University of Alberta

Multivariate Spatial Modeling of Metallurgical Rock Properties

by

Jared Luke Deutsch

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Mining Engineering

Department of Civil and Environmental Engineering
University of Alberta

Abstract

High resolution spatial numerical models of metallurgical properties constrained by geological controls and more extensively measured grade and geomechanical properties constitute an important part of geometallurgy. The spatial modeling of metallurgical rock properties has unique challenges. Metallurgical properties of interest may average nonlinearly, and the nonlinear behaviour may be unquantified due to substantial costs associated with sample collection and testing. The large scale of the samples presents an additional challenge in the modeling of these variables as the support volume for metallurgical properties may be 1-2 orders of magnitude larger than typical metal assays. Practical challenges including the highly multivariate nature of geometallurgical data sets, undersampling and complex optimization requirements complicate the problem.

Addressing these challenges requires an integrated statistical approach. In this thesis, a consistent framework for quantifying and modeling the nonlinear behaviour of metallurgical rock properties is introduced. This integrated approach is composed of three parts: a nonlinear modeling and inference strategy, a multivariate downscaling algorithm, and an integrated geostatistical approach to multivariate modeling of metallurgical properties.

The first contribution of this thesis is a novel semi-parametric Bayesian updating algorithm which has been developed to infer nonlinear behaviour given multiscale measurements of metallurgical rock properties and related linear properties. This approach may be applied to fit a power law which is demonstrated to be a flexible model for nonlinear modeling.

The second contribution addresses the challenge of highly multiscale data by the development of a direct sequential simulation method for the down-

scaling of metallurgical rock properties given highly multivariate information. The stochastic downscaling procedure developed is exact and respects intrinsic constraints, such as requirements for non-negativity.

The third contribution is the development of a consistent framework for geostatistical modeling of metallurgical variables in the presence of constraints, nonlinear variables, multiscale data, missing data, and complex relationships. This approach, and a number of the algorithms developed in this thesis are applied in a geometallurgical case study of a South American copper-molybdenum porphyry deposit. The thesis statement: an integrated statistical approach for the multivariate spatial modeling of metallurgical rock properties will lead to better mine and mill operation strategies to maximize mine value. Developments in this thesis facilitate the integrated approach which is applied to the case study demonstrating the value of this integrated statistical framework.

Preface

Parts of this thesis have been previously published, or are in the publication process.

Chapters 2, 5 and 6 are composed in part by Deutsch, J.L., Palmer, K., Deutsch, C.V., Szymanski, J. and Etsell, T.H., *Spatial Modeling of Geometallurgical Properties: Techniques and a Case Study*, published by Natural Resources Research as an original research article.

Chapter 3 is composed in part by Deutsch, J.L., Szymanski, J. and Etsell, T.H. (2014) Metallurgical Variable Re-expression for Geostatistics, presented at 16th Annual International Association for Mathematical Geology Conference in New Delhi, India, as an original work.

Chapter 4 is composed in part by Deutsch, J.L., Szymanski, J., Etsell, T.H. and Deutsch, C.V. (2015) Downscaling and Multiple Imputation of Metallurgical Variables, presented at 17th Annual International Association for Mathematical Geology Conference in Freiberg, Germany, as an original work.

Acknowledgements

I would like to thank my supervisors, Dr. Jozef Szymanski, Dr. Thomas Etsell and Dr. Hooman Askari-Nasab, for their expert guidance. The endless ideas and invaluable advice of Dr. Clayton Deutsch were inspiring and essential to this thesis. I am truly fortunate to have had the support of such enthusiastic and knowledgeable professors.

I also wish to thank Dr. Jeff Boisvert, Dr. Andreas Hamann and Dr. Jaime Gómez-Hernández for taking the time to review my work so thoroughly and provide valuable feedback on all aspects of this thesis. Practical advice and discussions with Bryan Rairdan, Kevin Palmer and Rodrigo Marinho are also much appreciated.

Financial support from sponsors of the Centre for Computational Geostatistics and the Natural Sciences and Engineering Research Council of Canada is gratefully acknowledged. Many researchers at the Centre for Computational Geostatistics helped refine rough ideas along the way; I would especially like to thank Dr. Ryan Barnett, Dr. Daniel Silva and Dr. Daniel Khan for so many insightful modeling discussions.

Finally, I never would have gotten this far without the support and encouragement of my family and friends. To Clayton and Pauline, Matthew and Amanda, Kalen and Adrienne, and most of all to Kelsey, thank you. This thesis is dedicated to you.

Table of Contents

List of Figures

List of Tables

List of Symbols

Symbol	Definition	First Use
k	a data type	10
K	number of data types	10
v	a discrete data scale (a volume or, less frequently, a mass)	10
$\mathbf{u}$	a location vector	10
A	a modeling or geologic domain	11
z	an outcome of the random variable Z	11
Z	a random variable	11
l	a realization (of the data, model, or otherwise)	11
L	number of realizations	11
n_k	number of sample locations of data type k	11
v_c	composite data scale	11
v_m	metallurgical test scale	11
v_{gm}	geomechanical data scale	11
E	expected value operator	11
m_Z	stationary mean of Z	11
$\mathbf{h}$	separation vector	11

C	covariance	11
ρ	correlation	11
λ	estimation weight	14
σ_E^2	estimation variance	14
P	a probability	18
f	a probability distribution function	22
F	a cumulative probability distribution function	25
ω	a power applied to the power transformation family	42
p	a proportion or variable of interest	44
y	linear variable outcome, or Gaussian outcome	45
Y	linear random variable re-expression	45
ϕ	linearizing re-expression function	45
x_i	a proportion of material in a blend	49
O	an objective function for optimization	67
β	optimization relaxation parameter	67
J	Jacobian matrix of partial derivatives	67
ϵ	error on a value	70
H	bandwidth or entropy or Hermite polynomial	88
I	information, or mutual information	90
U	uncertainty coefficient (normalized mutual information)	90
e	data event with a location, scale, type, and value	121
$P50$	value corresponding to the 50th percentile (median)	151

List of Abbreviations

Abbrv.	Definition	First Use
PPMT	projection pursuit multivariate transform	29
SCT	stepwise conditional transform	29
PCA	principal component analysis	29
MAF	minimum/maximum autocorrelation factors	29
RQD	rock quality designation	41

Chapter 1

Introduction

1.1 Problem setting and background

The last century has seen an increased focus on the mining of massive, low grade ore deposits. Increased metal demand and improved geologic understanding, spatial modeling, and mining and metallurgical advancements have made the exploitation of these low grade ores possible. With all mineral deposits, and particularly for low grade ores, the efficient planning and optimal design of mines and metallurgical plants is necessary to maximize metal recovery and mine economics. Metallurgical properties such as throughput are as important as grades to mine economics for these deposits. Geometallurgy is being put forward as an integrated approach that will improve mine and mill efficiency and increase mine value. Geometallurgy is a recently used term that refers to the characterization and understanding of the relationship of mineralogy and mineral textures to processing attributes (Lund et al., 2014), the characterization of metallurgical properties at a fine scale (Kuhar et al., 2011), and the spatial distribution and scaling of rock and metallurgical properties (Boisvert et al., 2013; Keeney and Walters, 2011). In this thesis, the focus is on the spatial modeling of metallurgical properties for variability and uncertainty quantification. The unifying vision of geometallurgy supported in this thesis is to integrate knowledge about the rock, deposit, process and mine value creation (Figure 1.1). The ideas of geometallurgy are not new, but increased knowledge of geological, mining, metallurgical and modeling processes

combined with decreasing average ore grades make this approach particularly attractive now.

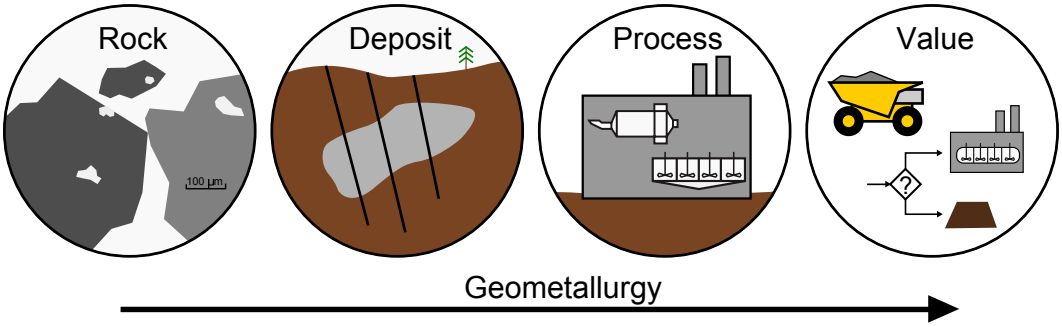

Figure 1.1: Cartoon with the geometallurgy vision proposed for this thesis.

Traditionally, research and work in these areas have been largely distinct. Microscopic rock properties are characterized by mineralogists. An emphasis is placed on understanding mineral associations and the scale of valuable minerals. Spatial modeling of the deposit is undertaken by geostatisticians and geologists who quantify the mineral resource in the deposit. Estimates and the quantification of uncertainty in metal grades and tonnages are emphasized. The processing plant for the concentration and extraction of valuable minerals is designed and operated by metallurgists. The plant is designed to optimize metal recovery based on a composite sample deemed representative of the mineral deposit. Value creation by mining and ore/waste classification is undertaken by mining engineers. The emphasis is on maximizing resource recovery and minimizing the cost of mining.

The geometallurgy vision supported in this research is to integrate these areas; an understanding of mineralogy and the variability of the deposit would be used to improve process design and operation. Metallurgical properties would be spatially modeled in the deposit and used for mine and mill planning and design. The mine would be designed to maximize mill throughput and ore optimally classified according to its metallurgical properties for multiple processing routes. Optimal campaigning and blending strategies would be used to maximize mine value accounting for the uncertainty in our understanding of the deposit. The vision of geometallurgy is compelling, but there are many engineering challenges to be addressed. In this thesis, the emphasis is on

challenges associated with the multivariate spatial modeling of metallurgical properties.

1.1.1 Metallurgical processes and properties

Mineral processing and metallurgical processes cover a very large range of operations from size reduction and liberation (comminution), mineral concentration, extraction and recovery. Although the unit operations used in metallurgical processes are diverse, there are a number of common process categories. Consider an open pit mine where ore is mechanically extracted and transported to the mill. A significant challenge is the separation of ore components by liberating and concentrating the valuable minerals, such as a copper bearing chalcopyrite ore, and converting them to a useful metallic form. One of the most common methods for this is to comminute the run-of-mine ore, concentrate the valuable particles using a concentration operation such as froth flotation, and convert the minerals using an aqueous chemical (hydrometallurgical), heat (pyrometallurgical) or electric (electrometallurgical) process.

These processes are expensive and energy intensive. Comminution in particular is the single largest energy consumer in almost all mining operations, with energy requirements on the order of 36% of all energy used by copper and gold mines in Australia (Ballantyne et al., 2012). The high energy usage of comminution represents a substantial operating cost. Maximizing ore throughput at minimal cost is critical for a low grade, high tonnage mine. Mineral concentration operations, including froth flotation, have a direct impact on metal recovery, as material not recovered in mineral processing operations will be treated as waste. The recovery in metallurgical processes, reagent usages for chemical processes and energy requirements for heat and electricity are also important factors.

These common processes point to some important metallurgical properties for economic characterization of a mineral ore. Energy requirements, throughput, flotation recovery, and reagent usage are all metallurgical properties which should be quantified. For the high resolution statistical approach proposed in

this thesis, these properties are modeled spatially at a high resolution to optimize the mining and processing operations.

1.1.2 The place of spatial models

Spatial models of rock properties are required for the evaluation of mineral deposits. These models are constructed with the goal of accurately quantifying spatial variability and joint uncertainty in rock properties conditional to available data. Spatial models are then used for the establishment of ore/waste boundaries, resource and reserve estimates, and mine planning. Limited drill hole data and a lack of knowledge about underlying geological phenomena make the construction of these models challenging. Uncertainty stemming from limited data about the subsurface is modeled and accounted for to make the best decisions under uncertainty.

Spatial models are commonly constructed of metal grades for mine planning, but these are not used as extensively in the subsequent design and optimization of mineral processing and extractive metallurgy plants. The design and performance of these plants is a function of a large number of metallurgical rock properties including comminution, flotation, and leaching indices. Spatial models used for mine design rarely include models for these metallurgical variables for a number of reasons. The tests and labwork required are expensive and, in many cases, require large volumes of rock from limited core samples. In addition, modeling process driven metallurgical properties such as the Bond mill work index and flotation recovery from intrinsic rock properties such as metal assays is complicated by multivariate, nonlinear relationships. These relationships are often not fully understood, and there are limited data available to infer them. Although the problem is complex, there is significant motivation to model these properties.

Unquantified variability in metallurgical properties is often compensated for by incorporating large engineering safety factors into mineral processing and extractive metallurgy operations. The current practice of over design, excessive material re-handling and blending, and human dynamic response factors is not optimal. Spatial models of metallurgical rock properties could

be used to improve mill design and operation by designing the mill to handle the expected range in properties, rather than designing the mill around a single homogeneous composite of the expected ore feed. The mine plan could also be designed to support mill requirements and increase ore throughput and metal recovery. The value in building spatial models is apparent, but there are a number of challenges which are the focus of this thesis.

Relating process driven metallurgical properties to geological rock properties is challenging. The metallurgical properties of interest are influenced both by process conditions and the intrinsic geological rock properties. The relationship is dependent on a large number of factors, so an empirical regression model is often developed to estimate the metallurgical response from intrinsic geologic rock properties. Although there is uncertainty in the metallurgical response, this is not often captured in the regression model. With a large number of rock properties available and relatively sparse metallurgical measurements, there is the danger of model overfitting and issues of scale. The metallurgical rock properties are undersampled relative to grade variables and sampled on supports which may differ by 1-2 orders of magnitude. Regularizing the geologic rock properties to the scale of the metallurgical properties may not be possible; the selective mining unit size may be smaller than the scale of metallurgical measurements.

The spatial modeling of metallurgical variables also has issues relating to the establishment of a representative distribution, the highly multivariate nature of the problem, missing data values due to undersampling and difficulties applying a linear geostatistics methodology. Due to the time, expense and scale associated with metallurgical tests, samples may be preferentially taken in areas of the deposit likely to be classified as ore in the mine plan. An unbiased estimate of the statistical distribution of these properties is required for modeling, so the preferential sampling must be accounted for. For the complete geometallurgical characterization of a deposit, many variables are important including assays, grinding indices, recoveries, specific gravity and geotechnical properties. Geostatistical techniques which will reproduce multivariate relationships in the data including multivariate transforms and

colocated cosimulation algorithms require complete data sets without missing values. Data imputation may be required in this case. Spatial modeling and geostatistical techniques which respect the nonadditivity are also required; linear geostatistics frameworks such as kriging cannot be applied to nonadditive metallurgical variables. Assuming a linear average would bias the resulting models.

Methods for the scaling and mixing of metallurgical properties are required to use the spatial models for improved mine and mill design and operation. The point scale geostatistical models must be blocked up to a selective mining unit scale, and these mining units combined to determine the blended response in the mill. For metallurgical variables which average nonlinearly, an averaging schema is required which relates a linearly scaling variable such as grade or ore type proportion to the metallurgical variable of interest. Metallurgical variables which average nonlinearly are referred to as nonlinear variables throughout this thesis; however, it is emphasized that the variables themselves are not nonlinear, averages of these variables are nonlinear. These averaging schema must be determined experimentally. Limited samples mean that there is substantial uncertainty and the mixed response of metallurgical properties should be evaluated probabilistically.

The goal of addressing these challenges is to improve mine planning and mill design and operation. This may include the creation of optimal blending or optimal campaigning strategies for specific high value ore types. If multiple processing paths are available, the ore should be optimally typed to maximize the net present value of the mine. All of the plans and designs must account for the uncertain, multivariate metallurgical rock properties.

1.2 Thesis statement

To address these challenges, an integrated statistical approach for the spatial modeling of metallurgical rock properties is proposed in this thesis. The proposed approach provides a framework for addressing the problems of 1) inferring and modeling nonlinear behaviour and mixing laws for metallurgical

variables, 2) the downscaling of large scale metallurgical samples, 3) the spatial modeling of metallurgical properties, and 4) challenges in the selection of multivariate modeling workflows for uncertainty evaluation with metallurgical properties. The thesis statement:

An integrated statistical approach for the multivariate spatial modeling of metallurgical rock properties which accurately quantifies the joint uncertainty in grade, geotechnical and metallurgical properties will lead to more effective mine and mill operation strategies.

The integrated statistical approach is composed of approaches for the spatial modeling of nonlinear metallurgical variables, large scale metallurgical samples and multivariate approaches for geostatistical modeling. The nonlinear behaviour of metallurgical variables with limited data is inferred using a semiparameteric Bayesian updating approach with nonlinear regression. Using models of the nonlinear behaviour, the metallurgical variables are re-expressed as linear variables which can be modeled using standard geostatistical algorithms, and back-transformed after modeling. Large scale metallurgical samples are first re-expressed as linear variables, then downscaled using direct sequential simulation with a stable downscaled histogram. These approaches are integrated with multivariate decorrelation and cosimulation algorithms for jointly realizing grade and metallurgical properties in the deposit.

1.3 Thesis outline

Chapter 2 reviews relevant literature and establishes the geometallurgical modeling prerequisites for this thesis. The focus is on techniques which form the basis for methods developed in this thesis, and geostatistical concepts that are used throughout. Chapter 3 introduces a consistent methodology for modeling nonlinear metallurgical rock properties, and several methods for inferring the mixing laws given varying degrees of information. A novel algorithm using semiparametric Bayesian updating to infer nonlinear behaviour given multiscale metallurgical data is developed. Chapter 4 focuses on the downscaling

of large scale metallurgical properties for scale consistent spatial modeling. An algorithm for stochastic downscaling using direct sequential simulation is developed with a focus on practical considerations given the challenging constrained, nonlinear, and multiscale problems. Chapter 5 addresses practical challenges for the construction and application of high resolution geometallurgical spatial models for improved mine planning and process operation. Chapter 6 applies a multivariate geostatistical workflow using metallurgical variables to a copper porphyry case study. Conclusions, limitations and future work are reviewed in chapter 7.

Each of the methods developed in this thesis are numerically and computationally intensive, requiring specialized software for practical application. A substantial contribution of this thesis is the development of numerical modeling software which implements the algorithms developed here. Computationally intensive algorithms are implemented in modern Fortran, and wrapped in Python for integration into geostatistical workflows.

Chapter 2

Geometallurgical modeling prerequisites

In chapter 1, geometallurgy was broadly defined to be the integration of data from numerous sources about the rock, deposit, process, and mining. In this thesis, this inclusive vision of geometallurgy is adopted from a spatial numerical modeling perspective. The regionalized variable framework of classical geostatistics is used as the foundation for spatial modeling of geometallurgical properties which quantify grade, metallurgical and geotechnical information. These spatial models form the basis of applications for mine value creation.

This chapter begins with a brief overview of the regionalized variable framework for geostatistics and defines notation used throughout this thesis. The focus then shifts to specific methodologies which are used, and which form the basis for new modeling algorithms and strategies developed. These methodologies are placed within the context of practical methods to build a spatial model of geometallurgical properties given the challenging multivariate, nonlinear, and multiscale nature of the problem. Techniques introduced here are developed and integrated into geostatistical workflows for the construction of high resolution geometallurgical models in chapter 5. Alternative techniques for spatial model construction are also discussed in chapter 5 in the context of workflow selection.

2.1 Regionalized variables and the geostatistical framework

With the goal of building spatial models of geometallurgical properties, the classical geostatistics framework is adopted. Geostatistics is a branch of applied statistics characterized by "the statistical study of natural phenomena" (Journel and Huijbregts, 1978). Specifically, the results of deterministic natural processes, such as the copper grade in a deposit, are numerically modeled by the application of probabilistic methods. A probabilistic model is adopted due to a lack of information about the deterministic processes and a lack of knowledge about the governing spatial law of the end result. This lack of information, and the engineering requirement to make the best decisions under uncertainty, motivate the adoption of the probabilistic approach of geostatistics.

The probabilistic statistical methods used are diverse including linear estimators (e.g., kriging), stochastic evaluation of conditional propabilites (e.g., sequential Gaussian simulation, Bayesian updating) and the application of spatial covariances (e.g., scaling laws) among others. In the context of geometallurgy, the focus of these methods is the spatial modeling of rock properties (grades, metallurgical and geomechanical properties) which are considered to be regionalized random variables. In this thesis, *properties* refers to the measured and modeled rock attributes, while *variables* refers to rock properties treated from a statistical perspective.

2.1.1 Rock properties as regionalized variables

The regionalized variable framework in this thesis is developed considering multiple data types, multiple data scales, and multiple realizations of the data. Rock properties may be continuous, categorical or mixed continuous-categorical variables. In this thesis, continuous variables including grades, work indices and geotechnical measurements are primarily the properties of interest. Categorical variables, including rock types and lithologies, and mixed

continuous-categorical variables, such as the degree and type of alteration may also be modeled.

Consider K data types, including grade, geotechnical and metallurgical properties enumerated as $k = 1, \ldots, K$. These data have an associated scale v, where v belongs to a set of discrete scales including the composite scale and larger metallurgical sample scale. Data events are considered to have a discrete scale v as data of a similar type are often collected on, or composited to, a common scale. Grade measurements may be composited to a bench scale of 15 m while metallurgical samples are often measured on 30 m intervals of core. The scale v is included to indicate the type of scale where many types are possible including:

$$v = \begin{cases} v_c \text{ composite scale} \\ v_m \text{ metallurgical test scale} \\ v_{gm} \text{ geomechanical scale} \end{cases} \tag{2.1}$$

Regionalized data events of type k are associated with a sample location vector $\mathbf{u}_i$ with sample locations indexed as $i = 1, \ldots, n_k$. Under a multiple imputation workflow, L realizations of the data are indexed as $l = 1, \ldots, L$. A random variable is denoted by the uppercase Z where the lowercase z denotes an outcome of this random variable. Outcomes include both measured and modeled values of the random variable. The random variables are regionalized, and restricted to be within a domain A. Data events are denoted:

$$z_k^l(\mathbf{u}_i; v), i = 1, \ldots, n_k, k = 1, \ldots, K, l = 1, \ldots, L$$

Unless realizations of the data or unknown locations are being used, equations in this thesis omit the data realization index l to simplify the notation. Using these data, stochastic and deterministic spatial models of the deposit are constructed. These models require statistical characterization and parameterization of the regionalized random variables for use with geostatistical models.

2.1.2 Parameterization of geostatistical models

A prerequisite for geostatistical modeling is the establishment of representative statistical relationships. This includes univariate statistics (the histogram

including the mean and variance), bivariate spatial statistics in the form of covariance functions or variograms, and multivariate statistics including correlation matrices. Statistics are calculated within a domain assuming second-order stationarity (Chilès and Delfiner, 2009). Two principal relationships are implied by second-order stationarity; the expected value of a random variable at a constant scale is the variable's mean which is constant within the domain:

$$E\{Z_k(\mathbf{u}); v\} = m_k(v), k = 1, \ldots, K, \mathbf{u} \in A$$

where E is the expected value operator and m denotes the mean. Additionally, the covariance $C_k(\mathbf{h}; v)$ within the domain follows a covariance function for a separation vector $\mathbf{h}$:

$$E\{[Z_k(\mathbf{u}; v) - m_k(v)][Z_k(\mathbf{u} + \mathbf{h}; v) - m_k(v)]\} = C_k(\mathbf{h}; v), k = 1, \ldots, K, \mathbf{u} \in A$$

For a standardized variable with unit variance, the correlation function $\rho(\mathbf{h}; v)$ is used in place of the covariance function. There is no requirement for the spatial covariance to be evaluated on data of the same scale. Multiscale covariances, referred to as average covariances, are numerically calculated using a discrete approximation. The discretization level is typically the smaller of the two scales, or a pseudo-point scale. The average covariance is calculated:

$$\overline{C}_k(\mathbf{u}_i, \mathbf{u}_j; v_i, v_j) = \frac{1}{n_{disc(i)} n_{disc(j)}} \sum_{\alpha=1}^{n_{disc(i)}} \sum_{\beta=1}^{n_{disc(j)}} C_k(\mathbf{u}_\alpha, \mathbf{u}_\beta; v_{disc})$$

Alternatively, average covariances may be calculated with the analytical solution (sextuple volume integral) of the prior equation (Journel and Huijbregts, 1978). The semivariogram $\gamma(\mathbf{h}; v)$, hereafter referred to as the variogram, is the most common statistic used to infer the covariance function. The variogram is defined:

$$\gamma_k(\mathbf{h}; v) = \frac{1}{2} E\{[Z_k(\mathbf{u}; v) - Z_k(\mathbf{u} + \mathbf{h}; v)]^2\}, k = 1, \ldots, K, \mathbf{u} \in A$$

Under the assumption of second-order stationarity the variogram and covariance are related by the variance $\sigma_k^2(v)$:

$$\gamma_k(\mathbf{h}; v) = \sigma_k^2(v) - C_k(\mathbf{h}; v), k = 1, \ldots, K$$

The decision of stationarity is a choice required to draw conclusions from limited data, and is explicitly made as soon as any statistics are calculated and modeled. Inference of these statistics is complicated by preferential sampling in the mining industry. This is a result of the sampling goal being to identify ore bodies and delineate ore-waste contacts, not to draw representative samples from a population.

In the case of preferential sampling and in the presence of spatial continuity, a declustering technique is required to infer a representative distribution (Deutsch, 1989; Journel, 1983). The cumulative distribution function F of each property is inferred in the presence of clustered samples:

$$F_k(z_k(v)), k = 1, \ldots, K$$

In the absence of spatial continuity, all samples are independent and there is no need for declustering (Bourgault, 1997). The physical and chemically driven metallurgical properties that are being modeled are taken from a natural deposit, so will exhibit a degree of spatial correlation. In the situation where the full range of variability has been sampled, but areas of the deposit have been preferentially sampled, a polygonal, cell or kriging based declustering method could be used (Bourgault, 1997; Kovitz and Christakos, 2004; Olea, 2007). If the data collected on the metallurgical variable of interest are very limited, but extensive secondary data such as metal assays are available, then declustering with secondary data could be considered (Deutsch et al., 1999; Pyrcz and Deutsch, 2002).

2.1.3 Theory of geostatistical estimation and simulation

Numerical models of a mineral deposit can be broadly classified as deterministic, where there is a single model of the deposit, or stochastic, where there are multiple realizations of the deposit. From a statistical perspective, these correspond to estimated (deterministic) and simulated (stochastic) models of the deposit. The model purpose and study aims will dictate whether the model constructed will be deterministic with a single best estimate or stochastic with multiple realizations of the domain. Modeling methods developed in this thesis

use a stochastic approach for multiple realizations; however, the principles of linear spatial estimation are ubiquitous even in the multiple realization context and are therefore reviewed here.

If the goal of the study is to infer a single best estimate of a set of rock properties over an entire domain for estimating ore tonnes, grade and for resource classification, then a deterministic approach such as ordinary kriging may be used as it provides a robust method for the spatial estimation of rock properties in a geologic domain (Rossi and Deutsch, 2013). As the kriging formalism forms the basis of many of the geostatistical techniques used in this thesis, including the stochastic methodologies, consider the linear estimation of z_k^* at location $\mathbf{u_0}$ and scale v given a set of nearby, spatially correlated data of the same data type k:

$$z_k^*(\mathbf{u}_0; v) - m_k(v) = \sum_{i=1}^{n_k} \lambda_i (z_k(\mathbf{u}_i; v) - m_k(v)) \tag{2.2}$$

where $\lambda_i, i = 1, \ldots, n_k$ are estimation weights assigned to the n_k sample locations with data of type k. The estimation variance $\sigma^2_{k,E}$ of this linear estimator is calculated given the weights and covariances $C_k(v)$ where $C_{i,j}(v)$ is shorthand for the covariance between two locations, $\mathbf{u}_i$ and $\mathbf{u}_j$ with scale v:

$$\sigma^2_{k,E}(\mathbf{u}_0; v) = C_{0,0}(v) - 2\sum_{i=1}^{n_k} \lambda_i C_{i,0}(v) + \sum_{i=1}^{n_k}\sum_{j=1}^{n_k} \lambda_i \lambda_j C_{i,j}(v) \tag{2.3}$$

which is minimized by the simple kriging equations (alternatively referred to as the normal equations):

$$\begin{aligned} \sum_{j=1}^{n_k} \lambda_j C_{i,j}(v) &= C_{i,0}(v), i = 1, \ldots, n_k \\ \sigma^2_{k,SK}(\mathbf{u}_0; v) &= C_{0,0}(v) - \sum_{i=1}^{n_k} \lambda_i C_{i,0}(v) \end{aligned} \tag{2.4}$$

Discussion and derivation of alternative kriging variants can be found in any standard geostatistical text (Chilès and Delfiner, 2009; Goovaerts, 1997; Journel and Huijbregts, 1978). Using variables in their original units, these deterministic approaches require an assumption that the variable averages linearly. Nonlinear estimation algorithms, including disjunctive, lognormal and

indicator kriging have been shown by Moyeed and Papritz (2002) to not perform better than linear techniques comparing the bias, relative mean squared error and other metrics of estimates. Moreover, the so-called nonlinear estimation algorithms are still aimed at variables that average linearly. For this reason, deterministic approaches, while applicable for linear variables where a single best estimate is required, should be avoided for metallurgical and geomechanical variables that do not average linearly.

For most geometallurgical applications the joint uncertainty of many variables is important, and there is uncertainty in model parameters. In these cases a stochastic approach using Monte Carlo simulation should be applied. This approach would be preferred for many applications including choosing an appropriate drill hole spacing and assessing plant and mine performance based on multiple rock, metallurgical and geomechanical properties. Stochastic methods are used as they can jointly model the spatial and multivariate uncertainty (Pyrcz and Deutsch, 2014; Rossi and Deutsch, 2013). Simulation makes no averaging assumption; all averaging is deferred until the end of the modeling workflow where point scale simulated values are averaged up to larger scales relevant for the mining operation. If the nonlinear behaviour of the metallurgical variables of interest is understood, nonlinear averaging can be applied at this point. Multiple high resolution realizations of the deposit are constructed which can then be post-processed for summary statistics and through a transfer function such as mining simulation for quantifying uncertainty in key parameters such as the uncertainty in mill throughput and head grade throughout the mines life. Techniques developed in this thesis are in the domain of the stochastic approach to deposit modeling. If required, realized values can be averaged across realizations for a single deterministic estimate at each location.

Sequential Gaussian simulation techniques which rely on the recursive decomposition of the multivariate Gaussian probability distribution are the most widely used stochastic modeling algorithms for continuous variables. The algorithm is reviewed here, although the interested reader is referred to any standard geostatistical text for a more detailed exposition. All data values,

$z_k^l(\mathbf{u}_i; v), i = 1, \ldots, k = 1, \ldots, K, n_k, l = 1, \ldots, L$ are transformed to Gaussian units, typically with the use of a normal score transform (also referred to as a probit or quantile function; Bliss, 1934; Deutsch and Journel, 1998). The Gaussian data are enumerated:

$$y_k^l(\mathbf{u}_i; v), i = 1, \ldots, n_k, k = 1, \ldots, K, l = 1, \ldots, L$$

Importantly, the data are assumed to be multivariate Gaussian after univariate transform. A number of check statistics for possible multivariate Gaussianity may be checked, however the verification of true multivariate Gaussianity is not possible without multiple realizations of the data which are not available for these physically sampled measurements. With the assumption of multivariate Gaussianity, model locations are visited sequentially and the normal equations (simple kriging) are applied to calculate the mean and variance as per Equations 2.2–2.4. A random residual $R(\mathbf{u}_0)$ is sampled from a Gaussian distribution with a mean of zero and variance equal to the estimation (kriging) variance. The result is a set of simulated values with the correct variance and conditional covariances. This is added to the estimate:

$$y_s^*(\mathbf{u}_0; v) = y^*(\mathbf{u}_0; v) + R(\mathbf{u}_0)$$

The simulation proceeds sequentially visiting each model location in turn. This process is repeated for each realization, for each variable, and for all locations within the stationary modeling domain A:

$$\{y_k^l(\mathbf{u}_i; v), i = 1, \ldots, n_{\mathbf{A}}, l = 1, \ldots, L, k = 1, \ldots, K, \mathbf{u} \in A\}$$

Data are back-transformed to original units, and upscaled to larger support volumes relevant for mining as required. This Gaussian approach is the primary algorithm for stochastic simulation of the deposit, and the primary approach used for modeling within this thesis. Alternative methods to the sequential approach include spectral methods (Fourier simulation) and turning bands (Journel and Huijbregts, 1978).

There are a number of problems with the sequential Gaussian simulation approach, principally the requirements for multivariate Gaussianity and maximum entropy leading to disconnected, destructured extreme values (Oz et al.,

2003). Multivariate Gaussianity can be enforced with a multivariate transfom; however, the property of maximum entropy is central to the use of the Gaussian distribution. Alternatives to the Gaussian approach, such as the multiple-point statistics approach advocated by Mariethoz and Caers (2014), require extensive training image construction. Limited sample data and the challenging geological, metallurgical and geotechnical nature of the problem limit the present applicability of these methods for multivariate metallurgical modeling. Therefore, the widely used multivariate Gaussian approach for stochastic modeling is adopted in this thesis.

2.1.4 Hierarchical modeling

Both deterministic and stochastic methods require the choice of a stationary domain for modeling. As previously discussed in the calculation of representative statistics, stationarity is a choice made to pool data together for the calculation of statistics and inference of properties. This includes all steps of the modeling workflow from the establishment of histograms and variograms to the simulation of metallurgical properties within the domain. It is rare that a deposit would only have one domain; numerous geologic controls, rock types, and intrusions are common. In these cases a hierarchical approach for modeling is adopted.

The hierarchical approach for modeling is the choice to pool data into specific subsets (rock types or domains) for statistic calculation and modeling. This may be deterministic using geologic knowledge to establish rock type boundaries and domains, or stochastic using multiple realizations of the deposit. Each domain is modeled separately (or by treating data from a separate domain as secondary data; McLennan, 2007) and the separate domain models merged. These merged models composed of multiple domains form the basis for model applications including mine planning and optimization.

Many techniques for domain modeling exist, including implicit modeling with distance functions, truncated Gaussian simulation, indicator simulation and multiple point geostatistics. The establishment of these domains, deterministically or stochastically, is the first step in the modeling process.

Techniques proposed and implemented in this thesis including semiparametric Bayesian updating for nonlinear inference and direct sequential simulation for downscaling operate within these stationary domains. All geostatistical and statistical algorithms used in this thesis are only applicable within a domain for which a decision of stationarity has been made, even if this limitation is not explicitly noted for each technique.

2.2 Bayesian updating

The basis of the multiscale nonlinear inference algorithm proposed in this thesis is Bayesian updating. Bayesian updating is the application of Bayesian inference in which a posterior, or updated, distribution is inferred given prior, likelihood and global distributions. The updated distribution is inferred using Bayes' law:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

where $P(A)$, $P(B)$ are the global probabilities of events A and B, $P(A|B)$ is the probability of event A given event B and $P(B|A)$ is the probability of event B given event A. In the context of Bayesian updating for geostatistics, an updated distribution $P(A|B)$ is calculated given the prior distribution $P(A)$, the likelihood distribution $P(B|A)$ and the global distribution $P(B)$. This requires an assumption of the relationship between events which is a primary basis for the classification of a Bayesian updating algorithm as parametric or non-parametric, not just the use of a parametric distribution.

Bayesian updating is commonly used in geostatistics (Deutsch and Zanon, 2004; Doyen and Pillet, 1996; Journel, 2002; Neufeld and Deutsch, 2004; Ren and Deutsch, 2006) for the integration of integrating secondary data in an estimate as an alternative to multivariate techniques such as cokriging. An estimate is made using data of the same type (same property and support volume) to establish the prior distribution. The likelihood distribution is established using secondary data, such as a correlated variable. Under a Markov assumption, Bayesian updating is applied on a colocated basis, where only

secondary data at the point where the estimate is being made are used, not nearby data.

2.2.1 Parametric Bayesian updating for geostatistics

Parametric Bayesian updating for geostatistics uses the multivariate Gaussian distribution, ubiquitous in geostatistical analyses, to define the relationship between the likelihood and prior distributions. Doyen et al. (1996), Deutsch and Zanon (2004) and Neufeld and Deutsch (2004) applied Bayesian updating assuming a bivariate Gaussian relationship between the likelihood and prior. Notation similar to that used by Neufeld and Deutsch is adopted in this thesis. Consider a Gaussian likelihood distribution parameterized by a mean m_L and variance σ_L^2. A correlated Gaussian prior distribution is parameterized by mean m_P and variance σ_P^2. The updated Gaussian distribution is parameterized by the updated mean, m_U, and variance, σ_U^2:

$$\begin{aligned} m_U &= \frac{m_L\sigma_P^2 + m_P\sigma_L^2}{(1-\sigma_L^2)(\sigma_P^2-1)+1} \\ \sigma_U^2 &= \frac{\sigma_P^2\sigma_L^2}{(1-\sigma_L^2)(\sigma_P^2-1)+1} \end{aligned} \tag{2.5}$$

The derivation of Bayesian updating for a bivariate Gaussian distribution can be found in any of the previously mentioned references on parametric Bayesian updating. For the common geostatistical approach of secondary data information with parametric Bayesian updating, the likelihood distribution quantifies the relationship between the primary and secondary variables. That is, given a colocated secondary value at a location being estimated, the likelihood distribution is the conditional probability distribution of the primary given the secondary variable. The prior distribution is established by kriging (normal equations) with nearby primary data. The updated distribution which merges these distributions is a function of the global distribution (the standard normal distribution), the likelihood and prior distributions. Note also that the updating equations are non-convex, and it is therefore possible to infer a mean higher than both the likelihood and prior means, and vice-versa.

Parametric Bayesian updating is one method for combining data from mul-

tiple sources, but there are others including colocated cokriging and (full) cokriging. Parametric Bayesian updating is considered easier to implement (Doyen and Pillet, 1996) and does not require modeling numerous cross variograms as required by cokriging variants. These advantages are offset by the requirement of a strong assumption of the distribution shape (bivariate Gaussianity) and potentially problematic variogram reproduction and variance inflation (Neufeld and Deutsch, 2004). Variance inflation here refers to the generation of realizations which have a variance significantly higher than the input data variance (variogram sill).

2.2.2 Non-parametric Bayesian updating for geostatistics

The multivariate Gaussianity assumption made by parametric Bayesian updating was modified by Neufeld and Deutsch (2006) with the proposal of non-parametric Bayesian updating. As discussed, parametric Bayesian updating explicitly assumes a bivariate Gaussian relationship between the likelihood and prior distributions. In non-parametric Bayesian updating, no assumption of multivariate Gaussianity is made. In place of the multivariate Gaussian assumption, the prior and likelihood distributions are assumed to either be 1) independent sources of information adopting an independence model, or 2) conditionally independent adopting a permanence of ratios model (Journel, 2002).

Non-parametric Bayesian updating discretizes the updating procedure, calculating the updated probabilities for discrete probability intervals. These intervals are analogous to bins used in the calculation of a histogram, although the choice of discretization interval is based on computational time, not data density.

Assuming independence between the likelihood $P(A|B)$ and prior distributions $P(A|C)$, updated probabilities are calculated:

$$P_I(A|B,C) = \frac{P(A|B)P(A|C)}{P(A)}$$

The assumption of independence is likely unrealistic, as a relationship between A and B, and between A and C implies a relationship between B and C. Of

course, there is no requirement that this relationship be causative. The alternative approach assuming conditional independence between the likelihood and prior distributions (Journel, 2002) and using the permanence of ratios equation:

$$P_{PR}(A|B,C) = \frac{\frac{1-P(A)}{P(A)}}{\left(\frac{1-P(A)}{P(A)}\right) + \left(\frac{1-P(A|B)}{P(A|B)}\right)\left(\frac{1-P(A|C)}{P(A|C)}\right)} \tag{2.6}$$

After probability combination, restandardization is required to ensure that the updated probabilities sum to one. As noted by Neufeld and Deutsch (2004), using the permanence of ratios approach does not magnify very low or very high probabilities as much as the independence approach. In addition to traditional modeling scenarios as a replacement for colocated cokriging, both parametric and nonparametric Bayesian updating have been applied for missing data imputation under a multiple imputation framework.

2.3 Missing data and the multiple imputation framework

Many multivariate techniques including decorrelation and multivariate transformation methods require completely equal sampling of the multivariate data. Frequently, the sampling of multiple metallurgical properties is unequal; that is, not all properties are measured at all data locations. Expensive metallurgical tests may only be performed on samples which are likely to be classified as ore, not lower grade samples. Newer metallurgical testing methods may not be available on legacy data collected earlier in the mine life. This unequal sampling is a significant problem for geostatistical modeling techniques that require equal sampling. Excluding all unequally sampled data locations from the modeling workflow would result in a substantial loss of information and may result in a bias for values which are not missing at random (Enders, 2010). Applying a regression technique to fill the missing data would artificially reduce the spatial and multivariate variability. As such, the multiple imputation framework for geologic data proposed by Barnett and Deutsch

(2015) is adopted. This technique replaces the missing values with stochastically imputed values which honor the multivariate and spatial distributions of the data. This multiple imputation approach is applied in this thesis for multiscale data imputation, and as such is reviewed here.

The multiple imputation framework adopts the geostatistical simulation paradigm of multiple model realizations and applies this to the data. Practically, a different set of imputed data will be used for each realization. The missing data are stochastically simulated conditional to all available colocated and spatially correlated samples resulting in a set of equally sampled data sets with the correct spatial and multivariate distributions. This requires the modeling of the conditional distribution of missing data given sampled data and drawing from this distribution. Methods for the establishment of this distribution available for the multiple imputation of missing data include both parametric (Gaussian) and non-parametric methods (Barnett and Deutsch, 2015). The parametric imputation framework applies the Bayesian updating approach discussed earlier using available colocated secondary values to parameterize a conditional multivariate Gaussian distribution. The non-parametric imputation framework applies kernel density estimation with an adapted Gibbs sampler to construct the conditional distribution. Once the conditional distribution is established either parametrically or non-parametrically, the missing value is drawn randomly from the conditional distribution.

In a case study conducted by Barnett and Deutsch, non-parametric multiple imputation was demonstrated to be the most effective at integrating complex multivariate features for spatial and colocated information. The non-parametric approach for multiple imputation is reviewed here following the development of Barnett and Deutsch. Consider the imputation of a missing value of type $k = p$ at location $\mathbf{u}_0$. As with the Gaussian simulation algorithms, all variables $Z_k(v), k = 1, \ldots, K$ are normal score transformed using the univariate cumulative distribution function to $Y_k(v), k = 1, \ldots, K$. A Gaussian mean and variance for the missing value $y_p(\mathbf{u}_0; v)$ is inferred using

the normal equations and all locally available primary data.

$$\begin{aligned}
\bar{y}_p(\mathbf{u}_0; v) &= \sum_{\alpha=1}^{n_\alpha} \lambda_{Y_p,\alpha} Y(\mathbf{u}_\alpha; v) \\
\sigma_p^2(\mathbf{u}_0; v) &= 1 - \sum_{\alpha=1}^{n_\alpha} \lambda_{Y_p,\alpha} \rho(\mathbf{u}_0 - \mathbf{u}_\alpha; v) \\
f(y_p(\mathbf{u}_0; v) | Y_p(\mathbf{u}_\alpha; v), \alpha = 1, \ldots, n_\alpha) &\sim N(\bar{y}_p(\mathbf{u}_0; v), \sigma_p^2(\mathbf{u}_0; v))
\end{aligned} \tag{2.7}$$

Secondary information is integrated using multivariate kernel density estimation with a correlated Gaussian kernel. In the method proposed by Barnett and Deutsch, a Gibbs sampler approach is used so that all $K-1$ secondary variables are available at $\mathbf{u}_0$. The conditional distribution inferred using kernel density estimation given all available secondary data at this colocated location:

$$\begin{aligned}
f(y_p(\mathbf{u}_0; v) | Y_k(\mathbf{u}_0; v), &k = 1, \ldots, K, k \neq p) \sim \\
&\frac{1}{\Delta} \int N_h(y_p(\mathbf{u}_0; v), y_k(\mathbf{u}_0; v), k = 1, \ldots, K, k \neq p) dy_p(\mathbf{u}_0; v)
\end{aligned} \tag{2.8}$$

where Δ is a normalizing constant, N_h is the bandwidth smoothed multivariate normal distribution and the integral is numerically evaluated using discrete integration. These two conditional distributions for $y_p(\mathbf{u}_0)$ are merged using Bayesian updating, which is the evaluation of the definition of a conditional probability distribution:

$$\begin{aligned}
f(y_p(\mathbf{u}_0; v) | Y_p(\mathbf{u}_\alpha; v), Y_k(\mathbf{u}_0)) = &\frac{f(y_p(\mathbf{u}_0; v) | Y_p(\mathbf{u}_\alpha; v)) f(y_p(\mathbf{u}_0; v) | Y_k(\mathbf{u}_0; v))}{f(y_p; v)}, \\
&k = 1, \ldots, K, k \neq p, \alpha = 1, \ldots, n_\alpha
\end{aligned} \tag{2.9}$$

This merged distribution is sampled using Monte Carlo simulation to draw a realization $y_p^l(\mathbf{u}_0; v)$. This procedure is applied to all missing data locations for each realization. The result is a set of L realizations of the data. For the stochastic modeling approach adopted, each realization of the block model uses a different realization of the data. The direct sequential simulation downscaling approach developed in this thesis is related to the multiple imputation workflow; multiple realizations of the small scale values are constructed.

2.4 Direct sequential simulation

The use of direct sequential simulation for stochastic downscaling is proposed in this thesis. Direct sequential simulation is a modification of sequential Gaussian simulation in which values are simulated in original units, instead of their Gaussian transform. Under the simple kriging principle, the spatial covariance of a variable is reproduced if the sampled distribution has the correct mean and variance (Deutsch, 2000; Gómez-Hernández et al., 2005; Oz et al., 2003; Soares, 2001). Direct sequential simulation forms the core of the downscaling algorithm developed in this thesis, so is briefly reviewed here.

2.4.1 Algorithm for direct sequential simulation

Direct sequential simulation is developed using notation and logical development similar to that used by Deutsch (2000). Consider a standardized random variable, $Y_k(\mathbf{u}; v), k = p$ such that:

$$y_p(\mathbf{u}; v) = \frac{z_p(\mathbf{u}; v) - m_p(v)}{\sigma_p(v)}$$

After standardization, the correlation function $\rho(\mathbf{h}; v)$ is used in place of the covariance function for $Y_p(\mathbf{u}; v)$, the mean is zero, and variance one. Direct sequential simulation of $Y_p(\mathbf{u}; v)$ proceeds in much the same manner as sequential Gaussian simulation. Given standardized data of the same type and scale, $y_p(\mathbf{u}_i; v), i = 1, \ldots, n_p$, and (typically gridded) locations to be simulated within the domain at the same scale, $\mathbf{u}_j, j = 1, \ldots, n_{\mathbf{A}}$, the algorithm proceeds for each realization $l = 1, \ldots, L$:

1. Visit simulation locations $\mathbf{u}_j, j = 1, \ldots, n_{\mathbf{A}}$ in a random order where the current location being simulated is indicated by $j = sim$ and locations previously simulation are indicated $j = 1, \ldots, sim - 1$:

2. Construct the conditional distribution $F(y_p(\mathbf{u}_{sim}; v) | y_p(\mathbf{u}_i; v), y_p(\mathbf{u}_j; v), i = 1, \ldots, n_p, j = 1, \ldots, sim-1)$. This distribution is conditional to both previously simulated values and data.

3. Realize a value $y_p^l(\mathbf{u}_{sim}; v)$ from the conditional distribution.

4. Continue visiting simulation locations and repeat for all realizations $l = 1, \ldots, L$.

Recall that sequential Gaussian simulation is equivalent to the recursive decomposition and sampling of the multivariate Gaussian distribution. Similarly, direct sequential simulation is equivalent to sampling from the data conditional distribution:

$$F(y_p(\mathbf{u}_j; v), j = 1, \ldots, n_{\mathbf{A}} | y_p(\mathbf{u}_i; v), i = 1, \ldots, n_p)$$

where the conditional density function of each sampled distribution is the product:

$$\begin{aligned} f(y_p(\mathbf{u}_{sim}; v) | y_p(\mathbf{u}_i; v), y_p(\mathbf{u}_j; v), i = 1, \ldots, n_p, j = 1, \ldots, sim-1) = \\ \prod_{J=1}^{sim-1} f(y_p(\mathbf{u}_{sim}; v) | y_p(\mathbf{u}_i; v), y_p(\mathbf{u}_j; v), i = 1, \ldots, j = 1, \ldots, J) \end{aligned} \tag{2.10}$$

This leverages the simple kriging principle which is the generalization of the multivariate Gaussian assumption in simple kriging for any parametric distribution which is defined by two parameters: a mean and homoscedastic variance.

2.4.2 The simple kriging principle

Two parameters are required according to the simple kriging principle: a conditional mean and homoscedastic conditional variance. The conditional mean is calculated (Deutsch, 2000):

$$\begin{aligned} E\{y_p(\mathbf{u}_{sim}; v) | y_p(\mathbf{u}_i; v), y_p(\mathbf{u}_j; v), i = 1, \ldots, n_p, j = 1, \ldots, sim-1\} = \\ \sum_{i=1}^{n_p} \lambda_i y_p(\mathbf{u}_i; v) + \sum_{j=1}^{sim-1} \lambda_j y_p(\mathbf{u}_j; v) \end{aligned} \tag{2.11}$$

The homoscedastic conditional variance is calculated:

$$\begin{aligned} Var\{y_p(\mathbf{u}_{sim}; v) | y_p(\mathbf{u}_i; v), y_p(\mathbf{u}_j; v), i = 1, \ldots, n_p, j = 1, \ldots, sim-1\} = \\ 1 - \sum_{i=1}^{n_p} \lambda_i \rho(\mathbf{u}_i - \mathbf{u}_{sim}; v) - \sum_{j=1}^{sim-1} \lambda_j \rho(\mathbf{u}_j - \mathbf{u}_{sim}; v) \end{aligned} \tag{2.12}$$

where the simple kriging equations are used for the weights (Equation 2.4). If the variables are Gaussian, this is equivalent to sequential Gaussian simulation and all probability density functions must be Gaussian. Even if the

distributions are not Gaussian, as in the case of direct sequential simulation, the covariance between the simulated location, and previously simulated locations and data values is correct (Deutsch, 2000). The correct sampling mean and variance are therefore known; however, the correct shape of the sampling distribution is unknown except in the Gaussian case. There have been many proposed methods for this; these are discussed further in chapter 4 in the context of the downscaling algorithm.

2.5 Multivariate geostatistical modeling

A number of geostatistical algorithms including kriging, sequential Gaussian simulation, Bayesian updating and direct sequential simulation have been reviewed, but these form only a small subset of algorithms used in a multivariate geostatistical model, such as the model constructed in chapter 6. One of the primary contributions of this thesis is the synthesis of many geostatistical algorithms into a consistent workflow for the multivariate spatial modeling of grade, metallurgical and geotechnical properties. The goal of this section is not to provide a laundry-list of algorithms used in the construction of a multivariate geostatistical model, but instead to focus on a few key geostatistical algorithms used throughout this thesis. These core algorithms include the bootstrap, intrinsic supersecondary cokriging and decorrelation algorithms. Bootstrap algorithms are used for the inference of uncertainty in statistics given limited data and no theoretical sampling distribution. Intrinsic supersecondary cokriging and decorrelation algorithms are two alternative classes of methods for modeling multivariate relationships. Cokriging constructs a multivariate model considering the correlations, while decorrelation algorithms remove the correlations prior to constructing the model and re-introduce the correlations after modeling.

2.5.1 Bootstrap algorithms

Bootstrap algorithms are robust techniques for calculating uncertainty in population statistics by resampling with replacement of sample values (Efron,

1982). At the core of bootstrap algorithms is the assumption that the relationship of samples to the population may be modeled by the relationship of resampled values to the samples. Random samples of the sample values are independently drawn with replacement and the statistic of interest calculated on these random drawings. Uncertainty is inferred using these values. The bootstrap is used when the distribution of the true samples is unknown, preventing inference of uncertainty directly. As the true distribution of values is typically unknown for grade, metallurgical and geotechnical properties, the bootstrap may be used to infer uncertainty in key statistics such as the histogram.

The spatial bootstrap (Pyrcz et al., 2006) is an extension of the bootstrap for calculation of the uncertainty in spatially correlated statistics. Rather than drawing from the sample values independently, correlated samples are drawn by simulation. For highly correlated samples, such as samples along a length of core, drawing independently would substantially underestimate the uncertainty in the statistic of interest. Consider the quantification of uncertainty in the mean of the experimental cumulative distribution function $F_k(z_k; v)$ for the random variable $Z_k(v)$ using LU simulation. The covariance matrix $\mathbf{C}_k(v)$ is the square matrix of spatial covariances $C_k(\mathbf{u}_\alpha - \mathbf{u}_\beta; v)$ where $\alpha, \beta = 1, \ldots, n_k$ and n_k is the total number of samples available in the domain of the random variable Z_k. The matrix is decomposed to symmetric lower triangular and upper triangular matrices using a Cholesky decomposition:

$$\mathbf{C}_k(v) = \mathbf{L}_k(v)\mathbf{U}_k(v)$$

L realizations of an independent standard normal random vector $\omega^l, l = 1, \ldots, L$ with length n_k are generated. These standard normal vectors may be correlated using the lower triangular matrix from the Cholesky decomposition:

$$\mathbf{y}_k^l(v) = \mathbf{L}_k(v)\omega^l, l = 1, \ldots, L$$

Using the standard normal distribution, G, to calculate probabilities associated with these correlated values, rank transformed realizations of the samples may be realized using the inverse of the experimental cumulative distribution function $F_k^{-1}(z_k; v)$:

$$\mathbf{z}_k^l(v) = F_k^{-1}(G(\mathbf{y}_k^l(v)); v)$$

Using the L vectors of resampled values $\mathbf{z}_k^l(v)$, the uncertainty in the mean can be calculated. Furthermore, uncertain histograms $F_k^l(z_k; v)$ may be inferred directly using the resampled values. In the models of uncertainty of metallurgical properties constructed in this thesis, the spatial bootstrap is used to quantify uncertainty directly.

2.5.2 Intrinsic supersecondary cokriging

There are many methods for jointly modeling the uncertainty in multiple variables including cokriging, colocated cokriging and the intrinsic framework. The intrinsic supersecondary approach (Babak and Deutsch, 2009b) is one approach for the joint modeling of a highly multivariate problem which enforces the bivariate correlations between variables in the final models. The intrinsic supersecondary model assumes that the spatial correlations among the secondary variables are proportional to the primary variable spatial correlation. Consider designating a normal score transformed random variable as the primary variable where $k = p$, $Y_p(v_c)$, and all variables are at a consistent scale v_c. All other variables are designated secondary variables, $Y_k(v_c)$, $k = 1, \ldots, K, k \neq p$. Using similar notation as Babak and Deutsch (2009b), the merged secondary (supersecondary, s) variable $Y_s(v_c)$ is calculated using multiple linear regression of the colocated secondary variables on the primary variable:

$$\begin{aligned} &\sum_{k=1}^{K,k\neq p} \tau_k \rho_{k,j}(v_c) = \rho_{k,k_p}(v_c), j = 1, \ldots, K, j \neq k_p \\ &Y_s(v) = \frac{\sum_{k=1}^{K,k\neq p} \tau_k Y_k(v_c)}{\rho_s(v_c)}, \rho_s(v_c) = \left[\sum_{k=1}^{K,k\neq p} \tau_k \rho_{k,k_p}(v_c) \right]^{\frac{1}{2}} \end{aligned} \tag{2.13}$$

The intrinsic supersecondary cokriging system uses the supersecondary variable with the intrinsic model to calculate weights for each variable. This system of equations is most easily understood using matrix notation. Let $\mathbf{R}$ be the square matrix of spatial correlations $\rho(\mathbf{u}_\alpha - \mathbf{u}_\beta; v_c)$ where $\alpha, \beta = 1, \ldots, n_\alpha$ and $\mathbf{r}$ be the vector of spatial correlations $\rho(\mathbf{u}_\alpha - \mathbf{u}_0; v_c)$ where $\alpha, \beta = 1, \ldots, n_\alpha$. The system of equations to be solved for the weight vectors is then:

$$\begin{bmatrix} \mathbf{R} & \rho_s\mathbf{R} & \rho_s\mathbf{r} \\ \rho_s\mathbf{R} & \mathbf{R} & \mathbf{r} \\ \rho_s\mathbf{r} & \mathbf{r} & 1 \end{bmatrix} \begin{bmatrix} \lambda_{Y_k} \\ \lambda_{Y_s} \\ \lambda_{Y_s,0} \end{bmatrix} = \begin{bmatrix} \mathbf{r} \\ \rho_s\mathbf{r} \\ \rho_s \end{bmatrix}, k = k_p \tag{2.14}$$

The conditional mean and variance are calculated using a linear estimator with the merged secondary variables:

$$\begin{aligned} \bar{y}_k(\mathbf{u}_0; v_c) &= \sum_{\alpha=1}^{n_\alpha} \lambda_{Y_k,\alpha} Y(\mathbf{u}_\alpha; v_c) + \lambda_{Y_s,0} Y_s(\mathbf{u}_0; v_c) + \sum_{\alpha=1}^{n_\alpha} \lambda_{Y_s,\alpha} Y(\mathbf{u}_\alpha; v_c) \\ \sigma_k^2(\mathbf{u}_0; v_c) &= 1 - \sum_{\alpha=1}^{n_\alpha} \lambda_{Y_k,\alpha} \rho(\mathbf{u}_0 - \mathbf{u}_\alpha; v_c) - \lambda_{Y_s,0}\rho_s + \sum_{\alpha=1}^{n_\alpha} \lambda_{Y_s,\alpha}\rho_s\rho(\mathbf{u}_0 - \mathbf{u}_\alpha; v_c) \end{aligned} \tag{2.15}$$

Assuming a positive definite correlation matrix and licit variogram model, the resulting intrinsic matrices will always be positive definite without modeling a linear model of coregionalization. This Markov assumption can be applied to any number of collocated samples and is appropriate for highly multivariate problems. Although multivariate relationships are controlled through the correlations after multiple univariate Gaussian transformation, there is no explicit control on the multivariate density with the supersecondary approach. The variables are assumed to be multivariate Gaussian after univariate transformation to Gaussian units. This is a common approach in multivariate geostatistical algorithms but unsuitable for very complex relationships (Barnett and Deutsch, 2015). In these cases, decorrelation algorithms may be used to independently model variables before re-introducing the complex relationships.

2.5.3 Decorrelation algorithms

Decorrelation methods including principal component analysis (PCA), minimum/maximum autocorrelation factors (MAF), the stepwise conditional transform (SCT) and the projection pursuit multivariate transform (PPMT) are a class of methods to multivariate transformation methods to model multivariate data using univariate techniques after decorrelation (Davis and Greenes, 1983; Vargas-Guzmán and Dimitrakopoulos, 2003). The rotated, linear combinations of the variables are uncorrelated for $h = 0$ for both PCA and MAF. For MAF the variables are also uncorrelated for $h = d$, where d is a separation

vector chosen for spatial decorrelation. The variables are then modeled independently and merged after modeling reversing the decorrelation procedure. Careful checking of the resulting spatial covariances is required to ensure that the original spatial covariances are preserved throughout this transformation.

Where there are numerous complex relationships, a multivariate transformation approach such as the PPMT (Barnett et al., 2014b) or SCT approach is advocated (Leuangthong and Deutsch, 2003). These multivariate transformation methods enforce multivariate normality in the transformed variables so that modeling under a multivariate normal assumption can be assumed and the modeled variables back-transformed with the multivariate transform to respect the original complex behaviour. A related class of transforms for removing complex behaviour for modeling are compositional transforms.

Compositional transforms, such as log ratios, can be applied to enforce additivity constraints that would not otherwise be enforced in the simulation workflow (Aitchison, 1982; Pawlowsky-Glahn and Olea, 2004). Compositional variables are modeled on the constrained hyper-plane and back transformed at the end of the modeling workflow to recover the original variables with the constraints applied. For full chemical compositional analyses this approach is particularly attractive. Size distributions present a unique challenge for modeling (Desbarats and Dimitrakopoulos, 2000). One approach for modeling size distributions is parameterization of the size distributions and modeling of the parameters. Other approaches such as the modeling of the distributions directly using a lookup table approach have also been developed. Both compositional transforms and the modeling of size distributions should be considered for geometallurgical workflows exhibiting these features. Algorithm selection is discussed further in chapter 5, where the focus is on workflows for geometallurgical modeling.

Many statistical and geostatistical algorithms have been introduced in this chapter. These algorithms form the basis for the nonlinear inference algorithms (chapter 3) and downscaling algorithms (chapter 4) developed in this thesis. These also form the basis for the proposed geometallurgical modeling work-

flows (chapter 5) and geometallurgical mineral deposit spatial characterization case study (chapter 6).

Chapter 3

Nonlinearity and mixing laws for metallurgical properties

Metallurgical properties encompass a wide range of properties relevant for mineral and metallurgical processing of metal ores. Examples include the Bond mill work index, froth flotation floatability and leaching rates. These properties often exhibit an intrinsically nonlinear averaging behaviour. Understanding and modeling this nonlinear behaviour is necessary for scaling and blending. Block models are constructed at a scale relevant for mining, but these models must be upscaled and blocks with different properties blended to understand how the material will move through the mill. The statistical inference and modeling of nonlinear metallurgical properties is the focus of this chapter.

Modeling the behaviour of mixtures of nonlinear metallurgical properties is of questionable importance if the mill is fed on a truck-by-truck basis, rather than from a blended stockpile. The importance of quantifying ore mixture properties in this case is evaluated using a small simulation study to demonstrate that even in a typical batch feeding case, a large amount of mixing occurs within the mill. After establishing the importance of evaluating mixture properties, a number of nonlinear and linear metallurgical properties are reviewed and potential reasons for the intrinsic nonlinearity in a number of metallurgical properties discussed.

To efficiently model nonlinear behaviour in metallurgical properties, a power transform re-expression framework is developed in the context of statistical re-

expression and previous research on the properties of mixtures. Alternative models for data-rich scenarios, and the numerical usage of the power transform are reviewed.

At the core of this chapter is the inference of nonlinear behaviour with varying amounts of data and varying degrees of experimental capacity. In the ideal scenario, the direct experimental evaluation of mixtures is carried out using ore blends in the lab. Inference in this case is straight forward using a regression model. Fitting the nonlinear behaviour in this case is demonstrated using experimental results from the literature. If there is no experimental capacity to directly evaluate the mixture, it was hypothesized that it may be possible to infer the nonlinear behaviour in a metallurgical variable using a correlated linear variable at the same scale. Using a Monte Carlo simulation study, this is demonstrated to be unreasonable and problems with the assumption of nonlinear behaviour on the basis of single-scale data highlighted.

The chapter concludes with nonlinear inference given multiscale measurements and multiscale correlated bivariate data. The multiscale scenario is common in a mining context; metallurgical samples are commonly measured on very large supports of up to 30 m, and assays taken on much smaller supports of 1 m or less. A semiparametric Bayesian updating algorithm for inferring nonlinear behaviour in this case is developed. Considerations for the application of the Bayesian updating algorithm including bandwidth selection for kernel density estimation, quantification of bivariate information, and regression attenuation are also documented.

3.1 Mixing in mineral processing equipment

Quantifying the nature of mixtures and the behaviour of geometallurgical properties in mixtures is central to this thesis. As such, the expected degree of mixing in mineral processing operations is briefly analyzed. Consider run-of-mine ore being fed into the mill. Two extreme scenarios for mill feed are sketched in Figure 3.1. In a batched mill feed, material is fed in sequentially with no

blending. The opposite case is a perfectly mixed ore feed after a blending operation.

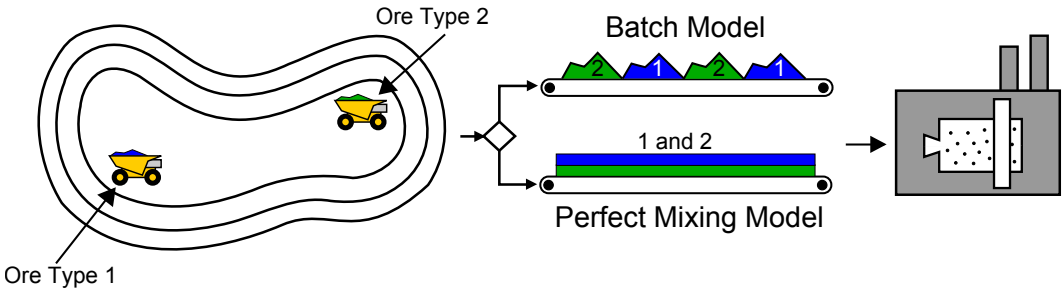

Figure 3.1: Sketch of two ore types being mined and fed into the mill in two different configurations: by batches sequentially or perfectly mixed after blending.

Consider the comminution of these ore feeds from the mine. In the case of a perfectly mixed ore feed, the properties of the ore mixture must be the relevant properties for the mill. When material is fed in batches, and in situations falling on the spectrum between these options, the properties of interest are unclear. If downstream operations in the mill act as plug-flow vessels with little to no mixing, properties of the individual ore types will dominate. As the ore is increasingly blended, properties of the mixture are increasingly relevant. With few exceptions, the first downstream processing step for ore is a comminution circuit, so expected blending behaviour in typical comminution circuits is considered here. The scale of mixing is important in these cases and is investigated with a small simulation study.

In a typical comminution process, ore is fed into a crusher, followed by a grinding circuit with post-classification (Napier-Munn and Wills, 2011). The grinding circuit will typically have at least one grinding operation with post-classification, and may have many more. The crushing operation will act as a plug-flow vessel where material fed into the vessel (the crusher here) leaves in the order it arrives (King, 2012). After crushing, consider a ball mill operating with post-classification (Figure 3.2). A set of typical parameters for a high throughput ball mill are considered (Table 3.1). Ball mills are among the most widely used grinding mills, so are used here. Alternatives including autogenous, semi-autogenous and rod mills are expected to behave comparably if operated in a recirculation mode with post-classification.

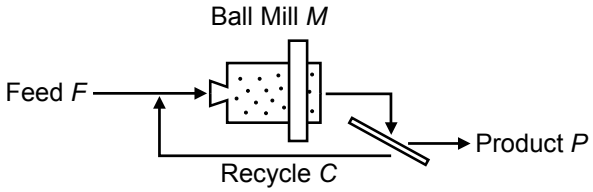

Figure 3.2: Sketch of ball mill configuration with post-classification. This flow diagram could represent either a grated ball mill with a discharge classifier, or an open ball mill with a post-classification.

Table 3.1: Ball mill parameters for a large, high throughput ball mill with post-classification.

Parameter	Value
Mill interior diameter	7 m
Mill length to diameter ratio	1.5
Charge volume fraction	0.45
Void volume fraction (of charge)	0.40
Slurry volume fraction (of void)	0.95
Slurry density	2.6 t/m^3
Pulp density	0.75
Steady state mass of solids in mill	135 t
Mill feed rate	1600 t/h
Recirculating load	400% (4800 t/h recycled)

Using these parameters and modeling the ball mill as a perfectly mixed vessel, a number of scenarios are considered. The goal of this modeling exercise is to evaluate the minimum amount of mixing that would reasonably be expected in a standard ball mill. This provides a measure of the importance of using the bulk mixed properties instead of the individual ore properties. The ball mill is treated as a single perfectly mixed vessel which is a reasonable model (King, 2012), although other models such as three-zoned mixing behaviour could be considered. Second-order effects including mill holdback are neglected in this analysis. Modeling the ball mill as a perfectly mixed vessel operating under steady state leads to the following mass balance equations, established with respect to the system as a whole and with respect to the ball mill:

$$\begin{aligned} \frac{\partial F}{\partial t} &= \frac{\partial P}{\partial t} \\ \frac{\partial M}{\partial t} &= \frac{\partial F}{\partial t} + \frac{\partial C}{\partial t} = 4\frac{\partial F}{\partial t} \end{aligned} \tag{3.1}$$

where the masses of the feed F, product P, recycle C and ball mill M are denoted. Time is denoted by t. This system was solved numerically using a finite difference method with 1 second time steps. Two scenarios corresponding to batch feeding of the mill are considered. In the first scenario, the mill feed is shocked with a second ore type. In the second scenario, a plug of material corresponding to one 400 tonne truck load is fed into the mill.

3.1.1 Effect of a mill feed shock

Consider a batch feeding operation where the mill feed is shocked and transitions entirely from one ore type to a second ore type. This corresponds to the case of sequential trucks loading a surge bin, or switching feed stockpiles. The mill feed for this scenario is shown in Figure 3.3. At t=30 minutes, the feed switches entirely from the blue ore type 1 to green ore type 2.

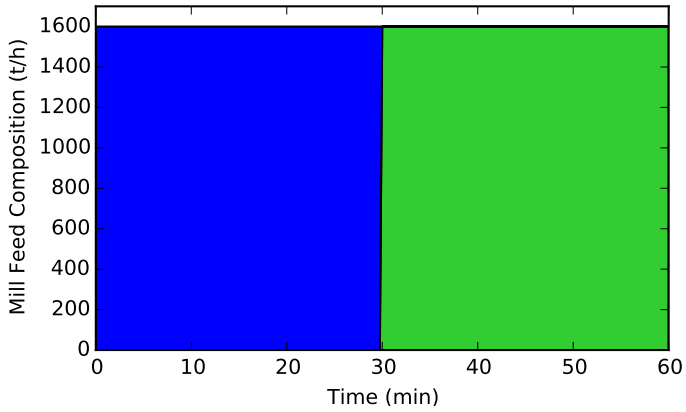

Figure 3.3: Composition of the mill feed where material is fed at a constant rate of 1600 tonnes per hour. At t=30 minutes, the feed switches from the blue ore type 1 to green ore type 2. This corresponds with a batch model of feeding the mill.

The composition of the product (Figure 3.4) changes exponentially. The amount of blue ore type 1 decreases exponentially and green ore type 2 increases exponentially after t=30 minutes. Due to the high recirculating load and large mass of ore contained in the ball mill, the product still contains 10% blue ore type 1 at t=45 minutes, 15 minutes after the shock. This indicates a substantial amount of mixing even under batch feeding conditions.

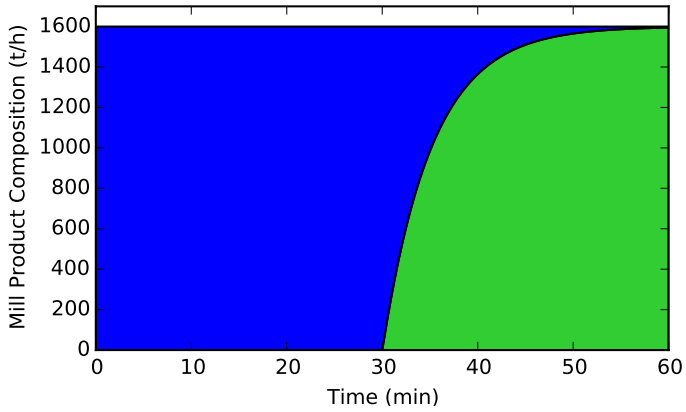

Figure 3.4: Composition of the product from the mill in Figure 3.4. Just past 30 minutes, the second ore type appears in the product assuming a well mixed mill. The amount of blue ore type 1 decreases exponentially with time.

3.1.2 Effect of a mill feed plug

Now consider a 400 tonne plug of ore fed into the mill corresponding to one truck load. The mill feed is shown in Figure 3.5. As the mill is processing 1600 tonnes per hour, the 400 tonne plug corresponds to a 15 minute mill feed period.

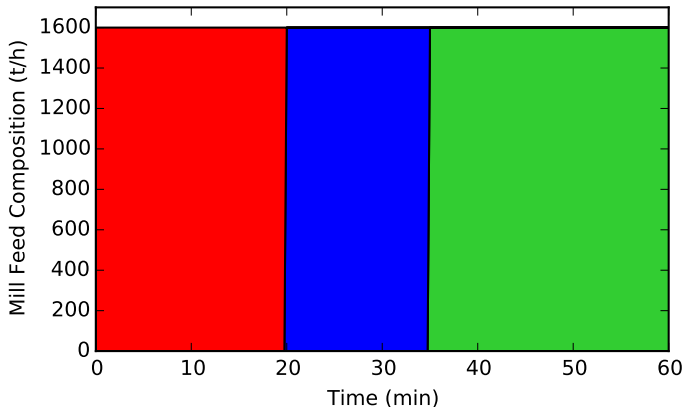

Figure 3.5: Composition of the mill feed where material is fed at a constant rate of 1600 tonnes per hour with multiple batches. At t=20 minutes, a plug of 400 tonnes of ore type 1 corresponding to one truck load are fed into the mill over 15 minutes before switching to ore type 2 at t=35 minutes. Red ore type 3 is fed into the mill first to visually separate the ore types in the product.

The composition of the mill product for this scenario, and the percent blue

(ore type 1) are shown in Figures 3.6 and 3.7. A high degree of mixing, with a peak concentration of 90% blue ore type, is observed for this scenario.

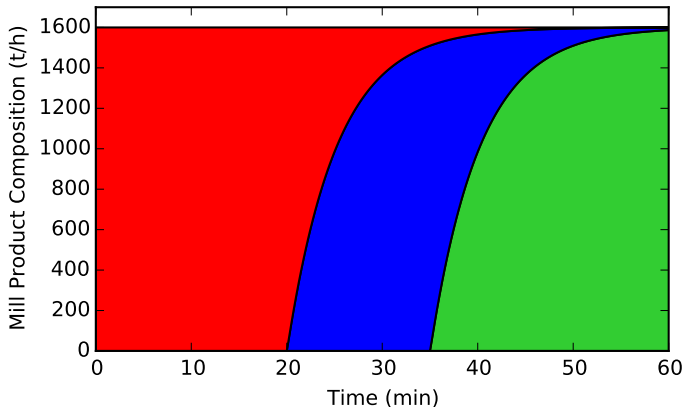

Figure 3.6: Composition of the mill product corresponding to the feed in Figure 3.5. Under the specified operating conditions, the mixing is substantive enough that the green ore type 2 (last material fed into the mill) and red ore type 3 (first material fed into the mill) are mixed.

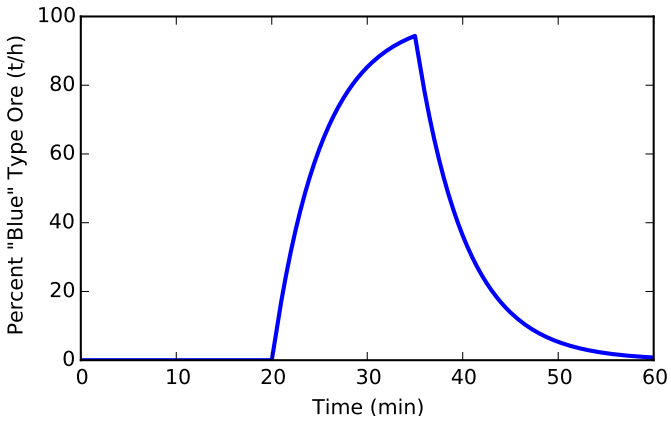

Figure 3.7: Percent blue (ore type 1, second ore type fed into the mill) within the this system as a function of time. The maximum percent of blue ore type contained within the product just exceeds 90% at t=35 minutes.

These scenarios indicate the high degree of mixing which occurs in early stage mineral processing operations. This high degree of mixing is supportive of common practice in metallurgical sampling to test a composite sample, rather than test individual ore types. Composite sample testing is useful to approximate the behaviour of blends in the mill with limited experimental work and sample collection. Additional comminution steps, such as a second milling

stage will further increase the degree of mixing among multiple ore types. The advantage of limiting experimental work and sample collection comes with a drawback of limited information on the nature of the mixing behaviour. Given high density sampling and testing, decisions about the degree of blending desired and optimal mining strategies to achieve the blending goal can be made.

The goal of optimizing mining strategies and the high degree of mixing expected emphasizes the importance of high resolution models of metallurgical properties. A high resolution spatial model of the metallurgical properties of the ore, combined with knowledge of the blending behaviour can be used to optimize and control the degree of blending. Consider the case where blending multiple ore types increases mill energy requirements (Yan and Eaton, 1994). As demonstrated by the small simulation study, the separation of ore types on a truck-by-truck basis in this case is insufficient to avoid the antagonistic blending behaviour. Further segregation would be required to avoid excessive blending in the mill. These decisions can only be made with both a high resolution spatial model and understanding of the behaviour of metallurgical properties of blends.

3.2 Nonlinearity and linearity in metallurgical variables

Only recently has substantial research in the area of understanding the mixing behaviour of mineral processing variables been published. This is perhaps unsurprising given the high degree of mixing within mineral processing operations. The Bond mill work index and froth floatability have been recently investigated and the experimental mixing behaviour documented (Conteras, 2013; Van Tonder et al., 2010; Yan and Eaton, 1994). There are no theoretical mixing and scaling laws available for these variables; instead an empirical relationship is developed based on the bench scale testing of mixtures of multiple ore types. The resulting empirical models are generally not linear mass or volume fraction weighted averages. The nonlinear behaviour of most metallurgical variables must be measured experimentally; complicated physi-

cal and chemical interactions make the establishment of a mixing relationship purely from theory difficult. Experimental results for the Bond mill work index and froth flotation grade (the grade of the flotation concentrate) are reviewed here. The strong nonlinearity of these experimental results motivates variable re-expression so that mixture behaviour through scaling and blending can be quantified.

The most widely used test for grindability in mineral processing mills is the Bond mill work index. The Bond mill work index (Bond, 1952) gives the work input required in kWh/t to grind a particle from a theoretical infinite size (practically a very large size) to a point where 80% of the product passes 100 microns. Many variations on the Bond equation have been proposed that improve the correlation between calculated energy requirements and observed plant work requirements (Morrell, 2008; Napier-Munn and Wills, 2011); however, the standard Bond mill work index remains widely used.

Yan and Eaton (1994) measured the Bond mill work index of blends of a more energy intensive ore (Mickey Doolan) and a less energy intensive ore (Vivians). They observed that the Bond mill work index of mixtures was not the volume average of the pure Bond work indices but was biased towards the more energy intensive ore (Figure 3.8). This behaviour was attributed to hold up of the more energy intensive ore which meant that the breakage more closely followed the behaviour of the Mickey Doolan ore.

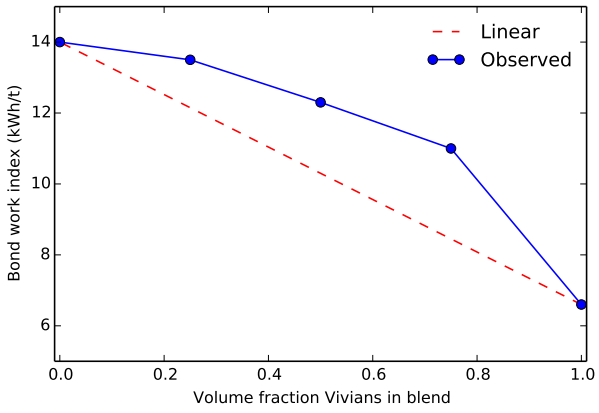

Figure 3.8: Bond mill work index for blends of two ores. Redrawn from Yan and Eaton (1994).

Froth flotation is a common strategy for the concentration of liberated minerals after comminution. The process is operated to maximize both metal recovery into the froth and the grade of the froth. Van Tonder et al. (2010) measured froth flotation recovery and grade using blends of four platinum ores. Three of these ores behaved similarly before blending while the Salene ore was a very poor floating ore. The effect of blending Salene ore with Townlands ore on grade (Figure 3.9) is reprinted from Van Tonder et al. The effect of blending on froth flotation grade is quite nonlinear and is fit well by a polynomial fit. Van Tonder et al. speculate that underlying physical/chemical interactions between the ores such as froth stabilization or destabilization as well as the combination of multiple grade-recovery curves are responsible for the nonlinearity in froth flotation parameters observed for blends. Froth destabilization can occur in the presence of extremely hydrophobic particles which may dewet and initiate bubble coalescence (Johansson and Pugh, 1992; Wang et al., 1999).

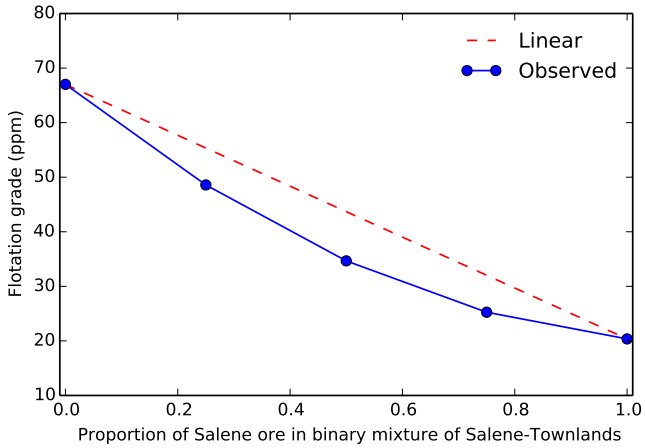

Figure 3.9: Flotation grade for mixtures of Salene and Townlands ores. Redrawn from Van Tonder et al. (2010).

Not all geotechnical and metallurgical variables are nonlinear and, importantly, a nonlinear relationship between variables does not imply that the variables themselves average nonlinearly (e.g., the relationship between a circle's area and radius is quadratic although both area and radius are linear variables). A widely used linear geotechnical/metallurgical property is the rock quality designation (RQD) which is a measure of the degree of fracturing in a

rock sample. Rock quality is described to be excellent (RQD=90-100%), good (75-90%), fair (50-75%), poor (25-50%) or very poor (0-25%) on the basis of measured RQD. RQD is defined to be (Deere and Deere, 1988; Palmstrom, 2005):

$$\text{RQD} = \frac{\sum \text{Length of core pieces} > 10 \text{ cm long}}{\text{Length of core sample}}$$

If RQD is being measured on very small core lengths then edge effects will play a large role in the resulting measurements. As such, the recommendation by Deere and Deere (1988) is to use the actual length of core runs when measuring RQD. Due to the small size of jointing (10 cm) used as the basis for RQD classification, these edge effects will be minimal for typical measurement lengths of 1.5 m. Under these conditions, RQD is a dimensionless linear measurement. Although RQD is linear, it is also a directional variable and as such care should be taken when modeling if drilling, and subsequent RQD measurements, are in multiple orientations.

3.3 Power transform re-expression framework

The nonlinearity of many metallurgical properties motivates a consistent method which accounts for the nonlinearity for spatial modeling. For modeling the nonlinear variables, a re-expression framework which will allow the use of linear averaging techniques is used. There are two primary goals associated with variable re-expression: data understanding and permitting the usage of available statistical techniques (Tukey, 1977). The requirements of additivity, homogeneity of error variance and symmetry of errors are all prevalent in traditional statistical and geostatistical analyses. In the context of scaling and blending metallurgical variables, the goal of variable re-expression is to linearize the variable of interest permitting the use of linear weighted averages.

3.3.1 Power transformation family

The power transformation family is a widely used and flexible re-expression framework. Consider re-expression of the random variable Z to Y. Tukey

(1957; 1977) defines a power transformation family with the general form:

$$y = (z + c)^{\omega}$$

where ω is the power of the transformation. The value c is a constant, often 0, which can be used to enforce $z + c \geq 0 \quad \forall z$ for variables that can be negative and are bounded. For the special case of $\omega = 0$, Tukey defines the transformation to be the logarithm leading to the expression:

$$y = \begin{cases} (z + c)^{\omega} & \omega \neq 0 \\ \log(z + c) & \omega = 0 \end{cases} \tag{3.2}$$

The usage of the logarithm for $\omega = 0$ is a choice in Tukey's power family and results in smoothly varying graphs with ω. The classic depiction of the power transforms from Tukey (1977) is recreated in Figure 3.10. For $p < 0$ the negative of the re-expression is commonly used so that the transform is rank preserving.

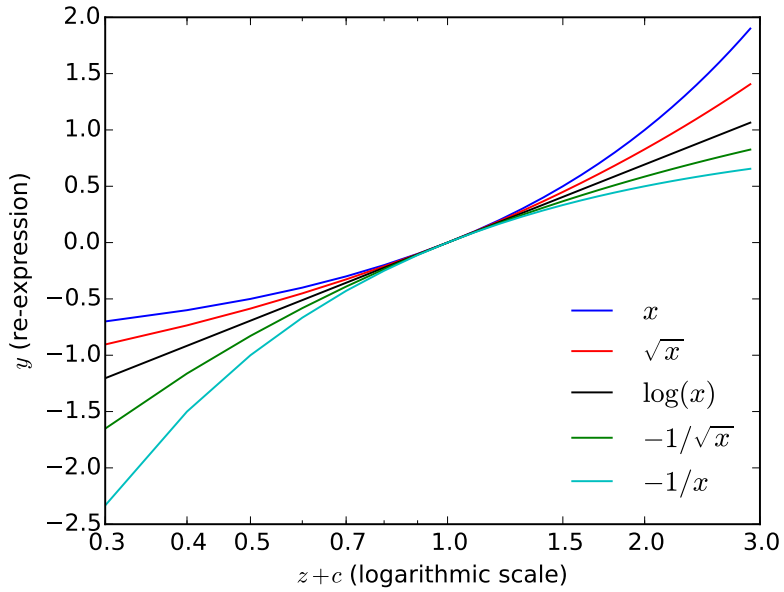

Figure 3.10: Plot of common power transforms recreated from Tukey (1977). Note that a linear transform of the labeled plots is actually plotted as in Tukey's original figure (eg: $-1 + x$ in place of x).

There have been a number of modifications made to the original Tukey transformation. The most notable change was by Box and Cox (1964) who modified the family slightly to account for the requirement of the choice of

the logarithm where $\omega = 0$, and preserve the rank ordering of the transformed variable for $\omega < 0$ within the transform itself. The Box-Cox transformation:

$$y = \begin{cases} \dfrac{(z+c)^{\omega} - 1}{\omega} & \omega \neq 0 \\ \log(z+c) & \omega = 0 \end{cases} \tag{3.3}$$

With the Box-Cox transformation, the case of $\omega = 0$ is the limiting case as $\omega \to 0$:

$$\lim_{\omega \to 0} \frac{(z+c)^{\omega} - 1}{\omega} = \lim_{\omega \to 0} (z+c)^{\omega} \log(z+c) = \log(z+c)$$

This property, and the rank preserving nature of the transform for $\omega < 0$ has led to the widespread usage of the Box-Cox transform. The equations of the lines plotted in Figure 3.10 are the Box-Cox transformations with a linear shift (additive constant for convenient visualization).

Other variations for negative variables including signed transformations have been proposed and are reviewed by Sakia (1992). Most metallurgical variables of interest are strictly positive or can be made positive using a constant c in the transformation expression so these techniques are not reviewed here.

3.3.2 Mixing frameworks

The power law transform can be related to mixing frameworks for a set of specific cases. Consider a binary mixture of two materials A and B with physical properties z_a and z_b. If the property mixes linearly and the materials are sufficiently well mixed, then the property of a mixture of the materials, z_{avg}, is a linear average weighted by the proportions, p_a and p_b:

$$z_{avg} = p_a z_a + p_b z_b = p_a z_a + (1 - p_a) z_b$$

Korvin (1982) considered binary mixtures of two materials, but with no requirement that the property average linearly. Instead of the requirement for linearity, if the property obeys eight physical constraints the binary mixture must follow an equation with the form:

$$z_{avg} = \begin{cases} [p_a z_a^{\omega} + p_b z_b^{\omega}]^{1/\omega} = [p_a z_a^{\omega} + (1 - p_a) z_b^{\omega}]^{1/\omega} & \omega \neq 0 \\ z_a^{p_a} z_b^{p_b} = z_a^{p_a} z_b^{(1-p_a)} & \omega = 0 \end{cases} \tag{3.4}$$

where ω is a finite power of the mixture rule. This equation is the linear average for $\omega = 1$, the harmonic average for $\omega = -1$ and the geometric average for $\omega = 0$.

Of the eight requirements for a mixture to follow the general mixture rule exactly, there are four of these properties which are generally applicable to mixtures of physical properties. These include reflexivity, idempotency, homogeneity with respect to the proportions and bi-symmetry. Reflexivity requires that a mixture of two materials with identical properties will have a mixed property equal to the identical materials. Idempotency requires that mixing a material with zero amount of a second material has no effect. Homogeneity with respect to the proportions requires that the properties of a mixture do not depend on the amount mixed, only on the ratio of the materials composing the mixture. Bi-symmetry requires that mixtures of mixtures are symmetric, for example in the ternary system with materials A, B and C, the properties of the mixture would be the same for $(A + B) + C$, $B + (A + C)$ and $A + (B + C)$. We consider these four properties to be generally applicable to the metallurgical variables of interest.

Other properties outlined by Korvin which do not necessarily apply to all mixtures of metallurgical properties include monoticity with respect to proportions, and monotonicity with respect to the properties; internity and homogeneity with respect the physical properties are also required for Korvin's law to apply. Monotonicity with respect to proportions and physical properties require a monotonic mixture model. Homogeneity with respect to the physical properties requires that the mixture law be linear with respect to the magnitude of the physical properties, that is, a mixture where $z_a = 10$ and $z_b = 20$ would have a z_{avg} exactly 10 times larger than a mixture where $z_a = 1$ and $z_b = 2$. Finally, internity requires that the properties of a mixture always lie between the properties of the pure component.

The framework proposed in this thesis is to re-express a nonlinear metallurgical variable $Z_k, k = p$ in terms of a linear variable, Y_p. This is done using a re-expression function:

$$Y_p = \phi(Z_p)$$

The linear variable Y_p can be averaged, scaled and estimated using standard linear weighted averages: $\bar{y} = \sum_i \lambda_i y_i$. The average nonlinear metallurgical variable is calculated by inverting the nonlinear transform function:

$$\bar{z} = \phi^{-1}(\bar{y})$$

This use of a power transform for the re-expression function ϕ is supported by the work of Korvin (1982), although other forms could be used.

Korvin (1982) has demonstrated that the general mixing rule (Eq. 3.4) is the correct nonlinear mixing rule to use for mixtures that follow a set of physical requirements. This mixing rule can be shown to be a weighted linear average of a set of variables re-expressed using Tukey's power transform (1957). Consider starting with Tukey's power transform equation (Eq. 3.2). A weighted linear average of a set of variables (letting $c = 0$) using Tukey's transform is:

$$w_{avg} = \begin{cases} \sum_{i=1}^{n} \lambda_i z_i^{\omega} & \omega \neq 0 \\ \sum_{i=1}^{n} \lambda_i \log(z_i) & \omega = 0 \end{cases} \tag{3.5}$$

where $\lambda_i, i = 1, \ldots, n$ are the weights. This equation can be written as:

$$\begin{cases} z_{avg}^{\omega} = \sum_{i=1}^{n} \lambda_i z_i^{\omega} & \omega \neq 0 \\ \log(z_{avg}) = \sum_{i=1}^{n} \lambda_i \log(z_i), \;\; z_{avg} = \prod_{i=1}^{n} z_i^{\lambda_i} & \omega = 0 \end{cases}$$

which is the general mixing rule (Eq. 3.4) when $n = 2$. This supports the usage of the power transform or Box-Cox transform for the physically based metallurgical variables. Equation 3.5 can be used as a general weighted average equation. The weights used, generally volume or mass proportions, should be consistent with how the mixing law was established. This average is equivalent to the generalized weighted means for the evaluation of mechanical properties of composites (Ji et al., 2004) and the power averages described by Jensen (1998).

3.3.3 Alternative models for mixtures

For specific classes of mixtures, the mixing framework of Korvin (1982) can be used to fit a model to the mixture property. Alternatively, a polynomial

fit may be used to generally fit properties which do not meet the power law requirements, and for properties with experimental data to support an alternative fit. In these cases, a simple polynomial fit may be used (Cornell, 2011; Scheffé, 1958). The canonical polynomials proposed by Scheffé are flexible polynomials over a simplex. For a mixture of n components with proportions $p_i, i = 1, \ldots, n$, the first and second degree polynomials are:

$$\begin{cases} \bar{z} = \sum_{i=1}^{n} \beta_i p_i & \text{degree} = 1 \\ \bar{z} = \sum_{i=1}^{n} \beta_i p_i + \sum_{1 \le i < j}^{n} \beta_{ij} p_i p_j & \text{degree} = 2 \end{cases}$$

Many alternative models have been formulated including the polynomial forms of Becker (1968) and ratio forms described by Snee (1973). These polynomial forms present an alternative to the power models which are focused on in this thesis, and are useful when the property is not monotonic. The power transform models discussed here are useful for their simplicity and, for binary mixtures, unique inversion to a linear variable.

3.3.4 Numerical application of the variable re-expression

Using the developed re-expression functions, variables can be scaled and blended. Both of these applications involve similar mathematical concepts, so blending is considered here with a synthetic case to demonstrate the application of the re-expression framework.

Consider the calculation of a blended Bond mill work index given two equal sized rock volumes ($p_1 = p_2 = 0.5$) with Bond mill work indices of $z_p(\mathbf{u}_1; v) = 11.0$ and $z_p(\mathbf{u}_2; v) = 13.5$ kWh/t. The ores are from the same region as the tested Vivians and Mickey Doolan ores. Using the re-expression framework the Bond mill work index of the blend can be calculated (using a fit power of 5.558, discussed further when fitting models):

$$\begin{aligned} \bar{z}_p(\mathbf{u}_{12}; 2v) &= \left[p_1 z_p(\mathbf{u}_1; v)^{5.558} + p_2 z_p(\mathbf{u}_2; v)^{5.558}\right]^{1/5.558} \\ &= \left[0.5 \cdot 11.0^{5.558} + 0.5 \cdot 13.5^{5.558}\right]^{1/5.558} \\ &= 12.5 \end{aligned}$$

This small example demonstrates the application of the mixing law for simple binary relationships. For the quaternary froth flotation grade relationship, the composition of each sample cannot be determined uniquely without additional information. In this case, knowing what geologic area the sample came from could reduce the space of possible compositions. Solving for the composition exactly will not be possible in all cases; in this case a stochastic mixing framework in which the composition is drawn randomly from possible compositions could be used.

3.4 Direct experimental evaluation of mixtures

Direct evaluation of mixture behaviour by experimental testing of blends is the ideal scenario for inferring mixture behaviour; ore with measured properties is blended and the properties of the blend measured. These blends are of known composition and, if desired, even the degree of mixing in the blended material can be controlled. These cases correspond to the previously shown experimental results of Yan and Eaton (1994) and Van Tonder et al. (2010).

3.4.1 Experimental design for mixture evaluation

There are a number of methods for the design of experiments with mixtures and modeling the results; the interested reader is referred to Cornell (2011). Many of the experimental design frameworks for mixtures require that the response is a function only of the proportion of the components mixed, and not on the amounts. This is equivalent to Korvin's homogeneity with respect to proportions requirement. Classical experimental methods with mixtures also require knowledge of the pure components of the mixture. In the case of metallurgical variables, pure rock types would be required. If the individual components are known, a standard simplex lattice design can be used.

Alternatively, blends of known composition and known properties can be evaluated directly and the mixed response measured. Regression can then be applied to directly fit the nonlinear behaviour. Both of these techniques for the experimental evaluation of mixtures are applied in this thesis.

3.4.2 Direct modeling of experimentally evaluated mixtures

The application of the metallurgical re-expression framework requires fitting the re-expression function $\phi(Z_p)$. This is demonstrated with both the Bond mill work index and froth flotation grade examples introduced earlier. The Bond mill work index tests of Yan and Eaton (1994) were reviewed earlier (Figure 3.8). In these tests, hard Vivians ore was blended with Mickey Doolan ore in various proportions and the Bond mill work index of the blend measured. This data was fit using the power transform framework with nonlinear regression; the equation for this fit, where p is the volume fraction of Vivians ore, is:

$$\text{Bond mill work index of blend} = \left(p6.6^{5.558} + (1-p)14^{5.558}\right)^{1/5.558}$$

Since there are only two components (Vivians and Mickey Doolan), the blended value is only a function of the Bond mill work indices of the two pure components and the power. The power in the fit (5.558) is high indicating that mixtures will favour the higher value. An exponent this high is more likely when the variance of the values is small, as is the case for the Bond mill work index values. The resulting fit is plotted in Figure 3.11 and a cross plot of the blended response using the power fit and experimentally measured Bond mill work index is shown in Figure 3.12.

The froth flotation study of Van Tonder et al. (2010) measuring froth grade was also fit using the power transform model. The four platinum ores used in the study: Salene (x_1), Waterval UG2 (x_2), Paardekraal (x_3) and Townlands (x_4) were blended using a simplex lattice mixture design. Experimental data from the experiments were not directly available, so the accurate ($R^2_{adj} = 0.94$) polynomial fit by Van Tonder et al. was used to back out the experimental froth flotation grade z_{avg}:

$$\begin{aligned} z_{avg} = {} & 20.36x_1 + 53.19x_2 + 71.81x_3 + 67.01x_4 - 31.92x_1x_2 - \\ & 35.59x_1x_3 - 36.1x_1x_4 + 284.6x_1x_3x_4 + 234.76x_2x_3x_4 \end{aligned}$$

where monomial coefficients (20.36, 53.19,...), are the pure component froth flotation grades. These pure component froth flotation grades were used with

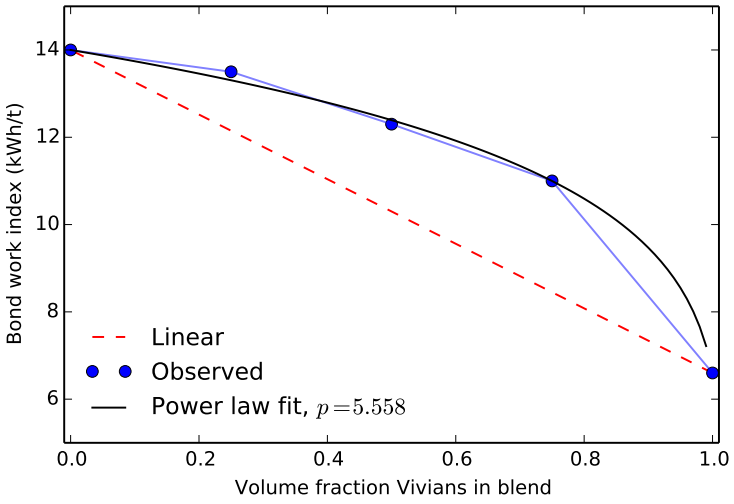

Figure 3.11: Plot of Bond mill work index for blends using the power law fit to data from Yan and Eaton (1994)).

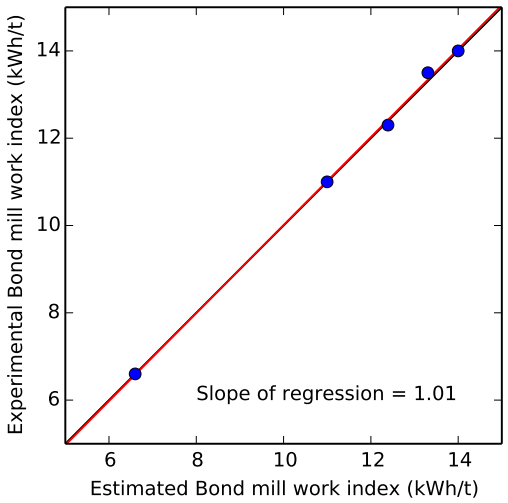

Figure 3.12: Cross plot of the estimated Bond mill work index for blends versus the experimentally measured Bond mill work index from Yan and Eaton (1994)).

the same 35 simplex design samples from Van Tonder et al. to fit the power law. The power fit (0.125) is low indicating that grades of mixtures will favour the lower value. A binary plot of the froth flotation grade for the low grade Salene and high grade Townlands ores is shown in Figure 3.13.

The visible nonlinearity in the flotation grade for blends of the very dis-

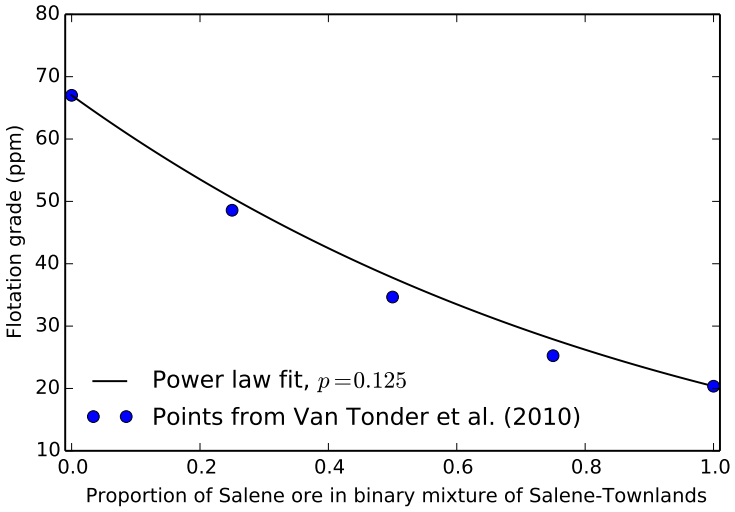

Figure 3.13: Plot of froth flotation grade for blends of the dissimilar Salene and Townlands ore using the power law fit to data from Van Tonder et al. (2010)).

similar Salene and Townlands ore is seen in blends of similar materials such as the Paardekraal and Townlands ores (Figure 3.14).

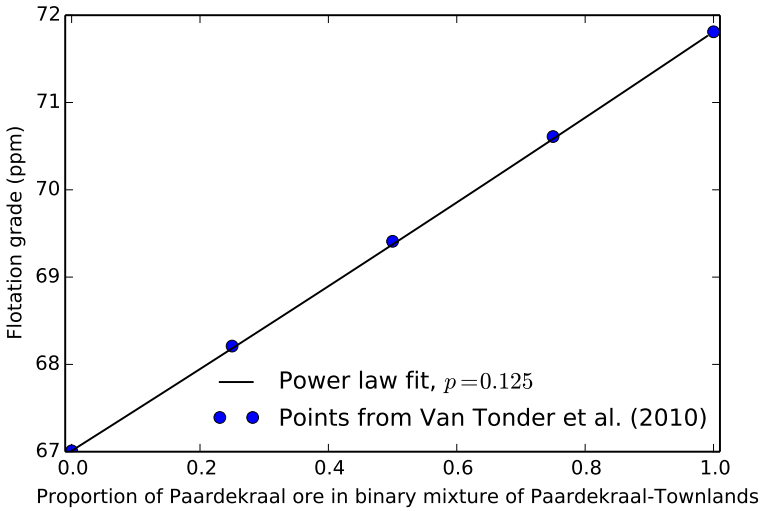

Figure 3.14: Plot of froth flotation grade for blends of the similar Paardekraal and Townlands ores using the power law fit to data from Van Tonder et al. (2010)).

The power law fits the data well for binary mixtures, but is bounded by the pure component grades, so does not fit data where ternary interactions resulted

in higher flotation grades than the individual components. This is reflected in a cross plot of the experimental and estimated blend values (Figure 3.15).

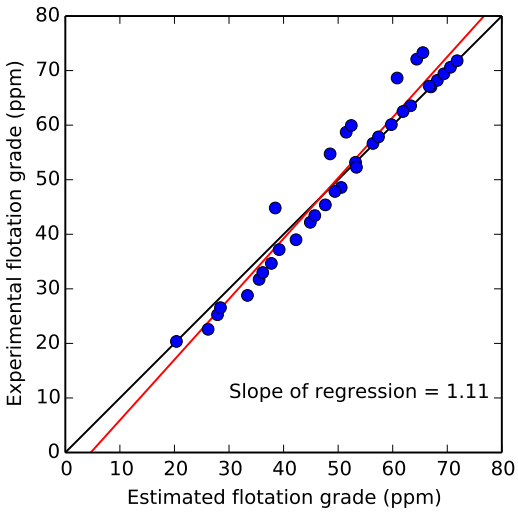

Figure 3.15: Cross plot of the estimated froth flotation grade for blends versus the experimentally measured froth flotation grade from Van Tonder et al. (2010)).

Two cases involving experimental results on the blending of ores have been considered. For the binary mixture of ores tested by Yan and Eaton, a power law re-expression accurately models the blending behaviour of the Bond mill work index. The quaternary mixture investigated by Van Tonder et al. exhibits significant ternary interactions between ores causing deviations from the requirements for a well-fitting power law as described by Korvin. In this case, an alternative re-expression such as a Scheffé polynomial is reasonable. Both of these cases are highly desirable from a re-expression perspective; blends of known composition have been tested and the nonlinear behaviour may be modeled directly. When only limited sample and experimental capacity are available, direct observation and modeling of metallurgical variables may not be possible or economically feasible.

3.5 Indirect nonlinear inference in the absence of direct experimental evaluation

If the direct experimental evaluation of mixtures is not possible due to limited sample for testing, experimental cost or constraints on sample conditions, then in specific cases the behaviour of the mixture may be evaluated indirectly. Three cases are considered for indirect inference in this thesis: inference in the presence of a 1) multiscale measurements of the metallurgical variable, 2) single-scale correlated bivariate measurements and 3) multiscale correlated bivariate measurements. When a bivariate distribution is leveraged, the distribution may be known theoretically, or observed experimentally, such as a relationship between iron content and grindability. The central idea of the bivariate indirect inference methods evaluated is that using a relationship between a nonlinear variable and linear variable it may be possible to infer the nature of the nonlinear variable through the bivariate distribution between the two.

In the case where measurements are only available on a single scale, nonlinear inference is demonstrated to be impractical even with a very large amount of data. This is demonstrated using a Monte Carlo simulation experiment which compares the bivariate distribution of the linear and nonlinear properties with a synthetic bivariate drawn from two linear properties.

3.5.1 Single-scale correlated bivariate inference

Consider the case of two variables, a nonlinear metallurgical property Z_m such as the Bond mill work index and an additive property Z_g such as a metal grade. These variables are measured on a single scale v and correlated at this scale. The proposed approach for indirect mixing law determination is to infer the form of the mixing law using the conditional distribution of $\{Z_m(v)|Z_g(v) = z_g\}$. In the absence of experimental data, it was hypothesized that it is possible to leverage correlated linear variables using conditional probabilities to infer the mixing behaviour between the linear and nonlinear variable.

The proposed approach for nonlinear inference given a single-scale corre-

lated bivariate distribution is to compare the conditional distribution $\{Z_m(v)|Z_g(v) = z_g\}$ with a conditional distribution drawn from a bivariate with the same correlation and univariate distributions drawn from linear samples which are rank transformed to have the same univariate distributions. The hypothesis is that a difference in the conditional percentiles and conditional expectations between these distributions will be observable in this situation. This would then be used to guide the inference of nonlinearity.

Consider calculating the correlation $\rho(Z_m(v), Z_g(v))$ and targetting this correlation for a univariate transformed bivariate normal distribution. This process is completed iteratively by drawing samples from a bivariate normal (BVN) distribution with correlation ρ(BVN) and rank transforming to have univariate distributions $F(Z_m(v))$ and $F(Z_g(v))$. This rank transform is analogous to the familiar normal score transformation. The correlation is chosen such that the univariate transformed distribution has the same correlation as the experimentally determined bivariate $Z_m(v), Z_g(v)$ distribution. Conditional expectations and percentiles are calculated for this bivariate distribution. These univariate transformed bivariate normal variables correspond to linear variables with the same univariate distribution and correlation as the nonlinear-linear relationship, but not the exact same bivariate distribution. A flowchart of this approach is shown in Figure 3.16.

Following the procedure outlined in Figure 3.16, spatially correlated random variables $Z_m(v_s)$ and $Z_g(v_s)$ were simulated using colocated sequential Gaussian simulation on a line. 50,000 samples spaced 1 m apart (i.e. $v_s = 1$ m) were simulated with a spherical variogram (range = 20 m) and correlation of 0.6. The linear grade variable $Z_g(v_s)$ was transformed to a moderately skewed lognormal distribution and the metallurgical variable $Z_m(v_s)$ was shifted and scaled to a non-standard normal distribution. The variables were upscaled to scale $v = 10$ m. The power law re-expression fit to data from Yan and Eaton (1994) was used for the metallurgical variable Z_m. Linear averaging was used for the grade variable Z_g. A section of the simulated data is shown in Figure 3.17. Histograms and scatterplots are shown in Figure 3.18.

The bivariate normal correlation was iteratively determined by testing a

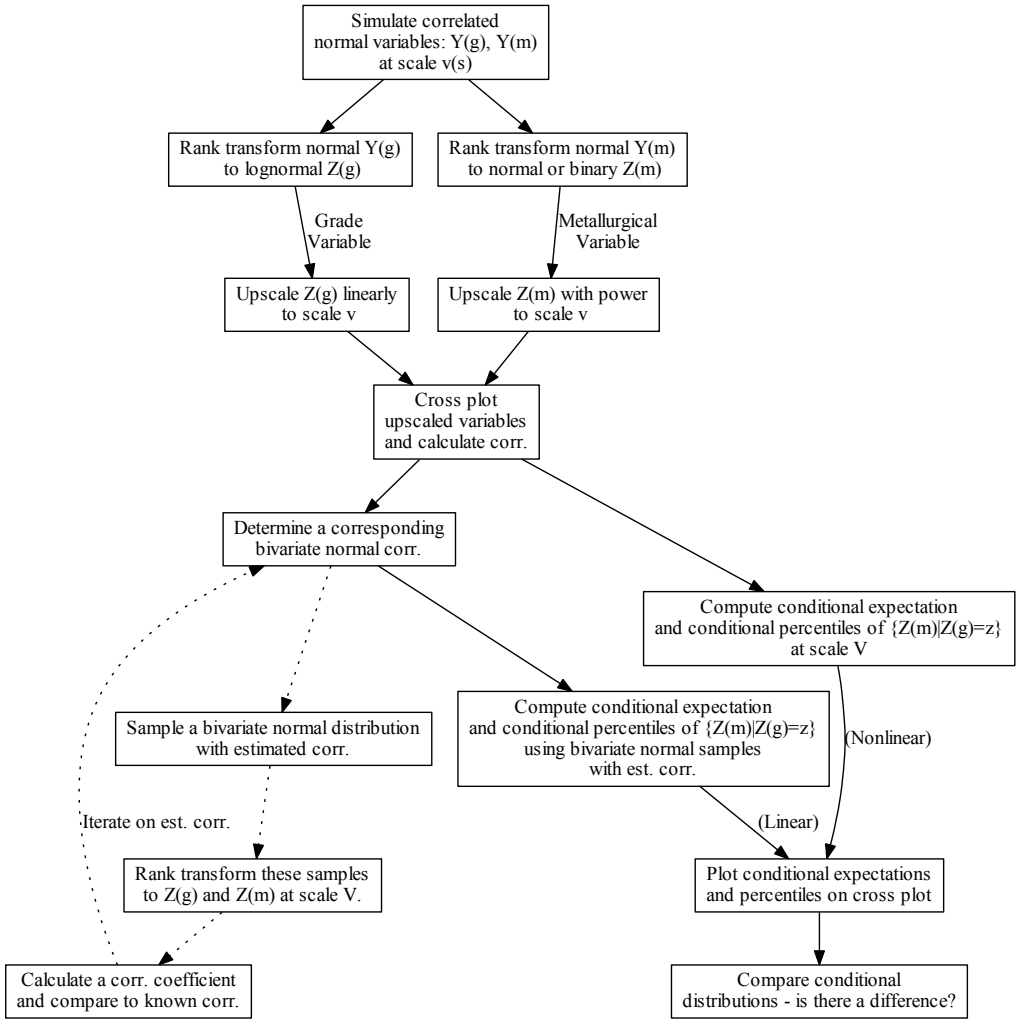

Figure 3.16: Flowchart of the Monte Carlo simulation study which was completed to examine if it is possible to determine the nature of nonlinear averaging using the single-scale bivariate distribution between a nonlinear and linear variable.

range of bivariate normal correlations and univariate transforming the simulated values to $Y(v)$ and $Z(v)$ distributions. A plot of the univariate transformed correlation against the bivariate normal correlation used is shown in Figure 3.19

For the desired correlation of 0.648 in Y_V, Z_V units, the bivariate standard normal correlation of 0.758 was determined. Using this correlation, the expected value of the univariate transforms and 10% and 90% percentiles were calculated for the conditional distribution $\{Z_V|Y_V = y\}$. These were also

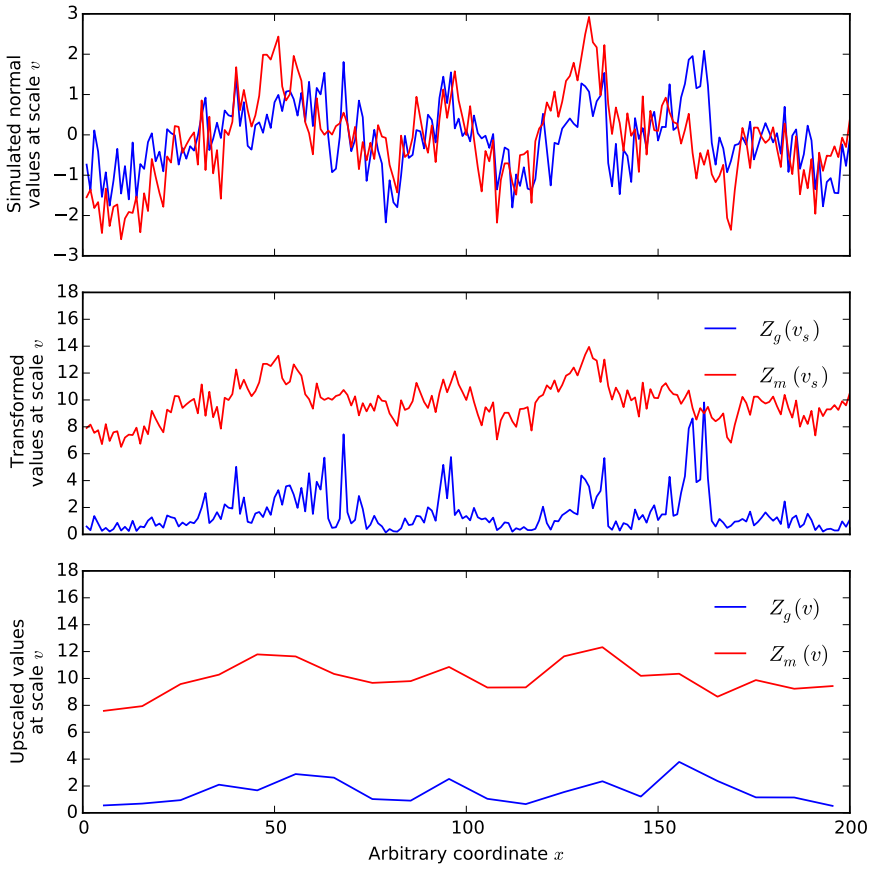

Figure 3.17: Simulated and upscaled Y and Z variables at scales v_s and v. The $Z_m(v)$ values are not at the linear average of the simulated $Z_m(v_s)$ values, but are higher due to the power law average favouring high values.

calculated for the original upscaled distributions. The results are plotted in Figure 3.20.

From this plot, no substantial difference can be see between the transformed bivariate standard normal values and the nonlinearly upscaled values. The same experiment was repeated for different power values (instead of the calculated 5.558 fit to the data from Yan and Eaton). The results when using a power of 0.5 instead of 5.558 are shown in Figure 3.21. As before, there is no discernible difference in the conditional distributions.

The more extreme case of a binary distributed small scale metallurgical variable was also tested in a Monte Carlo simulation study. The same procedure outlined in Figure 3.16 was used. A very small random component was added to the simulated binary variables so that the normal score (rank) trans-

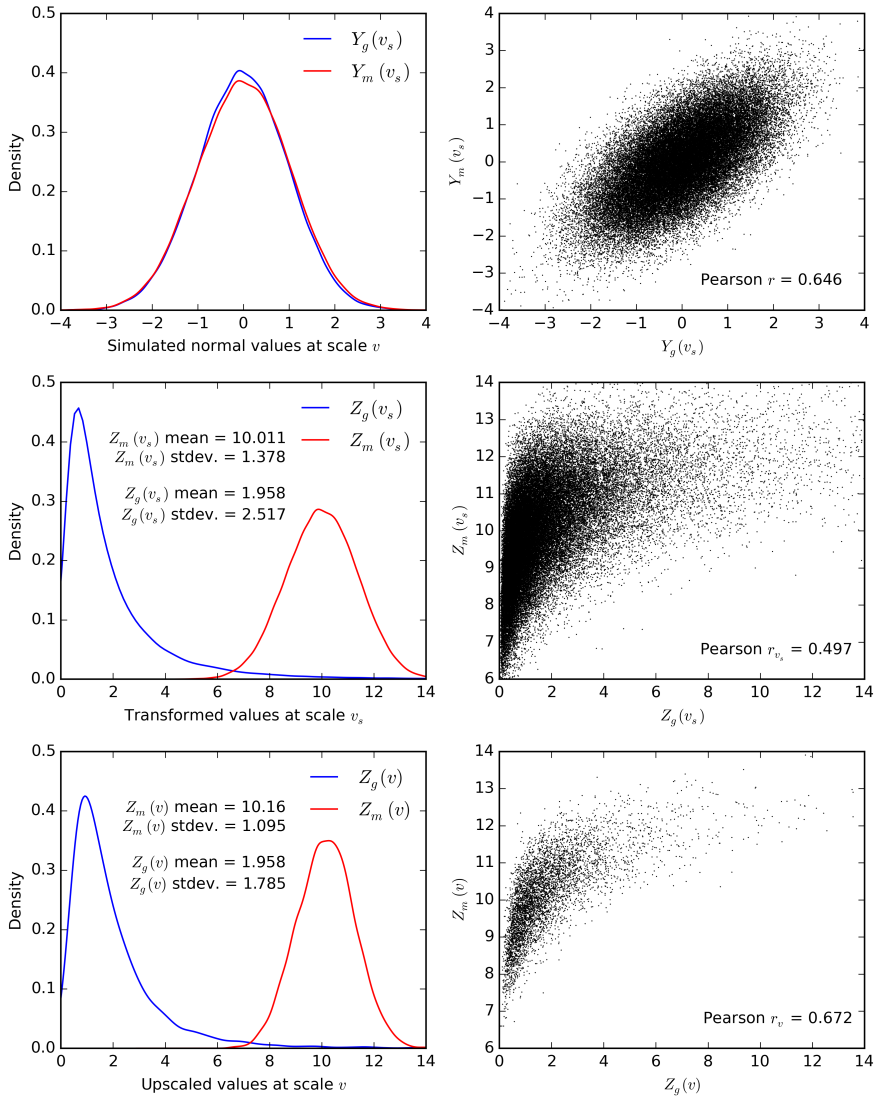

Figure 3.18: Histograms and scatterplots of the variables at scale v and upscaled variables at scale V.

formation would be unique. A number of the simulated variables are shown in Figure 3.22. Histograms and cross plots are shown in Figure 3.23.

The correlation was iteratively fit using the same procedure as before, and the conditional distributions calculated and plotted (Figure 3.24). There are differences between the linear and experimental conditional distributions visible in Figure 3.24, so linear averaging of the small scale metallurgical variable (i.e., $p = 1.0$) was also tested to determine if this was an artefact of simulation, or indicative of a substantial difference. The results are plotted in Figure 3.25.

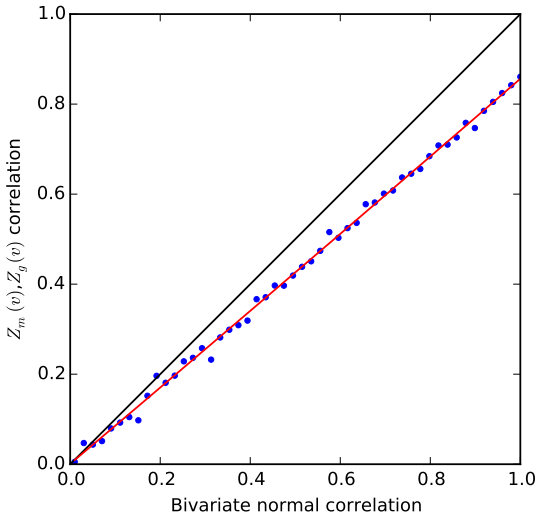

Figure 3.19: Correlation of univariate transforms of a standard bivariate normally distributed variable with specified correlation. The correlation in Y_V, Z_V units is lower, which follows from the heteroscedasticity and nonlinearity visible observed in the bivariate for these variables.

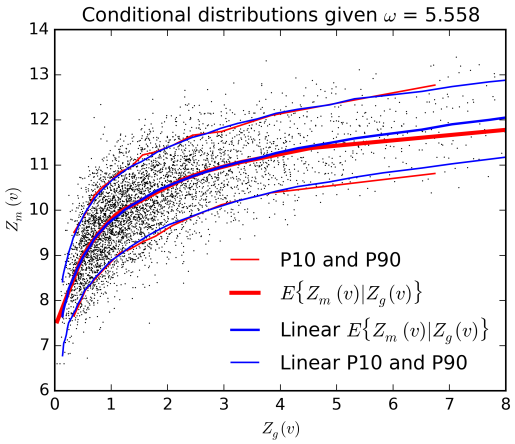

Figure 3.20: Percentiles and conditional expectation using the univariate transformed bivariate standard normal values (blue, linear) and upscaled values (red).

Based on the conditional distributions for linear averaging, no discernible different between the linear and nonlinear conditional distributions is visible in Figure 3.24.

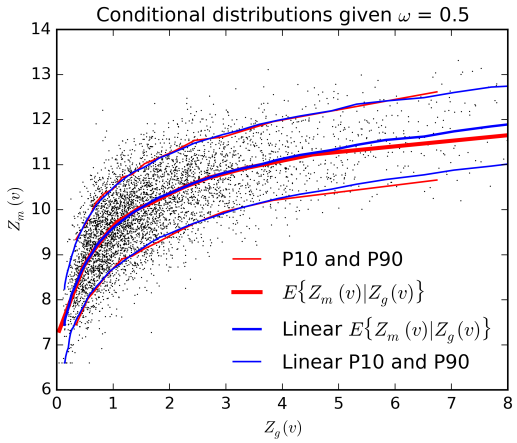

Figure 3.21: Percentiles and conditional expectation using the univariate transformed bivariate standard normal values (blue, linear) and upscaled values (red) where $p = 0.5$.

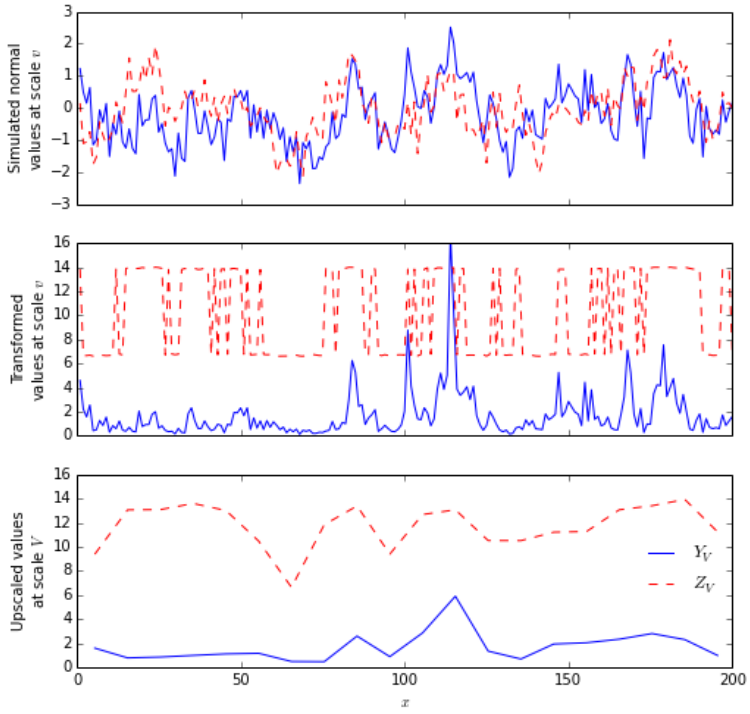

Figure 3.22: Simulated and upscaled Y and Z variables at scales v and V. The Z_V values fall on distinct lines due to the nature of upscaling; as 10 small scale Z_v values were combined to form the larger V scale the upscaled values must equal one of 11 values.

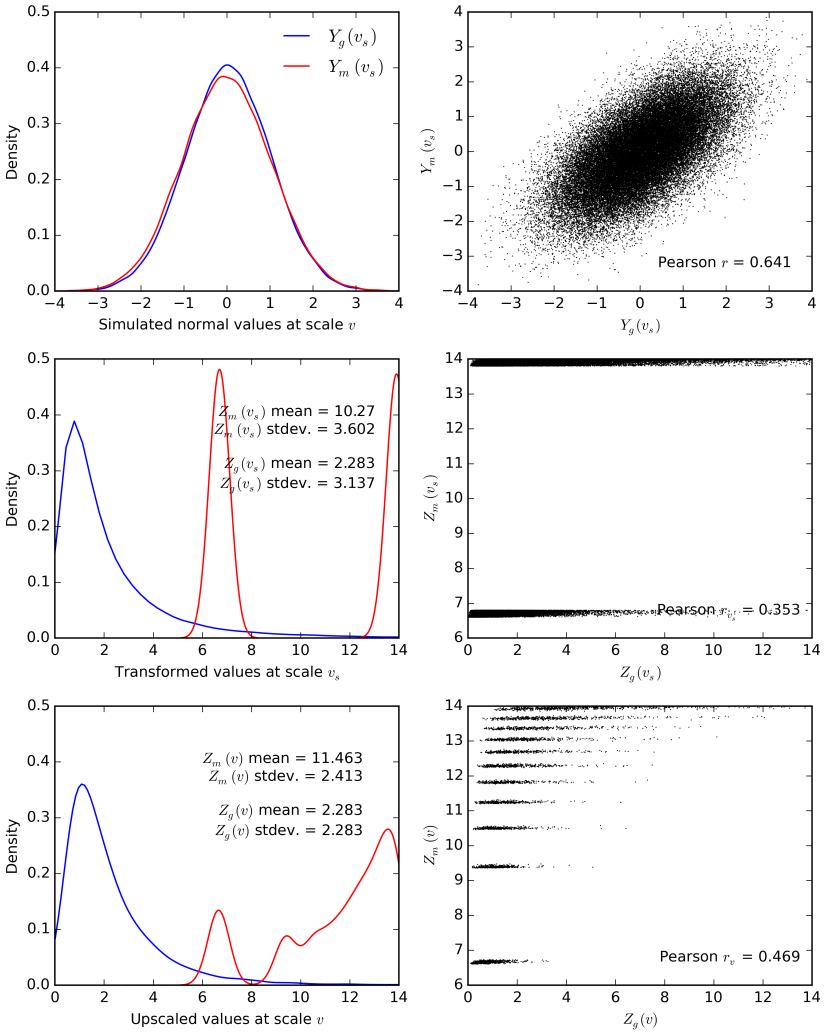

Figure 3.23: Histograms and scatterplots of the variables at scale v and upscaled variables at scale V with a binary Z.

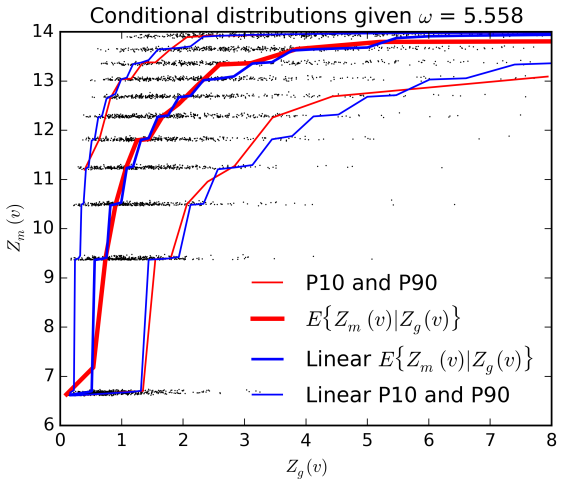

Figure 3.24: Percentiles and conditional expectation using the univariate transformed bivariate standard normal values (blue, linear) and upscaled values (red) for a binary distributed Z.

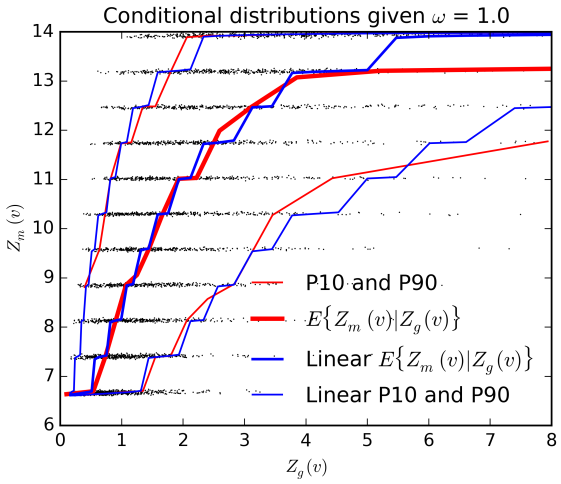

Figure 3.25: Percentiles and conditional expectation using the univariate transformed bivariate standard normal values (blue, linear) and upscaled values (red) for a binary distributed Z which was linearly upscaled.

3.5.2 Problems with the assumption of nonlinear behaviour

The challenges in assuming that nonlinear behaviour can be inferred solely on the basis of a single-scale bivariate distribution are more evident when the bivariate Gaussian distribution is considered as an example. Consider contouring the bivariate normal distribution. The probability density function f of the bivariate standard normal distribution with random variables X and Y (each variable has a mean of 0 and standard deviation of 1) with correlation ρ:

$$f(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right)$$

The marginal distributions of the bivariate standard normal distribution are exactly standard normal. Constant density contours are commonly used to show the shape of the bivariate distributions in a 2D plot. These contours can also be used to check the normality of a purported normal distribution (Deutsch and Deutsch, 2011). For a correlation coefficient of $\rho = 0.6$, the constant density contours for 25%, 50% and 90% are shown along with the marginal standard normal distributions in Figure 3.26.

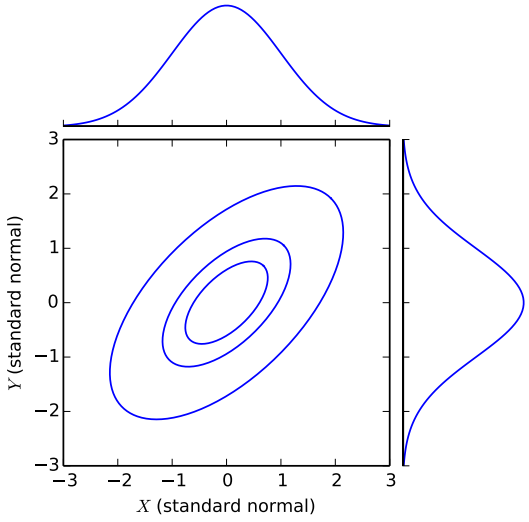

Figure 3.26: Constant density contours (25%, 50% and 90%) of the bivariate standard normal distribution. Histograms of the marginal standard univariate normal distributions are also shown.

The shape of conditional distributions $f(Y|X = x)$ can be illustrated using percentiles (here the 10%, 25%, 50% (median), 75% and 90% percentiles) and the conditional mean $E\{Y|X = x\}$ (Figure 3.27). As the multivariate normal distribution is linear and homoscedastic, these conditional statistics are linear. The bivariate standard normal distribution is considered, so the slope of these lines is equal to the correlation coefficient. Note that the median (50% percentile) is equal to the mean as the distributions are symmetric.

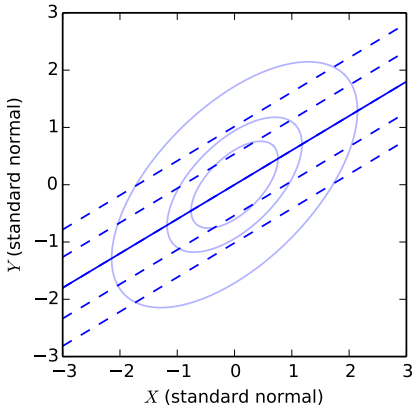

Figure 3.27: Percentiles of the conditional distributions (dashed lines, 10%, 25%, 50%, 75% and 90%) and the conditional mean $E\{Y|X = x\}$ (solid line) of the bivariate standard normal distribution. Constant density contours are shown in the background.

The properties of linearity and homoscedasticity of the multivariate normal distribution are illustrated in these figures, and are widely leveraged in geostatistics. Univariate transformations of these variables are now considered to demonstrate that although the underlying distribution may be multivariate normal, multivariate distributions of simply transformed variables may be nonlinear and heteroscedastic. This does not imply nonadditivity of the variables, and is purely a function of the univariate transformations applied.

A common bivariate relationship examined is the relationship between a grade variable, chosen to be the X variable here, and a metallurgical variable (Y for this example). Grade variables often appear lognormal while many metallurgical variables, such as the Bond mill work index, may appear more

symmetric. X was transformed to be lognormal as:

$$X_{\ln N} = \exp(\alpha + \beta X)$$

where X was originally a standard normal variable. The probability distribution function of $X_{\ln N}$ is then:

$$f(x) = \frac{1}{x\beta\sqrt{2\pi}} \exp\left(-\frac{(\ln(x) - \alpha)^2}{2\beta^2}\right)$$

For this example, $\alpha = 0.3$ and $\beta = 1.0$ were used. The random variable Y was transformed to be normally distributed with mean $\mu = 10$ and standard deviation $\sigma = 1.35$. The probability density function of the normal distribution parameterized by μ and σ, $N(\mu, \sigma)$ is:

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y - \mu)^2}{2\sigma^2}\right)$$

Contours of the bivariate distribution after these univariate transformations are shown in Figure 3.28. The conditional distribution statistics of $\{Y|X = x\}$ (Figure 3.29) show significant nonlinearity, although are homoscedastic as Y is normally distributed. This nonlinearity of the conditional statistics is not in itself sufficient evidence of nonlinear averaging behaviour, but instead may solely reflect differences in the univariate distributions.

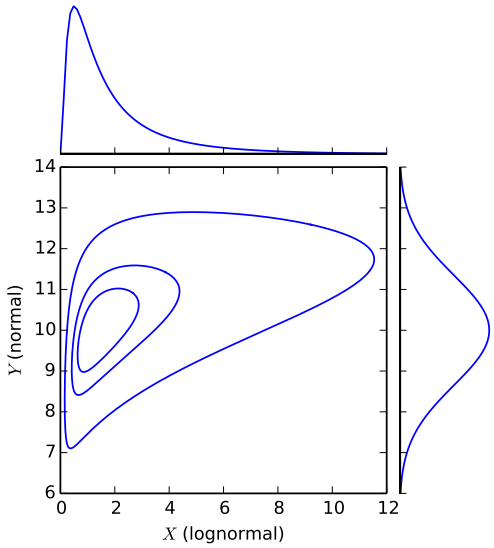

Figure 3.28: Constant density contours (25%, 50% and 90%) of a bivariate lognormal-normal distribution obtained by univariate transformations of a bivariate standard normal distribution. Histograms of the marginal distributions are also shown.

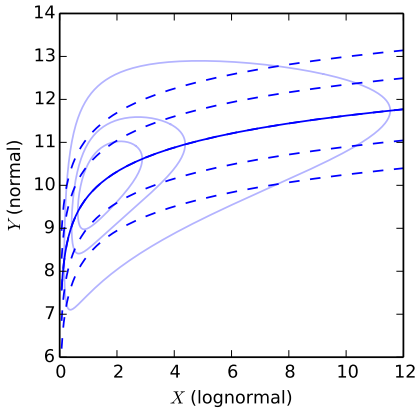

Figure 3.29: Percentiles of the conditional distributions (dashed lines, 10%, 25%, 50%, 75% and 90%) and the conditional mean $E\{Y|X = x\}$ (solid line) of the lognormal-normal distribution. Contours of the bivariate distribution are shown in the background.

3.6 Nonlinear inference with multiscale data

Single-scale correlated bivariate inference was demonstrated to be impractical, even with a highly correlated synthetic data set. Given sufficient multiscale, highly correlated data, it is demonstrated here to be possible to infer the nonlinear behaviour.

Consider the straight forward problem of estimating the nonlinear behaviour of a metallurgical random variable $Z_k, k = p$ given multiscale measurements for the variable. Bulk measurements on volumes V are available as well as n constituent samples at scale v per bulk sample where $V = nv$. Consider re-expressing the nonlinear variable Z_p as a linear variable Y using a power law transform. A power law transformation is a flexible re-expression framework for monotonic variables, as demonstrated earlier in this thesis.

In addition, an assumption is made that the error on the bulk samples is additive. This assumption is reasonable, and the correct re-expression is unknown at this point in any case. The form of this assumption of an additive, random error on the bulk sample measurements $z(\mathbf{u}_i; V)$ is expressed:

$$z_p(\mathbf{u}_i; V) = z_p(\mathbf{u}_i; V)' + \epsilon_p(\mathbf{u}_i; V), i = 1, \ldots, n$$

where $z_p(\mathbf{u}_i; V)'$ is the true value, and $\epsilon_p(\mathbf{u}_i; V)$ is the error associated with the measurement at the scale V. Therefore, linearly re-expressing as:

$$y_p(\mathbf{u}_i; V) = (z_p(\mathbf{u}_i; V))^\omega = (z_p(\mathbf{u}_i; V)' + \epsilon_p(\mathbf{u}_i; V))^\omega, i = 1, \ldots, n$$

and using a least squares regression method would introduce a bias. Instead of using a linear model, a nonlinear regression model is adopted to minimize the bias and maximize the covariance between the estimates and truth. The mean squared error criteria, which minimizes the variance of the estimator plus the bias squared is reasonable and used here.

This framework is demonstrated using data from Yan and Eaton (1994). A correlated Gaussian variable is simulated at the small scale v using sequential Gaussian simulation. This variable is normal score transformed to a linear volume fraction variable with a nominally uniform distribution between 0 and 1

(corresponding to the volume fraction of Vivian's ore). This variable can be converted to $z_p(v)$ values and upscaled to $z_p(V)$ values using experimental data from Yan and Eaton (Figure 3.30). Simulation study parameters are summarized in Table 3.2.

Table 3.2: Monte Carlo simulation study parameters for multiscale nonlinear inference.

Parameter	Value
Base nonlinear power for averaging	5.558
Gaussian variogram model	$0.2 + 0.8Sph(a = 25$ m$)$
Number of simulated small scale values	2500
Number of small scale values in large composite	5
Small scale value distribution	$Uniform$(6.6,14)

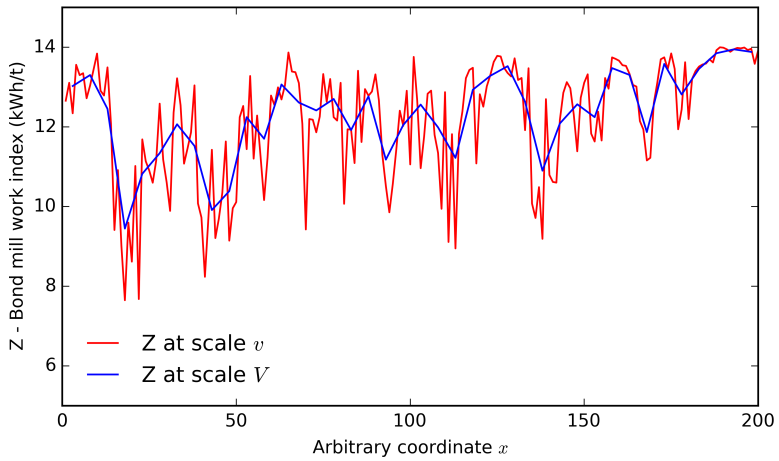

Figure 3.30: Synthetic multiscale samples of the Bond mill work index.

The relationship can be recovered using nonlinear least squares regression. The widely applied Levenberg-Marquardt algorithm (Marquardt, 1963) is effective for parameter inference problems such as the recovery of the nonlinear mixing law. The Levenberg-Marquardt algorithm is a hybrid algorithm using both the Gauss-Newton and gradient descent minimization algorithms. The nonlinear least squares problem is formulated as an objective function:

$$O(p) = \sum_{i=1}^{n} \left[z_p(\mathbf{u}_i; V) - \sum_{j=1}^{n} \lambda_j z_p(\mathbf{u}_j; v)^{\omega} \right]^2$$

where λ_j are the weights, equivalent to the volume fractions, of $z_p(\mathbf{u}_j; v)$ in the large scale composite samples $z_p(\mathbf{u}_i; V)$. The functional form presented is valid for $\omega \neq 0$ to simplify notation; however, the numerical implementation of this objective function includes the case where $\omega = 0$. For iteration t, given a relaxation parameter β, the Levenberg-Marquardt algorithm for this single parameter optimization problem is summarized as:

$$\begin{aligned} \omega_t &= \omega_{t-1} + \delta_t \\ \mathbf{J} &= \frac{d \sum_{j=1}^{n} \lambda_j z(\mathbf{u}_j; v)^{\omega_t}}{d\omega_t} \\ \delta_{t+1} &= \mathbf{J} \left[z(\mathbf{u}_i; V) - \sum_{j=1}^{n} \lambda_j z(\mathbf{u}_j; v)^{\omega_t} \right] / \left(\mathbf{J}^2 + \beta \mathbf{J}^2 \right) \end{aligned} \tag{3.6}$$

The Jacobian $\mathbf{J}$ is typically numerically approximated. The relaxation parameter is commonly determined using heuristics within the Levenberg-Marquardt implementation to balance slow convergence with robustness (Moré, 1978). The iterative algorithm is terminated after either a maximum number of iterations or once the objective function changes less than a provided threshold. For the simple example considered, where no error is associated with the measurements the mixing law is efficiently recovered using the nonlinear regression approach (Figure 3.31).

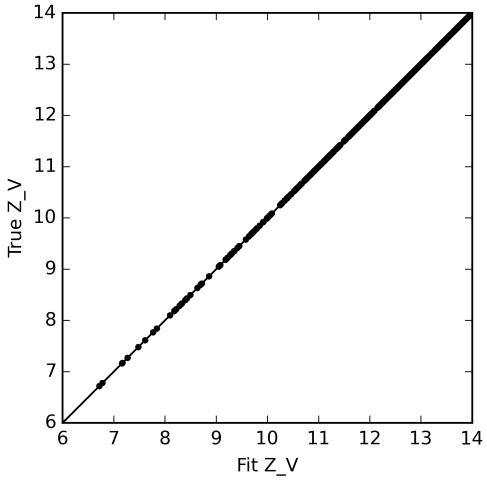

Figure 3.31: Cross plot of true large scale samples against fit large scale values using a power law fit with nonlinear regression.

This is the ideal case, analogous to the situation where multiple ore types are blended in known proportions and the responses measured. No measurement error was considered for this case, so the power law used was recovered exactly. In the case of no measurement error, the power fit to the Bond mill work index tests (5.56), was exactly recovered.

3.7 Measurement error and regression attenuation

Regression attenuation, also referred to as dilution, is the reduction in signal strength due to measurement error (Spearman, 1904). In the context of geometallurgy, attenuation is the reduction in predictive power and ability to construct accurate models in the presence of measurement error. This loss in accuracy accompanies a loss of precision in the estimates. Here, accuracy refers to estimates which reflect the true values and precision to the confidence interval ascribed to these estimates. Attenuation due to measurement error will affect all association measurements including covariances and regressions. The effect of attenuation for the nonlinear regression algorithms may result in a bias in models, and loss of accuracy, in the presence of measurement error or the usage of an imperfectly correlated secondary variable.

In the nonlinear regression model adopted for fitting power law models to multiscale samples, error is assumed to be concentrated on the large scale response measurements rather than the small scale predictor measurements. For a linear model under this assumption, and for error that is both unbiased and homoscedastic, explicitly accounting for attenuation due to measurement error is generally not required (Carroll et al., 2006). As demonstrated by Carroll et al., for a linear model with response measurement error, the variance of the fits will be increased, but will not be biased. This is not necessarily the case for nonlinear models. Error on the predictor variables, even if unbiased, homoscedastic and normally distributed, will attenuate both linear and nonlinear models. The effects of measurement error on the nonlinear model are investigated to identify potential sources of bias in the resulting models.

3.7.1 Potential attenuation due to measurement error

Consider adding an unbiased, homoscedastic measurement error with a standard deviation of 0.5 to the large scale samples: $\epsilon_p(\mathbf{u}_i; V) \sim N(0, 0.5)$ (corresponding to an approximately 5% error). The same nonlinear regression procedure was used to fit these large scale measurements with error to test the unbiasedness of the least squares nonlinear regressor (Figure 3.32). The slope of regression between the measured $Z_p(V)$ with error, and the fit $Z_p(V)$ is plotted; the slope of regression is approximately 1 with a value of 1.01 indicating that the fit has no conditional bias.

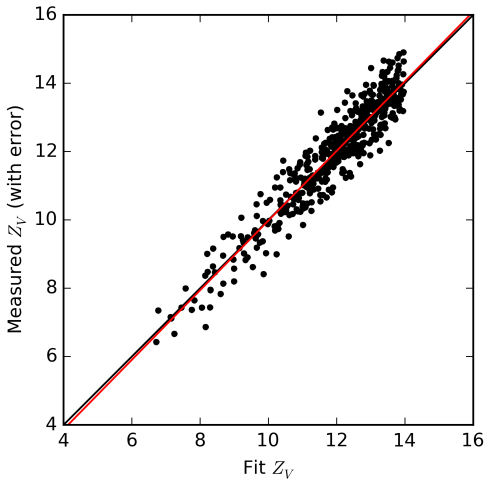

Figure 3.32: Cross plot of large scale samples with measurement error against fit large scale values using a power law fit with nonlinear regression. The slope of regression is overlaid in red.

The power fit to the samples with error is 4.16, lower than the original power of 5.56 indicating a potential bias in the power fit with respect to the true samples. In Figure 3.33, the same error term was also added to each of the small scale measurements (the predictor variables). The fit model is still conditionally unbiased with respect to the measured values in this case and is biased with respect to the true values.

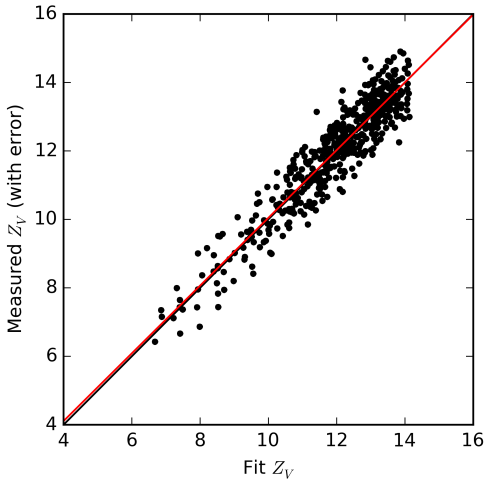

Figure 3.33: Cross plot of large scale samples with measurement error against fit large scale values using a power law fit with nonlinear regression applied to small scale values with measurement errors. The slope of regression is overlaid in red.

The regression model formulated in Equation 3.6 does not explicitly incorporate an error term where measurement errors may differ from sample to sample, but the same power attenuation is observed. If measurement errors differ from sample to sample, this error term could be incorporated as described in the original algorithm exposition (Marquardt, 1963). Even incorporating a measurement error term for the predictor, the problem of bias and model attenuation must still be addressed.

3.7.2 Theoretical framework for attenuation

Spearman (1904) documented a systematic reduction in the magnitude of calculated correlation coefficients when the correlation coefficients were inferred using data with sample error. This systematic reduction, termed correlation attenuation (also referred to as regression dilution), was observed even for an unbiased error. Critically for this thesis, this bias also occurs for linear models with predictor error, and may occur for nonlinear models with predictor or response error (Carroll et al., 2006; Frost and Thompson, 2000). Spearman proposed disattenuation of the attenuated correlation coefficient using the error associated with the samples. The linear model for correlation disattenuation, and equivalently linear regression disattenuation, with unbiased measurement error is derived here, as it forms the theoretical basis for the power law disattenuation model proposed in this thesis.

Disattenuation of correlation coefficients is derived using the mathematical framework of Spearman (1904). Consider inferring the correlation coefficient between two random variables X and Y. Consider independent, additive, normally distributed errors on both X and Y. Pairs of measurements of these random variables are therefore:

$$x_i^* = x_i + \epsilon_{x_i}, y_i^* = y_i + \epsilon_{y_i}, i = 1, \ldots, n$$

The experimental correlation coefficient is calculated on the random variables measured with error X^* and Y^*:

$$Corr\{X^*, Y^*\} = \frac{Cov\{X^*, Y^*\}}{\sqrt{Var\{X^*\}\, Var\{Y^*\}}} \tag{3.7}$$

Expanding the denominator, applying the assumption of independence between errors, and using the strong assumption of independence of the errors and their random variables:

$$\begin{aligned} Corr\{X^*, Y^*\} &= \frac{Cov\{X + \epsilon_X, Y + \epsilon_Y\}}{\sqrt{(Var\{X\} + Var\{\epsilon_X\})(Var\{Y\} + Var\{\epsilon_Y\})}} \\ &= Corr\{X, Y\} \frac{\sqrt{Var\{X\}\, Var\{Y\}}}{\sqrt{(Var\{X\} + Var\{\epsilon_X\})(Var\{Y\} + Var\{\epsilon_Y\})}} \end{aligned} \tag{3.8}$$

Disattenuation of the correlation under these assumptions therefore requires accurately measuring the true variance of the random variables and error components:

$$Corr\{X,Y\} = Corr\{X^*,Y^*\}\frac{\sqrt{(Var\{X\}+Var\{\epsilon_X\})(Var\{Y\}+Var\{\epsilon_Y\})}}{\sqrt{Var\{X\}\,Var\{Y\}}} \tag{3.9}$$

Disattenuation in the linear regression case is similar, (Carroll et al., 2006; Frost and Thompson, 2000). In the linear regression case, only measurement error in the predictor variables will result in a bias. The bias will result in an attenuated slope, with a magnitude closer to zero. Measurement error in the response variable will not result in a bias in the linear regression case. Carroll et al. claim that in the case of linear regression, an unbiased and homoscedastic error will increase the variance of the fit, but not introduce a bias. This is demonstrated with a simulation study. The additional claim is made that unbiased, homoscedastic errors on both linear and nonlinear regressors will increase the variability in fit parameters.

This increase in variability has a substantial effect on the nonlinear power law fits applied in this thesis. When nonlinear regression is used in these cases, error in the response variable may bias the regression. As demonstrated earlier with the power law, error in the predictor or response variable may attenuate the power fit. Carroll et al. attribute bias due to response error in nonlinear models to the increased variance biasing the Taylor approximation used by the nonlinear regressor and demonstrate this effect numerically for an exponential model. Numerical disattenuation of the power law models fit is considered here.

3.7.3 Attenuation of power law models

Using the Monte Carlo parameters defined earlier (Table 3.2), a set of realizations was constructed for varying degrees of response variable (large scale sample) and predictor variable (small scale sample) errors. This scenario was applied for a range of underlying power law models, from 1 (linear) to highly nonlinear cases with powers of -2 and 5. For this Monte Carlo experiment, 100

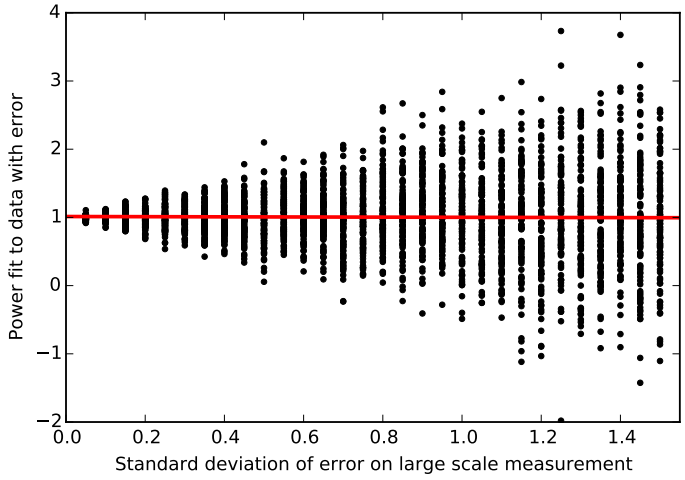

Figure 3.34: Power fit to data with a true underlying power of 1 (linear), with error in the response variable for varying degrees of error.

realizations of each scenario were generated and fit. The average power fit for each scenario was calculated.

Fit powers for the case where the true underlying model was 1 (linear) and the only error was in the response variable is shown in Figure 3.34. As expected for this linear case with response variable error, no bias is observed for any of the degrees of error, although the precision decreases substantially with increasing error. Results for the highly nonlinear case with a power of 5 are shown in Figure 3.35. No bias is observed in the results over the 100 realizations. Results were similar for all other powers tested; adding error to the large scale response variable in the case of the power law model does not introduce a bias to fit.

The case of predictor variable measurement error was also tested. As expected, substantial error introduced a bias, even in the linear case (Figure 3.36). As the nonlinearity of the power law model fit increased, the bias also increased (Figures 3.37, 3.38). In both of the nonlinear cases, increasing amounts of error substantially attenuated the power fit, driving it towards the linear case. If there is substantial measurement error in the predictor, or small scale, measurements when inferring mixing behaviour, then the nonlinear model fit will be biased with respect to the true values.

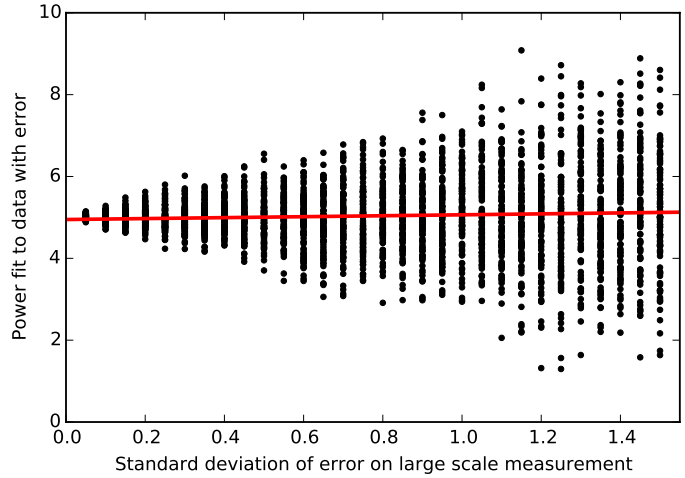

Figure 3.35: Power fit to data with a true underlying power of 5 (highly nonlinear) with error in the response variable for varying degrees of error.

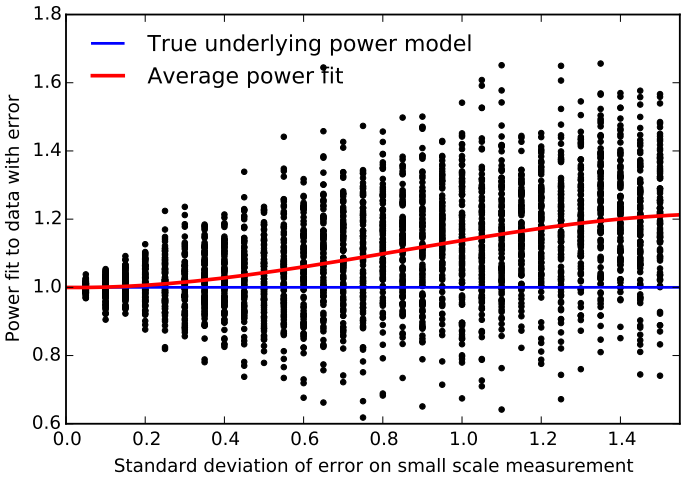

Figure 3.36: Power fit to data with a true underlying power of 1 (linear), with error in the predictor variable for varying degrees of error.

This bias introduces the possibility for regression disattenuation. If the measurement error is known, and quantified, then a similar Monte Carlo experiment could be conducted to understand the degree to which the measurement error attenuates the nonlinear model. Using this data, disattenuation could be applied to more accurately infer the underlying power without bias. The disattenuation procedure is demonstrated using the parameters from this Monte Carlo simulation study. The average powers fit, as a function of the

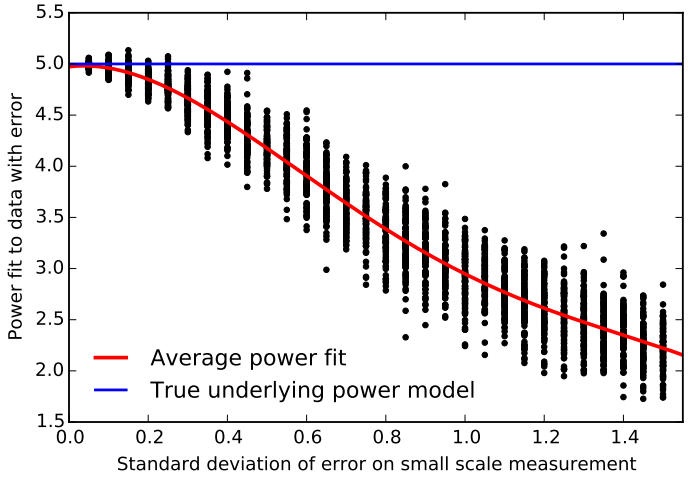

Figure 3.37: Power fit to data with a true underlying power of 5 (highly nonlinear), with error in the predictor variable for varying degrees of error.

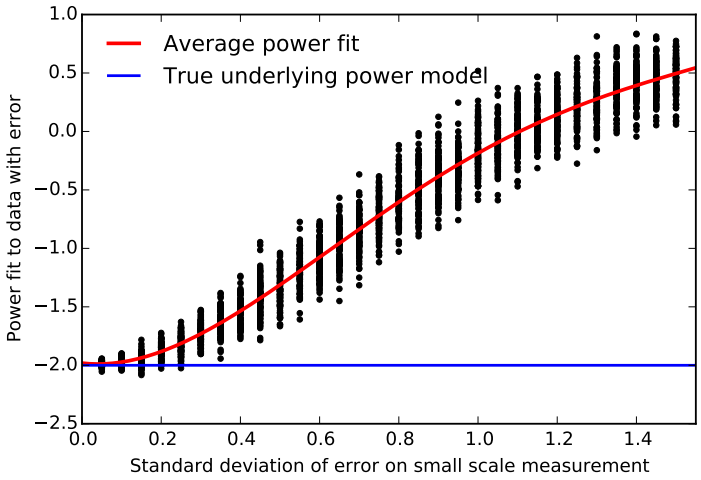

Figure 3.38: Power fit to data with a true underlying power of -2 (highly nonlinear), with error in the predictor variable for varying degrees of error.

underlying power and measurement error on the small scale measurements can be summarized in a nomograph (Figure 3.39). Given the measurement error, and biased fit of the power, a disattenuated power can be inferred. The degree to which the disattenuated power differs is a function of both the measurement error, and the degree of nonlinearity of the model.

Central to any discussion of regression disattenuation is the requirement to justify the need for disattenuation, particularly when the attenuated regres-

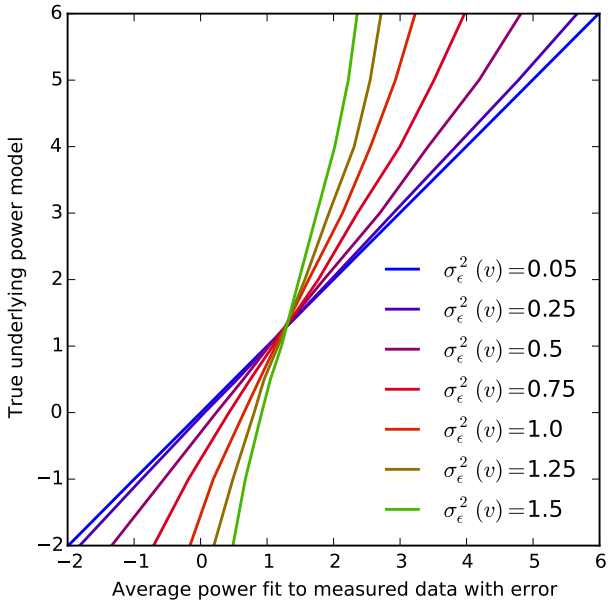

Figure 3.39: Nomograph of true underlying power as a function of the average power fit and error in small scale measured values.

sors are conditionally unbiased with respect to the measured values. If the regression is to be applied to values with the same experimental error, then using a disattenuated regressor could conditionally bias estimates. If, on the other hand, the regression is to be applied to values which are considered to have no error, then a disattenuated regressor is required to avoid biasing the results.

In the context of geometallurgical modeling, mixing behaviour is primarily quantified for the purposes of 1) modeling nonlinear metallurgical variables with linear frameworks and 2) modeling the behaviour of blends. This modeling will take the form of conditional geostatistical realizations of the deposit. Critically, these realizations are individually assumed to be without measurement error. Measurement error is incorporated into the uncertainty observed when considering all realizations together, which constitute the geometallurgical model. Therefore, for the purposes of geometallurgical modeling, the power law models of nonlinear behaviour should be disattenuated where measurement error can be quantified.

As demonstrated here, multiscale measurements may be used for fitting and inferring mixing laws of metallurgical variables. Single scale measurements are insufficient to infer the mixing law, even given a highly correlated linear variable. The hybrid of these situations are large scale measurements of the metallurgical variable along with small scale measurements of a correlated linear variable. A multiscale semiparametric Bayesian updating algorithm was developed to fit mixing models for this situation.

3.8 Multiscale correlated bivariate inference

The related, but more challenging, problem is the inference of nonlinear behaviour in a metallurgical random variable Z_p given bulk measurements for the variable and correlated, fine scale, constituent grades. This problem is addressed with a semiparametric Bayesian updating algorithm. Recall the semiparametric Bayesian updating algorithm used by Barnett and Deutsch (2015) for multiple imputation of missing data which was reviewed earlier. The case of missing data is analogous to this situation, although data are considered missing everywhere and the bivariate distribution is considered known, not strictly the conditional distributions. Bulk measurements of the metallurgical variable on volumes V are available as well as grade variable measurements on n constituent samples at scale v per bulk sample where $V = nv$. As before, consider re-expressing the nonlinear variable Z_p as a linear variable Y_p using a power law transform.

The semiparametric Bayesian updating algorithm may be implemented using a normal score framework, in which all calculations are done after normal score transformation and back-transformed after, or using a multigaussian kriging approach. Both approaches lead to the same end result so the choice is primarily driven by the available information. If bivariate samples of the metallurgical variable $Z_p(v)$ and linear variable $X_p(v)$ are available at the small scale, then the normal score approach is computationally faster. If no small scale samples are available, and therefore the small scale bivariate distribution has been inferred by downscaling, then the multigaussian approach is simpler

to implement. The semiparametric Bayesian updating algorithm for both approaches is shown considering each data collection scenario. The normal score transformation approach is summarized in Figure 3.40 and the multigaussian kriging approach in Figure 3.41.

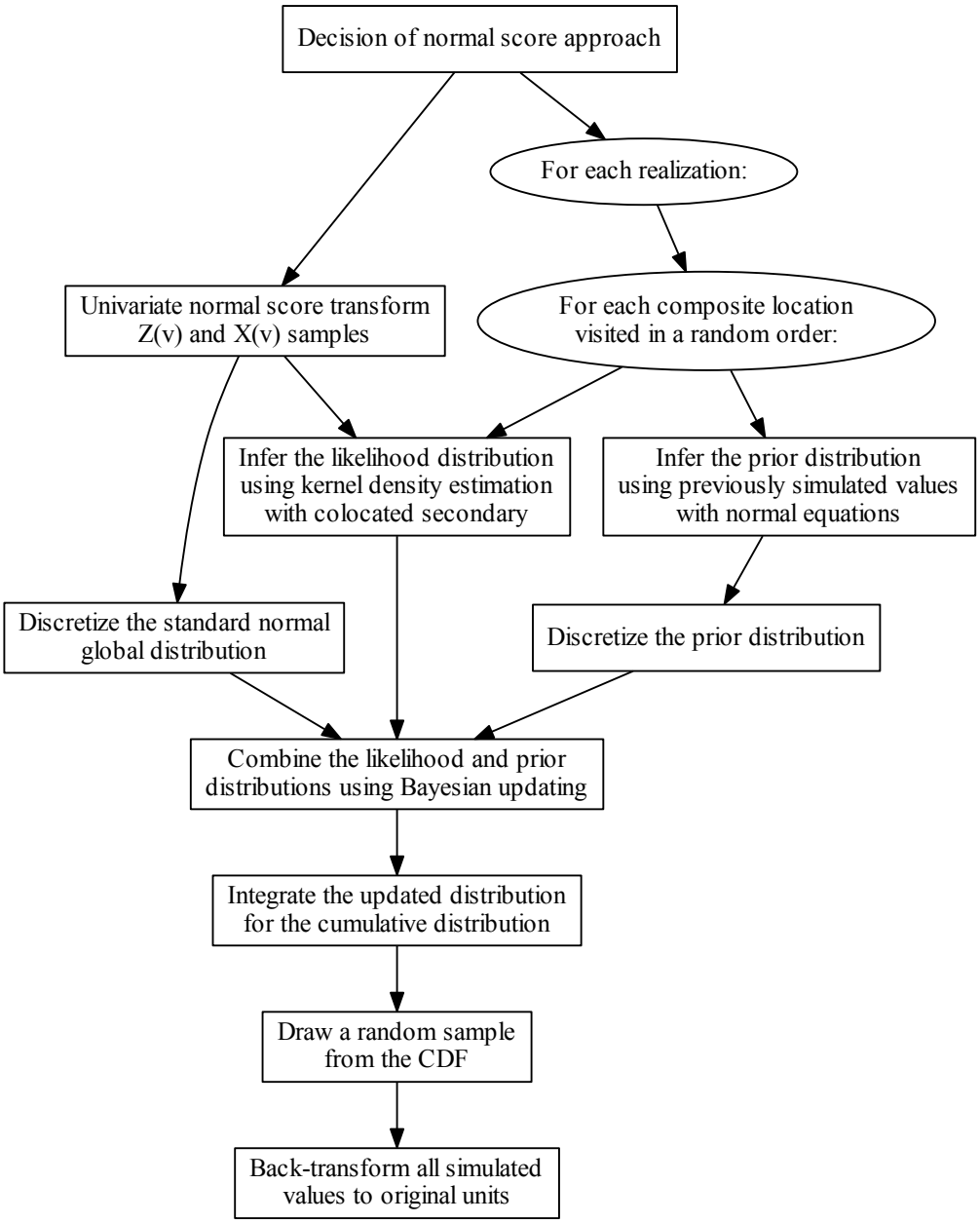

Figure 3.40: Flowchart of the semiparametric Bayesian updating algorithm using the normal score transformation approach where a number of small bivariate samples are available.

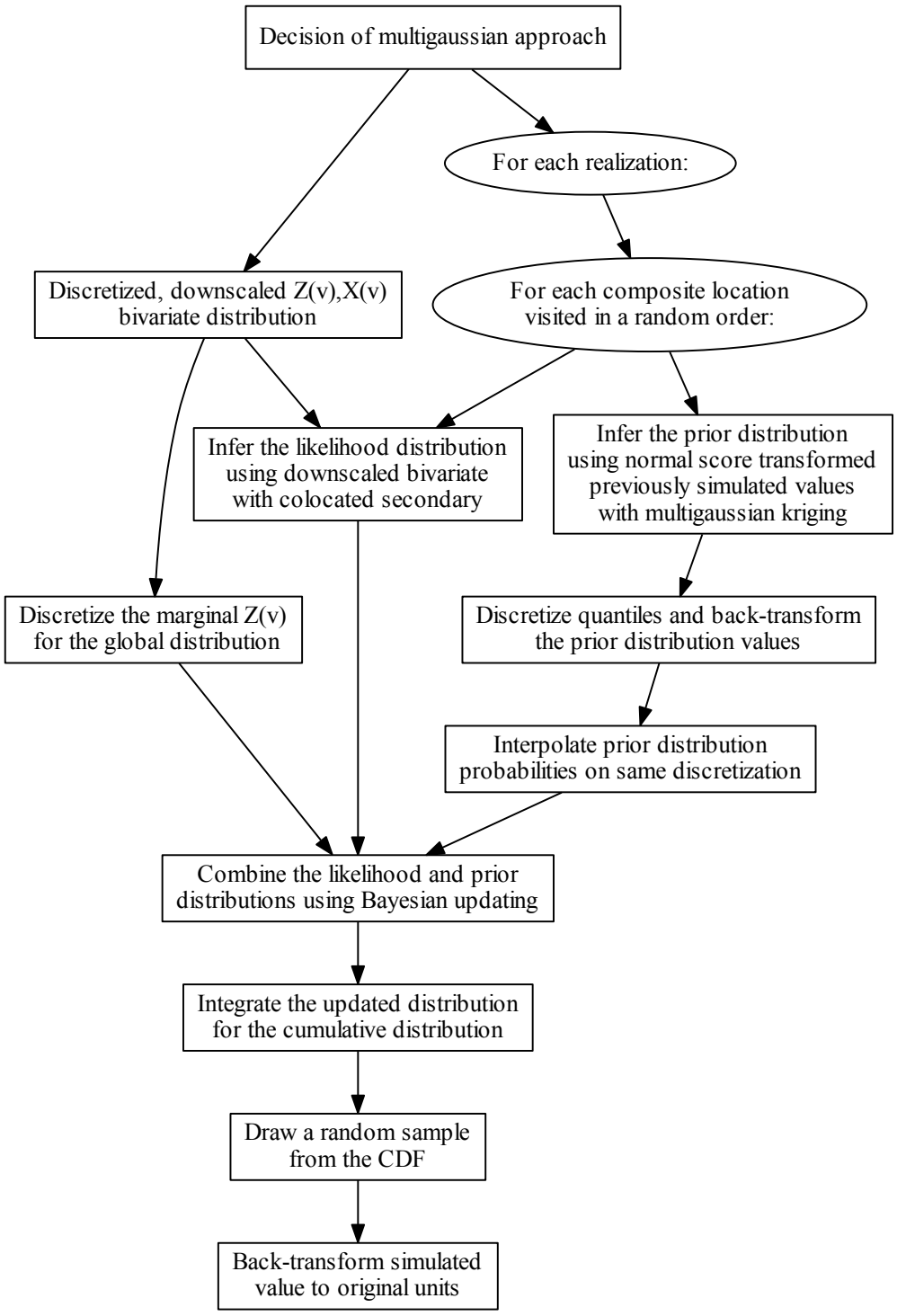

Figure 3.41: Flowchart of the semiparametric Bayesian updating algorithm using the multigaussian kriging approach where the small scale bivariate distribution is inferred by downscaling.

In the normal score approach, the bivariate distribution of $F(Z(v), X(v))$ is normal score transformed and fit using kernel density estimation. This small scale bivariate distribution is required for inferring the likelihood distributions. If samples of $F(Z(v), X(v))$ have not been directly sampled, then downscaling the bivariate distribution $F(Z(V), X(V))$ may be required and the multigaussian approach may be used directly. At this stage, it is assumed that samples of $F(Z(v), X(v))$ are available, or have been downscaled.

For each large scale composite sample available, multiple realizations of the small composite values are constructed. Multiple realizations are generated so that samples of the updated distributions can be fit, not the distributions themselves. This leverages the Monte Carlo simulation paradigm used throughout geostatistics, and this thesis. A random path through the small scale $X(v)$ samples contained in the composites is generated for each realization.

For each location along the random path, the likelihood distribution is inferred nonparametrically using kernel density estimation of the normal score transformed bivariate distribution of $F(Z(v), X(v))$. The likelihood distribution for location $\mathbf{u}_J$, where locations $\mathbf{u}_j, j = 1, \ldots, J-1$ have been previously simulated, for realization l using the kernel density estimate (Eq. 2.8) is therefore:

$$f(y_p^l(\mathbf{u}_J; v)|x(\mathbf{u}_J)) = \frac{1}{\Delta} \int N_h(y_p(\mathbf{u}_J; v), x(\mathbf{u}_J; v), k = 1, \ldots, K, k \neq p) dy_p(\mathbf{u}_J; v) \tag{3.10}$$

The prior distribution is estimated parametrically using multigaussian kriging (the normal equations) of the previously simulated values. This is the direct application of the normal equations (Eq. 2.2-2.4):

$$\begin{aligned} &\sum_{j=1}^{J-1} \lambda_j \rho_{i,j}(v) = \rho_{i,J}(v), i = 1, \ldots, J-1 \\ &\sigma^2_{k,SK}(\mathbf{u}_J; v) = \rho_{J,J}(v) - \sum_{i=1}^{J-1} \lambda_i \rho_{i,J}(v) \\ &y_p^*(\mathbf{u}_J; v) = \sum_{i=1}^{J} \lambda_i y_k(\mathbf{u}_i; v) \end{aligned} \tag{3.11}$$

The estimated prior probability distribution is discretized, and combined with the likelihood distribution using the nonparametric Bayesian updating equations under a permanence of ratios model (Eq. 2.6). The updated probability distribution is integrated for the cumulative distribution, and a realized value is drawn with Monte Carlo simulation. After simulation, all values are back-transformed to original units and simultaneously fit using nonlinear regression. This procedure is identical to the previously demonstrated nonlinear regression method for multiscale samples.

3.8.1 Demonstration of the algorithm

A synthetic example was constructed to demonstrate the semiparametric Bayesian updating algorithm for nonlinear inference using the normal score approach. The synthetic example was constructed by sequential Gaussian simulation of values for a nonlinear metallurgical variable Z and correlated linear variable X at a small scale v. The simulation study parameters are summarized in Table 3.3.

Table 3.3: Monte Carlo simulation study parameters used to demonstrate semiparametric Bayesian updating for nonlinear inference.

Parameter	Value
Base nonlinear power for averaging	2.5
Correlation at v in Gaussian units	0.80
Gaussian unit variogram model	$0.2 + 0.8Sph(a = 25$ m)
Number of simulated small scale values	2500
Number of small scale values in large composite	5
Z_v distribution	$Uniform$(6.6,14)
X_v distribution	$Uniform$(0.0,5.0)

Bivariate samples of the simulated values at scale v are shown in Figure 3.42. The high correlation between the variables can be observed in this bivariate scatterplot. A portion of the 1-dimensional simulated values and upscaled samples of the nonlinear variable Z is shown in Figure 3.43.

The bivariate samples at v were fit using kernel density estimation with a correlated Gaussian kernel. The kernel bandwidth (0.07) was selected by

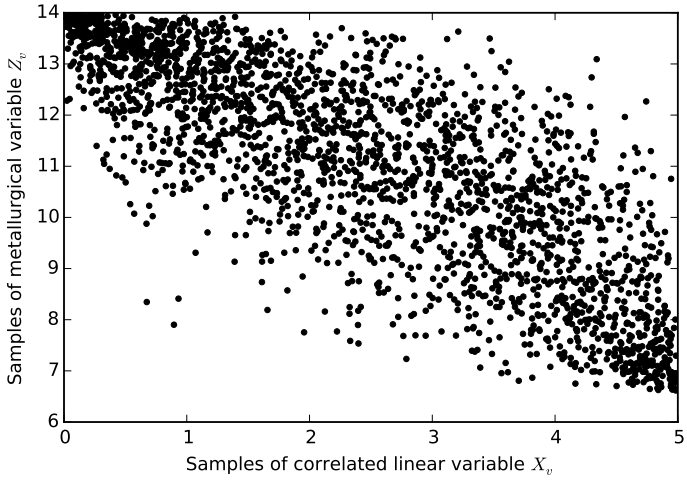

Figure 3.42: Bivariate samples of the simulated data values at v.

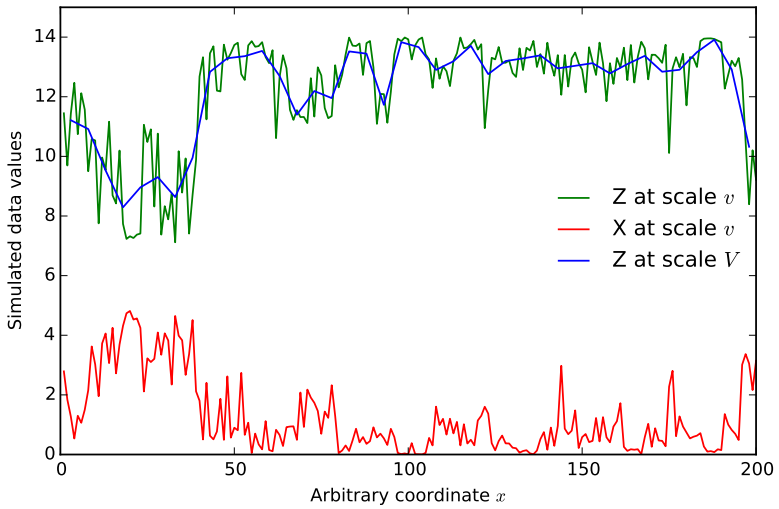

Figure 3.43: Portion of the nonlinear data simulated at scale v and for Z, upscaled to scale V.

choosing the smallest bandwidth which in the author's opinion did not overfit the data. The subjective choice of the kernel density estimation bandwidth selection is discussed later in this thesis. The estimated bivariate probability distribution contours calculated using a 100x100 discretization are shown in Figure 3.44.

Using the selected Gaussian kernel function, the conditional probability

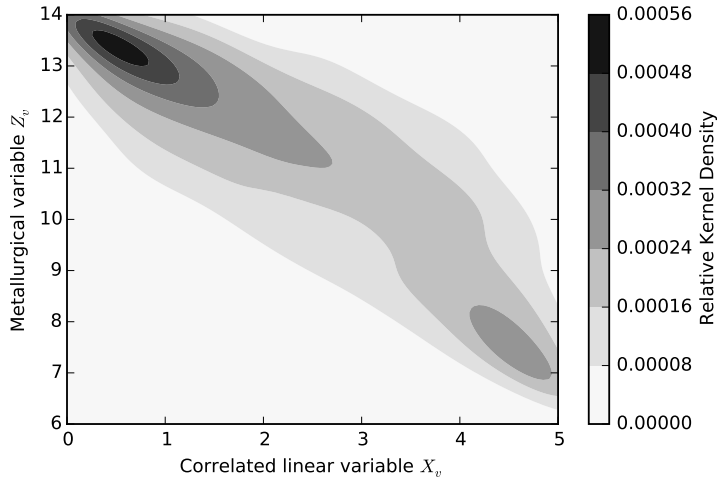

Figure 3.44: Kernel density estimation contours fit to bivariate samples of the simulated data values at v.

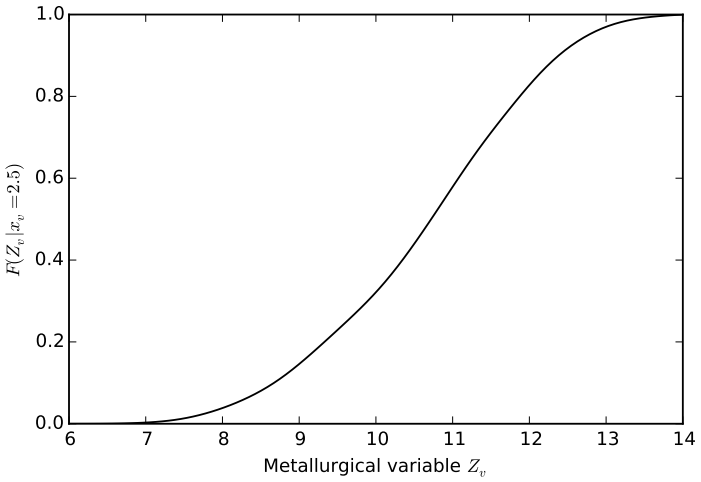

Figure 3.45: Conditional probability distribution function evaluated at $x_v = 2.5$.

distribution $F(Z_v|x_v = x)$ may be queried for any x_v. An example conditional probability distribution where $x_v = 2.5$ is shown in Figure 3.45.

The synthetic samples of X_v and Z_V were used with a semiparametric Bayesian updating workflow to infer the nonlinear behaviour. For this demonstration of the workflow, values of Z_v corresponding to the large Z_V samples are known. These samples were not used directly for fitting the nonlinear behaviour, as this would degenerate the case study to a multiscale inference problem previously demonstrated. The bivariate probability density function

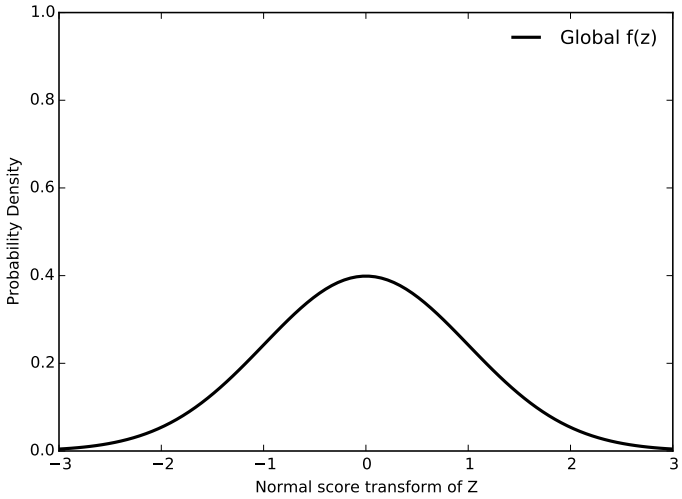

Figure 3.46: Global probability distribution function for Bayesian updating. Normal score units are used, so the global distribution is the standard normal distribution.

$f(X_v, Z_v)$ is required for the Bayesian updating workflow, so the kernel density estimation of the probability function fit to the samples was used. If bivariate samples were not available, distribution scaling (discussed in the scaling portion of this thesis) could alternatively be used.

Following the semiparametric Bayesian updating workflow, samples of Z_V and composite X_v samples were normal score transformed. The global distribution of the normal score transformed Z_V values is therefore the standard normal distribution shown in Figure 3.46.

Consider inferring the prior probability distribution in the algorithm. This prior probability distribution is inferred by multigaussian kriging of previously simulated Z_v samples within the Z_V composite. If no samples have been previously simulated, the prior degenerates to the global distribution (the standard normal distribution). Consider that previously simulated values are available, and the kriging mean and standard deviation are 1.1 and 0.6, respectively. The prior probability distribution function is therefore given by the normal distribution with a mean of 1.1 and standard deviation of 0.6, as shown in Figure 3.47.

Under a Markov assumption, collocated secondary values of X_v at the

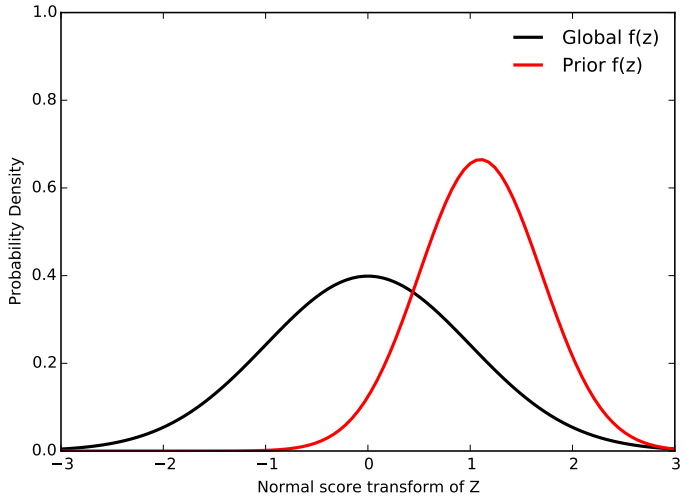

Figure 3.47: The prior probability distribution is the normal distribution with the mean and variance given by multigaussian kriging of previously simulated values.

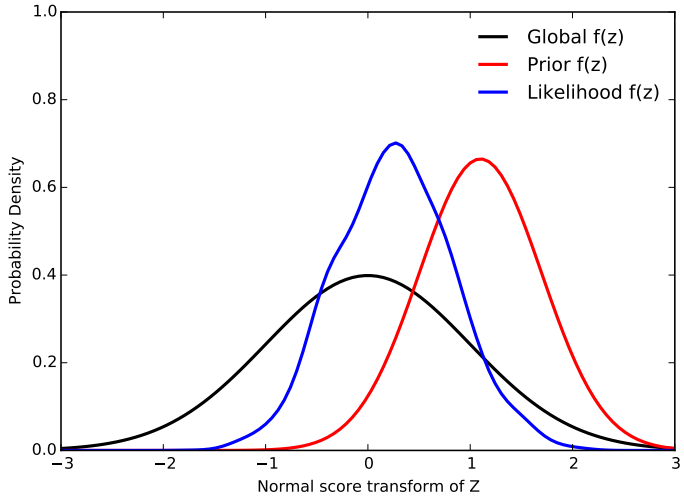

Figure 3.48: The likelihood distribution is evaluated using kernel density estimation with the secondary variable at the location being simulated under a Markov assumption.

location being simulated are used to infer the likelihood distribution. For this example, this is evaluated using kernel density estimation (as demonstrated previously in Figure 3.45). A sample likelihood distribution with a normal score $x_v = -0.3$ is shown in Figure 3.48.

These probability distributions are merged using the nonparametric Bayesian

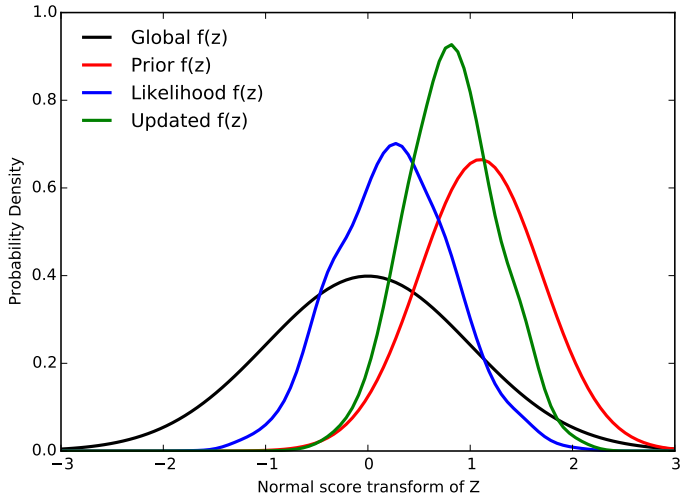

Figure 3.49: The updated probability distribution function is calculated using semiparametric Bayesian updating.

updating equations. The updated probability distribution is shown in Figure 3.49.

This updated probability distribution is integrated to a cumulative probability distribution function and sampled for a simulated value. The simulation algorithm then proceeds, and this procedure is repeated for each location for each realization required. Integrating to a cumulative probability distribution function is not strictly necessary to sample the distribution; an alternative algorithm such as rejection sampling could be used instead.

Once L realizations of Z_v^l have been constructed and back-transformed to original units, the nonlinear regression procedure used for multiscale inference may be directly applied. For this small example, a power of 3.8 was fit using nonlinear regression.

3.9 Parameters, the information effect and implementation for semiparametric Bayesian updating

A number of key areas with the application of semiparametric Bayesian updating are addressed using a series of Monte Carlo simulation experiments. The selection, and sensitivity, of a bandwidth for kernel density estimation is ap-

proached comparing the results of a large range of bandwidths. The selection of a variogram for the multigaussian kriging step is also introduced. Finally, the information contained in a secondary variable for inferring the nonlinearity and potential issues with nonlinear attenuation are considered in the context of challenges with measurement error previously discussed.

3.9.1 Kernel density estimation bandwidth selection

The general kernel density estimation criteria for bandwidth selection is the minimization of the expected mean integrated squared error (**MISE**) given the selection of a bandwidth H:

$$\mathbf{MISE}(H) = E\left\{\int \left[f_H^*(\mathbf{x}) - f(\mathbf{x})\right]^2 d\mathbf{x}\right\}$$

where $f_H^*(\mathbf{x})$ is the multivariate kernel density estimate at the multivariate vector $\mathbf{x}$ and $f(\mathbf{x})$ is the unknown, true multivariate distribution. For practical evaluation, this criteria is revised to consider an asymptotic or numerical approximation. Assuming a multivariate Gaussian distribution as the true underlying multivariate distribution, the optimal bandwidth for a multivariate Gaussian kernel is given by Silverman (1986). The square-root bandwidth used for the Gaussian kernel for dimension i for n samples, fit to d dimensions, with a sample standard deviation of σ_i is:

$$\sqrt{H_i} = \left(\frac{4}{2+d}\right)^{\frac{1}{4+d}} \sigma_i n^{\frac{-1}{4+d}}$$

Using the example demonstration of the nonlinear inference algorithm under these assumptions, the optimal square-root bandwidth is therefore 0.073. This rule of thumb, and the sensitivity of the bandwidth for the kernel density estimation was investigated using the previously demonstrated example, varying the bandwidth used. For each bandwidth tested, 10 synthetic realizations of the data were constructed and fit using the semiparametric Bayesian updating approach. The results are plotted in Figure 3.50.

Using too small of a bandwidth results in substantial instability in the likelihood distributions and problematic results for numerical solutions. Based on

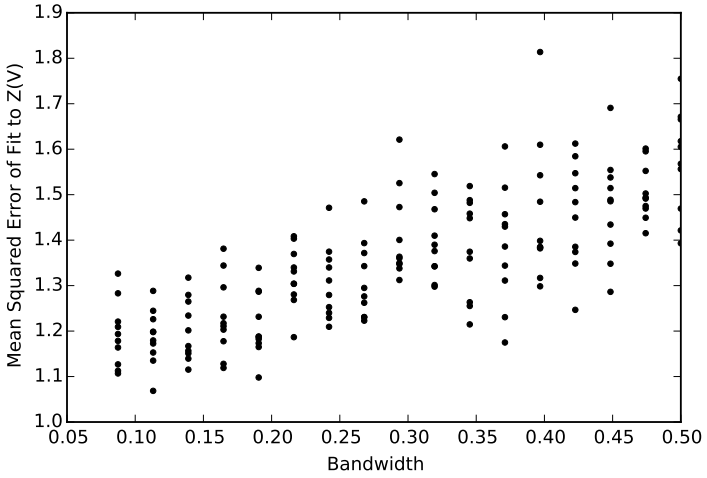

Figure 3.50: Mean squared error of power law fit to data as a function of the bandwidth used for kernel density estimation.

the results of this small test of bandwidth, for practical application of the semiparametric Bayesian updating approach there is a large range of acceptable bandwidths. Using the criteria of Silverman with a correlated multivariate Gaussian kernel works well.

3.9.2 Variogram inference for multigaussian kriging

The spatial covariance is required for the multigaussian kriging step for inferring the prior distribution. Using a variogram to infer the covariance requires a variogram at scale v for the normal score transform of Z. This is most easily accomplished by downscaling the variogram. Scaling algorithms for the variogram, and other statistics, are discussed in chapter 4.

3.9.3 The information effect

Central to the semiparametric Bayesian updating workflow is the assumption that there is a related linear variable at scale v which is related to the metallurgical variable at v. The limiting case of this assumption is a perfectly unrelated variable, in which case it is impossible to infer any nonlinear behaviour, or a perfectly correlated variable which reduces the multiscale sampling case. We are interested in quantifying our position on the spectrum of no information to perfect information for the expected nonlinear relationship. The correlation,

which is a linear measure of the relationship, is therefore not used to measure this information.

Instead, the relative information about Z_v given X_v is quantified using the mutual information and uncertainty coefficient. The mutual information is a measure of the relative information with the units of nats (the base of the natural logarithm), and the uncertainty coefficient is the dimensionless measures of the reduction in uncertainty in Z_v given X_v. The mutual information $I(X;Z)$ of Z and X is (Cover and Thomas, 2012):

$$\begin{aligned} I(X;Z) &= H(X) - H(X|Z) \\ &= -\int_x f(x)\ln(f(x))dx - \int_x\int_z f(x,z)\ln\left(\frac{f(z)}{f(x,z)}\right)dxdz \end{aligned} \tag{3.12}$$

where $H(X)$ is the entropy of X and $H(X|Z)$ is the conditional entropy of X, Z. The normalized mutual information is the uncertainty coefficient, defined as:

$$U(X;Z) = \frac{I(X;Z)}{H(X)}$$

The uncertainty coefficient, U, is 1 when X exactly determines Z, and 0 when X and Z are perfectly unrelated. For the purposes of generally quantifying the relative information contained in bivariate distributions, using the normalized uncertainty coefficient in place of the mutual information is reasonable. Three examples of contoured bivariate distributions, fit with kernel density estimation with a bandwidth determined by Silverman's rule, are shown in Figures 3.51–3.53. The weak relationship between X and Z in Figure 3.51 corresponds with an uncertainty coefficient of 0.109. The strong relationship in Figure 3.53 has an uncertainty coefficient of 0.234.

The simulation study which has been used to demonstrate bandwidth selection and the semiparametric Bayesian updating workflow was repeated using bivariate distributions with a range of uncertainty coefficients from very weakly related, to strongly related. The mean squared error of the power law fit is plotted for these scenarios (Figure 3.54). Multiple realizations were computed and fit for each scenario, each time using a different synthetic data set. As expected, an increasing uncertainty coefficient decreases the mean squared er-

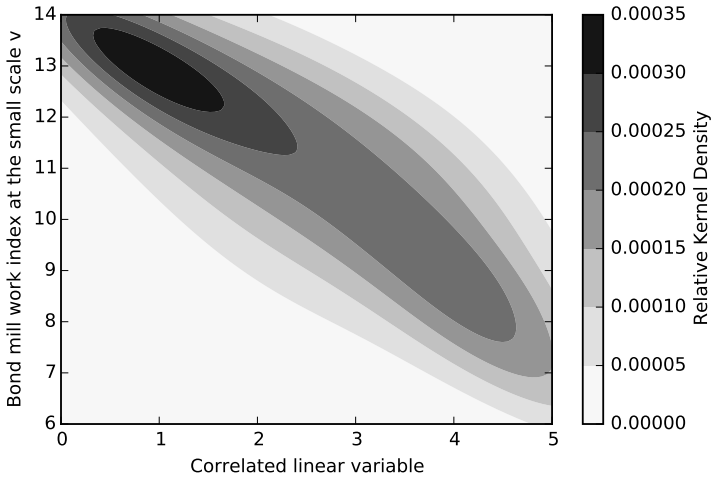

Figure 3.51: Contoured bivariate distribution with an uncertainty coefficient of 0.109.

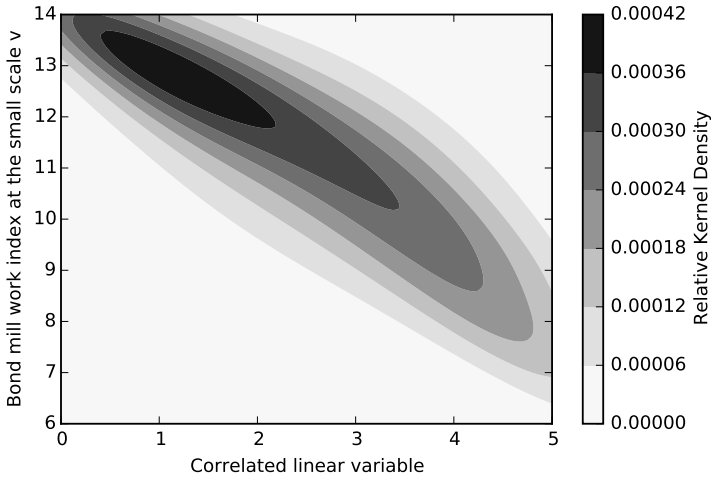

Figure 3.52: Contoured bivariate distribution with an uncertainty coefficient of 0.140.

ror of the fit. The limiting case of this relationship, where the uncertainty coefficient is 1 and $X = Z$ reduces the multiscale regression approach.

3.9.4 Application with no small scale samples

As part of this thesis, a large South American copper-molybdenum porphyry deposit was modeled. Modeling of this deposit is the focus of chapter 6; however, a portion of the data is used here to show potential application of

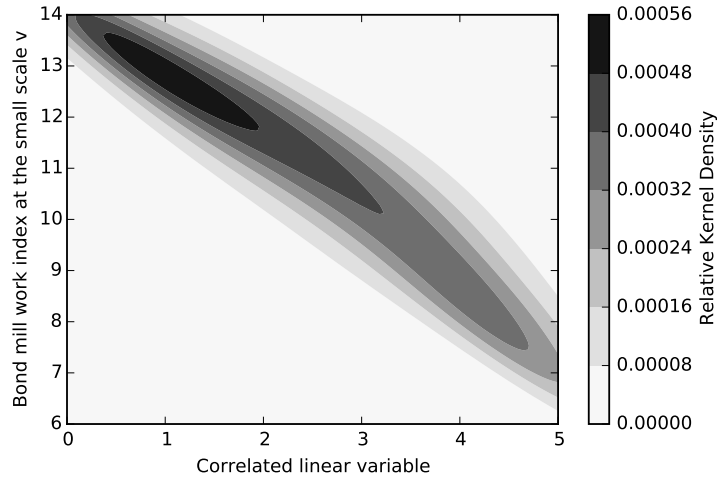

Figure 3.53: Contoured bivariate distribution with an uncertainty coefficient of 0.234.

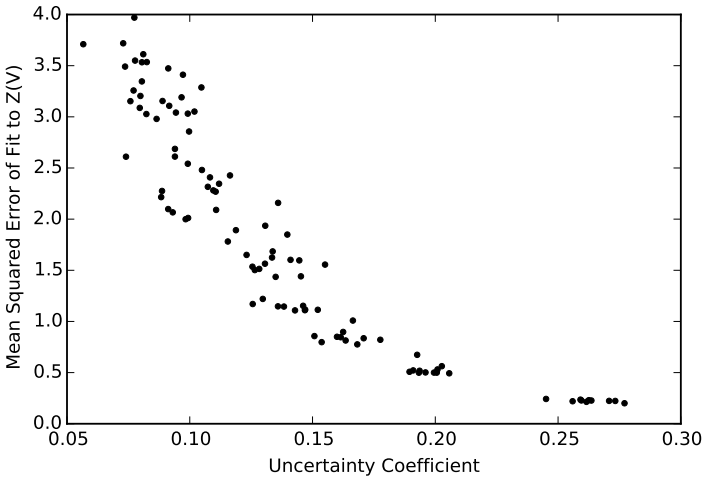

Figure 3.54: Mean squared error of power law fit to data as a function of the uncertainty coefficient of the small scale bivariate distribution.

the nonlinear re-expression framework. Iron grade, available at a composited scale of 15 m and Bond mill work index, available on nominally 30 m samples are available. A cross plot of iron grades, upscaled to the 30 m are cross plotted against Bond mill work index, and the bivariate distribution fit, in Figure 3.55. The variables are moderately correlated, with a correlation coefficient of 0.5. This correlation coefficient was too low to fit a reasonable re-expression, but

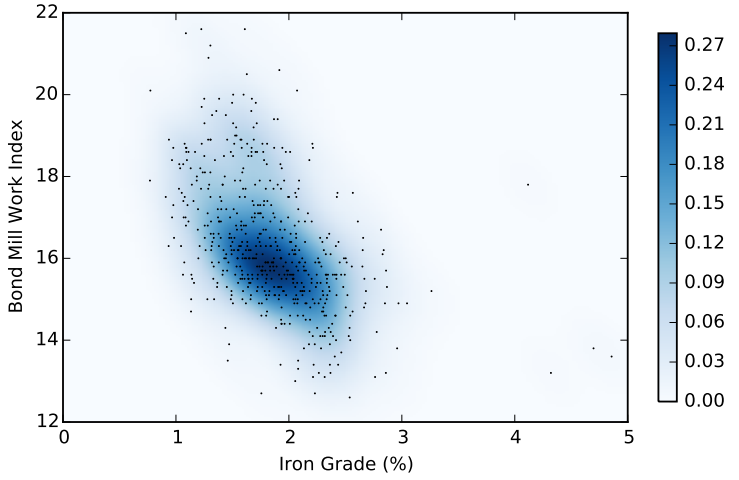

Figure 3.55: Kernel density estimation of bivariate distribution between Bond mill work index at original 30 m scale and upscaled iron grade. Individual bivariate samples are overlaid on the distribution.

the procedure for fitting the re-expression framework is demonstrated here for future use.

Statistics including the variogram, univariate histogram and bivariate distribution of the Bond mill work index are required at the 15 m scale. Downscaling statistics, and data, is the focus of chapter 4 so the reader is referred there for additional details. An assumption of either linearity, or a re-expression is required. As no data are available on any nonlinearity in the Bond mill work index for this deposit, linearity was assumed for statistic downscaling. This could be adjusted iteratively using the results of the nonlinear regression. The standardized variogram of the Bond mill work index was modeled at the 30 m scale, and downscaled using the variogram scaling laws to 15 m (Figure 3.56). Due to a large amount of short-range variability, the downscaled Bond mill work index has a variance 1.4 times as large as the original variance.

The univariate histogram of Bond mill work index was downscaled using the discrete Gaussian model (DGM) with a mosaic correction (Figure 3.57). The mosaic model was chosen as the distribution is not very close to zero which would have supported the usage of an indirect lognormal correction as

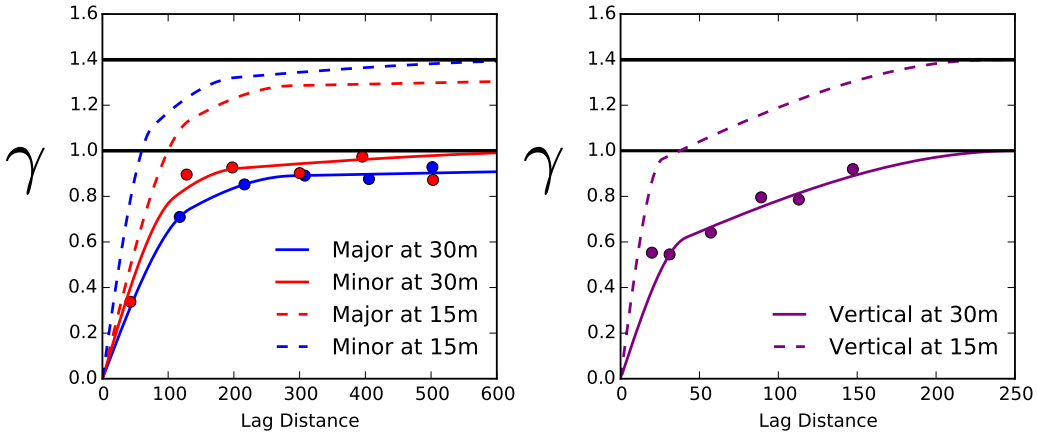

Figure 3.56: Standardized Bond mill work index variograms (experimental and modeled) at 30 m, and downscaled variograms with the new downscaled sill of 1.4 (relative to the standardized sill of 1).

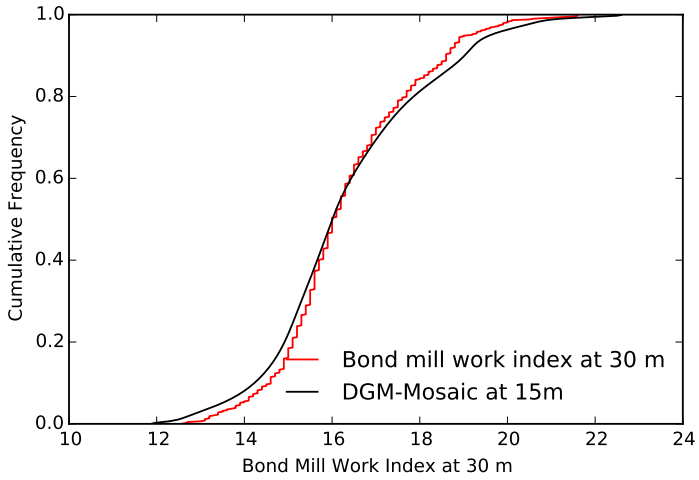

Figure 3.57: Downscaled cumulative distribution function of Bond mill work index using the discrete Gaussian model (DGM) with a mosaic correction, and original experimental cumulative distribution at 30 m.

in the case of the bitumen grades. This downscaled distribution has the correct variance matching the downscaled variogram.

The bivariate distribution was downscaled using a product of ratios approach. Univariate distributions which form the target for the downscaling are shown in Figure 3.58. After the iterative downscaling of the bivariate, the marginals exactly match (Figure 3.59). The downscaled bivariate distribution is shown in Figure 3.60.

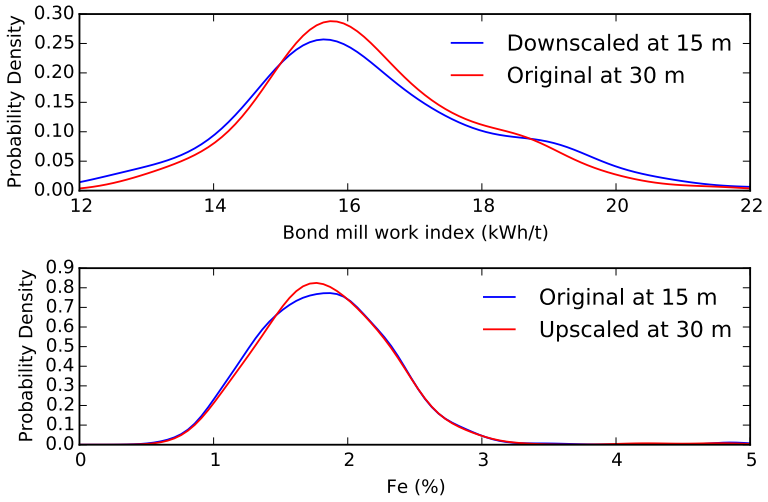

Figure 3.58: Univariate probability distributions for Bond mill work index and iron grade at a 15 m and 30 m scale.

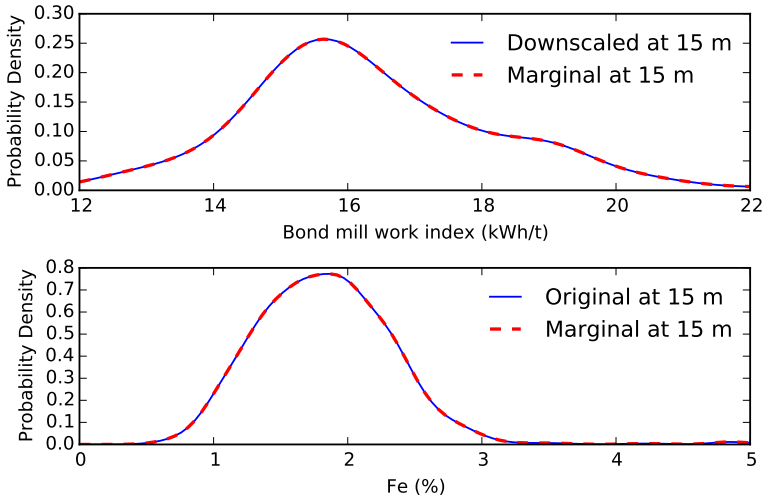

Figure 3.59: Marginal distributions of downscaled bivariate probability distributions for Bond mill work index and iron grade at a 15 m scale which exactly reproduce the target histograms.

The multigaussian approach for semiparametric Bayesian updating was used with this bivariate distribution; however, due to the low correlation no conclusions may be drawn on the nature of the nonlinear averaging of the Bond mill work index. Cross plots of simulated results using the same random number seed and powers of 0 and 3 are shown in Figures 3.61 and 3.62, respec-

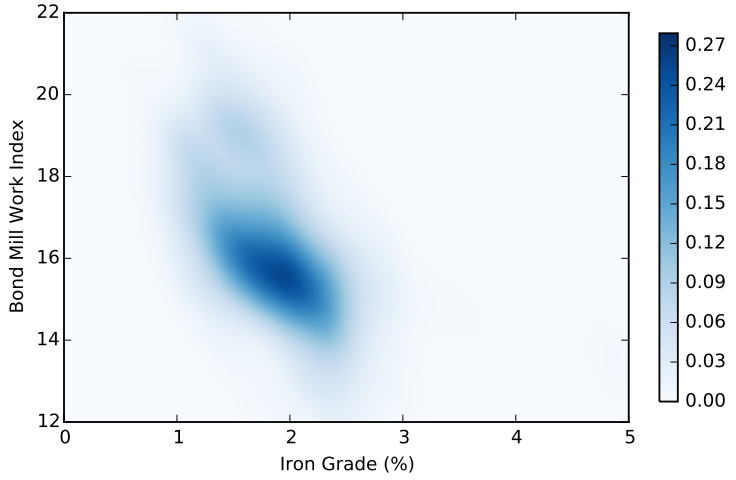

Figure 3.60: Downscaled bivariate distribution at 15 m using the product of ratios method.

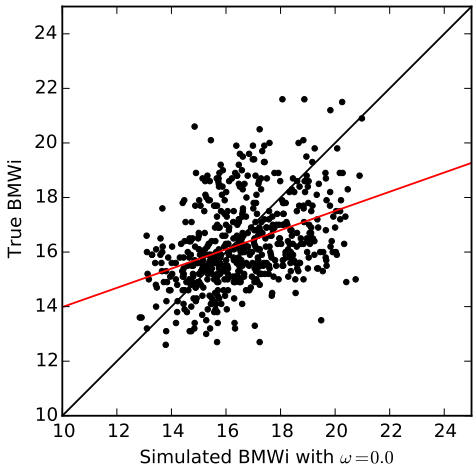

Figure 3.61: Cross plot of upscaled, simulated Bond mill work index values using $\omega = 0.0$ and true values.

tively. The slope of regression, and mean squared error are almost identical for these cases, so no conclusion on nonlinearity may be drawn.

3.10 Discussion of nonlinearity and mixing laws

This chapter has focused on the use and inference with nonlinear metallurgical properties for statistical modeling. The treatment of nonlinear variables is based not on the development of statistical techniques which are aimed

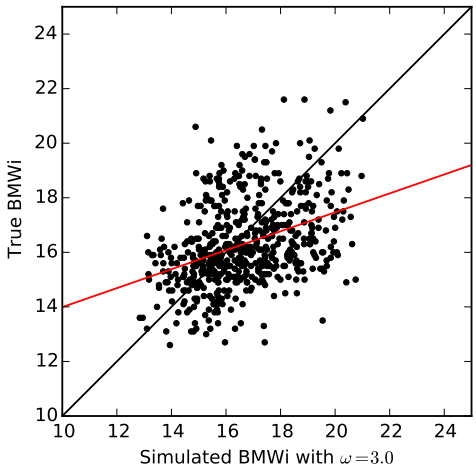

Figure 3.62: Cross plot of upscaled, simulated Bond mill work index values using $\omega = 3.0$ and true values.

at directly modeling nonlinear variables, but instead on the re-expression of nonlinear properties as linear variables which may be modeled using current statistical techniques. The ability to use current techniques is a substantial benefit to the use of re-expression as there is a very large body of algorithms and expert knowledge in the application of statistical and geostatistical models for linear variables. Three key developments of this chapter are discussed: the importance of modeling nonlinear behaviour, the use of a power law re-expression framework for fitting nonlinearity and methods to fit the nonlinear behaviour given varying amounts of information.

3.10.1 The importance of modeling nonlinear behaviour

The importance of modeling nonlinearity is emphasized when the small simulation study modeling expected mixing in a grinding stage, current practice in mineral processing, and mineral processing literature are considered. A high degree of mixing within the mill will occur even when material is batch fed on a truck-by-truck basis. This high degree of mixing emphasizes the requirement to quantify nonlinear mixing behaviour in metallurgical variables if possible to avoid biasing the resulting decisions.

Furthermore, models of nonlinear behaviour are required for spatial model

upscaling and blending. Upscaling from a point-scale model to a block-scale model requires a method to correctly average properties which average nonlinearly. Accurately determining the effect of blending multiple upscaled blocks also requires a model of nonlinear averaging.

3.10.2 The power law re-expression framework

The power law re-expression framework is proposed as the primary technique for nonlinear re-expression. The use of a power law model is supported both by its flexibility for modeling and correctness for specific types of mixtures. The theoretical correctness of the power law transform for specific types of mixtures, documented by Korvin (1982), does not apply to all metallurgical properties, but is applicable in certain cases. In many others, the power law still fits the mixture well.

Alternative frameworks, such as the Scheffé canonical polynomials are also reasonable possibilities for properties with more complex behaviour, such as the antagonistic recovery observed in froth flotation recovery of certain blends.

3.10.3 Methods for nonlinear inference

The majority of this chapter has focused on methods to fit re-expression models with varying degrees of information encountered in geometallurgical modeling. The ideal case, where direct experimental evaluation of mixtures is possible, can be fit by direct nonlinear regression of experimental mixing results, as demonstrated for two different literature cases. Direct measurement of the mixing phenomena and the development of a corresponding theory for the reason behind nonlinear behaviour is preferable to the proposed indirect methods, but this is no always possible. Restrictions on experimental work, such as the very large volume of rock required for a Bond mill work index test, high experimental cost, limited sample and legacy data may mean that this is not always possible.

If the capacity to directly experiment with mixtures is not available, but there are multiscale samples of the metallurgical variable then the nonlinearity may be fit indirectly using the multiscale data with nonlinear regression. In

the case where multiscale data are available, but there is measurement error associated with the small scale measurements, the power law model fit will be attenuated, and biased towards an approximately linear model. Given sufficient information about the nature of the error, and provided the error is unbiased and homoscedastic, the model may be disattenuated by construction of a nomograph with Monte Carlo simulation. This is important where the power law model is to be applied to measurements without error, such as the results of a geostatistical simulation study.

Decreasing the amount of information available to the geomodeler further, multiscale bivariate inference was considered for inferring the nonlinear behaviour given a large scale metallurgical variable, small scale linear variable and the bivariate relationship between the variables at the small scale. Semiparametric Bayesian updating with nonlinear regression was proposed for this case, and demonstrated using kernel density estimation to model samples from the small scale bivariate relationship. In the course of developing the semiparametric Bayesian updating method, LU simulation with correlated bivariate probabilities for the simulation of $Z(v)$ values was tested, but found to be highly dependent on the variogram used for simulation. The result of this dependency was comparatively poor characterization of the nonlinear behaviour relative to the proposed semiparametric Bayesian updating algorithm.

Methods for inferring nonlinear behaviour in metallurgical properties given varying degrees of information are summarized in Table 3.4. As demonstrated with the Bond mill work index earlier, a weak correlation, limited samples or excessive measurement error may reduce the predictive power of the proposed technique to the point where it is ineffective. In these cases, more data are required to reasonably infer nonlinearity in the metallurgical property. Furthermore, all inferred nonlinear behaviour should be reviewed for reasonableness given the physical nature of the properties.

Table 3.4: Summary of methods to infer nonlinear behaviour in metallurgical properties given varying degrees of information.

Available Information	Method
Measurements of blends of known ore types	Direct nonlinear regression
Single-scale measurements of metallurgical properties and correlated secondary	More data are required
Multiscale measurements of metallurgical properties	Nonlinear regression
Multiscale measurements of metallurgical properties with known error	Nonlinear regression with correction for attenuation
Multiscale correlated bivariate measurements	Semiparametric Bayesian updating with nonlinear regression

Chapter 4

Downscaling composite metallurgical properties

Multiscale sampling schemes, such as the sampling of comminution properties on 30 m intervals and assays on 1 m intervals, are common mining practices. These practices are driven by testing requirements and high costs associated with sample collection and experimental work. A typical comminution testing suite requires a large amount of rock mass for testing so cannot be sampled on high frequency intervals. Assays are cheaper, faster, and more critical for early stage decision making, so are sampled at a high resolution. The practice of multiscale sampling poses a substantial problem for geometallurgical modeling; the majority of statistical and geostatistical algorithms have been developed for linear, homotopically sampled data sets at a single scale. A multiple imputation procedure is proposed in this thesis for stochastically downscaling metallurgical variables. The resulting homotopically sampled data sets can then be used in a standard multivariate geostatistical modeling workflow.

In this chapter, multiscale sampling schemes are reviewed which further emphasizes the requirement for techniques which can use multiscale data. In the context of multiscale data, multiscale parameter inference techniques are reviewed and developed in the context of multivariate, metallurgical data. The scaling of variograms, correlations, and univariate and bivariate distributions using geostatistical techniques are included.

At the core of this chapter is the development of multiscale direct sequential simulation with intrinsic cokriging for stochastic downscaling of large samples.

As direct sequential simulation is used, the downscaled samples are exact, but there are challenges with histogram reproduction. The property of exactness is a function of correctly implemented sequential simulation (Gómez-Hernández et al., 2005), and the exact reproduction in original units of direct sequential simulation. Techniques for correctly reproducing the histogram are reviewed in this context. Implementation details of the algorithm, including dynamic resimulation to respect variable constraints, such as positivity, are developed. The algorithm is demonstrated, and alternative approaches to downscaling and multiscale data integration are reviewed.

4.1 Multiscale sampling schemes

In addition to the expense of many metallurgical tests limiting the number of samples tested, large masses of rock are typically required for testing. Sample requirements for rotary breakage tests to quantify the autogeneous and semi-autogeneous breakage parameters A and b require a minimum of 100 kg of sample from core (JKTech, 2015). Ball mill tests require a minimum of 10 kg of sample. Flotation tests require between 1 and 10 kg depending on the testing mode. These correspond to very large core lengths and are substantially more than the nominal 0.1 kg of sample required for most assays.

Multiscale sampling schemes are not unique to metallurgical sampling; seismic data is often collected in addition to petrophysical measurements on wells. Three-dimensional seismic is at a substantially lower resolution relative to high resolution lithology, porosity and permeability measurements. Micro-measurements, such as microresistivity, as collected in oilsands deposits, add an additional scale. Although not the focus of this thesis, the downscaling techniques developed could be applied in a petroleum context as well.

The multiple data scales of interest in a typical metallurgical sampling scheme are few; measurements are likely to be collected on a minimum of 10 cm of core in the case of high resolution assays and up to 30 m of core for comminution tests. For modeling of open pit deposits, standard practice is to composite high resolution measurements to nominal bench, or half-

bench, scales of 10–15 m. The changes of support for the mining applications of interest are therefore much less than comparable petroleum applications where seismic measurements may only be available on a nominal resolution of 10x10x10 m^3.

For spatial modeling using standard geostatistical algorithms, the large scale samples would ideally be downscaled to the modeling scale. This permits usage of the full suite of multivariate techniques requiring homotopic, single-scale sampling. With the addition of a nonlinear transform as described in chapter 3 linear techniques could be used as well. Downscaling the large scale measurements require parameters to be inferred at the small scale. Multiscale parameter inference is discussed in this context.

4.2 Multiscale parameter inference

Statistics inferred for all regionalized variables of interest in a geometallurgical study, including variograms, correlations and distributions, are relevant for a single scale of interest. If parameters may not be directly inferred on a different scale of interest, such as by upscaling the data and re-calculating the statistics, then parameter scaling is required. Scaling of the variogram, correlations, and univariate and bivariate histograms are discussed here in the context of downscaling.

The scaling of parameters is demonstrated and evaluated using drilling from an oil sands mining operation in Alberta, Canada. Measurements of bitumen grade and fines content (<45 microns) are available on 3 m intervals over a 24 square kilometer area. Bitumen grades were arithmetically upscaled to 9 m composites, and the correlated fines content were left at 3 m intervals. The degree of scale change (3x in this case) is therefore comparable to downscaling a metallurgical sample on 30 m of core to 10–15 m bench scale composites. Bitumen grade is an ideal candidate for evaluating parameter downscaling (and later in this chapter, data downscaling) as it averages linearly, and measurements are available on a fine scale so that downscaled statistics may be evaluated against known statistics. This is not a typical use case of parame-

ter downscaling; downscaling is not required if small scale measurements are available, but is illustrative for the purposes of this thesis. A plan view of the drill collar locations is shown in Figure 4.1.

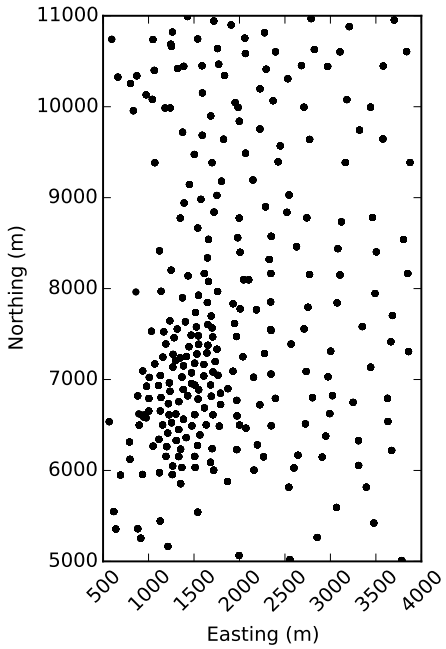

Figure 4.1: Plan view of collar locations for oil sands drilling. All drilling is vertical to an average depth of 80 m below the surface.

4.2.1 Variogram scaling

Variogram scaling is required for spatial inference, and for probability distribution and correlation scaling. Consider a variogram model of the random variable $Z_k(\mathbf{u}; v_1)$ at scale v_1 adopting a linear model of regionalization:

$$\gamma_k(\mathbf{h}; v_1) = C_k^0(v_1) + \sum_{i=1}^{nst} C_k^i(v_1)\Gamma^i(v_1)$$

where C^0 is the nugget effect, C^i are variance contributions, nst is the number of variogram structures and Γ^i are the variogram structures defined by a rotation (3 angles assuming a 3D model) and variogram ranges (3 orthogonal ranges $a_j^i, j = 1, \ldots, 3$). The variogram is to be scaled to scale v_2. For variogram scaling, the widely referenced scaling laws (Chilès and Delfiner, 2009; Oz et al., 2002), or an alternative approach such as the direct scaling method

of Babak et al. (2013) may be used. The variogram scaling laws are used here as the scale changes, and therefore the degree of variogram shape changes, are not very large. Using the variogram scaling laws, variance contributions are scaled from v_1 to v_2 as:

$$C_k^0(v_2) = C_k^0(v_1)\frac{\|v_1\|}{\|v_2\|}$$

$$C_k^i(v_2) = C_k^i(v_1)\frac{1-\bar{\Gamma}^i(v_2,v_2)}{1-\bar{\Gamma}^i(v_1,v_1)}, i = 1, ..., nst$$

$$\bar{\Gamma}^i(v,v) = \frac{1}{\|v\|^2}\int_{v_1}\int_{v_2}\Gamma_p^i(v_1-v_2)dv_1dv_2 \approx \bar{\Gamma}^i(v,v) = \frac{1}{n_{disc}^2}\sum_{i=1}^{n_{disc}}\sum_{j=1}^{n_{disc}}\Gamma_p^i(\mathbf{u}_i-\mathbf{u}_j)$$

where $\|x\|$ denotes the magnitude of the volume x, $|x|$ is the parallel length of the variogram scale along the range direction, and p denotes the pseudo-point scale variogram. This variogram is typically the variogram at the smaller of the two scales, although a nominal point scale may be used instead provided a correct discretization is used. The pseudo-point scale ranges are calculated using the variogram range scaling law:

$$a_j^i(v_2) = a_j^i(v_1) - |v_1| + |v_2|, i = 1, ..., nst, j = 1, \ldots, 3$$

The average variograms are calculated using a discrete approximation with n_{disc} points. The sill (variance), of the scaled variogram using the scaled variance contributions is then:

$$\sigma_k^2(v_2) = C_k^0(v_2) + \sum_{i=1}^{nst} C_k^i(v_2)$$

These variogram scaling laws assume constant shape, which holds for relatively small scale changes, but there will be substantial deviations for very large scale changes (Babak et al., 2013). Scale changes relevant for metallurgical samples, such as the downscaling of a 30 m comminution sample to 15 m bench-scale samples are comparatively small and the assumption of constant variogram shape is reasonable in most cases. If these criteria are not met, then the direct scaling approach of Babak et al. (2013) may be employed to account for the change of variogram shape.

For the oil sands example, variograms of bitumen grade at 9 m were calculated and modeled. As expected for this stratigraphic deposit, the horizontal

directions are much more continuous relative to the layered vertical directions. In addition, substantial zonal anisotropy is observed for an azimuth of 45 degrees. Experimental variograms and modeled variograms are shown in Figure 4.2.

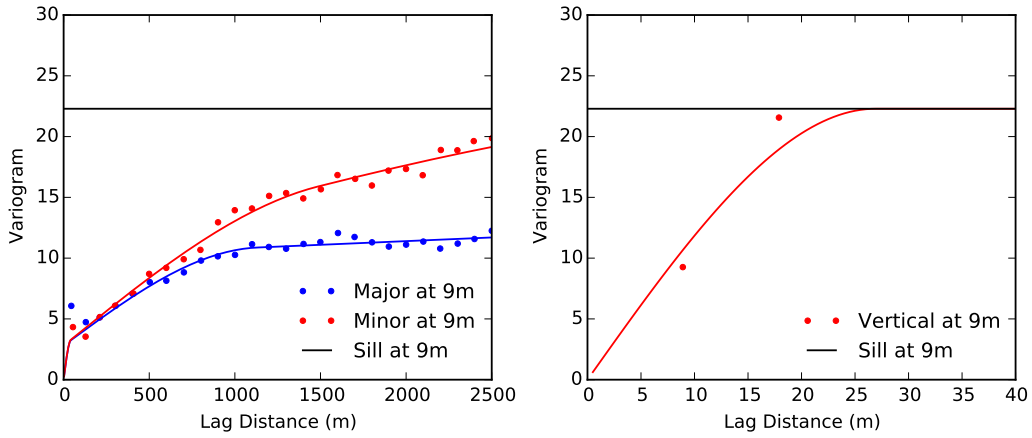

Figure 4.2: Modeled bitumen variograms at the 9 m scale. The major direction of continuity is 45 degrees East of North, minor is 135 degrees East of North and the vertical direction is perfectly vertical for this deposit.

The variogram model is expressed using variogram shorthand where *Sph* denotes the spherical variogram model, the major direction is oriented at an azimuth of 45 degrees, the minor at an azimuth of 135 degrees and the vertical direction is exactly vertical with a 90 degree dip.

$$\gamma_{Bit}(\mathbf{h}; 9\text{m}) = 2.81 Sph_{\substack{a_{major}=41.7 \\ a_{minor}=41.7 \\ a_{vert}=26.9}} + 7.38 Sph_{\substack{a_{major}=1132.0 \\ a_{minor}=1486.4 \\ a_{vert}=26.9}} + 12.1 Sph_{\substack{a_{major}=29945.8 \\ a_{minor}=4558.7 \\ a_{vert}=26.9}}$$

Applying the variogram scaling laws, the scaled variogram is now:

$$\gamma_{Bit}(\mathbf{h}; 3\text{m}) = 3.41 Sph_{\substack{a_{major}=41.7 \\ a_{minor}=41.7 \\ a_{vert}=20.9}} + 8.96 Sph_{\substack{a_{major}=1132.0 \\ a_{minor}=1486.4 \\ a_{vert}=20.9}} + 14.7 Sph_{\substack{a_{major}=29945.8 \\ a_{minor}=4558.7 \\ a_{vert}=20.9}}$$

This scaled variogram is plotted along with the true experimental variogram points and sill in Figure 4.3. The downscaled variogram is in good agreement with the experimental variogram points in the horizontal direction, but deviates from the vertical variogram slightly. The impact of this deviation will be examined when downscaling the bitumen grades.

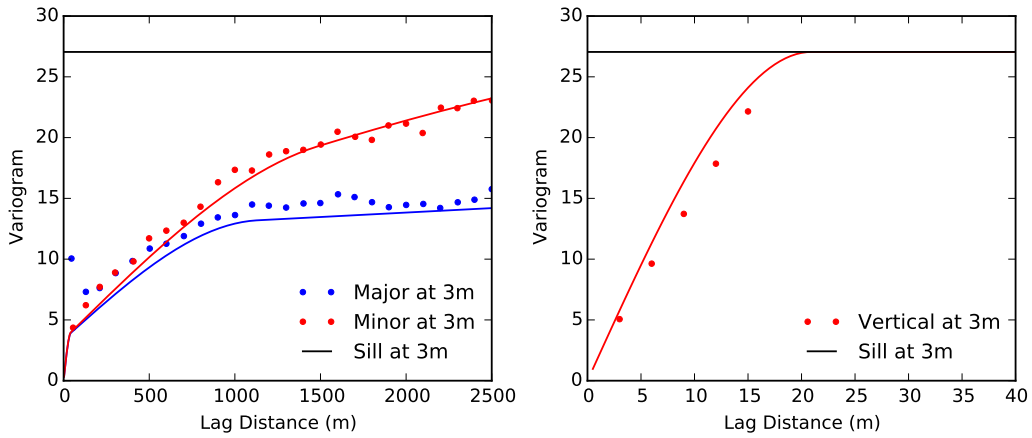

Figure 4.3: Downscaled bitumen variograms at the 3 m scale. Known variogram points at the 3 m scale are shown for comparison and evaluation of the downscaling procedure.

4.2.2 Correlation scaling

As the correlation is the standardized cross covariance at a lag distance of 0, applying the variogram scaling laws shows that the correlation must change with scale (Oz and Deutsch, 2002). The variogram scaling laws are applied where the covariance of interest is $C_{k,k'}$ instead of C_k. The covariance at lag distance 0 is the sill of the scaled cross variogram. The correlation coefficient of variable k to k' at scale v_2 is therefore:

$$\rho_{k,k'}(v_2) = \frac{C^0_{k,k'}(v_2) + \sum_{i=1}^{nst} C^i_{k,k'}(v_2)}{\sigma_k(v_2)\sigma_{k'}(v_2)}$$

Continuing with the oil sands deposit, the correlation between fines and bitumen is downscaled. The calculation of the cross variogram for downscaling is problematic from a practical perspective; there are three composite fines measurements within the 9 m large scale bitumen measurements being evaluated. In addition to software limitations, there is the practical challenge of difficult to model directions if variograms are calculated using the offset pairs directly. This challenge is addressed by upscaling the fines measurements at the 3 m scale to the 9 m scale, calculating the cross variogram and downscaling. The upscaling procedure is exact and permits the calculation of variograms at the scale of interest. Calculated and modeled cross variograms are shown in Figure

4.4 at the 9 m scale.

$$\gamma_{Bit,Fines}(\mathbf{h};9\text{m}) = -17.0 Sph_{\substack{a_{major}=299.3\\a_{minor}=299.3\\a_{vert}=26.7}} - 33.0 Sph_{\substack{a_{major}=2500.0\\a_{minor}=2173.0\\a_{vert}=26.7}} - 26.6 Sph_{\substack{a_{major}=40000.0\\a_{minor}=7425.1\\a_{vert}=26.7}}$$

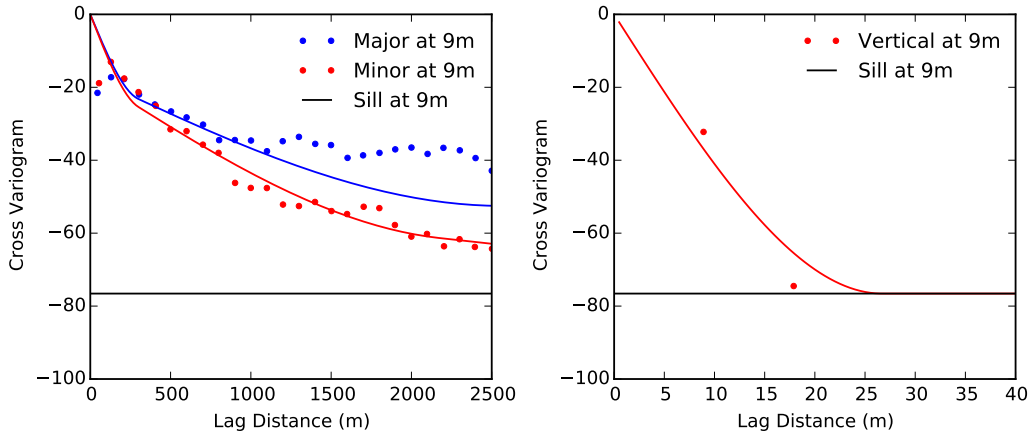

Figure 4.4: Modeled cross bitumen-fines variograms at the 9 m scale.

Using the same variogram scaling laws, the cross bitumen-fines variograms at 9 m may be downscaled (Figure 4.5). Note that in addition to the change in variance contributions, the vertical ranges also change.

$$\gamma_{Bit,Fines}(\mathbf{h};3\text{m}) = -20.7 Sph_{\substack{a_{major}=299.3\\a_{minor}=299.3\\a_{vert}=20.7}} - 40.1 Sph_{\substack{a_{major}=2500.0\\a_{minor}=2173.0\\a_{vert}=20.7}} - 32.3 Sph_{\substack{a_{major}=40000.0\\a_{minor}=7425.1\\a_{vert}=20.7}}$$

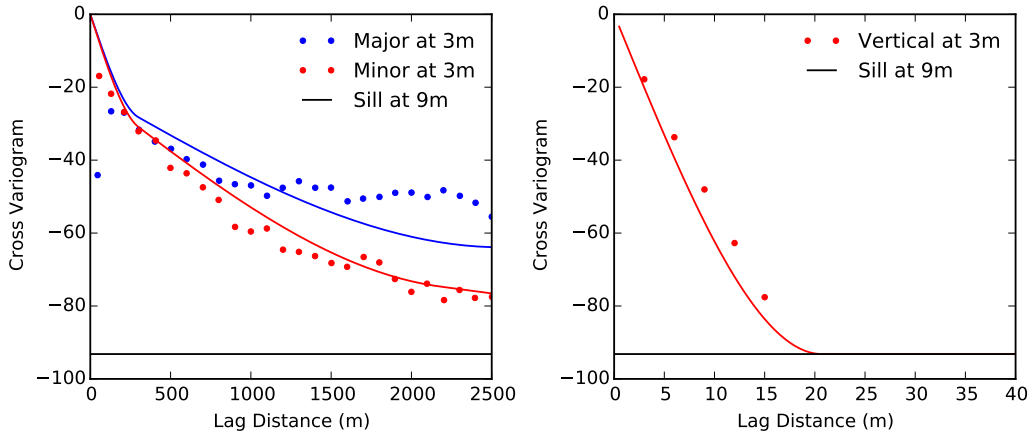

Figure 4.5: Downscaled cross bitumen-fines variograms at the 3 m scale.

The correlation coefficient is therefore calculated:

$$\rho_{Bit,Fines}(9\text{m}) = \frac{C^0_{Bit,Fines}(9\text{m}) + \sum_{i=1}^{nst} C^i_{Bit,Fines}(9\text{m})}{\sigma_{Bit}(9\text{m})\sigma_{Fines}(9\text{m})} = -0.842$$

4.2.3 Univariate distribution scaling

Using the variance (sill) of the downscaled variogram, the histogram may be downscaled. The three primary algorithms for histogram scaling are the affine (linear) correction, indirect lognormal and discrete Gaussian model. Of these algorithms, the discrete Gaussian model is the most flexible for a range of distribution shapes. Under an isofactorial model with the assumption of stationarity, the discrete Gaussian model may be used to transform the histogram from one scale to another (Chilès and Delfiner, 2009; Wackernagel, 2003). Primarily, the diffusion model is adopted for the transform; however, the mosaic and barycentric are alternatives (Machuca-Mory et al., 2008). Following the derivation of Machuca-Mory et al., consider the normal score transformation (Gaussian anamorphosis) of the variable to be downscaled at the original scale:

$$z_k\left(\mathbf{u}_i; v_1\right) = \varphi\left(y_k\left(\mathbf{u}_i; v_1\right); v_1\right) = F\left(z_k\left(\mathbf{u}_i; v_1\right)\right)^{-1} \circ G\left(y_k\left(\mathbf{u}_i; v_1\right)\right), i = 1, ..., n_k$$

where y denotes the Gaussian transformed variable, G is the standard normal distribution and F is the cumulative distribution function of the original variable. The anamorphosis is fit with orthonomal Hermite polynomials H such that:

$$H_0(y) = 1, H_1(y) = -y, H_{q+1}(y) = -\frac{y}{\sqrt{q+1}} H_p(y) - \sqrt{\frac{q}{q+1}} H_{q-1}(y)$$

$$\varphi(y; v_1) = \sum_{q=0}^{Q \to \infty} \varphi_q(v_1) H_q(y),\ \varphi_q(v_1) = E\left\{\varphi\left(Y(x; v_1)\right) H_q\left(Y(x)\right)\right\}$$

The new distribution is calculated under the assumption of either a diffusion, mosaic or barycentric model. Typically for upscaling the diffusion model is adopted. A limited number of Hermite polynomials is used such as $Q = 100$. The diffusion model assumes a Gaussian process for the variable of interest (Machuca-Mory et al., 2008). Under the diffusion model, the coefficients are adjusted:

$$\text{Diffusion model: } \varphi(y; v_2) = \sum_{q=0}^{Q} \varphi_q(v_1) r^q H_q(y),\ \sigma_k^2(v_2) = \sum_{q=0}^{Q} \varphi_q(v_1)^2 r^{2q}$$

where the constant r is the change of support coefficient solved using the second term.

Under a mosaic model, the variable is assumed to follow a dead leaves model, (Bordenave et al., 2006; Machuca-Mory et al., 2008). That is, the variable is composed of equal size domains which are sequentially placed in a Poisson process. Under the mosaic model:

$$\text{Mosaic model: } \varphi(y;v_2) = \sum_{q=0}^{Q} \varphi_q(v_1) r H_q(y),\ \sigma_k^2(v_2) = \sum_{q=0}^{Q} \varphi_q(v_1)^2 r^2$$

The barycentric model, used by Machuca-Mory et al. (2008), is a linear combination of the mosaic and diffusion models. The ratio of the variogram to the madogram (which is the expected value of squared absolute values) is used to determine the linear weight assigned to each model.

Alternative models to the discrete Gaussian model for histogram scaling include the affine (linear) and indirect lognormal corrections. The affine correction is unreasonable as the histogram shape does not change according to central limit theorem. The indirect lognormal correction is the correct change of shape model for a lognormal distribution, but requires iterative correction to avoid biasing the mean in all other cases (Zagayevskiy and Deutsch, 2011). For highly skewed variables which are approximately lognormal, the indirect lognormal correction is reasonable.

As the discrete Gaussian model is flexible for many variogram shapes, it is adopted in this thesis for most downscaling applications. The choice of a diffusion, mosaic or barycentric model is required for the discrete Gaussian model. The diffusion model is the most widely used for upscaling, and the only model implemented in most software packages (commercial or academic). When applied to downscaling histograms, the power q, which is the order of the Hermite polynomial in the diffusion model, results in extreme numerical instability where Q is large. This instability is observed when the mosaic and diffusion models are used to downscale the bitumen histogram in the oilsands example (Figure 4.6).

The large instability in the diffusion model results in a downscaled histogram which is not a licit cumulative distribution. This instability supports the use of the mosaic model for downscaling. The instability is intuitively understood when the weight (r) applied to higher order Hermite polynomials

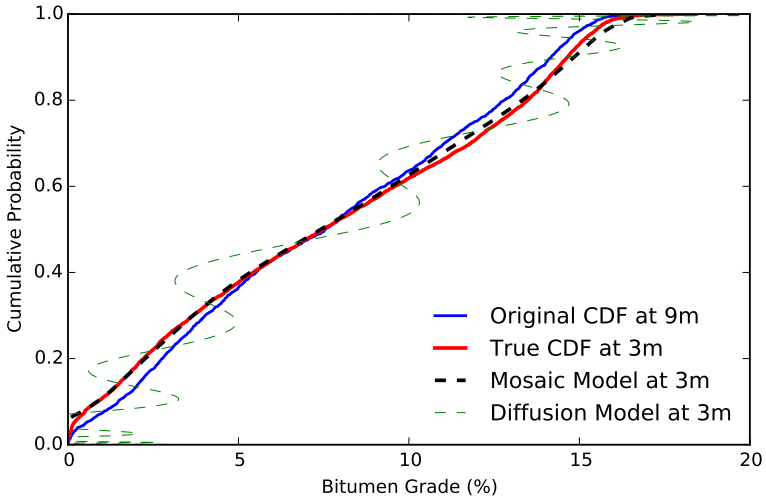

Figure 4.6: Downscaled bitumen histograms using the discrete Gaussian model with the mosaic and diffusion models at the 3 m scale.

(i.e.: for q large) is greater than the weight applied to low order polynomials. Furthermore, the relationships between the polynomials are complex which magnifies small numerical errors when computing the coefficients.

The downscaled mosaic model histogram is close to the true histogram, although is unbounded; bitumen grades below 0% have a non-zero probability under the downscaled histogram. This requires adjusting the downscaled histogram to meet the constraints. One approach for this adjustment would be to reset all negative grades to zero at the cost of a small bias in the resulting relationship. However, for highly skewed variables near a constraint, such as bitumen grade, an alternative correction could be applied.

As the bitumen grade is approximately lognormal, the indirect lognormal correction may be applied as an alternative to the discrete Gaussian model with the diffusion correction. The downscaled histogram is extremely close to the true histogram, and does not require adjustment to account for negative grades as the distribution is bounded by zero (Figure 4.7).

For downscaling an arbitrary distribution, the mosaic model is recommended due to its numerical stability and flexibility. For skewed, bounded distributions such as bitumen grade, the indirect lognormal correction may be applied for downscaling. These corrections apply to scaling the univariate

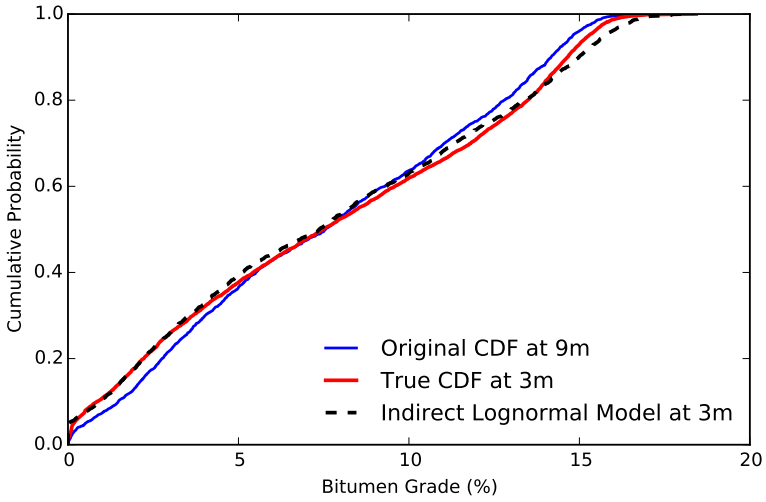

Figure 4.7: Downscaled bivariate probability distribution of fines and bitumen grade at the 3 m scale with pairs of bitumen and fines grades (at the original 3 m scale) plotted.

histogram, but no analogous corrections are available for bivariate and higher order distribution.

4.2.4 Bivariate distribution scaling

Under an intrinsic assumption, scaling of the bivariate distribution is not explicitly required for downscaling a metallurgical variable of interest. Only the univariate distribution must be downscaled in these cases. However, approaches such as the semiparametric Bayesian updating approach for nonlinear inference proposed in chapter 3 require a complete model of the bivariate distribution at the scale of interest. In these cases, scaling the bivariate distribution is required. Using the multivariate Gaussian distribution would be inappropriate in these cases; the distributions are complex and often incorporate constraints such as non-negativity. Scaling the bivariate (or a higher order) distribution is required in these cases.

Leuangthong et al. (2005) suggest an iterative approach to scale a multivariate distribution to exactly match the univariate distributions. The iterative procedure suggested by Leuangthong et al. is proposed in the context of direct sequential simulation using a cokriging system for integrating block data. In this thesis, this procedure for multivariate scaling is adapted

for multivariate distribution scaling. Consider scaling a bivariate distribution of variables k and k' with a probability density function $f_{k,k'}(z_k, z_{k'}; v_1)$ to the downscaled distribution $f_{k,k'}(z_k, z_{k'}; v_2)$. Using the univariate distribution scaling approaches described earlier, or known measurements, the univariate marginal distributions at scale v_2 are known: $f_k(z_k; v_2)$ and $f_{k'}(z_{k'}; v_2)$. The downscaling is iterative for steps $p = 1, \ldots, P$ which are carried out until the marginals of the downscaled bivariate distribution $f^{(p)}_{k,k'}(z_k, z_{k'}; v_2)$ match the known downscaled marginal distributions within a set tolerance. The iterative procedure adapted from Leuangthong et al.:

1. The starting condition ($p = 0$) for the iterative procedure is to set the downscaled bivariate distribution to the original distribution: $f^{(0)}_{k,k'}(z_k, z_{k'}; v_2) = f_{k,k'}(z_k, z_{k'}; v_1)$

2. For $p = 1, \ldots, P$ iterations, where P is determined by a tolerance on the difference between the marginal distributions:

   (a) Update the downscaled bivariate distribution using the product of the ratios of the marginal distributions:
   $f^{(p)}_{k,k'}(z_k, z_{k'}; v_2) = f^{(p-1)}_{k,k'}(z_k, z_{k'}; v_2) \prod_{k=1}^{K} \frac{f_k(z_k; v_2)}{f^{p-1}_k(z_k; v_2)}$

   (b) Compare the marginal distributions of the updated distribution for tolerance: **if** $f^p_k(z_k; v_2) - f_k(z_k; v_2) < tol \quad \forall \quad K$ **then finished**

Using the downscaled bitumen grade distribution, and known small scale fines distribution the bivariate distribution is downscaled. This is not required for the intrinsic approach to downscaling proposed in this thesis, but would be required for the nonlinear inference algorithm proposed in chapter 3. Consider the univariate probability density functions for bitumen and fines (Figure 4.8) at the 9 m and 3 m scale. The 3 m bitumen measurements have been downscaled with a lognormal distribution, and the fines measurements at 3 m are considered known.

The bivariate distribution at the 9 m scale fit is with kernel density estimation using a Gaussian kernel, and the previously described rules for bandwidth

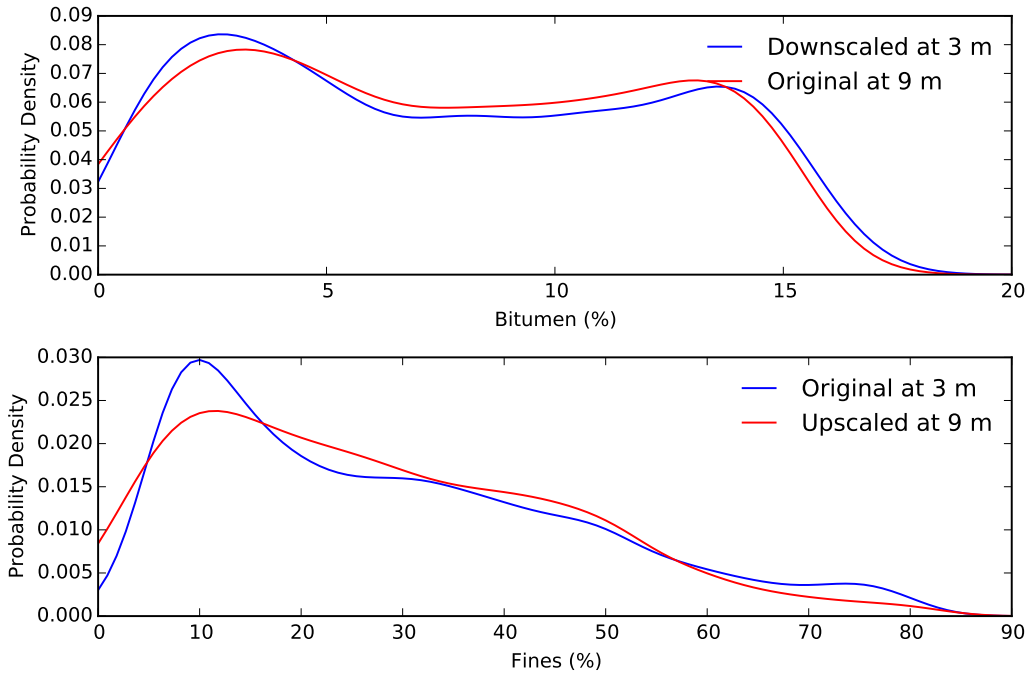

Figure 4.8: Downscaled bitumen histograms using the indirect lognormal correction at the 3 m scale.

selection. Values are calculated on a discrete grid at a 100x100 discretization. The bivariate probability distribution at the 9 m scale is plotted in Figure 4.9.

Using the adapted procedure for multivariate distribution scaling, and the target univariate histograms at 3 m, the bivariate distribution is downscaled to the 3 m scale. This downscaled bivariate distribution is plotted in Figure 4.10. Marginal distributions calculated from this downscaled bivariate distribution exactly match the target distribution (Figure 4.11).

To check bivariate distribution at the small scale, the known 3 m scale pairs of bitumen and fines grades are scatter plotted on to the bivariate density distribution in Figure 4.12. The points show good agreement with the downscaled distribution.

The approach proposed by Leuangthong et al. (2005) and adapted here for multivariate distribution scaling is severely limited by the curse of dimensionality. This nonparameteric approach to distribution scaling requires discretizing the distributions, and even at a discretization level of 100 steps per variable quickly becomes unmanageable. This is true for most nonparamet-

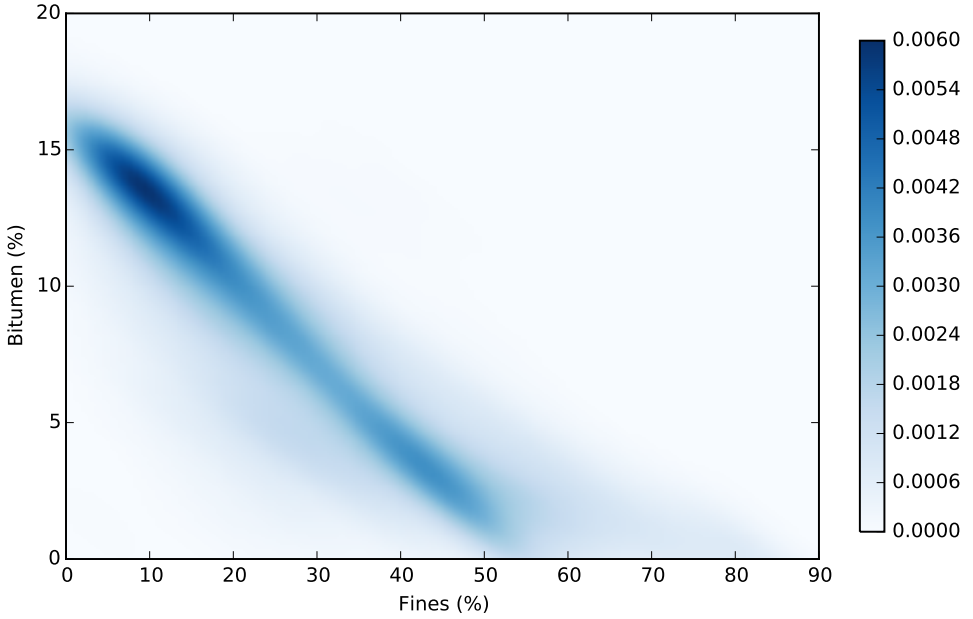

Figure 4.9: Bivariate probability distribution of fines and bitumen grade at the 9 m scale fit with kernel density estimation using a Gaussian kernel.

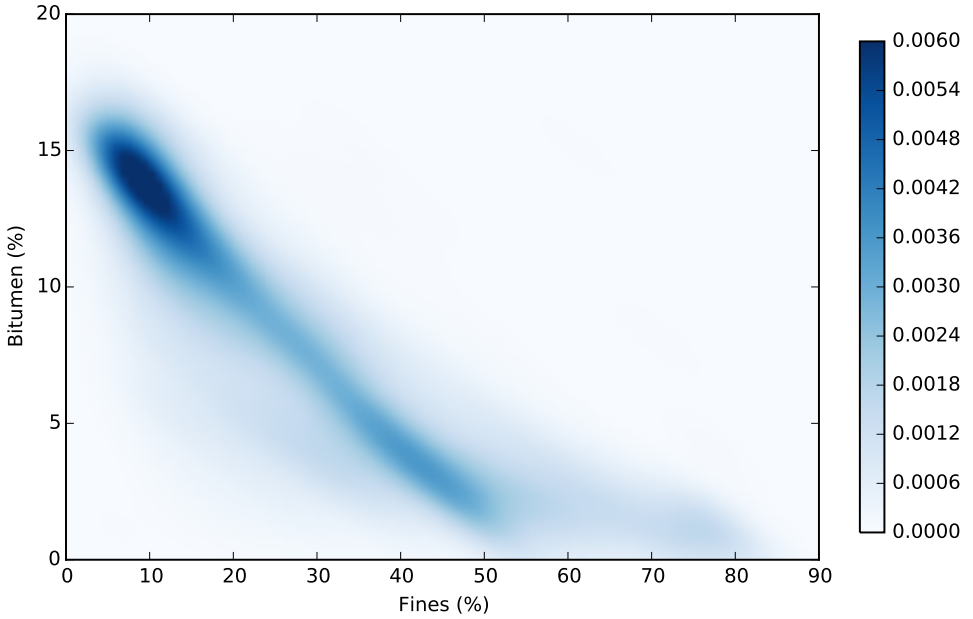

Figure 4.10: Downscaled bivariate probability distribution of fines and bitumen grade at the 3 m scale.

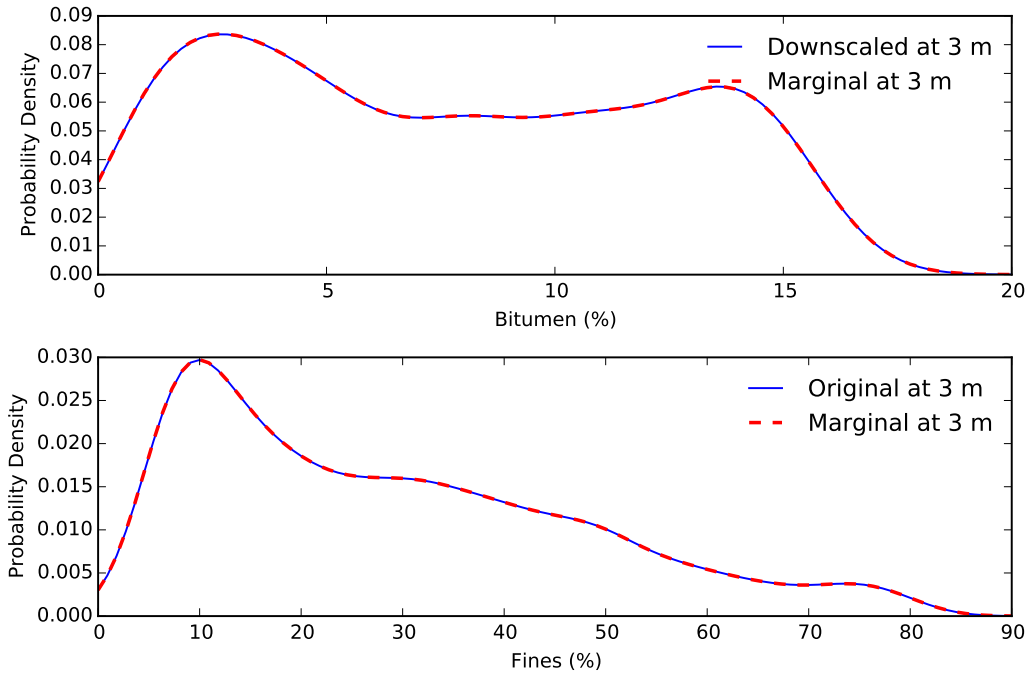

Figure 4.11: Marginal distributions calculated from the bivariate probability distribution of fines and bitumen grade which exactly match the targeted downscaled univariate distributions.

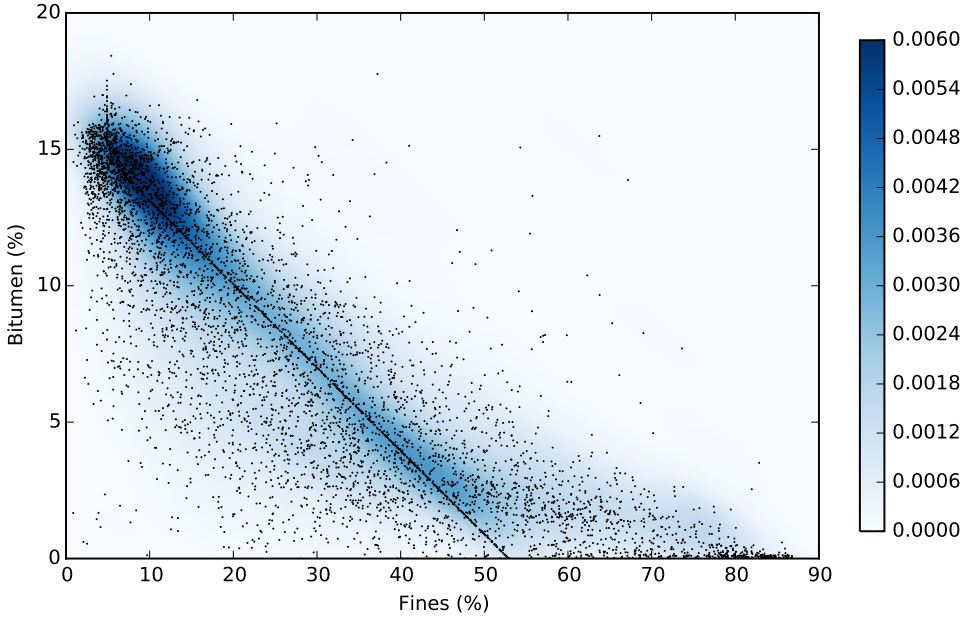

Figure 4.12: Downscaled bitumen histograms using the indirect lognormal correction at the 3 m scale.

ric techniques including kernel density estimation which is used to infer the multivariate distribution for discretization in the first step. This technique is therefore only applicable for few dimensions due to the curse of dimensionality.

4.3 Multiscale direct intrinsic sequential simulation

Downscaling of statistical parameters including the variogram, correlations, and univariate and bivariate distributions have been discussed with the goal of downscaling metallurgical measurements. Downscaling of large scale metallurgical measurements is proposed as the vast majority of (geo)statistical techniques require data at a common measurement scale in addition to requirements for stationarity, and commonly linearity. Upscaling all measurements to the largest data scale and modeling at a very large scale, such as 30 m blocks, is unreasonable when the scales of interest for mining equipment selectivity and resource classification are on a smaller scale. Downscaling large scale measurements to a smaller scale of interest would permit the usage of statistical techniques which require data at a common scale and high resolution modeling.

Philosophically, the downscaling approach proposed here adopts the multiple realization paradigm using multiple imputation in which there is no single true data set, but multiple data sets which have the correct statistical properties. This naturally fits with the simulation approach for geostatistical modeling in which multiple realizations of the deposit are constructed and analyzed. This downscaling procedure cannot be completed arbitrarily; as shown with the statistical scaling of parameters and observed during compositing, the properties will change with scale. Using the bitumen grade as an example, it would be innappropriate to arbitrarily downscale a 9 m measurement of bitumen grade of 9% to 3 m measurements of {9%,9%,9%} or {0%,9%,18%} without a statistical basis. Furthermore, correlated fine scale measurements could be used to improve the downscaling procedure.

With the statistical property scaling laws in mind, five required properties of a downscaling algorithm are proposed. These properties are:

1. Realized data values should have the correct spatial structure. The variogram of small scale values should accurately reproduce the downscaled variogram.

2. Realized data values should reproduce the correct small scale distribution. The histogram of downscaled values should accurately reproduce the downscaled probability distribution.

3. When upscaled, realized data values should exactly reproduce the large scale data. That is, the average of the small scale realized data values should exactly equal the large scale data value.

4. Realized data values should satisfy hard constraints, such as non-negativity.

5. Realized data values should have the correct multivariate structure. Downscaled correlations should be reproduced by the realized data.

These requirements motivate the use of direct sequential simulation, which by construction reproduces the variogram, is exact, may use multivariate data and with the appropriate selection of local histograms reproduces the univariate distribution. Using a Gaussian approach, such as sequential Gaussian simulation, would correctly reproduce the spatial structure and univariate distribution, but simulated small values would not average exactly to the true large scale values.

4.3.1 Direct sequential simulation

Direct sequential simulation is not typically used in geostatistical modeling. The reason for this is primarily due to problems inferring the local histogram which should be used to correctly reproduce the global histogram, and more challenging integration of multivariate data compared to Gaussian techniques. Compared to Gaussian algorithms, the primary advantage of direct sequential simulation is exactitude when using block data in original units. This exactitude is a property of correctly implemented sequential simulation (Gómez-Hernández et al., 2005), however is predicated on the correct choice of a search neighbourhood. If a search neighbourhood is chosen incorrectly, and simulated values

within the block data are excluded then exactness is not guaranteed. For typical modeling applications, particularly in a mining context, this is not an advantage as no block data are available. For the specific purpose of stochastically downscaling data where large scale measurements are available, the use of direct sequential simulation is supported. Here, two key areas of direct sequential simulation critical to its use as a downscaling technique are discussed: exactitude of simulation and histogram reproduction.

4.3.2 Exactitude of simulation

As shown by Ren (2007), the direct kriging or direct sequential simulation of discrete composite volumes within a large volume using the large scale value will exactly reproduce the large value. For a detailed derivation, the interested reader is referred to Ren; however, as the exactitude of direct sequential simulation is central to the downscaling proposal in this thesis, a small demonstration follows. Consider the direct sequential simulation of three values of type $k = p$ at locations $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ with scale v_c within a large scale composite centered at $\mathbf{u}_2$ with scale v_m. This situation is sketched in Figure 4.13.

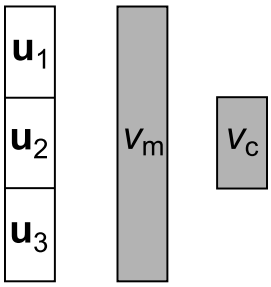

Figure 4.13: Sketch of a large scale composite composed of three smaller volumes.

The requirement for exactitude is that the simulated values at scale v_c average exactly to the large scale value at v_m, that is:

$$z_p(\mathbf{u}_2; v_m) = \frac{1}{3} \sum_{i=1}^{3} z_p^*(\mathbf{u}_i; v_c) \tag{4.1}$$

Consider that two of the three values, $z_p^*(\mathbf{u}_1; v_c)$ and $z_p^*(\mathbf{u}_2; v_c)$ have already been simulated. Therefore, the requirement for exactitude may then be ex-

pressed as:

$$z_p^*(\mathbf{u}_3; v_c) = 3z_p(\mathbf{u}_2; v_m) - z_p^*(\mathbf{u}_1; v_c) - z_p^*(\mathbf{u}_2; v_c) \tag{4.2}$$

This requirement is fullfilled by direct sequential simulation. The linear system of equations for the final location $\mathbf{u}_3$ are:

$$\begin{aligned}
&[C_{1,1}(v_c) + C_{1,2}(v_c) + \overline{C}_{1,2}(v_c, v_m)]\lambda_1 = C_{1,3}(v_c) \\
&[C_{2,1}(v_c) + C_{2,2}(v_c) + \overline{C}_{2,2}(v_c, v_m)]\lambda_2 = C_{2,3}(v_c) \\
&[\overline{C}_{2,1}(v_m, v_c) + \overline{C}_{2,2}(v_m, v_c) + \overline{C}_{2,2}(v_m)]\lambda_3 = \overline{C}_{2,3}(v_m, v_c)
\end{aligned} \tag{4.3}$$

where the average covariances are calculated:

$$\begin{aligned}
\overline{C}_{i,2}(v_c, v_m) &= \overline{C}_{2,i}(v_m, v_c) = \frac{1}{3}\sum_{\alpha=1}^{3} C_{\alpha,i}(v_c) \\
\overline{C}_{2,2}(v_m) &= \frac{1}{9}\sum_{\alpha=1}^{3}\sum_{\beta=1}^{3} C_{\alpha,\beta}(v_c)
\end{aligned} \tag{4.4}$$

Substituting Equation 4.4 into 4.3 and solving yields weights of $\lambda_1 = -1$, $\lambda_2 = -1$, and $\lambda_3 = 3$ which results in a simple kriging mean given by Equation 4.2, and a simple kriging variance of zero. Therefore, direct sequential simulation of the small scale values in this example is exact. This solution generalizes, and does not require contiguous volumes, only that the subvolumes exactly compose the large scale volume. Note also that the sequential aspect of this simulation is central to this property. Simultaneous simulation approaches, such as Fourier methods used in Gaussian simulation cannot be used with this approach.

4.3.3 Histogram reproduction

The challenge of histogram reproduction with direct sequential simulation has been researched extensively and a number of algorithms proposed (Caers, 2000; Oz et al., 2003; Robertson et al., 2006; Soares, 2001). Caers proposed the targetting of the global distribution by minimizing the deviation from the local conditional probabilities. The objective function can take the form of an absolute value (that is, L1 regularization) or squared differenced (L2 regularization). In addition to the proposal of an optimization approach for targetting

the global distribution, Caers emphasizes the requirement for consistency between the variogram and histogram. The example provided is that a variogram which is hyper-continuous near the origin (such as the Gaussian variogram) directly conflicts with a long-tailed histogram as isolated zones of high values are not consistent.

Soares (2001) uses a local portion of the global distribution near to the simple kriging estimate to establish the distribution. A Gaussian transformation to define the sampling intervals can be used with this approach; however, bias correction is required as the Gaussian transform is nonlinear. The related proposal of Oz and Deutsch (2003) applies a Gaussian transform with a lookup table. Shapes of the conditional distributions are established using the global normal score transform to Gaussian values resulting in a table of unique conditional distributions corresponding to discretized conditional mean and variance values.

Robertson et al. (2006) compared the proposed histogram reproduction algorithms of Caers (2000; L1 and L2 regularizers) and Oz (2003). In the case study conducted, Robertson found little difference in the results, with similar histogram and variogram reproduction across all correction algorithms. Using a histogram reproduction correction is required for direct sequential simulation; however, the choice of correction algorithm was not the most critical aspect of the simulation approaches. In this thesis, the approach of Oz and Deutsch (2003) is adopted as it is computationally efficient and simple to implement numerically.

4.3.4 Theoretical development

Consider a large scale metallurgical variable of interest, Z_p, such as a work index, that we are interested in downscaling. The random variable is regionalized as $Z_p(\mathbf{u}; v_m)$ within a deemed stationary domain. Measurements are available at $\mathbf{u}_i, i = 1, \ldots, n_p$. Regionalized secondary variables, such as metal assays, are also available at a composite scale as $Z_k(\mathbf{u}; v_c), k = k', \ldots, K$. The goal is to downscale measurements of the variable of interest at the metallurgical sample scale v_m to the composite scale v_c conditional to the secondary

values, that is, $l = 1, \ldots, L$ realizations are desired which satisfy the previously described constraints:

$$\{z_p^l(\mathbf{u}_j; v_c) | z_p(\mathbf{u}_i; v_m), z_k(\mathbf{u}_j; v_c)\}, i = 1, \ldots, n_k, j = 1, \ldots, n_{k'}, \\ k = k', \ldots, K, l = 1, \ldots, L$$

Adopting the local conditional distribution lookup table approach of Oz et al. (2003), the direct sequential simulation algorithm proposed in this thesis is first summarized:

1. Calculate a set of local conditional distribution lookup tables using the method of Oz et al. (2003)

2. For each downscaled data realization, $l = 1, \ldots, L$:

   (a) Queue the simulation locations using a random ordering, $\mathbf{u}_j, j = 1, \ldots, n_k$

   (b) Visit the next location in the queue, u_{sim}

   (c) Search for nearby data and previously simulated locations. These are data events e_j enumerated $j = 1, \ldots, n_{krige}$

   (d) Set up and solve the simple intrinsic cokriging system of equations $\mathbf{C}\lambda = \mathbf{R}$

   (e) Calculate the standardized estimate and variance, y_p^* and σ_p^2

   (f) Unstandardize the estimate and variance and lookup the corresponding local conditional distribution shape

   (g) Draw a simulated small scale value using Monte Carlo simulation from the local conditional distribution, z_{sim}

   (h) If this value violates a hard boundary constraint, such as a negative value, then reject the simulated values and add these locations back into the simulation queue

   (i) Standardize the simulated small scale value, y_{sim}

   (j) Move to the next location in the queue until we have reached the end

(k) Unstandardize all simulated values and move to the next data realization

For the direct sequential simulation implementation in this thesis, all variables are standardized to have zero mean, and standard deviation of one at the composite scale v_c. Standardization is used to simplify setting up of the simple kriging equations within the program, but is not mathematically required. The standardized variables, denoted here with y are:

$$y_p(\mathbf{u}_i; v_m) = \frac{z_p(\mathbf{u}_i; v_m) - m_p(v_c)}{\sigma_p(v_c)}, i = 1, \ldots, n_p$$
$$y_k(\mathbf{u}_j; v_c) = \frac{z_k(\mathbf{u}_j; v_c) - m_k(v_c)}{\sigma_k(v_c)}, j = 1, \ldots, n_{k'}, k = k', \ldots, K$$

Simple intrinsic cokriging is adopted as the multivariate model for downscaling. This does not require calculating and downscaling a full linear model of coregionalization provided that the downscaled correlation matrix is positive definite. The simple intrinsic cokriging equations are established considering each datum and previously simulated value as a data event e_j where each data event has a data type k, support volume (v_m or v_c here), standardized value y and location $\mathbf{u}$. The simple intrinsic cokriging equations can then be expressed as:

$$\mathbf{C}\lambda = \mathbf{R}, \mathbf{C} = \begin{bmatrix} \bar{C}(\mathbf{u}_1, \mathbf{u}_1) & \bar{C}(\mathbf{u}_1, \mathbf{u}_2) & \ldots \\ \bar{C}(\mathbf{u}_2, \mathbf{u}_1) & \bar{C}(\mathbf{u}_2, \mathbf{u}_2) & \ldots \\ \vdots & \vdots & \ddots \end{bmatrix}, \lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \end{bmatrix}, \mathbf{R} = \begin{bmatrix} \bar{C}(\mathbf{u}_1, \mathbf{u}_{sim}) \\ \bar{C}(\mathbf{u}_2, \mathbf{u}_{sim}) \\ \vdots \end{bmatrix}$$

with estimation weights denoted by λ and where the standardized average covariances between data events are calculated:

$$\bar{C}(\mathbf{u}_i, \mathbf{u}_j) = \rho\left(e_i, e_j; v_i, v_j\right)\left(1 - \bar{\gamma}_p\left(\mathbf{u}_i - \mathbf{u}_j; v_i, v_j\right)\right)$$
$$\bar{\gamma}_p\left(\mathbf{u}_i - \mathbf{u}_j; v_i, v_j\right) = \frac{1}{n_{disc(i)} n_{disc(j)}} \sum_{\alpha=1}^{n_{disc}(i)} \sum_{\beta=1}^{n_{disc}(j)} \gamma_p\left(\mathbf{u}_\alpha - \mathbf{u}_\beta; v_c\right)$$

where ρ is the correlation coefficient between the two data event types which has been previously inferred by downscaling of the cross variogram, and the variogram used in the calculation of the average variogram is the standardized, downscaled variogram for the variable of interest, $Z_p(\mathbf{u}; v_c)$. That is, the variogram used is the small scale variogram, or a variogram at a smaller

scale provided the average variogram values are calculated with an appropriate discretization. The average variogram is calculated such that discretized subvolumes of v_c within v_m are used. Therefore, if v_m is a 15 m core length, and v_c is a 5 m core length, a discretization level of $\{i = 3, j = 1\}$ would be used in the calculation of the average variogram. The standardized estimate mean and variance are calculated with the solved kriging weights:

$$y_p^* (\mathbf{u}_{sim}; v_c) = \sum_{j=1}^{n_{krige}} \lambda_j y_j$$

$$\sigma_p^2 (\mathbf{u}_{sim}; v_c) = 1 - \sum_{j=1}^{n_{krige}} \lambda_j \bar{C}(\mathbf{u}_j, \mathbf{u}_{sim})$$

As direct sequential simulation is used to exactly reproduce the large scale data (Ren, 2007), the local conditional distribution given the estimated local mean and variance is required. To establish the shape of the conditional distribution, the Gaussian transform lookup table approach of Oz et al. (2003) is used. Any of the histogram corrections previously described could be applied, but this approach was found to work well and be computationally efficient even for large data sets, such as the oil sands test case. Once the corresponding local distribution with the correct mean and variance are looked up, the simulated value is drawn from the distribution.

4.3.5 Constraints and dynamic resimulation

Using this method with carefully discretized large scale data will result in realizations of small scale values which average exactly to the large scale values and have the correct covariance structure. Careful discretization is emphasized as the average covariances are numerically approximated, not evaluated analytically. In the case of the oil sands example, in which the large 9 m samples are composed of three 3 m samples, the discretization of the 9 m should exactly match the locations of the 3 m samples it is composed of for exactness in the resulting simulations. If the discretization does not match exactly, then data will not be exactly reproduced due to numerical errors in the calculation of average covariances.

As simulation progresses and a large composite sample is filled in with simulated small scale samples, the kriging variance will decrease until the last piece to be simulated has a kriging variance of zero. However, there is no guarantee that the estimated mean or simulated value will satisfy any physical constraints, such as the requirement for non-negativity. This is not acceptable from a physical or practical standpoint, and setting the simulated value to zero would result in data sets which when upscaled did not exactly reproduce the large values introducing a bias.

The approach taken was to reject any set of simulated values which did not meet imposed physical constraints and draw new values. The rejected set of simulated values includes all realized values within the large scale sample which is being downscaled; not just the negative value itself. During the execution of the simulation algorithm, when imputing a volume results in a negative simulated value all previously simulated values coming from the large scale measurement are rejected and the locations pushed to the simulation queue. This dynamic resimulation method is similar to rejection sampling for Monte Carlo simulation from a distribution. This maintains the sequential nature of the simulation algorithm, but requires slightly more computational time if many values are rejected.

The ad hoc nature of the rejection sampling correction for downscaling is reasonable as it results in distributions which satisfy the physical constraints on the problem, and there is no usable method to determine the shape of a joint conditional distribution without adopting a parametric (Gaussian in this case) framework. The quantile transformation procedure used for Gaussian simulation is nonlinear, and any back-transformed realizations are not required to reproduce the input data on average. In practice, the issue of negative simulated values is uncommon as the conditional variance of a distribution near a constraint, such as zero for a positively constrained grade, is extremely small even if the global distribution of the variable has a high variance due to the proportional effect. This issue is most likely to occur if the histogram was downscaled using an unreasonable change of support method. Furthermore,

many metallurgical variables of interest, such as the Bond mill work index, have distributions which are not near a constraint.

4.3.6 Implementation

The described downscaling procedure using direct sequential simulation with intrinsic cokriging for downscaling was implemented for heterotopically sampled values with arbitrary scales in a Fortran program *datascale*. The implementation is as previously described where standardized data values are expected for all variables. This implementation permits the usage of any number of heterotopically sampled, multiscale variables. Every measurement, or data event, is associated to a scale and variable type, so scales may also differ within variable types. This flexibility means that no assumptions are made regarding discretization so this must be explicitly inputted by the user.

The dynamic resimulation queue is capped at a finite number of steps, after which values will be set to a constraint. In practice, if an appropriate change of support model is used for the histogram then few resimulation steps are required. If an inappropriate change of support correction is used, and the conditional variance is high near a constraint, then numerous resimulation steps will be required.

4.4 Demonstration of downscaling algorithm

As with statistical parameter downscaling, the downscaling algorithm is demonstrated using the oil sands data for which the true values are known so the results can be evaluated. The downscaling procedure was implemented using the program *datascale* with standardized fines (at 3 m), bitumen grade (at 9 m), the downscaled bitumen grade variogram, correlation coefficient and histogram using the indirect lognormal correction. The downscaled realizations were evaluated against the previously described criteria.

The primary criterion established for downscaling is that the downscaled values be exact. A downscaled realization of bitumen grade at 3 m was averaged up to 9 m and plotted against the input 9 m grade (Figure 4.14). Using

the direct sequential simulation approach, the large scale measurements are exactly reproduced.

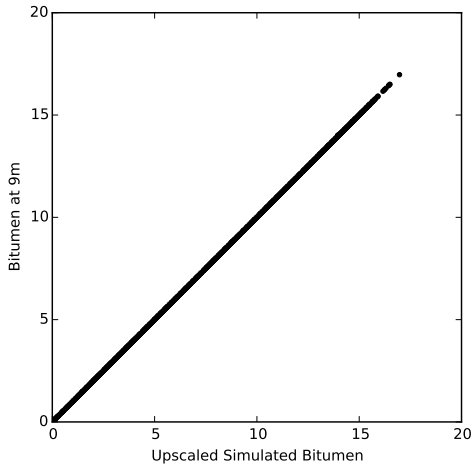

Figure 4.14: Reproduction of large scale bitumen grades at 9 m used to condition the downscaling. The large scale values are exactly reproduced.

Global histogram reproduction was evaluated considering both the true histogram at 3 m (which would typically be unknown) and the input downscaled histogram (Figure 4.15). Histogram reproduction is very close to the true histogram, and the constraint for positivity is satisfied. The lowest simulated bitumen grade is 0%.

Bivariate reproduction of the bitumen-fines relationship at the 3 m scale is evaluated by plotting the realized pairs with the downscaled bivariate distribution (Figure 4.16). The realized values have a slightly lower correlation (-0.79) compared to the true measurements (-0.84), but the simulated relationship is reasonable.

Reproduction of the spatial variability is evaluated by comparing the true experimental variogram points at 3 m (which would typically be unknown) with variograms calculated on the downscaled values (Figure 4.17). Variogram reproduction is nearly exact, likely due to the high degree of conditioning.

All criteria laid out for downscaling are satisfied using the proposed direct sequential simulation with intrinsic cokriging approach. Realizations are exact, have the correct multivariate structure, histogram and spatial variability. In

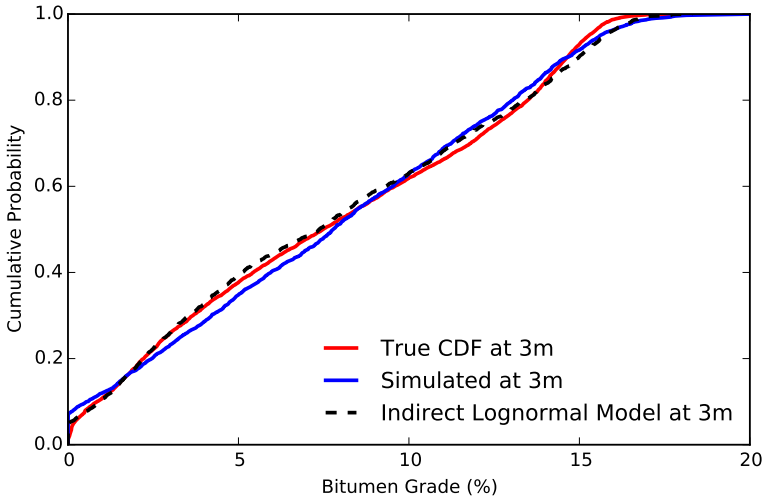

Figure 4.15: Univariate histogram reproduction of the input indirect lognormal scaled histogram at 3 m and the true histogram at 3 m. Histogram reproduction is very close for the simulated realization.

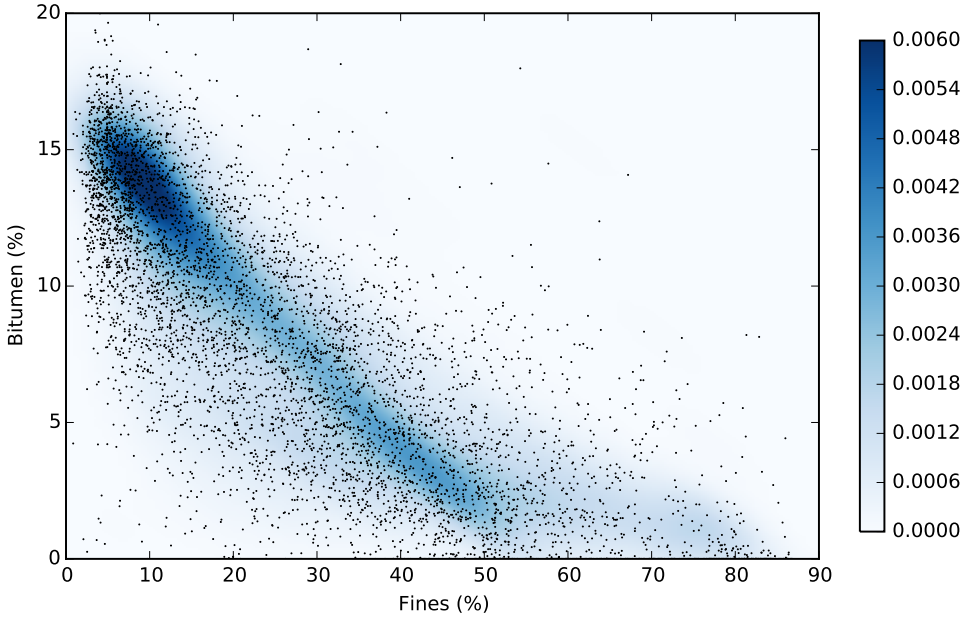

Figure 4.16: Downscaled bivariate distribution reproduction of the downscaled values at 3 m. The realized values have a slightly larger scatter (and corresponding lower correlation of 0.79) compared to the true bivariate distribution.

addition to these criteria, the effect of simulation order was investigated for this highly skewed data set to evaluate the resimulation approach (Figure 4.18). A slight bias in the final step of simulation, for which the kriging variance is zero,

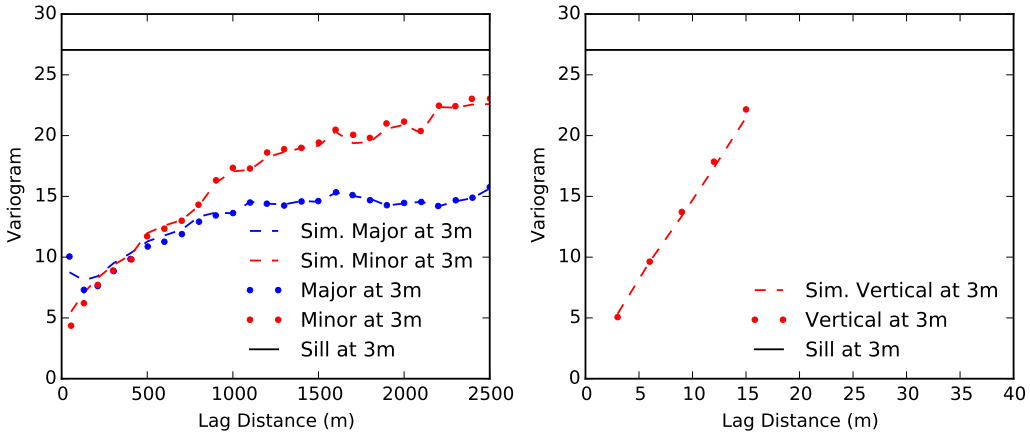

Figure 4.17: Downscaled variogram reproduction at the 3 m scale and experimental points from the true variogram.

is observed in the histograms. For the 3rd 3 m volume simulated within the 9 m volume, the probability of simulating a bitumen grade of 0% is very low. As multiple realizations, each of which have a different simulation path, are used for this stochastic approach this is unlikely to present a practical issue with modeling, but should be considered for highly skewed variables which are near a constraint, such as grade variables.

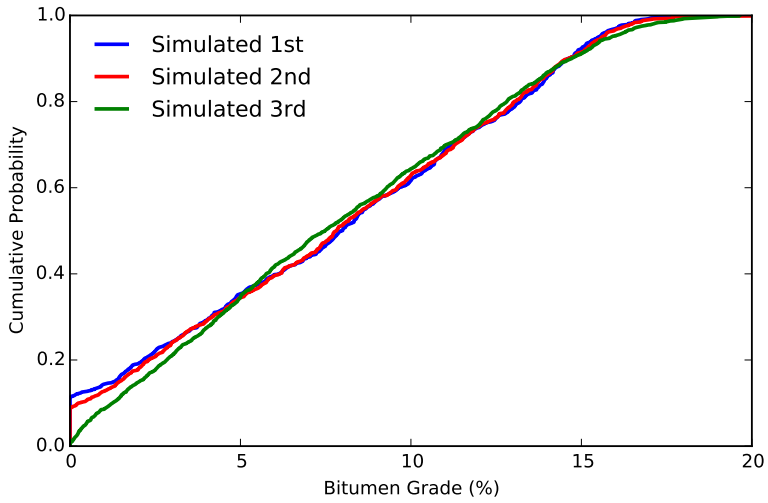

Figure 4.18: Histograms of 3 m scale values simulated 1st, 2nd and 3rd within the 9 m composite. Histograms are very close, although the 3rd simulated value has a lower chance to simulate a value very near to zero.

4.5 Discussion of downscaling

The downscaling procedure proposed, direct sequential simulation with intrinsic cokriging with dynamic resimulation, has been demonstrated in an atypical use case where the true small scale values are known. The resulting single-scale, homotopically sampled stochastic data sets may be used in further modeling operations such as intrinsic supersecondary colocated sequential Gaussian simulation, as will be demonstrated in the case study. Alternative frameworks for multiscale data integration, and downscaling, have been proposed by a number of researchers. These techniques are briefly reviewed, however the focus of this discussion is on the integration of downscaling into a simulation modeling workflow discussed.

4.5.1 Requirement for downscaling for multiscale data

Multiscale data presents a challenge for geostatistical modeling; the majority of modeling techniques have been developed for data at a consistent scale. For modeling techniques which do not explicitly require data at a single scale, available software is often limited to a single scale by design.

Using stochastic downsclaing, multiple non-unique realizations of the data are generated. These multiple data realizations are at a single scale, and suitable for a stochastic modeling workflow. However, as these realizations are non-unique, they are not suitable for deterministic modeling such as a kriging based estimation workflow.

4.5.2 Alternative algorithms for multiscale data

Multiscale data are very common in petroleum applications with the collection of seismic, well test, micro-resistivity and other measurements. As a result much of the research into multiscale data integration and downscaling has been for these applications. In addition to the work of Oz (2003) and Ren (2007), which forms the basis for the direct sequential simulation approach proposed in this thesis, a selection of alternative multiscale modeling and downscaling techniques are reviewed here.

Pardo-Igúzquiza et al. (2006) applied cokriging for image sharpening of remote sensing images. The approach requires modeling of a linear model of coregionalization as no Markov or intrinsic assumption is made for the variograms. This approach is deterministic, so necessarily will not reproduce the spatial structure and is not suitable for data downscaling for further modeling. Similarly, Goovaerts (2010) proposed a block kriging formalism for integrating multiscale data. This approach, also deterministic, is appropriate for estimation with multiscale data but not suitable for stochastic downscaling.

Jha et al. (2013) applied multiple point statistics using the direct sampling approach for downscaling climate models. This approach is very powerful making no assumptions of linearity or Gaussianity, but has the very strong requirement of a training image. As it is very unlikely that a geologically realistic training image is available for metallurgical properties, this approach has limited applicability for downscaling metallurgical measurements. Allard and Bourotte (2014) proposed a stochastic approach for disaggregating rainfall measurements. The sequential simulation method proposed is powerful if an empirical model of the generating process is available for fitting; however, this is unlikely to be available for metallurgical measurements.

Tran et al. (1999) combined block kriging and Bayesian updating for downscaling coarse block models of permeability to high resolution models. This is an efficient approach for a single variable of interest, but not straight forward to apply with multivariate data where the multiscale multivariate distributions, or an assumption of independence, is required to formulate the conditional distributions. Of the reviewed multiscale data integration approaches, this approach is the strongest alternative for downscaling large scale metallurgical measurements. Given a sufficient multivariate model for secondary measurements, this could be used as an alternative to the direct sequential simulation approach proposed.

4.5.3 Integration into a classical simulation workflow

The small demonstrative case study presented considered the downscaling of bitumen grades from a larger 9 m scale to the original 3 m scale that the data

were collected on. This is not a typical use case for the downscaling approach, but valuable as it permits the evaluation of the downscaling algorithm where the statistics of the true small scale values are known. Variogram and histogram reproduction at the small scale using this approach are reasonable. Due to the use of the resimulation step in the downscaling algorithm, no negative grades were simulated and simulated values average exactly to the original large scale values.

Using simple intrinsic cokriging, direct sequential simulation may be used to construct multiple downscaled realizations of the data suitable for integration into a traditional geostatistical simulation workflow. In a classical simulation workflow, multiple realizations of a model are built using a single data set. These realizations are then jointly processed for uncertainty assessment. Multivariate techniques such as sequential Gaussian simulation using colocated cokriging or intrinsic supersecondary cokriging require data to be at a consistent scale. Using the downscaling approach proposed in this thesis, multiple stochastic realizations of the data would be constructed. The placement of direct sequential simulation within a simulation workflow is highlighted in Figure 4.19.

As shown by the flow chart, integrating large scale metallurgical properties into a geostatistical workflow requires substantially more modeling effort; the variogram, histogram and correlation coefficients must be downscaled before imputing multiple data sets. The potential value for this effort is large. Realizations of metallurgical variables such as mill throughput at the scale of interest provide the information to quantify uncertainty in metallurgical properties as a function of grade.

In addition to the use of this downscaling technique for large scale metallurgical properties, other applications are possible. The use of a direct sequential simulation with intrinsic cokriging and dynamic resimulation is generally applicable to downscaling properties with the requirements discussed here. For data which are not vertically composited, as drill hole typically data are, additional care is required to correctly discretize the data and ensure that search neighbourhood includes all relevant data (Gómez-Hernández et al., 2005).

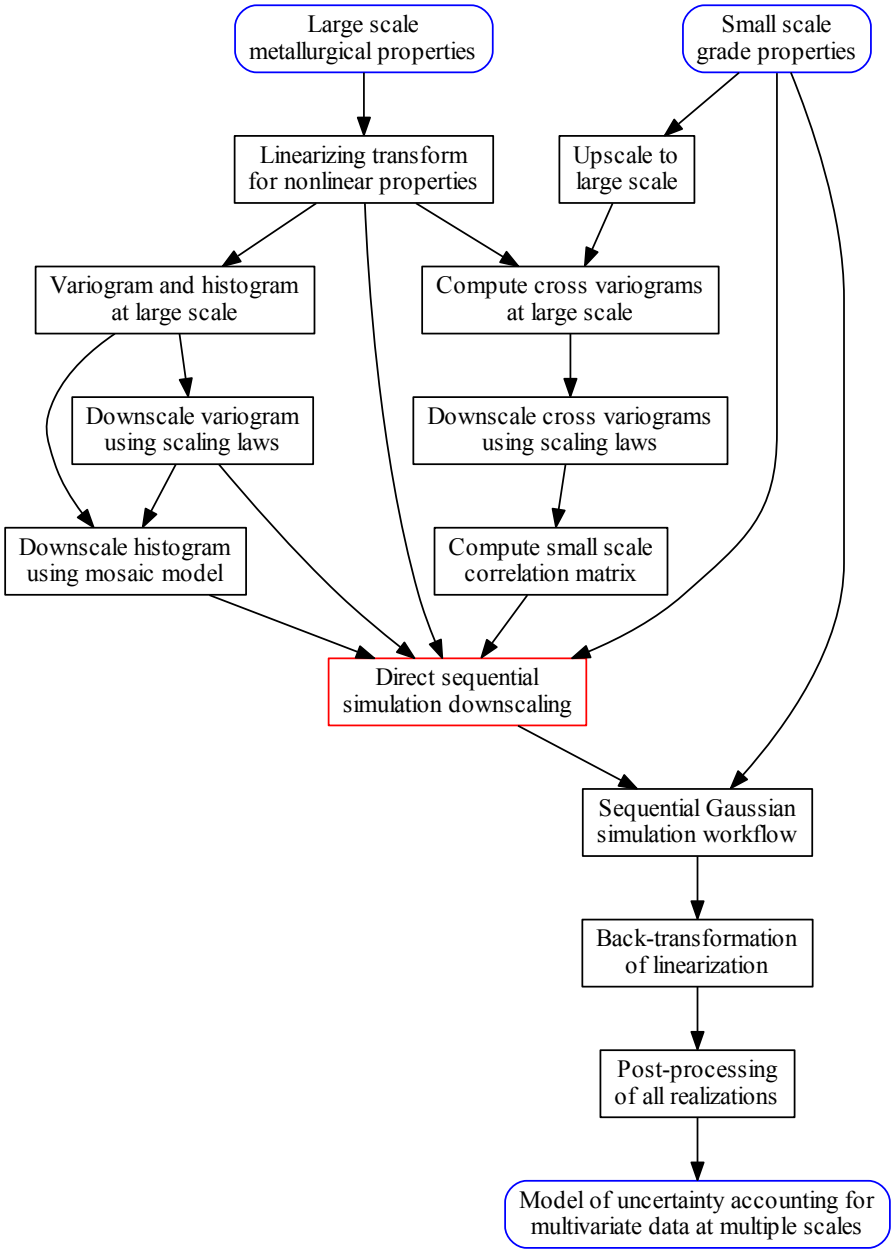

Figure 4.19: Integration of the downscaling workflow into a traditional geostatistical simulation workflow.

Chapter 5

Challenges for the spatial modeling of metallurgical properties

Challenges related to nonlinearity and multiple data scales for geometallurgical modeling have been addressed in chapters 3 and 4 of this thesis. The focus has been on the development of novel techniques and creative application of existing methodologies for the modeling of nonlinear variables and the downscaling of large scale metallurgical samples. In this chapter, the focus is not on the development of new techniques to address specific geometallurgical modeling challenges, but instead on the synthesis of existing spatial modeling methods into a cohesive workflow for geometallurgical modeling which is practical for engineering usage.

In addition to the selection and application of workflow techniques, a number of practical challenges for geometallurgical modeling including the inference and management of uncertainty in statistical parameters, the use of hierarchical transformation workflows, and the application of geometallurgical models for decision making are addressed.

5.1 Workflow selection

All geometallurgical modeling workflows are driven by specific study requirements, but there are common study goals and approaches. The quantification of uncertainty, decision of data collection rates, and optimal sequencing are

all common study goals. In these cases, many variables including grades, and metallurgical and geotechnical properties are often of interest driving the use of a multivariate workflow. Complex relationships, the previously discussed issues of nonlinearity and multiple scales, and a range of geologic conditions pose modeling challenges which require special care. Given the range of conditions encountered, there have been a number of approaches taken for high-resolution modeling of geometallurgical properties.

5.1.1 Literature proposed approaches

Powell (2013) emphasizes the requirement for high-fidelity models for process optimization on geometallurgical properties. Smoothly estimated kriged models are insufficient for optimizing mineral processing operations for a heterogeneous orebody. A model with too-low variability will understate the local variation in properties which should be considered in the design and operation of the mine and mill. In addition to estimates with too-low variability, kriging may introduce a bias for variables that do not average linearly, such as metallurgical properties (Coward et al., 2009; Lozano and Bennett, 2003). Kriging techniques classified as nonlinear methods including lognormal kriging, disjunctive kriging and indicator kriging were investigated by Moyeed and Papritz (2002) and found to perform relatively poorly for the estimation of skewed variables.

Numerous authors including Carrasco et al. (2008), Newton and Graham (2011), van den Boogaart et al. (2013), Boisvert et al. (2013) and others have proposed geostatistical simulation for modeling metallurgical properties as no requirement is placed on linear averaging. Simulation approaches also correctly model the joint uncertainty between variables at the cost of increased computational time compared to typical estimation algorithms such as ordinary kriging or multigaussian kriging. In the Olympic Dam expansion study by Boisvert et al. (2013), principal component analysis was used to generate uncorrelated variables which were simulated over the domain using sequential Gaussian simulation with a subsequent regression fit to predict plant performance. Redundant and unrelated variables were first eliminated to avoid over

fitting. A two stage regression model for predicting plant performance was used in which the variables were merged into supersecondary variables and then a linear regression model used to predict plant performance.

A number of researchers have proposed the establishment of geometallurgical domains within which metallurgical properties are estimated using a separate transfer function (Everett and Howard, 2011; Keeney and Walters, 2011; Newton and Graham, 2011). Domains are established on geological and metallurgical information, and the expected time during a mine's life when the rock will be mined. Within these domains variables are estimated or simulated using classical statistical and geostatistical algorithms such as multivariate regression and Gaussian simulation.

Modeling approaches advocated and used in this thesis are in the domain of stochastic simulation (discussed in chapter 2). These are the only methods which can correctly model the joint multivariate uncertainty in the rock properties of interest. The decision of model purpose and study aims will dictate the selection of variables to model, the domains to model and the required resolution of the model. Increased understanding of the spatial variability of metallurgical properties will lead to better decisions for mine planning and mill operation. This motivates high-resolution models of all relevant metallurgical properties, in addition to the grades and geomechanical properties required for mine planning. Optimization strategies, such as the optimization of production planning (Dimitrakopoulos and Jewbali, 2013), to maximize mill throughput and grades recovered require high-resolution models of multiple relevant variables including grades and metallurgical properties to accurately model the interactions among the numerous variables.

5.1.2 Domain definition and categorical variable modeling

All statistical techniques discussed in this thesis require a decision of stationarity. This decision of stationarity is typically applied to data within a geologically homogeneous domain. In the context of spatial modeling, stationarity is the decision of "geological homogeneity" for a rock unit (McLennan, 2007). Rock units are typically geologic, such as the selection of highly al-

tered porphyry quartz as a domain, and host granodiorite as a second domain. This decision permits the calculation of statistical values, such as histograms and variograms, and the spatial modeling of rock properties. The decision of stationarity requires balancing the number of domains with a mineral deposit's inherent geologic complexity. Modeling too few domains unacceptably pools differing rock types together, while too many domains is unreasonable for modeling where statistics and domain boundaries must be inferred (McLennan, 2007). The decision of the number, and type, of domains is subjective and requires geological and statistical input. The selection and modeling of stationary domains is discussed further in the case study in chapter 6.

Once rock units have been selected as stationary domains, domains must be spatially modeled and boundaries defined. Techniques for domain boundary modeling may be broadly classified as deterministic, such as manual and implicit modeling techniques, or stochastic such as indicator simulation and multiple point statistics. The choice of a deterministic or stochastic technique is driven both by how well defined the geologic structure is, and the extent to which a rock type impacts the resulting model. In petroleum applications where the distribution of highly impermeable rock units is typically the most important, and the distribution of permeabilities within these rock units secondary, stochastic approaches are frequently used. For mining applications where the distribution of grades and metallurgical properties is of upmost importance, deterministic approaches are commonly used.

Both approaches require the scale of the categorical variable of interest (such as rock type) to be at the modeling scale. Measurements of categorical variables are typically on a fine resolution as logged by the site geologist. Upscaling a categorical variable forms large compositional variables (Pawlowsky-Glahn and Olea, 2004) which cannot be modeled with standard categorical modeling techniques. Current practice in geostatistical modeling is to select the majority category as representative for the upscaled variable. This practice is adopted here, but could constitute an area for future work.

Using the categorical variables upscaled to the scale of interest, the choice of a stochastic or deterministic categorical variable modeling method is re-

quired. As with the choice of continuous variable modeling techniques, this choice is driven by the geology of the deposit and study goals. For a deposit with high frequency categorical variable changes where the properties of interest are strongly a function of the rock type, such as for a petroleum reservoir, then a stochastic technique is used. In many mining applications where the properties of interest are not driven purely by the rock type and particularly for large-scale massive deposits such as porphyry deposits, deterministic techniques are preferred. Widely used stochastic and deterministic categorical variable modeling techniques are briefly reviewed here, as they are a key part in all modeling workflows.

Deterministic techniques include manual definition, indicator kriging (Journel and Huijbregts, 1978), signed distance functions (Deutsch and Wilde, 2013; Silva and Deutsch, 2013) and boundary methods (McLennan, 2007; Pawlowsky et al., 1993; Srivastava, 2005). Manual definition of rock types is typically carried out by geologists who use expert judgement and available drilling data to manually define a 3D wireframe. Indicator kriging adopts the kriging formalism for binary-recoded categorical variables. This permits the usage of different variograms for each category. Signed distance functions calculate the signed distance of a location to each category, and are then modeled using inverse distance or kriging based techniques. Boundary methods, which model the boundary rather than the rock mass itself may be mixed with any of the previously discussed techniques, and may even be used in stochastic modeling if each boundary is allowed to flex.

Stochastic techniques for categorical variable modeling include sequential indicator simulation (Alabert, 1987; Deutsch, 2006; Journel and Isaaks, 1984), truncated Gaussian and plurigaussian simulation (Armstrong et al., 2011; Galli et al., 1994; Matheron et al., 1987) and multiple point statistics (Guardiano and Srivastava, 1993; Mariethoz and Caers, 2014; Strebelle, 2002). Sequential indicator simulation adopts the sequential simulation algorithm with indicator kriging as the internal estimator. Unlike sequential Gaussian simulation, which is the recursive decomposition of the multivariate Gaussian distribution, there is no statistical formalism for the choice of sequential indicator simulation,

but it has been demonstrated to work well and allows the easy incorporation of secondary variables and trends. Truncated Gaussian and plurigaussian techniques recode the categorical variable as a continuous variable permitting the usage of any of the continuous stochastic techniques, such as sequential Gaussian simulation. These techniques have been demonstrated to work particularly well for naturally ordered categorical variables, such as a set of facies consisting of sand, shaly-sand, and shale. Multiple point statistics, which may also be used for continuous variables, use a training image with the correct spatial distribution of categories for simulation with the application of Baye's law.

As discussed, the choice of either a deterministic or stochastic technique is driven by the geology and study goals. Within either the deterministic or stochastic classes, the specific method for categorical variable modeling is selected on the basis of the strengths and weaknesses of each approach. In chapter 6, the modeling of a copper-porphyry deposit including technique selection is discussed as an example of these choices. Within the modeled domains, continuous variables in this thesis are modeled stochastically after a set of multivariate transformations.

5.2 Sequential multivariate transformations

Adopting the stochastic simulation approach for geometallurgical spatial modeling of continuous variables, there are still a large number of methodology decisions to be made. Consider constructing a spatial geometallurgical model given heterotopic, multiscale, nonlinear, correlated multivariate data. Additional constraints on the final spatial model, such as positivity and compositional constraints complicate the choice of methodology. In order to meet model requirements using stochastic simulation, and maximize usage of the data, there are two evident modeling approaches. The first approach is to explicitly enforce all required constraints in the construction of the model. The second approach is to construct the model without explicitly enforcing constraints, and check the resulting model to see if it has the required proper-

ties. In this thesis, the first approach is the primary approach adopted where required constraints are explicitly built into the modeling approach.

The generalized stochastic simulation workflow for explicitly enforcing model constraints is shown in Figure 5.1. Consistent with the rest of this thesis, these steps follow domain definition and the decision of stationarity and are therefore applied separately for each domain. The proposed workflow may be viewed as a data transformation workflow, for which the goal is to generate multiple homotopic, single-scale, univariate Gaussian data which are either correlated, or decorrelated. Of the multiple transformation steps shown in Figure 5.1, only a number of these steps may be required for the deposit of interest and study goals.

The first data transformation step is the application of compositional transforms (Aitchison, 1982; Pawlowsky-Glahn and Olea, 2004). Compositional variables are typically linear variables in which the sum of the variables must equal a constraint. These types of variables are typically whole rock analyses where the sum of all measured components must equal 100% for a sample. Consider K compositional variables. The multivariate scatterplot of the data must lie on the hyperplane for which the sum is 100%, and therefore only have $K-1$ degrees of freedom, so the data can be transformed to $K-1$ variables. There are a number of additional practical considerations which have been discussed by Pawlowsky-Glahn and Olea (2004); the interested reader is referred for additional information.

As described in chapter 3, nonlinear variables may be transformed through the use of a linearizing transform such as a power law model for modeling. The nonlinear transform requires a model of the nonlinearity, so may require additional experimental measurements or a highly correlated component measurement for usage. If nonlinearity may not be inferred due to limited experimental information, then linearity is commonly assumed and all averaging is deferred as late into the modeling process as possible following stochastic simulation which makes no averaging assumption.

After linearizing, large scale metallurgical measurements may be downscaled to a consistent modeling scale such as a bench or half-bench scale as

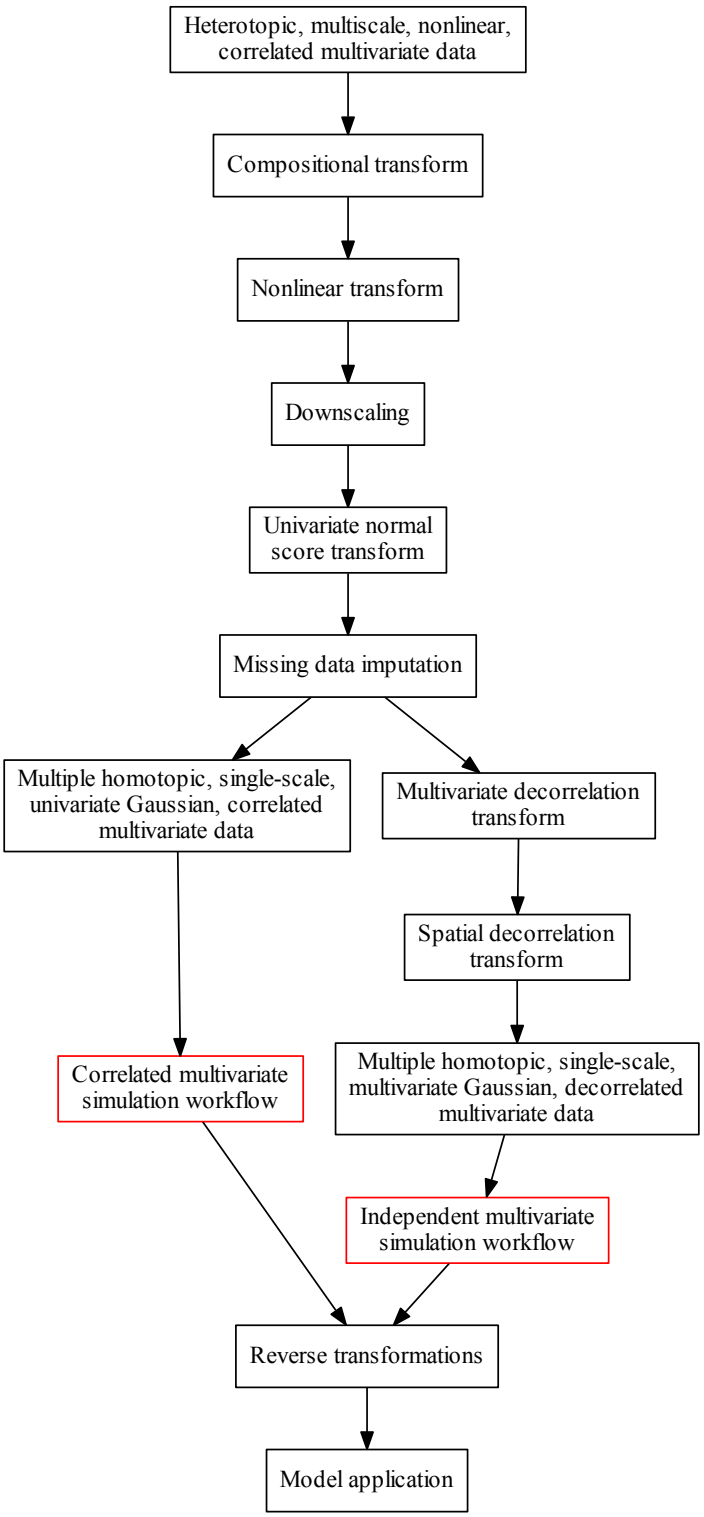

Figure 5.1: Generalized data transformation workflow for the spatial modeling of heterotopic, multiscale, nonlinear, correlated multivariate data.

described in chapter 4. After the application of compositional and nonlinear transforms, and downscaling of large scale measurements, the transformed data are heterotopic, single-scale, linear, correlated multivariate data. As downscaling is non-unique, the data are also multiple with L different data sets. For the application of Gaussian techniques, including missing data imputation and Gaussian simulation, the data are univariate normal score transformed.

Missing data are imputed using either a parametric or non-parametric imputation strategy (Barnett and Deutsch, 2015). As with the downscaling procedure, imputed values are non-unique and there are L distinct data sets. After data imputation, the multiple data sets are homotopic, single-scale, linear, correlated multivariate data. Under an assumption of multivariate Gaussianity, correlated multivariate techniques such as colocated cosimulation or super-secondary colocated cosimulation may be applied to construct L realizations, where each realization uses a distinct data set.

If the assumption of multivariate Gaussianity is unreasonable due to multivariate constraints, nonlinear features and heteroscedasticity, then a decorrelation transform may be applied. Decorrelation transforms, including the projection pursuit multivariate transform (Barnett et al., 2014b) and the stepwise conditional transform (Leuangthong and Deutsch, 2003) transform the complex, correlated multivariate data to multivariate Gaussian data which are decorrelated for $h = 0$. With the addition of a spatial decorrelation transform, such as minimum/maximum autocorrelation factors (Larsen, 2002), the data are additionally decorrelated for $h = \Delta$, where Δ is chosen as a lag of interest. Using the decorrelation approach, univariate simulation approaches such as sequential Gaussian simulation may be applied independently for each transformed variable to construct L realizations.

After L spatial models are constructed of the transformed variables, each transformation including decorrelation, normal score, nonlinear and compositional transforms are inverted in reverse order. The resulting models may be checked and verified, which will be discussed in greater detail in the case study, and the models used for the study goals.

5.3 Parameter inference and uncertainty

Geostatistical modeling parameters such as univariate histograms, multivariate correlation matrices and spatial covariance functions are inferred on limited spatially correlated data. As the spatial correlation of data is uncertain, the inference of reasonable parameter uncertainty is challenging.

5.3.1 Variogram uncertainty

The variogram is a central parameter for many geostatistical workflows. Variogram models must be fit to experimental variograms to ensure positive definiteness of calculated covariance matrices and permit calculation at distances with no experimental data. There is substantial uncertainty in the modeled variogram as it is fit to limited experimental data. The uncertainty in this model should be addressed. A number of methods for determining this uncertainty have been proposed over the years, but these methods are rarely applied and information on variogram uncertainty is not used in the modeling process.

The majority of proposed approaches for the calculation of variogram uncertainty are similar. Pardo-Igúzquiza and Dowd (2001) calculated the variance-covariance matrix of the experimental variogram using fourth-order moments. These moments were used to determine the variance-covariance matrix of the modeled variogram parameters using generalized least squares with the assumption that the variogram parameters were normally distributed. Marchant and Lark (2004) also documented the determination of the variance-covariance matrix of variogram model parameter estimates and iterative estimation of the variogram model parameters using the fourth order moment approach.

Ortiz and Deutsch (2002) calculated the fourth-order moments of the experimental variograms to determine the variance of each experimental point. Under a multivariate Gaussian assumption the uncertainty in each variogram point is known. The validity of this approach was demonstrated considering a local simulation and global simulation approach. This approach results in the uncertainty in each variogram point, but not a joint uncertainty in the

modeled variogram. The selection of extreme cases which honour this point uncertainty was proposed.

Koushavand et al. (2008) used the fourth-order moments with the assumption that each lag parameter was gamma-distributed to model variogram uncertainty. A global simulation approach using LU simulation to generate unconditional simulations of experimental variograms was also used. These methods produced comparable results when tested in a small case study.

These approaches result in theoretically correct distributions for the uncertainty in variogram model parameters under a multivariate normal distribution assumption. The assumption of multivariate normality, uniquitous in much of geostatistics, may be appropriate for variogram uncertainty determination in specific cases, but is unlikely to be for parameters near constraints. If a variogram shape is assumed, the information matrix on the variogram model parameters calculated from the fourth order moments fully parameterizes the joint distribution of the model parameters if the distribution is multivariate normal. Unfortunately, inferring the information matrix is numerically unstable in the common case of very limited experimental data in a 3D deposit. Constraints imposed on the variogram parameters, such as the requirement for positive structure ranges and contributions further complicate the problem and render the multivariate normal error assumption for parameters invalid in practical cases.

There are a number of theoretical and implementation challenges for the application of the information matrix approach in a realistic 3D setting with few data. Variogram ranges and structure contributions are constrained to be positive, and the structure contributions have a sum constraint. For parameters such as the nugget effect which may be very close to zero, the multivariate normal approximation for uncertainty becomes intractable. This problem is exacerbated when the parameter uncertainty is very high due to limited calculation points and many parameters.

There is also substantial numerical instability in both estimating and inverting the Jacobian matrix. The partial derivatives range from zero for modifying the range on a direction perpendicular to the experimental lag to very

large values for the structure contributions. Fixing the variogram model angles at a slight offset to the experimental directions reduces the problem slightly, but not enough to allow stable estimation of the Jacobian in cases tested. For these reasons, the simulation approaches to variogram uncertainty of Ortiz and Deutsch (2002) and Koushavand et al. (2008) are recommended. These approaches consider the joint uncertainty in experimental lags and will always result in positive, licit variogram values.

5.3.2 Mean uncertainty

Uncertainty in the probability distribution of a variable, particularly in the distribution of the mean, is an important parameter in a geostatistical workflow. This applies for grade variables as well as metallurgical variables, such as grinding indices where a higher than expected average power requirement will lead to poor plant performance and throughput. Many techniques have been developed and applied for the incorporation of mean (and distribution) uncertainty in geostatistical workflows including the bootstrap, conditional finite-domain and global kriging.

The bootstrap, and more generally from a spatial modeling perspective, the spatial bootstrap (Deutsch, 2004; Efron, 1979; Pyrcz et al., 2006) are statistical resampling techniques which can be used to resample a distribution for mean uncertainty. By resampling the distribution of a variable within a stationary domain, multiple possible realizations of the histogram and mean may be constructed. The spatial bootstrap is widely used for mean uncertainty due to its robustness and ease of implementation, but does not account for domain extents in the quantification of uncertainty.

The conditional finite-domain (Babak and Deutsch, 2009a) is an alternative to the bootstrap for quantifying mean uncertainty. By simulating under rotations and translations of the sampling configuration within a stationary domain, multiple realizations of the mean may be generated. The conditional finite-domain approach is not as widely used as the spatial bootstrap due to complexity of implementation and the requirement for many sampling configurations for stable results.

Global kriging my also be used to quantify uncertainty in the mean (Deutsch and Deutsch, 2010). Under the strong assumption of stationarity, the kriging system is solved using the average covariance of the sampled data locations to the entire stationary domain. The average covariance is typically inferred using Monte Carlo integration due to complex domain shapes. Global kriging effectively accounts for data conditioning to the domain, but does not incorporate the shape of the distribution for assessment of uncertainty.

The bootstrap, conditional finite-domain and global kriging approaches may be used to quantify uncertainty, but there is still the requirement to enforce the calculated mean uncertainty on simulated realizations. Relying on ergodic fluctuations from simulation is an unreliable method for enforcing mean uncertainty. The two primary options for integrating mean uncertainty are post processing (Journel and Xu, 1994), or transforming pre-simulation subject to the sampled distribution (Barnett et al., 2014a).

Post-processing matches quantiles from the simulated realizations with the expected distribution from one of the three techniques described using a procedure analogous to the normal score transform. This guarantees that the final distributions exactly match the calculated mean uncertainty, but artificially alters covariances between simulated locations. In the pre-simulation approach data are normal score transformed using different uncertain distributions for each realization. This does not modify covariances between simulated locations, but does not guarantee that realizations will reproduce the input mean uncertainty. The decision of mean uncertainty quantification method, and method for targetting the mean uncertainty will depend on the specific study goals.

5.4 Large scale uncertainty inference

The emphasis in this thesis has been on methodologies for spatial modeling with metallurgical variables and the selection of these methodologies. Relatively little discussion has been given to applications of the spatial model constructed. In this section, the very common engineering application of the

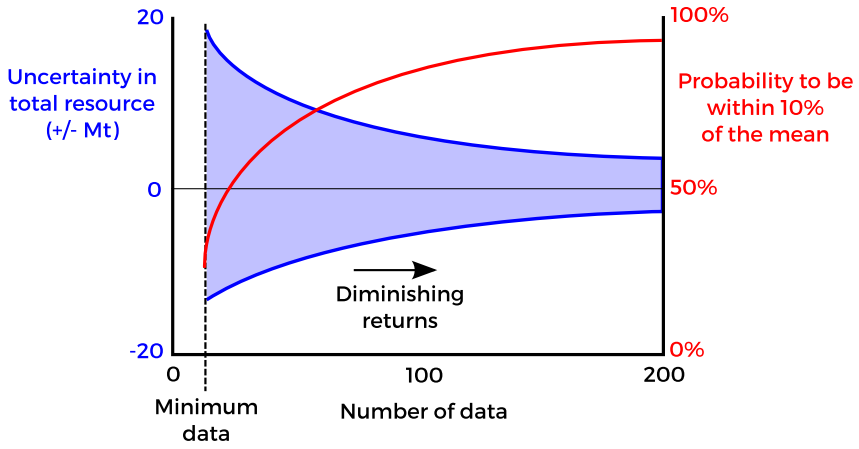

Figure 5.2: Sketch of global uncertainty and the probability to be within a percentage of the mean as a function of the number of data collected. Adapted from Khan et al. (2014).

spatial models for large scale uncertainty inference is discussed. Large scale uncertainty inference is the quantification of uncertainty on a scale much larger than typical models constructed on a selective mining unit basis.

Models of large scale uncertainty are typically used in the decision of drilling and sample collection requirements in the early stages of mineral deposit assessment. Of interest is uncertainty in key parameters, such as global resource tonnage or the probability to be within a specified percentage of the mean. These metrics are useful in the decision of an acceptable number of data to collect. The goal is a quantitative measurement of these uncertainties as a function of data collection (Figure 5.2).

Consider the typical sequence of events in the exploration and assessment of a mineral deposit. Initially, few data are available and the deposit is primarily modeled conceptually based on geologic knowledge and limited sampling. Quantitative assessment of uncertainty using geostatistical methods is impractical at this stage as there are too few data to quantify the spatial distribution or statistical distributions of measurements. A weights of evidence approach may be applied at this very early stage (Agterberg and Bonham-Carter, 1990; Agterberg et al., 1990; Cheng, 2015).

As more samples are collected, uncertainty in key large scale parameters may be quantified. This is commonly accomplished using a resimulation ap-

proach (Khan et al., 2014). Collecting more data yields diminishing returns for global resource assessment, and is primarily of interest for local uncertainty quantification and the delineation of ore/waste boundaries.

5.4.1 Measurement of data spacing

Many decisions on the number of data to collect are based on an average drill hole spacing, not the total number of drill holes. Drill hole spacing is typically defined to be the distance between neighbouring drill holes if drilling is on a rectangular grid. Practically, drilling is rarely on a rectangular grid as the goal is to delineate ore/waste boundaries and maximize drill rig utility, not collect a statistically representative sample. In order to compare the existing drill spacing with potential scenarios for drill spacing, measurements of data spacing are required.

In two dimensions, such as cases where vertical drilling is employed over an area, data spacing may be evaluated by calculating the equivalent spacing given a circle containing a constant number of data, N (Wilde, 2010).

$$\text{Data Spacing} = \sqrt{\frac{\pi(r_N + r_{N+1})}{2N}}$$

This is schematically shown in Figure 5.3. Using N data, where N is typically 10-20 samples, the average radius of circles just containing the N and $N+1$ nearest data to the location for which the data spacing is being calculated are used. If this process is repeated for every cell in a map over a 2D area with a set of samples (Figure 5.4), the result is a map of the data spacing (Figure 5.5).

In three dimensions, the logical extension is from a circle to disc or sphere. Using the equivalent radius method with a constant number of samples with either of these shapes yields sub-optimal results. An alternative approach was developed for quantification of data spacing in 3D. Consider generating a set of synthetic drill holes perpendicular to the plane of major continuitity (Figure 5.6a), a common approach to drilling a mineral deposit. These synthetic drill holes are on a rectangular grid with drill hole spacing x. These drill holes are intersected with an ellipsoid as shown in Figure 5.6b. The length of drilling

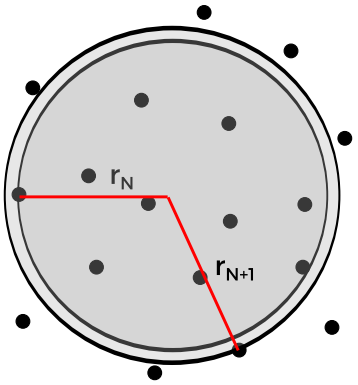

Figure 5.3: Sketch of constant number of data method for data spacing calculation in 2D. Circles with radii equal to the distance to the Nth nearest data, and $N+1$th nearest data are calculated and used in the data spacing calculation.

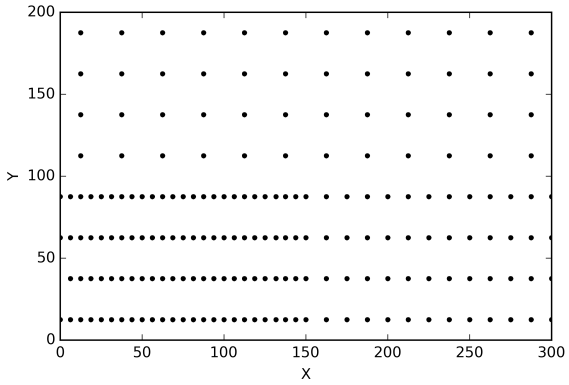

Figure 5.4: Synthetic data locations used to demonstrate data spacing calculation in 2D.

intersecting the ellipsoid is calculated for many drill hole spacings, and many offsets relative to the ellipsoid. The drill hole spacing, or data spacing, may then be plotted as a function of the drill hole length intersecting the ellipsoid (Figure 5.6c). For a location of interest in the deposit, the drill hole or data spacing may be calculated by using the same ellipsoid and calculating the length of drilling within the ellipsoid.

5.4.2 Workflow for uncertainty assessment

Using knowledge of the current data spacing and uncertainty (quantified through simulation), alternative data spacings may be evaluated. A general workflow for uncertainty assessment is the resimulation workflow (Boucher et al., 2005;

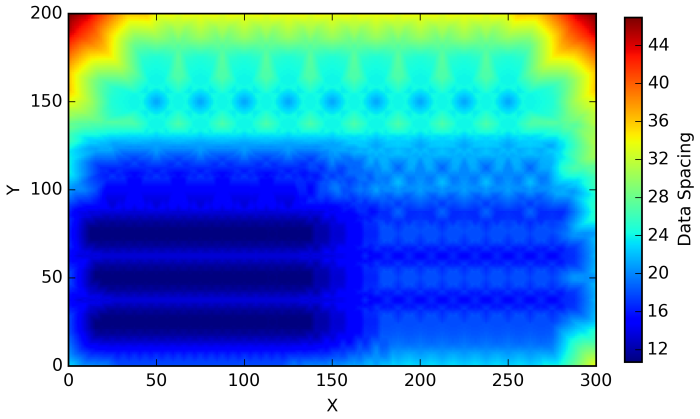

Figure 5.5: Results of the data spacing calculation for the synthetic case shown in Figure 5.4.

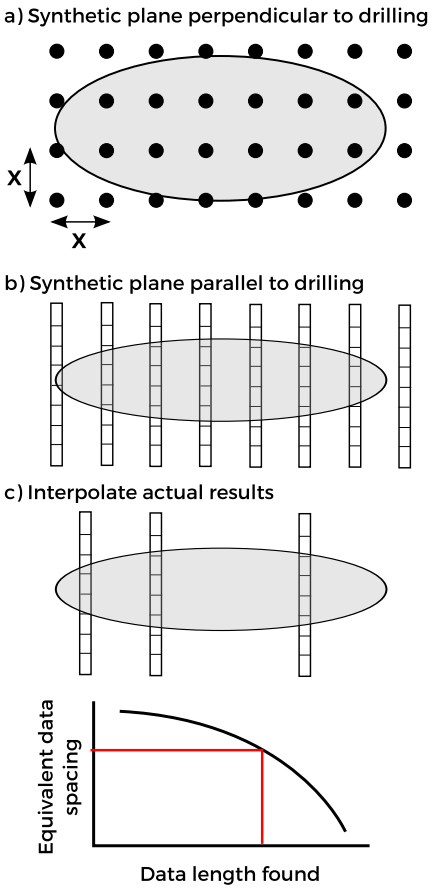

Figure 5.6: Data spacing calculation in 3D. Multiple synthetic drilling configurations are used to construct a plot of the equivalent data spacing as a function of the data length found within an ellipsoid oriented along the major plane of drilling.

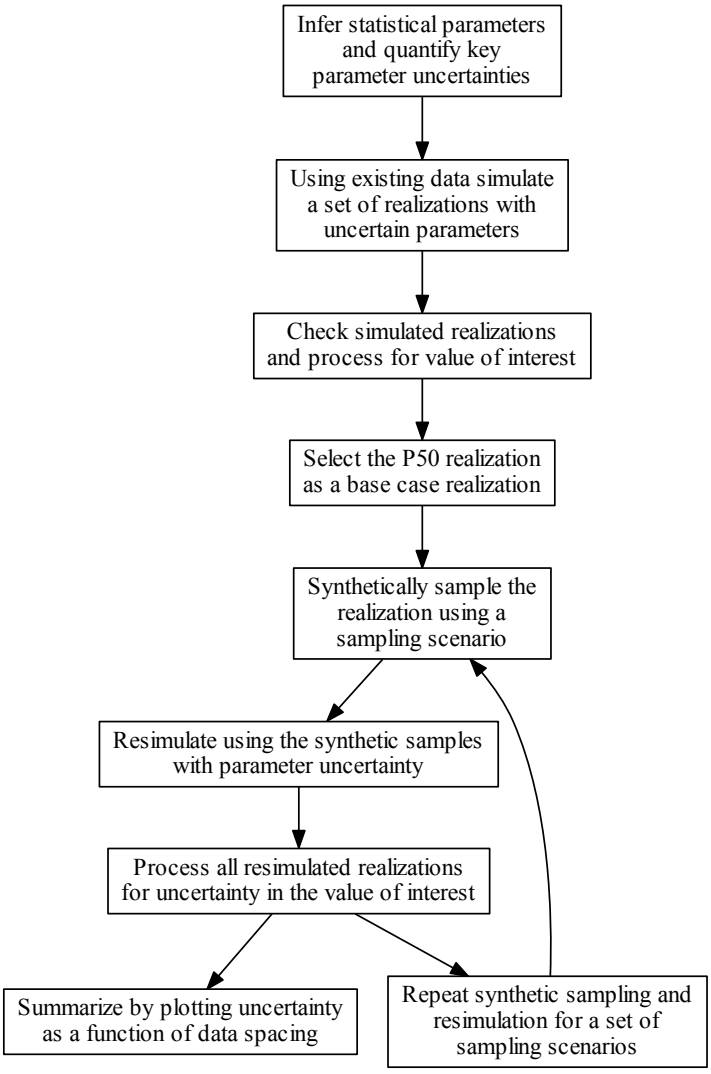

Figure 5.7: Resimulation workflow for calculating uncertainty as a function of data spacing.

Deutsch et al., 1999; Wilde, 2010; Wilde and Deutsch, 2013), which is summarized in Figure 5.7. The basis for the resimulation workflow is to calculate the variability associated with realizations of the domain of interest given uncertain parameters and variable drill spacings.

The assessment of uncertainty relies on knowing the "true" distribution of values. As the truth is unknown, a synthetic true distribution is generated to

have the correct spatial distribution of values conditional to currently available data. The P50 (median) realization is typically selected as a base case for the resampling procedure, but any realization (or a set of realizations) could alternatively be chosen. The basis for the selection of the P50 realization is that it is a moderate realization, not that it is more theoretically correct relative to other realizations.

The selected realization is synthetically sampled for a number of drill spacings, similar to the approach for 3D data spacing calculation, and the simulation procedure repeated. The result is uncertainty as a function of data spacing (Figure 5.2). Many factors influence the relationship between data spacing and uncertainty. Decreasing the data spacing (increasing the number of data collected) will decrease uncertainty. An increased range of spatial correlation (variogram range) will typically increase large scale uncertainty, but decrease local uncertainty. An increased domain size relative to the same sampling pattern will typically increase uncertainty.

The quantification of uncertainty as a function of data spacing has been discussed as an example application of high resolution spatial geometallurgical models. There are many other applications which are not discussed in detail here including simulated mining, resource/reserve definition and optimizing blending strategies. Models for these applications can effectively be built using the data transformation workflow presented (Figure 5.1). In the following chapter, a multivariate spatial model is built for a large copper-molybdenum porphyry deposit and simulated mining presented as an example application.

5.5 Discussion

This chapter has focused on the integration of spatial modeling workflows for a consistent multivariate modeling approach and the application of these workflows for value generation in a mine. The classical example of the quantification of uncertainty as a function of data collection or drill hole spacing was used as an example application of the techniques presented here. These models of uncertainty may be directly used to make decisions about the re-

quirement for additional data collection or the possibility to scale back on the amount of data collected to save substantial amounts on drilling costs. Making this decision requires a model which accurately quantifies the uncertainty; the simulation based workflows presented here may be used to generate these models.

Additional value generating applications of these models are possible, and are discussed further in chapter 6. A key application of high resolution spatial models of metallurgical properties is the evaluation of the potential variation in properties using a mine plan. Using the proposed mine plan, the daily variability in feed to the mill may be calculated and uncertainty quantified leading to better mill operation strategies. The mine plan could be refined to better manage feed variability for the mill. This application, and other applications are discussed for a copper-molybdenum porphyry case study.

Chapter 6

Case study: South American copper porphyry deposit

The geometallurgical mineral deposit characterization approach outlined in chapter 5 was applied for the spatial modeling of a large tonnage South American copper-molybdenum porphyry deposit. For this large tonnage deposit, quantifying mill throughput is critical. The optimal open pit mining operation would maximize both mill throughput and metal recovered. Therefore, a spatial model of geometallurgical properties including grinding indices and rock properties as well as copper and molybdenum grades is required. In this chapter, a multivariate spatial model of the porphyry deposit is constructed using supplied exploration data, validated and applied for quantifying uncertainty in expected daily mill properties.

6.1 Background

A major mining corporation provided 18.5 km of drilling for this study. The drilling data provided includes logs of lithology (13 distinct lithologies) and alteration (8 distinct alterations) where each alteration is classified by intensity as weak, medium or high (for a total of 24 alteration classifications). Assays of copper, molybdenum, silver, iron, sulfur, acid soluble copper and cyanide soluble copper were supplied on 2 m intervals. Interval data with rock quality designations (RQD), core recovery, point load test strength and fracture frequencies were also supplied. Metallurgical data in the form of comminution and whole rock analyses on nominally 30 m intervals were supplied. The

13 comminution indices include Bond mill work indices and semi-autogenous grinding (SAG) mill indices. Both the ball and SAG mill indices are relevant to the study as a mineral processing plant utilizing both mill types is planned. Whole rock chemical analysis including assays on 68 different elements is available on the same intervals as the comminution indices. The data were divided into two types: fine scale assay data and large scale comminution and whole rock analysis data.

Averaging all data up to the 30 m scale of the comminution measurements would be impractical and result in an unacceptable loss of information. Instead, data available on 2 m intervals were composited to 15 m lengths (bench scale compositing) and modeled separately from the 30 m intervals for the first part of this study. To demonstrate the application of downscaling, the critical semi-autogenous grinding index Axb was downscaled to the 15 m lengths and modeled at this scale. Only exploration diamond drilling is available for this deposit. Supplied drilling is spaced approximately 100 m apart with a -70° inclination at a bearing of 270°. An oblique view of the data configuration is shown in Figure 6.1. Locations for the 15 m composited samples are shown in the figure.

6.2 Model purpose and study aims

A high resolution geometallurgical model of critical comminution indices, geotechnical measurements and metal grades is required to make optimal decisions for this deposit. Drilling data including logs of lithology, alteration, assays of copper, molybdenum and silver (along with other elements), and comminution tests of the semi-autogenous grinding mill indices (such as Axb) and Bond mill work indices were used. Using geostatistical methods described in this thesis, a stochastic, multivariate model of all rock, geotechnical and geometallurgical properties of interest was constructed.

There are numerous applications of the high resolution multivariate model. For this thesis, two applications are considered. The first application is the quantification of uncertainty in the optimal ultimate pit using the Lerchs-

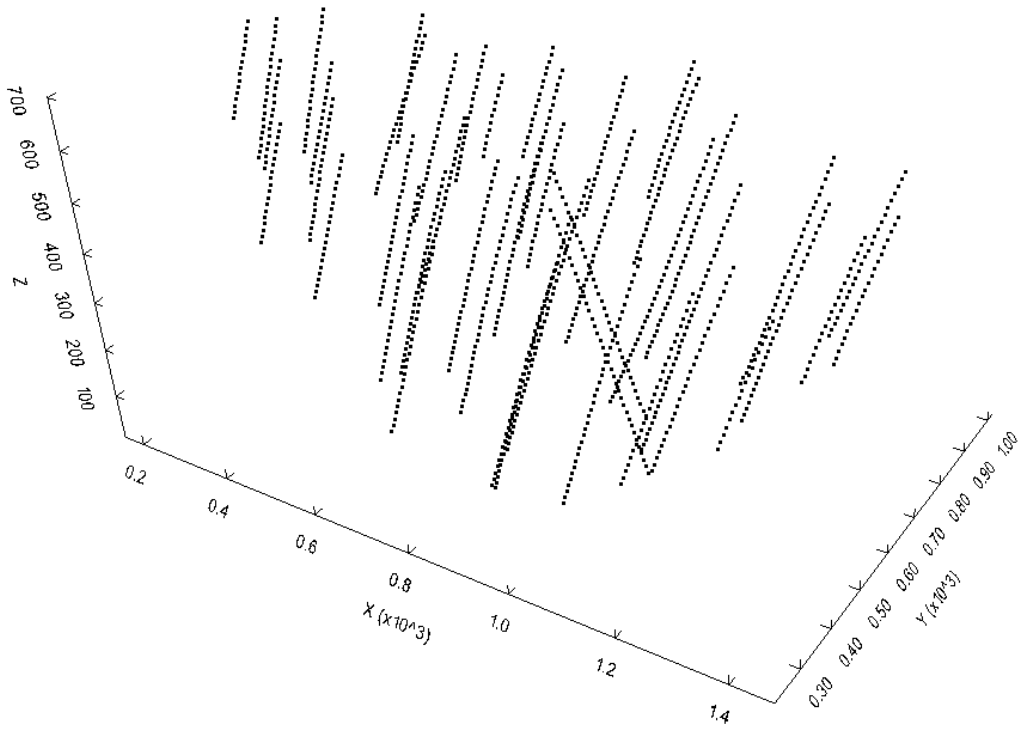

Figure 6.1: Oblique view of drill hole sampling locations with the 15 m composites locations.

Grossmann algorithm (Dimitrakopoulos, 1998; Lerchs and Grossmann, 1964). The Lerchs-Grossmann algorithm belongs to a set of graph theory algorithms for calculating the smallest maximum valued closure of a directed acyclic graph (Deutsch and Deutsch, 2013). In a mining context, given a block model with assigned economic values and a set of geotechnical constraints on pit slopes, the blocks which should be mined for optimal mine value are calculated. Due to the long periods of mining, changes in commodity prices and processing costs and continuous collection of information, the optimal pit calculated by Lerchs-Grossmann is unlikely to be the true ultimate pit, but is a reasonable approximation. For the purposes of this thesis, a synthetic economic model was created which included costs for crushing and grinding as a function of grinding indices and density, and prices for the final concentrate as a function of both copper and molybdenum grade.

The second application is the approximation of expected feed variability to the mill. Using a portion of the optimal pit from the Lerchs-Grossmann analysis, mining and classification of material as ore or waste, and subsequent routing to the mill for grinding may be approximated. Variation, and uncertainty, in the expected feed to the mill may then be calculated. In keeping with the ultimate pit analysis and stochastic methodologies employed in this thesis, the simulated mining adopts the modern geostatistical view of "all realizations all the time". All realizations constructed of the pit were processed using this procedure.

6.3 Modeling workflow

Stochastic simulation was chosen as the primary approach for modeling as it places no requirements on variable linearity and effectively quantifies the joint uncertainty among variables. A supersecondary approach was implemented here with sequential Gaussian simulation. This method requires homotopic sampling of all variables, so this approach was augmented with data imputation (Barnett and Deutsch 2014) to realize complete homotopic data sets. The supersecondary workflow that followed from incorporating data imputation is outlined in Figure 6.2. This workflow is generally applicable for multivariate modeling of continuous variables in the presence of missing data; however, additional considerations such as compositional variables would alter the workflow for different deposits. Details of the workflow follow.

Representative histograms were established for each of the variables by declustering, and the multivariate data univariate normal score transformed using the representative distributions. A sample correlation matrix for the normal score data was inferred, in addition to univariate variograms for each variable. This workflow requires a multivariate Gaussian assumption; data are assumed to be multivariate Gaussian after transformation. There are commonly violations of multivariate normality after the standard univariate normal score quantile to quantile transformation. Proceeding with the multivariate Gaussian assumption, simulating and checking the results is a reasonable approach.

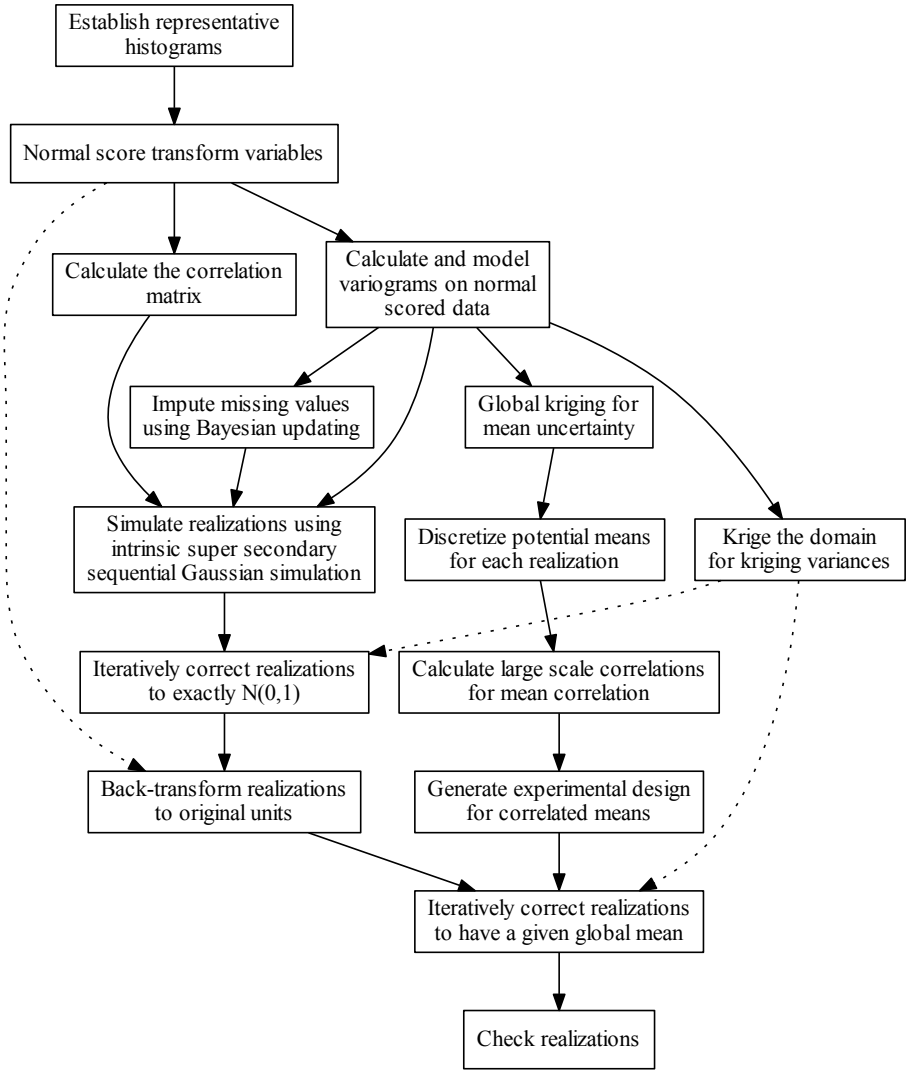

Figure 6.2: Flowchart of high-resolution supersecondary modeling methodology.

Methods involving the multivariate transformation to a multivariate Gaussian distribution such as the stepwise conditional transform and projection pursuit multivariate transform could be implemented as an alternative approach.

Missing data were imputed using the non-parametric multiple imputation approach of Barnett and Deutsch. Using the imputed data sets, sequential Gaussian simulation under an intrinsic supersecondary variable approach was applied for simulation with the program usgsim (Manchuk and Deutsch 2012).

The entire gridded volume was simulated, not just the grid cells located within the implicitly modeled domains. This was carried out to avoid boundary effect issues on the edges of the domain. A different data realization from data imputation was used for each realization. The simulated model over the entire volume was then clipped within the domain. For many variables, the variogram range was large relative to the size of the domain. This, combined with the blending of multiple variograms from the use of the supersecondary approach resulted in realizations with variances lower than 1. Back-transforming these Gaussian realizations to the original units and using these realizations would result in biases in the final models. For this reason, the simulated realizations were iteratively corrected to be exactly standard normal.

The iterative correction of each variable was done by kriging the domain in normal score units and using the kriging variances to weight the iterative normal score correction of each realization. Global kriging was used to avoid any search artifacts in the kriging variance from a limited search. Using the kriging variances to weight the iterative normal score correction enforces precise local data reproduction, since the kriging variance for nearby cells is approximately zero. The result of this iterative correction is a suite of exactly standard normal realizations which can be back-transformed to original units (grades).

The back-transformed realizations will match the declustered histograms by construction; however, there is no variation in global statistics such as the mean of the realizations. The estimate of the global mean within each domain is uncertain and this uncertainty should be reflected in the simulated realizations. Uncertainty in the mean for each variable was determined by global kriging of the domain with the data. The distribution of possible mean values from the calculation of the global kriging variance was discretized and evenly spaced quantiles corresponding to mean values taken from this distribution. This assumes that the distribution of mean values is Gaussian, which is a very reasonable assumption from central limit theorem.

The multivariate relationship of the mean grades must be accounted for when correcting the realizations to have the correct mean uncertainty distribu-

tions. Using perfectly correlated quantiles would result in final uncertainties which were much too high and assuming uncorrelated means would understate the global mean uncertainty. Large scale correlations between variables were determined by averaging each data value with a large number of nearest neighbours using a moving window approach. As these calculated correlations are on a much larger scale, they approximate the true global mean correlations. Latin hypercube sampling was used to generate an experimental design used with the quantiles to sample the multivariate mean uncertainty correctly. This approach results in the correct global mean uncertainty with the correct correlations across realizations on average and, therefore, the correct final multivariate mean uncertainty distributions. Global means were considered uncorrelated across domains as there was no geologic reason to support the use of correlated means.

6.3.1 Nonlinearity and multiple data scales

No explicit correction for nonlinear metallurgical variables is made in this workflow. This is due to limited information; no experimental data on nonlinearity was available for the metallurgical variables and the correlations between large scale metallurgical variables to small scale linear variables were too low to accurately infer a nonlinear relationship as shown in chapter 3. Simulation was used which makes no averaging assumption; however, all block averaging to larger scales assumes linearity. For nonlinear variables, this will introduce a bias into the large scale estimates which may not be quantified without additional experimental measurements to infer nonlinearity. If additional experimental data are collected, or additional measurements which support a strong multiscale relationship are made, then the nonlinear modeling strategies discussed in chapter 3 may be applied.

In addition, the multiple data scales are not explicitly addressed in this workflow. Recall the two primary scales of data: small scale assay measurements which are composited to the bench (15 m) scale, and large scale metallurgical measurements on 30 m intervals. In the presented workflow, data are required to be at a single scale. Therefore, in the absence of a downscaling

procedure the two data scales are modeled separately, and the assumption is made that the change in variance is minimal between the scales. The workflow was carried out with this assumption, and additionally using the downscaling developments in chapter 4 for the critical Axb parameter. The impact of the downscaling may then be observed in the simulated mining section of the case study.

6.3.2 Multivariate stochastic methodologies

The intrinsic supersecondary model was chosen as the multivariate stochastic simulation method for this case study. Under the intrinsic model, covariances are assumed to be scaled variograms rather than requiring an explicit model of coregionalization such as the linear model of coregionalization. Adopting the intrinsic supersecondary approach also makes the strong assumption of multivariate normality. Using a test of bivariate normality (Deutsch and Deutsch, 2011), the bivariate distributions between the assay variables were checked. Of the 45 scatterplots, 9 failed the test for bivariate normality. This is a substantial percentage which deviated from the assumption of normality, so alternative methods including multivariate transformations were examined as alternative modeling strategies.

Early tests with multivariate transformations effectively removed all non-normal features of the bivariate relationships, but simulated realizations had poor variogram reproduction and transformed variables had very limited spatial continuity. The intrinsic supersecondary methodology was adopted in place of the multivariate transforms for this reason, and cross plots after simulation checked. Future research in the area of multivariate transforms could be employed to improve variogram reproduction and permit the usage of multivariate transforms.

6.3.3 Mixed continuous, categorical variables

The supplied alterations and extent of alteration may be viewed as a mixed continuous, categorical variable. Alterations were supplied as alteration type, such as potassic-feldspar alteration, classified as weakly, moderately or in-

tensely altered. These mixed continuous, categorical alterations were used for domain definition as grades and metallurgical properties corresponded very strongly with degree and type of alteration for this porphyry deposit. High degrees of alteration, particularly potassic-feldspar, and to a lesser extent potassic-biotite, were a strong indicator of high copper and molybdenum mineralization. These rules were used in domain definition for grouping dissimilar lithologies which shared a high degree of alteration.

6.4 Parameter and data definition

Geometallurgical modeling domains (rock types) were established using a multiple signed distance function approach. These modeling domains were used in the choice of stationary statistics including variable distributions, multivariate relationships and variograms. The multiple lithologies and alterations supplied with degree of alteration are too numerous to model separately. Modeling domains were established by grouping lithologies and alterations using a defined set of domain definition rules. The domain definition rules include:

1. Rock types to be grouped together should be geologically linked or similar
2. Rock types to be grouped together should be statistically similar
3. Rock types to be grouped together should be spatially similar; the data types would ideally occur close to each other in the deposit and drill hole logs
4. Rock types with a very low prevalence over the deposit should be considered as candidates for grouping with other rock types
5. Finely layered rock types should be considered as candidates for grouping with other rock types to account for our ability to discriminate when estimating with limited data
6. Each domain should have enough data to reasonably estimate modeling statistics including statistical distributions and variograms

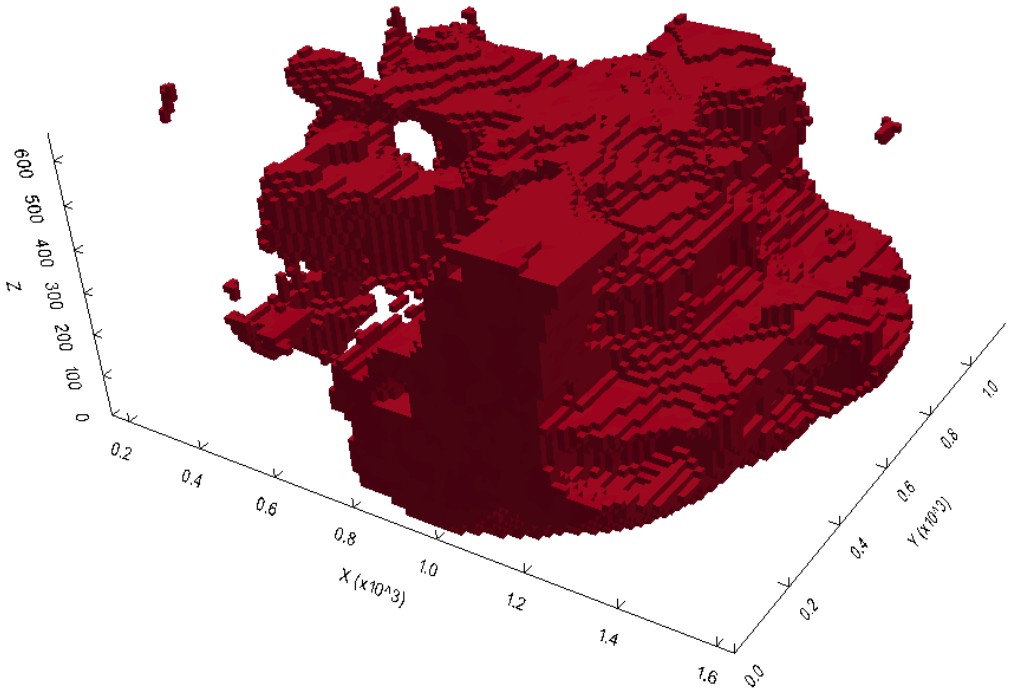

Figure 6.3: Oblique view of domain associated with highly altered, high grade region and immediately surrounding host rock. The view is clipped to only show the high grade domain.

Multiple category signed distance functions were used for domain modeling over the area of interest using a servosystem correction to enforce reproduction of the global proportions (Silva and Deutsch, 2013). Modeled blocks were 15 m x 15 m x 15 m corresponding to approximately 8600 tonnes on average. The domain model was further refined by clipping the model to any points greater than 300 m from the nearest drill hole sample. Three domains were modeled. This avoids emphasizing our ability to model these exactly, but these rock domains could be refined further with additional geologic knowledge. The largest and economically most important domain modeled (approximately 50% of the deposit) is shown in Figure 6.3. This domain will be used to show statistics and statistic reproduction for the modeling workflow.

Within each domain, representative distributions of each variable were estimated by declustering. Cell declustering was chosen as it is a common best-practice technique and simple to apply to any three dimensional data set.

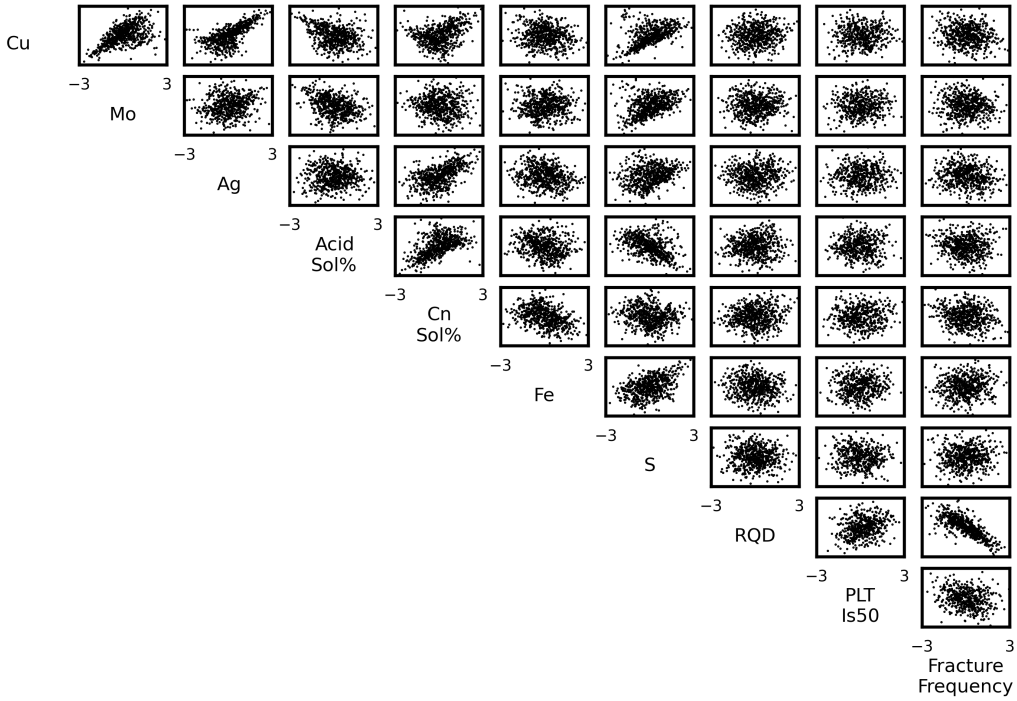

Figure 6.4: Scatterplot matrix of the normal score transformed assay scale measurements.

Declustering was applied to all variables and the bivariate relationships examined. A matrix of the normal score transformed bivariate scatterplots of data composited to 15 m for the domain of interest is shown in Figure 6.4. Many of the variables are uncorrelated, but there are some strong relationships among copper, molybdenum and silver. As expected, RQD and fracture frequency are also highly correlated.

Using the declustered distributions as representative histograms, variables were normal score transformed. The correlation matrix and normal score variograms were calculated. Missing data are imputed using nearby data of the missing type and other correlated variables. These imputed data sets are stochastic, and a different data set is used for every simulation (Barnett and Deutsch, 2015). The imputed variables have the correct histograms and spatial correlations for simulation. Data values were imputed using Bayesian updating with the correlation matrix calculated using the heterotopic data values. Although there is no requirement that the heterotopic correlations be posi-

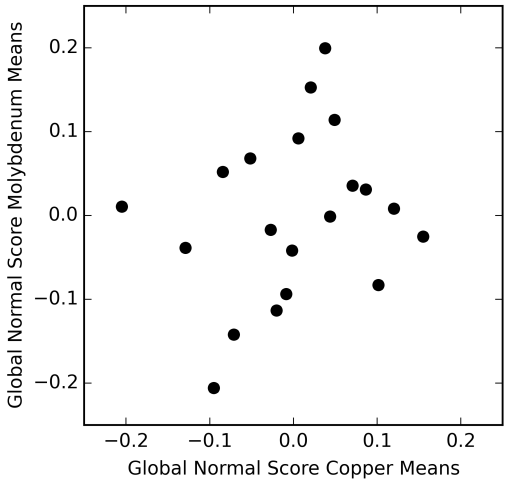

Figure 6.5: Correlated global means for copper and molybdenum drawn with Latin hypercube sampling with a global correlation of 0.22.

tive definite, the correlation matrices in these cases were positive definite and could be used without projection to a nearby positive definite correlation matrix. The implementation of Barnett and Deutsch was used with the normal score transformed data values to generate data realizations for each simulation.

6.5 Model execution and checking

The modeling workflow was executed and post-processed to enforce correctly correlated global means with uncertainty. An example of the correlated normal score global means for copper and molybdenum is shown in Figure 6.5. The samples were drawn with correlated Latin hypercube sampling for approximately uniform sampling across the constructed realizations.

Final realizations were clipped to not extrapolate far away from available drilling or substantially below the bottom of drill holes. Smooth e-type models were constructed by averaging all realizations on a block-by-block basis. The simulated realizations were assembled and visualized in 3D and with a number of cross sections as an initial check of the model. An oblique view of simulated Bond mill work indices for the domain of interest is shown in Figure 6.6. The simulated distribution of hard and soft rock is reasonable. Cross sections of a sample realization and the e-type mean for the normal

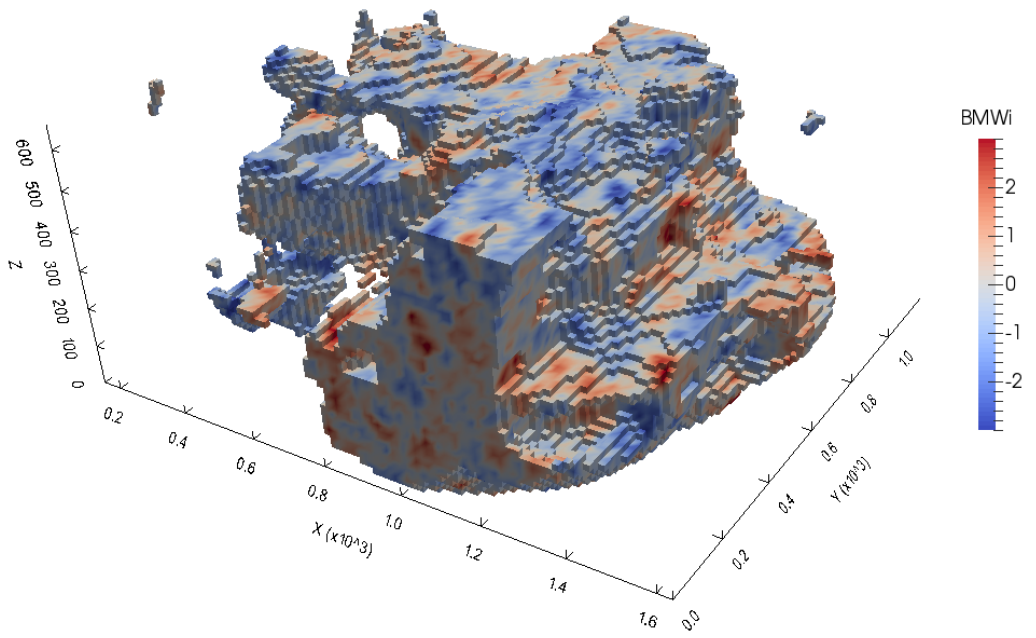

Figure 6.6: Oblique view of normal score Bond mill work index realization.

score transformed copper grade are shown in Figure 6.7. As hard boundaries were considered for the simulations, and the mean grades and hardness vary significantly among the rock types, the domain boundaries are visible in both the simulated realizations and e-type means.

The models are visually reasonable given knowledge of the deposit, but there are many statistics which could be checked. With the histogram transformation workflow adopted, histograms of the simulated realizations should exactly match the shape of the reference distribution and have global means which follow the expected global mean variance. This is observed for all variables; histogram reproduction for copper grade is shown in Figure 6.8 as an example. The reference histogram in red is exactly in the center of the simulated realizations.

The reproduction of the bivariate distributions was also assessed. Here, bivariate scatterplots for a sample realization are shown in Figure 6.9 for the composited variables. The overall relationships are similar to the original scatterplots in Figure 6.4, but there are some notable differences. The wedge shaped distribution between copper and molybdenum grades and the strong

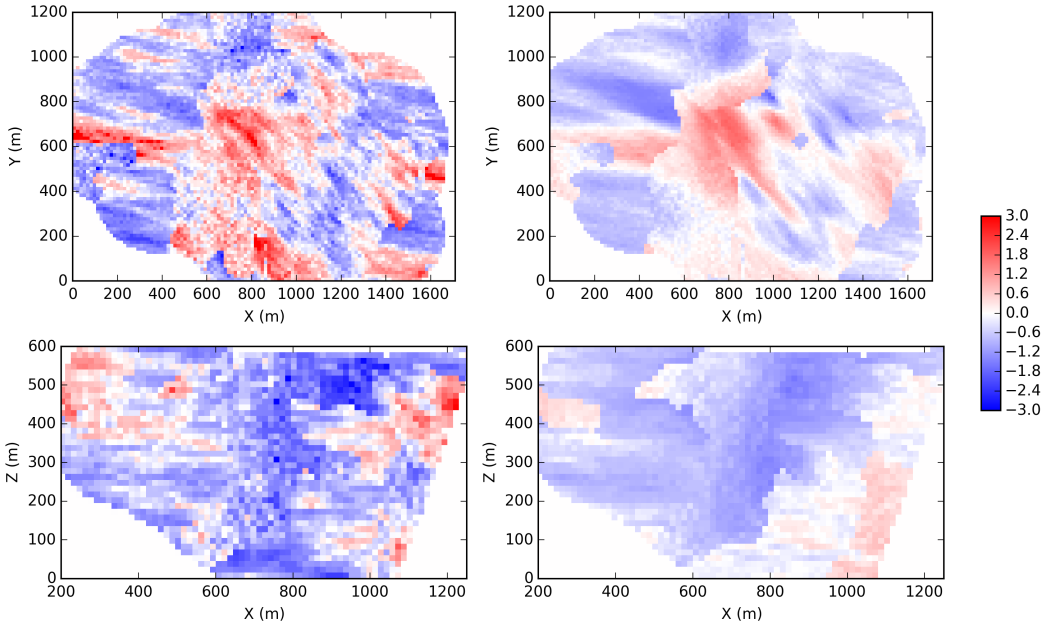

Figure 6.7: Plan view (upper) and cross section (lower) of a sample realization (left) and e-type mean (right) for normal score transform copper.

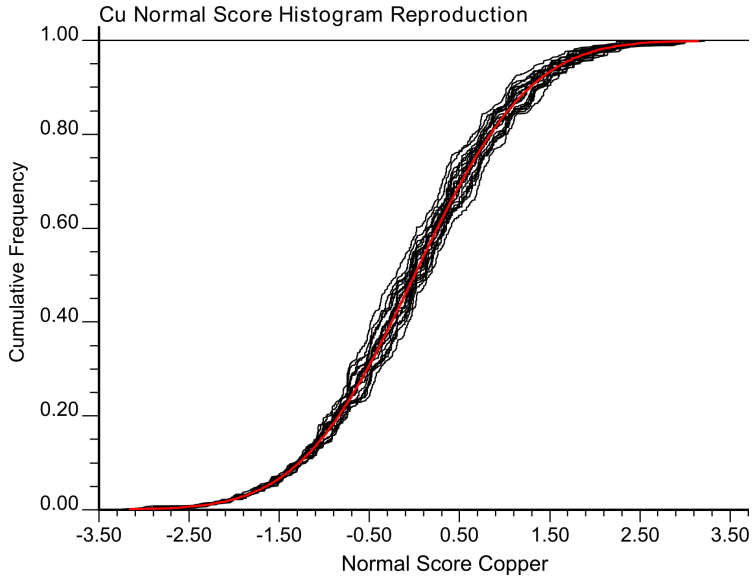

Figure 6.8: Histogram reproduction for normal score transform copper for the domain of interest. The correct histogram (the univariate normal distribution) is overlaid on the multiple black histograms of the realizations as a red line.

relationship between RQD and fracture frequency are not exactly reproduced, although the correlations are correct. This is due to the use of the supersecondary method for simulation; these relationships are not explicitly controlled. Other bivariate relationships are well reproduced. Methods such as the step-

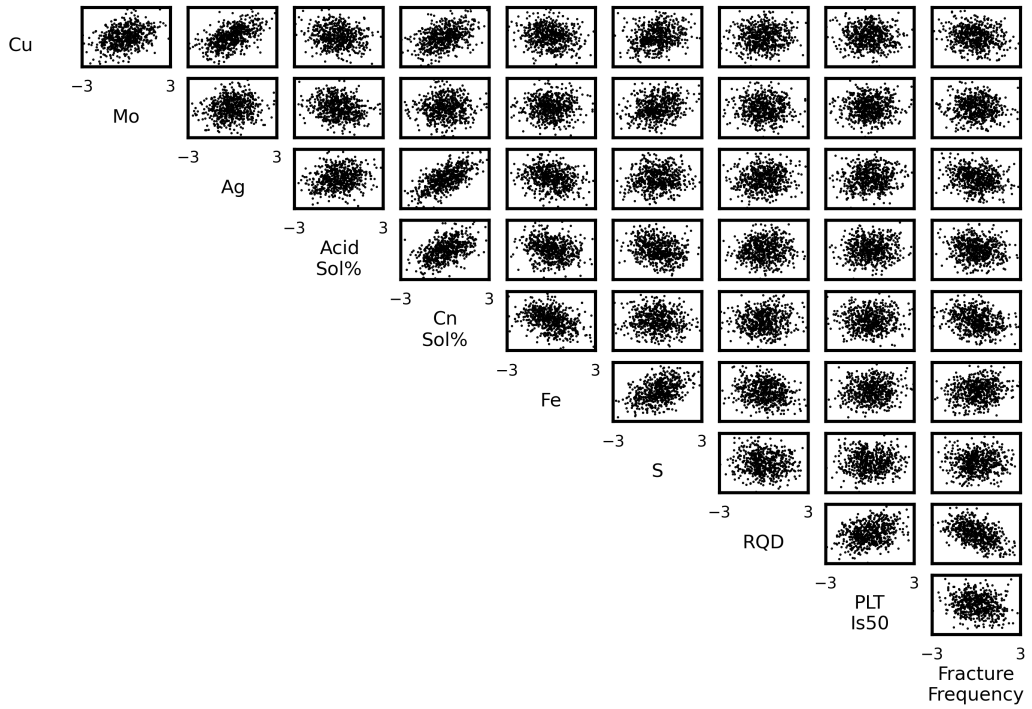

Figure 6.9: Bivariate reproduction matrix for the domain of interest. A random sample of 2500 simulated blocks was used for this matrix.

wise conditional transform or projection pursuit multivariate transform could be used as alternatives which explicitly control the bivariate relationship.

Variogram reproduction was assessed in both normal score and original units. Variogram reproduction in normal score units for copper and the bond mill work index in normal score units is shown in Figure 6.10. Short range variogram reproduction is reasonable although the simulated long range structure is too continuous for copper compared to the true variogram. This is attributed to the blending of multiple variograms when using a multivariate methodology which assumes an intrinsic cross variogram, not an explicit model of coregionalization. Other model checks including correlation coefficient calculation and ensuring exact data reproduction were also conducted. Data were exactly reproduced and, due to the nature of the supersecondary simulation method, correlation reproduction was within 1%.

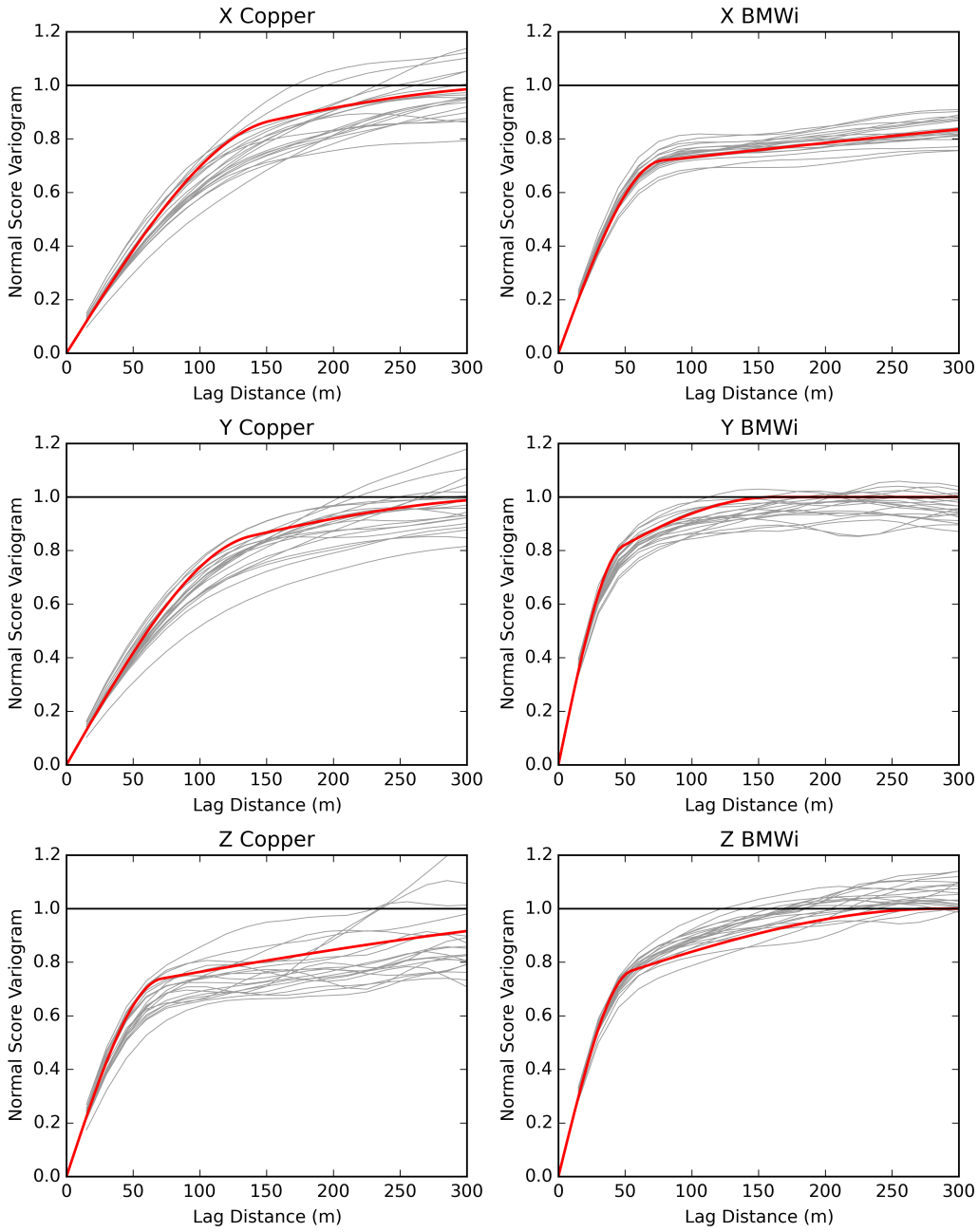

Figure 6.10: Variogram reproduction for copper and the bond mill work index for the domain of interest. The normal score variogram model is shown in red as well as 20 variograms from the transformed realizations.

6.6 Model post-processing and application

Using the verified models, a number of applications were considered to quantify uncertainty relevant for mine and mill operation. Using the Lerchs-Grossmann

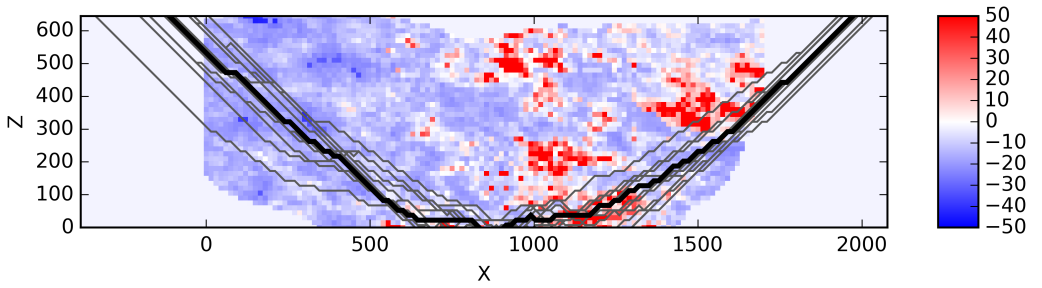

Figure 6.11: Multiple realizations of the ultimate pit applied to different realizations (grey). The 50% probability contour to be contained within the ultimate pit is also shown (bold black). These realizations are overlaid on a slice through the center of the deposit of the deviation from the global mean of the semi-autogenous grinding index Axb. The model is trimmed by an average Axb, as this material would not be processed.

algorithm and a simplified economic model which penalizes low-throughput ore corresponding to low Axb values and high Bond mill work indices, the ultimate pit was estimated for 20 realizations. A number of these pits are shown in Figure 6.11. Using these pits, the probability that a block would be optimally mined in 50% of the pits was contoured to provide a simplified measure of the optimal pit on average. A realization of the deviation of the semi-autogenous grinding property Axb from the global mean is shown here as well. This analysis can be used to quantify the uncertainty in the tonnage of the deposit, and expected variability in all properties including metallurgical properties such as Axb.

Within the ultimate pits, substantial regions of high Axb corresponding to less energy intensive ore and low Axb (more energy intensive ore) are observed. The temporal variation in Axb was examined as a second application which analyzed an intermediate pit expansion over a period of 300 operating days. Using a planned extraction sequence applied to 20 high-resolution realizations, the average daily Axb was calculated for each day of the intermediate mine plan, for each of the realizations (Figure 6.12). The expected deviation from the average Axb provides a measure of the throughput which can be achieved and power requirements. Using all realizations simultaneously, the daily expected Axb can be calculated (indicated on the figure) as well as the expected 80% probability interval (indicated on the figure). This analysis indicates that

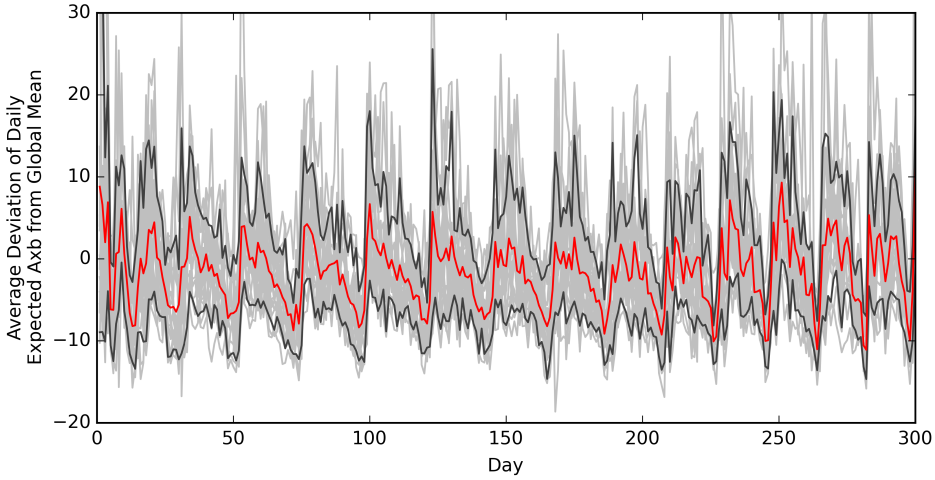

Figure 6.12: Average deviation of the daily expected Axb from the global mean based on the simulated mining of 20 realizations (light grey lines). The 10th and 90th percentiles (black lines) and daily expected value (red line) are overlaid.

there is substantial uncertainty in the daily characteristics of ore processed by the mill, and that the Axb of the ore will vary substantially on a daily timescale.

This model of Axb was constructed under the assumption of linear averaging due to limited experimental data and a lack of highly correlated secondary small scale measurements, and additionally using the 30 m measurements of Axb, not measurements on a 15 m scale. Under an assumption of linearity, the Axb parameter may be downscaled to the 15 m scale using the procedure proposed in chapter 4.

6.6.1 Comparison with downscaling

The daily uncertainty in Axb modeled using the comminution variables at the original 30 m scale was compared with Axb downscaled using the procedure proposed in chapter 4. Experimental and modeled variograms of the original 30 m measurements of Axb are shown in Figure 6.13. The standardized correlogram was used in place of the traditional experimental variogram for variogram inference due to the highly skewed Axb distribution with values from 20 to over 200.

Assuming linearity, as there was not enough experimental evidence to infer

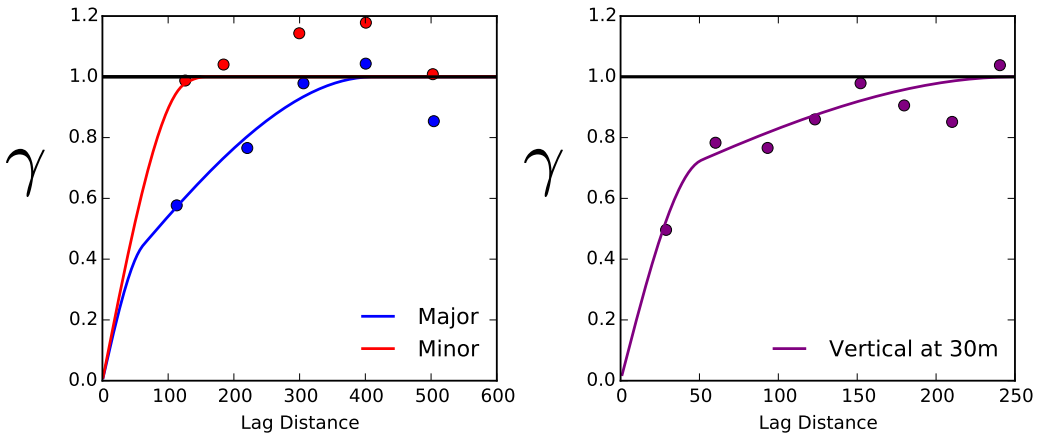

Figure 6.13: Experimental and modeled variograms of 30 m measurements of Axb. The correlogram was used to infer the variogram.

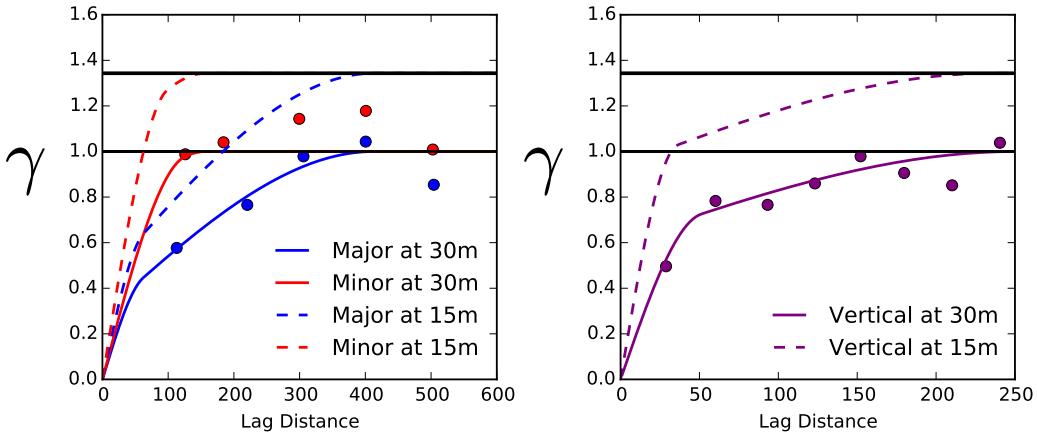

Figure 6.14: Downscaled variograms of the Axb at 15 m. Downscaling resulting in an increase in variance of 30%.

a nonlinear averaging schema, the variogram was downscaled using the downscaling laws (Figure 6.14). The downscaled variogram resulted in an increase in variance of 30% from the 30 m measurements to 15 m measurements due to the short range of the vertical variogram. This is a substantial increase in variance, and is expected to substantially increase the observed average deviation of daily Axb feed compared to Figure 6.12.

The univariate distribution of Axb (histogram) was downscaled adopting the discrete Gaussian model with the mosaic model (Figure 6.15), as the distribution of Axb is skewed, but not near a constraint. The downscaled histogram is stable and reproduces the increase in variance of 30%.

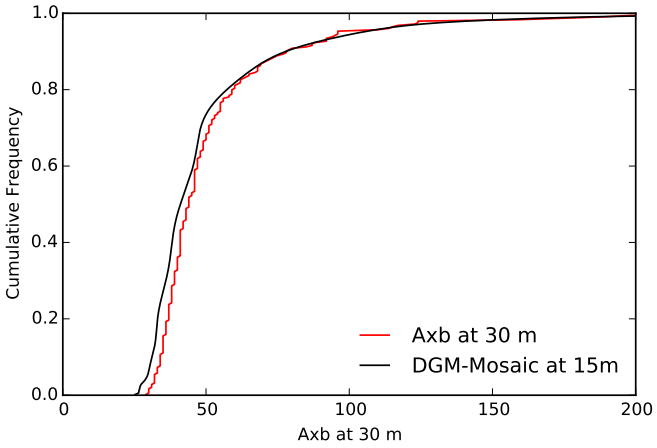

Figure 6.15: Downscaled 15 m histogram of Axb using the discrete Gaussian model with a mosaic model.

Using the direct sequential simulation procedure illustrated in chapter 4, measurements of Axb were stochastically downscaled. As the Axb measurements are not near a constraint, only limited resimulation was required for the downscaled measurements to meet the required conditions. The downscaled Axb realizations were used in the simulation workflow and the simulated mining procedure repeated for the same area. The resulting realizations of Axb, shown in Figure 6.16, show the effect of the substantially increased variance. Trends from the 30 m calculation are still visible in the 15 m plot, although the exact conditioning near data, such as at approximately 27 days is reduced due to the stochastic approach to downscaling.

Without measurements of Axb on 15 m intervals, and without operational data to reconcile the models, the variance increase resulting from the use of downscaled Axb data realizations may not be validated. In addition, potential nonlinearity in Axb has not been modeled, and the variogram downscaling approach which supports the 30% increase in variance assumes linearity. If the Axb parameter is highly nonlinear, then the results would be expected to change. However, the downscaling methodology presented in chapter 4 is a statistically consistent methodology for modeling multiscale metallurgical measurements, and has been demonstrated in cases such as the oil sands bitumen grade example to be a reasonable model for multiscale modeling.

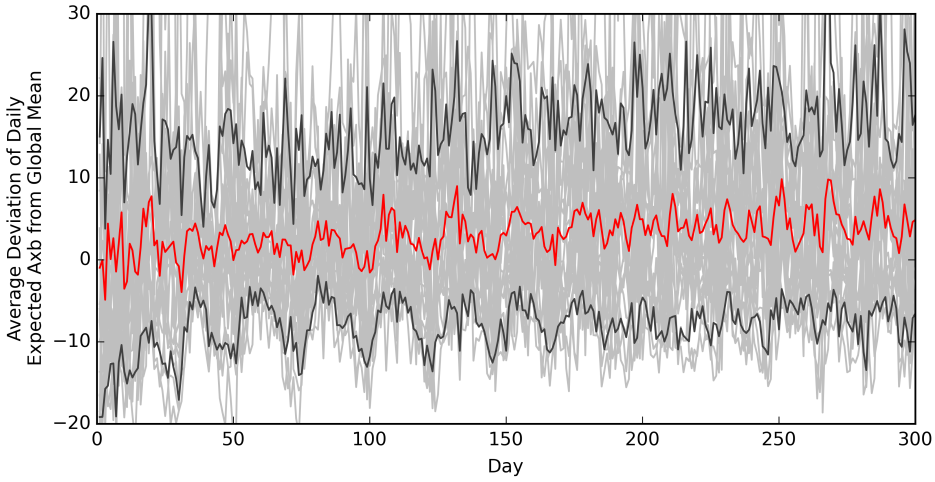

Figure 6.16: Average deviation of the daily expected Axb from the global mean based on the simulated mining of 20 realizations (light grey lines). The 10th and 90th percentiles (black lines) and daily expected value (red line) are overlaid using the 15 m downscaled data.

Therefore, even in the absence of validation data for this deposit, the increase in variance is supported by the Axb variogram and should be used for planning instead of the original variance at 15 m.

6.7 Discussion

A comprehensive methodology for geometallurgical mineral deposit characterization has been presented and applied to a South American copper-molybdenum porphyry deposit. Multivariate geostatistical simulation using an intrinsic supersecondary approach was used to create high-resolution realizations of the deposit which can be used to quantify the uncertainty and optimize mining and processing decisions. Stochastic data imputation was applied to impute missing values and allow for the application of full multivariate techniques. Using stochastic data imputation and a simulation workflow imposes no requirements for linearity on the variables during modeling. All averaging operations are deferred until the end of modeling where the nonlinear behaviour of variables can be explicitly accounted for. In addition to these techniques, there are a number of methodologies including compositional data analysis and size distribution modeling which are very relevant for geometallurgical modeling,

but were not reviewed as these features were not measured in the case study deposit.

The multivariate supersecondary approach used has a number of limitations, primarily in the handling of multivariate data which are not multivariate normal after univariate Gaussian transformations. In the case study presented, these deviations were relatively few. For mineral deposits with substantive deviations from multivariate normality, multivariate normal transformations such as the stepwise conditional transform (Leuangthong et al. 2006) or projection pursuit multivariate transform (Barnett and Deutsch 2012) could be applied. These transformed variables may be decorrelated using an algorithm such as min/max autocorrelation factors (Switzer and Green 1984). The advantage of these approaches over the supersecondary approach applied is better reproduction of the multivariate histogram and constraints among variables. The primary disadvantage is the increased model complexity and increased challenges with input model statistic reproduction associated with a large chained transformation workflow.

Using this workflow, two applications were demonstrated in quantifying the uncertainty in the optimal ultimate pit, and estimating the uncertainty in the metallurgical properties of ore fed into the mill on a daily basis. These applications demonstrate the value in the high-resolution geometallurgical modeling approach. Multiple realizations which correctly reproduce the spatial and multivariate relationships among variables are required for both of these applications.

Applying the downscaling procedure for the Axb samples, which are critical to predicting ore throughput in semi-autogenous grinding, the variance of Axb increased by 30%. This large increase in variance resulted in a substantially high daily mill feed variability relative to calculations done using the same procedure, but with 30 m samples. Without data at 15 m, the proposed increase in variance for daily throughput cannot be validated for this deposit. However, the downscaling approach proposed is a statistically consistent approach which has been validated on other data sets, such as the oil sands

data shown here. Use of the downscaling approach is therefore supported for planning with the increased deposit variability.

Due to limited data, nonlinearity in metallurgical variables was not explicitly accounted for in this case study. The weakly correlated small scale samples were not enough to infer nonlinearity for this deposit using the semiparametric Bayesian updating approach proposed in this thesis (chapter 3). Adopting the sequential Gaussian simulation approach defers averaging for the nonlinear variables; however, upscaling the realized models requires an assumption of linearity. As additional data are collected for this exploration project, the nonlinearity could be modeled and the spatial model presented refined with the addition of a nonlinear transform step.

The comprehensive multivariate spatial modeling methodology applied in this case study demonstrates the value of the proposed integrated approach. Only two applications, optimum pit uncertainty calculation and expected daily variability in mill feed, were demonstrated in this case study, yet they clearly demonstrate the value inherent in a complete geometallurgical approach. Expected throughput, approximated using grinding indices, can be accounted for when calculating uncertainty in the ultimate pit. The expected daily variability in daily ore grinding indices can be accounted for when designing the mill and during short and long term mine planning. Many other applications including data spacing analysis and resource/reserve definition are possible.

Chapter 7

Conclusions

This final chapter summarizes, discusses the limitations of, and proposes future work for the developments in this thesis. This thesis makes three primary contributions to the field of multivariate spatial modeling of geometallurgical properties. Techniques are developed in this thesis for the inference and modeling of nonlinear variables, the downscaling of large scale metallurgical composites are integrating these steps into data transformation workflows. The transformation workflows developed incorporate a series of transforms designed to remove compositional, nonlinear, missing data and multiscale challenges associated with metallurgical data. These contributions were motivated by the case study of a large copper-molybdenum porphyry deposit. Each of these three contributions are briefly reviewed here.

7.1 Nonlinear re-expression

Many metallurgical variables, such as the Bond mill work index and froth flotation recoveries average nonlinearly. This nonlinear behaviour is due to complex chemical-physical interactions, such as ore holdback in a ball mill or froth destablization. Even when ore is fed to the comminution circuit in 400 tonne batches, a very large amount of mixing occurs in a mineral processing circuit emphasizing the importance of modeling nonlinear relationships to avoid a bias in predicted behaviours.

The heart of the proposed methodology for nonlinear modeling is to re-express nonlinear metallurgical variables as linear variables which may be

blended, scaled and modeled using available linear techniques. After modeling, the re-expressed variables may be back-transformed to the original metallurgical values. The power law transform used throughout this thesis is a very flexible re-expression framework, suitable when relationships are monotonic and supported if the metallurgical property meets a set of physical requirements as discussed in chapter 3. If the metallurgical property exhibits substantial synergistic or antagonistic behaviour resulting in a non-monotonic relationship, then alternate re-expression frameworks such as Scheffé polynomials may be used instead. Nonlinear re-expression laws are most easily fit when direct experimental results of blended ores are available, but it may be possible to fit the nonlinearity in specific circumstances where direct evaluation is not available.

In the absence of direct experimental evaluation, multiscale measurements are required. Single scale measurements do not provide enough evidence for a nonlinear relationship, as demonstrated with a Monte Carlo simulation study. Given sufficient multiscale measurements of the variable of interest, nonlinear regression may be used to fit the relationship. The effect of measurement error on nonlinear regression was demonstrated to be problematic, particularly for highly nonlinear variables. If the measurement error is known, then it was demonstrated to be possible to remove attenuation in the model fit with error using a simulation study.

Alternatively, using multiscale measurements where a highly related small scale linear variable and large scale metallurgical variable are available, it may be possible to infer the nonlinearity. The proposed approach is to apply semiparametric Bayesian updating with nonlinear regression in these cases. This technique leverages the relationship between the correlated small scale linear variable and the large scale metallurgical variable. Semiparametric Bayesian updating uses the nonparametric fit to the bivariate distribution between the small scale variable and the large scale variable and the parametric normal equations to infer conditional distributions. Using Monte Carlo simulation, values are drawn from the updated distributions and fit with nonlinear regres-

sion. The result is a flexible technique for multiscale nonlinear inference, but with a number of limitations.

The primary limitation of the proposed semiparametric Bayesian updating technique is the requirement for a very strong relationship to infer nonlinearity. In the case of the Bond mill work index presented in chapter 3, the relationship with iron content was not strong enough to reliably infer a nonlinear relationship. This limitation is additionally increased by the requirement to downscale required statistics including the variogram and bivariate distributions unless sufficient small scale measurements are available. This is primarily a time-cost for modeling, which is made relatively easier with semi-automatic fitting tools.

Future work in the area of nonlinear re-expression is two fold. The first area of research would be extending methods for nonlinear inference with limited data. Research in the field of small scale correspondence measurements (Kuhar et al., 2011) is very promising for future research efforts in the inference of nonlinear behaviour. Small scale correspondence measurements are effectively multiscale measurements of the nonlinear variable where the error has been quantified. The application of disattenuation strategies discussed in this thesis are very promising for this area and would be a logical area of future work.

Related to this area of research is the use of a probabilistic re-expression framework. Given sufficient measurements of nonlinear relationships, uncertainty in the regression model could be accurately quantified and used. Possible methods for the probabilistic re-expression include transformations evaluated using a probabilistic conditional distribution or multiple regression models. These probabilistic methods will require additional experimental data, however, they could be very useful for accurately quantifying uncertain nonlinear relationships.

The second area of future nonlinear re-expression research is further investigation of highly antagonistic and synergistic relationships. This is particularly relevant for froth flotation where trace contaminants may substantially change froth properties and recovery. In this thesis, the focus has been on capturing all information on blending behaviour in a single re-expressed variable which is attractive for spatial modeling, but situations such as these might be better

served with the introduction of an auxiliary variable to represent the degree of contamination.

7.2 Downscaling large scale metallurgical composites

Multiscale metallurgical measurements pose a spatial modeling challenge as most geostatistical algorithms and implementations require data at a single scale. The proposal in this thesis is to stochastically downscale large scale metallurgical samples to the modeling scale. The downscaling procedure proposed is direct sequential simulation with intrinsic cokriging and dynamic resimulation. Intrinsic cokriging leverages bivariate relationships with correlated small scale variables, and dynamic resimulation ensures that constraints such as non-negativity are met.

The intrinsic assumption is both a strength and limitation of the algorithm. Adopting the intrinsic model is a strength as the number of cross variograms which must be downscaled is reduced and there is no requirement for an explicit linear model of coregionalization. This is also a limitation, as the spatial model of coregionalization used will be incorrect in cases where variograms substantially differ for coregionalized variables. Similarly, the dynamic resimulation method effectively enforces constraints during simulations, but will introduce an order-bias for variables near a constraint as demonstrated with bitumen grades. This order-bias is unlikely to be a cause for concern across all realizations together, but is an undesirable feature of the algorithm.

Future work in the area of downscaling large scale metallurgical composites will focus on streamlining the procedure. The stochastic downscaling algorithm requires downscaling the variogram, histogram, and if using multivariate data, cross variograms. If there are numerous metallurgical variables of interest to be downscaled, this poses a substantial barrier to practical usage. This procedure could be streamlined with semi-automatic implementations which are checked for internal consistency, such as positive-definiteness in the downscaled correlation matrix. In addition, the iterative procedure adopted for downscaling bivariate distributions was effective for the included exam-

ples, but has no theoretical basis so requires further development and research before widespread application.

7.3 Geometallurgical mineral deposit characterization

Specific challenges for spatial geometallurgical modeling of nonlinear and multiscale variables were covered in chapters 3 and 4, while chapter 5 focused on the integration of these techniques into a general spatial modeling workflow and the selection of techniques. The general data transformation workflow was applied in chapter 6 for the spatial modeling of a large copper-molybdenum porphyry deposit.

The data transformation workflow is subject to the standard geostatistical limitations, principally the requirements for stationarity, assumptions of multivariate Gaussianity as a covariance model, and usage of two-point statistics. Within these limitations, it is possible to check and verify models with cross validation, production data and conformance to modeling assumptions. This minimum criteria approach which was applied in the case study can be used in the decision to rely on a model.

The integration of parameter uncertainty, including histogram, correlation, and variogram uncertainty, is an area of future research with substantial potential for geometallurgical mineral deposit characterization. Further research into the selection of methods for integrating parameter uncertainty, and critically evaluating the effect of parameter certainty inclusion is required. This is a general problem applicable to all areas of geostatistical research, but is particularly relevant for the spatial modeling of metallurgical variables where there is often very limited data.

7.4 Software

Software is important to geometallurgical modeling; it is the tool which enables the construction of large 3D multivariate models which account for the many challenges discussed. A number of specialized programs were developed as part of this thesis for nonlinear modeling and multiscale modeling, and to facilitate

the construction of these models. Major software developed as part of this thesis are listed here and are available from the author.

- Python package for implementing the integrated workflow, *pygeostat*, which is a mixed fortran-python package for integrating geostatistical software with python, paraview, GSLIB and scripting workflows together in an easy to use, parallel framework. Nonlinear inference and statistics downscaling are bundled with this package which is compatible with Python 3.4 and higher.
- Nonlinear fitting package, *spbusim*, which performs semiparametric Bayesian updating for fitting with any nonlinear regression software for nonlinear inference.
- Downscaling software, *datascale* and an updated version of *histscale* for downscaling histograms and data using the workflow in chapter 4.

7.5 Final comments

Recall the thesis statement proposed in chapter 1: *An integrated statistical approach for the multivariate spatial modeling of metallurgical rock properties which accurately quantifies the joint uncertainty in grade, geotechnical and metallurgical properties will lead to more effective mine and mill operation strategies.*

In this thesis, an integrated data transformation workflow was proposed for the multivariate spatial modeling of metallurgical rock properties with geostatistical techniques. The integrated approach, applied to a copper-molybdenum porphyry deposit, is viable for many applications including data spacing analysis, quantification of uncertainty in the ultimate pit and simulated mining operations for expected mill feed variability. Two key challenges of metallurgical variable modeling were addressed: modeling nonlinear metallurgical variables and the integration of multiscale data. The third contribution, the integration of these strategies for a comprehensive geometallurgical spatial modeling workflow, completes this thesis.

A re-expression framework, permitting the modeling and usage of nonlinear metallurgical variables and a novel semiparametric Bayesian updating approach for nonlinear inference were developed. The nonlinear inference method was demonstrated to be reasonable for highly related multiscale samples, but requires a strong relationship. A multivariate direct sequential simulation with intrinsic cokriging and dynamic resimulation technique was developed for downscaling large scale metallurgical samples. The proposed stochastic technique for downscaling is exact, multivariate and meets sample constraints permitting the usage of standard geostatistical techniques and software with the downscaled measurements. The integrated approach, and specific methodologies for nonlinear and multiscale samples quantify joint uncertainty in geometallurgical properties which may be used for effective mine and mill operation planning.

References

Agterberg, F. and Bonham-Carter, G. (1990). Deriving weights of evidence from geoscience contour maps for the prediction of discrete events. In *Proceedings of the 22nd APCOM Symposium, Berlin, Germany*, volume 2, pages 381–395.

Agterberg, F., Bonham-Carter, G. F., Wright, D., et al. (1990). Statistical pattern integration for mineral exploration. *Computer applications in resource estimation prediction and assessment for metals and petroleum*, pages 1–21.

Aitchison, J. (1982). The statistical analysis of compositional data. *Journal of the Royal Statistical Society. Series B (Methodological)*, 44(2):139–177.

Alabert, F. (1987). *Stochastic imaging of spatial distributions using hard and soft information*. MSc thesis, Stanford University.

Allard, D. and Bourotte, M. (2014). Disaggregating daily precipitations into hourly values with a transformed censored latent gaussian process. *Stochastic Environmental Research and Risk Assessment*, 29(2):453–462.

Armstrong, M., Galli, A., Beucher, H., Loc'h, G., Renard, D., Doligez, B., Eschard, R., and Geffroy, F. (2011). *Plurigaussian simulations in geosciences*. Springer Science & Business Media.

Babak, O., Cuba, M., and Leuangthong, O. (2013). Direct upscaling of semivariograms and cross semivariograms for scale-consistent geomodeling. *Transactions of the Society for Mining, Metallurgy and Exploration*, 334:544–552.

Babak, O. and Deutsch, C. V. (2009a). Accounting for parameter uncertainty in reservoir uncertainty assessment: The conditional finite-domain approach. *Natural Resources Research*, 18(1):7–17.

Babak, O. and Deutsch, C. V. (2009b). Collocated cokriging based on merged secondary attributes. *Mathematical Geosciences*, 41(8):921–926.

Ballantyne, G., Powell, M., and Tiang, M. (2012). Proportion of energy attributable to comminution. In *Proceedings of the Eleventh Mill Operators' Conference*, pages 25–30. Australasian Institute of Mining and Metallurgy.

Barnett, R. M. and Deutsch, C. V. (2015). Multivariate imputation of unequally sampled geological variables. *Mathematical Geosciences*, pages 1–27. http://dx.doi.org/10.1007/s11004-014-9580-8.

Barnett, R. M., Deutsch, J. L., and Deutsch, C. V. (2014a). Practical workflows for geostatistical modeling with mean uncertainty. *Sixteenth annual report of the centre for computational geostatistics. University of Alberta, Edmonton*.

Barnett, R. M., Manchuk, J. G., and Deutsch, C. V. (2014b). Projection pursuit multivariate transform. *Mathematical Geosciences*, 46(3):337–359.

Becker, N. (1968). Models for the response of a mixture. *Journal of the Royal Statistical Society. Series B (Methodological)*, 30(2):349–358.

Bliss, C. I. (1934). The method of probits. *Science*, 79(2037):38–39.

Boisvert, J., Rossi, M., Ehrig, K., and Deutsch, C. (2013). Geometallurgical modeling at olympic dam mine, South Australia. *Mathematical Geosciences*, 45:901–925.

Bond, F. C. (1952). The third theory of comminution. *Trans. AIME*, 193(2):484–494.

Bordenave, C., Gousseau, Y., and Roueff, F. (2006). The dead leaves model: A general tessellation modeling occlusion. *Advances in Applied Probability*, 38(1):31–46.

Boucher, A., Dimitrakopoulos, R., and Vargas-Guzmán, J. (2005). Joint simulations, optimal drillhole spacing and the role of the stockpile. In *Geostatistics Banff 2004*, pages 35–44. Springer.

Bourgault, G. (1997). Spatial declustering weights. *Mathematical Geology*, 29(2):277–290.

Box, G. E. and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 26(2):211–252.

Caers, J. (2000). Adding local accuracy to direct sequential simulation. *Mathematical Geology*, 32(7):815–850.

Carrasco, P., Chiles, J., and Seguret, S. (2008). Additivity, metallurgical recovery and grade. In *Proceedings of the Eighth International Geostatistics Congress*, pages 465–476. The University of Chile.

Carroll, R. J., Ruppert, D., Stefanski, L. A., and Crainiceanu, C. M. (2006). *Measurement error in nonlinear models: a modern perspective*. CRC press.

Cheng, Q. (2015). Boostwofe: A new sequential weights of evidence model reducing the effect of conditional dependency. *Mathematical Geosciences*, 47(5):591–621.

Chilès, J.-P. and Delfiner, P. (2009). *Geostatistics: modeling spatial uncertainty*, volume 497. John Wiley & Sons.

Conteras, A. (2013). *Effect of Ore Blends on Flotation of Copper and Molybdenum in Porphyry Ores*. MSc thesis, University of Queensland.

Cornell, J. A. (2011). *Experiments with mixtures: designs, models, and the analysis of mixture data*, volume 895. Wiley.

Cover, T. M. and Thomas, J. A. (2012). *Elements of information theory*. John Wiley & Sons.

Coward, S., Vann, J., Dunham, S., and Stewart, M. (2009). The primary-response framework for geometallurgicalvariables. In *Proceedings of the Seventh International Mining Geology Conference*, pages 109–113. The Australasian Institute of Mining and Metallurgy.

Davis, B. M. and Greenes, K. A. (1983). Estimation using spatially distributed multivariate data: an example with coal quality. *Journal of the International Association for Mathematical Geology*, 15(2):287–300.

Deere, D. and Deere, D. (1988). The rock quality designation (rqd) index in practice. In *Symposium on Rock Classification Systems for Engineering Purposes, 1987, Cincinnati, Ohio, USA*.

Desbarats, A. and Dimitrakopoulos, R. (2000). Geostatistical simulation of regionalized pore-size distributions using min/max autocorrelation factors. *Mathematical Geology*, 32(8):919–942.

Deutsch, C. V. (1989). Declus: a fortran 77 program for determining optimum spatial declustering weights. *Computers & Geosciences*, 15(3):325–332.

Deutsch, C. V. (2000). MIN E 614 lecture notes. University of Alberta, Canada.

Deutsch, C. V. (2004). A statistical resampling program for correlated data: spatial_bootstrap. *Centre for Computational Geostatistics*, 6.

Deutsch, C. V. (2006). A sequential indicator simulation program for categorical variables with point and block data: Blocksis. *Computers & Geosciences*, 32(10):1669–1681.

Deutsch, C. V., Frykman, P., and Xie, Y.-L. (1999). Declustering with seismic or soft geological data. *Centre for Computational Geostatistics*, 1:1–27.

Deutsch, C. V. and Journel, A. G. (1998). *GSLIB: geostatistical software library and users guide*. Oxford University Press New York.

Deutsch, C. V. and Wilde, B. J. (2013). Modeling multiple coal seams using signed distance functions and global kriging. *International Journal of Coal Geology*, 112:87–93.

Deutsch, C. V. and Zanon, S. (2004). Direct prediction of reservoir performance with Bayesian updating under a multivariate gaussian model. In *Petroleum Societys 5th Canadian International Petroleum Conference (55th Annual Technical Meeting), Calgary, Alberta, Canada*.

Deutsch, J. L. and Deutsch, C. V. (2010). Some geostatistical software implementation details. *Centre for Computational Geostatistics*, 12:1–10.

Deutsch, J. L. and Deutsch, C. V. (2011). Plotting and checking the bivariate distributions of multiple gaussian data. *Computers & Geosciences*, 37(10):1677–1684.

Deutsch, M. V. and Deutsch, C. V. (2013). An open source 3d lerchs grossmann pit optimization algorithm to facilitate uncertainty management. *Centre for Computational Geostatistics*, 15:1–6.

Dimitrakopoulos, R. (1998). Conditional simulation algorithms for modelling orebody uncertainty in open pit optimisation. *International Journal of Surface Mining, Reclamation and Environment*, 12(4):173–179.

Dimitrakopoulos, R. and Jewbali, A. (2013). Joint stochastic optimisation of short and long term mine production planning: method and application in a large operating gold mine. *Mining Technology*, 122(2):110–123.

Doyen, P. and Pillet, W. (1996). Seismic porosity mapping in the ekofisk field using a new form of collocated cokriging. In *1996 Society for Petroleum Engineers Annual Technical Conference and Exhibition*, pages 21–30.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, 7(1):1–26. Institute of Mathematical Statistics.

Efron, B. (1982). *The jackknife, the bootstrap and other resampling plans*. Society for Industrial and Applied Mathematics.

Enders, C. K. (2010). *Applied missing data analysis*. Guilford Publications.

Everett, J. E. and Howard, T. J. (2011). Predicting finished product properties in the mining industry from pre-extraction data. In *Proceedings of the First AUSIMM International Geometallurgy Conference*, pages 205–215. The Australasian Institute of Mining and Metallurgy.

Frost, C. and Thompson, S. G. (2000). Correcting for regression dilution bias: Comparison of methods for a single predictor variable. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 163(2):173–189.

Galli, A., Beucher, H., Le Loch, G., Doligez, B., et al. (1994). The pros and cons of the truncated gaussian method. In *Geostatistical Simulations*, pages 217–233. Springer.

Gómez-Hernández, J. J., Froidevaux, R., and Biver, P. (2005). Exact conditioning to linear constraints in kriging and simulation. In *Geostatistics Banff 2004*, pages 999–1005. Springer.

Goovaerts, P. (1997). *Geostatistics for natural resources evaluation*. Oxford University Press.

Goovaerts, P. (2010). Combining areal and point data in geostatistical interpolation: applications to soil science and medical geography. *Mathematical Geosciences*, 42(5):535–554.

Guardiano, F. B. and Srivastava, R. M. (1993). Multivariate geostatistics: beyond bivariate moments. In *Geostatistics Troia92*, pages 133–144. Springer.

Jensen, J. L. (1998). Some statistical properties of power averages for lognormal samples. *Water Resources Research*, 34(9):2415–2418.

Jha, S. K., Mariethoz, G., Evans, J. P., and McCabe, M. F. (2013). Demonstration of a geostatistical approach to physically consistent downscaling of climate modeling simulations. *Water Resources Research*, 49(1):245–259.

Ji, S., Wang, Q., Xia, B., and Marcotte, D. (2004). Mechanical properties of multiphase materials and rocks: a phenomenological approach using generalized means. *Journal of Structural Geology*, 26(8):1377–1390.

JKTech (2015). JKTech Laboratory Services. www.jktech.com.au. Accessed: 2015-07-09.

Johansson, G. and Pugh, R. (1992). The influence of particle size and hydrophobicity on the stability of mineralized froths. *International Journal of Mineral Processing*, 34(12):1–21.

Journel, A. and Huijbregts, C. (1978). *Mining Geostatistics*. Blackburn Press.

Journel, A. G. (1983). Nonparametric estimation of spatial distributions. *Journal of the International Association for Mathematical Geology*, 15(3):445–468.

Journel, A. G. (2002). Combining knowledge from diverse sources: an alternative to traditional data independence hypotheses. *Mathematical Geology*, 34(5):573–596.

Journel, A. G. and Isaaks, E. H. (1984). Conditional indicator simulation: application to a saskatchewan uranium deposit. *Journal of the International Association for Mathematical Geology*, 16(7):685–718.

Journel, A. G. and Xu, W. (1994). Posterior identification of histograms conditional to local data. *Mathematical Geology*, 26(3):323–359.

Keeney, L. and Walters, S. (2011). A methodology for geometallurgical mapping and orebody modelling. In *Proceedings of the First AUSIMM International Geometallurgy Conference*, pages 217–225. The Australasian Institute of Mining and Metallurgy.

Khan, D. K., Deutsch, J. L., and Deutsch, C. V. (2014). A refresher on large scale uncertainty for resource estimation. *Sixteenth annual report of the centre for computational geostatistics. University of Alberta, Edmonton*.

King, P. (2012). *Modeling and simulation of mineral processing systems*. Society for Mining, Metallurgy and Exploration.

Korvin, G. (1982). Axiomatic characterization of the general mixture rule. *Geoexploration*, 19(4):267–276.

Koushavand, B., Ortiz, J., and Deutsch, C. (2008). A methodology to quantify and transfer variogram uncertainty through kriging and simulation. *Tenth annual report of the centre for computational geostatistics. University of Alberta, Edmonton*.

Kovitz, J. L. and Christakos, G. (2004). Spatial statistics of clustered data. *Stochastic Environmental Research and Risk Assessment*, 18(3):147–166.

Kuhar, L. L., Jeffrey, M. I., McFarlane, A. J., Benvie, B., Botsis, N. M., Turner, N., and Robinson, D. J. (2011). The development of small-scale tests to determine hydrometallurgical indices for orebody mapping and domaining. In *Proceedings of the First AUSIMM International Geometallurgy Conference*, pages 335–345. The Australasian Institute of Mining and Metallurgy.

Larsen, R. (2002). Decomposition using maximum autocorrelation factors. *Journal of Chemometrics*, 16(8-10):427–435.

Lerchs, H. and Grossmann, I. (1964). Optimum design of open-pit mines. *Transactions of the Canadian Institute of Mining, Metallurgy and Petroleum*, 68:17–24.

Leuangthong, O. and Deutsch, C. V. (2003). Stepwise conditional transformation for simulation of multiple variables. *Mathematical Geology*, 35(2):155–173.

Leuangthong, O., Ortiz, J., and Deutsch, C. (2005). On the scaling and use of multivariate distributions in geostatistical simulation. *Centre for Computational Geostatistics*, 7:107:6.

Lozano, C. and Bennett, C. (2003). Geometallurgical modeling applied to production forecasting, plant design and optimisation. In *Proceedings Procemin: Mineral Processing Conference*, pages 1–11.

Lund, C., Lamberg, P., and Lindberg, T. (2014). A new method to quantify mineral textures for geometallurgy. In *Process Mineralogy, Cape Town, South Africa*. Minerals Engineering Institute.

Machuca-Mory, D., Babak, O., and Deutsch, C. (2008). Flexible change of support model suitable for a wide range of mineralization styles. *Mining Engineering Magazine*, February:10.

Marchant, B. and Lark, R. (2004). Estimating variogram uncertainty. *Mathematical Geology*, 36(8):867–898.

Mariethoz, G. and Caers, J. (2014). *Multiple-point Geostatistics: Stochastic Modeling with Training Images*. Wiley-Blackwell.

Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial & Applied Mathematics*, 11(2):431–441.

Matheron, G., Beucher, H., De Fouquet, C., Galli, A., Guerillot, D., Ravenne, C., et al. (1987). Conditional simulation of the geometry of fluvio-deltaic reservoirs. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.

McLennan, J. (2007). *The Decision of Stationarity*. PhD thesis, University of Alberta.

Moré, J. J. (1978). The Levenberg-Marquardt algorithm: implementation and theory. In *Numerical Analysis*, pages 105–116. Springer.

Morrell, S. (2008). A method for predicting the specific energy requirement of comminution circuits and assessing their energy utilisation efficiency. *Minerals Engineering*, 21(3):224–233.

Moyeed, R. A. and Papritz, A. (2002). An empirical comparison of kriging methods for nonlinear spatial point prediction. *Mathematical Geology*, 34(4):365–386.

Napier-Munn, T. and Wills, B. A. (2011). *Wills' Mineral Processing Technology: An Introduction to the Practical Aspects of Ore Treatment and Mineral Recovery*. Butterworth-Heinemann, 7th edition.

Neufeld, C. and Deutsch, C. (2004). Incorporating secondary data in the prediction of reservoir properties using bayesian updating. *Centre for Computational Geostatistics*, 6:114:12.

Neufeld, C. and Deutsch, C. (2006). Data integration with non-parametric Bayesian updating. *Centre for Computational Geostatistics*, 8:105:18.

Newton, M. J. and Graham, J. M. (2011). Spatial modelling and optimisation of geometallurgical indices. In *Proceedings of the First AUSIMM International Geometallurgy Conference*, pages 247–261. The Australasian Institute of Mining and Metallurgy.

Olea, R. A. (2007). Declustering of clustered preferential sampling for histogram and semivariogram inference. *Mathematical Geology*, 39(5):453–467.

Ortiz, J. and Deutsch, C. V. (2002). Calculation of uncertainty in the variogram. *Mathematical Geology*, 34(2):169–183.

Oz, B. and Deutsch, C. V. (2002). Size scaling of cross correlation between multiple variables. *Natural Resources Research*, 11(1):1–18.

Oz, B., Deutsch, C. V., Tran, T. T., and Xie, Y. (2003). Dssim-hr: A fortran 90 program for direct sequential simulation with histogram reproduction. *Computers & Geosciences*, 29(1):39–51.

Oz, B., V Deutsch, C., and Frykman, P. (2002). A visualbasic program for histogram and variogram scaling. *Computers & Geosciences*, 28(1):21–31.

Palmstrom, A. (2005). Measurements of and correlations between block size and rock quality designation (rqd). *Tunnelling and Underground Space Technology*, 20(4):362–377.

Pardo-Igúzquiza, E., Chica-Olmo, M., and Atkinson, P. M. (2006). Downscaling cokriging for image sharpening. *Remote Sensing of Environment*, 102(1):86–98.

Pardo-Igúzquiza, E. and Dowd, P. (2001). Variance–covariance matrix of the experimental variogram: assessing variogram uncertainty. *Mathematical Geology*, 33(4):397–419.

Pawlowsky, V., Olea, R. A., and Davis, J. C. (1993). Boundary assessment under uncertainty: a case study. *Mathematical Geology*, 25(2):125–144.

Pawlowsky-Glahn, V. and Olea, R. A. (2004). *Geostatistical Analysis of Compositional Data*. Oxford University Press.

Powell, M. S. (2013). Utilising orebody knowledge to improve comminution circuit design and energy utilisation. In *Proceedings of the Second AUSIMM International Geometallurgy Conference*, pages 27–35. The Australasian Institute of Mining and Metallurgy.

Pyrcz, M. and Deutsch, C. (2002). Declustering and debiasing. *Centre for Computational Geostatistics*, 4:1–25.

Pyrcz, M. J. and Deutsch, C. V. (2014). *Geostatistical reservoir modeling*. Oxford University Press.

Pyrcz, M. J., Gringarten, E., Frykman, P., and Deutsch, C. V. (2006). Representative input parameters for geostatistical simulation. *Stochastic modeling and geostatistics: Principles, methods, and case studies, volume II: AAPG Computer Applications in Geology*, pages 123–137.

Ren, W. (2007). *Exact Downscaling in Reservoir Modeling*. PhD thesis, University of Alberta.

Ren, W. and Deutsch, C. (2006). Bayesian updating with locally varying correlation. *Centre for Computational Geostatistics*, 8:104:23.

Robertson, R. K., Mueller, U. A., and Bloom, L. M. (2006). Direct sequential simulation with histogram reproduction: A comparison of algorithms. *Computers & Geosciences*, 32(3):382–395.

Rossi, M. and Deutsch, C. (2013). *Mineral resource estimation*. Springer Science & Business Media.

Sakia, R. (1992). The Box-Cox transformation technique: a review. *The Statistician*, pages 169–178.

Scheffé, H. (1958). Experiments with mixtures. *Journal of the Royal Statistical Society. Series B (Methodological)*, 20(2):344–360.

Silva, D. and Deutsch, C. (2013). Correcting distance function models to correct proportions. *Centre for Computational Geostatistics*, 15(116):10.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis*, volume 26. CRC press.

Snee, R. D. (1973). Techniques for the analysis of mixture data. *Technometrics*, 15(3):517–528.

Soares, A. (2001). Direct sequential simulation and cosimulation. *Mathematical Geology*, 33(8):911–926.

Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1):72–101.

Srivastava, R. M. (2005). Probabilistic modeling of ore lens geometry: An alternative to deterministic wireframes. *Mathematical Geology*, 37(5):513–544.

Strebelle, S. (2002). Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical Geology*, 34(1):1–21.

Tran, T. T., Wen, X.-H., Behrens, R. A., et al. (1999). Efficient conditioning of 3d fine-scale reservoir model to multiphase production data using streamline-based coarse-scale inversion and geostatistical downscaling. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.

Tukey, J. W. (1957). On the comparative anatomy of transformations. *The Annals of Mathematical Statistics*, 28(3):pp. 602–632.

Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley.

van den Boogaart, K. G., Konsulke, S., and Tolosana Delgado, R. (2013). Non-linear geostatistics for geometallurgical optimisation. In *Second AUSIMM International Geometallurgy Conference*. The Australasian Institute of Mining and Metallurgy.

Van Tonder, E., Deglon, D. A., and Napier-Munn, T. J. (2010). The effect of ore blends on the mineral processing of platinum ores. *Minerals Engineering*, 23(8):621–626.

Vargas-Guzmán, J. A. and Dimitrakopoulos, R. (2003). Computational properties of min/max autocorrelation factors. *Computers & Geosciences*, 29(6):715–723.

Wackernagel, H. (2003). *Multivariate geostatistics*. Springer.

Wang, G., Pelton, R., Hrymak, A., Shawafaty, N., and Heng, Y. M. (1999). On the role of hydrophobic particles and surfactants in defoaming. *Langmuir*, 15(6):2202–2208.

Wilde, B. (2010). *Data Spacing and Uncertainty*. MSc thesis, University of Alberta.

Wilde, B. and Deutsch, C. (2013). Methodology for quantifying uncertainty versus data spacing applied to the oil sands. *CIM Journal*, 4(4):211–219.

Yan, D. and Eaton, R. (1994). Breakage properties of ore blends. *Minerals Engineering*, 7(23):185–199.

Zagayevskiy, Y. and Deutsch, C. (2011). A short note on an update to the change of support program. *Centre for Computational Geostatistics*, 13(313):12.