity risk reduction by predictive modelling with the LASSO" in the Journal of Commodity Markets (JCM), Elsevier. Financial econometrics has been omitted in the chapter.

Solving problems of uncertainty quantification and predicting the distribution of the underlying mineral content, especially precious metals require the application of complex cross-disciplinary methods. Recent developments in high-dimensional statistics (HDS) and machine learning provide a variety of methods, which can be used in mine evaluation and for efficient extraction. More specifically, some HDS methods are capable of taking the prediction accuracy over control by the reduction of the so-called generalisation error, which refers to a cumulative result deterioration (due to uncertain/ misspecified model and its estimation errors) when applying predictions to new observations.

This chapter is devoted to the development of an HDS-based optimisation technique for modelling the distribution of gold at the production stage. A penalised regression is applied to address the problem of geological uncertainty and determine worthwhile targets in the unexplored areas of a mine. Due to built-in strict parameter selection and enhanced predictive performance, the presented technique provides statistically valid and accurate modelling infrastructure.

It will be shown that geological assumptions can be combined with advanced penalised regression to obtain a sparse statistical model with enhanced predictive ability. The presented method helps in obtaining answers to questions related to the distribution of gold-bearing ore within a real-world production dataset.

5.1 Feature Selection

Inability to accurately delineate a mineral deposit is considered a major impediment to success. Several published sources (in chapter 2) reveal that the presence of structural and functional heterogeneity in the ground require the application of increasingly advanced quantitative optimisation methods to predict anomalies, detect worthwhile targets, minimise exploration time and the cost of discovery.

To provide better justification for a second optimisation objective to be investigated, additional literature on HDS was reviewed.

Statistical spatial heterogeneity (Fotheringham et al., 2002) is a condition where a global regression model cannot describe the relationship between the response variable and explanatory variables because of the variation of characteristics among the observations' regions.

According to Bermingham et al. (2015) a statistical technique, known as the feature selection (FS) is used in situations in which the data contains some features that are either redundant or irrelevant and can thus be removed without incurring much loss of information (Bühlmann & Geer, 2011; James et al., 2015, Hastie et al., 2017).

More recent literature on the methods for regression (e.g. Hastie et al., 2017, Ch.3) suggests that various types of optimisation problems can be solved with the so called ℓ_1-norm as a regulariser. An overview of some basic optimisation algorithms for convex problems, with an emphasis on aspects of particular relevance to regularised estimators such as LASSO can be found in James et al. (2015) and Hastie et al. (2017).

This chapter is looking at the evaluation from a different angle of view and presents an approach that integrates HDS and mathematical theory and is suitable for finding global spatial optimums in real-world datasets.

Recent developments in HDS have established that the prediction accuracy can be improved by

1. Shrinking the values of the regression coefficients (ridging)
2. Setting some coefficients to zero (LASSO)

Several surveys report that both Lasso and ridge reduce the variance of the predicted values and may improve the overall prediction accuracy in terms of the mean-squared error (James et al., 1996; Tibshirani et al., 2010; James et al., 2015; Hastie et al., 2017).

Seeking an appropriate method that can be used to treat the prediction of minerals revealed that the FS methodology is a strong alternative to existing global optimisation methods.

The main reasons why the FS is interesting are

1. The FS reduces the size of the problem to enable algorithms to work faster.
2. It makes the output easier to interpret.
3. It provides an accurate prediction on new data by minimising overfitting on the training set (90%).
4. It provides better sparse data handling.

Feather review of modern statistical methods described by Simon et al. (2011), Friedman et al. (2008), Tibshirani et al. (2010) and James et al. (2013) revealed that some ℓ_1-regularisation techniques such as sparse regression, known as LASSO perform FS as part of their overall operation. Proposed in 1996 by Tibshirani (1996), this penalisation method is built on the ℓ_1–norm, and is now widely used to tackle simultaneously variable estimation and feature selection in sparse problems.

As previously discussed (chapter 2), the prediction of mineral variability is often viewed as function interpolation methodology that has been criticised for the requirement of a number of assumptions (section 2.4.1). Classical geostatistical methods presume a certain probabilistic structure obtained from variograms. Loosely speaking, it is assumed that the observation of a mineral concentration at location $S \in \mathcal{O}$ follows a random field $(Z(S))_{S\in\mathcal{O}}$ indexed by points of a given geological structure $\mathcal{O} \subset \mathbb{R}^3$.

Having adopted appropriate assumptions on this random field such as stationarity, covariance structure, etc. and given measurements $(Z(S_i)_{i=1}^I$ at locations $S_i \in \mathcal{O}$, $i = 1, \dots I$, the estimation of $Z(S_0)$ at an unobserved point $S_0 \notin \{S_1, \dots, S_I\}$ is defined in terms of conditional expectations

$$\mathbb{E}(Z(S_0) \,|\, Z(S_1), \dots, Z(S_I)) \tag{5.1}$$

While a similar probabilistic framework and the estimation in terms of conditional expectation as in (5.1) is followed in this chapter, the proposed approach is different.

Instead of relying on stationarity with an inconvenient need to specify a certain dependence type, it is supposed that the dependence between spatial observations is modelled naturally, in terms of a hidden structure, which contains a certain concentration of commercially worthwhile content.

The shape, size, and content of a deposit are random and determined by a superposition of the so-called *prime deposits*, which are modelled by *ball-shaped* structures with uncertain radius, locations and the rock mass. Having assumed that such deposits may virtually occur at any location of the domain, an "over-determined" spatial model is obtained, which can be estimated using modern HDS techniques.

The idea is that the FS can bring spatial data to global spatial optimality by removing irrelevant values from blasthole assay (5.4.1) without occurring much loss of information. The proposed estimation is performed by the selection of a most appropriate model from the viewpoint of its predictive accuracy.

5.1.1 LASSO versus ridge

Data sets containing more features than observations are referred to as high-dimensional statistics (Bühlmann & Geer, 2011). A high-dimensional case can be defined as the case where the number of features p is larger than the number of observations n, that is $p \gg n$.

Some statistical inference such as the least squares linear regression cannot be applied for solving HDS problems because the conventional statistical framework for fitting many parameters is based on assuming structural smoothness, enabling the estimation of smooth functions.

The high-dimensional inference is based on the assumption of the existence of certain "sparsity". Since a key property of the ℓ_1-constraint is its ability to yield sparse solutions, the term "sparse" is used in this chapter for models with few nonzero coefficients. Other definitions and quantification of sparsity in data can be found in Bühlmann & Geer (2011).

Linear regression is often met in many applications because this is one of the standard statistical tools, which is usually characterised by the matrix and vector notation. Linear regression attempts to model a dependent variable using the best straight line fit to a set of predictor variables.

The best fit is usually taken to be one that minimises the RMSE, which is the sum of square of the differences between the actual and predicted values of the dependent variable.

A linear regression model is characterised by the relation

$$Y = \mathbf{X}\beta^* + \varepsilon \tag{5.2}$$

Where

$Y = (Y_i)_{i=1}^n$ are observed response variables, $n \in \mathbb{N}$

$\mathbf{X} = (\mathbf{X}_{i,j})_{i=1,j=1}^{n,p}$ is a given matrix

$\beta^* = (\beta_j^*)_{j=1}^p \in^p$ is unknown vector of coefficients (oracle)

$\varepsilon = (\varepsilon_i)_{i=1}^n$ is a random variable (zero-mean, with unknown variance σ^2)

n is the sample size

p denotes number of predictors

Note that the matrix entries $\mathbf{X} = (\mathbf{X}_{i,j})_{i=1,j=1}^{n,p}$ are non-random by assumption. This situation is referred to *fixed design*. The columns $\mathbf{X}_{\cdot,j}$ for $j = 1, \ldots, p$ are called *predictor realisations*. In practice, the data Y, $\mathbf{X}$ of a linear model occur in the following settings:

Suppose that a phenomenon is modelled in terms of response variable Y_1 and predictor variables $X_1 = (X_{1,j})_{j=1}^{p}$ with a hypothetical relation

$$Y_1 = \sum_{j=1}^{p} X_{1,j}\beta_j^* + \varepsilon_1 \tag{5.3}$$

such that ε_1 and $(X_{1,j})_{j=1}^{p}$ are independent and the coefficients $\beta^* = (\beta_j^*)_{j=1}^{p}$ are not known.

A sequence of n independent observations of this random phenomena is modelled by n independent copies of $(X_i, Y_i)_{i=1}^{n}$ and (X_1, Y_1).

Having observed the realisations $\mathbf{X}_{i,\cdot}$ of predictor variables X_i for $i = 1, \ldots, n$, the design matrix $\mathbf{X}$ is obtained and the relation $Y = \mathbf{X}\beta^* + \varepsilon$ is claimed.

The explanatory variables $(X_{1,j})_{j=1}^{p}$ can usually be associated with functions $f_1, \ldots, f_p$ of some other (multivariate) explanatory random variable, say U_1 in the sense $(X_{1,j} = f_j(U_1))_{j=1}^{p}$.

That is, the design matrix in this case, is given by a realisation of

$$X = (f_j(U_i))_{i=1,j=1}^{n,p} \tag{5.4}$$

for independent copies $U_1, \ldots, U_n$ of U_1. The model in (5.2) is linear because it is linear in the coefficients $\beta^* = (\beta_j^*)_{j=1}^{p}$.

As mentioned above, the predictor (p-variables) can be obtained via non-linear transformation of some (multivariate) explanatory factors. Similarly, the response variable can also be described by some non-linear transformation of the actual observations, for instance by their measurement on a different (logarithmic) scale. On this account, linear models are able to capture non-linear dependencies between random phenomena.

However, the entire methodology relies on a correct choice of transforming functions, i.e. $f_1, \ldots, f_p$. In the framework of linear models, this task must be performed manually.

In applications of linear models, the estimation of β^* from observation of $(X_i, Y_i)_{i=1}^{n}$ is typically unstable, if the entries of the random vector X_1 are correlated. In this case, the least-squares estimation may be ill-posed (small eigenvalues in the matrix $\mathbf{X}^\top\mathbf{X}$).

Furthermore, there are applications where with more potential explanatory variables than observations ($n < p$). In such situations the ridge regression (see James et al., 2013, p.215) penalises the size of the regression coefficients, suggests solving a penalised least squares problem

$$\hat{\beta}(\lambda) = \operatorname{argmin}_{\beta\in\mathbb{R}^p} \left(\|Y - \mathbf{X}\beta\|^2 + \lambda\|\beta\|^2 \right) \tag{5.5}$$

with an appropriate penalisation parameter $\lambda > 0$. Note that even in the case that $\mathbf{X}$ does not have full rank, this problem possesses a unique solution, which is obtained from the equation

$$(\mathbf{X}^\top\mathbf{X} + \lambda\mathbf{1})\hat{\beta}(\lambda) = \mathbf{X}^\top Y, \tag{5.6}$$

Where **1** stands for the identical matrix. The ridge penalty λ causes a shrinkage of coefficients

$$\hat{\beta}(\lambda) \to 0 \quad \text{for } \lambda \to \infty.$$

This usually yields a variance reduction

$$\text{Var}(\hat{\beta}(\lambda)) < \text{Var}(\hat{\beta}(\lambda')) \quad \text{for } \lambda > \lambda' > 0 \tag{5.7}$$

but introduces a systematic error (bias)

$$\mathbb{E}(\hat{\beta}(\lambda)) \neq \beta^*, \quad \lambda > 0. \tag{5.8}$$

The LASSO is described in James et al. (1996) as a regularisation method for regression that uses an ℓ_1 penalty to achieve a sparse solution and the name "LASSO" was also introduced as an acronym for Least Absolute Selection and Shrinkage. The goal of this process is to minimise the prediction error.

This method is a variation of ridge regression because the both methods have similar formulations. Both regularisation work by adding a penalty term to the log-likelihood function. For further reading it is suggested a study by Bühlmann & Geer (2011).

In the case of ridge regression, the penalty term is β_1^2 and in case of the LASSO the penalty is $|\beta_1|$. The quantity to be minimised in the two cases as follows

$$L + \lambda \sum \beta_1^2 - \text{in case of ridge regression}$$

$$L + \lambda \sum |\beta_1| - \text{in case of LASSO regression}$$

Where λ is a free parameter which is usually selected in such a way that the resulting model minimises the out of sample error, β is regression coefficient and L is log likelihood function.

Figure 5.1 shows the difference between the LASSO and ridge estimations. The solid blue areas are the constraint regions, the red ellipses are the contours of the residual-sum-of-squares function (or least squares error function). The point $\hat{\beta}$ on Figure 5.1 depicts the unconstrained least-squares estimate. In the LASSO the constrained region has corners, i.e. if the first point is in the proximity of the corner, then it has one coefficient β_j equal to zero.

If one have relaxed conditions on the coefficients, then the constrained regions on Figure 5.1 are getting bigger and eventually they will hit the centres of the ellipses. Both LASSO and ridge determine coefficients by finding the first point where the elliptical contours hit the region of constraints.

In the case of ridge regression, the effect of the penalty term ℓ_2 is to shrink the coefficients that contribute most to the error, i.e. the ℓ_2 reduces the magnitude of the coefficients that contribute to increasing L, i.e. the log-likelihood function.

In the case of LASSO regression (Figure 5.1), the effect of the penalty term ℓ_1 is to set these coefficients exactly to zero, i.e. the LASSO selects the most predictive coefficients that provide the lowest p-values. This unique feature of the LASSO brings a major improvement from the viewpoint of generalisation error reduction since it determines $\widehat{\beta}(\lambda)$ as "minimise".

Hence, the LASSO performs variable selection and models generated from the LASSO are generally easier to interpret than those produced by ridge regression.

The only difference from ridge regression is that the regularisation term is in an

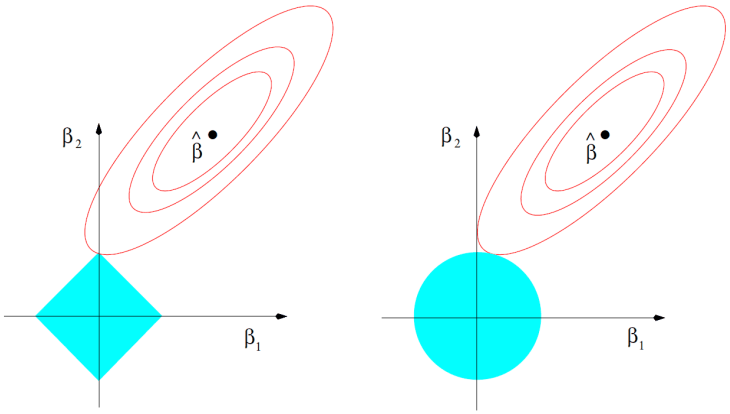

Figure 5.1. An illustration of 2-D case of estimation for the LASSO (left) and ridge regression (right). Adapted from James et al. (2015)

absolute value. LASSO regression can lead to feature selection whereas ridge can only shrink coefficients close to zero. This feature is the reason of our interest to the LASSO. This method overcomes the disadvantage of ridge regression by not only punishing high values of the coefficients β but actually setting them to zero if they are not relevant.

With the goal of the variable selection in mind, the LASSO brings a major improvement from the viewpoint of the generalisation error reduction since LASSO determines $\hat{\beta}(\lambda)$ as a minimiser

$$\hat{\beta}(\lambda) = \text{argmin}_{\beta \in \mathbb{R}^p} \left(\|Y - \mathbf{X}\beta\|^2 + \underbrace{\lambda \|\beta\|_1}_{\text{penalisation}} \right) \tag{5.9}$$

with l_1-norm (instead of the Euclidean norm) on the coefficients

$$\|\beta\|_1 = \sum_{j=1}^{p} |\beta_j|, \qquad \beta = (\beta_j)_{j=1}^{p} \in \mathbb{R}^p. \tag{5.10}$$

*Note that in contrast to the ridge regression, this minimisation problem is not solvable by simple linear algebra but rather needs for quadratic programming or related algorithms. More importantly, such a solution is more appropriate because usually many coefficients in $\hat{\beta}(\lambda)$ decay automatically (which yields a kind of model selection by variables reduction).

Below is the answer to why $\hat{\beta}(\lambda)$ is a sparse vector. Finding the unconstrained minimum

$$\beta \mapsto \|Y - \mathbf{X}\beta\|^2 + \lambda \|\beta\|_1 \tag{5.11}$$

is equivalent to finding the constrained minimum

$$\beta \mapsto \|Y - \mathbf{X}\beta\|^2 \quad \text{subject to} \quad \|\beta\|_1 \leq R \tag{5.12}$$

for an appropriate boundary parameter $R = R(\lambda)$. Since the ball $\|\beta\|_1 \leq R$ has sharp edges, the solution $\widehat{\beta}(\lambda)$ is a sparse vector, which can be determined by a very efficient algorithm - coordinate descent. The coordinate descent finds the minimum so fast, that it is possible to determine $\widehat{\beta}(\lambda)$ for many (all relevant) $\lambda \in \mathbb{R}_+$.

Because of this, one can choose the best $\widehat{\lambda}$ through `nfold` cross-validation.

Steps for the LASSO procedure are as follows

1. Divide all observations into train and test sample (90 % train, 10 % test). Do this randomly in 10 different ways.
2. Determine $\widehat{\beta}(\lambda)$ on train data, and with this $\widehat{\beta}(\lambda)$ calculate the density on test data. Average the result over all 10 data sets. Do it for (all relevant) $\lambda \in \mathbb{R}_+$.
3. Determine that penalty parameter $\widehat{\lambda}$, which yields the best result (averaged density realisation) on test data.
4. Obtain the coefficients as $\widehat{\beta}(\widehat{\lambda})$.

To assess the quality of the model, the model parameters are estimated using LASSO for the training and test data sets.

5.1.2 R-`glmnet` package

Optimisation with `glmnet`: for all models, the `glmnet` provides a range of elastic-net penalties ranging from ℓ_2 (ridge) to ℓ_1 (LASSO). The regularisation path is computed for the LASSO at a grid of values for the regularisation parameter λ and solves the following problem over a grid of values of λ covering the entire range:

$$\min_{\beta_0,\beta} \frac{1}{N} \sum_{i=1}^{N} \omega_i \iota \left(y_1, \beta_0 + \beta^T x_i\right) + \lambda \left[\left(1 - \alpha \parallel \beta \parallel_2^2 /2 + \alpha \parallel \beta \parallel_1\right)\right]. \tag{5.13}$$

Where $l(y, \eta)$ is the negative lag-likelihood contribution for observation i. For the Gaussian case it is $\frac{1}{2}(y - \eta)^2$.

The elastic net penalty is controlled by parameter α and bridges the gap between LASSO $\alpha = 1$ by default and ridge $\alpha = 0$. The tuning parameter λ (regularisation coefficient) determines the overall complexity of the model and controls the overall strength of the penalty.

The developers of `glmnet` (Friedman et al., 2008) state the following:

The ridge penalty shrinks the coefficients of correlated predictors towards each other while the lasso tends to pick one of them and discard the others. The elastic-net penalty mixes these two: if predictors are correlated in groups, an $\alpha = 0.5$ tends to select the groups in or out together. This is a higher level parameter, and users might pick a value upfront, else experiment with a few different values. One use of α is for numerical stability: for example, the elastic-net with $\alpha = 1 - \epsilon$ for some small $\epsilon > 0$ performs much like the LASSO, but removes any degeneracies caused by extreme correlations.

The ridge penalty ℓ_2 shrinks the coefficients of correlated predictors towards each other, while the LASSO ℓ_1 results in sparsity and tends to pick one of them and discard the others.

By combining a squared ℓ_2-penalty with the ℓ_1-penalty the elastic net penalty is obtained (Zou & Hastie, 2005; James et al., 2015, p.56; Hastie et al., 2017, p.73). For example, if predictors are correlated in groups, an α=0.5 tends to select the groups in or out together.

In section 5.5 the `glmnet` algorithms use cyclical coordinate descent, which optimises the objective function over each parameter with others fixed, and cycles repeatedly until convergence. Due to highly efficient updates and techniques such as warm starts and active-set convergence, the `glmnet` usually computes the solution path very fast.

In the R-source code (Appendix C) the main function is `glmnet` which uses `glmnet()` command to fit LASSO 5.10 model. In particular, the LASSO has been executed to

recover gold concentrations (Appendix C) in g/m^3 from `marvin.csv` (5.4.1) dataset.

Since the package takes only numerical variables, a function `make model matrix` is required for creating `x` matrix, corresponding to the all available predictors and transforming qualitative variables into dummy variables.

5.2 Some Peculiarities of Design

Suppose that within a predetermined gold-bearing rock mass $\mathcal{O} \subset \mathbb{R}^3$ there are sparsely distributed gold observations at locations which are based on samples obtained during blast-hole grade control program. Then, the distribution is being estimated at each spatial point within the targeted range. To address such problems, a number of techniques have been suggested, which are usually based on interpolation (Table 2.1).

The proposed approach originates from the observation that the occurrence of worthwhile content is associated with a geological structure which contains a distinct (compared to the average area) concentration of minerals. Specifically, it is proposed to predict and reconstruct the orebody, the shape and distribution of which should give the best possible explanation of the taken samples.

The core idea is that having determined a notably high gold concentration $y \in \mathbb{R}_+$ at a location $x \in \mathcal{O} \subset \mathbb{R}^3$, the conditioned probability distribution for the gold concentration (potentially valuable) at any other point $x' \in \mathcal{O} \subset \mathbb{R}^3$ depends on the Euclidean distance $\|x - x'\|$ to the location and the concentration of measured gold. Although this dependence is difficult to quantify, it is assumed that the higher y and the smaller the distance $\|x - x'\|$, higher chance to find high concentration at point x'.

This idea is represented in terms of a kernel density.

$$\kappa : \mathbb{R}^3 \to \mathbb{R}_+, \quad x \mapsto \kappa(x), \quad \int_{\mathbb{R}^3} k(x)dx = l.$$

with the assumption that given a location $l \in \mathcal{O}$, the function $y \mapsto k(x - l)$ represents a prime deposit (placer) concentrated around a point $l \in \mathcal{O}$.

This statement can be quantified as follows: given the occurrence of a prime gold deposit located at point $l \in \mathcal{O}$, the expectation of the averaged gold concentration within each rock mass volume $V \subset \mathcal{O}$ is represented by a random variable, where the (conditioned) expectation is given by

$$\int_V k(v - l)dv. \tag{5.14}$$

To reflect the expected concentration decay depends only on distance without taking into account any particular direction. It is assumed that

$$k(x) \quad \text{depends on Euclidean distance } \|x\| \text{ only and decreases with } \|x\|.$$

Suppose further that such deposit can be located at $p \in \mathbb{N}$ different points $(l_j)_{j=1}^p \subset \mathcal{O}$ and has different intensities $(\beta_j)_{j=1}^p$. To resemble an ore body by a superposition of such deposits, it is supposed that the expectation of the (averaged) concentration within each rock mass volume $V \subset \mathcal{O}$ is represented by a random variable whose (conditioned) expectation is given by

$$\sum_{j=1}^{p} \beta_j \int_V k(x - l_j)dx. \tag{5.15}$$

Finally, it is agreed to explicitly construct such random variables for each rock mass volume $V \subset \mathcal{O}$ in terms of

$$Y(V) = \frac{1}{\eta(V)} \sum_{j=1}^{p} \beta_i \int_V k(x - l_j)dx + \frac{\sigma}{\eta(V)}\varepsilon(V), \quad V \in \mathcal{B}(\mathcal{O}) \tag{5.16}$$

indexed by all Borel sets $\mathcal{B}(\mathcal{O})$ of $\mathcal{O}$. Here, $\sigma \in \mathbb{R}_+$ is a predetermined parameter and the family $\{\varepsilon(V) : V \in \mathcal{B}(\mathcal{O})\}$ represents a so-called *Gaussian white noise*. This is a family of Gaussian random variables which satisfy

$$\varepsilon(V) \sim N(0, \eta(V)), \quad \mathrm{Cov}(\varepsilon(V), \varepsilon(V')) = \eta(V \cap V'), \quad V, V' \in \mathcal{B}(\mathcal{O}), \tag{5.17}$$

where η denotes the Lebesgue measure[1].

Having agreed on the assumptions (5.16) and (5.17), determining the shape of the ore body is being relaxed to the estimation of the parameters

$$l = (l_j)_{j=1}^p \text{ and } \beta = (\beta_j)_{j=1}^p$$

given the observations $(y_i)_{i=1}^n$ of the gold concentration within a number $n \in \mathbb{N}$ of volume samples $(V_i)_{i=1}^n$.
Therefore, a generic method is the maximisation of the log-likelihood density

$$L_{l,\beta}(y) = \log\left(\Pi_{i=1}^n \exp(-\frac{\|y_i - \mu_i(\beta, l)\|^2}{2\sigma^2\eta(V_i)^{-1}})/\sqrt{2\pi\sigma^2\eta(V_i)^{-1}}\right). \tag{5.18}$$

Here the numbers $(\mu_i(\beta, l))_{i=1}^n$ is the expectation of gold concentration within the sample volume V_i under assumption of deposit locations $l = (l_j)_{j=1}^p$ with intensities $\beta = (\beta_j)_{j=1}^p$

$$\mu_i(\beta, l) = \sum_{j=1}^{p} \beta_j \int_{V_i} k(x - l_j)dx, \qquad i = 1, \ldots, n. \tag{5.19}$$

Following the above maximum-likelihood principle, a maximum of the function is determined.

$$(\mathbb{R}^3)^p \times \mathbb{R}^p \to \mathbb{R}_+, \quad (l, \beta) \mapsto L_{l,\beta}(y) \tag{5.20}$$

However, this approach has a number of significant drawbacks:

- Since the log-likelihood function is not convex, the dimension $4p$ of its arguments is high.
- There are numerical issues determining its global maximum.
- The choice of the number $p \in \mathbb{N}$ of prime deposits is arbitrary and there is no obvious way to determine this number exactly. Choosing p as too low results in a poor model, whereas if p is too high, then a large number of parameters may cause over-fitting and deteriorate the predictive performance of the model.

[1] URL http://mathworld.wolfram.com/LebesgueMeasure.html

5.3 Proposed Design

In this section, a different approach is suggested. The idea is based on advances in high-dimensional statistics and ℓ_1 penalties. Utilising these tools, the described above obstacles can be overridden. To increase the chance of the problem solvability, from an abstract perspective, a moderate-dimensional non-convex problem is reformulated into an extremely high-dimensional convex problem. Furthermore, a built-in penalisation of parameters with the focus on the minimisation of the generalisation errors leads to dimension reduction while maximising the model's predictive performance.

Having supposed that the volume samples $(V_i)_{i=1}^n$ have equal sizes

$$\eta(V_i) = \nu_0 \in \mathbb{R}_+ \quad i = 1, \dots, n \tag{5.21}$$

and assuming that the locations $(l_j)_{j=1}^p$ of all prime deposits are fixed, it can be observed that the log-likelihood function must be maximised with respect to intensity parameters $\beta = (\beta_j)_{j=1}^p$ only. However, the problem of finding a maximum of

$$\mathbb{R}^p \to \mathbb{R}, \quad \beta \mapsto L_{l,\beta}(y) \tag{5.22}$$

can be formulated with the the standard framework of linear models, where the unknown coefficients $\beta \in \mathbb{R}^p$ are estimated under the assumption that $Y = (y_i)_{i=1}^n$ are observations of the response variable $\mathbf{Y} = \mathbf{X}\beta + \varepsilon$ with independent, identically normally distributed random variables $\varepsilon = (\varepsilon_i)_{i=1}^n$ and a non-random design matrix

$$\mathbf{X} = (\mathbf{X}_{i,j})_{i,j=1}^p, \quad \mathbf{X}_{i,j} = \int_{V_i} k(x - l_j)dx, \quad i = 1, \dots, n, \;\; j = 1, \dots, p. \tag{5.23}$$

Note, that the centres $l = (l_j)_{j=1}^p$ of prime deposits are fixed. To address this issue, it is proposed that their total number p is sufficiently high, which allows a potential prime deposit to exist virtually at every location.

At the same time, it is suggested to use a penalised regression. Instead of solving the ordinary regression

$$\hat{\beta} = \text{argmin}_{\beta \in \mathbb{R}^p} \|Y - \mathbf{X}, \beta\|$$

an adequate formulation is addressed in the form of the LASSO (5.9) with l_1-norm (5.10) on the coefficients

$$\hat{\beta}(\lambda) = \text{argmin} \left(\|Y - \mathbf{X}\beta\|^2 + \underbrace{\lambda \|\beta\|_1}_{\text{penalization}} \right)$$

with l_1-norm on the coefficients

$$\|\beta\|_1 = \sum_{j=1}^{p} |\beta_j|, \quad \beta = (\beta_j)_{j=1}^p \in \mathbb{R}^p.$$

This methodology has several advantages. It allows a view from the perspective of HDS, whose methods are designed to solve linear models under the assumptions that the number of explanatory variables exceeds (by magnitudes) the number of observations, having in mind that only a few (not knowing which) of the explanatory variables have non-zero coefficients.

Furthermore, the penalised regression in the LASSO formulation performs an automated variable selection, by determining those non-zero coefficients which maximise the predictive

performance through an estimation of the generalisation error via `nfold=10` (set by default) cross-validation.

Note that this method is appropriate the proposed statistical estimation since it is supposed that at virtually every location there could be a prime deposit of unknown intensity. Realistically, almost all intensities will vanish, but it is not known which deposits contribute to non-zero intensity.

Finally, let us emphasise that the goal of a traditional model is not to fit/smooth a given sample of observations, but to infer a mineral concentration in the entire domain with enhanced predictive ability while taking into account all relevant physical considerations (for instance, the mineral clustering described in terms of the ore body concept). These goals are achievable by the proposed methodology.

5.4 Experiment

This section presents an extract from a real-world blasthole data, also known as “Marvin", which was collected at the production phase during the grade control program. Marvin deposit was discovered in late 1996. Geologically, it relates to porphyry deposits and represents an important resource and the dominant source of copper and gold. It is located ≈100 km North of Orange and contains two major elements as a source of revenue: copper Cu, gold Au and triuranium octoxide U_3O_8 as sub-product.

5.4.1 Geological settings

In 2004 the Marvin mine carried out a grade control exercise to define ore and waste using gold samples obtained from vertical blast holes in order to delineate the gold-bearing ore for the next four benches to be dug from the RL455.

The 119 horizontally spaced as 5×2.5 metres blastholes ≈10m long were drilled from the surface by rotary rigs (RC method) and sampled in 2.5m vertical intervals as shown in Figure 5.2. Geometrical limits of the blasthole drilling grid are presented in Table 5.1.

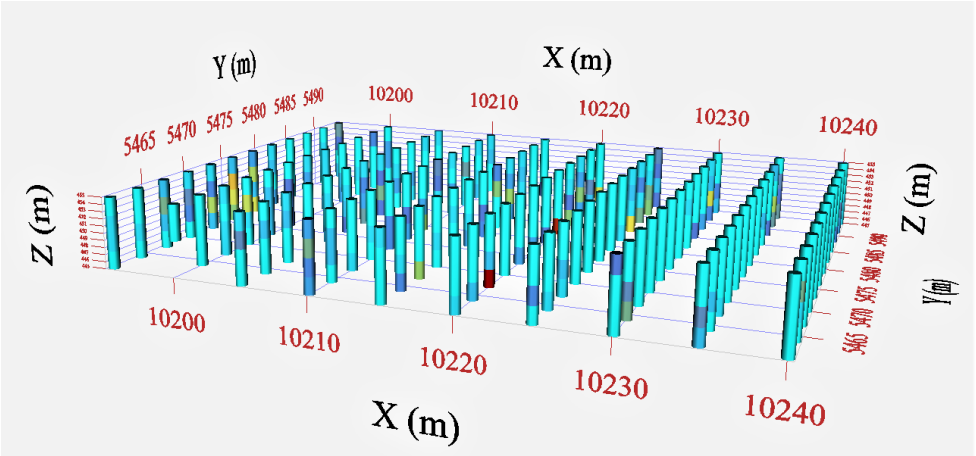

Figure 5.2. Blasthole grid: number of developed blastholes=119, ∠Dip=90°(vertical); horizontal spacing=5m×2.5m, samples=5975; vertical sampling intervals=2.5m; spatial limits: [X=10195m×10240m]; [Y=5461m×5494.00m]; [Z=10.1m×10.9m]

.

On-site analysis revealed the presence of multiple clusters of gold close to the vertical blast holes =10 metres long each. No accurate engineering assumptions regarding gold distribution were made at the time of the grade control program due to the lack of reliable data in between the 2.5 metres intervals.

The existence of gold within the flitch (the area is 245m×33m, or 8085m^2) provided an opportunity for a test the proposed LASSO/`glmnet` system for predicting the distribution of gold at unsampled locations (i.e between blastholes and the vertical sampling intervals as shown in Figure 5.12). Using actual drilling data shown in Figure 5.2 and Tables 5.1; 5.2 coordinates and grades were assigned to 5975 valid samples of gold.

5.4.2 Blasthole sampling data

The used parameters have been as follows: blasthole ID, $x; y; z$ Cartesian coordinates, the positions of drilling collars on surface at level 455 m, the length of the sampled sections of each blasthole from/to, the true depth (m) and gold grade in g/tonne.

Lithology, stratigraphy and lithochemistry have not been considered in this experiments because these required the application of more complex AI techniques. The geometry and descriptive statistics for Marvin blasthole assay is summarised in Tables 5.1–5.3.

The presence of heterogeneity in Marvin indicates that even at best scenario, there will be discrepancy between prediction and reality.

Taking into account standard processing cycle as: blasting → excavation → transport → crushing → grinding → separation → concentration → smelting, any accurate reconciliation between the mined gold grades and predicted gold distribution becomes problematic.

Table 5.1. Bench blasthole array geometry

	X min	X max	Y min	Y max	Z min	Z max	Field
Geometry	10195.0m	10240.0m	5461.0m	5494.0m	445.0m	455.0m	25523m^3

Table 5.2. Statistics reporting on numerical variables by quantiles in g/t

Total	Min	Max	Mean	SD	σ^2	Q:0.05	Q:0.10	Q:0.25	Q:0.75	Q:0.95
5975	0.01	19.22	1.97	2.95	8.71g/t^2	0.04	0.07	0.23	2.48	8.73

Basic description of Marvin data is summarised in Table 5.2. The data suggests the existence of a certain trend in terms of spatial development of grades towards East. This trend is evident in Figure 4.4 and Figure 4.5 and can be also identified from the horizontal semivariance shown in Figure 5.3.

Table 5.3. Statistics reporting on validated numerical variables by quartiles dividing the range of a probability distribution into continuous sectors in Au g/t

Sample	Min	Max	Mean	St.Dev	Var	Q0.05	Q0.10	Q0.25	Q0.75	Q0.95
5975	0.01g/t	19.2g/t	1.97	2.9g/t	8.6	0.04g/t	0.07g/t	0.2g/t	2.5g/t	8.7g/t

The shape of the curve in the variogram in Figure 5.3 supports the idea that the distribution of gold within Marvin is not that erratic. Low grades are present in all directions and high grades are not present in all directions. The variorgam settings are set as follows:

i. Omni-directional in horizontal plane with `lag=1m`
ii. Domain=whole field. The chosen partition=whole field
iii. The Cut-off grades set: 0; 2; 4; 6; 8; 10; 12; 14; 16; 18 g/t.
iv. Compositing Length=1.25m with minimum length=0.1m.
v. Maximal distance=10.5m; slicing height=0.1m

Since the average length of vertical blast holes has been =10 metres, the vertical variogram is much harder to accurately fit and interpret. Therefore, pre-processing estimation needs to be performed. To do this, a spherical model is applied with no-nugget effect involved. Figure 5.4 presents experimental curve of the vertical spherical variogram:

i. The experimental variance= 8.59213 $(g/t)^2$
ii. Vertical Lag = 0.6m.
iii. Maximum distance=6.24m, slicing radius=0.5m
iv. Vertical Sill = 3.95875 $(g/t)^2$

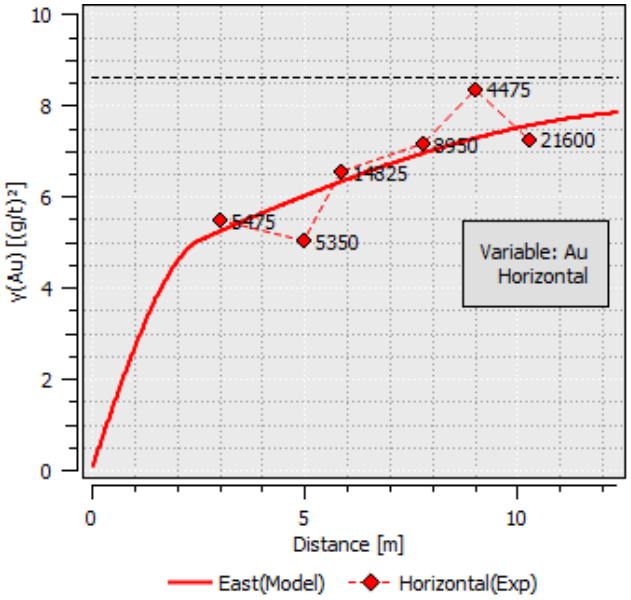

Figure 5.3. Marvin's horizontal spherical variogram.

Next, a swath variational analysis is applied which compares between sample points and estimated values to identify a bias towards underestimation or overestimation and the presence of any smoothing in the results.

The swath plots shown in Figures 5.5–5.7 show the average grade for the blocks in the swath along with the averaged sample values in the swath. Figures illustrate the averaged values against the averaged block grades for a series of swaths slicing through the block model in the North-East direction.

A swath plot shown in Figure 5.5 illustrates that the apparent grade increases considerably with an increase of depth. Figure 5.6 detects a shift of higher grades towards the North.

Figure 5.7 shows the presence of a stable trend in terms of a consistent increase of grades with the depth. It appears that the plots in Figures 5.5–5.11 provide enough evidence for the existence of the trend in increasing high grades with decreasing of RL.

Kriging estimates in Figures 5.8–5.11 indicate that the development of underlying flitch can be viewed as (potentially) economically profitable.

To explore the distribution, four gold favourability maps are created using kriging. The realisations shown in Figures 5.9–5.10 indicate that depending on the depth, the flitch 10m may contain gold concentrations $\geq$19.00 ppm. or 19 Au g/t. The cut-off grade

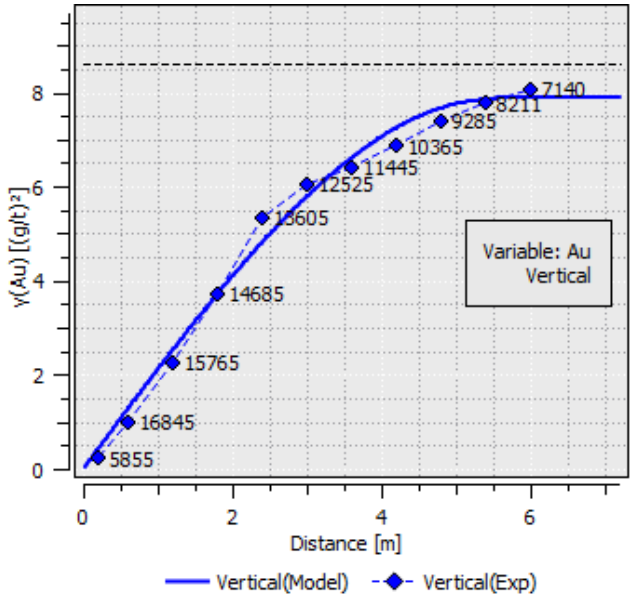

Figure 5.4. Marvin's vertical spherical variogram (10m) shows that gold variability is not increasing with the increase of distance.

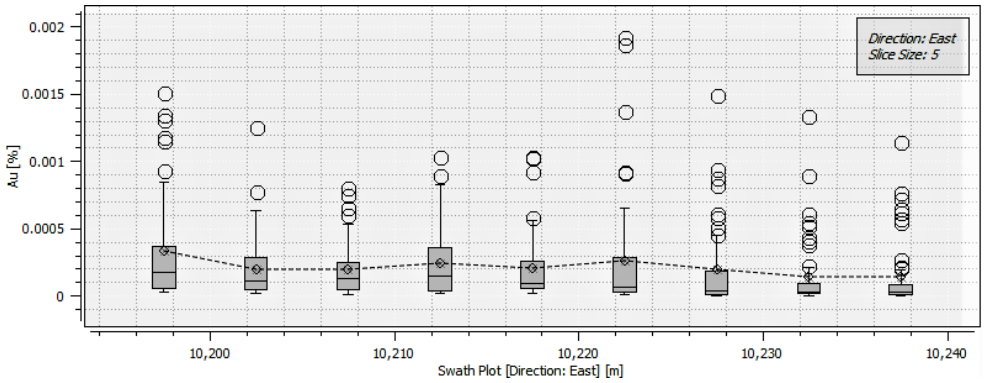

Figure 5.5. Swath plot reveals considerable increase of high grades towards the East. Axis Y shows associated Au g/t values. Declustered values are marked as black open circles above the boxes

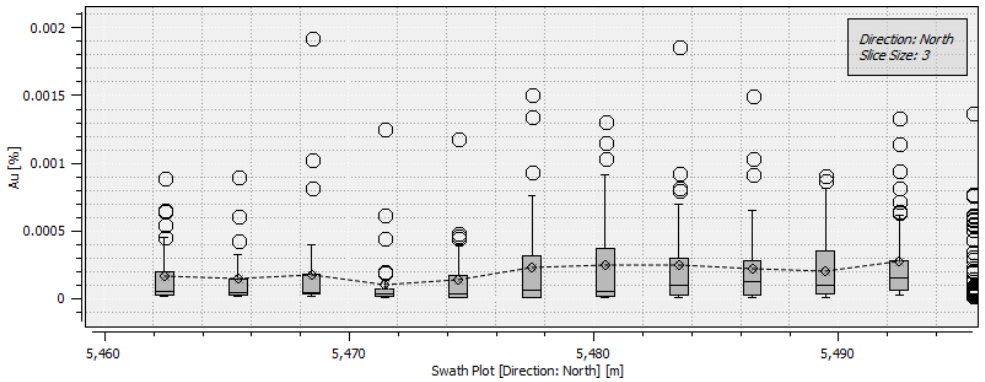

Figure 5.6. Swath plot reveals the increase of high grades in the Northern direction

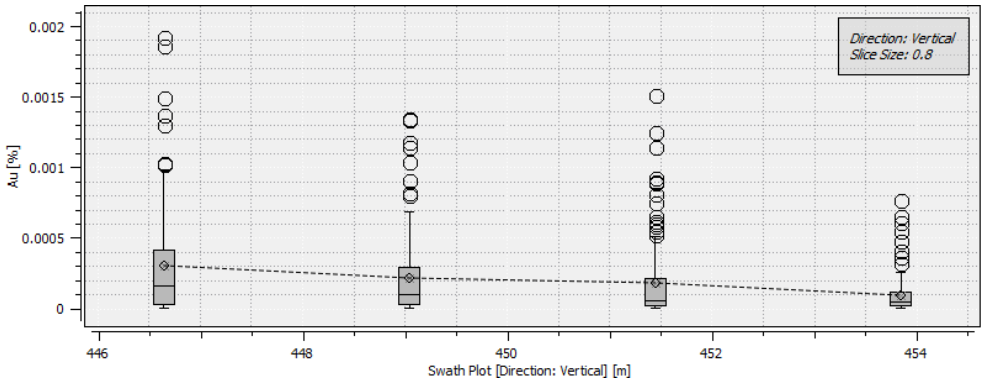

Figure 5.7. Swath plot showing a developing trend in increasing gold grades with depth and the existence of high localised grades below RL445.

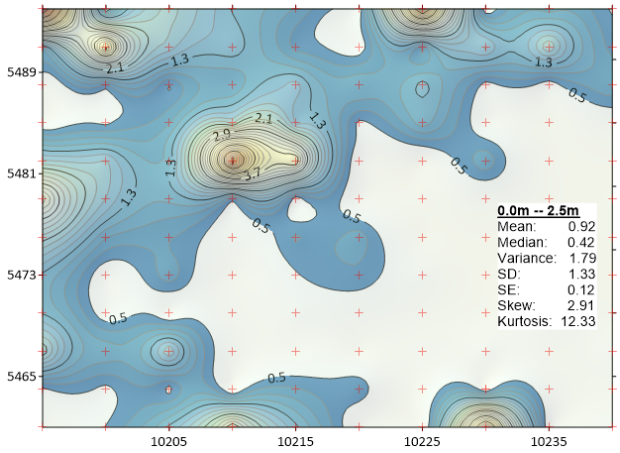

Figure 5.8. Distribution of ore in the form of isopath from RL455 to RL452.5 obtained from kriging. RC sampling interval (length) 0.0m–2.5m., collars are marked as red crosses.

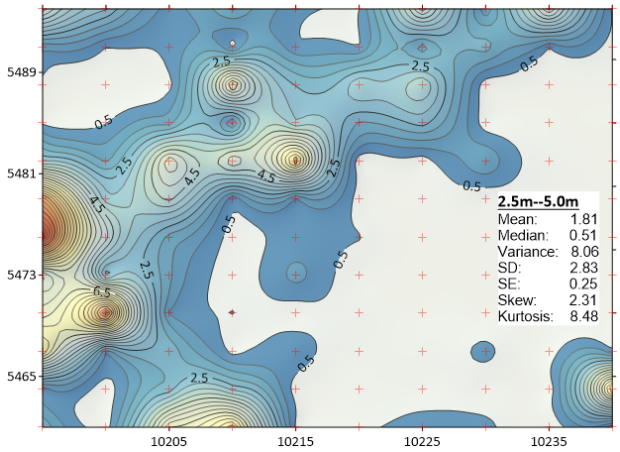

Figure 5.9. Distribution of ore from RL452.5 to RL450, kriging, sampling interval 2.5m–5.0m. The trend in higher grades is evident.

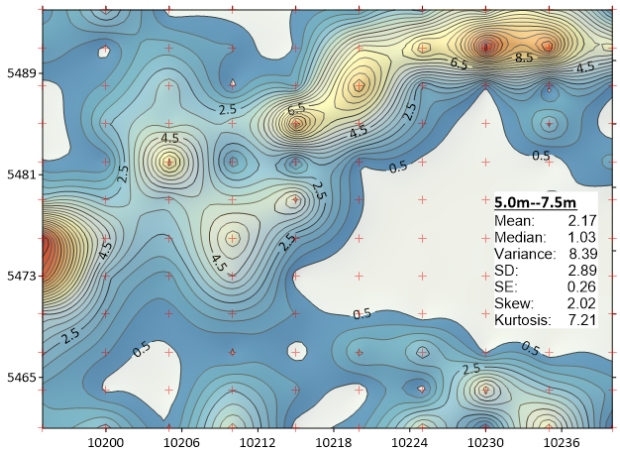

Figure 5.10. Distribution of ore from RL450 to RL447.5, kriging, sampling interval 5.0–7.5m

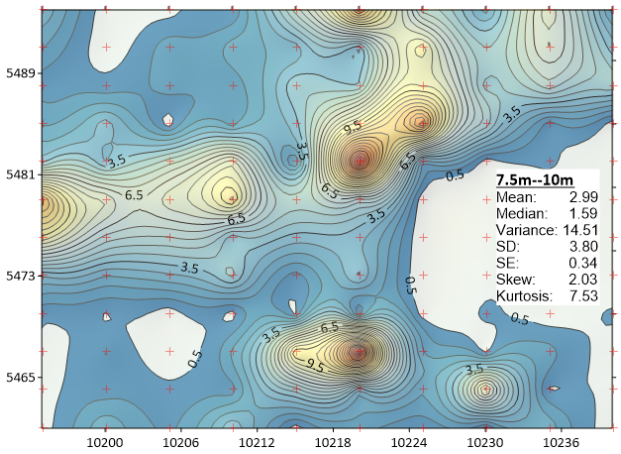

Figure 5.11. Distribution of ore from RL447.5 to RL445, sampling interval 7.5–10m, consistent increase of grades with the depth.

=0.5 g/t has been applied to all models by default to display the contrast between low and higher grades.

Low grades are present at the locality 5470N/10230E. Relatively high grades may occur at 5485N/10210E. The cross-sections from kriging indicate the presence of developing trend in increasing grades with depth. The extraction and processing of rock mass of underlying flitch can be economic.

However, the Eastern region shows a very low reading. Realising that the grade control data has been derived from only a part of the mine, gold concentrations may be different elsewhere.

From an engineering point of view, the depth below RL445 should be worthy for close attention because the underlying horizontal levels are likely to have larger amount of gold. Comparing four cross-sections side-by-side (Figures 5.8–5.11), one may conclude that the distribution is not gradational. The sampling intervals 2.5 metres are too large to make an assumption on gold potential.

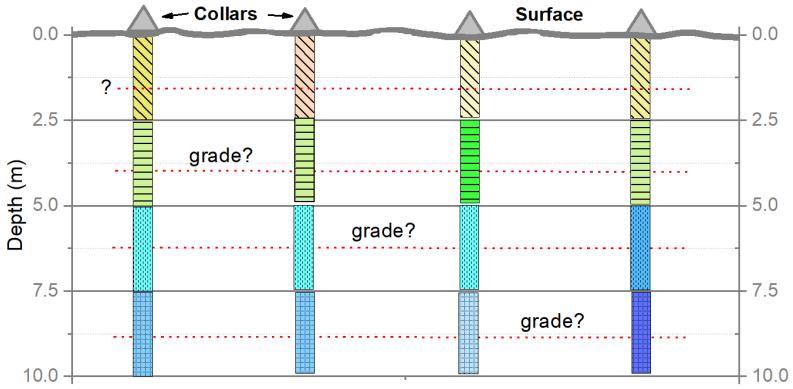

Figure 5.12. Figure shows blastholes developed by RC. Samples taken in 2.5 metres vertical interval. The levels to be predicted are marked as red dotted lines.

Marvin is porphyry-related deposit and hosts gold and copper. The blast holes were developed with reverse circulation drilling (RC). This type of drilling produces small size rock chips. Then, the chips taken in intervals 2.5 metres (see Figure 5.12) were brought to the surface by compressed air and then collected in bags.

Sampling in 2,5 metre intervals (Figure 5.12) provides enough information to define copper. However, the length 2.5 is not enough for accurately predicting the gold.

The knowledge of the variation of gold grades along 2.5m intervals is a lot more commercially interesting than just knowing the averaged grade of a 2.5m thick level. Therefore, the task is to predict “what is going in between these 2.5m intervals?".

1. Gold content is represented by several, potentially narrow structures.
2. Exploratory drilling cannot meaningfully take place.
3. The edges of the structures are poorly defined, making their definition at the time of extraction difficult.
4. Economic cut-off grade is unknown. Lowering the cut-off grade will result in a steady increase in the tonnage to be extracted and milled.
5. If cut-off grade increased, the impact of dilution may be large.

Marvin is the case when it is unclear which factor is the most important in defining the distribution - economic cut-off grade or the shapes. In any scenario, the knowledge

of the content between the sampled intervals is highly desirable.

An alternate way is to predict ore outlines by in between the sampling intervals, enabling gold projection from one horizontal section to another for presenting better defined, high-priority structures before actual extraction commence. Here is the engineer's dilemma

1. how to predict the distribution of gold in between 2.5m intervals?
2. how to define the ore shapes, positions and boundaries for efficient production?

The lines marked in red dots in Figure 5.12 are the levels at which the LASSO will predict the concentrations and gold content in g/m^3 of rock mass.

5.5 Method Implementation

In this section the estimation of model parameters through iterative application of the ℓ_1-penalty is implemented. The source code with comments is presented in Appendix C.

Predicting gold variability within Marvin is a moderate-dimensional nonconvex problem. To make it solvable, the problem is moderately reformulated into a high-dimensional convex problem.

a) At the beginning, a number ($p = 700$) of deposit centres is randomly scattered around the area to be investigated. Each kernel is represented by a normal density whose mean equals the corresponding deposit centre, with the same variance.

b) The coefficients are estimated using standard LASSO method with 10-fold cross validation for penalisation parameter selection.

c) Each deposit centre, which has non-zero coefficients is disturbed (its kernel variance and kernel centre are changed randomly). The deposit with zero coefficients are removed and those obtained via disturbance are added to the potential set of deposits.

d) The steps b) and c) are repeated (five to seven times, lines 65–68 in source code in Appendix C). The results are depicted in Figures 5.13–5.16.

5.6 Results

It is fair to suggest that the blasted flitch is likely to be mined in one 10 metre lift with truck-and-shovel method. In this scenario, the cost of extraction is high. The cross-sections from kriging and LASSO do not provide enough evidence that the gold concentrations may not be considered as anomalous.

The flitch is rather marginal and depends on the efficiency of ore treatment. Logically, low-cost processing by simple gravity treatment may cover the expenditure of the development. Importantly, the depth below RL445 (the underlying flitch) should be worthy for a more closed attention because the underlying levels are likely to have a larger amount of gold. The adjacent flitch at the Easterly locality is suggested for careful grade control.

The results of the extraction of "non-zero" coefficients and deleting irrelevant values are shown in the form of horizontal cross sections in Figures 5.13–5.16, which illustrate the overall predictive performance of ℓ_1 estimator for matrix completion and show gold variation per m^3 of rock mass.

The favourability maps obtained from multiple program runs appropriately represent the boundaries between barren rock and the ore closely associated with gold.

It is apparent from Figures 5.13–5.16 that gold distribution is erratic. It clearly indicates low continuity with higher gold grade potential at the northwestern locality.

The maps obtain from LASSO and kriging indicate that gold in the localised and disseminated form exists at all levels (1.875m; 3.75m; 5.625m and 7.5 metres) below the drilling collars. There is an apparent trend of increasing middle/ high grades with depth.

Relatively intense Au concentrations appear in the vicinity of locality 5475N/ 10190E which is nearly the maximum anomaly contrast. The models obtained from kriging and LASSO clearly indicate that

1. The extraction and processing of the underlying flitch can be economically profitable.
2. The flitch adjacent to the Eastern part may have some economic potential.

Commands `dat`, `centres`, `obs`, `resp`, the `glmnet` fit and the extraction of nonzero coefficients are available via the environment and R-console during running the code (Appendix C). Each R-code execution generates set of centres C1, C2,.., C500, etc. as shown in Listings 5.1.

Then, `nfold=10` cross-validation is applied, irrelevant values are removed and "nonzero" coefficients are extracted. It has to be noted that multiple runs (the loop from line 79 to 181) provide the same set of "non-zero" coefficients which form the basis for predictive mapping.

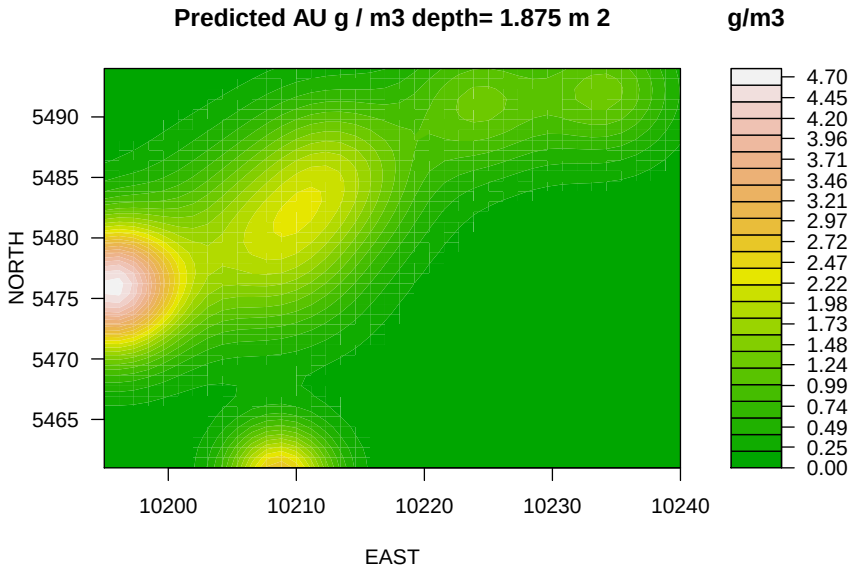

Figure 5.13. Predicted production target favorability map in Au g/m^3 at depth =1.875m, where high grades contrasting low grades. Model indicates the presence of a developing trend towards North-East direction with overall increase of grade and the existence of internal dilution in SE and E directions assigned the colour green. Note: "2" in the upper title =warm start.

Listing 5.1. Set of computed centres

```
 [1] "C3"   "C8"   "C30"  "C74"  "C100" "C110" "C168" "C173" "C200" "C232"
[11] "C236" "C260" "C275" "C314" "C338" "C339" "C346" "C462" "C472" "C494"
[21] "C558" "C578" "C647" "C666" "C687" "C695"
 [1] 3.000000 3.000000 3.000000 3.000000 3.000000 3.000000 3.000000
 [8] 3.000000 3.000000 3.000000 3.000000 3.000000 3.127384 2.800745
[15] 3.144847 2.716084 3.265503 3.264992 2.844915 3.252711 3.118251
[22] 3.161754 3.212044 3.273708 3.285459
 [1] "C1"  "C4"  "C7"  "C8"  "C9"  "C10" "C12" "C15" "C19" "C21" "C23"
[12] "C24" "C28" "C29" "C37" "C39" "C40" "C41" "C42" "C43" "C46" "C47"
[23] "C48" "C51" "C52"
 [1] 3.000000 3.000000 3.000000 3.000000 3.000000 3.000000 3.000000
 [8] 3.000000 3.000000 3.127384 2.800745 3.144847 3.265503 3.264992
[15] 3.252711 3.161754 3.212044 3.273708 3.285459 2.813359 2.812365
[22] 3.221475 2.910224 2.757218 2.856110 3.282293 3.276477 3.542000
 [1] "C1"  "C2"  "C5"  "C7"  "C8"  "C9"  "C10" "C11" "C12" "C13" "C14"
[12] "C15" "C17" "C18" "C20" "C22" "C23" "C24" "C25" "C27" "C28" "C29"
[23] "C31" "C37" "C41" "C46" "C48" "C50"
 [1] 3.000000 3.000000 3.000000 3.000000 3.000000 3.000000 3.000000
 [8] 3.000000 3.000000 3.127384 2.800745 3.144847 3.265503 3.264992
[15] 3.252711 3.161754 3.273708 3.285459 2.813359 2.812365 3.221475
[22] 2.910224 2.757218 2.856110 3.282293 3.276477
 [1] "C1"  "C2"  "C3"  "C4"  "C5"  "C6"  "C7"  "C8"  "C9"  "C10" "C11"
[12] "C12" "C13" "C14" "C15" "C16" "C18" "C19" "C20" "C21" "C22" "C23"
[23] "C24" "C25" "C26" "C27"
 [1] 3.000000 3.000000 3.000000 3.000000 3.000000 3.000000 3.000000
 [8] 3.000000 3.000000 3.127384 2.800745 3.144847 3.265503 3.264992
[15] 3.252711 3.161754 3.273708 3.285459 2.813359 2.812365 3.221475
[22] 2.910224 2.757218 2.856110 3.282293 3.276477
 [1] "C1"  "C2"  "C3"  "C4"  "C5"  "C6"  "C7"  "C8"  "C9"  "C10" "C11"
[12] "C12" "C13" "C14" "C15" "C16" "C17" "C18" "C19" "C20" "C21" "C22"
[23] "C23" "C24" "C25" "C26"
 [1] 3.000000 3.000000 3.000000 3.000000 3.000000 3.000000 3.000000
 [8] 3.000000 3.000000 3.127384 2.800745 3.144847 3.265503 3.264992
[15] 3.252711 3.161754 3.273708 3.285459 2.813359 2.812365 3.221475
[22] 2.910224 2.757218 2.856110 3.282293 3.276477
 [1] "C1"  "C2"  "C3"  "C4"  "C5"  "C6"  "C7"  "C8"  "C9"  "C10" "C11"
[12] "C12" "C13" "C14" "C15" "C16" "C17" "C18" "C19" "C20" "C21" "C22"
[23] "C23" "C24" "C25" "C26"
 [1] 3.000000 3.000000 3.000000 3.000000 3.000000 3.000000 3.000000
 [8] 3.000000 3.000000 3.127384 2.800745 3.144847 3.265503 3.264992
[15] 3.252711 3.161754 3.273708 3.285459 2.813359 2.812365 3.221475
[22] 2.910224 2.757218 2.856110 3.282293 3.276477
 [1] "C1"  "C2"  "C3"  "C4"  "C5"  "C6"  "C7"  "C8"  "C9"  "C10" "C11"
[12] "C12" "C13" "C14" "C15" "C16" "C17" "C18" "C19" "C20" "C21" "C22"
[23] "C23" "C24" "C25" "C26"
```

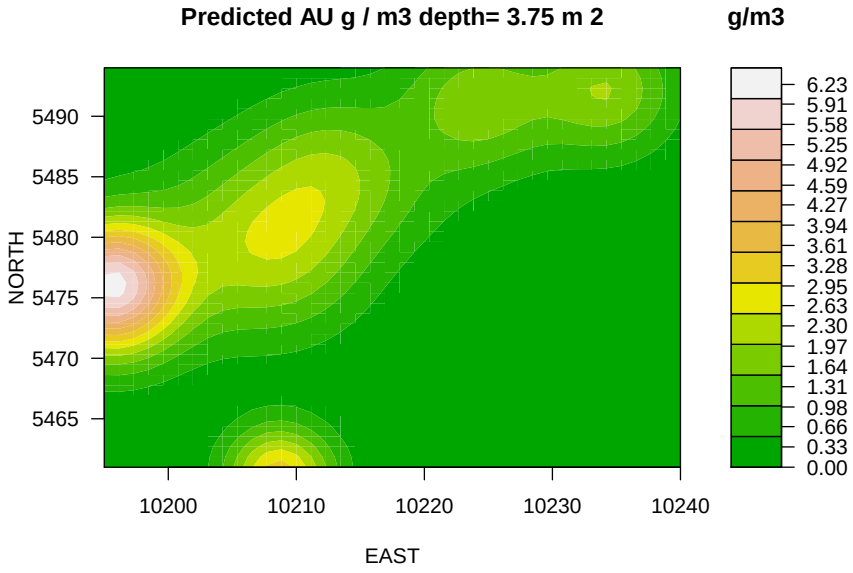

Figure 5.14. Production target favorability map at depth =3.75m in Au g/m^3. Stable trend identified, model indicates significant geometrical changes accompanied by the mean and median displacements.

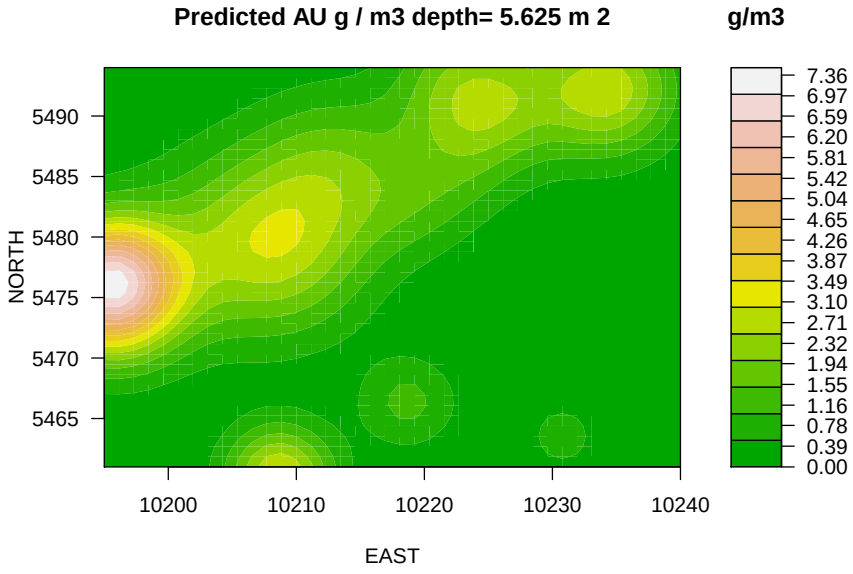

Figure 5.15. A realisation of cross-section at depth =5.625m. Model indicates consistent increase of Au grades with the depth.

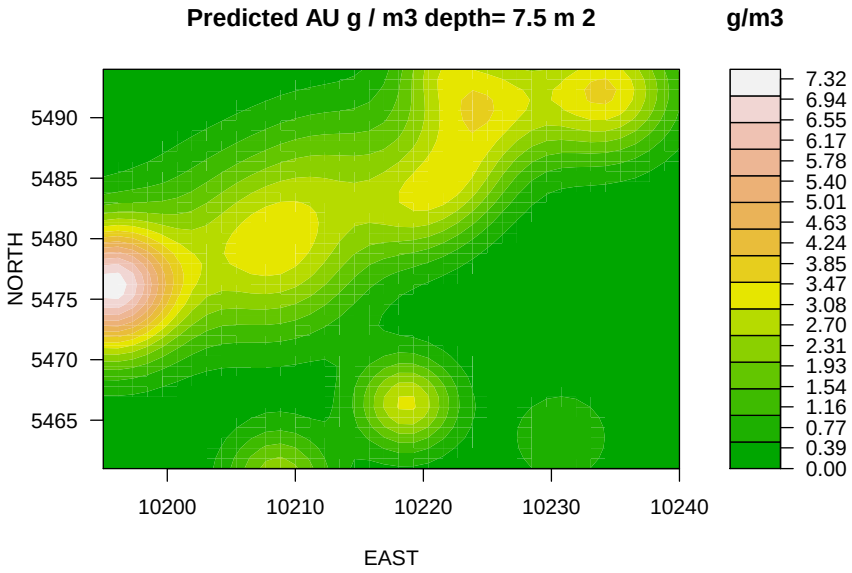

Figure 5.16. A realisation of cross-section at depth =7.5 metres. The apparent Au grade increases considerably with the increase of depth. Contours illustrate a preferential trend in the northeasterly direction. The extraction of the underlying level, i.e. below RL445 can be economically profitable. The adjacent flitch at the Easterly locality requires for careful grade control.

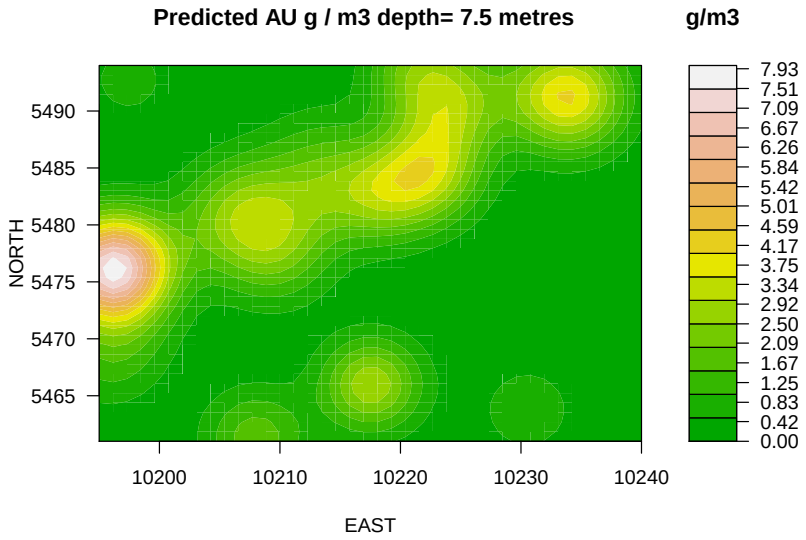

Figure 5.17. The effect of randomly removing subsets of the data for cross-validation. The cross-section at depth=7.5m in PDF obtained from second program restart from line=78. The same graphical output from one run (sf. Figure 5.16) to the next is unlikely to happen. Adobe starts vectorisation from the first emerged centre.

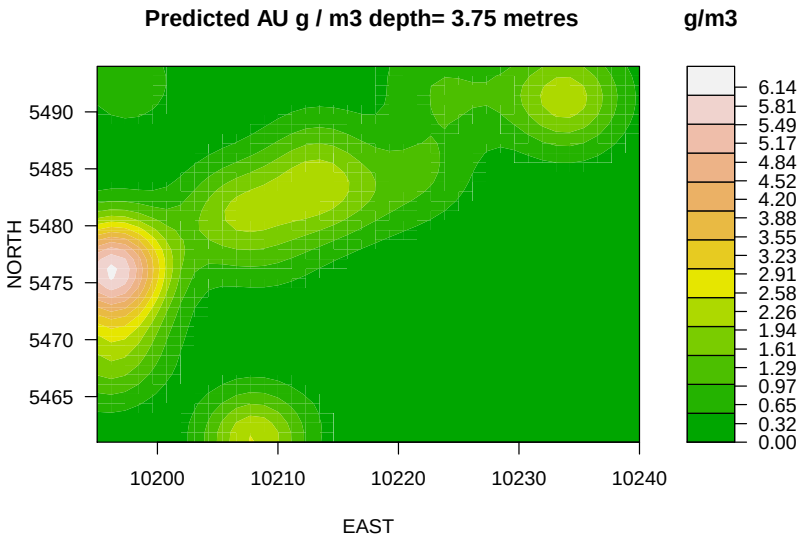

Figure 5.18. The effect of "randomness" at depth =3.75m generated from 2nd program restart from line=78 (sf. Figure 5.14).

5.7 Conclusion

The undertaken in Chapters 2 and 5 surveys revealed that

1. Developing predictive tools that are both user-friendly and efficient is an extremely challenging task.
2. Since the nature of gold distribution within the flitch is not well defined (too large sampling intervals), it is not possible to create the best `global` method for making predictions, because all existing methods have limitations to a greater or lesser extent.
3. In general, as the flexibility of a method increases, its interpretability decreases.

The most important aspects that need to be considered for design

a. A method should reflect the nature of ore extraction.

b. Realistic applicability (not just theoretical) of the developed method to real-world scenarios.

The existence of heterogeneity and the lack of reliable data between RC samples inspired the idea for developing a `convex` optimisation method in which the prediction would not be derived from variograms and be capable of solving nonconvex tasks, simultaneously providing globally optimal and statistically-valid solutions. This idea got its continuation in a formal attempt to develop a method based on penalisation principle.

5.7.1 Indirect suggestions found in the statistical literature

Statistical literature suggested that a regularisation by penalty:

- for Akaike information criterion (AIC), one needs to exclude insignificant variables. For instance, the penalty

$$\lambda\|\beta\|_0 \qquad (\|\beta\|_0 \text{ number on non-zero entries}),$$

is computationally demanding and in high dimensional settings is infeasible. For example, Velasco & González-Salazar (2019) suggest that AIC should not be used if R-users interested in prediction more than explanation.

- the l_2 norm $\|\beta\|_2$ does not exclude variables.
- the l_1 norm $\|\beta\|_1$ is working
- for elastic net: use a penalisation $\lambda(\alpha\|\beta\|_1 + (1-\alpha)\|\beta\|_2)$ (for instance, with $\alpha = 1/2$).

A number of studies (section 5.1) suggest that the ℓ_1-norm may provide a computationally attractive, theoretically justified form of regularisation. The LASSO method was admitted as more restrictive in estimating the coefficients because it sets a number of them to exactly zero. Due to built-in strict parameter penalisation, the LASSO methodology appears to be less sensitive to the presence of heterogeneity in data. The built-in focus on the predictive performance of the LASSO can provide accurate spatial analysis.

5.7.2 Design considerations

The idea was that a global predictive model can be be fitted via penalised maximum likelihood. It was proposed that the regularisation path needs to be computed by `glmnet` R-package for the LASSO penalty at a grid of values for the regularisation parameter `lambda`.

To reduce computational complexity, a convex nonlinear programming optimisation code with nonconvex objective function has been executed. Having only a small number of "non-zero" parameters the code selects predictors and shrinks their coefficients toward zero relative to the least-squares estimates.

"Marvin" dataset provided a context to explore optimisation problem of gold distribution within the deposit. The performance of the ℓ_1-penalty LASSO `glmnet` system has been tested on this data. The predicted gold distribution between the sampled intervals has been visualised with R as the distribution of gold content per m^3 of rock mass.

5.7.3 Observations

An interesting phenomenon can be observed from graphical output. Because the fitting is based on randomly removing subsets of the data for cross-validation, the same graphical output from one run to the next is unlikely to happen.

This effect can be seen by comparing Figure 5.16 with Figure 5.17 and Figure 5.14 with Figure 5.18 rendered in PDF. Figures reflect the distributions of the same cross-sections at 7.5 and 3.75 metres depth but obtained from different program restarts from line 78 (Appendix C).

Because of randomness, the centres are not plotted simultaneously. Adobe starts rendering from the first emerged centre. As a consequence, there are some insignificant graphical discrepancies between the models obtained from program restarts from line 78. The application of .png or .jpg partially resolves that problem and the output become graphically similar. Despite a lower resolution, .png and .jpg provide better graphical consistency from all subsequent runs.

This phenomena relates to other cross-validated with `10fold` function estimators (kriging, IDW). Ten `R`-runs will result in ten insignificantly different renders.

5.7.4 LASSO versus kriging - which method is more predictive?

The separate application of either kriging or ℓ_1 estimations does not guarantee the expected realistic accuracy.

The methods are not competitive, they just look at the prediction from different angles of view. Describing similar matter, they represent different statistical concepts and utilise different metrics. For example, both methods indicate the presence of trend (which is not weak) in increasing grades with depth. A firm confirmation that LASSO prevails over the kriging in terms of predictive accuracy has not been established.

From an engineering perspective, the rate of change of gold grades in the "kriged" models, especially between RL455–RL447.5 (Figures 5.9–5.11 display sharp changes in grades and ore/waste contacts) is too rapid and the borders of the extraction targets are not well defined.

From the other hand, not much evidence available to suggest that the flitch is anomalous. The shape of dilution (barren rock) is more critical during plant operation than the location of insignificant concentrations of the gold. The Cu/Au deposits such as "Marvin" are usually extracted by blocks, in a 10m lift by a shovel with bucket capacity >20 m^3. That means the impact of the excessive amount of dilution may be large. From a production supervisor's viewpoint, the borders rendered by LASSO and a g/m^3 metrics are more convenient to utilise during levelling with on-board GPS.

It is important to realise that Method 2 is not a more predictive than kriging. It only has a built-in tendency to estimate coefficients as zero - the larger the penalty coefficient $log\lambda$ is, the greater is that tendency. Method 2 does not provide significantly more realistic representation of the distribution. However, ore/ waste mineable shapes and the extraction targets are better presented and understood.

It needs to be admitted, however, that due to strict parameter penalisation, Method 2 is less susceptible to the erratic distribution and the presence of heterogeneity. Since the ℓ_1 penalties were convex, the assumed sparsity in the available data lead to significant computational efficiency. From a practical point of view, the patterns plotted by ML better represent extraction rationality. Method 2 does not require an end-user to be familiar with basic programming to execute the prediction.

5.7.5 Computational efficiency

The program (Appendix C) was stable during all executions and demonstrated the ability to compute full regularisation paths. Method 2 fulfils the task with minimum amount of time, consuming a low amount of RAM and did not overload the memory. The maximal time taken for running LASSO/ `glmnet` from cold start was 1 min 50 sec.

It has also been observed that the the ℓ_1-norm is a less flexible approach than some other regression methods and more interpretable than linear regression because in the final model the response variable was only related to a small subset of the predictors.

Method 2 highlights: The most important aspects of the prediction with the application of the LASSO-regularised classifier for finding the `global` optimum have been assembled as experienced through our research and work in the mining industry.

- The concept of feature selection is found as potentially applicable to evaluation. It has been proposed that the prediction accuracy can be improved with ℓ_1-norm by setting regression coefficients to zero.
- The problem of predicting the distribution of gold has been tackled by reformulating it into a high-dimensional convex optimisation problem which has been solved with the application of the ℓ_1 penalised regression.
- The "non-zero" coefficients were extracted using ℓ_1 penalisation with built in `nfold=10` cross validation. The regularisation path has been computed for the LASSO penalty at a grid of values for the regularisation parameter `lambda`.
- Each program run provides one global optimum in the form of gold concentrations in g/m^3.
- Results from both LASSO and kriging indicate the presence of the trend towards North-East direction with overall increase of grade with depth and the existence of internal dilution in SE direction. Both methods suggest the existence of localised occurrence of middle/ high localised grades withing the underlying flitch, the development of which can be economically profitable.
- Method 2 is fast fulfils the task with minimum amount of time consuming, low amount of RAM and does not overload the memory. The average time taken for running the system have been from 110 sec to 160 sec.
- Standard implementation of `glmnet` can deal efficiently with spatial problems.

Chapter 6

Discussion

The problem of predicting gold is as old as mining. All known methods based on statistical models and mathematical concepts do not guarantee exact prediction and realistic representation of the gold that exists in deposits.

The developed methods 1 and 2 have been the result of research, engineering design and complex decision-making process in which mathematics, high-dimensional statistics, programming, 3-D modelling, geoscience and the authors experience in the geophysical exploration have been applied to convert the available data optimally to meet predictive objectives. Both predictive methods are currently not in existence and have never been previously described in the literature.

6.1 Method 1

Gold distribution prediction is a very complex procedure that requires the simultaneous consideration of a number of geophysical, spatial, structural and economic constraints. Some conventional methods often fail to deal with nonconvex geological sets. Accurate prediction of gold in highly heterogeneous and structurally complex geological environment, in fact, relates to a class of NP-hard problems that can not be reliably solved with currently existing techniques.

The investigated drilling data was limited and extremely erratic, ranging from high to low spatial continuity. Therefore, it was impossible to define reliable gold distribution everywhere in the domain. Basis analysis revealed that the mineralisation had the form of erratic and localised occurrences. The recovered grades ranged between 0.0 and 355.5 g/t.

Through a survey of the literature, it has been found that clusterwise linear regression is one of the most promising methods to be applied to the problem of predicting Cartesian x, y, z coordinates of the gold consentrations and the attributed grades between widely spaced drillholes.

The problem has been tackled by reformulating the prediction into a CLR problem. The developed UML 3.7 system has been tuned for solving complex geological and topological constraints as well as nonlinear problems related to the presence of heterogeneity in the underlying geology. An attempt was made to identify hidden targets, big footprints of ore and the localised occurrences of gold between widely spaced drillholes.

To improve the accuracy of prediction, the local search has been conducted with the discrete gradient method, in which the data was learned through class-related hyperplanes and computed four sets of local optimums using minimal regression mean-squared error (RMSE).

6.1.1 Key observations

The UML system constructed consistent approximations in the form of multi-object predictive structures (prototypes). The outcome appeared as multiple structures whose long axis radiated from the surface. All generated structures were located away from the actual drillholes. In contrast to expectations, the trajectories of the structures tended to radiate from a single point(s) on the surface (drill collars) winding down in a continuous and gradually widening curves to form helix-like shapes. The angular deviations between trajectories were observed as from 0.0061 to 0.0366 radian.

It was found that the structures had tended to mimic the behaviour of the actual trajectories and the extracted drill-cores. The synthesised 5601 sub-structures imitated the actual drilling parameters such as length of drill cores, dip°, azimuth°, maximal depth and grades.

It was observed that two test sets demonstrate remarkable resemblance to the actual drilling data. Side-by-side comparison of the sets supported an idea that the sets can perform similarly well on predicted values. The feature that makes predictive structures similar is near-similar distribution of the densities of the means.

It was found that the increase of the number of hyperplanes in CLR model does not lead to any significant increase of the accuracy of prediction.

The system performed satisfactorily during all conducted runs, utilising a low amount of RAM. The process of prediction was one that was entirely AI controlled subject to available drilling data with humans able only to observe and interpret upon output.

6.1.2 Limitations

Method 1 was found to have the following limitations:

1. The obtained locally optimal solutions only point towards locations where the gold might be located. A compensation for the absence of geostatistical assumptions they, though not unexpected, has been inability of Method 1 to create 3-D solid(s) and provide them with volumetric information.
2. An attempt to create a solid with a relatively small amount of predicted data may yield erroneous results. An attempt to reflect isolated high-grade samples, other than group of clusters with convectional CAD system may lead to biases in 3-D modelling.
3. A source of insignificant limitation is that each program restart (loop re-run) generates new unique set of solutions which can not be repeated.
4. Requires investments in time and mathematical development.
5. It is computationally demanding and requires a user to be familiar with basic programming to execute the simulation.
6. Due to the nature of the data to analyse, statistical description is a difficult and resource-intensive process.

6.1.3 Visualisation

The most effective way to visualise locally optimal solutions is point cloud processing. The representation of nonconvex data can be considered as a 3-D raster, which supports adequate extraction of geometrical features of multiple, potentially rough objects. In this case, one can get a 3-D box filled with points, and the grades with unique colours.

More generally, a geological point cloud becomes a database containing points in the Cartesian coordinate system x, y, z and Au grade.

Considering the fact that more than one million solutions can be generated for one hour, the density of a sparse point cloud can be 0.5–1 pts/m^2. A medium density point cloud of 2–5 pts/m^2 can be achieved by the involvement of 1.5 million points.

In this scenario, geological point clouds can be used to represent volumetric data, as is sometimes done in medical imaging and the visualisation of point clouds at full resolution allows modelling of multiple solids at the highest level of detail, unachievable by conventional GIS systems.

There is a large number of commercial packages that can import x, y, z data, convert the data into point clouds and to pre-process the data for 3-D modelling. Some applications, such as Trimble RealWorks and Pix4D compute optimal point cloud densification. Other packages can convert data into accurate 3D meshing (Leica Cyclone), and provide point classification (Pointools).

Some recently introduced applications can automatically adjust colouring (Bentley Descartes) and texturing (RealityCapture) of 3-D models. Unfortunately, all these applications are expensive and the most important features are not accessible to most PhD students.

6.1.4 Potential end-users

The strength of Method 1 is that it can suggest the existence of gold occurrence at specific regions where multiple low-grade intersections are misleading.

This can be done by setting geometrical limits (the box) with the “search dialogue" which defines maximum and minimums in a particular area of interest. This approach could improve initial targeting and significantly reduce the number of drill holes required to discover a resource.

Apart from important drilling attributes such as target depth, pull-down, rate of penetration and number of rods being used, the crew gains a feel for the gold potential to facilitate efficient targeting. Since the built-in incremental algorithm constantly refines the accuracy of prediction as new drilling data become available, the method can be viewed as a decision-support system for designing optimal drilling grids, and calculation of sufficient amount of exploration drilling. Method 1 can also be involved in testing larger areas.

Method 1 can be effectively utilised in situations when

1. an exploration lease is adjacent to a promising property which has not been explored but potentially worthwhile and
2. the value of this property is based on the exploration potential.

As a decision-support system, Method 1 can help security institutions and the banking community in establishing the fair market value of grassroots exploration projects in the open market. The system can allow an investor to identify future potential problems and to recover losses before actual mine development and production begins.

6.2 Method 2

The existence of heterogeneity in a part of a mine inspired the idea for developing a `convex` method in which the prediction would be capable of solving nonconvex tasks

simultaneously providing statistically-valid global solutions. A number of studies suggested that the feature selection approach would provide a computationally attractive, theoretically justified form of regularisation, which makes the output easier to interpret.

6.2.1 Why feature selection is used?

Predictions of minerals are not precise, being dependent on the limited information available on location, shape and distribution of the occurrence and on the limited sampling results. In selecting the LASSO as the basis, the goals were:

1. The applicability of the method to real-world extractive scenarios.
2. The nature of mining should be considered.

LASSO regularisation was selected for the ability to add a penalty term to the log-likelihood function. This method not only punishes high values of the coefficients β, but also sets them to zero if they are not relevant. This unique feature brought a major improvement from the viewpoint of generalisation error reduction.

6.2.2 Key observations

Due to strict parameter penalisation, Method 2 is less susceptible to the erratic gold distribution and the presence of heterogeneity in the blasthole data. The 10 cross-validation method was applied to assessing the accuracy of the predicted 2-D cross-sections within the investigated flitch. Since the ℓ_1 penalties were convex, the assumed sparsity in the data lead to significant computational efficiency.

The system was stable during each execution (loop re-runs) and demonstrated the ability to compute full regularisation paths. The average time taken for running the code from cold start has been 1 min 50 sec. It was found that the built-in focus on predictive accuracy can provide spatial analysis of significantly bigger spatial data, >50 thousand samples and to serve as a platform for further evaluation techniques.

6.2.3 Advantages

The advantages of the LASSO are as follows

1. RC samples are large, often longer than 2.5 metres and much less variable. In such situations prediction of the gold in between the sampling intervals by kriging is complex. LASSO creates predictive models significantly faster.
2. In kriging, the anisotropy is depicted by computing the variogram in different directions. Since the LASSO is the penalisation/ cross-validation system, it does not require variography.
3. Standard R-studio is an open source environment for R. It is accessible, does not require the application of commercial GIS packages and can be used off-line, at the extraction site.
4. The blasted flitch is to be mined in one 10 metre lift with a mining shovel with bucket capacity (heaped 1:1) from 20 to 50 m^3. The variation of density of rock in oxide Cu/Au deposits (such as Marvin) is usually large and may range from 2.1 to 3.8 g/cm^3 (or 2.1–3.8 t/m^3). At truck-and-shovel extraction stage, gold distribution in g/m^3 is a more convenient metric for the extraction crew to comprehend.

5. The investigated orebody is unstructured, the shapes of low-grade boundaries are gradational. LASSO assists in fast decision making on equipment set up for efficient extraction. The location of high-priority extraction targets is better defined.

To sum, the ℓ_1-norm and `glmnet` provided useful insight on the ore distribution and the location of high-priority production targets within the flitch. Prediction with the LASSO is the first practical step to illustrate this approach in a modern high-dimensional programming language. It was found that the built-in focus on predictive accuracy can provide a spatial analysis of significantly larger data sets and serve a platform for further statistically-valid evaluation techniques.

6.2.4 Limitations

Method 2 was found to have three limitations:

1. The fitting is based on randomly removing subsets of the data for cross-validation. 2-D rendering (plotting) is the last stage, which begins from line 172. The same graphical output from one loop run to the next is unlikely to happen.
2. Selecting PDF for rendering is not recommended. Regardless of an operating system, the application of Adobe vectorisation during multiple runs may result in the appearance of insignificant geometrical differences between models (see Figure 5.17). Instead of PDF (Appendix C, line 159), it is suggested .png or .jpg to be used because these formats are less sensitive to the sequence of pixelisation.
3. Smoothness and the requirement to average 2-D graphics. Prior to making a decision, a comparison of the cross-sections obtained from subsequent runs is highly desirable.

However, these limitations are not significant and do not have an impact on the output.

6.2.5 Potential end-users

The application of Method 2 is not limited to the production stage. In the grassroots exploration phase, the most useful feature of the LASSO is the automatic graphical representation of cross-sections through short (pre-set) intervals. Depending on PC or MAC hardware, hundreds of predictive cross-sections can be plotted in a reasonably short time.

Apart from important drilling attributes such as target depth, pulldown, rate of penetration and number of rods being used, the exploration crew gains a feel for gold potential. The method allows assuming a sufficient amount of drilling and facilitates efficient targeting. LASSO makes it possible to point to the likelihood of the existence of exploration targets at the beginning of the program.

The method is a user-friendly, accessible, can be used offline, at remote areas and does not require online licence verification. The only requirement is regular drilling database updates and revisions. Through the use of the program, the exploration and extraction crews can see in near-real time the effect of the taken decisions in the search for the presence, the absence and the extension of high-priority mineralisation. Since the method is fast and the cost is not prohibitive, it is integrable into any drilling software (PC, Unix, Linux, etc.).

Method 2 would also be beneficial to the mine consulting professionals who are working in machine learning tasks.

6.3 Future Work

Gold extraction is going deeper. The average depth of discovery for gold deposits in Australia is rising 10 metres per decade. By 2032, half of all Australia's gold production will come from deposits yet to be discovered. To predict and discover a world-class Tier 1 or smaller economical deposits, the mining industry will require more efficient exploration and the development of tools that can provide greater precision in less time and for less money.

Extensions to the work done in this thesis may be focused on the development of free and user-friendly predictive applications with convex and nonconvex objective functions that are capable of taking into account geological and structural uncertainty. Future research will be devoted to the development of an affordable and accessible applications that are capable of representing 3-D shapes of the underlying mineral content defined by a selected coordinate's system.

It may be helpful to undertake additional studies on the automatic convergence of solutions into point clouds and provide the required density for a more accurate 3-D resource modelling in semi-transparent view.

References

Abedi, M., Norouzi, G., & Fathianpourb, N. (2013). Fuzzy outranking approach: A knowledge-driven method for mineral prospectivity mapping. *Intl. J. of Applied Earth Observation and Geoinformation*, *21*, 556–567.

Adamu, M. & Brandon, B. (2014). Point-based model for predicting mineral deposit using gis and machine learning. *First Conference on Systems Informatics; Modelling and Simulation; IEEE*, 65–70.

Ahlberg, J., Nilson, E., & Walsh, J. (1969). The theory of splines and its applications. *SIAM*, *11*(1), pp.91–92.

Akteke-Öztürk, B., Weber, G.-W., & Güiser, K. (2017). Optimization of generalized desirability functions under model uncertainty. *Optimization*, *66*(12), 2157–2169.

Annels, A. & Dominy, S. (2003). Core recovery and quality: Important factors in mineral resource estimation. *Applied Earth Science : IMM Transactions section B*, *112*, 305–312.

Annels, E. (2004). Mineral deposit evaluation: A practical approach, *Chapman and Hall London, New York*, xv., ill, 436p.

Armstrong, M. (1994). Is research in mining geostats as dead as dodo? in geostatistics for the next century; r. dimitrakopolous. *Kluwer, Dordrecht, Netherlands*, 303–312.

Armstrong, M. & Champigny, N. (1989). A study on kriging small blocks. *CIM Bulletin, Centre de Géostatistique, France*, *82*(923), 128–133.

Azpurua, M. & Ramos, K. D. (2010). A comparison of spatial interpolation methods for estimation of average electromagnetic field magnitude. *Progress In Electromagnetics Research*, *14*, 135–145.

Bagirov, A. (2008). Modified global k-means algorithm for minimum sum-of-squares clustering problems. *Pattern Recognition, Elsevier Science Inc., New York*, *41*(10), 3192–3199.

Bagirov, A., Karmitsa, N., & Mäkela, M. (2014). *Introduction to nonsmooth optimization: Theory, practice and software*. Springer International Publishing.

Bagirov, A., Ugon, J., & Mirzayeva, H. (2013). Nonsmooth nonconvex optimization approach to clusterwise linear regression problems. *European Journal of Operational Research*, *229*(1), 132–142.

Barnett, C. & Williams, P. (2006). Mineral exploration using modern data mining techniques. *Society of Economic Geologists*, (12), 285–310.

Baxter, E. (2004). *Distributed hydrologic modeling using gis* (Ed.2). Springer Netherlands.

Berkman, D. A. (2001). *Field geologist's manual* (4th ed). Carlton, Vic. : Australasian Institute of Mining and Metallurgy.

Bermingham, M. L., Pong-Wong, R., Spiliopoulou, A., Hayward, C., Rudan, I., Campbell, H., & Haley, C. S. (2015). Application of high-dimensional feature selection: Evaluation for genomic prediction in man. *Scientific Reports*, *5*(10312).

Bivand, R., Pebesma, E., & Gómez-Rubio, V. (2013). Applied spatial data analysis with r. second edition. *Springer*, 405p.

Bonham-Carter, G., Agterberg, F., & Wright, D. (1998). Integration of geological datasets for gold exploration in nova scotia. *Digital Geologic and Geographic Information Systems*, 15–23.

Bradley, P. (1925). Mining methods in alaska-juneau mine. *Trans. American Institute Mining and Metallurgy*, *72*, 103–109.

Bühlmann, P. & van de Geer, S. (2011). Statistics for high-dimensional data: Methods, theory and applications. *Springer-Verlag, Berlin, Heidelberg*, XVIII, 558p.

Burrough, P. & McDonnell, R. (1998). Principles of geographical information systems. *New York, Oxford*.

Campbell, T. & Fiske, W. (1959). Convergent and discriminant validation by the multitrait multimethod matrix. *Psychological Bulletin*, *56*, 81–105.

Cardon, H. & Hoogstraten, R. (1995). Key issues for successful industrial neural-network applications: An application in geology. *Artificial Neural Networks, Springer*, 235–245.

Carranza, E. (2011). Geocomputation of mineral exploration targets. *Computers & Geosciences*, *37*(12), 1907–1916.

Carras, S. (1990). Ore reserve estimation at the feasibility study stage -the period of greatest risk. *AusIMM, Melbourne Australia*, 99–108.

Çevik, A., Weber, G.-W., B.M., E., & Oğuz, K. (2017). Voxel–mars: A method for early detection of alzheimer's disease by classification of structural brain mri. *Annals of Operations Research*, *258*(1), 31–57.

Chilès, J. & Delfiner, P. (2012). *Geostatistics: Modeling spatial uncertainty* (2 Edition). Wiley.

Clarici, E., Owen, D., & Durucan, S. (1993). Modeling and representation of spatial data using artificial neural networks. *Proceedings, Symposium on Innovative Mine Design for the 21st Century*, 199–204.

Clow, G. (1991). Why gold mines fail. *Northern Miner*, 31–34.

Converse, A. (1970). *Optimization*.

Cox, D. & Miller, H. (1965). The theory of stochastic process. *London, Chapman and Hall*, pp.378–385.

Cracknell, M. & de Caritat, P. (2017). Catchment-based gold prospectivity analysis combining geochemical, geophysical and geological data across northern australia. *Geochemistry-exploration environment analysis*, *17*(3), 204–216.

Cressie, N. (1989a). Geostatistics. *The American Statistician*, *43*(4).

Cressie, N. (1989b). The origins of kriging. *Mathematical Geology*, *22*(3), 239–253.

David, M. (1977). *Geostatistical ore reserve estimation*. Elsevier Scientific Pub. Co.

Dershowitz, W. & Einstein, H. (1989). Characterizing rock joint geometry with joint system models. *Rock Mechanics and Rock Engineering*, *21*, 21–51.

DeSarbo, W., Oliver, R., & Rangaswamy, A. (1989). A simulated annealing methodology for clusterwise linear regression. *Psychometrika*, *54*(4), 707–736.

Desharnais, G., Paiement, J., Hatfield, D., & Poupart, N. (2017). Mining big data: The future of exploration targeting using machine learning. *Conference Paper; Exploration targeting*.

Diggle, P. & Ribeiro, P. (2007). Model-based geostatistics. *Springer-Verlag, New York*, XIV, 232.

Dimitrakopoulos, R. (1993). Artificially intelligent geostatistics: A framework accommodating qualitative knowledge-information. *in Mathematical Geology*, *25*(3), 261–279.

Dowding, C. (1976a). Comparison of predicted and encountered geology for seven colorado tunnel. *Society for Mining, Metallurgy & Exploration*, 21p.

Dowding, C. (1976b). Geologic investigation and prediction. *RETC proceedings*, 19–38.

Ducart, D., Silva, A., Toledo, C., & de Assis, L. (2016). Mapping iron oxides with landsat-8/oli and eo-1/hyperion imagery from the serra norte iron deposits. *Brazilian Journal of Geology*, *46*(3), 332–349.

Everitt, B., Landau, S., Leese, M., & Stahl, D. (2011). Cluster analysis. *Wiley series in probability and statistics*, 848p.

Falivene, O., Cabrera, L., Tolosana-Delgado, & Sáez, A. (2010). Interpolation algorithm ranking using cross-validation and the role of smoothing effect. *Computers and Geosciences*, *36*(4), 512–519.

Fletcher, R. (1987). Practical methods of optimization. *Optimization Software, Inc. Chichester-New York-Brisbane-Toronto, sec. ed.*

Fotheringham, A., Brunsdon, C., & Charlton, M. (2002). Geographically weighted regression: The analysis of spatially varying relationships. *John Wiley & Sons*, *13*.

Friedman, J., Hastie, T., & Tibshirani, R. (2008). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, *33*(1), pp.1–22.

Gaetan, G. & Guyon, X. (2010). Spatial statistics and modelling. *Springer*, *14*, 302p.

Galit, S. (2010). To explain or to predict? *Statistical Science*, *25*(3), 289–310.

Ghosh, A., Harpalani, S., & Daemon, J. (1987). Expert system for coal mine roof bolt design. *Proc. 28th US Rock Mechanics Symposium, Tuscon, Arizona*, 1137–1144.

Goovaerts, P. (1997). Geostatistics for natural resources evaluation. *Oxford University Press, New York*, 483p.

Gouws, F. & Aldrich, C. (1996). Rule based characterization of industrial flotation processes with inductive techniques and genetic algorithms. *Industrial and Engineering Chemistry Research*, *35*(11), 4119–4127.

Grange, S. K., Lewis, A. C., & Carslaw, D. C. (2016). Source apportionment advances using polar plots of bivariate correlation and regression statistics. *Atmospheric Environment*, *145*, 128–134.

Grivas, D. & Reagan, J. (1988). An expert system for the evaluation and treatment of earth slope instability. *Proc. 5th Int. Symp. on landslides, Lausanne, Brookfield VT: Balkema*, 649–654.

Groves, D., Condie, K., Goldfarb, R., Hronsky, J., & Vielreicher, R. (2005). Secular changes in global tectonic processes and their influence on the temporal distribution of gold-bearing mineral deposits. *Economic Geology*, *100*(2), 203–224.

Guilford, P., Kettner, W., & Christensen, R. (1956). The nature of the general reasoning factor. *Psychological review*, *63*, 169–72.

Gy, P. (1989). Pierre gy's sampling theory and sampling practice. *Boca Raton, Fla. : CRC Press. 2v.*

Hagens, A. & Doveton, J. (1991). Application of a simple cerebellar model to geologic surface mapping. *Computers & Geosciences*, *17*(4), 561–567.

Hale, A. (1881). Memoranda on the analysis of statistics. *Transactions of the American Institute of Mining Engineers*, (9), 608–620.

Hales, L. & Ynchausti, R. (1992). Neural networks and their use in the prediction of sag mill power. *Communition Theory and Practice*, 495–504.

Harrington, P. (2011). *Machine learning in action*. JRC Institute for Environment and Sustainability.

Harris, B. (1989). Beyond geographic information systems: Computers and the planning professional. *Journal of American Planning Association*, *44*, 85–92.

Harrison, A. & Hanover, N. (1929). Extension of oreshoots with comments on the art of ore finding. *Transactions, AIMME*, (1), 318–324.

Harvey, A. & Fotopoulos, G. (2016). Geological mapping using machine learning algorithms. *XIII ISPRS Congress*, *B8*(6).

Hastie, T., Tibshirani, R., & Friedman, J. (2017). *The elements of statistical learning*. Springer Series in Statistics. Springer-Verlag New York.

Hazen, S. (1967). *Some statistical techniques for analyzing mine and mineral-deposit sample and assay data*. Bullet.621. U.S. Dep.of the Interior, Bureau of Mines.

Heale, R. & Forbes, D. (2013). Understanding triangulation in research. *Evidence-Based Nursing*, *16*(4), 98–98.

Hekmat, A., Osanloo, M., & Moarefvand, P. (2013). Block size selection with the objective of minimizing the discrepancy in real and estimated block grade. *Arabian Journal of Geosciences*, *6*(1), 141–155.

Hengl, T. (2007). *Practical guide to geostatistical mapping of environmental variables*. Manning Publications.

Henley, S. (2001). The importance of being stationary. *Earth Science Computer Applications*, *16*(12), 1–3.

Henley, S. & Watson, D. (1998). Possible alternatives to geostatistic. *Computer applications in the minerals industries. International symposium*, pp.337–354.

Henning, C. (2016). *Handbook of cluster analysis*. Chapman & Hall/ CRC Handbooks of Modern Statistical Methods.

Hodouin, D., Flament, F., & Bazin, C. (1989). Reliability of material balance calculations - a sensitivity approach. *Minerals Engineering*, *2*(2), 157–169.

Hoerl, A. & Kennard, R. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*(1), 55–67.

Isaaks, E. & Srivastava, R. (1989). *Applied geostatistics*. Oxford University Press, ISBN 978-0- 19-605013-4.

Jackson, C. & Knaebel, J. (1934). Sampling and estimation of ore deposits. *U.S. Department of Commerce*, (356), 166p.

James, G., Witten, D., Hastie, T., & Tibshiran, R. (2013). An introduction to statistical learning with applications in r. *Springer*, 204p.

James, G., Witten, D., Hastie, T., & Tibshiran, R. (2015). *Statistical learning with sparsity the lasso and generalizations*.

James, G., Witten, D., & Tibshiran, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B*(58), 267–288.

Joly, A., Porwal, A., & McCuaig, T. (2012). Exploration targeting for orogenic gold deposits in the granites-tanami orogen: Mineral system analysis, targeting model and prospectivity analysis. *Ore Geology Reviews*, *48*, 349–383.

JORC. (2012). Australasian code for reporting of exploration results, mineral resources and ore reserves. *AIGMCA*, 44p.

Journel, A. & Huijbregts, C. (1978). *Mining geostatistics*. Blackburn Press.

K., K. & A., G. (2014). Pragmatic bayesian kriging for non-stationary and moderately non-gaussian data, 61–64.

Kara, G., Özmen, A., & Weber, G.-W. (2019). Stability advances in robust portfolio optimization under parallelepiped uncertainty. *Central European Journal of Operations Research*, *27*(1), 241–261.

Karmitsa, N., Bagirov, A., & Makela, M. (2012). Comparing different nonsmooth minimization methods and software. *Optimization Methods and Software*, *27*(1), 131–153.

Kiaei, H., Sharghi, Y., Ilkhchi, A., & Naderi, M. (2015). 3d modeling of reservoir electrofacies using integration clustering and geostatistic method in central field of persian gulf. *Journal of Petrolium Engineering*, *135*, 152–160.

King, H. & McMahon, D. (1982). A guide to understanding of ore reserve estimation. *AusIMM, Supplement to Proceedings 281, Parkville, Victoria*, *p.1*.

Kinnicutt, P. (1994). Extended visualisation and geostatistical functionality in nomad, a 3-d subsurface modeler. *Proc. 8th Int. Conf. Computer Methods and Advances in Geomechanics, Balkema*, *1*, 459–564.

Knight, R. H., Lane, R. G., Ross, H. J., Abraham, A., & Cowan, J. (2007). Implicit ore delineation. *In Proceedings of Exploration 07: Fifth Decennial International Conference on Mineral Exploration*, 1165–1169.

Knoll, K. (1989). And now the bad news. *Northern Miner*, *6*, 48–52.

Krige, D. (1951). A statistical approach to some basic mine valuation problems on the witwatersrand. *Journal of the Chemical, Metallurgical & Mining Society of South Africa*, *52*(6), 119–139.

Krige, D. (1976). Some basic considerations in the application of geostatistics to gold ore valuation. *Future trends in Geomathematics*, *Pion Ltd.*

Lee, C. & Sterling, R. (1992). Identifying probable failure models for underground openings using a neural network. *Int. Journal of Rock Mechanics and Mining Sciences & Geomechanical Abstracts*, *29*(1), 49–67.

Li, J. & Heap, A. (2008). A review of spatial interpolation methods for environmental scientists. *Geoscience Australia*, *5*(2), xvi, 137 p.

Loquin, K. & Dubois, D. (2010). Kriging and epistemic uncertainty: A critical discussion. *Methods for Handling Imperfect Spatial Information, Springer*, *256*(3), 269–300.

Lucas, J. (2018). Australia's major gold miners digging deeper underground in hunt for major discoveries. *ABC News, 4 March.*

Mackenzie, B. & Doggett, M. (1992). Worldwide trends in gold exploration. *Kingston, Ont.: Centre for Resource Studies*, *Queen's University*, xi, 67p.

Matheron, G. (1963). Principles of geostatistics. *Economic Geology*, *58*(8), pp.1246–1266.

Matheron, G. & Kleingeld, G. (1987). The evolution of geostatistics. *Geostat, ARCOM 87*, *3*, pp.9–12.

McCammon, R. (1989). Prospector ii-the redesign of prospector. *AI Systems in Government; Washington, DC.* 86–92.

Melkumyan, A. & Ramos, F. (2011). Non-parametric bayesian learning for resource estimation in the autonomous mine. *The Australasian Institute of Mining and Metallurgy*, *35th APCOM Symposium, Carlton.*

Merks, J. (1992). Geostatistics or voodoo science? *Northern Miner*, *April 20*, p.A4.

MIAC. (2012). Mineral exploration drilling, code of practice, 56p.

Millar, D. & Hudson, J. (1994). Performance monitoring of rock engineering systems using neural networks. *Trans. Inst. Mining Metallurgy*, A3–A16.

Miskelly, N. (1982). Towards a better understanding of ore reserves. *The Securities Institute Journal*, (2), 12–15.

Mitchell, T. (1997). *Machine learning, 1st ed.* Mcgraw-Hill Series in Computer Science.

Mueller, F. (1878). *Ueber die vortheile der mimicry bei schmetterlingen* (4th ed). Digitized in Harvard University, 18 Jul 2008.

Myers, D. (1989). To be or not to be ... stationary? *Mathematical Geology*, *21*(3).

Noble, A. (2011). Mineral resource estimation; sme mining engineering handbook. (4), pp.03–217.

Oliver, M. & Webster, R. (2015). *Basic steps in geostatistics: The variogram and kriging.* Springer.

Ord, J. (1983). Encyclopedia of statistical sciences. *Wiley, New York*, *4*, pp.709–714.

Paciorek, B. (2008). Why kriging in arcgis may be a bad idea: A statistician's perspective. *Technical Vignette; Harvard School of Public Health*, (1).

Pan, G. & Harris, D. P. (1993). Delineation of intrinsic geological units. *Mathematical Geology*, *25*(1), 9–39.

Parker, H. (2012). Reconciliation principles for the mining industry. *Transactions of the IMM: Section A*, *121*(3), 160–176.

Patel, M., Chaudhary, S., & Garg, S. (2018). Improved pre-copy algorithm using statistical prediction and compression model for efficient live memory migration. *International Journal of High Performance Computing and Networking*, *11*, 55.

Philip, G. & Watson, D. (1986). Geostatistics and spatial data analysis. *Math. Geol. 18*(5), 505–508.

Pimpler, E. (2017). Spatial analysis with arggis. *Packt*, 290p.

Pitard, F. (1993). Sampling theory and sampling practice, 488p.

Polyak, B. (1987). Introduction to optimization. *Optimization Software, Inc., Publications Division*.

Poole, D. & Mackworth, A. (2017). Artificial intelligence: Foundations of computational agents. *Cambridge University press*, 820p.

Postolski, T. & Sinclair, A. (1998). Quantitative estimation of dilution and ore loss resulting from block estimation errors and a specified cutoff grade. *Exploration and Mining*, *7*(1), 91–98.

Rehak, D., Christiano, P., & Norkin, D. (1985). An expert system component of a geotechnical site characterization workbench. *App. of Knowledge-based Syst to Eng Analysis and Design, Am. Soc. Mech. Eng., New York*, 117–133.

Rendu, J. (1994). Mining geostatistics forty years passed: What lie ahead? *Society for Mining Metallurgy and Exploration. Jackling Lecture*.

Richert, G. (1929). Mining methodsat mines de matahambre. *Pinar del Rio, Bureu of Mines Cuba, Inf. Circ 6145*.

Robinson, C. & Lee, F. (1967). Open file map of geology encountered in straight creek pilot bore, colorado. *USGS*, *59*(1).

Rockafellar, T. (1994). *Nonsmooth optimisation*.

Rodriguez-Galiano, V. & Chica-Rivas, M. (2015). Evaluation of different machine learning methods for land cover mapping of a mediterranean area using multi-seasonal landsat images and digital terrain models. *International Journal of Digital Earth*, *7*(6), 492–509.

Rossi, M. & Deutsch, C. (2014). *Mineral resource estimation*. 332p.

Sarkar, B., O'Leary, J., & Mill, A. (1988). An integrated approach to geostatistical evaluation. *Mining Magazine*, *159*(3), 611p.

Schodde, R. (2017a). Long-term forecast of australia's mineral production and revenue outlook for gold: 2017–2027. *Keynote presentation at the CET Members Day, Perth*.

Schodde, R. (2017b). Recent trends and outlook for global exploration. *Presentation to the PDAC*.

Selley, R. C., Cocks, L. R. M., & Plimer, I. R. (2005). *Encyclopedia of geology*. Elsevier Academic Amsterdam ; London.

Shi, B., Bloom, L. M., & Mueller, U. A. (2000). Applications of conditional simulation to a positively skewed platinum mineralization. *Natural Resources Research*, *9*(1), 53–64.

Shurtz, R. (1985). A critique of journel's "the deterministic side of geostatistics. *Math. Geol. 17*(8), 861–868.

Shurtz, R. (1994). Spatial least squares approximation and geologic plausibility. *Proc.1994 Inter. Assoc. Math. Geol. Ann. Conf. 3-5*, 319–324.

Sichel, H. (1947). An experimental and theoretical investigation of bias error in mine sampling with special reference to narrow gold reefs. *Trans.Inst.Min.Metall.* (56), 403–443.

Simon, N., Friedman, J., Hastie, T., & Tibshirani, R. (2011). Regularization paths for cox's proportional hazards model via coordinate descent. *Journal of Statistical Software*, *39*(5), 1–13.

Sinclair, A. & Blackwell, G. (2002). Applied mineral inventory estimation. *Cambridge University Press*, (2), 95–108.

Sinclair, A. & Blackwell, G. (2004). Applied mineral inventory estimation. *Cambridge University Press*, (1), 401p.

Sinclair, A. & Vallee, M. (1993). Improved sampling control and data gathering for improved mineral inventories and production control. *Geostatistics for the Next Century, Kluwer*, *June 3-5*, 253–259.

Singer, D. & Koude, R. (1997). Use of neural network to integrate geoscience information in the classification of mineral deposits and occurrences. *Proceedings of Exploration 97; 4th. ICME*, 127–123.

Skabar, A. (2007). Mineral potential mapping using bayesian learning for multilayer perceptrons. *Mathematical Geology*, *39*(5), 439–451.

Solderberg, A. (1930). Mining methods and costs at the utah copper co., bingham canyon, utah. *Bureau of Mines*, *Inf. Circ. 5234*, p.23.

Späth, H. (1979). Algorithm 39 clusterwise linear regression. *Computing*, *22*(4), 367–373.

Tajdar, V. (2006). Prediction of ultimate pile bearing capacity using artificial neural networks. *10th. International Conference on Piling and Deep Foundations, Amsterdam.*

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society*, *B*(58), 267–288.

Tibshirani, R., Bien, J., Friedman, J., Hastie, T., Simon, N., Taylor, J., & Tibshirani, R. (2010). Strong rules for discarding predictors in lasso-type problems. *Journal of the Royal Statistical Society*, *74*(2), 245–266.

Tolosana-Delgado, R., Mueller, U., & van den Boogaart, K. G. (2019). Geostatistics for compositional data: An overview. *Mathematical Geosciences*, *51*(4), 485–526.

Tolwinski, B. & Underwood, R. (1992). An algorithm to estimate the optimal evolution of an open pit mine. *Optimisation techniques for open pit mining scheduling*, *Proc. of ARCOM XXII, vol. 25*(3), 399–409.

Uria-Tellaetxe, I. & Carslaw, D. C. (2014). Conditional bivariate probability function for source identification. *Environmental Modelling & Software*, *59*, 1–9.

Velasco, J. & González-Salazar, C. (2019). Akaike information criterion should not be a "test" of geographical prediction accuracy in ecological niche modelling. *Ecological Informatics*, *51*, 25–32.

Wagenmakers, E., Ratcliff, R., Gomez, P., & Iverson, G. (2004). Assessing model mimicry using the parametric bootstrap. *J. Math. Psychololy*, *48*(1), 28–50.

Wahlstrom, E. (1964). The validity of geological projection: A case history. *Econ. Geology*, *59*(3), 465–474.

Wattermeyer, G. (1919). The application of the theory of probability in the determination of ore reserves. *J. Chem. Met. and Min. Soc. South Africa*, *19*(7), 102–107.

Webster, R. & Oliver, M. (1992). Sample adequately to estimate variograms of soil properties. *Journal of Soil Science*, *43*, 177–192.

Webster, R. & Oliver, M. (2007). Geostatistics for environmental scientists. *John Wiley and Sons*, 317p.

Williams, P. (2002). Probabilistic learning models, chapter, from foundations of bayesianism. *Kluwer*, pp.117–134.

Wu, X. & Zhou, Y. (1993). Reserve estimation using neural network techniques. *Computers & Geosciences, Pergamon Press*, *19*(4), 567–575.

Yama, B. & Lineberry, G. (1999). Artificial neural network application for a predictive task in mining. *Mining Engineering*, *51*(2), 59–64.

Yerlikaya-Özkurt, F., Askan, A., & Weber, G.-W. (2016). A hybrid computational method based on convex optimization for outlier problems: Application to earthquake ground motion prediction. *Informatica, Lith. Acad. Sci. 27*, 893–910.

Yu, T. & Vongpaisa, S. (1996). New blast damage criteria for underground blasting. *CIM Bulletin*, *89*, 139–145.

Zhang, M. & Bhattacharyya. (1995). Characterisation of undermined strata in predicting coal mining subsidence. *The AusIMM Annual Conference, Newcastle*, *23*(25), 45–50.

Zou, H. & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Stanford University, USA*, *67*(2), pp.301–320.

Zuo, E. & Carranza, J., E. (2011). Support vector machine: A tool for mapping mineral prospectivity. *Computers & Geosciences*, *37*(12), 1967–1975.

Appendix A: Raw data transformation source code `fortran`

Execution code: transformation of raw drill-core Snd assay to input format. Syntax adjusted to LaTeX2

```
c=======================================================================
c      Transformation from actual to input format syntax adjusted to LaTeX2e
c=======================================================================
c      main programm
       implicit double precision(a-h,o-z)
        PARAMETER(maxrec=100000, maxnft=30)
         double precision a1(maxrec,maxnft),a2(maxrec,maxnft)
         1,a3(maxrec,maxnft),a4(maxrec,maxnft),a(maxrec,maxnft)
         open(40,file='Results.txt')
        open(78,file='datainput1.txt',status='old',form='formatted')
        open(79,file='datainput2.txt',status='old',form='formatted')
       open(80,file='datainput3.txt',status='old',form='formatted')
c=======================================================================
       nft1=3
        nft2=3
        nft3=3
       do i=1,maxrec
         read(78,*,END=901,ERR=900) (a1(i,k),k=1,nft1)
         nrecord1=i
       end do
  900 stop 'Error in input file'
  901 WRITE(40,*) 'Input complete. Number of records: ',nrecord1
      WRITE(40,*)

       do i=1,maxrec
          read(79,*,END=903,ERR=902) (a2(i,k),k=1,nft2)
          nrecord2=i
       end do
  902 stop 'Error in input file'
  903 WRITE(40,*) 'Input complete. Number of records: ',nrecord2
      WRITE(40,*)

       do i=1,maxrec
         read(80,*,END=905,ERR=904) (a3(i,k),k=1,nft3)
         nrecord3=i
       end do
  904 stop 'Error in input file'
  905 WRITE(40,*) 'Input complete. Number of records: ',nrecord3
      WRITE(40,*)
c=======================================================================
c            STEP 1
       n1=1
       n2=1
       do i=1,nrecord1
    1   do j=1,nft1
          a4(n2,j)=a1(i,j)
        end do
         do j=1,nft2
          a4(n2,j+nft1)=a2(n2,j)
        end do
         n2=n2+1
          if(n2.gt.nrecord2) go to 49
         if(a2(n2,1).le.a2(n2-1,1)) then
          go to 3
         else
          go to 1
         end if
    3  end do
```

```
84     49   continue
85            n3=nft1+nft2
86   c=========================
87   c           STEP 2
88          n6=0
89          do i=1,nrecord3
90           c1=a3(i,1)
91           do j=1,nrecord2-1
92            c2=a4(j,4)
93            c3=a4(j+1,4)
94            if((c1.ge.c2).and.(c1.lt.c3)) then
95             n6=n6+1
96             do k=1,6
97              a(n6,k)=a4(j,k)
98             end do
99             do k=1,3
100             a(n6,k+6)=a3(i,k)
101            end do
102            go to 8
103           end if
104          end do
105    8   end do
106         do i=1,n6
107           a(i,9)=dmax1(0.0d+00,a(i,9))
108         end do
109         write(40,*)
110         do i=1,n6
111          write(40,43) (a(i,j),j=1,9)
112         end do
113   43    format(9f12.2)
114         close(40)
115         close(78)
116         close(79)
117         close(80)
118       stop
119      end
```

Appendix B: UML Source code `fortran`

Method 1 №1: Computation of local optimums for drill-core data. Original syntax adjusted to LaTeX2

```
71
72
73  c     implicit double precision (a-h,o-z)
74        PARAMETER(maxvar=2000)
75        double precision xl(maxvar),xu(maxvar),x0(maxvar),x(maxvar)
76  c=======================================================================
77  c  Input data:
78  c  n        - number of variables
79  c  nlineq   - number of linear equalities
80  c  neq      - number of nonlinear equalities
81  c  nlinineq - number of linear inequalities
82  c  nineq    - number of nonlinear inequalities
83  c=======================================================================
84     n = 10
85      nlineq = 0
86      neq = 0
87      nlinineq = 0
88     nineq = 0
89  c=======================================================================
90  c  lower and upper bounds for variables:
91  c  xl - lower bound
92  c  xu - upper bound
93  c  If they are not given take xl < -10^5 and xu > 10^5
94  c=======================================================================
95        do i=1,n
96          xl(i) = -1.d+08
97          xu(i) = 1.d+08
98        end do
99        do i=1,n
100         xl(i) = -32.78
101         xu(i) = 32.78
102       END do
103 c=======================================================================
104 c  Starting point
105 c=======================================================================
106       do i=1,n
107        x0(i)=1.0d+01
108       END do
109 c=======================================================================
110 c  Calling DFHM1
111 c=======================================================================
112       call dgm01(n,neq,nlineq,nlinineq,nineq,xl,xu,x0
113         1 ,x,fvalue,convalue,cputime)
114       STOP
115       END
116 c=======================================================================
117 c#  In this subroutine we describe the problem entering
118 c#  objective function and constraints
119 c=======================================================================
120       subroutine func(x,objf,conlineq,conlinineq,coneq,conineq)
121       implicit double precision (a-h,o-z)
122       PARAMETER(maxvar=2000, maxcon=2000)
123       double precision x(maxvar),coneq(maxcon),conineq(maxcon)
124      1 ,conlineq(maxcon),conlinineq(maxcon)
125 c=======================================================================
126 c       coneq - array of nonlinear equalities
127 c       conlineq - array of linear equalities
128 c       conlinineq - array of linear inequalities
129 c       conineq - array of nonlinear inequalities
130 c=======================================================================
```

```
131         pi=3.141592654d+00
132           f1=0.d+00
133            f2=0.d+00
134             do i=1,10
135            f1=f1+x(i)*x(i)
136           f2=f2+dcos(2.0d+00*pi*x(i))
137         END do
138         objf=-2.d+01*dexp(-2.d-01*dsqrt(f1/1.0d+01))-
139        1    dexp(f2/1.0d+01)+2.d+01+dexp(1.0d+00)
140   c=======================================================================
141         return
142         end
143   c=======================================================================
144   c# This subroutine calls both local and global solvers: OPTIMUM and OPTIMUM1.
145   c# It also calls subroutines to solve the system of linear equations and identify box constraints.
146   c# This generates output files: results.txt and local_solutions.txt
147   c=======================================================================
148         subroutine dgm01(nvar,neq,nlineq,nlinineq,nineq,xl,xu,x0
149        1  ,xfinal,ob,fpenglob,cputime)
150         implicit double precision (a-h,o-z)
151         PARAMETER(maxvar=2000, maxcon=2000)
152         double precision x(maxvar),x3(maxvar),x0(maxvar)
153        1  ,xl(maxvar),xu(maxvar),xlow(maxvar),xup(maxvar),coneq(maxcon)
154        2  ,conineq(maxcon),conlineq(maxcon),conlinineq(maxcon)
155        3  ,c1(maxcon,maxvar),xfinal(maxvar),x2final(maxvar)
156         INTEGER l1(maxvar),l3(maxvar),indlow(maxvar),indup(maxvar)
157         COMMON /cbound/xlow,xup,/csize/m,/citer/niter,/cindbound/indbound
158        1  ,/cconeq/numeq,/cconin/numineq,/cpenf/fconst,/cpen/penalt
159        2  ,/cl3/l3,n2,/cc1/c1,/cnshift/nshift,/cconlineq/numlineq,/cl1/l1
160        3  ,/csize1/m1,/cnumneq/numneq,/cindicator/indlow,indup
161        4  ,/cobstart/objstart,/cconlinineq/numlinineq
162        5  ,/cnconst/nconst,/cnf/nf,/cnewpoint/newpoint
163         character*30 outfi/'results.txt'/
164         open(40,file=outfi)
165         call cpu_time(time1)
166   c=======================================================================
167         WRITE(40,*)
168         WRITE(40,112)
169    112  FORMAT('      **** DISCRETE GRADIENT METHOD ******')
170         WRITE(40,*)
171         WRITE(40,115)
172     115 FORMAT('                (c) FoST')
173         WRITE(40,*)
174         WRITE(40,15)
175     15  FORMAT('                (c) Federation University Australia')
176         WRITE(40,*)
177         WRITE(40,*)
178         WRITE(40,*)
179         WRITE(40,1113)
180    1113 FORMAT(' 'Solver System for drill assay nonsmooth optimisation')
181         WRITE(40,*)
182         WRITE(40,*)
183         WRITE(40,113)
184     113 FORMAT('
185         WRITE(40,*)
186         WRITE(40,*)
187   c=======================================================================
188   c accuracy for solution feasibility
189         eps1=1.d-06
190   c accuracy for acceptance of solution
191         eps2=1.d-04
192         m=nvar
193         m1=m
194   c box constraints
195         do i=1,m
196           xlow(i)=xl(i)
197           xup(i)=xu(i)
198         end do
199          numlineq=nlineq
200           numeq=neq
201            numlinineq=nlinineq
202            numineq=nineq
203           numneq=neq
```

```
204          nconst=numeq+numlinineq+numineq
205         nconst1=nconst+numlineq
206   c nf -total number of function calls and niter - total number of iterations
207         nf=0
208         niter=0
209   c=======================================================================
210   c# Printing information on input data in file results.txt
211   c=======================================================================
212         WRITE(40,*)
213         write(40,1)
214    1    FORMAT('                          Input data')
215         write(40,*)
216         write(40,11) m1
217      11 FORMAT('Number of variables',i50)
218         WRITE(40,*)
219         write(40,1112) numlineq
220   1112  FORMAT('Number of linear equality constraints',i32)
221         WRITE(40,*)
222         write(40,1114) numlinineq
223   1114  FORMAT('Number of linear inequality constraints',i30)
224         WRITE(40,*)
225         write(40,12) numeq
226      12 FORMAT('Number of nonlinear equality constraints',i29)
227         WRITE(40,*)
228         write(40,13) numineq
229      13 FORMAT('Number of nonlinear inequality constraints',i27)
230         WRITE(40,*)
231         WRITE(40,*)
232         WRITE(40,*)
233         WRITE(40,1118)
234   1118  FORMAT('Initial point given by user:')
235         WRITE(40,*)
236         do i=1,m1
237          write(40,100) i,x0(i)
238         end do
239         WRITE(40,*) '_____________________________________________________________'
240   c=======================================================================
241   c# Projecting onto box (if starting point is outside the box)
242   c=======================================================================
243         fbox1=0.d+00
244         do i=1,m
245          fbox1=fbox1+dmax1(0.d+00,xlow(i)-x0(i))
246        1  +dmax1(0.d+00,x0(i)-xup(i))
247         end do
248         if (fbox1.gt.0.d+00) then
249           WRITE(40,*)
250           WRITE(40,1014)
251           WRITE(40,*)
252   1014    FORMAT('The initail point is outside the box'
253        1          ' and it is projected onto it')
254         end if
255         do i=1,m
256          IF(x0(i).lt.xlow(i)) x0(i)=xlow(i)
257          IF(x0(i).gt.xup(i)) x0(i)=xup(i)
258         end do
259   c=======================================================================
260         call varind
261
262         penalt=1.d+03
263
264         call linearep
265   c=======================================================================
266         do i=1,m
267          x(i)=x0(l1(i))
268         end do
269   c=======================================================================
270         nshift=2
271          call fv(x,f,objf)
272          objstart=1.d+02*dabs(objf)
273         fpen2=fconst
274   c=======================================================================
275         WRITE(40,*)
276          WRITE(40,77) f
```

```
277    77  FORMAT('Value of the aggregate function at initial point',f21.8)
278   c=======================================================================
279   c#  Calculations of constraint violations at initial point and their
280   c#  description in output file: results.txt
281   c=======================================================================
282         if (nconst1.gt.0) then
283          WRITE(40,*)
284          WRITE(40,2118)
285   2118   FORMAT('Constraint violations at initial point:')
286          WRITE(40,*)
287         end if
288           call func(x0,objf,conlineq,conlinineq,coneq,conineq)
289         vialtot=0.d+00
290         if (numlineq.gt.0) then
291           WRITE(40,*)
292          WRITE(40,2119)
293   2119   FORMAT('Linear equalities:')
294          WRITE(40,*)
295          do i=1,numlineq
296           convial=dabs(conlineq(i))
297           vialtot=vialtot+convial
298           write(40,101) i,convial
299          end do
300         end if
301         if (numlinineq.gt.0) then
302          WRITE(40,*)
303          WRITE(40,2120)
304   2120   FORMAT('Linear inequalities:')
305          WRITE(40,*)
306          do i=1,numlinineq
307           convial=dmax1(0.d+00,conlinineq(i))
308           vialtot=vialtot+convial
309           write(40,101) i,convial
310          end do
311         end if
312         if (numeq.gt.0) then
313          WRITE(40,*)
314          WRITE(40,2121)
315   2121   FORMAT('Nonlinear equalities:')
316          WRITE(40,*)
317          do i=1,numeq
318           convial=dabs(coneq(i))
319           vialtot=vialtot+convial
320           write(40,101) i,convial
321          end do
322         end if
323         if (numineq.gt.0) then
324          WRITE(40,*)
325          WRITE(40,2122)
326   2122   FORMAT('Nonlinear inequalities:')
327          WRITE(40,*)
328          do i=1,numineq
329           convial=dmax1(0.d+00,conineq(i))
330           vialtot=vialtot+convial
331           write(40,101) i,convial
332          end do
333         end if
334   c=======================================================================
335         IF(vialtot.gt.eps1) THEN
336          write(40,*)
337          write(40,*)
338          write(40,122)
339    122   FORMAT('The starting point is infeasible')
340         END if
341         IF(vialtot.le.eps1) THEN
342          write(40,*)
343          write(40,*)
344          write(40,123)
345    123   FORMAT('The strating point is feasible')
346         END if
347         write(40,*)
348         write(40,*)
349         write(40,777) vialtot
```

```
350   777  FORMAT('Sum of constraint violations at initial point:',f20.8)
351        write(40,*)
352        write(40,*)
353        WRITE(40,*)
354   c=======================================================================
355        nshift=1
356        call optimum(x)
357   c=======================================================================
358        nshift=2
359        penalt=1.d+01
360   102  call optimum(x)
361        call fv(x,f,objf)
362        if (fconst.ge.eps1) then
363          penalt=1.d+01*penalt
364          IF(penalt.le.2.d+04) GO TO 102
365        end if
366        call fv(x,f,objf)
367        call restorepoint(x,x3)
368   c=======================================================================
369     7  FORMAT('Value of objective function at final point:',f20.8)
370   771  FORMAT('Value of objective function:',f20.8)
371   118  FORMAT('Sum of constraint violations:',f20.10)
372    18  FORMAT('Variables:')
373   100  FORMAT('Variable',i4,f20.10)
374   101  FORMAT('Constraint',i4,f20.10)
375   c=======================================================================
376         WRITE(40,*)
377         WRITE(40,*)
378         WRITE(40,*)
379         WRITE(40,*) '             FINAL SOLUTION:                '
380         WRITE(40,*)
381         WRITE(40,*)
382         WRITE(40,*)
383         write(40,7) f
384         write(40,*)
385         WRITE(40,*)
386         WRITE(40,18)
387         WRITE(40,*)
388         do i=1,m1
389         WRITE(40,100) i,x3(i)
390         end do
391         WRITE(40,*)
392   c=======================================================================
393   c  Constraint violations at final point
394   c=======================================================================
395         if (nconst1.gt.0) then
396         WRITE(40,*)
397         WRITE(40,2123)
398   2123  FORMAT('Constraint violations at final point:')
399         WRITE(40,*)
400          end if
401            call func(x3,objf,conlineq,conlinineq,coneq,conineq)
402            vialtot=0.d+00
403           if (numlineq.gt.0) then
404           WRITE(40,*)
405          WRITE(40,2124)
406   2124    FORMAT('Linear equalities:')
407           WRITE(40,*)
408            do i=1,numlineq
409            convial=dabs(conlineq(i))
410            vialtot=vialtot+convial
411            write(40,101) i,convial
412           end do
413         end if
414
415         if (numlinineq.gt.0) then
416          WRITE(40,*)
417          WRITE(40,2125)
418   2125    FORMAT('Linear inequalities:')
419           WRITE(40,*)
420            do i=1,numlinineq
421             convial=dmax1(0.d+00,conlinineq(i))
422            vialtot=vialtot+convial
```

```
423           write(40,101) i,convial
424          end do
425         end if
426
427         if (numeq.gt.0) then
428          WRITE(40,*)
429          WRITE(40,2126)
430  2126    FORMAT('Nonlinear equalities:')
431          WRITE(40,*)
432            do i=1,numeq
433             convial=dabs(coneq(i))
434             vialtot=vialtot+convial
435            write(40,101) i,convial
436          end do
437         end if
438         if (numineq.gt.0) then
439          WRITE(40,*)
440          WRITE(40,2127)
441  2127    FORMAT('Nonlinear inequalities:')
442           WRITE(40,*)
443           do i=1,numineq
444             convial=dmax1(0.d+00,conineq(i))
445             vialtot=vialtot+convial
446            write(40,101) i,convial
447           end do
448         end if
449         write(40,*)
450         WRITE(40,118) vialtot
451         WRITE(40,*)
452  c=======================================================================
453         WRITE(40,*)
454         write(40,99) niter
455   99    FORMAT('Total number of iterations',i42)
456         write(40,*)
457         write(40,*)
458         write(40,222) nf
459  222    FORMAT('Total number of objective function evaluations',i22)
460         write(40,*)
461         write(40,*)
462         write(40,*)
463  c      WRITE(40,1034)
464  c1034   FORMAT('No feasible solution has been found')
465         call cpu_time(time2)
466         cputime=time2-time1
467         WRITE(40,1036) cputime
468  1036   FORMAT('Total CPU time in seconds:',f10.4)
469  c=======================================================================
470        CLOSE(40)
471        return
472        end
473  c=======================================================================
474  c Local solver: OPTIMUM
475  c parameters: pw - power (z(\lambda)),
476  c step0   - constant from Armijo algorithm
477  c slinit  - initial value of step in the approximation subgradients
478  c slmin   - final value of the step in the approximation subgradients
479  c maxiter - maximum number of iterations
480  c maxdisc - maximum number of appreoximated subgradients
481  c=======================================================================
482        subroutine optimum(x)
483        implicit double precision (a-h,o-z)
484        PARAMETER(maxvar=2000, maxcon=2000, maxdg=1000,maxit=100000)
485        double precision x(maxvar),x1(maxvar),g(maxvar),v(maxvar)
486       1 ,w(maxdg,maxvar),prod(maxdg,maxdg),z(maxdg),fvalues(maxit)
487       2 ,gprev(maxvar)
488        INTEGER ij(maxdg)
489        common /csize/m,/citer/niter,/cpw/pwt,/cpenf/fconst
490       1 ,/cij/ij,jvertex,/cz/z,/ckmin/kmin,/ceps4/eps4
491       2 ,/cnumneq/numneq,/cfpen5/fpen5,/cfpen1/fpen1,/cnshift/nshift
492  c=======================================================================
493        pw=7.d+00
494        dist1=1.d-07
495        step0=-2.d-01
```

```
      div=6.d-01
      slinit=5.d+00
      eps4=1.d-05
      eps6=1.d-07
      slmin=1.d-08*slinit
      maxiter=50000
      maxdisc=MIN(m+3,200)
      sdif=1.d-05
      IF(m.le.3) mturn=4
      IF(m.gt.3) mturn=5
C=====================================================================
      sl=slinit
      ncycles=dlog(slmin/slinit)/dlog(div)+1
      ncycles1=2*ncycles/3
      ncycles=ncycles+200
      call fv(x,f2,objf)
      fpen2=fconst
      pwt=1.d+00
C=====================================================================
      do ncycle=1,ncycles
       IF(ncycle.gt.1) sl=div*sl
       IF(sl.lt.slmin) go to 4
C=====================================================================
       IF(sl.gt.1.d-01) step0=-2.d-01
       IF((sl.le.1.d-01).AND.(sl.gt.1.d-02)) step0=-3.d-01
       IF((sl.le.1.d-02).AND.(sl.gt.1.d-04)) step0=-4.d-01
       IF(sl.le.1.d-04) step0=-5.d-01
C=====================================================================
       if (numneq.gt.0) then
        IF(ncycle.le.ncycles1) eps4=1.d-04
        IF(ncycle.gt.ncycles1) eps4=1.d-05
       end if
       IF(ncycle.gt.1) pw=9.d-01*pw
       if(pw.lt.1.5d+00) pw=1.5d+00
       t=sl**pw
       pwt=dmin1(pwt,t)
       do i=1,m
        g(i)=1.d+00/dsqrt(DBLE(m))
       end do
       niter1=niter+1
       n4=0
C=====================================================================
       do niter=niter1,maxiter
        n4=n4+1
        f1=f2
        fvalues(niter)=f1
        IF((nshift.eq.1).AND.(f1.lt.1.d-04)) GO TO 4
        fpen1=dmin1(1.d+01*eps4,fpen2+eps4)
C=====================================================================
         if (n4.gt.mturn) then
          mturn2=niter-mturn+1
          ratio1=(fvalues(mturn2)-f1)/(dabs(f1)+1.d+00)
          IF(ratio1.LT.sdif) GO TO 1
         end if
         if (n4.ge.10) then
          mturn2=niter-9
          ratio1=(fvalues(mturn2)-f1)/(dabs(f1)+1.d+00)
          IF(ratio1.LT.1.d-04) GO TO 1
         end if
C=====================================================================
         do ndg=1,maxdisc
             call dgrad(x,sl,g,v,f1,f4,ndg)
             dotprod=0.d+00
             do i=1,m
              dotprod=dotprod+v(i)*v(i)
             end do
             r=dsqrt(dotprod)
             IF(r.lt.eps6) GO TO 1
             IF(ndg.eq.1) then
                          rmean=r
                          kmin=1
                          rmin=r
             END if
```

```
569               IF(ndg.gt.1) then
570                            rmin=dmin1(rmin,r)
571                            IF(r.eq.rmin) kmin=ndg
572                            rmean=((ndg-1)*rmean+r)/ndg
573               END if
574               toler=dmax1(1.d-07,dist1*rmean)
575               do i=1,ndg-1
576                prod(ndg,i)=0.d+00
577                do j=1,m
578                 prod(ndg,i)=prod(ndg,i)+w(i,j)*v(j)
579                end do
580                prod(i,ndg)=prod(ndg,i)
581               end do
582               prod(ndg,ndg)=dotprod
583  c=============================================================================
584               do i=1,m
585                w(ndg,i)=v(i)
586               end do
587               call wolfe(ndg,prod)
588  c=============================================================================
589               do i=1,m
590                v(i)=0.d+00
591                do j=1,jvertex
592                 v(i)=v(i)+w(ij(j),i)*z(j)
593                END do
594               END do
595  c=============================================================================
596               r=0.d+00
597               do i=1,m
598                r=r+v(i)*v(i)
599                end do
600               r=dsqrt(r)
601               if(r.lt.toler) GO TO 1
602  c=============================================================================
603                do i=1,m
604                 g(i)=-v(i)/r
605                 x1(i)=x(i)+sl*g(i)
606                end do
607  c=============================================================================
608                if(ndg.gt.1) then
609                          rdif=0.d+00
610                          do i=1,m
611                           rdif=rdif+(gprev(i)-g(i))**2
612                          end do
613                          rdif=dsqrt(rdif)
614                          if(rdif.LT.eps6) GO TO 1
615                          rdif=rold-r
616                          IF(rdif.LT.toler) GO TO 1
617                END if
618  c=============================================================================
619                do i=1,m
620                gprev(i)=g(i)
621                end do
622                rold=r
623                call fv(x1,f4,objf)
624                f3=(f4-f1)/sl
625                decreas=step0*r
626                if((fconst.le.fpen1).AND.(f3.lt.decreas)) then
627                           fpen3=fconst
628  c=============================================================================
629                           call armijo(x,g,f1,f5,f4,sl,step,r,fpen3)
630                           f2=f5
631                           fpen2=fpen5
632                           do i=1,m
633                            x(i)=x(i)+step*g(i)
634                           end do
635                           print 21, niter,ndg,f2,sl,fpen2
636                           GO TO 2
637  21                       format(I6,I4,3f16.8)
638                end if
639              END do
640  c=============================================================================
641            go to 1
```

```
642   2      END do
643   1     END do
644   4     return
645         end
646   c=======================================================================
647   c#  Subroutines Wolfe and Equations solves quadratic problem to find descent direction.
648   c # The number of iteration is restricted by J MAX
649   c#  The vector Z contains weights of extreme points.
650   c#  The matrix A is main matrix in the system of linear equations from Wolfe algorithm.
651   c#  Matrix PROD contains inner products of subgradients
652   c=======================================================================
653         subroutine wolfe(ndg,prod)
654         implicit double precision (a-h,o-z)
655         PARAMETER(maxvar=2000, maxcon=2000, maxdg=1000)
656         common /csize/m,/w01/a,/cij/ij,jvertex,/cz/z,/ckmin/kmin
657         INTEGER ij(maxdg)
658         double precision z(maxdg),z1(maxdg),a(maxdg,maxdg)
659        1 ,prod(maxdg,maxdg)
660         j9=0
661         jmax=200*ndg
662         jvertex=1
663         ij(1)=kmin
664         z(1)=1.d+00
665   c=======================================================================
666   c#  To calculate X
667   c=======================================================================
668    1    r=0.d+00
669         do i=1,jvertex
670          do j=1,jvertex
671           r=r+z(i)*z(j)*prod(ij(i),ij(j))
672          end do
673         end do
674         IF(ndg.eq.1) GO TO 5
675   c=======================================================================
676   c  To calculate <X,P_J> and J
677   c=======================================================================
678         t0=1.d+12
679         do i=1,ndg
680           t1=0.d+00
681           do j=1,jvertex
682             t1=t1+z(j)*prod(ij(j),i)
683           end do
684            if(t1.lt.t0) then
685                         t0=t1
686                         kmax=i
687           end if
688         end do
689   c=======================================================================
690   c#  First stopping criterion
691   c=======================================================================
692         rm=prod(kmax,kmax)
693         do j=1,jvertex
694          rm=dmax1(rm,prod(ij(j),ij(j)))
695         end do
696         r2=r-1.d-12*rm
697         if(t0.gt.r2) GO TO 5
698   c=======================================================================
699   c#  Second stopping criterion
700   c=======================================================================
701         do i=1,jvertex
702          if(kmax.eq.ij(i)) GO TO 5
703         end do
704   c=======================================================================
705   c# Step 1(e) from Wolfe's algorithm
706   c=======================================================================
707         jvertex=jvertex+1
708         ij(jvertex)=kmax
709         z(jvertex)=0.d+00
710   c=======================================================================
711    2    do i=1,jvertex
712          do j=1,jvertex
713           a(i,j)=1.d+00+prod(ij(i),ij(j))
714          end do
```

```
715         end do
716         j9=j9+1
717         if(j9.gt.jmax) GO TO 5
718         call equations(jvertex,z1)
719         do i=1,jvertex
720          if(z1(i).le.1.d-10) go to 3
721         end do
722         do i=1,jvertex
723          z(i)=z1(i)
724         end do
725         go to 1
726     3   teta=1.d+00
727         do i=1,jvertex
728          z5=z(i)-z1(i)
729          if(z5.gt.1.d-10) teta=dmin1(teta,z(i)/z5)
730         end do
731         do i=1,jvertex
732          z(i)=(1.d+00-teta)*z(i)+teta*z1(i)
733          if(z(i).le.1.d-10) then
734                             z(i)=0.d+00
735                             kzero=i
736          end if
737         end do
738         j2=0
739         do i=1,jvertex
740          IF(i.ne.kzero) then
741                         j2=j2+1
742                         ij(j2)=ij(i)
743                         z(j2)=z(i)
744          END if
745         end do
746         jvertex=j2
747         go to 2
748     5   return
749         end
750          subroutine equations(n,z1)
751           implicit double precision (a-h,o-z)
752           PARAMETER(maxvar=2000, maxcon=2000, maxdg=1000)
753          common /w01/a
754         double precision a(maxdg,maxdg),z1(maxdg),b(maxdg,maxdg)
755         do i=1,n
756          do j=1,n
757           b(i,j)=a(i,j)
758          end do
759          b(i,n+1)=1.d+00
760         end do
761         do i=1,n
762          r=b(i,i)
763          do j=i,n+1
764           b(i,j)=b(i,j)/r
765          end do
766          do j=i+1,n
767           do k=i+1,n+1
768            b(j,k)=b(j,k)-b(i,k)*b(j,i)
769           end do
770          end do
771         end do
772         z1(n)=b(n,n+1)
773         do i=1,n-1
774           k=n-i
775           z1(k)=b(k,n+1)
776           do j=k+1,n
777            z1(k)=z1(k)-b(k,j)*z1(j)
778           END do
779         end do
780         z2=0.d+00
781         do i=1,n
782          z2=z2+z1(i)
783         end do
784         do i=1,n
785          z1(i)=z1(i)/z2
786         end do
787         return
```

```
788         end
789   c=======================================================================
790   c# Subroutine discrete gradient: calculates approximate subgradients
791   c=======================================================================
792         subroutine dgrad(x,sl,g,dg,f1,f4,ndg)
793         implicit double precision (a-h,o-z)
794         PARAMETER(maxvar=2000, maxcon=2000, maxdg=1000)
795         double precision x1(maxvar),g(maxvar),x(maxvar),dg(maxvar)
796         common /csize/m,/cpw/pwt,/cnshift/nshift
797         a1=0.d+00
798         do k=1,m
799          d1=dabs(g(k))
800          if (a1.lt.d1) then
801            a1=d1
802            imax=k
803          end if
804         end do
805         do k=1,m
806          x1(k)=x(k)+sl*g(k)
807         end do
808         IF(ndg.gt.1) r2=f4
809         IF(ndg.eq.1) call fv(x1,r2,objf)
810         flambda=r2
811         dsum=0.d+00
812         do k=1,m
813           if(k.ne.imax) then
814              r3=r2
815              x1(k)=x1(k)+pwt
816               call fv(x1,r2,objf)
817              dg(k)=(r2-r3)/pwt
818              dsum=dsum+dg(k)*sl*g(k)
819           END if
820         end do
821         dg(imax)=(flambda-f1-dsum)/(sl*g(imax))
822         return
823         end
824   c=======================================================================
825   c Line search (Armijo-type)
826   c=======================================================================
827         subroutine armijo(x,g,f1,f5,f4,sl,step,r,fpen3)
828         implicit double precision (a-h,o-z)
829         PARAMETER(maxvar=2000, maxcon=2000, maxdg=1000)
830         common /csize/m,/cpenf/fconst,/cfpen5/fpen5,/cfpen1/fpen1
831        1 ,/cnstep/nstep
832         double precision x(maxvar),g(maxvar),x1(maxvar),fline(1000)
833        1 ,sline(1000),flpen(1000)
834         k=1
835         fline(k)=f4
836         sline(k)=sl
837         flpen(k)=fpen3
838         step=sl
839     1   step=step+sl
840         do i=1,m
841          x1(i)=x(i)+step*g(i)
842         end do
843         call fv(x1,f5,objf)
844         nstep=nstep+1
845         fpen5=fconst
846         f3=f5-f1+1.d-03*step*r
847         IF((fpen5.gt.fpen1).OR.(f3.gt.0.d+00)) GO TO 3
848         IF((fpen5.le.fpen1).AND.(f3.le.0.d+00)) GO TO 2
849   2     k=k+1
850         fline(k)=f5
851         sline(k)=step
852         flpen(k)=fpen5
853         IF(k.gt.100) GO TO 3
854         GO TO 1
855   3     continue
856         step=sline(1)
857         f5=fline(1)
858         fpen5=flpen(1)
859         do i=2,k
860          if (f5.gt.fline(i)) then
```

```
861             f5=fline(i)
862             step=sline(i)
863             fpen5=flpen(i)
864           end if
865         end do
866         return
867         end
868   c=======================================================================
869   c# This subroutine gives values of aggregate function
870   c=======================================================================
871         subroutine fv(x1,f,objf)
872         implicit double precision (a-h,o-z)
873         PARAMETER(maxvar=2000, maxcon=2000, maxdg=1000)
874         double precision x(maxvar),x1(maxvar),x2(maxvar),coneq(maxcon)
875        1 ,conineq(maxcon),xlow(maxvar),xup(maxvar),c1(maxcon,maxvar)
876        2 ,conlinineq(maxcon)
877         INTEGER l1(maxvar),l3(maxvar),indlow(maxvar),indup(maxvar)
878         COMMON /cnf/nf,/csize/m,/cconeq/numeq,/cc1/c1,/cconin/numineq
879        1 ,/cpenf/fconst,/cpen/penalt,/cobstart/objstart,/cnshift/nshift
880        2 ,/cbound/xlow,xup,/cl3/l3,n2,/cl1/l1,/csize1/m1
881        3 ,/cindicator/indlow,indup,/cindbound/indbound
882        4 ,/cconlinineq/numlinineq
883   c=======================================================================
884         do i=1,m
885            x(l1(i))=x1(i)
886         end do
887         do i=n2,1,-1
888            x(l3(i))=c1(i,m1+1)
889         do j=1,m1
890            x(l3(i))=x(l3(i))+c1(i,j)*x(j)
891          end do
892         end do
893         ind1=0
894         IF(indbound.eq.0) THEN
895           do i=1,m1
896            IF((x(i).lt.xlow(i)).OR.(x(i).gt.xup(i))) then
897               ind1=1
898               GO TO 1
899            END if
900           end do
901         END if
902       1 IF(ind1.eq.0) then
903            call func(x,objf,conlineq,conlinineq,coneq,conineq)
904              nf=nf+1
905         END if
906         IF(ind1.eq.1) then
907               do i=1,m
908                x2(i)=x(i)
909                IF(x2(i).lt.xlow(i)) x2(i)=xlow(i)
910                IF(x2(i).gt.xup(i)) x2(i)=xup(i)
911               end do
912              call func(x2,objf,conlineq,conlinineq,coneq,conineq)
913              nf=nf+1
914              objstart=objf
915              call func(x,objf,conlineq,conlinineq,coneq,conineq)
916              nf=nf+1
917              objf=objstart
918         END if
919   c=======================================================================
920         f3=0.d+00
921         f4=0.d+00
922         IF(indbound.eq.0) THEN
923          do i=1,m
924           IF((indlow(l1(i)).eq.0).AND.(indup(l1(i)).eq.0))
925        1    f3=f3+dmax1(0.d+00,xlow(l1(i))-x(l1(i)))
926        2     +dmax1(0.d+00,x(l1(i))-xup(l1(i)))
927              IF((indlow(l1(i)).eq.1).AND.(indup(l1(i)).eq.0))
928        1    f3=f3+dmax1(0.d+00,x(l1(i))-xup(l1(i)))
929              IF((indlow(l1(i)).eq.0).AND.(indup(l1(i)).eq.1))
930        1    f3=f3+dmax1(0.d+00,xlow(l1(i))-x(l1(i)))
931           end do
932           do i=1,n2
933            IF((indlow(l3(i)).eq.0).AND.(indup(l3(i)).eq.0))
```

```
      1    f4=f4+dmax1(0.d+00,xlow(l3(i))-x(l3(i)))
      2     +dmax1(0.d+00,x(l3(i))-xup(l3(i)))
          IF((indlow(l3(i)).eq.1).AND.(indup(l3(i)).eq.0))
      1    f4=f4+dmax1(0.d+00,x(l3(i))-xup(l3(i)))
          IF((indlow(l3(i)).eq.0).AND.(indup(l3(i)).eq.1))
      1    f4=f4+dmax1(0.d+00,xlow(l3(i))-x(l3(i)))
        end do
       END if
       f1=0.d+00
       do i=1,numeq
        f1=f1+dmax1(0.d+00,coneq(i)-1.d-07,-coneq(i)-1.d-07)
       end do
       f2=0.d+00
       do i=1,numineq
        f2=f2+dmax1(0.d+00,conineq(i))
       end do
       f6=0.d+00
       do i=1,numlinineq
        f6=f6+dmax1(0.d+00,conlinineq(i))
       end do
       fbox=f4+f3
       f5=f1+f2
       fconst=f5+fbox+f6
       IF(nshift.eq.1) THEN
         f=f5+f6+fbox
         fconst=0.d+00
       END if
       IF(nshift.eq.2) f=objf+penalt*(f5+fbox+f6)
c=======================================================================
       return
       end
c=======================================================================
c This subroutine solves the system of linear equations and transforms the search space.
c=======================================================================
       subroutine linearep
       PARAMETER(maxvar=2000, maxcon=2000)
       implicit double precision (a-h,o-z)
       DOUBLE PRECISION conlineq(maxcon),x(maxvar),x2(maxvar)
      1 ,a(maxcon,maxvar),a1(maxcon,maxvar),c1(maxcon,maxvar)
      2 ,coneq(maxcon),conineq(maxcon),conlinineq(maxcon)
       INTEGER l1(maxvar),l2(maxvar),l3(maxvar)
       COMMON /cconlineq/numlineq,/csize/m,/csize1/m1,/cl3/l3,n2
      1 ,/cl1/l1,/cc1/c1
       h=1.d-01
        do i=1,m
         x(i)=0.d+00
        end do
       n=numlineq
       do i=1,n
         do j=1,m+1
          a(i,j)=0.d+00
         end do
       end do
       do i=1,n
        call func(x,objf,conlineq,conlinineq,coneq,conineq)
        a(i,m+1)=conlineq(i)
        do j=1,m
         x(j)=h
         call func(x,objf,conlineq,conlinineq,coneq,conineq)
         a(i,j)=(conlineq(i)-a(i,m+1))/h
         x(j)=0.d+00
        end do
       end do
       do i=1,n
        do j=1,m+1
         a1(i,j)=a(i,j)
        end do
       end do
       n1=0
       n2=0
       do j=1,m
        do i=1,n
         do k=1,n2
```

```
1007            IF(i.eq.l2(k)) GO TO 4
1008           end do
1009           IF(dabs(a1(i,j)).ge.1.d-10) GO TO 1
1010     4     end do
1011           n1=n1+1
1012           l1(n1)=j
1013           GO TO 2
1014     1     n2=n2+1
1015           l2(n2)=i
1016           l3(n2)=j
1017           ncurrent=i
1018           r=a1(ncurrent,j)
1019           do k1=1,m+1
1020              a1(ncurrent,k1)=a1(ncurrent,k1)/r
1021           end do
1022           dpen0=0.d+00
1023           do k1=1,m+1
1024            c1(n2,k1)=-a1(ncurrent,k1)
1025            end do
1026           c1(n2,j)=0.d+00
1027           do i=1,n
1028             do k1=1,n2
1029              IF(i.eq.l2(k1)) GO TO 5
1030             end do
1031            b1=a1(i,j)
1032            do k1=1,m+1
1033              a1(i,k1)=a1(i,k1)-b1*a1(ncurrent,k1)
1034            end do
1035     5     end do
1036     2     end do
1037           m1=m
1038           m=n1
1039           return
1040           end
1041     c=======================================================================
1042     c# This subroutine makes reverse transformation returning points to the original space.
1043     c=======================================================================
1044           subroutine restorepoint(x1,x)
1045           implicit double precision (a-h,o-z)
1046           PARAMETER(maxvar=2000, maxcon=2000, maxdg=1000)
1047           double precision x(maxvar),x1(maxvar),c1(maxcon,maxvar)
1048           INTEGER l1(maxvar),l3(maxvar)
1049           common /csize/m,/cl3/l3,n2,/cl1/l1,/csize1/m1,/cc1/c1
1050           do i=1,m
1051             x(l1(i))=x1(i)
1052           end do
1053           do i=n2,1,-1
1054             x(l3(i))=c1(i,m1+1)
1055              do j=1,m1
1056               x(l3(i))=x(l3(i))+c1(i,j)*x(j)
1057             end do
1058            end do
1059           return
1060           end
1061     c=======================================================================
1062     c# This subroutine identifies variables with box constraints
1063     c=======================================================================
1064           subroutine varind
1065           implicit double precision (a-h,o-z)
1066           PARAMETER(maxvar=2000)
1067           double precision xlow(maxvar),xup(maxvar)
1068           INTEGER indlow(maxvar),indup(maxvar)
1069           COMMON /csize1/m,/cbound/xlow,xup,/cindicator/indlow,indup
1070          1 ,/cindbound/indbound
1071           do i=1,m
1072            IF(xlow(i).LT.-1.d+05) indlow(i)=1
1073            IF(xlow(i).GE.-1.d+05) indlow(i)=0
1074            IF(xup(i).gt.1.d+05) indup(i)=1
1075            IF(xup(i).le.1.d+05) indup(i)=0
1076           end do
1077           indbound=1
1078           do i=1,m
1079             IF((indlow(i).eq.0).or.(indup(i).eq.0)) THEN
```

```
1080            indbound=0
1081          GO TO 1
1082         END if
1083        end do
1084    1   RETURN
1085        END
```

Appendix C: Lasso/glmnet source R-code.

Method №2: Computation of LASSO penalisation with `nfold=10` for Marvin blasthole data, packages: `mvtnorm,glmnet,matrix,foreach`.

```
14
15
16 rm(list=ls(all=TRUE))   # clean all data
17 #install.packages("mvtnorm") # install packages
18 #install.packages("glmnet")  # if necessary
19 require(mvtnorm) # load packages
20 require(glmnet)
21 ##############################################################################
22 # set your working directory
23 setwd("/Igor/data/windows/Documents/Codes/Rcodes/")
24 #
25 set.seed(123)
26 ##############################################################################
27 dat<-read.table("Marvin1.csv") # load data from file
28 dat<-matrix(data=dat$V1, ncol=7, byrow=TRUE) # process data
29 dat<-matrix(data=as.numeric(dat[, -1]), ncol=6, byrow=FALSE)
30 dat<-dat[, -3]
31 colnames(dat)<-(c("x", "y", "z1", "z2", "au")) # the observations consist of
32 #  x, y, z1  beginning of segment
33 #  x, y, z2  end of segment
34 # au  gold concentration
35 ##############################################################################
36 resp<-dat[, 5] # take gold concentration as response variable
37 obs<-dat[, 1:4]
38 ##############################################################################
39 centers<-matrix(data=0, nrow=700, ncol=3)     # data for ore bodies
40 # first three columns for mean
41 # other nine for variance matrix
42 sigmas<-rep(1, nrow(centers))
43 ##############################################################################
44 # create potential ore bodies randomly
45 ##############################################################################
46 centers[, 1]<-runif(n=nrow(centers), min=min(obs[,1]), max=max(obs[, 1]))
47 centers[, 2]<-runif(n=nrow(centers), min=min(obs[,2]), max=max(obs[, 2]))
48 centers[, 3]<-runif(n=nrow(centers), min=min(obs[,3:4]), max=max(obs[, 3:4]))
49 #
50 sigmas<-runif(n=length(sigmas), min=3, max=3)
51 #
52 ##############################################################################
53 make_model_matrix<-function(obs, centers, sigmas)
54 {
55 ##############################################################################
56 # create model matrix
57 ##############################################################################
58 # each segment contains gold concentration derived from random normal
59 # density originated from ore body
60 # rows stand for observations (gold concentration in segments)
61 # columns stand for ore bodies
62 # explanatory coefficient are obained from density integral along each segment
63 # which is calculated by trapezoidal rule
64 ##############################################################################
65 mod_mat<-matrix(data=0, nrow=nrow(obs), ncol=nrow(centers))
66 for (j in 1:ncol(mod_mat)) {
67 center<-matrix(centers[j, 1:3], byrow=TRUE, nrow=nrow(obs), ncol=3)
68 p1<-obs[, c(1,2,3)]-center
69 p2<-obs[, c(1,2,4)]-center
70 mod_mat[,j]<-(apply(FUN=dmvnorm, X=p1,MARGIN=1,sigma=diag(nrow=3,ncol=3,x=sigmas[j]^2))+
71 apply(FUN=dmvnorm, X=p2, MARGIN = 1, sigma=diag(nrow=3, ncol=3, x=sigmas[j]^2) ) )/2
72 }
73 return(mod_mat)
```

```
74  }
75  #
76  library(glmnet) # load library
77  #
78  #*****************************************************************************
79  #  BEGIN RUN THIS SEVERAL TIMES
80  #*****************************************************************************
81  #
82  max_iter<-7
83  #
84  for (k in 1:max_iter)   {
85  if (k>1)
86  {
87  sigmas<-sigmas_new
88  centers<-centers_new
89
90  if (k<max_iter-3) {
91
92  sigmas<-c(sigmas, sigmas*runif(n=length(sigmas), min=0.9, max=1.1))
93
94  centers<-rbind(centers, centers+ cbind(
95  runif(n=nrow(centers), min=-1, max=1),
96  runif(n=nrow(centers), min=-1, max=1),
97  runif(n=nrow(centers), min=-1, max=1)
98  )
99  )
100 }
101 }
102 ###############################################################################
103 #
104 # Create model matrix
105 #
106 ###############################################################################
107 model_mat<-make_model_matrix(obs=obs, centers=centers, sigmas=sigmas)
108 ###############################################################################
109 #
110 #  Make explanatory matrix
111 #
112 ###############################################################################
113 EXPL<-cbind(1, model_mat)   # define model matrix
114 colnames(EXPL)<-c( "i0", paste("C", rep(1:ncol(model_mat)), sep=""))
115 index<-1:ncol(EXPL)-1
116 names(index)<-colnames(EXPL)
117 ###############################################################################
118 #
119 #  Make glm fit
120 #
121 ###############################################################################
122 #
123 cvobj<-cv.glmnet(x=EXPL,y=resp, intercept=FALSE) # try with gaussian model
124 ###############################################################################
125 #
126 # Extract non-zero coefficients
127 #
128 ###############################################################################
129 co<-coef(cvobj)      # most of the coefficients are zero
130 non_zero<-which(co!=0)
131 non_zero_names<-names(co[non_zero,])
132 data_index<-index[non_zero_names]
133 sigmas_new<-sigmas[data_index]
134 centers_new<-centers[data_index,]
135 ###############################################################################
136 plot(centers_new[,1], centers_new[,2])
137 print(sigmas_new)
138 print(non_zero_names)
139 ###############################################################################
140 # Watch the animation run
141 ###############################################################################
142 }
143 betas<-co[names(co[,1 ])!="(Intercept)", 1]
144
145 zlevels<-seq(from=min(dat[,3]), to=max(dat[, 3]), length=5)
146 res_dim1<-40
```

```
147 res_dim2<-40
148
149 xgrid<-seq(from=min(dat[,1]), to=max(dat[, 1]), length=res_dim1)
150 ygrid<-seq(from=min(dat[,2]), to=max(dat[, 2]), length=res_dim2)
151
152 for (i in 1:length(zlevels))
153 {
154 L<-cbind(expand.grid(X=xgrid, Y=ygrid), zlevels[2], zlevels[i] )
155 M<-cbind(1, make_model_matrix(obs=L, centers=centers, sigmas=sigmas))
156
157 zgrid<-matrix(nrow=res_dim1, ncol=res_dim2, data=M%*%betas, byrow=FALSE)
158
159 pdf(paste("/Igor/data/windows/Documents/Codes/Rcodes/
160 ", i, ".pdf", sep="" ) , width=7, height=5, paper='special')
161 filled.contour(x = xgrid,y = ygrid,
162 z = zgrid, color.palette = terrain.colors,
163 plot.title = title(main =paste("gold concentration at depth",zlevels[i],"m"),
164 xlab = "meters x-direction",ylab = "meters y-direction"),
165 key.title = title(main="gold \n  g/m3"),
166 key.axes = axis(4, round(seq(from=min(zgrid), to=max(zgrid), length=20),
167
168 digits = 2) )
169 )
170 dev.off()
171
172 filled.contour(x = xgrid,y = ygrid,
173                z = zgrid, color.palette = terrain.colors,
174                plot.title = title(main = paste("Predicted AU g / m3 depth=",
175                zlevels[i], "metres", cex.main=""),
176                                   xlab = "EAST",ylab = "NORTH"),
177                                   key.title = title(main="g/m3"),
178                key.axes = axis(4, round(seq(from=min(zgrid),
179                                   to=max(zgrid), length=20), digits = 2) )
180 )
181 }
```

Appendix D: Matrices of polar Kernel densities

Comparison of matrices before conversion into polar Kernels.

1. Matrix 1 is a comparison of Prototype P5 (5 hyperplanes) versus Prototype P10A (after adding 5 more hyperplanes to existing 5).
2. Matrix 2 is a comparison of Prototype P7 (7 hyperplanes) versus Prototype P10A (after adding 3 more hyperplanes to existing 7).

The results indicate that the compared pairs do not have a significant change in the shape of their matrices. The effect of adding extra hyperplanes is insignificant.

Matrix 1. P5 versus P10A: matrices of mean density distribution

P5 KERNEL DENSITY MATRIX

7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06	7.71E-06
1.25E-05	8.50E-06	2.53E-06	4.11E-07	5.08E-08	6.92E-09	1.49E-09	6.24E-10	3.98E-10	4.12E-10	1.07E-09	5.76E-09	4.17E-08	2.77E-07	1.21E-06	2.67E-06	2.59E-06	1.10E-06	2.32E-07	3.17E-08	3.87E-09	6.00E-10	1.61E-10	9.47E-11	1.34E-10	4.44E-10	2.86E-09	2.72E-08	2.68E-07	1.93E-06	7.43E-06	1.25E-05
1.47E-05	3.08E-06	5.49E-08	1.18E-08	1.85E-08	4.76E-09	5.64E-10	7.71E-11	2.19E-11	1.52E-11	1.90E-11	2.15E-11	8.58E-12	3.59E-11	1.48E-08	4.02E-07	3.85E-07	1.27E-08	2.46E-11	7.82E-15	1.51E-18	7.24E-22	3.03E-24	2.77E-25	8.46E-25	7.52E-23	9.83E-20	6.19E-16	4.73E-12	1.06E-08	2.04E-06	1.47E-05
1.49E-05	6.58E-07	3.82E-08	2.28E-08	1.08E-10	1.29E-11	4.50E-13	1.05E-14	4.83E-16	3.83E-17	7.06E-18	1.23E-16	2.00E-14	1.01E-13	1.33E-11	2.47E-08	2.36E-08	1.10E-11	8.47E-18	1.07E-25	4.30E-34	1.28E-41	4.88E-47	1.93E-49	1.88E-48	3.67E-44	3.09E-37	9.25E-29	4.48E-20	1.43E-12	1.86E-07	1.49E-05
1.49E-05	1.33E-07	5.70E-08	1.73E-09	9.74E-11	5.30E-11	3.90E-12	4.74E-14	7.76E-16	2.36E-17	6.89E-19	1.29E-21	2.49E-21	6.84E-17	1.01E-15	6.00E-10	5.75E-10	6.82E-16	9.06E-27	7.88E-41	6.83E-44	3.67E-34	3.56E-29	5.97E-28	5.40E-30	2.10E-35	7.61E-45	9.38E-47	2.44E-31	5.17E-18	6.26E-09	1.49E-05
1.49E-05	6.80E-08	8.72E-09	3.59E-09	2.83E-09	1.24E-10	7.00E-12	1.03E-12	2.19E-14	3.60E-17	1.09E-20	2.17E-23	1.41E-25	1.36E-22	8.89E-19	5.65E-12	5.43E-12	2.99E-21	2.96E-38	7.91E-41	7.01E-26	6.19E-18	9.23E-15	1.23E-14	1.02E-16	3.37E-21	2.76E-29	6.81E-43	8.21E-46	5.49E-25	8.40E-11	1.49E-05
1.45E-05	1.80E-07	7.56E-08	2.82E-08	6.91E-09	3.48E-09	1.88E-10	2.05E-13	5.50E-16	1.19E-17	1.45E-19	2.65E-24	2.50E-28	1.10E-30	3.63E-22	2.04E-14	1.96E-14	9.08E-28	4.48E-50	2.39E-26	4.85E-14	1.44E-09	1.98E-09	5.72E-11	6.81E-13	1.54E-15	8.77E-21	1.18E-31	1.12E-51	1.84E-33	4.49E-13	1.45E-05
1.36E-05	1.88E-07	8.51E-08	1.43E-08	3.65E-08	2.88E-10	6.44E-17	2.78E-23	3.87E-24	3.36E-27	3.77E-25	8.59E-25	5.40E-31	4.57E-34	1.01E-26	2.80E-17	2.69E-17	1.90E-35	1.11E-37	3.65E-16	2.26E-08	4.67E-09	3.50E-13	6.03E-17	1.59E-18	2.03E-18	2.44E-19	1.17E-25	1.64E-42	1.79E-43	8.58E-16	1.36E-05
1.31E-05	8.56E-08	1.19E-08	6.94E-08	7.96E-09	2.19E-18	2.32E-14	1.04E-11	4.43E-13	1.10E-17	1.64E-27	1.94E-31	3.38E-31	3.55E-38	1.88E-32	1.45E-20	1.40E-20	2.74E-44	8.11E-28	2.83E-10	7.08E-09	2.09E-16	5.13E-26	1.44E-32	1.30E-33	7.63E-30	5.96E-25	6.77E-25	9.91E-37	4.21E-55	5.43E-19	1.31E-05
1.32E-05	9.95E-08	1.93E-08	8.87E-08	1.19E-14	7.91E-13	3.79E-08	2.88E-09	1.15E-11	2.59E-14	7.33E-20	4.36E-34	1.53E-35	1.15E-42	2.36E-39	2.83E-24	2.73E-24	2.69E-54	1.74E-20	1.11E-08	1.49E-15	1.29E-31	1.62E-43	4.80E-33	1.48E-29	3.99E-34	1.27E-37	2.26E-29	2.50E-34	2.12E-61	1.29E-22	1.32E-05
1.34E-05	1.88E-07	1.22E-09	4.24E-09	1.10E-18	3.00E-08	1.77E-10	2.79E-16	6.75E-20	5.01E-20	4.53E-20	1.92E-27	3.52E-44	5.21E-44	2.01E-47	2.07E-28	2.00E-28	8.04E-52	1.09E-15	2.22E-11	2.12E-28	1.09E-47	8.00E-32	5.32E-23	3.58E-18	1.09E-18	4.69E-30	4.34E-39	2.62E-35	2.05E-57	1.85E-26	1.34E-05
1.35E-05	2.19E-07	8.34E-09	9.85E-14	3.34E-12	2.45E-10	4.58E-20	1.72E-32	8.98E-38	8.03E-35	3.87E-28	5.71E-27	1.00E-43	3.10E-47	9.97E-56	5.70E-33	5.51E-33	1.11E-45	2.00E-13	2.24E-18	2.03E-47	6.24E-40	3.27E-29	1.34E-22	3.03E-16	8.47E-12	3.97E-16	6.06E-40	1.14E-39	3.24E-55	1.65E-30	1.35E-05
1.36E-05	9.04E-08	2.73E-08	3.99E-16	2.86E-09	9.27E-15	9.14E-25	1.06E-31	1.88E-42	4.53E-55	4.57E-44	1.14E-32	1.09E-39	5.41E-53	5.94E-61	5.86E-38	5.66E-38	1.04E-40	1.08E-13	1.14E-29	2.15E-55	4.93E-40	1.10E-35	7.60E-32	9.01E-24	1.89E-13	2.95E-09	2.52E-24	2.07E-47	8.39E-55	4.54E-35	1.36E-05
1.32E-05	1.09E-08	1.56E-09	6.82E-19	2.99E-08	9.76E-19	2.82E-22	1.66E-30	3.27E-42	4.04E-55	7.49E-68	1.54E-44	6.04E-40	2.76E-61	2.72E-67	2.25E-43	2.18E-43	6.54E-37	1.70E-16	2.97E-45	1.14E-53	5.39E-48	3.08E-51	9.78E-51	9.39E-41	1.21E-23	1.91E-09	6.06E-14	1.40E-55	3.55E-56	3.01E-40	1.32E-05
1.30E-05	7.49E-10	2.38E-12	7.20E-14	1.91E-12	1.33E-14	1.18E-21	9.77E-37	1.13E-50	1.23E-60	6.74E-73	1.40E-62	1.69E-44	2.88E-68	2.37E-74	3.24E-49	3.12E-49	2.78E-34	7.87E-22	3.90E-65	4.06E-58	8.17E-64	7.12E-76	2.85E-79	3.43E-67	2.21E-42	1.09E-16	8.41E-09	1.22E-40	2.45E-59	5.54E-46	1.30E-05
1.29E-05	1.21E-09	1.10E-09	6.41E-10	6.36E-13	3.74E-13	2.01E-29	5.09E-45	3.05E-60	9.71E-65	3.41E-79	8.52E-87	2.40E-53	1.35E-65	1.67E-78	1.73E-55	1.67E-55	7.94E-33	1.07E-29	4.38E-77	9.75E-69	1.71E-87	1.36E-109	1.88E-117	4.40E-103	1.16E-69	5.44E-31	6.75E-09	4.44E-29	2.78E-64	3.86E-52	1.29E-05
1.26E-05	9.02E-10	3.68E-08	2.70E-09	2.57E-09	1.71E-16	1.14E-30	9.92E-43	9.92E-39	5.17E-39	1.05E-63	5.00E-99	1.72E-66	1.84E-65	4.87E-83	3.46E-62	3.34E-62	1.53E-32	4.23E-40	2.02E-79	1.58E-85	4.97E-119	2.15E-152	2.81E-165	1.98E-148	1.73E-105	2.38E-52	3.13E-14	6.70E-21	5.13E-71	1.03E-58	1.26E-05
1.20E-05	1.38E-10	2.04E-08	6.49E-11	8.26E-09	1.54E-19	3.79E-28	1.31E-41	7.53E-27	2.28E-22	3.37E-38	4.07E-95	6.28E-84	7.38E-68	9.65E-89	2.58E-69	2.49E-69	1.99E-33	4.92E-53	4.75E-86	1.72E-108	2.00E-158	2.80E-204	9.49E-223	3.12E-203	7.44E-150	9.13E-81	8.37E-25	4.21E-16	1.55E-79	1.04E-65	1.20E-05
1.13E-05	5.18E-12	1.88E-10	4.56E-08	6.95E-09	1.02E-17	3.08E-33	1.56E-38	1.29E-24	8.29E-15	1.49E-20	1.23E-65	1.16E-105	8.67E-73	1.29E-95	7.16E-77	6.89E-77	1.74E-35	1.67E-68	5.64E-97	1.26E-137	1.11E-205	3.02E-265	7.26E-290	1.73E-267	9.14E-203	3.07E-116	1.29E-40	1.10E-14	5.79E-76	1.89E-72	1.13E-05
1.11E-05	1.64E-12	1.76E-10	1.49E-08	4.71E-14	2.12E-16	7.16E-47	6.55E-45	5.03E-32	2.49E-16	9.10E-11	2.50E-42	1.08E-131	2.98E-80	1.16E-103	7.39E-85	7.12E-85	1.03E-38	1.67E-86	3.39E-112	6.23E-173	8.56E-261	0	0	0	3.21E-264	9.03E-159	1.16E-61	1.20E-16	3.49E-66	5.36E-79	1.11E-05
1.17E-05	2.89E-12	1.39E-08	4.49E-08	4.49E-10	4.10E-22	4.74E-66	9.63E-61	4.43E-49	6.21E-27	7.71E-09	3.42E-25	3.38E-116	3.01E-90	7.03E-113	2.84E-93	2.74E-93	4.10E-43	4.88E-107	1.04E-131	2.07E-214	0	0	0	0	0	2.33E-208	5.96E-88	5.40E-22	3.45E-58	3.68E-86	1.17E-05
1.20E-05	1.10E-11	3.23E-08	5.90E-09	1.00E-09	6.94E-35	4.98E-82	4.97E-86	8.83E-76	1.28E-46	9.03E-15	3.15E-14	4.44E-90	8.89E-103	2.88E-123	4.07E-102	3.92E-102	1.10E-48	2.21E-129	1.60E-155	4.64E-262	0	0	0	0	0	5.27E-265	1.78E-119	1.02E-30	5.56E-52	4.95E-94	1.20E-05
1.11E-05	1.03E-11	4.20E-09	5.96E-09	1.41E-08	1.03E-54	1.50E-106	8.98E-121	3.99E-112	2.17E-75	1.47E-28	1.95E-09	2.96E-68	7.69E-118	1.88E-132	2.17E-111	2.09E-111	1.99E-55	5.88E-137	1.25E-183	0	0	0	0	0	0	0	3.06E-156	7.94E-43	1.47E-47	1.29E-102	1.11E-05
9.42E-06	1.88E-12	1.47E-11	6.20E-08	1.90E-12	1.34E-81	1.29E-139	5.70E-165	4.07E-158	3.05E-113	3.29E-50	8.17E-11	9.98E-51	1.95E-135	4.71E-135	4.30E-121	4.13E-121	2.43E-63	5.66E-147	4.98E-216	0	0	0	0	0	0	0	3.05E-198	2.58E-58	6.32E-45	6.51E-112	9.42E-06
7.15E-06	6.67E-14	9.48E-16	3.11E-08	1.48E-21	1.52E-115	3.19E-181	1.27E-218	9.42E-214	3.54E-160	1.02E-79	2.30E-18	1.71E-37	1.45E-155	8.00E-139	3.17E-131	2.78E-127	2.00E-72	1.60E-159	1.00E-252	0	0	0	0	0	0	0	1.75E-245	3.49E-77	4.46E-44	6.37E-122	7.15E-06
4.03E-06	8.10E-15	4.88E-16	2.13E-08	6.64E-36	1.52E-156	2.25E-231	9.89E-282	4.93E-279	3.40E-216	4.41E-117	4.35E-32	1.48E-28	1.44E-177	9.17E-144	8.70E-142	1.46E-130	1.11E-82	1.32E-174	1.02E-293	0	0	0	0	0	0	0	0	1.96E-99	5.15E-45	1.21E-132	4.03E-06
1.37E-06	6.25E-14	1.33E-13	9.25E-11	1.72E-55	3.00E-201	4.55E-290	0	0	2.70E-281	2.63E-162	5.55E-52	6.49E-24	4.31E-156	7.08E-150	8.87E-153	2.85E-134	4.13E-94	3.20E-192	0	0	0	0	0	0	0	0	0	4.60E-125	9.71E-48	4.43E-144	1.37E-06
2.52E-07	9.88E-14	5.94E-13	6.35E-13	2.59E-80	5.19E-245	0	0	0	0	2.17E-215	4.77E-78	1.44E-23	4.85E-137	3.68E-157	3.36E-164	2.07E-138	1.04E-106	2.27E-212	0	0	0	0	0	0	0	0	0	4.47E-154	2.99E-52	3.15E-156	2.52E-07
2.34E-08	3.03E-14	4.33E-14	1.64E-14	2.24E-110	7.86E-296	0	0	0	0	2.48E-276	2.76E-110	1.63E-27	1.60E-120	1.29E-165	4.74E-176	5.55E-143	1.77E-120	4.72E-235	0	0	0	0	0	0	0	0	0	1.81E-186	1.51E-58	4.35E-169	2.34E-08
1.07E-09	1.80E-15	5.17E-17	1.86E-19	1.12E-145	0	0	0	0	0	0	1.07E-148	9.31E-36	1.54E-106	3.05E-175	2.49E-188	5.55E-148	2.02E-135	2.87E-260	0	0	0	0	0	0	0	0	0	3.05E-222	1.25E-66	5.12E-181	1.07E-09
2.34E-11	2.06E-17	1.01E-21	8.76E-28	3.25E-186	0	0	0	0	0	0	2.82E-193	2.70E-48	4.36E-95	4.87E-186	4.84E-201	2.06E-153	1.56E-151	5.13E-288	0	0	0	0	0	0	0	0	0	2.13E-261	1.68E-76	9.67E-179	2.34E-11
2.45E-13	4.59E-20	2.48E-25	1.72E-39	5.44E-232	0	0	0	0	0	0	4.97E-244	3.96E-65	3.61E-86	5.23E-198	3.51E-214	2.83E-159	8.11E-169	0	0	0	0	0	0	0	0	0	0	0	3.71E-88	3.62E-177	2.45E-13

P10A KERNEL DENSITY MATRIX

6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06	6.40E-06
1.10E-05	8.31E-06	3.50E-06	9.42E-07	2.00E-07	4.24E-08	1.13E-08	4.63E-09	3.27E-09	4.13E-09	8.87E-09	2.80E-08	1.05E-07	3.75E-07	1.01E-06	1.73E-06	1.70E-06	9.70E-07	3.49E-07	9.43E-08	2.38E-08	7.12E-09	3.12E-09	2.33E-09	3.19E-09	7.74E-09	2.96E-08	1.47E-07	7.39E-07	2.96E-06	7.64E-06	1.10E-05
1.25E-05	4.07E-06	1.61E-07	3.96E-09	5.93E-09	4.71E-09	1.41E-09	2.92E-10	6.27E-11	1.60E-11	4.12E-12	7.26E-13	1.47E-12	2.78E-10	1.71E-08	1.62E-07	1.60E-07	1.63E-08	2.52E-10	1.16E-12	3.88E-15	2.44E-17	6.72E-19	1.52E-19	3.78E-19	9.45E-18	1.50E-15	7.00E-13	3.69E-10	8.16E-08	3.20E-06	1.25E-05
1.24E-05	1.26E-06	1.69E-08	5.30E-08	6.16E-09	6.39E-11	3.40E-13	3.89E-15	2.39E-16	1.12E-16	2.55E-16	9.45E-16	1.31E-15	2.58E-15	2.86E-11	4.78E-09	4.73E-09	2.75E-11	2.17E-15	1.10E-20	2.59E-26	1.55E-26	3.60E-25	1.13E-25	8.06E-28	1.49E-31	8.69E-28	7.82E-22	9.54E-16	1.66E-10	5.98E-07	1.24E-05
1.21E-05	3.86E-07	6.30E-08	5.94E-09	5.04E-11	1.03E-13	5.18E-16	2.62E-17	1.13E-19	6.64E-23	6.73E-26	4.69E-23	2.44E-20	2.83E-19	4.57E-15	4.27E-11	4.22E-11	4.40E-15	2.15E-22	1.19E-24	1.07E-18	3.89E-16	5.76E-16	2.14E-17	5.31E-20	8.49E-24	3.16E-29	2.18E-34	1.36E-23	2.72E-14	5.61E-08	1.21E-05
1.24E-05	5.77E-08	1.85E-08	5.06E-09	2.73E-12	1.07E-11	2.83E-13	4.70E-15	7.91E-18	1.65E-21	3.22E-26	3.26E-31	3.18E-28	6.32E-24	6.79E-20	1.12E-13	1.11E-13	6.54E-20	7.14E-26	6.17E-16	1.17E-11	2.69E-11	3.80E-13	6.99E-16	8.11E-19	5.01E-22	2.62E-26	5.26E-33	1.14E-33	4.16E-19	3.10E-09	1.24E-05
1.22E-05	8.18E-08	4.82E-08	5.95E-10	2.82E-09	5.71E-10	3.70E-12	7.56E-16	1.49E-19	7.65E-23	2.08E-26	6.79E-32	9.33E-39	1.51E-30	9.34E-26	8.51E-17	8.42E-17	8.90E-26	9.69E-19	2.23E-10	4.87E-09	5.13E-12	1.04E-16	3.94E-21	2.87E-24	2.82E-26	2.07E-28	9.08E-33	7.63E-42	6.20E-25	9.27E-11	1.22E-05
1.10E-05	2.53E-07	1.87E-07	1.23E-08	2.57E-08	9.26E-11	8.21E-17	2.91E-23	6.58E-26	1.57E-30	8.00E-32	2.96E-34	4.42E-41	3.87E-39	2.41E-32	1.86E-20	1.83E-20	6.33E-31	1.41E-13	5.59E-08	7.66E-11	2.70E-18	1.17E-26	3.83E-33	2.35E-36	1.51E-36	1.56E-35	2.52E-36	1.56E-42	7.16E-32	1.21E-12	1.10E-05
1.02E-05	2.05E-07	2.56E-08	1.10E-07	6.94E-09	1.45E-16	9.02E-17	1.40E-13	6.34E-15	3.01E-20	8.35E-31	4.99E-41	8.16E-44	1.06E-49	1.41E-38	1.15E-24	1.14E-24	1.51E-26	2.20E-10	9.75E-09	4.55E-17	3.91E-30	5.42E-43	1.55E-51	5.86E-47	2.89E-50	1.13E-47	1.13E-43	9.22E-46	5.47E-40	7.21E-15	1.02E-05
1.09E-05	6.85E-08	2.49E-10	2.94E-08	3.12E-13	3.91E-16	1.95E-09	1.30E-08	1.06E-10	5.53E-15	3.79E-23	6.62E-39	1.14E-49	7.17E-58	1.36E-45	2.04E-29	2.01E-29	3.24E-23	3.66E-09	1.18E-12	1.02E-27	1.56E-47	4.32E-49	4.95E-37	4.83E-31	5.36E-32	7.14E-43	8.09E-55	1.56E-51	3.33E-49	2.68E-17	1.09E-05
1.11E-05	1.26E-07	2.82E-09	1.21E-09	1.27E-12	1.42E-10	4.02E-08	2.79E-10	3.07E-13	4.20E-16	4.74E-21	7.11E-33	6.72E-58	5.75E-62	1.17E-53	1.02E-34	1.01E-34	6.18E-21	6.55E-10	1.00E-19	8.68E-43	4.25E-58	1.46E-40	4.67E-29	9.22E-22	9.48E-20	7.94E-27	9.10E-49	7.60E-60	2.24E-59	6.02E-20	1.11E-05
1.09E-05	1.94E-07	5.80E-09	1.65E-13	3.51E-11	5.09E-10	7.91E-13	1.39E-18	1.53E-22	1.31E-23	1.64E-24	2.89E-31	6.35E-52	6.23E-68	8.98E-63	1.44E-40	1.42E-40	1.04E-19	1.25E-12	5.87E-30	2.78E-62	3.61E-53	2.05E-38	7.60E-28	4.08E-19	1.60E-13	8.46E-16	3.43E-33	1.06E-70	2.95E-70	8.43E-23	1.09E-05
1.11E-05	9.61E-08	2.46E-08	2.72E-10	1.58E-13	1.75E-14	1.48E-23	1.61E-33	1.31E-38	1.70E-37	1.56E-33	4.44E-34	5.01E-49	7.22E-76	6.10E-73	5.76E-47	5.68E-47	1.56E-19	2.57E-17	2.40E-43	1.41E-73	8.45E-54	1.18E-42	2.13E-33	4.17E-23	2.57E-13	8.61E-10	2.08E-21	2.18E-58	1.04E-77	6.69E-26	1.11E-05
1.04E-05	5.49E-08	5.75E-09	4.98E-08	5.34E-17	6.89E-23	1.64E-27	8.15E-42	5.87E-59	9.06E-58	4.09E-48	2.58E-41	2.75E-49	1.28E-84	3.67E-84	6.48E-54	6.39E-54	2.08E-20	5.64E-24	6.81E-60	5.04E-74	5.46E-60	2.83E-53	1.03E-45	9.90E-34	3.94E-19	8.39E-09	2.03E-13	2.42E-45	3.95E-86	2.13E-29	1.04E-05
1.03E-05	5.41E-08	4.61E-11	2.62E-08	4.74E-16	3.66E-17	4.11E-29	8.07E-49	3.52E-68	2.70E-84	2.96E-68	5.66E-53	1.05E-52	1.77E-82	1.96E-96	2.05E-61	4.26E-60	2.45E-22	1.33E-32	1.35E-79	6.80E-79	9.73E-72	2.78E-70	8.60E-65	5.43E-51	5.75E-31	7.81E-13	3.18E-09	7.69E-35	5.12E-96	3.37E-33	1.03E-05
1.06E-05	1.34E-08	1.22E-10	3.95E-11	1.84E-14	2.05E-16	9.88E-37	6.62E-60	2.46E-83	3.27E-99	5.91E-94	4.70E-69	2.78E-59	2.83E-82	9.54E-110	1.83E-69	4.71E-62	2.55E-25	3.34E-43	1.85E-102	3.47E-88	4.78E-89	1.13E-93	1.24E-90	6.90E-75	8.01E-49	6.96E-22	8.03E-09	7.03E-27	2.23E-92	3.28E-37	1.06E-05
1.06E-05	1.22E-09	1.01E-08	7.41E-13	2.07E-10	1.15E-20	6.83E-39	7.86E-60	8.25E-85	3.67E-107	3.29E-125	1.47E-89	5.13E-69	4.85E-84	9.83E-120	4.59E-78	1.54E-64	2.36E-29	9.03E-56	1.90E-111	6.69E-102	6.48E-112	1.89E-123	3.06E-123	2.03E-105	1.06E-72	5.93E-36	3.26E-12	1.85E-21	9.97E-83	1.14E-41	1.06E-05
1.03E-05	2.77E-09	6.84E-08	9.39E-10	1.40E-08	1.14E-24	6.42E-39	4.96E-60	3.13E-56	1.02E-76	2.69E-127	1.75E-114	6.59E-82	8.90E-88	3.03E-129	3.23E-87	1.40E-67	1.94E-34	2.61E-70	1.97E-119	4.88E-120	2.42E-140	1.31E-159	1.31E-162	1.38E-142	1.35E-102	4.83E-55	2.13E-19	1.39E-18	1.52E-74	1.38E-46	1.03E-05
9.04E-06	2.17E-08	5.72E-08	5.23E-08	1.92E-10	7.28E-21	5.76E-45	2.99E-37	2.22E-34	1.16E-52	8.65E-99	7.83E-144	5.88E-98	1.75E-93	8.29E-140	6.39E-97	3.57E-71	1.41E-40	8.08E-87	1.43E-130	1.34E-142	2.49E-174	3.74E-202	9.63E-209	2.19E-186	1.62E-138	3.77E-79	2.25E-30	3.02E-18	7.93E-68	1.00E-51	9.04E-06
8.90E-06	4.00E-08	2.26E-09	2.29E-08	5.92E-16	4.43E-22	3.71E-43	4.17E-21	2.72E-19	5.47E-35	7.66E-76	3.22E-158	3.65E-117	3.68E-101	2.01E-151	3.55E-107	2.55E-75	9.09E-48	2.68E-105	7.16E-145	1.40E-169	7.09E-214	4.41E-251	1.22E-261	7.99E-237	1.87E-180	2.81E-108	3.81E-45	1.88E-20	1.41E-62	5.56E-57	8.90E-06
1.06E-05	1.34E-08	2.11E-09	6.17E-11	1.31E-11	2.59E-28	1.16E-28	1.35E-11	5.74E-11	1.06E-23	1.87E-58	8.51E-134	1.58E-139	8.31E-111	1.40E-159	5.54E-118	5.11E-80	5.20E-56	9.52E-126	2.50E-162	5.51E-201	5.55E-259	0	0	6.76E-294	2.05E-228	2.00E-142	1.04E-63	3.38E-25	8.56E-59	5.00E-62	1.06E-05
1.13E-05	8.61E-10	2.15E-08	2.44E-10	3.90E-09	1.45E-39	3.45E-20	1.01E-08	2.09E-09	8.55E-19	1.26E-46	8.49E-114	4.73E-165	2.01E-122	5.33E-164	2.42E-129	2.86E-85	2.63E-65	3.62E-148	6.08E-183	8.20E-237	0	0	0	0	2.14E-282	1.37E-181	4.55E-86	1.74E-32	1.77E-56	2.47E-67	1.13E-05
9.99E-06	4.13E-11	7.91E-08	1.07E-08	1.94E-10	3.61E-48	9.81E-18	1.75E-12	1.31E-14	2.83E-20	2.34E-40	3.21E-98	9.89E-194	5.19E-136	1.81E-169	2.98E-141	4.49E-91	1.18E-75	1.48E-172	1.03E-206	4.61E-277	0	0	0	0	0	8.92E-226	3.22E-112	2.57E-42	1.25E-55	2.87E-73	9.99E-06
8.20E-06	3.35E-10	1.59E-08	6.59E-08	1.63E-15	1.33E-41	2.66E-21	7.04E-23	1.41E-26	3.87E-28	1.20E-39	4.57E-87	6.97E-205	1.44E-151	5.42E-176	1.03E-153	1.97E-97	4.72E-87	6.45E-199	1.21E-233	0	0	0	0	0	0	5.56E-275	3.65E-142	1.09E-54	3.03E-56	7.16E-80	8.20E-06
7.10E-06	1.17E-09	1.11E-10	1.16E-09	2.36E-24	4.67E-40	6.86E-31	6.55E-40	2.63E-45	2.19E-42	1.69E-44	2.47E-80	4.27E-188	4.27E-169	1.44E-183	9.92E-167	2.42E-104	1.67E-99	3.21E-222	9.87E-264	0	0	0	0	0	0	0	6.68E-176	1.33E-69	2.49E-58	4.06E-87	7.10E-06
4.16E-06	7.41E-10	2.65E-14	1.00E-12	6.00E-37	1.57E-43	1.69E-46	1.41E-63	8.45E-71	5.09E-63	6.59E-55	5.03E-78	1.82E-174	1.36E-188	3.40E-192	2.69E-180	8.34E-112	5.22E-113	6.85E-239	0	0	0	0	0	0	0	0	1.97E-213	4.68E-87	7.01E-62	8.46E-95	4.16E-06
1.09E-06	8.46E-11	7.16E-17	8.82E-12	2.66E-53	5.06E-52	3.96E-68	7.02E-94	4.68E-103	4.90E-90	7.08E-71	3.88E-80	5.37E-164	4.62E-210	7.12E-202	2.04E-194	8.04E-120	1.45E-127	1.57E-257	0	0	0	0	0	0	0	0	9.31E-255	4.72E-107	6.72E-67	1.22E-102	1.09E-06
1.10E-07	1.75E-12	4.32E-15	2.42E-13	2.03E-73	1.56E-65	8.86E-96	8.11E-131	4.46E-142	1.94E-123	2.10E-92	1.13E-86	1.11E-156	1.69E-233	1.32E-212	4.36E-209	2.17E-128	3.57E-143	3.83E-278	0	0	0	0	0	0	0	0	0	1.37E-129	2.20E-73	6.30E-111	1.10E-07
3.94E-09	6.50E-15	8.91E-15	1.92E-17	1.20E-87	4.59E-84	1.89E-129	2.17E-174	7.35E-188	3.18E-163	1.71E-119	1.24E-97	1.58E-152	6.59E-259	2.17E-224	2.60E-224	1.64E-137	7.78E-160	0	0	0	0	0	0	0	0	0	0	1.14E-154	2.46E-81	7.24E-120	3.94E-09
4.89E-11	4.37E-18	6.28E-16	4.41E-24	1.50E-93	1.29E-107	3.84E-169	1.34E-224	2.08E-240	2.15E-209	3.85E-152	5.18E-113	1.58E-151	2.76E-286	3.16E-237	4.36E-240	3.46E-147	1.51E-177	0	0	0	0	0	0	0	0	0	0	2.73E-182	9.38E-91	1.65E-129	4.89E-11
2.06E-13	5.31E-22	1.51E-18	2.91E-33	3.03E-103	3.48E-136	7.44E-215	1.92E-281	0	6.00E-262	2.39E-190	8.16E-133	1.09E-153	0	4.09E-251	2.05E-256	2.05E-157	2.58E-196	0	0	0	0	0	0	0	0	0	0	1.88E-212	1.22E-101	7.23E-140	2.06E-13
2.90E-16	1.16E-26	1.24E-22	5.55E-45	9.83E-117	8.98E-170	1.37E-266	0	0	0	4.10E-234	4.85E-157	5.25E-159	0	4.68E-266	2.69E-273	3.40E-168	3.93E-216	0	0	0	0	0	0	0	0	0	0	3.71E-245	5.43E-114	5.94E-151	2.90E-16

Matrix 2. P7 versus P10: matrices of mean density distribution

P7 KERNEL DENSITY MATRIX

2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05	2.45E-05
3.63E-05	2.09E-05	4.35E-06	1.00E-06	2.98E-07	7.25E-08	1.87E-08	7.09E-09	4.88E-09	6.49E-09	1.50E-08	4.57E-08	1.34E-07	4.93E-07	3.25E-06	1.08E-05	1.01E-05	2.55E-06	2.14E-07	8.87E-09	3.05E-10	1.50E-11	1.75E-12	6.89E-13	1.08E-12	6.29E-12	1.05E-10	3.22E-09	1.05E-07	2.12E-06	1.65E-05	3.63E-05
4.38E-05	5.35E-06	2.58E-07	7.47E-09	5.14E-09	1.19E-09	8.38E-11	7.11E-12	1.80E-12	1.70E-12	3.66E-12	5.95E-12	2.85E-12	7.31E-10	3.90E-08	2.36E-06	2.12E-06	9.34E-09	4.72E-13	1.35E-18	1.74E-24	9.18E-30	1.44E-33	2.86E-35	1.38E-34	1.31E-31	8.17E-27	5.83E-21	5.48E-15	7.51E-10	2.27E-06	4.38E-05
4.73E-05	1.08E-06	3.59E-08	1.25E-08	4.49E-12	2.64E-17	1.59E-22	3.53E-26	1.38E-27	2.31E-26	5.69E-23	1.14E-18	4.41E-15	5.95E-14	2.35E-10	2.47E-07	2.15E-07	1.13E-12	2.45E-22	8.07E-35	4.27E-48	1.64E-51	7.37E-44	1.95E-40	6.69E-41	2.14E-45	1.27E-53	1.43E-40	3.51E-27	1.14E-15	6.98E-08	4.73E-05
4.85E-05	2.02E-07	1.13E-07	1.36E-12	6.04E-21	1.11E-20	2.57E-22	4.43E-25	1.83E-27	4.41E-29	1.00E-30	4.26E-34	4.41E-24	4.80E-17	7.47E-14	1.23E-08	1.03E-08	4.38E-18	2.93E-35	1.02E-56	5.02E-38	1.62E-26	3.60E-20	2.22E-17	2.12E-17	1.61E-20	5.28E-28	8.14E-42	2.80E-44	7.61E-24	4.97E-10	4.85E-05
4.86E-05	2.98E-07	3.97E-09	6.31E-18	1.46E-17	1.92E-22	2.43E-30	2.19E-37	4.09E-41	6.38E-41	3.73E-38	7.51E-36	4.14E-38	8.77E-24	7.23E-18	2.87E-10	2.30E-10	5.41E-25	7.96E-52	2.14E-37	2.58E-21	1.42E-13	1.83E-10	3.43E-09	1.81E-08	8.94E-09	2.82E-12	2.24E-22	5.35E-43	2.16E-34	8.05E-13	4.86E-05
4.81E-05	2.39E-07	4.07E-12	5.69E-15	1.54E-20	5.39E-35	1.70E-51	1.41E-62	9.37E-69	9.63E-67	1.23E-57	6.75E-47	9.31E-41	3.50E-34	1.04E-20	3.21E-12	2.37E-12	2.10E-33	5.51E-50	2.00E-24	5.33E-14	1.10E-12	9.73E-15	7.16E-16	4.19E-14	3.67E-10	2.44E-07	7.37E-11	7.95E-27	2.63E-47	3.01E-16	4.81E-05
4.74E-05	2.66E-07	1.86E-13	1.06E-14	1.94E-31	1.14E-50	3.39E-64	1.60E-77	7.69E-86	6.39E-88	2.62E-86	2.44E-67	7.95E-50	3.99E-46	5.17E-25	1.76E-14	1.12E-14	2.55E-43	5.40E-38	6.81E-18	4.44E-16	7.54E-24	5.38E-33	2.03E-37	2.62E-34	1.11E-24	3.41E-13	2.88E-07	1.19E-15	1.09E-46	2.81E-20	4.74E-05
4.61E-05	1.62E-07	1.46E-12	2.00E-19	1.18E-42	3.15E-61	4.74E-86	4.99E-107	9.84E-118	1.45E-116	5.23E-107	9.82E-97	2.48E-65	5.31E-50	7.85E-31	4.95E-17	2.43E-17	9.60E-55	1.15E-29	8.48E-18	1.49E-27	4.57E-47	3.11E-65	7.78E-74	4.40E-69	2.48E-52	7.75E-30	1.34E-11	1.79E-09	6.78E-35	6.70E-25	4.61E-05
4.39E-05	1.07E-07	4.03E-10	4.12E-29	8.80E-49	1.41E-82	4.89E-121	4.20E-151	1.71E-164	3.43E-159	9.24E-140	1.84E-116	2.82E-87	1.56E-57	3.64E-38	7.51E-20	2.39E-20	1.42E-64	5.37E-25	3.85E-24	2.01E-48	2.45E-82	1.87E-111	4.04E-125	2.00E-118	4.10E-93	2.85E-57	7.45E-24	2.71E-08	1.38E-25	3.72E-30	4.39E-05
4.24E-05	1.64E-07	3.66E-10	5.33E-33	7.82E-63	1.03E-114	3.72E-169	9.53E-210	4.01E-226	8.45E-216	1.44E-184	2.19E-145	1.75E-115	9.96E-69	5.15E-47	6.14E-23	1.07E-23	3.10E-57	5.46E-24	6.38E-37	1.10E-78	1.16E-129	1.17E-171	2.84E-191	2.46E-182	5.00E-147	1.70E-95	4.93E-44	4.13E-12	9.25E-19	4.31E-36	4.24E-05
4.27E-05	1.06E-07	1.12E-12	6.79E-36	8.29E-85	1.22E-157	2.09E-230	2.62E-241	1.91E-228	1.34E-253	2.00E-241	1.05E-183	4.46E-135	1.39E-83	2.22E-57	2.58E-26	2.16E-27	2.07E-51	1.21E-26	3.86E-56	2.41E-118	4.88E-189	7.68E-246	2.71E-272	8.14E-261	4.50E-214	1.64E-144	3.88E-72	6.34E-21	2.03E-14	1.01E-42	4.27E-05
4.41E-05	1.59E-07	2.57E-17	8.71E-44	1.05E-114	2.32E-211	1.25E-216	6.30E-168	1.08E-155	8.14E-178	4.03E-241	2.02E-231	1.25E-160	4.23E-102	4.81E-69	5.26E-30	1.98E-31	4.22E-47	5.87E-33	8.51E-82	2.12E-167	1.81E-260	0	0	0	2.99E-294	2.57E-204	3.64E-108	9.78E-35	1.46E-12	5.29E-50	4.41E-05
4.50E-05	1.53E-07	3.67E-22	1.12E-56	1.57E-152	5.29E-241	1.55E-152	4.08E-109	8.26E-98	5.16E-116	4.66E-172	2.17E-283	1.28E-192	2.81E-124	5.82E-75	5.04E-34	8.23E-36	2.62E-44	6.19E-43	6.85E-114	7.56E-226	0	0	0	0	0	6.53E-275	4.07E-152	1.52E-53	3.43E-13	9.09E-58	4.50E-05
4.46E-05	2.78E-08	7.80E-20	1.46E-74	2.81E-198	1.86E-180	1.42E-101	7.14E-65	8.57E-55	3.41E-68	4.77E-115	2.46E-216	4.78E-231	4.08E-150	5.52E-82	2.21E-38	1.54E-40	4.99E-43	1.43E-56	2.01E-152	1.08E-293	0	0	0	0	0	0	5.42E-204	2.38E-77	2.65E-16	3.50E-53	4.46E-05
4.36E-05	1.90E-09	5.80E-20	1.91E-97	6.00E-252	1.06E-130	9.58E-64	3.38E-35	1.20E-26	2.35E-34	4.31E-70	1.12E-158	6.52E-276	5.73E-171	1.60E-90	4.41E-43	1.31E-45	2.89E-43	7.17E-74	2.16E-197	0	0	0	0	0	0	0	8.60E-264	3.74E-106	6.70E-22	9.98E-49	4.36E-05
4.29E-05	1.53E-08	1.41E-22	2.51E-125	2.62E-212	9.78E-92	4.78E-39	4.31E-20	2.29E-13	1.69E-14	3.45E-37	2.06E-110	2.14E-272	4.62E-193	1.41E-100	3.98E-48	4.99E-51	5.13E-45	7.85E-95	8.44E-249	0	0	0	0	0	0	0	0	5.92E-140	5.55E-30	4.58E-45	4.29E-05
4.34E-05	7.05E-08	1.13E-27	3.33E-158	2.64E-167	1.46E-63	1.76E-27	1.48E-19	5.91E-15	1.27E-08	2.44E-16	1.52E-71	1.63E-218	8.10E-219	3.81E-112	1.62E-53	8.60E-57	2.78E-48	1.88E-119	0	0	0	0	0	0	0	0	0	9.44E-179	1.51E-40	3.39E-42	4.34E-05
4.46E-05	2.30E-07	2.95E-35	4.44E-196	3.17E-130	3.55E-46	4.79E-29	1.38E-33	2.06E-31	9.89E-17	1.53E-07	4.54E-42	4.51E-171	3.10E-248	3.14E-125	2.96E-59	6.67E-63	4.59E-53	9.80E-148	0	0	0	0	0	0	0	0	0	1.51E-222	1.34E-53	4.03E-40	4.46E-05
4.53E-05	2.24E-07	2.53E-45	5.95E-239	4.53E-101	1.40E-39	9.61E-44	3.46E-62	9.76E-63	8.06E-39	8.48E-11	5.44E-22	4.56E-130	2.59E-281	7.90E-140	2.44E-65	2.34E-69	2.32E-59	1.12E-179	0	0	0	0	0	0	0	0	0	2.44E-271	3.92E-69	7.73E-39	4.53E-05
4.46E-05	4.92E-08	7.10E-58	9.96E-248	7.72E-80	8.89E-44	1.43E-71	2.35E-105	6.25E-109	6.85E-75	4.16E-26	2.63E-11	1.68E-95	0	6.07E-156	9.14E-72	3.68E-76	3.57E-67	2.77E-215	0	0	0	0	0	0	0	0	0	0	3.74E-87	2.39E-38	4.46E-05
4.24E-05	2.49E-09	6.54E-73	4.39E-212	1.57E-66	9.17E-59	1.56E-112	4.29E-163	5.43E-170	6.07E-125	1.80E-53	5.13E-10	2.27E-67	0	1.42E-173	1.61E-78	2.61E-83	1.68E-76	1.50E-254	0	0	0	0	0	0	0	0	0	0	1.17E-107	1.19E-38	4.24E-05
3.90E-05	2.55E-11	1.97E-90	1.94E-181	3.79E-61	1.53E-84	1.26E-166	2.12E-235	6.38E-246	5.60E-189	6.92E-93	4.02E-18	1.11E-45	4.01E-261	1.02E-192	1.60E-85	8.36E-91	2.41E-87	0	0	0	0	0	0	0	0	0	0	0	1.21E-130	9.50E-40	3.90E-05
3.39E-05	1.74E-11	1.95E-110	8.67E-156	1.09E-63	4.16E-121	7.52E-234	0	0	5.40E-267	2.35E-144	1.27E-35	2.00E-30	9.87E-224	2.23E-213	1.34E-92	1.20E-98	1.06E-99	0	0	0	0	0	0	0	0	0	0	0	4.06E-156	1.23E-41	3.39E-05
2.61E-05	2.24E-09	6.33E-133	3.89E-135	3.75E-74	1.83E-168	0	0	0	0	7.06E-208	1.62E-62	1.31E-21	5.30E-190	1.49E-235	9.25E-100	7.79E-107	1.42E-113	0	0	0	0	0	0	0	0	0	0	0	4.47E-184	2.55E-44	2.61E-05
1.63E-05	4.65E-08	6.72E-158	1.76E-119	1.53E-92	1.30E-226	0	0	0	0	1.88E-283	8.30E-99	3.12E-19	6.22E-160	3.04E-259	3.58E-107	2.27E-115	5.80E-129	0	0	0	0	0	0	0	0	0	0	0	1.62E-214	8.52E-48	1.63E-05
7.56E-06	1.56E-07	2.34E-185	8.00E-109	7.47E-119	1.50E-295	0	0	0	0	0	1.71E-144	2.72E-23	1.59E-133	1.91E-284	6.50E-115	2.98E-124	7.24E-146	0	0	0	0	0	0	0	0	0	0	0	1.92E-247	4.59E-52	7.56E-06
2.48E-06	8.38E-08	2.67E-215	3.66E-103	4.33E-153	0	0	0	0	0	0	1.43E-199	8.65E-34	8.87E-111	0	5.34E-123	1.76E-133	2.76E-164	0	0	0	0	0	0	0	0	0	0	0	7.44E-283	3.98E-57	2.48E-06
5.53E-07	7.27E-09	9.99E-248	1.69E-102	3.00E-195	0	0	0	0	0	0	4.78E-264	1.00E-50	1.08E-91	0	1.97E-131	4.73E-143	3.22E-184	0	0	0	0	0	0	0	0	0	0	0	0	5.57E-63	5.53E-07
8.24E-08	1.02E-10	1.22E-282	7.81E-107	2.47E-245	0	0	0	0	0	0	0	4.25E-74	2.86E-76	0	3.28E-140	4.56E-148	1.14E-205	0	0	0	0	0	0	0	0	0	0	0	0	1.25E-69	8.24E-08
8.08E-09	2.29E-13	6.89E-285	3.65E-116	0	0	0	0	0	0	0	0	6.56E-104	1.66E-64	0	2.45E-149	1.57E-151	1.24E-228	0	0	0	0	0	0	0	0	0	0	0	0	4.54E-77	8.08E-09
5.17E-10	8.28E-17	6.82E-266	1.71E-130	0	0	0	0	0	0	0	0	3.70E-140	2.10E-56	0	8.20E-159	2.43E-155	4.12E-253	0	0	0	0	0	0	0	0	0	0	0	0	2.65E-85	5.17E-10

P10 KERNEL DENSITY MATRIX

2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05	2.44E-05
3.62E-05	2.12E-05	4.63E-06	9.93E-07	2.45E-07	5.24E-08	1.25E-08	4.60E-09	3.13E-09	4.20E-09	1.02E-08	3.42E-08	1.17E-07	5.01E-07	3.27E-06	1.07E-05	1.00E-05	2.51E-06	2.08E-07	8.55E-09	2.91E-10	1.43E-11	1.65E-12	6.49E-13	1.02E-12	5.98E-12	1.01E-10	3.11E-09	1.03E-07	2.09E-06	1.63E-05	3.62E-05
4.36E-05	5.73E-06	2.00E-07	8.37E-09	4.56E-09	5.50E-10	2.39E-11	1.53E-12	3.57E-13	3.85E-13	1.16E-12	3.17E-12	2.54E-12	4.51E-10	3.57E-08	2.31E-06	2.05E-06	8.86E-09	4.35E-13	1.20E-18	1.50E-24	7.69E-30	1.19E-33	2.35E-35	1.15E-34	1.11E-31	7.15E-27	5.28E-21	5.13E-15	7.23E-10	2.23E-06	4.36E-05
4.71E-05	1.04E-06	4.69E-08	5.91E-09	6.36E-13	1.38E-18	1.50E-22	4.38E-24	9.56E-26	2.96E-27	1.67E-24	7.51E-20	8.28E-16	3.88E-14	1.48E-10	2.36E-07	2.02E-07	1.02E-12	2.07E-22	6.32E-35	3.10E-48	9.78E-52	4.78E-44	1.33E-40	4.65E-41	1.46E-45	9.35E-54	1.18E-40	3.14E-27	1.10E-15	6.91E-08	4.71E-05
4.85E-05	1.63E-07	7.88E-08	1.50E-13	1.42E-18	8.96E-21	1.98E-23	8.76E-27	2.48E-29	1.11E-30	1.25E-31	5.35E-33	1.41E-25	8.29E-18	3.62E-14	1.13E-08	9.28E-09	3.68E-18	2.20E-35	5.57E-57	3.15E-38	1.17E-26	2.83E-20	1.85E-17	1.81E-17	1.36E-20	4.25E-28	5.89E-42	2.41E-44	7.49E-24	5.13E-10	4.85E-05
4.86E-05	3.22E-07	1.02E-09	7.15E-16	3.24E-18	8.51E-25	6.00E-34	9.60E-42	1.11E-45	3.81E-45	1.68E-41	7.93E-38	8.22E-38	3.92E-25	3.35E-18	2.50E-10	1.97E-10	4.16E-25	5.15E-52	1.24E-37	1.81E-21	1.13E-13	1.59E-10	3.14E-09	1.73E-08	8.61E-09	2.64E-12	1.92E-22	3.87E-43	2.25E-34	8.92E-13	4.86E-05
4.81E-05	4.46E-07	7.37E-14	5.11E-15	4.31E-23	1.37E-39	1.61E-50	2.15E-58	1.51E-64	4.54E-68	1.85E-63	4.70E-51	9.70E-43	3.91E-36	2.04E-21	2.59E-12	1.91E-12	1.46E-33	2.60E-50	1.24E-24	3.91E-14	8.94E-13	8.43E-15	6.54E-16	4.03E-14	3.67E-10	2.46E-07	7.06E-11	6.55E-27	2.94E-47	3.61E-16	4.81E-05
4.71E-05	3.69E-07	3.80E-12	3.09E-16	1.77E-35	3.26E-47	1.39E-62	8.24E-77	4.87E-85	4.87E-86	2.09E-82	1.05E-73	1.26E-53	8.43E-47	4.14E-26	1.27E-14	8.39E-15	1.57E-43	2.62E-38	4.27E-18	3.18E-16	5.74E-24	4.25E-33	1.67E-37	2.29E-34	1.05E-24	3.45E-13	2.91E-07	1.07E-15	7.46E-47	3.53E-20	4.71E-05
4.51E-05	7.73E-08	2.76E-11	2.01E-22	3.56E-39	7.38E-61	8.17E-88	7.74E-110	1.92E-120	4.96E-118	3.81E-106	9.06E-93	5.65E-71	3.83E-52	2.53E-32	3.04E-17	1.67E-17	5.14E-55	5.57E-30	5.12E-18	9.76E-28	3.01E-47	2.03E-65	5.21E-74	3.20E-69	2.05E-52	7.26E-30	1.35E-11	1.70E-09	5.01E-35	8.40E-25	4.51E-05
4.15E-05	3.72E-09	2.33E-10	5.98E-29	5.32E-48	2.51E-85	3.23E-126	1.78E-157	9.27E-171	4.79E-164	5.68E-142	6.25E-115	8.88E-95	4.56E-61	4.62E-40	3.69E-20	1.50E-20	4.35E-65	2.51E-25	2.14E-24	1.13E-48	1.28E-82	9.17E-112	1.99E-125	1.09E-118	2.68E-93	2.30E-57	7.05E-24	2.61E-08	1.07E-25	4.54E-30	4.15E-05
3.83E-05	3.30E-10	2.18E-11	5.95E-30	8.96E-65	1.28E-120	8.64E-178	9.97E-220	5.48E-236	4.39E-224	6.87E-190	1.63E-146	6.96E-112	1.15E-73	2.54E-49	2.34E-23	6.02E-24	9.07E-58	2.38E-24	3.10E-37	4.91E-79	4.44E-130	3.94E-172	9.33E-192	9.09E-183	2.36E-147	1.09E-95	4.13E-44	3.87E-12	7.38E-19	5.32E-36	3.83E-05
3.80E-05	1.26E-08	9.03E-15	4.72E-35	1.70E-89	9.84E-167	1.56E-242	4.09E-242	3.83E-229	2.58E-254	6.78E-250	1.59E-187	4.02E-133	6.12E-90	4.17E-60	7.65E-27	1.09E-27	5.68E-52	4.79E-27	1.56E-56	8.04E-119	1.25E-189	1.60E-246	5.35E-273	1.85E-261	1.41E-214	7.80E-145	2.72E-72	5.55E-21	1.62E-14	1.47E-42	3.80E-05
4.01E-05	1.18E-07	2.90E-17	3.62E-45	3.61E-122	1.14E-223	1.80E-217	1.56E-168	3.47E-156	2.59E-178	9.20E-242	5.88E-238	8.09E-161	6.89E-110	2.87E-71	1.24E-30	8.80E-32	1.06E-47	2.04E-33	2.74E-82	4.97E-168	2.87E-261	0	0	0	5.66E-295	8.37E-205	2.02E-108	7.70E-35	1.14E-12	1.14E-49	4.01E-05
4.21E-05	2.60E-07	4.18E-18	2.69E-60	8.63E-163	4.71E-242	3.25E-153	1.45E-109	3.86E-98	2.47E-116	1.68E-172	3.89E-284	5.65E-195	3.37E-132	9.91E-78	9.54E-35	3.19E-36	5.99E-45	1.83E-43	1.67E-114	1.16E-226	0	0	0	0	0	1.35E-275	1.68E-152	1.03E-53	2.56E-13	2.53E-57	4.21E-05
4.20E-05	3.45E-07	1.86E-17	1.93E-80	2.32E-211	2.28E-181	3.96E-102	3.31E-65	5.26E-55	2.23E-68	2.50E-115	6.80E-217	1.37E-235	1.61E-148	1.12E-85	3.39E-39	5.15E-41	1.01E-43	3.48E-57	3.53E-153	1.01E-294	0	0	0	0	0	0	1.58E-204	1.34E-77	1.84E-16	1.58E-53	4.20E-05
4.06E-05	3.02E-07	3.21E-19	1.34E-105	1.49E-266	1.66E-131	3.25E-64	1.85E-35	8.77E-27	1.91E-34	3.02E-70	4.48E-159	1.16E-282	1.71E-168	3.83E-95	5.46E-44	3.72E-46	5.11E-44	1.40E-74	2.59E-198	0	0	0	0	0	0	0	1.66E-264	1.68E-106	4.24E-22	4.29E-49	4.06E-05
4.00E-05	4.90E-07	1.77E-23	9.00E-136	1.51E-213	1.82E-92	1.80E-39	2.52E-20	1.79E-13	1.55E-14	2.97E-37	1.11E-110	3.83E-273	3.83E-192	3.91E-106	3.96E-49	1.20E-51	7.74E-46	1.18E-95	6.64E-250	0	0	0	0	0	0	0	0	2.04E-140	3.12E-30	1.85E-45	4.00E-05
4.11E-05	3.63E-07	3.12E-30	5.85E-171	1.81E-168	3.00E-64	6.74E-28	8.38E-20	4.48E-15	1.19E-08	2.38E-16	1.04E-71	4.00E-219	1.82E-219	1.20E-118	1.28E-54	1.72E-57	3.52E-49	2.13E-120	0	0	0	0	0	0	0	0	0	2.40E-179	7.35E-41	1.27E-42	4.11E-05
4.34E-05	9.46E-08	1.76E-39	3.67E-211	2.44E-131	7.40E-47	1.70E-29	6.83E-34	1.37E-31	8.69E-17	1.55E-07	3.66E-42	1.46E-171	1.82E-250	1.10E-132	1.86E-60	1.10E-63	4.80E-54	8.06E-149	0	0	0	0	0	0	0	0	0	2.72E-223	5.53E-54	1.38E-40	4.34E-05
4.50E-05	8.07E-08	3.18E-51	2.23E-256	3.69E-102	2.75E-40	2.89E-44	1.36E-62	5.14E-63	6.00E-39	8.25E-11	4.86E-22	1.84E-130	3.86E-285	3.02E-148	1.20E-66	3.15E-70	1.96E-60	6.47E-181	0	0	0	0	0	0	0	0	0	2.99E-272	1.33E-69	2.40E-39	4.50E-05
4.45E-05	8.82E-08	1.84E-65	1.64E-249	6.29E-81	1.53E-44	3.33E-72	6.63E-106	2.36E-109	3.93E-75	3.56E-26	2.43E-11	8.09E-96	0	2.49E-165	3.46E-73	4.02E-77	2.40E-68	1.10E-216	0	0	0	0	0	0	0	0	0	0	1.03E-87	6.59E-39	4.45E-05
4.16E-05	2.35E-07	3.39E-82	7.59E-214	1.20E-67	1.29E-59	2.58E-113	7.90E-164	1.33E-170	2.44E-125	1.25E-53	4.58E-10	1.24E-67	0	6.17E-184	4.47E-80	2.28E-84	8.83E-78	3.93E-256	0	0	0	0	0	0	0	0	0	0	2.53E-108	2.88E-39	4.16E-05
3.72E-05	3.49E-07	2.01E-101	3.40E-183	2.59E-62	1.62E-85	1.35E-167	2.30E-236	9.14E-247	1.44E-189	3.59E-93	3.25E-18	6.56E-46	4.78E-262	4.58E-204	2.64E-87	5.75E-92	9.73E-89	0	0	0	0	0	0	0	0	0	0	0	2.00E-131	2.00E-40	3.72E-05
3.17E-05	3.47E-07	3.79E-123	1.47E-157	6.26E-65	3.06E-122	4.76E-235	0	0	8.02E-268	8.35E-145	8.68E-36	1.21E-30	1.37E-224	1.02E-225	7.72E-95	6.46E-100	3.21E-101	0	0	0	0	0	0	0	0	0	0	0	5.04E-157	2.20E-42	3.17E-05
2.40E-05	1.85E-07	2.29E-147	6.18E-137	1.70E-75	8.71E-170	0	0	0	0	1.58E-208	8.74E-63	7.77E-22	8.33E-191	4.20E-245	1.40E-102	3.24E-108	3.18E-115	0	0	0	0	0	0	0	0	0	0	0	4.07E-185	3.85E-45	2.40E-05
1.47E-05	1.06E-07	4.44E-174	2.51E-121	5.20E-94	3.72E-228	0	0	0	0	2.44E-284	3.31E-99	1.73E-19	1.07E-160	9.69E-264	1.91E-110	7.21E-117	9.46E-131	0	0	0	0	0	0	0	0	0	0	0	1.05E-215	1.07E-48	1.47E-05
6.58E-06	3.02E-08	2.75E-203	9.83E-111	1.79E-120	0	0	0	0	0	0	4.74E-145	1.34E-23	2.90E-134	6.70E-284	1.66E-118	7.15E-126	8.43E-148	0	0	0	0	0	0	0	0	0	0	0	8.67E-249	4.74E-53	6.58E-06
2.05E-06	1.49E-09	5.44E-235	3.73E-105	6.91E-155	0	0	0	0	0	0	2.55E-200	3.61E-34	1.66E-111	0	7.24E-127	3.15E-135	2.25E-166	0	0	0	0	0	0	0	0	0	0	0	2.29E-284	3.33E-58	2.05E-06
4.30E-07	1.17E-11	3.44E-269	1.37E-104	3.00E-197	0	0	0	0	0	0	5.17E-265	3.38E-51	2.02E-92	0	1.45E-135	6.22E-145	1.80E-186	0	0	0	0	0	0	0	0	0	0	0	0	3.71E-64	4.30E-07
5.91E-08	1.46E-14	0	4.85E-109	1.47E-247	0	0	0	0	0	0	0	1.10E-74	5.17E-77	0	1.29E-144	1.42E-150	4.33E-208	0	0	0	0	0	0	0	0	0	0	0	0	6.57E-71	5.91E-08
5.29E-09	2.90E-18	4.72E-288	1.67E-118	0	0	0	0	0	0	0	0	1.25E-104	2.80E-65	0	5.13E-154	3.60E-154	3.12E-231	0	0	0	0	0	0	0	0	0	0	0	0	1.85E-78	5.29E-09
3.06E-10	9.13E-23	3.94E-269	5.53E-133	0	0	0	0	0	0	0	0	4.89E-141	3.21E-57	0	9.05E-164	4.04E-158	6.73E-256	0	0	0	0	0	0	0	0	0	0	0	0	8.27E-87	3.06E-10

Appendix E: Output format of prediction

The below is an extract of output file obtained from Method 1. File contains information on each segment, Cartesian coordinates, spatial orientation and the length of each sample interval.

Segm	x	y	Azim°	Dip°	from (m)	to (m)	Predict	Interval (m)
6	503549.22	2113865.32	318.10	-58.40	210.00	210.80	0.10	0.80
	503549.22	2113865.32	318.10	-58.40	210.80	211.40	0.00	0.60
	503549.22	2113865.32	318.10	-58.40	211.40	212.00	0.10	0.60
	503549.22	2113865.32	318.10	-58.40	212.00	222.00	0.66	10.00
	503549.22	2113865.32	318.10	-58.40	222.00	223.00	0.00	1.00
	503549.22	2113865.32	318.10	-58.40	223.00	224.00	0.00	1.00
7	503549.22	2113865.32	317.60	-58.10	225.00	226.00	0.00	1.00
	503549.22	2113865.32	317.60	-58.10	226.00	226.60	0.10	0.60
	503549.22	2113865.32	317.60	-58.10	226.60	227.20	0.10	0.60
	503549.22	2113865.32	317.60	-58.10	227.20	228.30	0.10	1.10
	503549.22	2113865.32	317.60	-58.10	228.30	229.15	0.10	0.85
	503549.22	2113865.32	317.60	-58.10	229.15	230.00	0.10	0.85
	503549.22	2113865.32	317.60	-58.10	230.00	231.00	0.90	1.00
	503549.22	2113865.32	317.60	-58.10	231.00	232.00	0.11	1.00
	503549.22	2113865.32	317.60	-58.10	232.00	233.00	0.11	1.00
	503549.22	2113865.32	317.60	-58.10	233.00	253.40	0.00	20.40
15	503549.22	2113865.32	318.90	-57.00	345.00	346.00	0.15	1.00
	503549.22	2113865.32	318.90	-57.00	346.00	347.00	0.15	1.00
	503549.22	2113865.32	318.90	-57.00	347.00	348.00	1.31	1.00
	503549.22	2113865.32	318.90	-57.00	348.00	348.70	0.15	0.70
	503549.22	2113865.32	318.90	-57.00	348.70	349.30	1.31	0.60
	503549.22	2113865.32	318.90	-57.00	349.30	350.00	0.15	0.70
	503549.22	2113865.32	318.90	-57.00	350.00	351.00	0.15	1.00
	503549.22	2113865.32	318.90	-57.00	351.00	352.00	0.15	1.00
	503549.22	2113865.32	318.90	-57.00	352.00	353.00	1.27	1.00
	503549.22	2113865.32	318.90	-57.00	353.00	354.00	1.26	1.00
	503549.22	2113865.32	318.90	-57.00	354.00	355.00	0.16	1.00
	503549.22	2113865.32	318.90	-57.00	355.00	356.00	0.16	1.00
	503549.22	2113865.32	318.90	-57.00	356.00	357.00	0.16	1.00
	503549.22	2113865.32	318.90	-57.00	357.00	358.00	0.16	1.00
	503549.22	2113865.32	318.90	-57.00	358.00	358.40	0.16	0.40
33	503549.22	2113865.32	318.10	-58.40	210.00	210.50	0.10	0.50
	503549.22	2113865.32	318.10	-58.40	210.50	211.00	0.10	0.50
	503549.22	2113865.32	318.10	-58.40	211.00	211.50	0.10	0.50
	503549.22	2113865.32	318.10	-58.40	211.50	212.00	0.10	0.50
	503549.22	2113865.32	318.10	-58.40	212.00	212.50	0.10	0.50
	503549.22	2113865.32	318.10	-58.40	212.50	213.00	0.10	0.50
	503549.22	2113865.32	318.10	-58.40	213.00	213.50	0.10	0.50
	503549.22	2113865.32	318.10	-58.40	213.50	214.00	0.92	0.50
	503549.22	2113865.32	318.10	-58.40	214.00	214.50	0.10	0.50
	503549.22	2113865.32	318.10	-58.40	214.50	215.00	0.10	0.50
	503549.22	2113865.32	318.10	-58.40	215.00	227.00	0.00	12.00