

# **Practical Geostatistics 2000**

**Isobel Clark  
Geostokos (Ecosse) Limited, Scotland**

**William V Harper  
Otterbein College, Westerville OH, USA**



**Ecosse North America Llc, Columbus Ohio, USA**



Second Reprint, August 2001

Copyright © 2000 by Geostokos (Ecosse) Limited, Scotland.

All rights reserved. Published simultaneously worldwide on CD.

Reproduction or translation of any part of this work beyond that permitted by Section 107 or 108 of the 1976 United States Copyright Act without the permission of the copyright owner is unlawful. Requests for permission or further information should be addressed to Permissions Department, Geostokos (Ecosse) Limited, Alloa Business Centre, Whins Road, Alloa FK10 3SA, Central Scotland.

Tel: +44 (0)773 663 1632 Fax: +44 (0)1259 215274

e-mail: [ecosse@stokos.demon.co.uk](mailto:ecosse@stokos.demon.co.uk)

<http://geoecosse.bizland.com>

<http://geoecosse.hypermart.net>

Cataloging in Publication Data:

Clark, Isobel and Harper, William V.

Practical geostatistics 2000 / Isobel Clark & William V. Harper.

p. cm.

ISBN: 0-9703317-0-3; 0-9703317-1-1 (CD-Rom); 0-9703317-2-X

(combination pack)

1. Geology—Statistical methods. 2. Mining engineering. 3. Environmental engineering I. Harper, William V. II. Title

QE33.2.S82 C43 2000

519.5'02'4553—dc21

Ecosse North America Llc, publishers.

Typeset using Scienti...c Word 4.0, MacKichan Software Inc,

<http://www.mackichan.com>

Printed by Kinkos, Westerville, Ohio USA. <http://www.kinkos.com>

This book is dedicated to

**Anne Anderson**

The 1979 *Practical Geostatistics* started with a quotation from the poem “The Chorus” in Edwin Morgan’s anthology *The New Divan*, published by Carcanet/Manchester in 1977:

*“What is obscure is seldom clear”*

For this new book, longer and (we hope) much improved, we use a longer — but just as ironic — quotation:

*“Let us all be quiet in unison  
and hear what the oracle has to tell us.*

*The oracle says, when winter comes  
the autumn is over. Are we comforted?”*

### **Special Request for our Readers**

We have a favour to ask of you, the readers - we want your guidance on what changes should be made in subsequent editions. We hope to update this book fairly frequently as we are publishing this ourselves and have total control over what it contains, how many copies are printed, etc. In addition to sending us questions and typos, we hope that you will take the time to specify what new topics we should add (or what topics should be expanded). Make sure Isobel gets a copy of all such comments. E-mail her at [PG2000@stokos.demon.co.uk](mailto:PG2000@stokos.demon.co.uk) and she will post all typos and major comments on the web page

[http://uk.geocities.com/geocosse/PG2000\\_2nd\\_reprint](http://uk.geocities.com/geocosse/PG2000_2nd_reprint).

Thanks (in advance) for all help in this effort.

# Contents

Preface

Notation

<b>1. Introduction</b>	<b>Page 1</b>
1.1 Expectations	
1.2 The problem to be solved	
1.3 Data sets	
1.4 Software	
<b>2. Why a statistical approach?</b>	<b>Page 7</b>
2.1 Investigating the sample data	
2.2 Measures of central tendency	
2.3 Measures of spread or variability	
2.4 Graphical descriptions of the data	
2.5 Other useful descriptive statistics	
2.6 Discrete data	
2.7 Into the unknown	
2.8 Worked examples	
2.8.1 Coal project data, calori...c values	
2.8.2 Iron ore example	
2.8.3 Wolfcamp data	
2.8.4 Scallops, total catch	
2.9 Exercises	
<b>3. Normal (Gaussian) distributions</b>	<b>Page 31</b>
3.1 The gap between data and population	
3.2 Is it a Normal distribution?	

- 3.3 Estimating population parameters
    - 3.3.1 estimating the population average
    - 3.3.2 estimating the standard deviation
    - 3.3.3 confidence intervals for standard deviation
    - 3.3.4 confidence intervals for mean
  - 3.4 Selection (grade/tonnage) calculations
  - 3.5 Summary of chapter
  - 3.6 Worked examples
    - 3.6.1 Coal project, calorific values
    - 3.6.2 Iron ore example
    - 3.6.3 Wolfcamp data
    - 3.6.4 Scallop data, total catch
  - 3.7 Exercises
4. Lognormal distributions (and others) Page 67
- 4.1 The lognormal distribution
    - 4.1.1 estimating the mean of a lognormal population
    - 4.1.2 confidence intervals on the population mean
  - 4.2 The three parameter lognormal
  - 4.3 Selection (grade/tonnage) calculations
    - 4.3.1 two parameter lognormal — reef widths
    - 4.3.2 three parameter lognormal — gold grades
  - 4.4 More complex distributions and mixtures
    - 4.4.1 mixtures of Normal or lognormal populations
  - 4.5 Worked examples
    - 4.5.1 Scallops, total caught
    - 4.5.2 Organic matter in soil
    - 4.5.3 Calcium in limestone
    - 4.5.4 Geevor Tin mine, Cornwall
  - 4.6 Exercises
5. Discrete distributions Page 103
- 5.1 Review of Discrete Moments
  - 5.2 Bernoulli and Binomial Distributions
  - 5.3 Negative Binomial and Geometric Distributions

- 5.4 Poisson Distribution
- 5.5 Mixtures of Poisson Distributions (Compound Poisson)
  - 5.5.1 Oswego Zircon data
  - 5.5.2 Other examples
- 5.6 Spatial Considerations
- 5.7 Solved Problems
- 5.8 Exercises

**6. Hypothesis testing** **Page 135**

- 6.1 Single sample tests
  - 6.1.1 test on sample mean
  - 6.1.2 test on sample standard deviation
- 6.2 Two sample tests
  - 6.2.1 test on standard deviations
  - 6.2.2 test on means
  - 6.2.3 paired sampling
  - 6.2.4 test for sample distribution
- 6.3 Worked examples
  - 6.3.1 Heights of students
  - 6.3.2 Geevor tin mine – development versus stope
- 6.4 Exercises

**7. Relationships** **Page 147**

- 7.1 Straight line relationships
  - 7.1.1 quantifying the strength of the relationship
  - 7.1.2 Predicting one variable from the other
  - 7.1.3 Calori...c Value versus Ash Content
  - 7.1.4 Calori...c Value versus Sulphur Content
- 7.2 Other worked examples
  - 7.2.1 Gold grade versus reef width
  - 7.2.2 Scallops caught
  - 7.2.3 Application — Krige's Regression Effect
- 7.3 Relationships involving more than two variables
  - 7.3.1 Predicting Sulphur from Calori...c Value and Ash Content
  - 7.3.2 Application — Krige's moving average template

7.3.3 Curvilinear Regression

7.3.4 Application — Polynomial Trend Surface Analysis

7.4 Exercises

8. The spatial element Page 185

8.1 Including location as well as value

8.2 Spatial relationships

8.3 Inverse distance estimation

8.4 Worked examples

8.4.1 Coal project, calorific values

8.4.2 Iron ore project

8.4.3 Wolfcamp data

8.4.4 Scallops caught

8.5 Exercises

9. The semi-variogram Page 207

9.1 The experimental semi-variogram

9.1.1 Irregular sampling

9.1.2 Cautionary notes

9.2 Modelling of the semi-variogram function

9.2.1 The linear model

9.2.2 The generalised linear model

9.2.3 The Spherical model

9.2.4 The exponential model

9.2.5 Gaussian model

9.2.6 The hole effect model

9.2.7 Paddington mix model

9.2.8 Judging how well the model fits the data

9.2.9 equivalence to covariance function

9.2.10 the nugget effect

9.3 Worked examples

9.3.1 Silver example from Practical Geostatistics 1979

9.3.2 Coal project: calorific values

9.3.3 Wolfcamp aquifer

9.4 Exercises

## 10. Estimation and Kriging

Page 247

### 10.1 Estimation error

- 10.1.1 one sample estimation
- 10.1.2 another single sample
- 10.1.3 two sample estimation
- 10.1.4 another two sample estimation
- 10.1.5 three sample estimator

### 10.2 Choosing the optimal weights

- 10.2.1 three sample estimation
- 10.2.2 the general form for the 'optimal' estimator
- 10.2.3 confidence levels and degrees of freedom
- 10.2.4 simple kriging

### 10.3 Ordinary kriging

- 10.3.1 'optimal' unbiased estimator
- 10.3.2 alternate form: matrices
- 10.3.3 alternate form: covariance
- 10.3.4 three sample estimation

### 10.4 Worked examples

- 10.4.1 Coal project, calorific values
- 10.4.2 Iron ore example, (Page95)
- 10.4.3 Wolfcamp, residuals from quadratic surface

### 10.5 Cross validation

- 10.5.1 cross cross validation

### 10.6 Exercises

## 11. Areas and volumes

Page 295

### 11.1 The impact on the distribution

- 11.1.1 Iron ore example, Normal example
- 11.1.2 Gevor Tin Mine, lognormal(ish) example

### 11.2 The impact on kriging

- 11.2.1 the use of auxiliary functions
- 11.2.2 Iron ore example, Page 95
- 11.2.3 Wolfcamp aquifer, quadratic residuals

**12. Other kriging approaches** **Page 315**

**12.1 Universal kriging**

**12.1.1 Wolfcamp aquifer**

**12.2 Lognormal kriging**

**12.2.1 the proportional effect**

**12.2.2 the lognormal transformation**

**12.2.3 Geevor Tin Mine, grades**

**12.2.4 SA Gold Mine**

**12.3 Indicator kriging**

**12.4 Rank uniform kriging**

**12.5 Summary of chapter**

**13. Bibliography** **Page 339**

**Tables**

**Data Sets**

**Index**

# Preface

This is the bit where the authors traditionally thank everyone who helped in the production of the book and set the tone for the rest of the tome.



Isobel Clark



Bill Harper

## Isobel:

Thanks guys and girls. There are way too many people who have made contributions over the years to expect to cover them all. The 11 years at the Royal School of Mines in London and the 10 at the University of the Witwatersrand form the basis for the teaching material and exercises in the bulk of these chapters. Open short courses and in-house courses to all manner of audiences added to and enhanced both the direction of the teaching and the exercises provided. So to all those ex-undergraduates, ex-postgraduates and course participants on four continents, thanks.

Thanks to Dick Minnit, Fred Cawood and Hennie Greyvenstein in Johannesburg for bugging me until the book actually got written. Thanks to Audun Clark for unconditional support. Thanks to Gavin Lind for assistance in courses and ancillary material. And, of course, endless thanks to my co-author who made all this possible.....

This book is intended for self-study, especially in its CD form. We hope to develop this into a full interactive teaching medium in the future. All comments are welcome although positive criticism is likely to get more response!

Contact me: Dr Isobel Clark, Geostokos (Ecosse) Limited,  
Alloa Business Centre, Whins Road, Alloa, FK10 3SA, Central Scotland  
Tel: 0773 663 1632 Fax: 01259 210254. e-mail: drisobelclark@yahoo.co.uk

**Bill:** Many moons ago in the early 1980s Howard Hume and I were trying to dig into the seemingly strange world of geostatistics. Much was not intuitive and seemed to hide behind a strange wall of jargon. Howard and I decided to contact those individuals that we felt might be able to shed light down the borehole. Fortunately we heard from Isobel (whose 1979 book was one of the few readable geostatistics publications at that time) and she changed her next U.S. trip to come to Columbus to visit us at Battelle in 1984. As they say, the rest is history. Isobel and I collaborated on numerous articles in the 1980s. Perhaps my biggest claim to geostatistical fame was publishing the Wolfcamp data in my first geostatistical analysis. This resulted in many publications by others using this particular set of data.

In the late 1980s I moved away from environmental geostatistical applications after the Battelle high-level nuclear waste program was terminated by the U.S. government. Until early 1999 most of my 1990s efforts have been focused on more classical statistical methods, operations research, quality management, and discrete event simulation. In Spring 1999 Isobel was staying at our place for a few days' visit. After several enjoyable beverages, I raised the issue of her many years of promising to update her book. The result of that discussion is now before your eyes. As the Grateful Dead used to sing, *"What a long, strange trip it's been"*.

I would like to recognize the invaluable support from my lovely wife, Paula, and daughter, Rachel. I am also thankful of the continued good support of Otterbein College. Finally, I want to thank Isobel for being Isobel.

Please send us any and all feedback. We hope to update this often to meet your requests.

Contact me: William V. Harper, Otterbein College, Mathematical Sciences  
Towers Hall 136, 1 Otterbein College, Westerville, OH 43081-2006 U.S.A.  
Phone: 614-823-1417 Fax: 614-823-3201  
WHarper@otterbein.edu <http://geocosse.hypermart.net>

# Notation

<i>a</i>	the range of influence or distance scaling parameter in a model semi-variogram
<i>b</i>	breadth of a panel or block under study
<i>b<sub>k</sub></i>	coefficients estimated in a regression or trend surface analysis
<i>c</i>	cut-off or 'pay' value used in a grade/tonnage calculation
<i>d</i>	distance between a sample and a location at which an estimate is required
<i>e</i>	2.718281828..... the natural anti-logarithm function
<i>f</i>	a secondary or ancillary variable measured on a sample
<i>g</i>	the primary value measured at a sample location
<i>h</i>	the distance between any two specified locations within the study area
<i>i</i>	a subscript identifier
<i>j</i>	a subscript identifier
<i>k</i>	a subscript or integer power
<i>l</i>	length of a panel or block under study
<i>m</i>	number of local samples in an estimation
<i>n</i>	number of samples available in the current study
<i>p</i>	slope of the line in a linear semi-variogram model
<i>r</i>	estimate for the correlation coefficient between two variables
<i>s</i>	unbiased estimate for the standard deviation of a population
<i>t</i>	Student's distribution for the arithmetic mean of Normal samples Sichel's estimator for the average of a lognormal population
<i>v</i>	a volume
<i>w</i>	weighting factor
<i>x</i>	standardised Normal variate, mean 0 standard deviation 1
<i>y</i>	logarithm of <i>g</i> value
<i>z</i>	residual from a trend surface
<b>A</b>	an area or volume whose value is to be estimated
<b>C</b>	nugget effect $C_0$ or sill $C_i$ on a semi-variogram model
<b>F</b>	auxiliary function for the variance within a panel or block
<b>H</b>	auxiliary function for a corner point with a 2d panel
<b>M</b>	a specified point within an area or volume
<b>N</b>	number of pairs of samples found in a semi-variogram calculation
<b>P</b>	proportion of the population above a specified cut-off value or pay limit may also be used to represent percentage rather than proportion
<b>Q</b>	an auxiliary function between two 'lines' at right angles
<b>S</b>	sum of squared residuals from a regression exercise
<b>T</b>	the unknown value at an unsampled location the distribution of Sichel's <i>t</i> estimate for the lognormal mean
<b>V</b>	the variance of samples around their arithmetic mean
<b>X</b>	co-ordinate for a location, usually east/west
<b>Y</b>	co-ordinate for a location, usually north/south
<b>Z</b>	co-ordinate for a location, usually elevation or depth

## Notation continued

### *lower case Greek letters*

⊙	alpha	power to which distance is raised in a semi-variogram model the significance level of a hypothesis test
-	beta	additive constant used in a three parameter lognormal distribution
◦	gamma	Sichel's correction factor for the lognormal mean estimation a model semi-variogram or function calculated from such
±	delta	a small change
²	epsilon	the error of estimation, actual value; estimated value
μ	theta	an angle usually expressed in degrees
λ	lambda	Lagrangian multiplier
μ	mu	arithmetic mean of values in a Normal population
◦	nu	degrees of freedom
»	eta	damping or decay parameter in a hole effect semi-variogram model
¼	pi	3.141592654.... the ratio of the perimeter of a circle to its diameter
½	rho	the population correlation coefficient between two variables
¾	sigma	the standard deviation of values in a Normal population
ι	tau	the arithmetic mean of values in a lognormal population
Á	phi	the standard Normal density function
Â	chi	the distribution of sample variances from Normal samples
!	omega	the standard deviation of values in a lognormal population

### *Upper case Greek letters*

§	Sigma	a summation sign, e.g. $\sum_{i=1}^n$ is "sum for series $i = 1; 2; 3; \dots; n$   $1; n$ "
©	Phi	cumulative proportion under the standard Normal distribution
ª	Psi	Sichel's confidence factors for the lognormal distribution