ANALYSING SPATIAL DATA VIA GEOSTATISTICAL METHODS

Craig John Morgan

A dissertation submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfillment of the requirements for the degree of Master of Science.

Johannesburg, 2005

DECLARATION

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

Morgan

5th day of December 2005

ABSTRACT

This dissertation presents a detailed study of geostatistics. Included in this work are details of the development of geostatistics and its usefulness both in and outside of the mining industry, a comprehensive presentation of the theory of geostatistics, and a discussion of the application of this theory to practical situations. A published debate over the validity of geostatistics is also examined.

The ultimate goal of this dissertation is to provide a thorough investigation of geostatistics from both a theoretical and a practical perspective. The theory presented in this dissertation is thus tested on various spatial data sets, and from these tests it is concluded that geostatistics can be effectively used in practice provided that the practitioner fully understands the theory of geostatistics and the spatial data being analyzed. A particularly interesting conclusion to come out of this dissertation is the importance of using additive regionalized variables in all geostatistical analyses.

ACKNOWLEDGEMENTS

First of all, I wish to express thanks to my two supervisors, Professor Paul Fatti and Professor Dick Minnitt, for their assistance over the past three years. I would also like to thank Christina Dohm for supplying a sizeable spatial data set taken from the South African gold mining industry, and for allowing me to attend the three courses

1. Statistical Valuation of Ore Reserves (MINN 511)
2. Geostatistical Evaluation of Mineral Resources (MINN 510)
3. Practical Implementation of Geostatistical Evaluation (MINN 572)

given at the University of the Witwatersrand in 2003. Special thanks also goes out to Isobel Clark for making available the geostatistical computer package Geostokos Toolkit (c.2003), and for her general eagerness to help. Finally, I wish to thank the School of Statistics and Actuarial Science at the University of the Witwatersrand for financing the purchase of the geostatistical add-on package S+SpatialStats (2001).

CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

LIST OF SYMBOLS AND COMMON NOTATION

	indicate the conclusion of a proof
	indicates the conclusion of an example
	indicates the conclusion of a note
$\alpha(L\,;\,l)$	an auxiliary function
β	the threshold parameter of a three-parameter lognormal distribution
β_1	the square of the population skewness
β_2	the population kurtosis
c_0	a nugget-effect
$C(0)$	the covariance value at a lag of zero; equal to the variance of a second-order stationary random function
$C(h)$	a covariance function
$\chi(L)$ and $\chi(L\,;\,l)$	auxiliary functions
$C_V(h)$	the regularized covariance function of support *V*, i.e. the covariance function of a random function of support *V*
$\overline{C}(V,v)$	the mean value of $C(h)$ when one extremity of vector *h* describes the domain $V(x)$ and the other extremity independently describes the domain $v(y)$
$C(x,y)$ or $\mathrm{cov}(x,y)$	the covariance between the random variables $Z(x)$ and $Z(y)$

$D^2(v/V)$	the dispersion variance of blocks $v(y)$ within $V(x)$
$E\{Z(x)\}$	the expectation of a random function or a random variable
$F(L)$ and $F(L\,;\,l)$	auxiliary functions
$f(x)$	generally indicates a function of the spatial co-ordinate x, or a probability density function
$\gamma(h)$	a semi-variogram. $2\gamma(h)$ is the notation for a variogram.
$\gamma_V(h)$	the regularized semi-variogram of support V, i.e. the semi-variogram of a random function of support V
$\gamma(x,y)$	the semi-variogram value between the random variables $Z(x)$ and $Z(y)$
$\bar{\gamma}(V,v)$	the mean value of $\gamma(h)$ when one extremity of the vector h describes the domain $V(x)$ and the other extremity independently describes the domain $v(y)$
h	a separation vector of modulus $\|\|h\|\|$
$H(L\,;\,l)$	an auxiliary function
λ_i , λ_j , …	generally refers to kriging weights or the weights of a linear combination
μ	the population mean
$N(h)$	indicates the set containing all sample pairs separated by the vector h
OK	ordinary kriging
OLK	ordinary lognormal kriging
Ω	a spatial study area of two- or three-dimensions

$\Phi(y)$	the cumulative distribution function of the standard normal distribution
^{1}, ^{2}, ^{3}	one-, two- and three dimension Euclidean space respectively
σ^2	the population variance
σ_E^2	the estimation variance
$\sigma_E^2(v,V)$	the extension variance of v to V
σ_{OK}^2	the ordinary kriging variance
σ_{OLK}^2	the ordinary lognormal kriging variance
σ_{SK}^2	the simple kriging variance
σ_{UK}^2	the universal kriging variance
SK	Simple kriging
t	Sichel's t-estimator
τ	the mean value of a two- or three-parameter lognormal distribution
UK	universal kriging
var$\{Z(x)\}$	the variability of a random function or a random variable
$V(x)$	A volume (or area) V centred at the point location x
$v(x)$	a volume (or area) v, usually smaller than V, centred at the point location x
ϖ, ϖ_i	Lagrange multipliers
x, y, …	generally indicates a point co-ordinate in one-, two- or three-dimensional space

$Z_V(x)$	a block support random function or random variable, with a support volume (or area) of V
$z_V(x)$	a block support regionalized variable or a realization of a block support random variable, with a support volume (or area) of V
$Z(x)$, $Y(x)$, …	a random function or a random variable at the point location x
$z(x)$, $y(x)$, …	a regionalized variable or a realization of a random variable at the point location x