1-1-1996

Variograms and kriging in the analysis of spatial data

Suresh Tripathi
Edith Cowan University

Edith Cowan University

Copyright Warning

You may print or download ONE copy of this document for the purpose of your own research or study.

The University does not authorize you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following:

- Copyright owners are entitled to take legal action against persons who infringe their copyright.

- A reproduction of material that is protected by copyright may be a copyright infringement. Where the reproduction of such material is done without attribution of authorship, with false attribution of authorship or the authorship is treated in a derogatory manner, this may be a breach of the author's moral rights contained in Part IX of the Copyright Act 1968 (Cth).

- Courts have the power to impose a wide range of civil and criminal sanctions for infringement of copyright, infringement of moral rights and other offences under the Copyright Act 1968 (Cth). Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

USE OF THESIS

The Use of Thesis statement is not included in this version of the thesis.

VARIOGRAMS AND KRIGING

IN THE ANALYSIS OF SPATIAL DATA

by

Suresh Tripathi

A thesis submitted

in partial fulfilment of the

requirements for the award of

Master of Science : Mathematics and Planning

Department of Mathematics

Edith Cowan University.

June 1996

TABLE OF CONTENTS

Declaration

I certify that this thesis does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any institution of higher education; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text.

Signature

Date 25/10/96

Acknowledgments

First and foremost I wish to express my profound gratitude to my supervisor Dr Lyn Bloom for her patience, guidance, enthusiasm, support and encouragement. I wish to thank also the Chairperson of the Mathematics Department Dr David McDougall for providing me with some financial support in the way of sessional teaching during my study. I am also grateful to Associate Professor Byron Lamont and Dr Lyn Bloom for providing me respectively with the Banksia data and the Soil data analysed in this thesis. I am deeply indebted to my friend Andrew Mehnert, my family and other colleagues for their support and words of encouragement. Finally I thank my father for providing me with continuous moral support, love and ceaseless encouragement during my study.

Abstract

This research is in the area of geostatistics and consists essentially of two parts. The first is an investigation of the variogram and cross variogram and the associated kriging and cokriging methods of spatial prediction and the second is an application of these in the analysis of two (original) data sets.

In the first part (chapter 1 to chapter 5), the focus is on summarising and illustrating the various techniques of Exploratory Data Analysis (EDA) and some methods used to estimate and model the experimental variograms and cross variograms for a given data set, together with some of the geostatistical methods of kriging and cokriging used for prediction purposes. The research also illustrates some of the many applications of this theory in the earth and environmental sciences.

The second part of the thesis (chapter 6) is an application of these geostatistical techniques to the analysis of two (new) data sets. The first data set consists of the Available Phosphate (in ppm) and Potassium (in ppm) from two fields (one cropped and one uncropped) in the Jimperding Brook area of Western Australia. The second data set consists of the number of species of Banksia at various locations within a region of Southwestern Australia.

1. INTRODUCTION

1.1. Background and Significance

The term *Geostatistics* was first introduced by Matheron (1962) in connection with the analysis of problems arising in the mining industry. It was subsequently extended to geology (Matheron 1963a, 1963b, 1971a) in general and to other earth sciences such as geography (Boots and Dufournaud 1994) and soil science (McBratney and Webster 1986 ; Yost et al 1982). However the theory involved is now being applied in a wide variety of other areas. These include hydrology (de Kwaadsteniet 1990), pollution control (Wartenberg, Uchrin, and Coogan 1991), and atmospheric science (Hass 1990). Some of the ideas introduced by Matheron had appeared earlier in other forms. For instance, what Matheron refers to as the variogram was called a structure function in physics by Kolmorogov (1941), in probability by Yaglom (1962), and in meteorology by Gandin (1963), and a mean-squared difference by Jowett (1952) in time series. However it is Matheron's terminology that has persisted and which will be used throughout this thesis.

The name *kriging* was introduced by Matheron in honour of the South African mining engineer D. G. Krige who (Krige 1951, 1966) developed methods for determining the true ore-grade distributions from the distributions based on the sampled ore grades. However the formulation of spatial linear prediction did not come from Krige's work. The contributions of Weiner (1949), Wold (1938), and Kolmogorov (1941) all contain optimal linear prediction methods.

1

In geostatistics one is dealing with spatial data, where the locations of the samples, as well as their values, are important. Such data frequently exhibit spatial autocorrelation, which, loosely speaking, means that nearby values will be more alike than values located far away from one another. One is interested in obtaining a profile of the distribution of the data values and in using this to predict the value at a particular location. The variogram is an important tool in modelling the spatial continuity (the spatial dependence) of the data while kriging is the method, or more accurately methods, of spatial interpolation used for prediction purposes. The type of kriging used depends on the nature of the data and the underlying assumptions of the mathematical model to be used. One advantage of this approach is that it is possible, via the kriging variance, to obtain an idea of the accuracy of such predictions. When multivariate data are analysed, methods of co-kriging are sometimes used as the prediction accuracy of one variable can often be improved by knowledge of values of another variable measured at the same locations.

1.2. Regionalized Variables

Matheron (1962, 1971b, 1973) introduced the term regionalized variable $z(\mathbf{s})$ where the location index $\mathbf{s}$ is assumed to vary continuously over the domain D, a fixed subset of d-dimensional Euclidean Space $\mathbf{R}^d$. Regionalized variables and their relevant applications in constructing models are very important in geostatistics. A regionalized variable is simply a random variable where the values vary with temporal or spatial location, defined in space-time to exhibit certain types of spatial or temporal

2

characteristics. A regionalized variable exhibits two main features, one of which is a structure aspect, and other is a random aspect. The first is based on theoretical grounds which express structural and geometrical properties in an adequate form, and the second is based on practical grounds which solves the problem of estimation of a regionalized variable from fragmentary sampling data. These two purposes are related to each other because the error of estimation depends on the structural characteristics; for example, the error of estimation becomes greater when the regionalized variable is more discontinuous. Regionalized variable theory focusses attention on means and variances of the variables under consideration. In particular, under the intrinsic stationarity assumptions, increments between all pairs of points, situated at a given distance **h**, have a mean zero and a finite variance that remains the same in the various parts of the area under study. This makes it possible to use the variogram.

1.3. Notation and Stationarity Assumptions

In general, although not exclusively, throughout this thesis we shall follow the notation used by Cressie (1991). Geostatistical models are handled with the help of a probabilistic framework which is traditionally presented in terms of continuous random processes. In this thesis we consider a point s in d-dimensional Euclidean space R^d and suppose that a random variable $\{z(s):s \in D\}$ at spatial location s is a realization of a stochastic process $\{Z(s):s \in D\}$. For simplicity we will give examples from R^2 throughout this thesis.

The process $\{Z(\mathbf{s})\}$ is said to exhibit *strict (strong) stationarity,* if for any integer m, and locations $\mathbf{s}_1,\ldots,\mathbf{s}_m$, and any vector $\mathbf{h} \in R^d$ and all $m \geq 1$

$$P[Z(\mathbf{s}_1) \leq z_1, \cdots, Z(\mathbf{s}_m) \leq z_m] = P[Z(\mathbf{s}_1 + \mathbf{h}) \leq z_1, \cdots, Z(\mathbf{s}_m + \mathbf{h}) \leq z_m] \qquad (1.1)$$

This means the joint distribution of this process evaluated at any set of points is unchanged if all the points are moved the same distance in the same direction. In particular, the expected value (mean) of $Z(\mathbf{s})$ is constant and does not depend on the actual position. Then the process is said to be *stationary in the mean* and this can be expressed as

$$E[Z(\mathbf{s}) - Z(\mathbf{s+h})] = 0 \qquad (1.2)$$

where $\mathbf{h}$ (lag) is a vector in R^d. If the covariance cov $(Z(\mathbf{s}), Z(\mathbf{s+h}))$ is finite and depends upon only $\mathbf{h}$, then the function $C(\mathbf{h})$, called the *covariogram,* can be defined by

$$C(\mathbf{h}) = \text{cov}[Z(\mathbf{s}), Z(\mathbf{s+ h})] \qquad (1.3)$$

If $C(\mathbf{h})$ is a function of only $\|\mathbf{h}\|$ then $C(\mathbf{h})$ is called *isotropic.*

The process $\{Z(\mathbf{s})\}$ is said to be *second order (weak) stationary* if it satisfies (1.2) and (1.3). Strong stationarity implies second-order stationarity whenever $P\{Z(s) \leq z\}$ yields a finite second moment.

The process {$Z(\mathbf{s})$} is said to be *increment stationary* if it satisfies equation (1.2) and

for any integer m, locations $\mathbf{s}_1, \ldots, \mathbf{s}_m$, and vector $\mathbf{h}$

$$P\left[Z(\mathbf{s}_2) - Z(\mathbf{s}_1) \le z_1, \ldots, Z(\mathbf{s}_m) - Z(\mathbf{s}_{m-1}) \le z_{m-1}\right] =$$

$$P\left[Z(\mathbf{s}_2 + \mathbf{h}) - Z(\mathbf{s}_1 + \mathbf{h}) \le z_1, \ldots, Z(\mathbf{s}_m + \mathbf{h}) - Z(\mathbf{s}_{m-1} + \mathbf{h}) \le z_{m-1}\right] \qquad (1.4)$$

If the variance of the difference between $Z(\mathbf{s})$ and $Z(\mathbf{s+h})$ is finite, and depends only

on $\mathbf{h}$ then the *variogram* $2\gamma(\mathbf{h})$ can be defined as

$$2\gamma(\mathbf{h}) = \text{var } [Z(\mathbf{s})\text{-}Z(\mathbf{s+h})] \qquad (1.5)$$

The process {$Z(\mathbf{s})$} is said to be *intrinsic stationary* or, alternatively, to satisfy the

intrinsic hypothesis if it satisfies equations (1.2) and (1.5). In this case the variogram

can be written as the expected squared difference between random variables $Z(\mathbf{s})$ and

$Z(\mathbf{s+h})$ as below.

$$2\gamma(\mathbf{h}) = E\left[(Z(\mathbf{s}) - Z(\mathbf{s} + \mathbf{h}))^2\right] \qquad (1.6)$$

Note that $\gamma(0) = 0$, $\gamma(\mathbf{h}) \ge 0$ and $\gamma(-\mathbf{h}) = \gamma(\mathbf{h})$. If var $[Z(\mathbf{s})] = C(\mathbf{0})$ (finite), and

second-order stationarity is satisfied, then the variogram can be written in terms of the

covariogram as

$$2\gamma(\mathbf{h}) = \text{var}\left[Z(\mathbf{s}) - Z(\mathbf{s} + \mathbf{h})\right]$$

$$\therefore \quad 2\gamma(\mathbf{h}) = \text{var}\,(Z(\mathbf{s})) + \text{var}\,(Z(\mathbf{s}+\mathbf{h})) - 2\,\text{cov}\,(Z(\mathbf{s}), Z(\mathbf{s}+\mathbf{h}))$$

$$\therefore \quad 2\gamma(\mathbf{h}) = 2\big(C(\mathbf{0}) - C(\mathbf{h})\big) \tag{1.7}$$

which implies that $Z(\mathbf{s})$ is also intrinsically stationary. In this case the variogram is bounded and $C(\mathbf{0})$ is known as the sill of the variogram.

For processes whose means and variances exist, strict stationarity implies second order stationarity but the converse is not always true. Increment stationarity implies intrinsic stationarity but the converse is not always true. Second order stationarity implies intrinsic stationarity, but the converse is not always true. Strict stationarity implies increment stationarity, but the converse is not always true. A familiar example here is Brownian motion which is increment stationary but not strictly stationary (Christensen 1990).

1.4. Probability Model

In the stationarity framework, the regionalized variable $Z(\mathbf{s})$ can be modelled as the sum of a deterministic part $\mu(\mathbf{s})$, a mean function called the drift, and a zero-mean stochastic process $\delta(\mathbf{s})$ called the residual random part. The mean function can be thought of as the large scale variation representing the variable's global trend over D. The zero-mean stochastic process can be thought of as the small scale variation representing the spatial dependence after the trend is removed. This type of model is known as a stochastic model or a probabilistic model (Cressie 1991) and can be expressed as equation (1.8)

6

$$Z(\mathbf{s}) = \mu(\mathbf{s}) + \delta(\mathbf{s}), \qquad \mathbf{s} \in D \tag{1.8}$$

If $\mu(\mathbf{s})$ varies slightly over a region, or the expected value of the drift is not constant but varies over the region D then the drift may possibly be expressed as a linear combination of suitable base functions and we can write

$$Z(\mathbf{s}) = \sum_{j=1}^{p+1} \beta_{j-1} \, f_{j-1}(\mathbf{s}) + \delta(\mathbf{s}) \tag{1.9}$$

where $\mathbf{s}$ is a location index, β_{j-1}, j = 1,...,p+1 are called the unknown drift coefficients, and $f_{j-1}(\mathbf{s})$, j=1,...,p+1 are known base functions at $\mathbf{s}$. The model expressed in equation (1.9) is known as the *universal kriging* model in geostatistics. A special case of the universal kriging model is the *ordinary kriging* model, when $\mu(\mathbf{s})$ is unknown but constant and the *simple kriging* model when the mean function is known. This will be discussed in detail in chapter 4.

2. Exploratory Data Analysis

2.1. Overview

Exploratory Data Analysis (EDA) is an important step in the analysis of any data. It can be used to detect outliers (atypical observations), symmetric, non-Gaussian and skew distributions and also to check stationarity assumptions and look at the spatial continuity of the data. One of the aims of EDA is to detect these effects before the modelling and prediction. The idea is to apply EDA in order to obtain a better geostatistical model and hence improve the description of the underlying spatial structure.

EDA deals with the identification of distributional properties of the data (such as location, spread, skewness, and tail properties), and provides resistant methods and graphical tools for better description, modelling and interpolation. The methods of EDA are also useful in model evaluation, hypothesis formulation and methodology for hypothesis testing (Haining 1990).

EDA should be distinguished from Confirmatory Data Analysis (CDA) which is concerned with significance and hypothesis testing, estimation and prediction after the model has been fitted and any underlying model assumptions have been checked. The use of EDA may be particularly important in areas where formal CDA test procedures

are difficult to implement or are of unknown reliability. It may detect not only the data problems mentioned above but also problems that are associated with the spatial organization of the data and which may affect the fitting of the spatial model. Resistant estimation procedures are designed to handle the situations where the data satisfy the assumptions made but contain a small number of local or global outliers, whereas robust methods are designed to handle the situations where the data violate the model assumptions and contain atypical observations. These outliers may produce errors and lead to the selection of inappropriate models.

For example, to remove the problem of trend in the data, Cressie (1991) suggests the use of the median polish. The approach is to apply median polish to the data to obtain a set of residuals that are without trend. The residuals from this resistant fit then can be used for further analysis to obtain a good model for interpolation.

Some methods of EDA will be illustrated here on a data set given in a different context by Cressie (1986). The measurements are taken from an ore body in Australia and consist of Iron Ore ($\% Fe_2 O_3$) data modified by multiplying by and adding unspecified constants. The data were sampled on a regulary spaced grid of 50 metres by 50 metres. Along with their spatial locations, the modified data, called here The Iron Ore Data, are given as Appendix A.1

2.2. Stem-and-Leaf plots

A stem and leaf plot is one of the simplest ways of providing a graphical representation of the overall structure of the data. It provides some idea of shape, distribution and the location of global outliers. However a stem-and-leaf plot does not provide any indication of local outliers or the local configuration of the data. The stem-and-leaf plot given in figure 2.1 below is for the iron ore data and was obtained by using MINITAB. The second and the third columns give the stem and leaves respectively. For example in the first row 45 stem and 46 leaves indicate 45.4% and 45.6% iron ore respectively. The first column provides a count of the number of leaves.

Stem and leaf plot of iron ore data

```
2     45  46
3     46  9
8     47  12558
9     48  0
14    50  03455
23    51  233455699
31    52  01125689
38    53  1255557
47    54  012333348
(11)  55  00111477899
54    56  111122345667778999
37    57  11345668889
26    58  112344566777999
11    59  001116677
2     60  38
```

Figure 2.1

We can see from Figure 2.1 that the distribution is negatively skewed. Because the actual data are retained in a stem and leaf plot, the data values are easily viewed and global outliers can be noted. There are no global outliers in this particular data set.

However, a global outlier can be seen in the stem-and-leaf plot in Figure 2.2 which uses the coalash data (Cressie 1991) given in Appendix A.3. The second and the third columns give the stem and leaves respectively. For example in the first row 7 stem and 003 leaves indicate 7.0%, 7.0% and 7.3% coalash data respectively. The first column provides a count of the number of leaves.

Stem and leaf plot for coalash data

```
   3   7   003
  14   7   66678888999
  27   8   0011122222223
  51   8   55666666678888889999999
  84   9   00000000111112222223333333344444
 (36)  9   555555666666667778888888888999999999
  88  10   00000000111111122222233333344444444
  55  10   56666677777788888899999
  32  11   00001111122222223344
  13  11   5666789
   6  12
   6  12   578
   3  13   11
   1  13
   1  14
   1  14
   1  15
   1  15
   1  16
   1  16
   1  17
   1  17   6
```

Figure 2.2

11

2.3 Box Plots

A boxplot is a graphical display of the five-number summary consisting of maximum, minimum, median, upper and lower quartiles. In the case of the iron ore data this five-number summary is 60.8, 45.4, 55.4, 57.8 and 52.5 respectively and the box plot of this data was obtained using MINITAB and is given in figure 2.3.
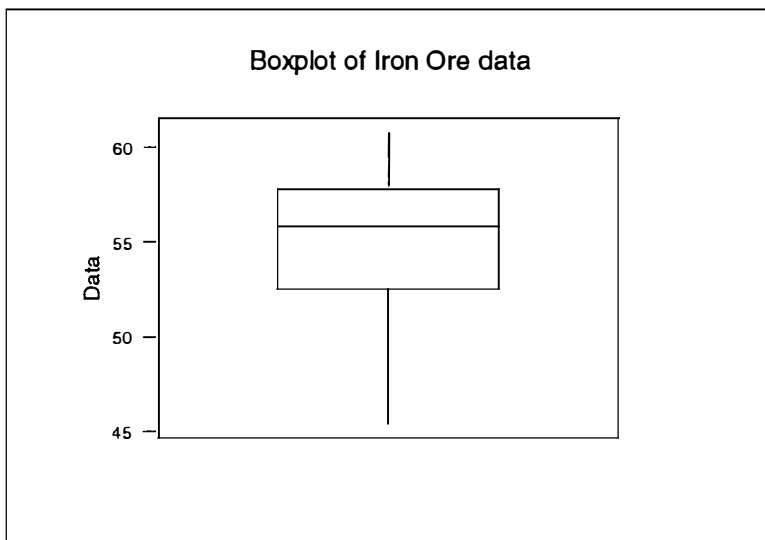
Figure 2.3

2.4 Surface and Contour Plots

A surface plot displays a three-dimensional surface. The data values are displayed as the z co-ordinate while the data locations are given as the x and y coordinates respectively. The surface plot of the cropped field Available Phosphate data (listed in Appendix A.4 and analysed in detail in section 6.2) was obtained using MINITAB and is given in figure 2.4.
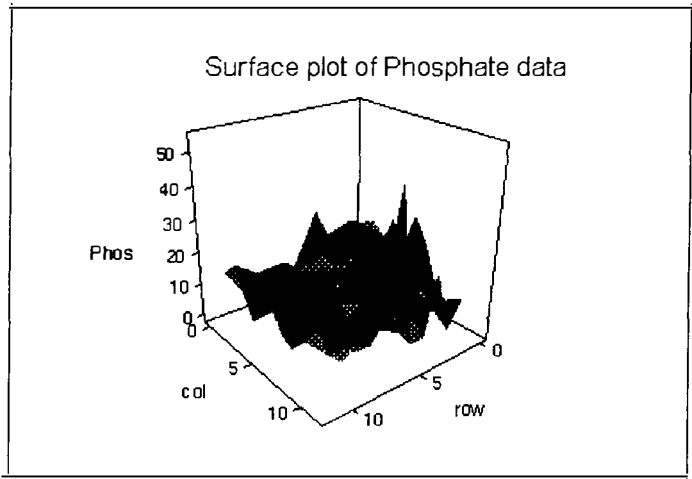
Figure 2.4

A contour plot depicts contour curves for the relevant surface. A contour plot for the iron ore data was obtained using MINITAB and is displayed in figure 2.5.

Figure 2.5

13
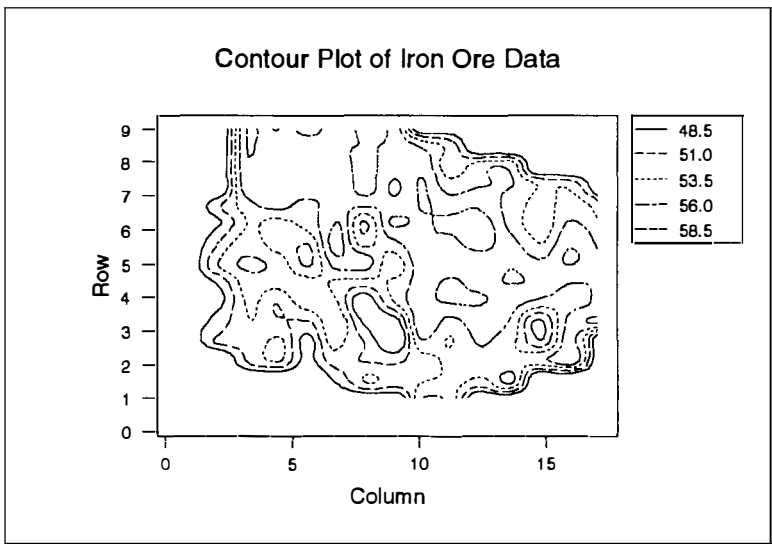
2.5 Post Plots

Post plots provide a graphical display of the data, in which we can see not only the values but also the actual locations of the sample points. The sample locations can be regularly or irregularly spaced. Post plots allow us to get an idea of the general trend of the values in different directions. Figure 2.6 below shows the post plot of the iron ore data. It has been obtained by using GEOEAS. The four different symbols represent intervals between adjacent quartiles. It is possible via GEOEAS to obtain a plot with the actual values listed at each location. From Figure 2.6 we can visualise a trend in the N-S direction, which is gradually increasing from the South to North (from bottom to top). The distance between two locations was scaled to lag 1 from 50 metres in this graph.

Post plot of Iron Ore Data

1st Quartile:	$45.400 \leq + \leq 52.500$
2nd Quartile:	$52.500 < \times \leq 55.800$
3rd Quartile:	$55.800 < \square \leq 57.800$
4th Quartile:	$57.800 < \ast \leq 60.800$
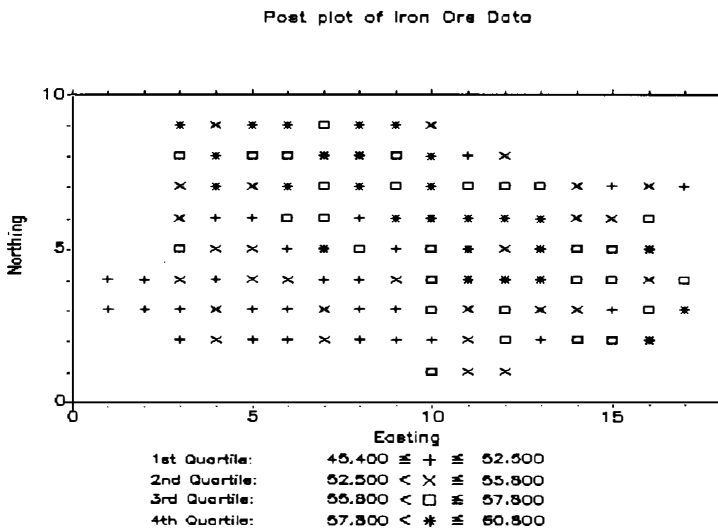
Figure 2.6

2.6 Bivariate Plots

A bivariate plot is a graphical description of pairs of sample values. Each pair can be represented as a point (x,y), and one plots Z(**s**+h**e**) against Z(**s**) for a particular lag h and unit vector **e**. Bivariate plots are useful to draw attention to atypical observations and also exhibit the spatial continuity of the data. This technique is illustrated here using the iron ore data with h=1 and **e**=(1,0) a unit vector in the E-W direction. There are no global outliers evident in this plot.

Bivariate plot of Iron Ore Data
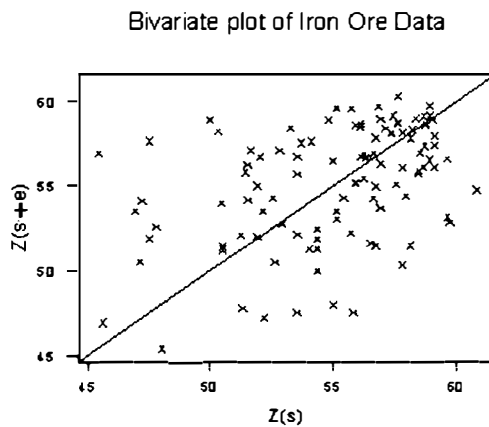
Figure 2.7

2.7 Row and column means and medians plots

When the data are given on a rectangular or square grid then this method is appropriate to provide graphical to help to get an idea of trend by using both mean and median plots across the rows and across the columns (Cressie 1991). It is important in the preliminary steps to look for data outliers, spatial pattern and if possible to remove

15

any trend or drift present in the data. The presence of possible outliers is indicated by large (positive and negative) values on row or column mean and median plots. If outliers appear in the data then the effect of these may be able to be minimised by an appropriate transformation. There are two important aspects of using median and mean. From the EDA point of view the median is a resistant statistic whereas the mean is a nonresistant statistic (Cressie 1991). The comparison between mean values and median values may highlight atypical observations in rows or columns. For example if the differences between the mean values and the median values are too large or too small then these rows or columns may contain atypical observations.

Column means and medians as well as row means and medians were plotted in order to check the directional trend of the iron ore data. The column means and the column medians were plotted in figure 2.8 and the row means and medians in figure 2.9. In these the median is denoted by a solid circle and the mean is denoted by a cross. These correspond to calculations in the E-W and the N-S directions respectively. In the E-W direction there are two atypical observations (bottom left) and there is no trend. In figure 2.9, the estimated row mean and row median plots, the data exhibit nonstationarity and a strong linear trend exists in the N-S direction. This nonstationarity in the mean, or trend, can be minimised by applying a resistant technique (for example median polish) to obtain a new data set (the residuals) which hopefully will provide a symmetric distribution which also satisfies the intrinsic

hypothesis. These plots are produced by using MINITAB and are given in Figures 2.8

and 2.9 respectively.
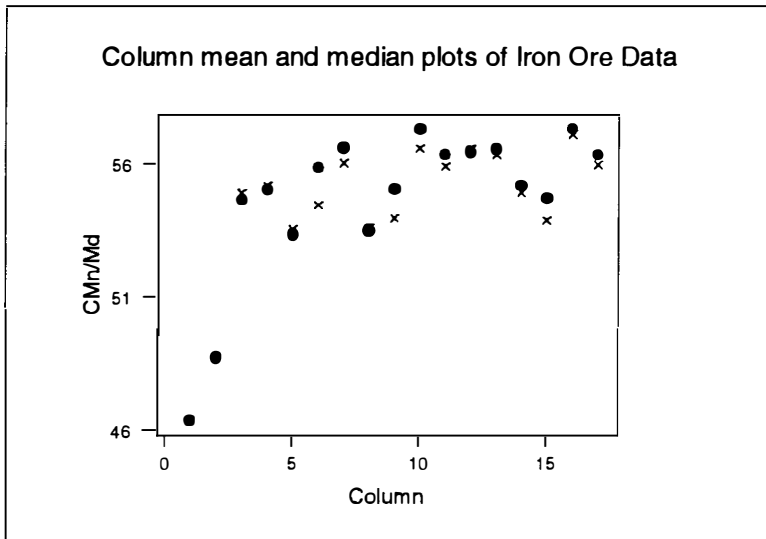
Figure 2.8
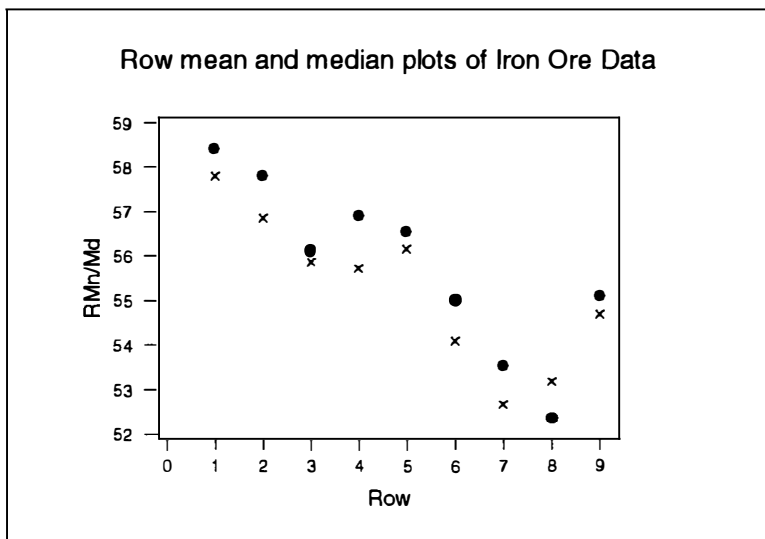
Figure 2.9

2.8 Median Polish

Median polish (Velleman and Hoaglin 1981) is similar to two-way analysis of variance and both use additive models. The difference between these two is that two-way analysis of variance fits the model by finding row and column means whereas median polish fits the model by finding row and column medians. The additive model is one in which each cell is represented as the sum of four components: a constant common term, row effects, column effects and residuals.

Median polish is resistant to atypical observations and is used when the data are nonstationary in the mean. Such spatial outliers could have a serious distorting effect on the values of non-resistant statistical estimators of the statistical properties. It would be inappropriate to fit a variogram in the presence of trend in the data. It may be possible to minimise the trend by performing a sweep across the rows and another sweep down the columns. The subtraction of these row medians from the data values is called half sweep and subtraction of these column median from the numbers in its column obtained from the half sweep called one complete sweep. Then the data denoted by D_{ij} can be modelled as

$$D_{ij} = C + RE_i + CE_j + R_{ij} \qquad (2.2)$$

where C, RE_i CE_j and R_{ij} are common term, row effects, column effects and residuals respectively.

The residuals from median polish for the iron ore data were obtained for fifteen iterations using MINITAB and is given in Appendix A.2. These residuals satisfy a Gaussian distribution exhibiting no systematic pattern and appear to be compatible with the intrinsic hypothesis.

3.   Variograms

3.1. Overview

In this chapter we shall consider a number of variogram models and also look at ways of estimating $\gamma(\mathbf{h})$ from the data. We shall refer to such an estimate as the experimental or empirical variogram. We shall consider mainly the so-called classical estimator introduced by Matheron (1963b) but we shall consider also one of the robust estimators introduced by Cressie and Hawkins (1980). The former is usually denoted by $\hat{\gamma}(\mathbf{h})$ and the latter by $\bar{\gamma}(\mathbf{h})$ but, when the meaning is clear, we shall continue to write simply $\gamma(\mathbf{h})$. The experimental variogram is used mainly for obtaining an idea of the spatial continuity in different directions by measuring the spatial autocorrelation of the data. These experimental variograms are also useful for the identification of drift and the determination of anisotropy of the data (Furr and Hampen 1986). The (theoretical) variogram model will be denoted by $2\gamma(\mathbf{h},\theta)$. The type of theoretical model depends upon the parameters θ. Once an appropriate theoretical model is chosen then the parameters of this model can be estimated to fit the points of the experimental variogram. The estimated parameters in the fitted models play an important role in the subsequent kriging methods of interpolation. Details will be given in chapter 4.

If the variogram at first increases with lag but later levels off then we have what is known as a bounded model. In this case θ takes the form $\theta = (c_0, c, a)$ where

c_0, c and a are called the nugget effect, partial sill and range respectively. In addition, $c_0 + c$ is called the sill and is the least upper bound of the model values. If the sill in direction $\dfrac{\mathbf{h}}{\|\mathbf{h}\|}$ is attained by $2\gamma(\mathbf{h})$ at a finite distance a, then $2\gamma(\mathbf{h})$ is said to have range a in that direction. However the variogram may have a sill but attain it only asymptotically. In this case we take as the (effective) range the distance at which 95% of the sill is reached (Zimmerman 1993). The range of the variogram is important because it marks the limit of the spatial autocorrelation of the variable concerned. As noted in section 1.3 we have $\gamma(\mathbf{0}) = 0$. If $\gamma(\mathbf{h}) = c_0 > 0$ as $\mathbf{h} \to 0$ then c_0 is the nugget effect (Matheron 1962). It is generally thought that this discontinuity at the origin is due to microscale variation. However Cressie (1991) suggests that this is an over-simplification and that c_0 should be split up as $c_0 = c_{MS} + c_{ME}$ where c_{MS} is due to microscale variation and c_{ME} is the measurment error.

In the unbounded model, the variogram increases with increase in lag. In this case θ is given by $\theta = (c_0, b^{'})$ where c_0 is again the nugget effect and b is the slope of the model.

For measuring the spatial continuity of the data, the behaviour of variogram near the origin is very informative for subsequent interpolation. Matheron (1971b) categorized some important types of such behaviour and these have been summarised by Cressie (1991) as follows.

21

(i)   The variogram is continuous at the origin.   Then $Z(\mathbf{s})$ is L_2-continuous ($E[Z(\mathbf{s}+\mathbf{h})-Z(\mathbf{s})]^2 \to 0$ if and only if $2\gamma(\mathbf{h}) \to 0$, as $\|\mathbf{h}\| \to 0$).

(ii)   The variogram does not approach zero as the lag approaches zero. Then $Z(\mathbf{s})$ is not L_2 continuous. This discontinuity of $\gamma(\mathbf{h})$ at the origin is the nugget effect discussed above.

(iii) $\gamma(\mathbf{h})$ is a positive constant (except of course at $\mathbf{h}=0$).   Then $Z(\mathbf{s}_1)$ and $Z(\mathbf{s}_2)$ are uncorrelated for any $\mathbf{s}_1 \neq \mathbf{s}_2$. Then $Z(\mathbf{s})$ is called white noise.

Also the variogram $2\gamma(\mathbf{h})$ must satisfy the property of conditional negative definiteness which means that for any locations s_i i=1, 2, ..., m and real numbers a_i, i=1,2 , ..., m satisfying $\sum_{i=1}^{m} a_i = 0$, it follows that

$$\sum_{i=1}^{m}\sum_{j=1}^{m} a_i a_j \gamma(\mathbf{s}_i - \mathbf{s}_j) \leq 0$$

3.2. Variogram Models

In the bounded case there are three forms of simple model generally considered. They can be used singly or the model used can be a combination of two or more simple models (of the same or different types). The combined model is a complex model and is used to get better fit. The three simple bounded models in frequent use are the Spherical model, the Exponential model and the Gaussian model. In each of the first

22

two models the variogram exhibits linear-type behaviour near the origin and approaches its sill asymptotically.

The spherical model appears to be model most commonly used and is given by

$$\gamma(\mathbf{h},\theta) = \begin{cases} 0 & , \quad \mathbf{h}=0 \\ c_0 + c_s \left\{ \frac{3}{2}\left(\frac{\|\mathbf{h}\|}{a_S}\right) - \left(\frac{1}{2}\right)\left(\frac{\|\mathbf{h}\|}{a_S}\right)^3 \right\} & , \quad 0 < \|\mathbf{h}\| < a_s \\ c_0 + c_s & , \quad \|\mathbf{h}\| > a_s \end{cases} \tag{3.1}$$

where the parameter $\theta = (c_0, c_s a_s)'$ where $c_o \ge 0, c_s \ge 0, and\, a_s \ge 0$ and the partial sill c and the range a have the subscript s to indicate the spherical model is being used.

The exponential model has become very useful to statisticians because of its wide application in first-order auto-regressive and Markov processes in time series (McBratney and Webster 1986) and is given in equation (3.2).

$$\gamma(\mathbf{h};\theta) = \begin{cases} 0 & , \quad \mathbf{h}=0 \\ c_0 + c_e \left\{ 1 - \exp\left(\frac{\|\mathbf{h}\|}{a_e}\right) \right\}, & \mathbf{h} \ne 0 \end{cases} \tag{3.2}$$

$\theta = (c_0, c_e a_e)'$ where $c_o \ge 0, c_e \ge 0, and\, a_e \ge 0.$ In this case c and a have the subscript e. However Isaaks and Srivastava (1989) write the exponential model with 3h in place of h in equation (3.2) and this is used by GEOEAS package also.

The Gaussian model also approaches its sill asymptotically but it exhibits parabolic-type behaviour near the origin. The Gaussian model is given by

$$\gamma(\mathbf{h};\theta) = \begin{cases} 0 & , \quad \mathbf{h}=0 \\ c_0 + c_g \left\{ 1 - \exp\left(-\dfrac{\mathbf{h}^2}{a_g^2} \right) \right\}, & \mathbf{h} \neq 0 \end{cases} \tag{3.3}$$

One example of the unbounded case is the linear model. This is often used in combination with one or more of the simple bounded models. The linear model is given by

$$\gamma(\mathbf{h};\theta) = \begin{cases} 0, & \mathbf{h}=0 \\ c_0 + b_l \|\mathbf{h}\|, & \mathbf{h} \neq 0 \end{cases} \tag{3.4}$$

$\theta = (c_{o,} b_l')$, *where* $c_{o,} b_l \geq 0$. Here b is the slope of the model with subscript l.

Sometimes more complex forms of theoretical model can be considered while fitting the experimental points. Any linear combination of variograms with positive coefficients is itself a valid variogram. The most useful combinations of the models have been applied in soil sciences by incorporating two or more different ranges of the bounded variogram model.

Single and Double Fitted Spherical Models
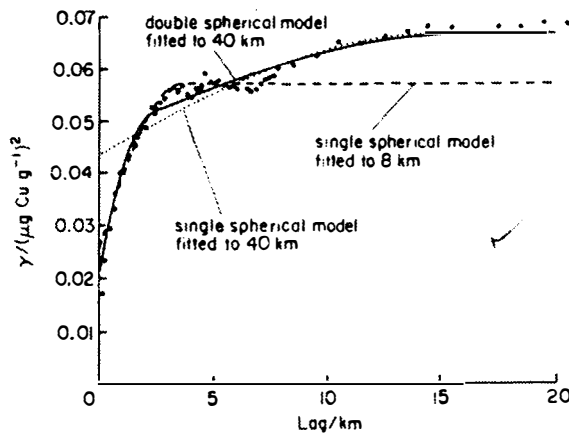
Figure 3.1

Webster and McBratney (1989) have applied such a combination in mineral exploration and soil research to describe the variation in copper and cobalt sampled from South East Scotland. Single models (dashed and dotted lines) and a double spherical model (solid line) were fitted to the experimental semivariogram of $\log_{10}$ copper concentration in the top-soil. This is illustrated in figure 3.1 which was scanned from Webster and McBratney (1989).

3.3. Sample Variograms

In practice the variogram must be estimated from the data. As mentioned in section 3.1 we shall consider two estimators in detail and refer to these as the classical estimator and the robust estimator. The former is the one used almost exclusively in practice and is in the standard geostatistical software. However, it is very sensitive to the effect of outliers. As its name suggests the robust estimator is more resistant to outliers and was introduced for this reason.

25

The classical estimator was introduced by Matheron (1962) and is given by

$$2\hat{\gamma}(\mathbf{h}) = \frac{1}{|N(\mathbf{h})|} \sum_{N(\mathbf{h})} (Z(\mathbf{s}_i) - Z(\mathbf{s}_j))^2 \qquad (3.5)$$

where the sum is over $N(\mathbf{h}) \equiv \{(i,j): \mathbf{s}_i - \mathbf{s}_j = \mathbf{h}\}$, and $|N(\mathbf{h})|$ is the number of distinct pairs in $N(\mathbf{h})$. On a regular grid, we can estimate the classical variogram for various discrete values of **h** for each pair of observations. Geostatistical software packages such as GEOEAS and GEOPACK use the classical estimator. In addition, we have written MINITAB macros for the four directions E-W, N-S, EW-NS, and NE-SW and these are listed in Appendix B.1. The classical estimator of the variogram was estimated in the E-W direction for the iron ore data using one of these MINITAB macros and is given in figure 3.2 below.
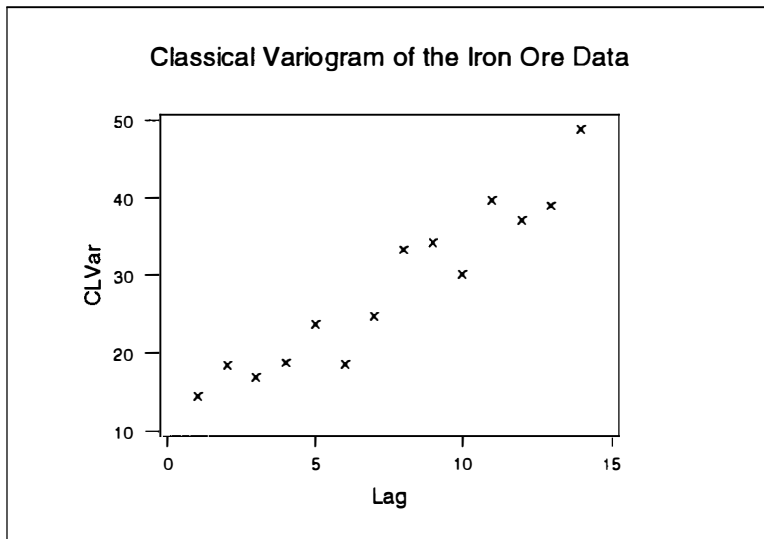
Figure 3.2

The experimental variogram obtained here shows that the variogram values are steadily increasing as the lags are increasing. This indicates a possible trend

26

in the data. However the corresponding experimental variogram estimated for the

residuals of the iron ore data obtained by median polish shows no such trend (figure

3.3). The values of the classical estimator for the original and residual iron ore data

are given in Appendices B.2 and B.3 respectively.
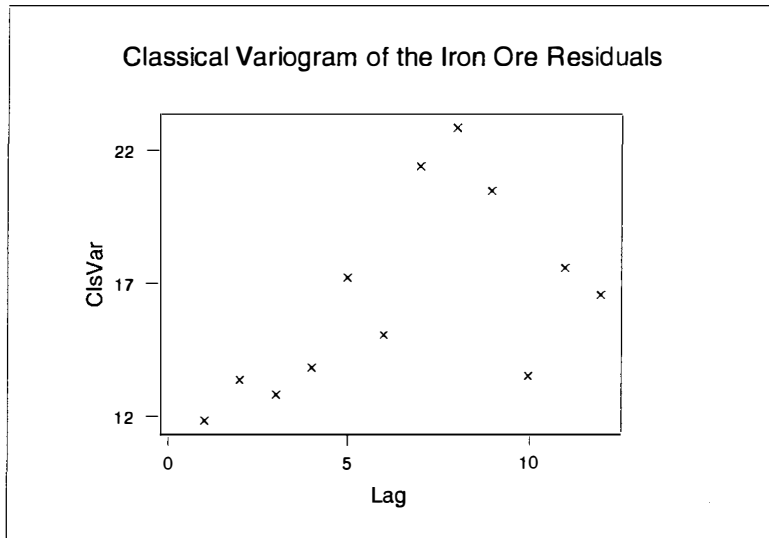
Figure 3.3

The GEOEAS package calculates the experimental semivariogram ($\gamma(\mathbf{h})$) and uses

directions and tolerances along with bandwidth to capture a resonable number of data

pairs. The tolerances use both plus and minus on the distance and on the direction of

$\mathbf{h}$. Bandwidth is used to define the width of two parallel lines and is perpendicular to

the direction of centerline to the second point in a pair. More detailed information

about the GEOEAS software package is given in the GEOEAS 1.2.1 USER GUIDE

(Englund and Sparks 1991). The variogram plot of the iron ore data was also

produced by using GEOEAS with tolerance, bandwidth, and direction 1, 20, and 0

respectively and is given in figure 3.4. The variogram values were obtained by using

GEOEAS. The values of classical and robust variograms for the original iron ore data

are given in Appendix B.4.
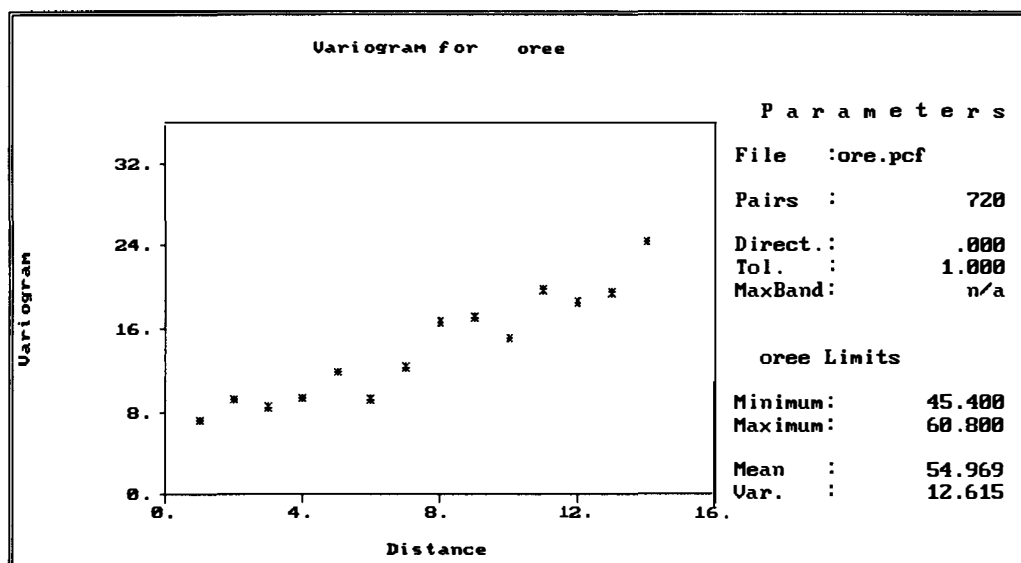
Figure 3.4

The classical estimator is very senisitive to the effect of outliers. It gives an unbiased

estimate when Z(**s**) satisfies the intrinsic hypothesis. However the estimate has $O(\frac{1}{n})$

bias when Z(**s**) is second order stationary (Cressie 1991). Some improvement was

made to deal with these outliers. For example, Cressie and Hawkins (1980)

introduced several robust methods of estimating the variogram; Unlu et al (1990)

compared various methods of fitting a model to the empirical variogram. The word

robust is used to describe inferences of procedures that are stable when the model

assumptions depart from those of a central model (Cressie 1991). The robust

estimators mentioned above are given in (3.6) and (3.7).

28

$$2\bar{\gamma}(\mathbf{h}) \equiv \frac{\left\{\frac{1}{|N(\mathbf{h})|} \sum_{N(\mathbf{h})} \left|Z(\mathbf{s}_i) - Z(\mathbf{s}_j)\right|^{\frac{1}{2}}\right\}^4}{\left(\frac{0.457 + 0.494}{N(\mathbf{h})}\right)} \tag{3.6}$$

$$\text{and} \quad 2\tilde{\gamma}(\mathbf{h}) \equiv \frac{\left[med\left\{\left|Z(\mathbf{s}_i) - Z(\mathbf{s}_j)\right|^{\frac{1}{2}} : (\mathbf{s}_i, \mathbf{s}_j) \in N(\mathbf{h})\right\}\right]^4}{B(\mathbf{h})} \tag{3.7}$$

where med{.} denotes the median of the sequence {.} and B(**h**) is a correction for bias. Thus various location estimators can be applied to $\left\{\left|Z(\mathbf{s}_i) - Z(\mathbf{s}_j)\right|^{\frac{1}{2}} : (\mathbf{s}_i, \mathbf{s}_j) \in N(\mathbf{h})\right\}$, which after normalising for bias, yield robust variogram estimators. Cressie and Hawkins (1980) proved that absolute transformation produced more Gaussian-like process in comparison to other transformations, such as log. A Gaussian process Z(**s**), (Z(**s**+**h**)-Z(**s**))² is distributed as $2\gamma(\mathbf{h}). \chi^2$ where χ^2 is a chi-square random variable with 1 degree of freedom. Thus $2\gamma(\mathbf{h})$ is the first moment of a highly skewed random variable. The robust variogram for the residuals of the iron ore data was computed using a MINITAB macro given in Appendix B.1 in the E-W direction and is given in figure 3.5. The robust estimator of the variogram in figure 3.6, for the original ore data, has lower values than the classical variogram estimator because it downweights potentially atypical observations. However changing to the robust variogram does not remove the trend present in this data. It is very obvious that there is no trend in the figure 3.6, the robust estimator for the residuals of the or edata. The values for figure 3.5

29

and 3.6 are given in Appendices B.2 and B.3 respectively.

Robust Variogram of the Iron Ore Data

Figure 3.5

Robust Variogram of the Iron Ore Residuals
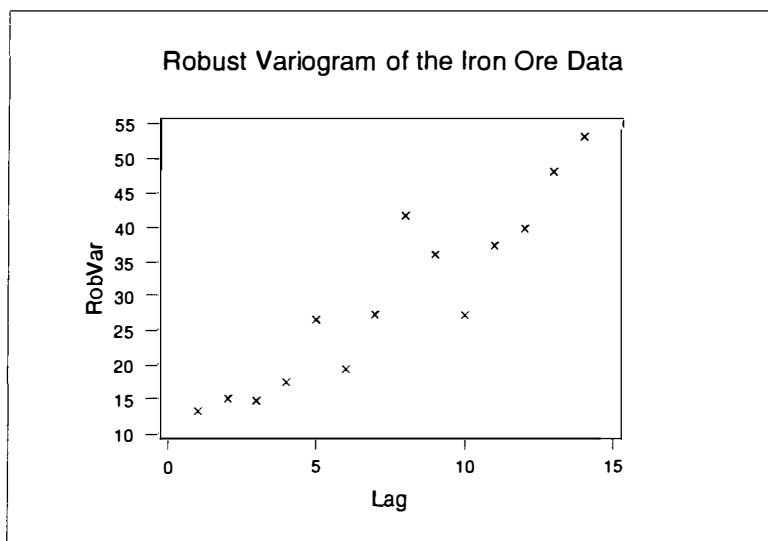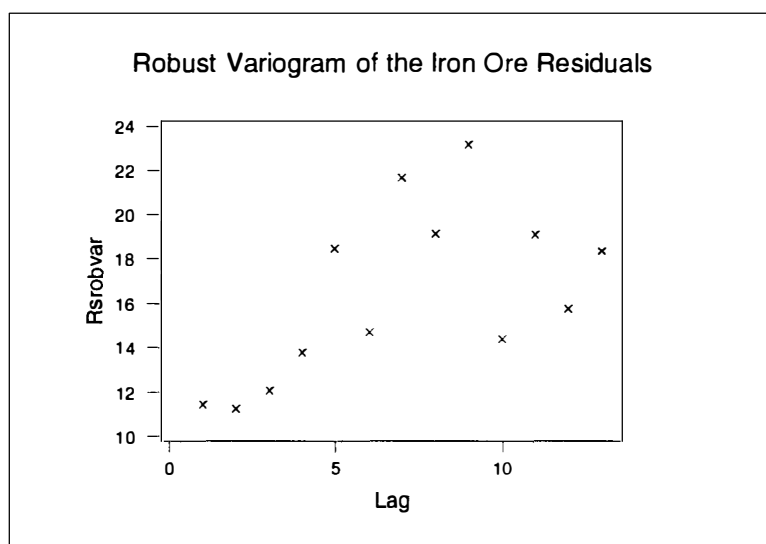
Figure 3.6

Resistant and exploratory techniques were also used to improve the estimation of variograms (Cressie, Hamlett and Horton 1986).

3.4. Fitting a Variogram Model

Fitting a model is an important process in all applied mathematics and is no less so in geostatistics. A good fit of the variogram near the origin is especially important. (Cressie 1991; Stein 1988). In earlier papers Burgess and Webster (1980) and Webster and Oliver (1990) have illustrated experimental variograms of soil variables in several regions and fitted theoretical variograms to them. There are some important statistical procedures for fitting the theoretical model to the experimental variograms as given below.

Fitting a variogram 'by eye' (trial and error method) is a commonly used method generally available in geostatistical software such as GEOEAS and GEOPACK. This method requires an experienced person to fit the variogram model accurately. As used in GEOEAS it provides estimates for the parameters and gives a general picture of the fitted model. The values obtained can be used as the initial values when fitting the model by other (iterative) methods.

To produce a variogram plot in GEOEAS the *Prevar* module is used to create a file for data pairs and these pairs are used in the *Vario* module to produce the variogram plot. The *Vario* module is also used to superimpose variogram models on the experimental variogram until a reasonable fit is obtained. The iron ore data were used once again. The omnidirectional semivariogram plot of the iron ore data was produced using GEOEAS and the values are given in Appendices B.5 and B.6. In Figures 3.7 and 3.8 spherical and Gaussian models respectively have been fitted to this

31

omnidirectional semivariogram.
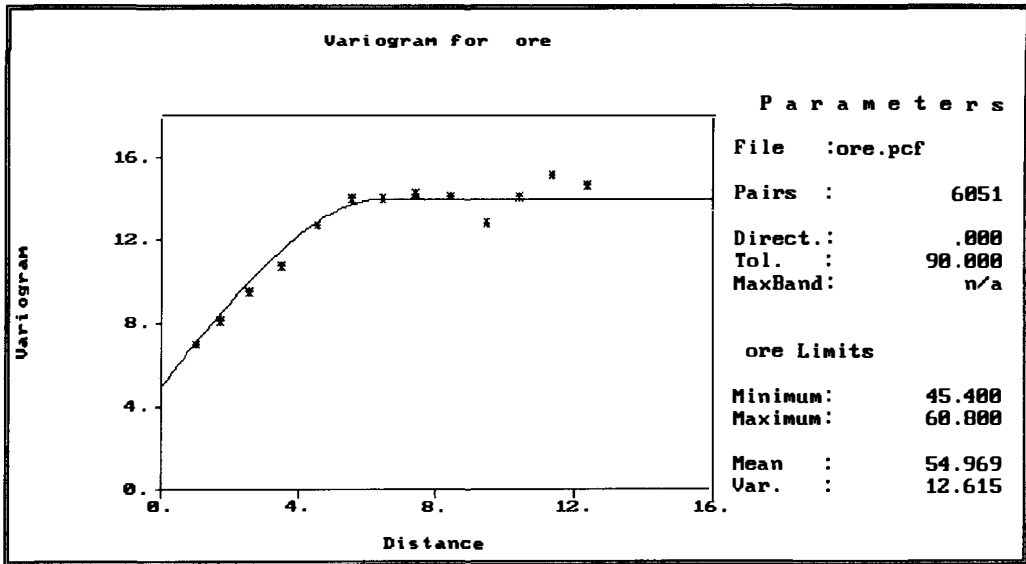
Figure 3.7
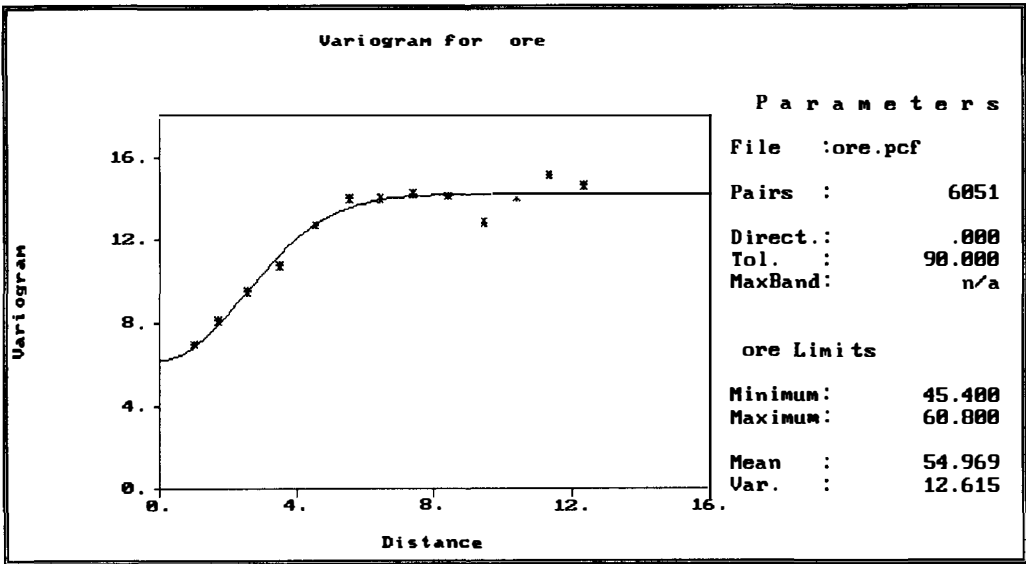
Figure 3.8

The choice of model type is particularly important and the final choice is often

determined by specialist knowledge of characteristics of the actual variable under consideration.

Several other procedures for fitting a variogram model have been proposed. The simplest is ordinary least squares. Suppose the theoretical variogram model is denoted by $2\gamma(\mathbf{h}_j;\theta)$, and the experimental variogram model is denoted by $2\hat{\gamma}(\mathbf{h}_j)$. The objective of ordinary least squares is to find the value of θ will minimise the error variance as given in (3.8) below.

$$\sum_{j=1}^{k}\left[\hat{\gamma}(\mathbf{h}_j)-\gamma(\mathbf{h}_j;\theta)\right]^2 \qquad (3.8)$$

The estimation of parameters based on least squares methods can be improved in this context by using weighted least squares which is a non-linear optimisation process for evaluating the optimal set of variogram parameters. The weighted least squares (cost criterion) approach here is to estimate θ to minimize

$$\sum_{j=1}^{k}\frac{|N(\mathbf{h}(j))|}{(\gamma(\mathbf{h}(J);\gamma))^2}\left\{\gamma^\wedge(\mathbf{h}(j))-\gamma(\mathbf{h}(j);\theta)\right\}^2 \qquad (3.9)$$

The weights chosen here are those given by Cressie (1991). One of the advantages of this option is that it automatically gives more weight to early lags with maximum number of pairs and downweights lags with a small number of pairs to produce unbiased estimated minimum variance.

33

To get a weighted least squares estimate of the parameters, the weight can be assigned to each sample. When it is included, the expression is evaluated for each observation in the data set to be analysed, and the values obtained are taken as inverse elements of the diagonal variance-covariance matrix of the dependent variable. As mentioned by McBratney and Webster (1986) Laslett has an alternative choice with the weights defined by

$$\frac{N(\mathbf{h}_j)\gamma(\mathbf{h}_j)}{\{\gamma(\mathbf{h}_j;\theta)\}^3}$$

The cost criterion suggested by Cressie in equation (3.9) is very reasonable in the sense that the smaller the theoretical variograms, the more weight the residual assigns to it. Also, the more pairs of observations $N(\mathbf{h}_j)$ there are, the more weight is assigned to it to obtain a good fit of the variogram model (Cressie 1991; Stein 1988). This is important to obtain a good fit of the variogram model near the origin. However Zhang, Eijkeren and Heemink (1995) pointed out some drawbacks in Cressie's cost criterion. These are the following.

(i) Weighted least squares uses a non-linear optimisation process to evaluate the optimal set of parameters and the sum of the weights differs in each iteration. This may lead to wrong conclusions when the cost of two iterations with different weighting factors are compared. In this case the minimization procedure may converge to local minimum or even diverge.

(ii) Weighted least squares does not yield the same cost for a positive or negative deviation. To check this, one experimental variogram $\gamma^* = 1$ with N=1 was considered. Now let if $\gamma(\mathbf{h}(i);\lambda) = .5,$ then the cost is 1, while if $\gamma(h(j) = 1.5,$ the cost is only 1/9. They proposed a new cost criterion, given in (3.10) below, to overcome these problems.

$$\sum_{j=1}^{k} \frac{|N(\mathbf{h}(j))|}{(\mathbf{h}(J))^2} \left[\hat{\gamma}(\mathbf{h}(j)) - \gamma(\mathbf{h}(j);\theta) \right]^2 \qquad (3.10)$$

It is clear that the weights are constant throughout the iterations because the weights do not depend on the parameters in (3.10), and this criterion gives the same cost for positive and negative deviations. So it removes the drawbacks while keeping the merits of Cressie's method.

The DUD and Marquardt methods are two algorithms used for fitting a variogram model by the weighted least squares method (Gotway 1992). They assign the weighting scheme to each individuals sample depending upon the numbers of distinct pairs N(**h**). The weights are defined as the ratio of N(**h**) (number of distinct pairs) and the variogram model. The sum of weights assigned to each sample must be such that it gets an unbiassed estimate. In the Marquardt method the partial derivative of the model with respect to each parameter must be defined. In any particular example one frequently needs to use these algorithms interchangebly to obtain the best estimates.

For example, Cressie (1991) computed the robust variogram from the original coalash

data in the N-S direction and fitted a spherical model to the experimental variogram using the weighted least squares method. The values of the parameters obtained are $c_0 = 0.89$, $c_s = 0.14$ and $a_s = 4.31$.

3.5. Anisotropy

We have already noted in section (1.3) that if the $2\gamma(\mathbf{h})$ ($or\ 2\,C(\mathbf{h})$) is a function only of $\|\mathbf{h}\|$ then $2\gamma(\mathbf{h})$ ($or\ 2\,C(\mathbf{h}))$ is called isotropic. Now we will consider anisotropy. If the distance between $Z(\mathbf{s})$ and $Z(\mathbf{s}+\mathbf{h})$ depends upon the direction of $\mathbf{h}$, as well as its magnitude then $2\gamma(\mathbf{h})$ ($or\ 2\,C(\mathbf{h}))$ is anisotropic. In other words, anisotropic directional variograms are obtained if the structural character of the regionalized variable (for example, grade of mineral deposit) differs in various directions. It is important to explore the possible pattern of anisotropy to understand the structural character of the data. In this case this structural character of the regionalized variable may exhibit major changes in the range or the sill or the nugget as the direction changes.

Geometric anisotropy is determined by looking at the contour map of the variogram values when the range changes with direction, while the sill remains constant. If these form an elliptical pattern then we have geometric anisotropy (Isaaks and Srivastava 1989). Sometimes this anisotropy can be corrected by a linear transformation of the lag vector $\mathbf{h}$. The anisotropic semivariogram may be detected either to the gradient of an unbounded model or the distance parameter of a bounded model and is defined by

$$\gamma(\mathbf{h},\theta) = \sqrt{\left\{A^2\cos^2(\theta-\varphi) + B^2\sin^2(\theta-\varphi)\right\}} \qquad (3.11)$$

where A is the gradient of the variogram in the direction of maximum variation, B is the gradient in the direction of minimum variation, ϕ is the angle of maximum gradient and θ is the angle between the maximum variation and the minimum variation. When $\theta = \varphi$ the equation (3.11) becomes $\gamma_1(\mathbf{h}) = A\mathbf{h}$ and when $\theta = \varphi + \pi/2$ it becomes $\gamma_2(\mathbf{h}) = B\mathbf{h}$. To measure the anisotropy ratio in this case the degree of anisotropy can be defined as the ratio of the maximum to minimum gradients as given below.

$$q = A/B = \gamma_1(\mathbf{h})/\gamma_2(\mathbf{h}) \qquad (3.12)$$

In the case of the bounded model Isaaks and Srivastava (1989) suggested handling geometric anisotropy by estimating the ratio of maximum range to the minimum range with help of a rose diagram in two dimensions. The ratio $\dfrac{A}{B}$ is a measure of the anisotropy. Once the isotropy is decided then we fit the isotropic variogram model to the omnidirectional experimental variogram. Generally a more efficient way of scaling the geometric variogram model in the bounded case is to combine more than two directional variogram models from different directions into a single function that is consistent in all directions and describes the spatial continuity. Apart from geometric anisotropy the other term one finds used is zonal anisotropy. However the definition of this differs somewhat from author to author. A common approach is to define zonal anisotropy as any kind of anisotropy except geometric anisotropy. However Isaaks and Srivastava (1989) define zonal anisotropy as the case where the sill

37

varies with direction but the range does not. Myers and Journel (1990) computed and interpreted variograms with zonal anisotropies. For the unbounded model slope anisotropy is obtained when nugget effect is constant but slope is different in different directions. A more specific classification of anisotropy, and the one we shall follow here, was introduced by Zimmerman (1993) and refers essentially to range, sill and nugget anisotropy, with slope anisotropy to be considered the unbounded model equivalent of range anisotropy.

Range (slope) anisotropy is observed when the sill and nugget effect are the same but the range (slope) varies with direction. Geometric anisotropy is thus a special case of range (slope) anisotropy. The range (slope) anisotropy which is not geometric could be called non-geometric range (slope) anisotropy (Zimmerman 1993). The graphical illustrations of slope and range anisotropy are given in figure 3.9 and figure 3.10 respectively.

Non-geometric slope anisotropy
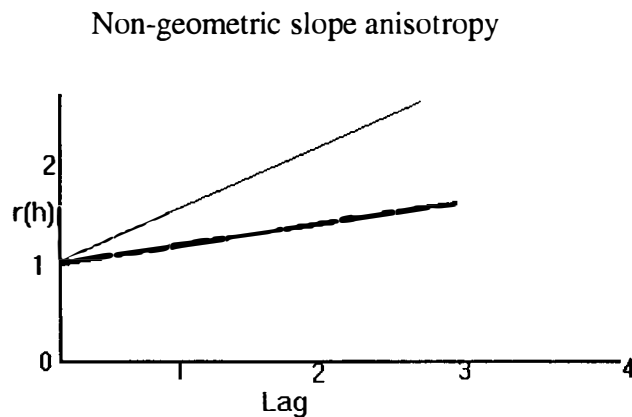
Figure 3.9
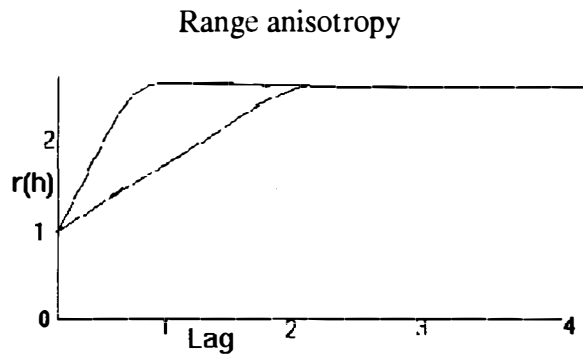
Range anisotropy

Figure 3.10

Sill anisotropy is observed when range and nugget effect are constant and the sill

varies with direction. An illustration of sill anisotropy is given in figure 3.11.

Sill anisotropy
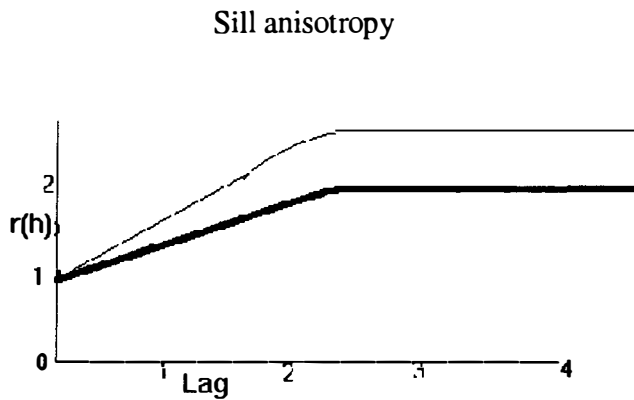
Figure 3.11

If the variogram model has a sill then the process $Z(\mathbf{s})$ is second order stationary, and

the variogram and the covariogram are related by

$$\gamma(\mathbf{h}) = C(\mathbf{0}) - C(\mathbf{h}) \tag{3.13}$$

If the sill is direction-dependent, then using (3.13), we can write

$$\lim_{\alpha \to \infty} \gamma(\alpha\mathbf{h}) = C(\mathbf{0}) - \lim_{\alpha \to \infty} C(\alpha\mathbf{h}) \tag{3.14}$$

39

Thus, if the sills are different in at least two directions then there exist vectors $\mathbf{h}_1$ and

$\mathbf{h}_2$ such that $\lim\limits_{\alpha \to \infty} C(\alpha\mathbf{h}_1) \neq \lim\limits_{\alpha \to \infty} C(\alpha\mathbf{h}_2)$ which implies further

that $\lim\limits_{\alpha \to \infty} C(\alpha\mathbf{h}) \neq 0$ for some $\mathbf{h}$. Hence if the sill varies, then the spatial correlation

between two sample points does not vanish in every direction as the distance between

two data locations increases. Zimmerman (1993) has described two possibilities:

either a second order stationarity is appropriate but the spatial correlation does not

vanish in every direction as distance increases or the second order stationary model is

simply not appropriate. In order to illustrate these two possibilites Zimmermann

(1993) has given two (constructed) examples on R^2 of processes with direction-

dependent sills. In the first of these the sill anisotropy is due to a process where the

stationarity assumptions are appropriate but the spatial correlation does not vanish in

every direction as lag increases. In the second of these the sill anisotropy is due to the

process exhibiting nonstationarity in the mean.

Despite these comments comments, geostatisticians normally deal with direction-

varying sills by representing the semivariogram as a nested model of the form

$$\gamma(\mathbf{h}) = \sum_{i=1}^{m} \gamma_i(\|A_i\mathbf{h}\|) \qquad (3.15)$$

where $A_1, ..., A_m$ are matrices that are defined in an appropriate directions and

$\dfrac{A_1(\mathbf{h})}{\|A_1\mathbf{h}\|}, ..., \dfrac{A_m\mathbf{h}}{|A_m\mathbf{h}|}$ and $\gamma_1(\mathbf{h}_1), ..., \gamma_m(\mathbf{h}_m)$ are isotropic semivariograms.

If the nugget variance varies by changing direction while the range and the sill remain constant then we have nugget anisotropy. This situation is illustrated in the figure 3.12. A pure nugget effect model entails a complete lack of spatial correlation.

Nugget effect
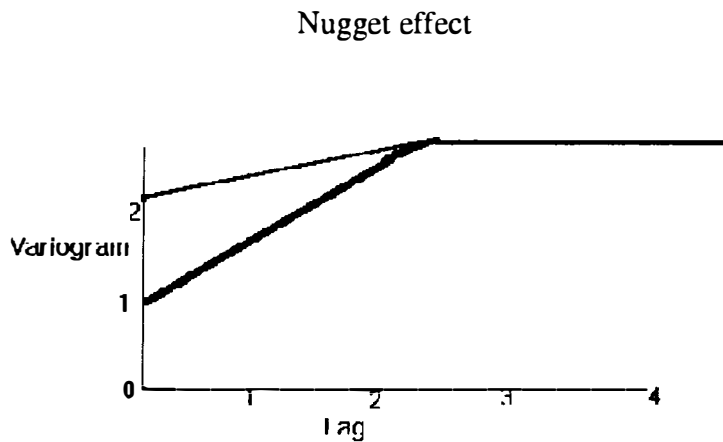
Figure 3.12

To establish anisotropy, variograms must be estimated in at least two or three different directions. In many practical studies there is some prior information about the axes of the anisotropy. For example a contour map may be able to offer some clues to the directions of minimum and maximum continuity. Once the directions of maximum and minimum continuity have been established, one needs to choose directional and distance tolerances that are large enough to allow sufficient pairs for estimating precise variograms.

We shall illustrate anisotropy here by use of the residuals of the iron ore data. Using GEOEAS nine directional semivariograms were estimated and these are given in figures 3.13 to 3.21. The directions chosen were 0 (N90E), 20 (N70E), 40 (N50E),

41

60 (N30E), 80 (N10E), 100 (N10W), 120 (N30W), 140 (N30W) and 160 (N70W). For each direction an angular tolerance of $\pm 20^0$ was chosen to capture a reasonable number of data pairs in each direction. A variogram value of 7 was fixed for each direction and the corresponding range estimated. These values are given in the appendices B.7 to B.16. A rose diagram was then prepared to check for geometric anisotropy.
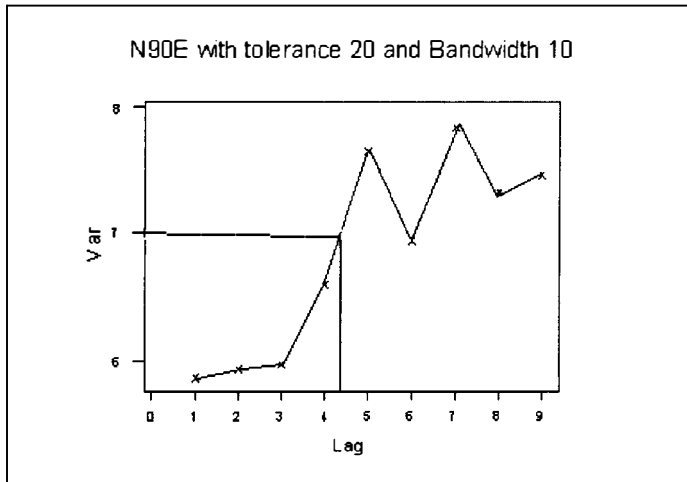
Figure 3.13
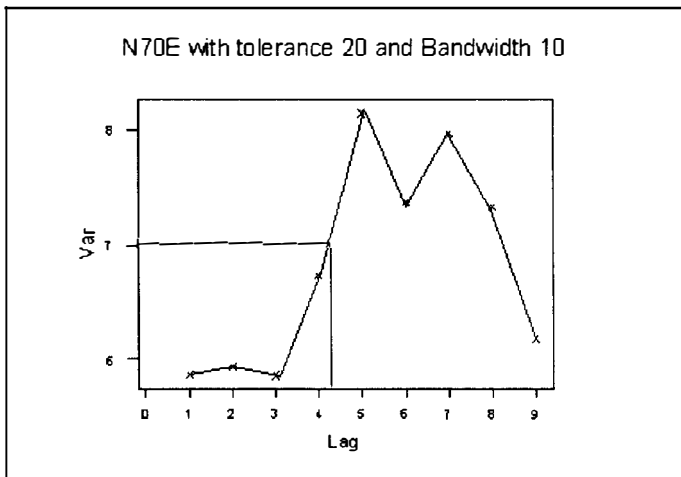
Figure 3.14
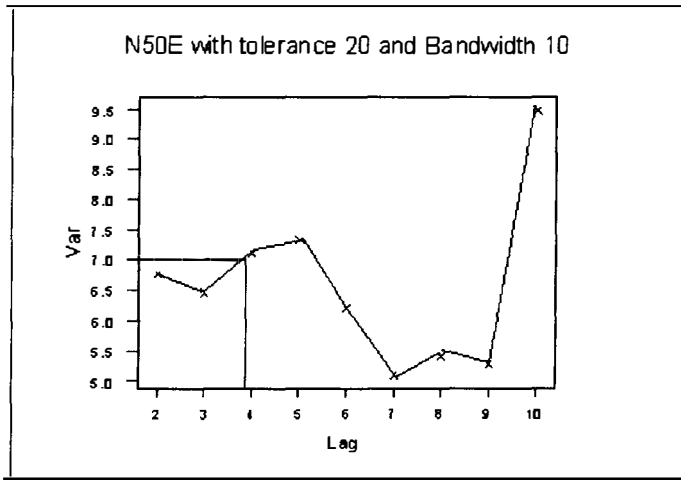
Figure 3.15
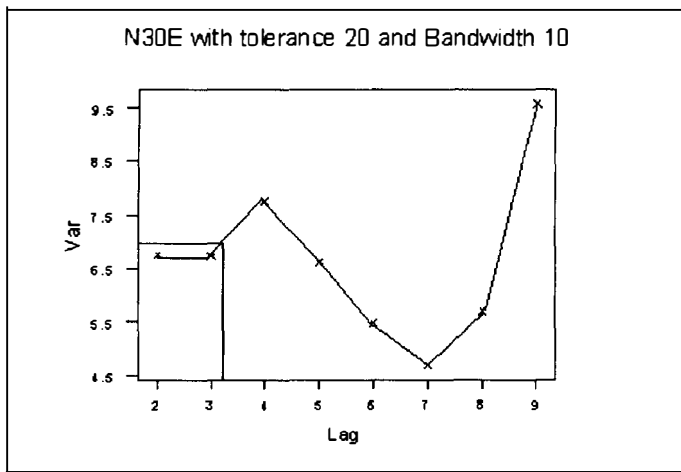
Figure 3.16

Figure 3.17

43

Figure 3.18

Figrue 3.19

Figure 3.20

44
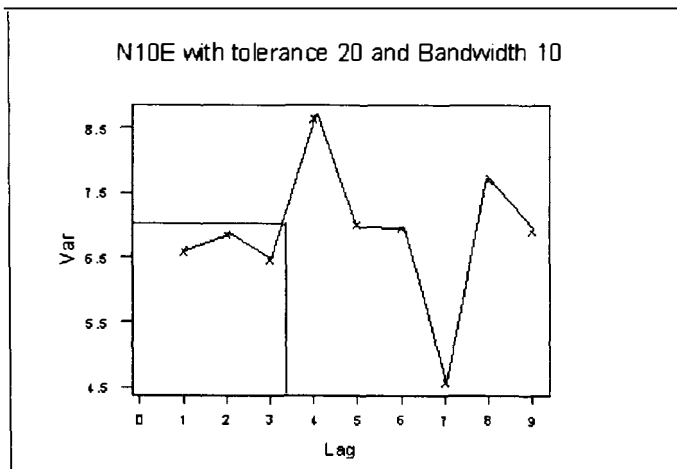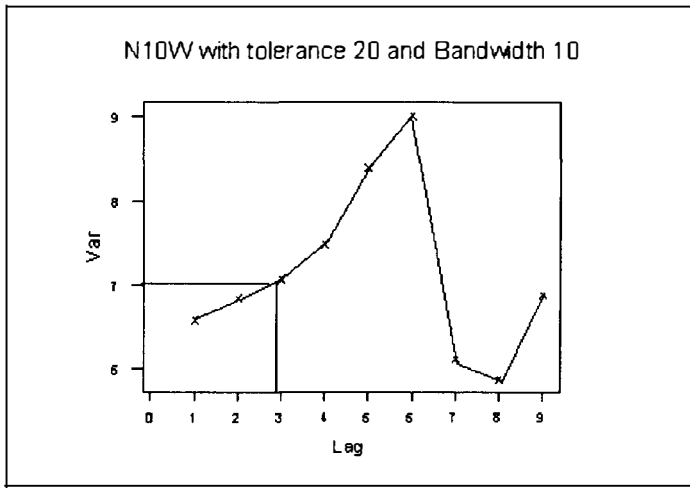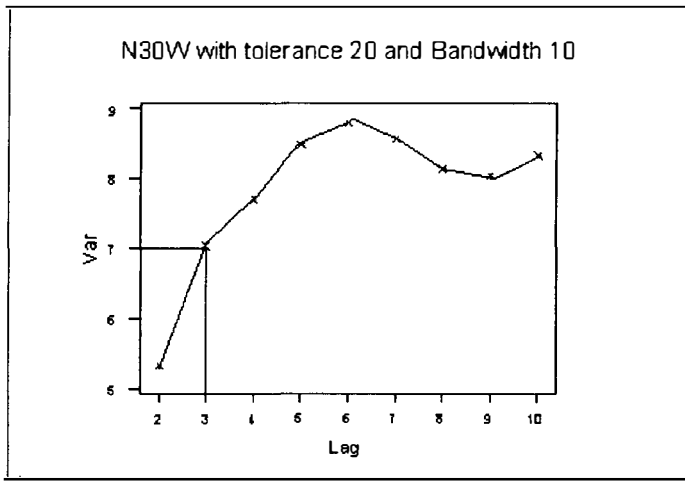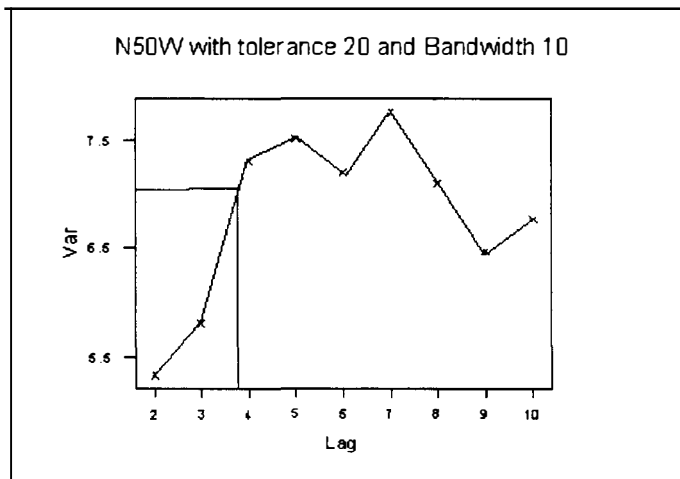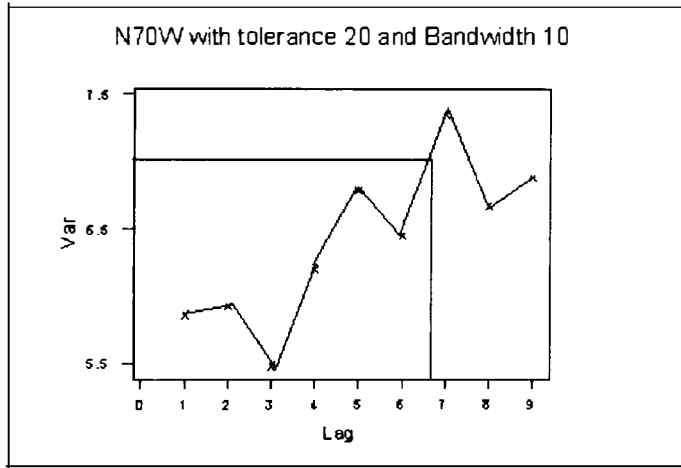
Figure 3.21

The rose diagram is shown below in figure 3.22. An ellipse can be fitted to this rose diagram with its major axis 160^0 from the East and its minor axis 70^0 from the East. This elliptical pattern indicates geometric anisotropy.

Rose diagram
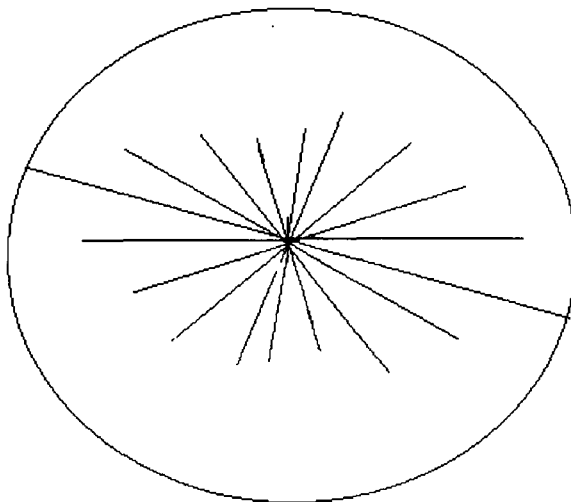
Figure 3.22

Another example of range anisotropy is given for soil data from the Riverland area of South Australia (Brooker, Winchester and Adams 1993). The purpose of the experiment was to assist in crop selection and irrigation design and scheduling. Soil data were collected from a square grid of 75 m by 75 m covering an area about 6 square kilometres. This study concentrates on two regionalized variables, the depth of topsoil in centimetres and the root zone readily available water in millimetres and determines values of them for irrigation valve areas.

The experimental semivariograms were estimated for both regionalized variables in the N-S, E-W, NW-SE and NE-SW directions (see figure 3.23). These exhibited anisotropic variation. A spherical model was fitted to the scaled experimental variogram for the both regionalized variables (see figures 3.24 and 3.25) and the relevant parameters were estimated. From figure 3.23, it can be seen that variation in the N-S direction is much more rapid than in the E-W direction. The scaling of the semivariogram can be done in several ways (by changing range, direction, and sill). In this paper it was done by scaling the distance in the E-W direction by the ratio of the range N-S to the range E-W. In other words, the ranges which are 995 metres in E-W, and 220 metres in N-S direction, obtained from two experimental semivariogram models were combined into a single function that describes the spatial variation as given in figure 3.25. The anisotropic ratio was calculated in the E-W direction by dividing maximum range by its minimum range which is equivalent to 995/220=4.5

metres (ratio of larger values of range to smaller values). Figures 3.23, 3.24 and 3.25

were scanned and reproduced from Brooker, Winchester and Adams (1993).
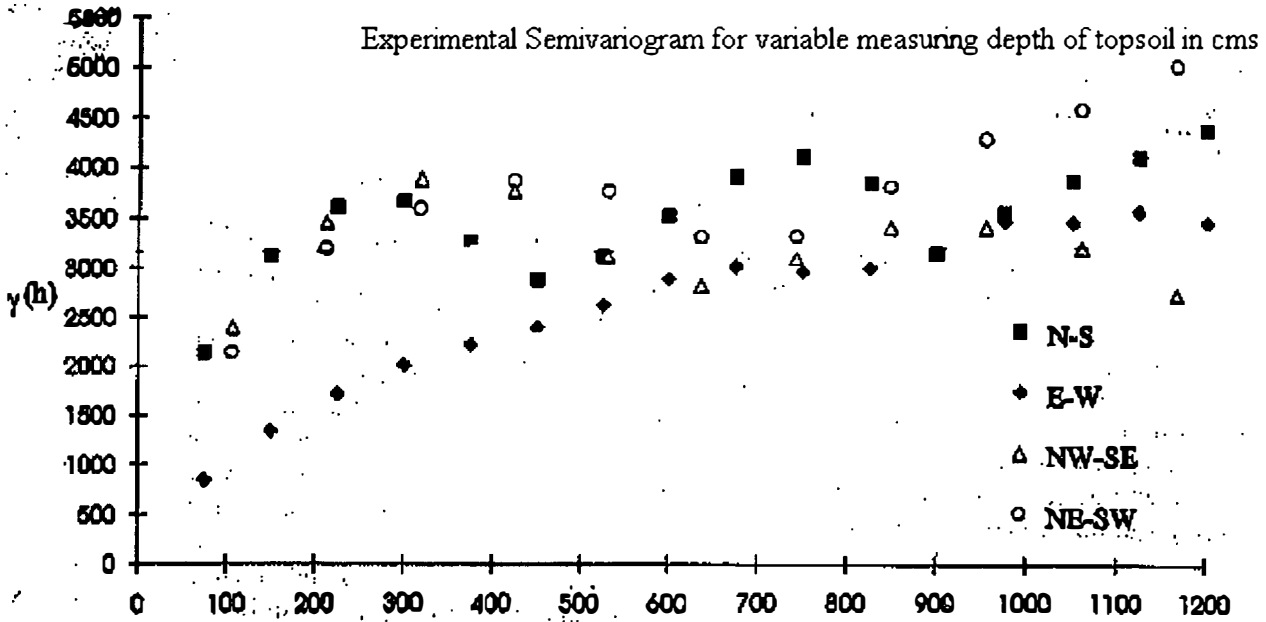
Figure 3.23
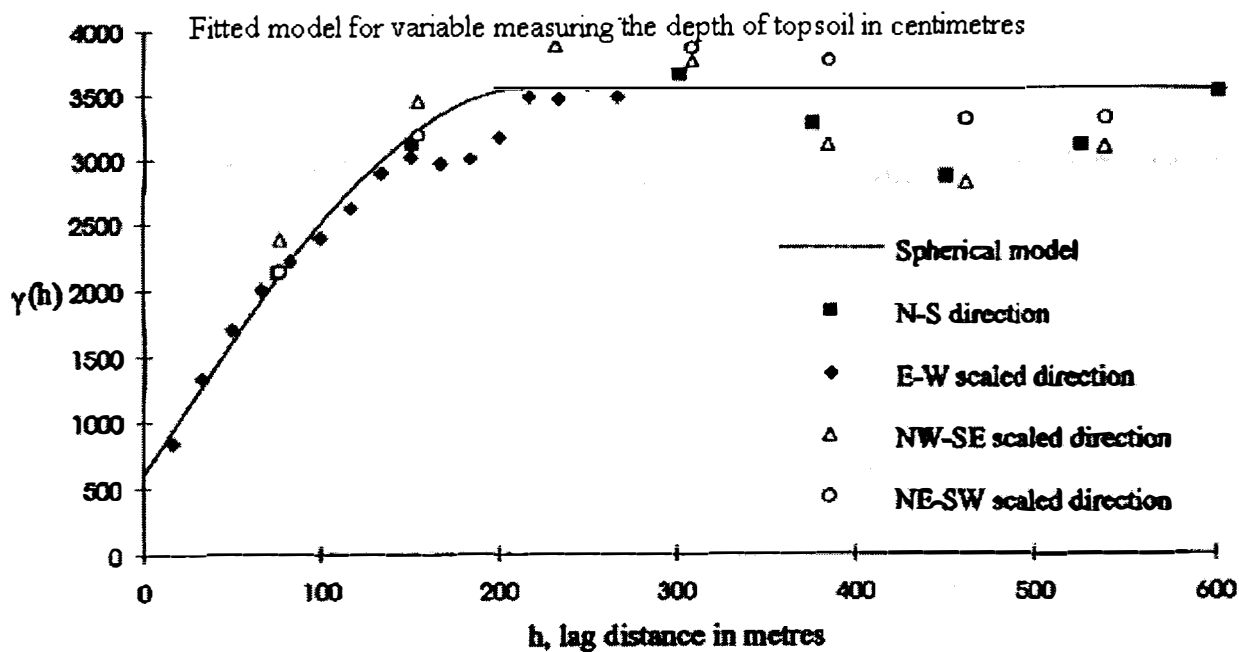
Figure 3.24
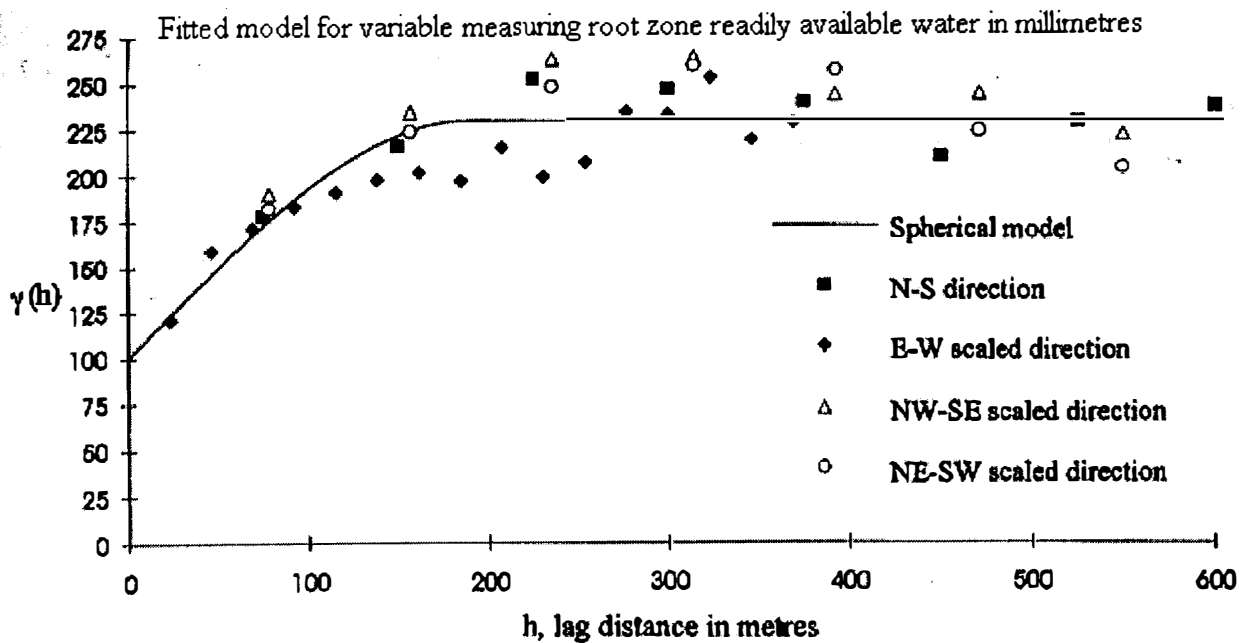
Figure 3.25

To illustrate range anisotropy, Huijbregts and Matheron (1970) computed the experimental covariograms in the E-W and the N-S directions for data measuring the gravimetric recognition in Val d'Or. This is given in Appendix A.10. A spherical covariogram model was then superimposed on each of these experimental covariograms. The range in the E-W direction was different from that in the the N-S direction so range anisotropy was confirmed. This range anisotropy was modelled overall as the sum of two spherical covariogram models.

3.6 Relative Variograms

The relative variogram is useful when a subregion of region D has a higher variance than the average variance at other subregions. In this case, the variograms can be estimated in various subregions to minimise the whole complex structure of the regions. The use of the relative variogram becomes essential when the structural region D is very complex and the data violate the stationarity assumptions (Cressie 1991). Then the region D can be divided into different subregions in each of which intrinsic stationarity exists and the variograms can be estimated for every subregion. The estimated variograms obtained from each subregion are combined and scaled to obtain a nearly constant sill. Let the region D be a disjoint union of subregions $\{ D_j : j = 1,...,k \}$ and assume that in the jth subregion the process Z is intrinsically stationary with mean μ_j and variogram $2\gamma_Z^{(j)}(\mathbf{h})$, j=1,...,k. Consider the function $Y(\mathbf{s})=g(Z(\mathbf{s}))$ where $g(Z(\mathbf{s}))$ is a smooth transformation and contains at least two

continuous derivatives. This function can be expanded in a Taylor series about $E[(Z(\mathbf{s})]$ by δ-methods (Kendall and Stuart 1969).

It was suggested (Journel and Huijbregts 1978) that some extent of nonstationarity in the mean can be modelled by the relative variogram $\dfrac{2\gamma_z^{(j)}(\mathbf{h})}{f(\mu_j)}$ for some positive function f, independent of j, where its estimation can be combined over various subregions of D with common mean μ_j. The most common relative variogram in this case is $\dfrac{2\gamma_z^{(j)}(\mathbf{h})}{\mu_j^2}$.

Isaaks and Srivastava (1989) have given an illustration of the use of relative variograms. Because of the complexity of the data the Walker Lake region was subdivided into seven subregion to investigate the consistency of the experimental variogram and the spatial estimates. They estimated relative omnidirectional variograms for each subregion.

4. Kriging

4.1. Overview

The particular methods of spatial prediction known collectively as kriging will be considered here. Kriging exploits second order stationarity properties to produce unbiased minimum variance predictions. Kriging is the prediction of unobserved values of a random variable based on a weighted average of observed values of that variable. This average function can be expressed as the second order moment of the random function in terms of a linear combination of observed values $\{Z(\mathbf{s}_i)\,i=1...,n\}$.

It is optimal because it produces an estimate of the value at an unsampled location with minimum error variance.

There are two types of kriging methods that will be considered in this chapter. These are parametric and non-parametric methods. The parametric methods include the linear techniques of simple, ordinary, universal and median-polish kriging. These linear techniques, particularly ordinary kriging, are used extensively in spatial data analysis and we shall consider them in detail in sections 4.2, 4.3 and 4.4. The non-linear techniques were developed to handle outliers but being based on normal related hypothesis, they are not always applicable in practice. The most widely used non-parameteric method, particularly in the area of ore-reserve estimation, is indicator kriging and we shall describe this method in detail in section 4.5. A comparison of different kriging estimators has been made by Boufassa and Armstrong (1989) and Laslett and McBratney (1990).

4.2. Ordinary Kriging

When the intrinsic hypothesis holds then kriging with mean unknown but constant is simply a minimum variance method of predicting certain average values of the random function Z(s) on the basis of observed values $\{Z(\mathbf{s}_i), i = 1,...,n\}$. We wish to estimate $Z(s_0)$ via a weighted linear combination of the data at n neighbouring sample points. We write

$$p(\mathbf{Z};\mathbf{s}_0) \quad = \quad \sum_{i=1}^{n} \lambda_i Z(\mathbf{s}_i) \tag{4.1}$$

Also, we need the constraint

$$\sum_{i=1}^{n} \lambda_i = 1 \tag{4.2}$$

to ensure an unbiased estimate. In this case, and using (1.2), and writing

$$R(\mathbf{s}_0) = p(\mathbf{Z};\mathbf{s}_0) - Z(\mathbf{s}_0) \tag{4.3}$$

we have

$$E[R(\mathbf{s}_0)] = E\left[\sum_{i=1}^{n} \lambda_i Z(\mathbf{s}_i) - Z(\mathbf{s}_0) \sum_{i=1}^{n} \lambda_i \right]$$

$$= \sum_{i=1}^{n} \lambda_i E[Z(\mathbf{s}_i) - Z(\mathbf{s}_0)]$$

$$= 0 \tag{4.4}$$

The kriging variance is denoted by

$$\sigma_R^2 = \text{var } (R(\mathbf{s}_0))$$

$$= E\left[(p(Z;\mathbf{s}_0) - Z(\mathbf{s}_0))^2 \right],$$

52

using (4.4) from which it follows that

$$\sigma_R{}^2 = -\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j\gamma(\mathbf{s}_i-\mathbf{s}_j)+2\sum_{i=1}^{n}\lambda_i\gamma(\mathbf{s}_i-\mathbf{s}_0) \qquad (4.5)$$

We wish to minimise σ_R^2 subject to (4.2). To this end we apply the method of Lagrange multipliers to the function

$$\varphi(\lambda_1,\lambda_2,\ldots,\lambda_n,m) = \sigma_R^2 - 2m(\sum_{i=1}^{n}\lambda_i - 1) \qquad (4.6)$$

where m is the Lagrange parameter. Setting $\dfrac{\partial\varphi}{\partial\lambda_i}$, i=1, 2, ..., n and $\dfrac{\partial\varphi}{\partial m}$ equal to zero

we obtain the ordinary kriging system of equations

$$\begin{bmatrix} \gamma(\mathbf{s}_1-\mathbf{s}_1) & \cdot & \cdot & \gamma(\mathbf{s}_1-\mathbf{s}_n) & 1 \\ & \cdot & & \cdot & \cdot \\ & \cdot & & \cdot & \cdot \\ \gamma(\mathbf{s}_n-\mathbf{s}_1) & \cdot & \cdot & \gamma(\mathbf{s}_n-\mathbf{s}_n) & 1 \\ 1 & \cdot & \cdot & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 \\ \cdot \\ \cdot \\ \lambda_n \\ m \end{bmatrix} = \begin{bmatrix} \gamma(\mathbf{s}_1-\mathbf{s}_0) \\ \cdot \\ \cdot \\ \gamma(\mathbf{s}_n-\mathbf{s}_0) \\ 1 \end{bmatrix} \qquad (4.7)$$

The system can be written in matrix notation as

$$\Gamma\lambda = \gamma$$

from which it follows that the system has the solution

$$\lambda = \Gamma^{-1}\gamma$$

The kriging variance or minimum mean square prediction error can be written as

$$\sigma_R^2(\mathbf{s}_0) = 2\sum_{i=1}^{n}\lambda_i\gamma(\mathbf{s}_i-\mathbf{s}_0) - \sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j\gamma(\mathbf{s}_i-\mathbf{s}_j) \qquad (4.8)$$

Ordinary kriging is an exact interpolation method which means that this optimal

linear predictor, written as $\hat{Z}(\mathbf{s}_0)$, satisfies

53

$$\hat{Z}(\mathbf{s}_0) = Z(\mathbf{s}_i), \quad i = 1, 2, \ldots, n \tag{4.9}$$

We have defined kriging in terms of the semivariogram and now we will give the corresponding system in terms of the covariogram. Suppose equation (1.3) and the model (1.7) hold. In this case

$$\sigma_R^2 = C(\mathbf{0}) + \sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j C(\mathbf{s}_i - \mathbf{s}_j) - 2\sum_{i=1}^{n}\lambda_i C(\mathbf{s}_i - \mathbf{s}_0) \tag{4.10}$$

and the ordinary kriging system of equations can be written as

$$\begin{bmatrix} C(\mathbf{s}_1 - \mathbf{s}_1) & . & . & C(\mathbf{s}_1 - \mathbf{s}_n) & 1 \\ . & . & . & . \\ . & . & . & . \\ C(\mathbf{s}_n - \mathbf{s}_1) & . & . & . & C(\mathbf{s}_n - \mathbf{s}_n) & 1 \\ 1 & & . & . & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ . \\ . \\ . \\ \lambda_n \\ m \end{bmatrix} = \begin{bmatrix} C(\mathbf{s}_1 - \mathbf{s}_0) \\ . \\ . \\ . \\ C(\mathbf{s}_n - \mathbf{s}_0) \\ 1 \end{bmatrix} \tag{4.11}$$

The above system of equations can be written in matrix notation as

$$\mathbf{C}\lambda = c$$

from which the system has the solution as given below

$$\lambda = \mathbf{C}^{-1}c$$

The kriging variance or minimum mean square prediction error can be written as

$$\sigma_R^2(\mathbf{s}_0) = C(\mathbf{0}) - 2\sum_{i=1}^{n}\lambda_i C(\mathbf{s}_i - \mathbf{s}_0) + \sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j C(\mathbf{s}_i - \mathbf{s}_j) \tag{4.12}$$

When the local mean is known but constant, $E[Z(\mathbf{s})] = \mu$ say, and equal to the population mean then ordinary kriging is known as simple kriging, the covariogram must be used and the constraint (4.2) no longer applies. In this case, if second order

54

stationarity holds, we only need to model the residuals with the help of the covariance function (Cressie 1991).

$$p(Z; s_0) = \mu + \sum_{i=1}^{n} \lambda_i (Z(s_i) - \mu) \qquad (4.13)$$

The residuals can be obtained below as

$$R(s_0) = p(Z; s_0) - Z(s_0) \qquad (4.14)$$

Then it follows that

$$E[R(s_0)] = E\left[\sum_{i=1}^{n} \lambda_i . Z(s_i) - Z(s_0) \sum_{i=1}^{n} \lambda_i \right]$$

$$= \sum_{i=1}^{n} \lambda_i E[Z(s_i) - Z(s_0)]$$

$$= 0 \qquad (4.15)$$

The kriging variance is denoted by

$$\sigma_R^2 = \text{var}\ (R(s_0))$$

$$= E\left[(p(Z; s_0) - Z(s_0))^2\right],$$

using (4.15) from which it follows that

$$\sigma_R^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j C(s_i - s_j) + C(0) - 2 \sum_{i=1}^{n} \lambda_i C(s_i - s_0) \qquad (4.16)$$

We wish to minimise σ_R^2 without any constraint.

Setting $\dfrac{\partial \varphi}{\partial \lambda_i}$, i=1, 2, ..., n equal to zero this leads to the simple system of equations

55

$$\begin{bmatrix} C(\mathbf{s}_1 - \mathbf{s}_1) & . & . & C(\mathbf{s}_1 - \mathbf{s}_n) & 1 \\ & . & & . & \\ & . & . & . & \\ & . & . & . & \\ C(\mathbf{s}_n - \mathbf{s}_1) & . & . & . & C(\mathbf{s}_n - \mathbf{s}_n) & 1 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 \\ . \\ . \\ . \\ . \\ \lambda_n \end{bmatrix} = \begin{bmatrix} C(\mathbf{s}_1 - \mathbf{s}_0) \\ . \\ . \\ . \\ C(\mathbf{s}_n - \mathbf{s}_0) \end{bmatrix} \qquad (4.17)$$

This system can be written in matrix notation as

$$\mathbf{C}\lambda = c$$

from which it follows that the system has the solution

$$\lambda = \mathbf{C}^{-1}c$$

The kriging variance or minimum mean square prediction error can then be written as

$$\sigma_R^2 = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j C(\mathbf{s}_i - \mathbf{s}_j) + C(0) - 2 \sum_{i=1}^n \lambda_i C(\mathbf{s}_i - \mathbf{s}_0) \qquad (4.18)$$

The practical application of simple kriging is limited so we shall not consider it further in this thesis.

The iron ore data were considered once again simply to illustrate ordinary kriging. GEOEAS was used to produce the ordinary kriging estimates. We first used the *Prevar* and *Vario* options to estimate and model the variogram and then used the *Krige* option for the actual kriging. A spherical model with parameters $a_s = 7$, $c_s = 9.5$ and $c_0 = 4.6$ were chosen to fit the experimental semivariogram. The value at each location within the specified region can be kriged and the kriging values are

produced, in this case in terms of appropriate quartile symbols. The values obtained

by GEOEAS is given in Appendix C.1.
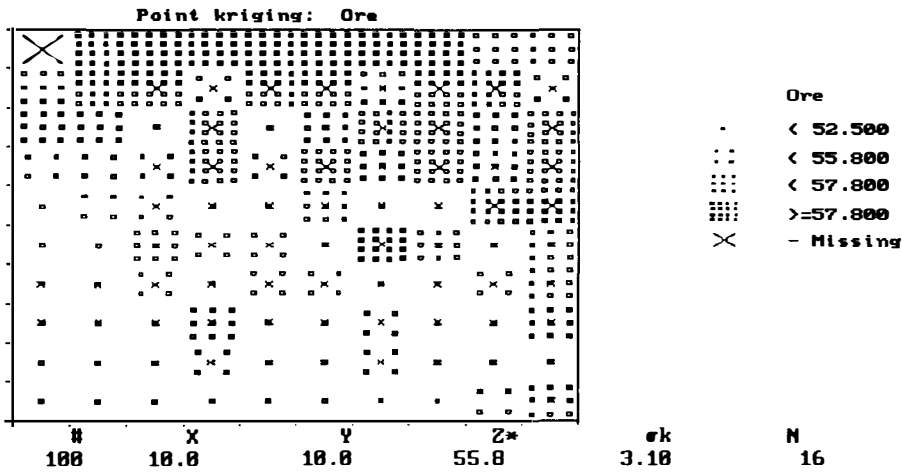
Figure 4.1

For example, the values of the kriging estimate and the kriging variance at location

(6, 1) were obtained as 50.7 and 8.96 respectively, using the surrounding

neighbourhood of iron ore data as shown in figure 4.2.
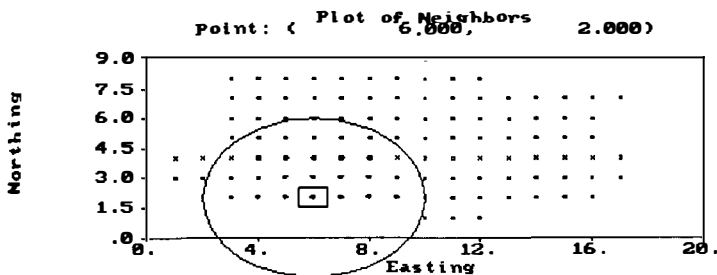
Figure 4.2

Isaaks and Srivastava (1989) gave a very simple example using seven sample data

points to illustrate the mechanics of ordinary kriging. They used the covariogram

and an exponential model in this example.

4.3. Universal Kriging

As mentioned earlier in section 1.4 in equation (1.8) universal kriging is appropriate

when data are nonstationary in the mean. In this case the order k of the polynomial

and the variogram of the residuals must be known (Cressie and Helterbrand 1994).

Suppose $Z(\mathbf{s}_0)$ is to be predicted from data $\{Z(\mathbf{s}_i)$, i=1,...,n$\}$ and f_{j-1} is a

polynomial of order j-1 (j-1=1,..., k) where f_{j-1} and k were defined in section 1.4.

Then the unbiasedness condition becomes

$$E\left(\sum_{i=1}^{n}\lambda_i Z(\mathbf{s}_i)\right) = \sum_{j=1}^{k} a_{j-1}f_{j-1}(\mathbf{s}_0) \tag{4.19}$$

and a set of sufficient conditions for this is

$$\sum_{i=1}^{n}\lambda_i f_{j-1}(\mathbf{s}_i) = f_{j-1}(\mathbf{s}_0) \qquad \text{j-1=1,...,k} \tag{4.20}$$

We minimise σ_R^2, given below in equation (4.21), subject to the constraints in (4.19)

and (4.20).

$$\sigma_R^2 = -\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j\gamma(\mathbf{s}_i - \mathbf{s}_j) + 2\sum_{i=1}^{n}\lambda_i\gamma(\mathbf{s}_0 - \mathbf{s}_j) \tag{4.21}$$

To do this we apply the method of Lagrange multipliers to the function

58

$$\varphi(\lambda_1, \lambda_2, \ldots, \lambda_n, m_1, \ldots, m_p) = \sigma_R^2 - 2\sum_{j=1}^{p+1} m_{j-1} \{\sum_{i=1}^{n} \lambda_i f_{j-1}(\mathbf{s}_i) - f_{j-1}(\mathbf{s}_0)\} \quad (4.22)$$

where $m_1, \ldots, m_p$ are the Lagrange parameters, to obtain the universal kriging system of equations

$$\begin{bmatrix} \gamma(\mathbf{s}_1 - \mathbf{s}_1) & \cdot & \cdot & \gamma(\mathbf{s}_1 - \mathbf{s}_n) & f_{i-n-1}(\mathbf{s}_j) \\ \cdot & & & \cdot & \\ \cdot & & & \cdot & \\ \cdot & & & \cdot & \\ \gamma(\mathbf{s}_n - \mathbf{s}_1) & \cdot & \cdot & \gamma(\mathbf{s}_n - \mathbf{s}_n) & f_{i-n-1}(\mathbf{s}_j) \\ f_{i-n-1}(\mathbf{s}_j) & & \cdot & \cdot & f_{i-n-1}(\mathbf{s}_j) \quad 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 \\ \cdot \\ \cdot \\ \cdot \\ \lambda_n \\ m_1 \\ \cdot \\ \cdot \\ m_p \end{bmatrix} = \begin{bmatrix} \gamma(\mathbf{s}_1 - \mathbf{s}_0) \\ \cdot \\ \cdot \\ \gamma(\mathbf{s}_n - \mathbf{s}_0) \\ f_0(\mathbf{s}_0) \\ \cdot \\ \cdot \\ f_k(\mathbf{s}_0) \end{bmatrix} \quad (4.23)$$

The above system equation can be then written in symbolic form as

$$\Gamma\lambda = \gamma$$

from which it follows that the system has the solution

$$\lambda = \Gamma^{-1}\gamma \quad (4.24)$$

The kriging variance or minimum mean square prediction error can be written as

$$\sigma_R^2 = -\sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j \gamma(\mathbf{s}_i - \mathbf{s}_j) + 2\sum_{i=1}^{n} \lambda_i \gamma(\mathbf{s}_0 - \mathbf{s}_j) - 2\sum_{j=1}^{p+1} m_{j-1}\{\sum_{i=1}^{n} \lambda_i f_{j-1}(\mathbf{s}_i) - f_{j-1}(\mathbf{s}_0)\} (4.25)$$

A sample of 85 potentiometric data was collected and measured in irregularly spaced intervals over a 250 by 200 metre region of the Palo Duro Basin, Texas, for a high level nuclear waste repository in salt (Furr and Hampen 1986). This is, in fact, the original Wolfcamp-aquifer data set considered differently by Cressie (1991). They used the semivariogram to quantify the spatial relationships of the data and universal

59

kriging was used to estimate the potentiometric surface of the Wolfcamp-aquifer. They also used the semivariogram to check for trend and anisotropy. These were done by plotting the semivariogram in the E-W, N-S, EW-NS and EN-SW directions. They found that the four directional semivariograms appear almost the same for the first 75 metres indicating confirmation of isotropy but spreading after 75 metres shown the sign of anisotropy. However they found that the spreading was caused by drift dominating the semivariograms after 75 metres. They also used cross validation to identify those data values that did not fit the pattern of the rest of the data. The degree of drift was specified and determined in order to obtain a good model for the variogram. One method of doing this involves fitting a least squares polynomial trend surface to the original data. Then the parameter of the drift will be locally estimated as part of the kriging estimation. Statistical linear regression techniques were used to fit a trend surface to the 85 potentiometric values using SAS. The linear function was sufficient to remove the trend in this case and the residuals were then analysed. If the kriging neighbourhood used is less than 75 metres then the assumption of isotropy is reasonable.

4.4. Median Polish Kriging

Median polish kriging is a useful alternative to universal kriging when the data exhibit directional trend. This trend can often be removed by applying median polish to the original data. If the residuals satisfy stationarity assumptions then ordinary kriging can be applied to the residuals and the final prediction at point s_0 can be

obtained by adding the estimated large scale variation (the fit) with the prediction obtained by small scale variation. Cressie and Glonek (1984) proved that median based removal of the trend gave a less biased estimation of the stationary error covariance than mean-based.

The deterministic component of the surface $\mu(\mathbf{s})$ is the normal form of trend surface and is defined by.

$$\mu(\mathbf{s}) = a + c(x) + r(y) \qquad \mathbf{s} = (x\ y)' \in D \tag{4.26}$$

The residuals $R(\mathbf{s})$ of the stochastic component obtained from the median polish is also a random variable and its estimate is a weighted average of the data of the form

$$\hat{R}(\mathbf{s}_0) = \sum_{i=1}^{n} \lambda_i R(\mathbf{s}_i) \qquad \mathbf{s}_0 \in R^2 \tag{4.27}$$

If these residuals satisfy the intrinsic hypothesis then ordinary kriging can be applied on them. The final result can be obtained as

$$\hat{Z}(\mathbf{s}_0) \equiv \overline{\mu}(\mathbf{s}_0) + \hat{R}(\mathbf{s}_0) \tag{4.28}$$

where $\overline{\mu}(s_0)$ the estimated trend surface and $\hat{R}(\mathbf{s}_0)$ is the estimate of the weighted average of the data under consideration.

Cressie (1988) gave an example of median polish kriging on the residuals of the iron ore data. In this analysis, Cressie found that data are nonstationary in mean and a trend exists in the N-S direction. He used median polish to remove this trend and the residuals then appeared to satisfy the intrinsic hypothesis. The variograms obtained from the residuals exhibited range anisotropy. He doubled the scale in the N-S direction and used an isotropic variogram model. The predictor of actual Fe_2O_3 was

obtained simply by adding the predicted residuals to the trend estimated by interpolating planes between the median polish fit. In this case we analysed the data by using GEOEAS and applied ordinary kriging on the residuals of the iron ore data obtained from the median polish which is given in Appendix A.2. These residuals are treated as a new data set which satisfy the intrinsic hypothesis. The variogram is estimated and modelled using GEOEAS and the values of kriging estimates can be seen below in figure 4.3. The values obtained by GEOEAS are given in Appendix C.2.
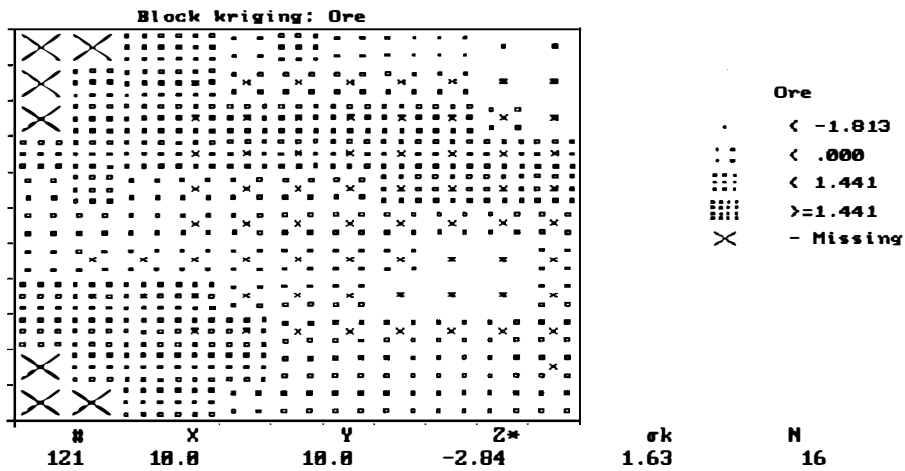
Figure 4.3

The value at location (4.556, 3.667) was kriged by considering the surrounding neighbourhood of the iron ore residuals and the estimate and kriging variance were obtained as -.670 and .647 respectively.
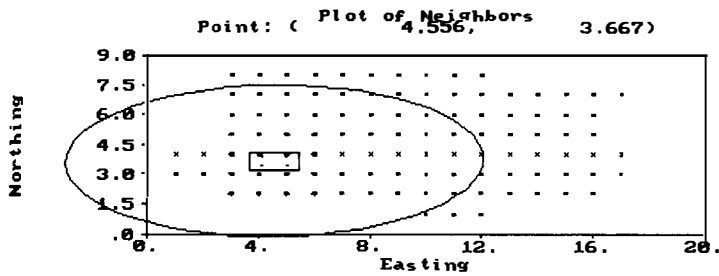
Figure 4.4

The original value can be obtained at location (4.556, 3.667) by adding the common

effect, row effect, cloumn effect and residuals is 53.67 as described in section 2.8 in

equation 2.2.

4.5. Indicator Kriging

Indicator kriging is a nonparametric geostatisitcal technique used when the

distribution of the variable under consideration is highly skewed. It has advantages

over ordinary kriging or simple kriging because it enables us to have a confidence

interval to confine our risk for the estimated values and works on the transformed

data (0,1) according to several cut off properties. Indicator kriging is data dependent

and takes account of any extreme values. Indicator kriging is more efficient in

estimating particular irregular properties such as mineralization. However its

disadvantage is that it requires the estimation and modelling of many (indicator)

variograms thus making kriging equations very complex.

Indicator kriging works on simple binary transformation whereby each datum is

transformed into an indicator before variography or kriging is performed

63

(Journel 1990 ; Solow 1993). The indicator variogram estimated from the transformed data set based on several cut off properties can then be interpolated at other locations. The final result of indicator kriging is a cumulative probability density function for every block, enabling the probability of that block exceeding the specific cut off values to be calculated (Fytas, Chaouai and Lavigne 1990). Indicator kriging has been applied in the many areas, such as the estimation of recoverable reserves (Lemmer 1984) and the estimation of spatiotemporal distributions of hydrogen-ion deposition (Bilonick 1988).

Suppose the indicator random variables are defined as

$$I(z,s) \equiv \begin{cases} 1, & if\ Z(s) \leq z \\ 0, & otherwise,\ s \in D, z \in R \end{cases} \tag{4.29}$$

where $Z(s)$ is the observed value at location s and z is a cut off value.

If second order stationarity holds for $Z(s)$ then we can write

$$F(z)=E[I(s,z)]=E[(I(s+h),z)]$$

and then the Indicator variogram $2\gamma_z(h)$ can be written as

$$2\gamma_z(h) = var(I(s+h,z) - I(s,z)), \quad h \in R^d, z \in R \tag{4.30}$$

which depends only on h.

Knowledge of $2\gamma_z(h)$ can be obtained from the bivariate distribution of $Z(s+h)$ and $Z(s)$. The (ordinary) indicator kriging predictor of $I(s_0,z)$, based on data $(I(s_1,z),\ldots,I(s_n,z))$, is

$$\hat{I}(s_0,z)= \sum_{i=1}^{n} \lambda_i(z)I(s_i,z), \tag{4.31}$$

64

where $\sum_{i=1}^{n} \lambda_i (z) = 1$ (4.32)

Writing $R(\mathbf{s}_0, z) = \hat{I}(\mathbf{s}_0, z) - I(\mathbf{s}_0, z)$, we have

$$E[R(\mathbf{s}_0), z)] = E\left[\sum_{i=1}^{n} \lambda_i (z) I(\mathbf{s}_i, z) - I(\mathbf{s}_0, z) \sum_{i=1}^{n} \lambda_i (z) \right]$$

$$= \sum_{i=1}^{n} \lambda_i E[I(\mathbf{s}_i, z) - I(\mathbf{s}_0, z)]$$

$$= 0 \qquad (4.33)$$

The kriging variance is denoted by

$$\sigma_R^2 = \text{var } (R(\mathbf{s}_0))$$

$$= E\left[(\hat{I}(\mathbf{s}_0, z) - I(\mathbf{s}_0, z))^2 \right]$$

using (4.33), from which it follows that

$$\sigma_{\mathbf{R}}^2 = -\sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i (z) \lambda_j (z) \gamma_z (\mathbf{s}_i - \mathbf{s}_j) + 2 \sum_{i=1}^{n} \lambda_i (z) \gamma_z (\mathbf{s}_i - \mathbf{s}_0) \qquad (4.34)$$

We wish to minimise σ_R^2 subject to (4.32). We will apply the method of Lagrange

multipliers to the function

$$\varphi(\lambda_1, \lambda_2, \ldots, \lambda_n, m) = \sigma_R^2 - 2m(z)(\sum_{i=1}^{n} \lambda_i (z) - 1) \qquad (4.35)$$

where m(z) is the Lagrange parameter. Setting $\dfrac{\partial \varphi}{\partial \lambda_i}$, i=1, 2, ..., n and $\dfrac{\partial \varphi}{\partial m}$ equal to

zero we obtain the ordinary kriging system of equations

$$
\begin{bmatrix}
\gamma_z(\mathbf{s}_1 - \mathbf{s}_1) & . & . & \gamma_z(\mathbf{s}_1 - \mathbf{s}_n) & 1 \\
. & & & . & \\
. & & . & . & \\
. & & . & . & \\
\gamma_z(\mathbf{s}_n - \mathbf{s}_1) & . & . & \gamma_z(\mathbf{s}_n - \mathbf{s}_n) & 1 \\
1 & & . & . & 1 & 0
\end{bmatrix}
\cdot
\begin{bmatrix}
\lambda_1(z) \\
. \\
. \\
. \\
\lambda_n(z) \\
m(z)
\end{bmatrix}
=
\begin{bmatrix}
\gamma_z(\mathbf{s}_1 - \mathbf{s}_0) \\
. \\
. \\
. \\
\gamma_z(\mathbf{s}_n - \mathbf{s}_0) \\
1
\end{bmatrix}
\qquad (4.36)
$$

The above system equation can be written in symbolically as

$$
\Gamma(z)\lambda(z) = \gamma_z
$$

from which it follows that the system has solution

$$
\lambda(z) = \Gamma^{-1}(z)\gamma_z
$$

The kriging variance or minimum mean square prediction error can then be written as

$$
\sigma_R^2(\mathbf{s}_0) = 2\sum_{i=1}^{n}\lambda_i(z)\gamma_z(\mathbf{s}_i - \mathbf{s}_0) + \sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i(z)\lambda_j(z)\gamma_z(\mathbf{s}_i - \mathbf{s}_j) \qquad (4.37)
$$

The expression $\hat{I}(\mathbf{s}_0, z)$ is actually an approximation to

$\Pr(Z(\mathbf{s}_0) \le z \mid I(\mathbf{s}_1, z), \ldots, I(\mathbf{s}_n, z))$, based on the knowledge of F(z). If F(z) is known then

we have a case of simple kriging which yields a better approximation to this

conditional probability.

Fytas, Chaouai and Lavigne (1990) studied ore reserve estimation in the new Pascalis

gold deposit located in the Abitibi region (Canada). In this case, six cut off grades

were selected for each mineralization and each composite was transferred to 0 and 1.

Indicator kriging was performed on the transformed composite values (0,1). For

example all the composites below .2g Au/tonne are transformed to 1 and the rest to 0

and then indicator kriging was carried out for each of the six cut off grades selected. Three experimental semivariograms were estimated for each group of transformed grades. These were taken in the E-W (horizontal) N-S (vertical) and EW-NS directions respectively. Spherical variogram models were fitted to these experimental semivariograms and tested by cross-validation.

4.6. Cross Validation

Cross-validation is a standard statistical technique and is used to test the goodness of fit of the variogram model that has been used. Before applying cross-validation it is important that the theoretical variogram fits the empirical variogram well near the origin. Cross-validation is applied by removing each datum from the data set, one at a time, and estimating its value based on the surrounding data, using kriging techniques with the proposed theoretical semivariogram model (Burgess and Webster 1980). Repeating cross-validation many times allows an assessment of the variability of prediction error. This can help to check the adequacy of the proposed theoretical model and diagnose those data values that do not fit the overall pattern. It also detects atypical observations when the absolute values are too large. Atypical observations can be detected by evaluating the cross-validation residuals, and are then standardized to their distribution. The standard residuals should have mean 0 and standard deviation 1. This emphasises the need for visual analysis such as a plot of standardised residuals (cross-validation). Thus, identification of potentially anomalous data is possible.

McBratney, Webster, McLaren and Spiers (1982) performed Cross-validation on data sampled from South East Scotland. The fits shown in figure 3.1 were tested by cross-validation and confirmed that the double spherical model was a good fit and the single spherical model was a poor fit.

5. Multivariate Geostatistics

5.1. Overview

Essentially for completeness we shall give, in this chapter, a brief overview of some of the main ideas of multivariate geostatistics. This area was developed, initially as Factorial Kriging Analysis by authors such as Wackernagel (1988, 1989, 1995), Royer (1984), Wartenberg (1985), Goulard (1989), Goulard and Voltz (1992) and Goovaerts (1991, 1992, 1994). We shall describe the standard multivariate technique of principal component factor analysis as well as a range of spatial multivariate techniques. In section 5.2 we outline the main details of principal component analysis. In section 5.3 we give a brief overview of the cross variogram and the cross covariogram. In section 5.4 we describe the types of correlation structure used to describe a range of interrelationships between the variables. In section 5.5 we shall consider the particular method of multivariate spatial prediction known as ordinary cokriging which exploits the appropriate joint stationary properties to obtain an unbiased minimum variance prediction estimate based on the values of two or more variables observed here at the same sample locations.

Multivariate geostatistics has been applied in several areas such as geophysics (Chiles and Guillen 1984; Galli 1984) geology (Jaquet 1989 ; Cressie and Hoef 1993) remote sensing (Ma and Royer 1988), soil science (Wackernagel 1988; Goulard 1989; Goovaerts 1991) and hydrogeology (Rouhani and Wackernagel 1990).

5.2. Principal Component Analysis

In this section we consider principal component analysis. This is a standard multivariate analysis technique used to transform a set of correlated variables $Z_1, Z_2, \ldots, Z_k$ into a set of uncorrelated factors $Y_1, Y_2, \ldots, Y_k$. The criterion used is to select the set of orthogonal factors which extracts successively a maximum proportion of the overall variance on the original set. For example, the first transformed variable $Y_1(\mathbf{s})$, or first principal component has maximum variance. The second principal component $Y_2(\mathbf{s})$ has the second largest variance and so on. The idea is to explain the variance-covariance structure of the original variables by means of a reduced set of principal components or factors. These principal components (factors) are suitable linear combinations of the original variables.

Let $Z(\mathbf{s}) = [Z_1(\mathbf{s}), \ldots Z_k(\mathbf{s})]$ where the original variables have been centred by the subtraction of the relevant means. The principal component decomposition of $V = \text{cov}(Z(\mathbf{s}))$ is given by

$$V = Q \Lambda Q' \tag{5.1}$$

$$V = \left(Q \Lambda^{1/2}\right)\left(Q \Lambda^{1/2}\right)' \tag{5.2}$$

$$= AA' \tag{5.3}$$

where $Q = [\mathbf{e}_1, \ldots, \mathbf{e}_k]$ and $\Lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \cdot & \\ & & \cdot \\ 0 & & \lambda_k \end{bmatrix}$

where $(\lambda_i, \mathbf{e}_i)$ with $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k$ are the eigenvalue-eigenvector pairs of V, and

$A = Q\Lambda^{1/2} = \left[\sqrt{\lambda_1}\mathbf{e}_1, \ldots \sqrt{\lambda_k}\mathbf{e}_k\right].$ In this case the principal components

$Y(\mathbf{s}) = [Y_1(\mathbf{s}), \ldots, Y_k(\mathbf{s})]$ are given by

$$Y(\mathbf{s}) = Z(\mathbf{s})A \qquad (5.4)$$

and hence

$$Z(\mathbf{s}) = Y(\mathbf{s})A' \qquad (5.5)$$

That is,

$$Z_j(\mathbf{s}) = \sum_{p=1}^{k} a_{pj} Y_{pj}(\mathbf{s}) \quad \text{where} \quad j = 1, 2, \ldots, k. \qquad (5.6)$$

One disadvantage in using classical principal component analysis for spatial data is that it does not take into account any spatial correlation which may exist between observations.

5.3. Cross Covariograms and Cross Variograms

The cross variogram and the cross covariogram are two ways of measuring the spatial relationship between two variables.

Let $Z(\mathbf{s}) = (Z_1(\mathbf{s}), \ldots, Z_k(\mathbf{s}))$ $\mathbf{s} \in D$ be a multivariate stochastic process. If $Z(\mathbf{s})$ satisfies the conditions

$$E[Z_i(\mathbf{s})] = \mu_i \qquad (5.7)$$

where μ_i is constant, and

$$E\left[(Z_i(\mathbf{s}) - \mu_i)(Z_j(\mathbf{s}+\mathbf{h}) - \mu_j\right] = C_{ij}(\mathbf{h}) \qquad (5.8)$$

where $C_{ij}(\mathbf{h})$ is called the cross covariogram, then the process $\{Z(\mathbf{s})\}$ is said to be

joint second-order stationary. Note that, in general,

$C_{ij}(\mathbf{h}) \neq C_{ji}(\mathbf{h})$ and $C_{ij}(-\mathbf{h}) \neq C_{ij}(\mathbf{h})$ but $C_{ij}(\mathbf{h}) = C_{ji}(-\mathbf{h})$.

$$E\big[Z_i(\mathbf{s}+\mathbf{h}) - Z_i(\mathbf{s})\big] = 0 \qquad (5.9)$$

$$\mathrm{cov}\big[\big(Z_i(\mathbf{s}+\mathbf{h}) - Z_i(\mathbf{s})\big),\big(Z_j(\mathbf{s}+\mathbf{h}) - Z_j(\mathbf{s})\big)\big] = 2\gamma_{ij}(\mathbf{h}) \qquad (5.10)$$

where $2\gamma_{ij}(\mathbf{h})$ is called the cross covariogram, then the process $\{Z(\mathbf{s})\}$ is said to be

joint intrinsic stationary. In this case the cross semivariogram $\gamma_{ij}(\mathbf{h})$ can be written

as

$$\gamma_{ij}(\mathbf{h}) = \frac{1}{2}E\big[Z_i(\mathbf{s}+\mathbf{h}) - Z_i(\mathbf{s})\big]\big[Z_j(\mathbf{s}+\mathbf{h}) - Z_j(\mathbf{s})\big], \ \mathbf{h} \in R^d \qquad (5.11)$$

and $\gamma_{ij}(\mathbf{h})$ is an even function with $\gamma_{ii}(\mathbf{h})\gamma_{jj}(\mathbf{h}) \geq |\gamma(\mathbf{h})|^2$.

When the process $\{Z(\mathbf{s})\}$ satisfies joint second order stationary then

$$\gamma_{ij}(h) = C_{ij}(0) - \frac{1}{2}\big(C_{ij}(-\mathbf{h}) + C_{ij}(\mathbf{h})\big) \qquad (5.12)$$

An example of the cross semivariogram has been given by Webster, Atteia and

Dubois (1994). Their aim was to determine the cross dependence of the

concentration of seven potentially toxic metals contained in their soil data. The

experimental cross semivariograms were computed, interpreted and finally modelled.

5.4. Analysis of the Correlation Structure

In the analysis of the correlation structure there are three matrices which represent a relationship between the variables. The variance-covariance matrix V represents the interrelationship between variables at the same location. The variogram matrix $\Gamma(\mathbf{h}) = \left(\gamma_{ij}(\mathbf{h})\right)$ describes how the relationship between the spatial increments alters with $\mathbf{h}$. The coregionalization matrix B_u represents the relationship between variables at the spatial scale defined by the basic variogram model and will be considered in detail in section 5.6.

If the correlation between two variables does not depend on the spatial scale then these variables exhibit *intrinsic correlation*. In this case we can write

$$C(\mathbf{h}) = V\rho(\mathbf{h}) \tag{5.13}$$

and so

$$\frac{\sigma_{ij}\rho(\mathbf{h})}{\sqrt{\sigma_{ii}\rho(\mathbf{h})\sigma_{jj}\rho(\mathbf{h})}} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}} = r_{ij} \tag{5.14}$$

This happens when $C_{ij}(\mathbf{h}) = b_{ij}\rho(\mathbf{h})$ and then we can interpret b_{ij} as the σ_{ij}.

For variograms the corresponding model is

$$\Gamma(\mathbf{h}) = B\gamma(\mathbf{h}) \tag{5.15}$$

for some (positive semi-definite) matrix B called the *coregionalization matrix*. In this case we can write

$$\gamma_{jj'}(\mathbf{h}) = b_{jj'}\gamma(\mathbf{h}) \quad jj' = 1,2,\ldots,k \tag{5.16}$$

for some basic variogram model $\gamma(\mathbf{h})$.

We then carry out a principal component analysis on B, as described in section 5.2.

Hence $B = AA'$ with $A = \left(\sqrt{\lambda_1}\mathbf{e}_1, \ldots, \sqrt{\lambda_k}\mathbf{e}_k\right)$, where $\left(\lambda_p, \mathbf{e}_p\right)$ are the eigenvalue-eigenvector pairs of B and $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k$. Thus

$$A = (a_{pj}) = (\sqrt{\lambda_{pj}}\,\mathbf{e}_{pj}), \ p\,j = 1,2,\ldots,k \tag{5.17}$$

$$\text{and } b_{jj} = \sum_{p=1}^{k} a_{pj}a_{pj'}, \ j\,j' = 1,2,\ldots,k \tag{5.18}$$

The resulting factors are given by $Y(\mathbf{s}) = Z(\mathbf{s})A$ and associated linear model of intrinsic correlation is defined by

$$Z_j(\mathbf{s}) = \sum_{p=1}^{k} a_{pj}Y_p(\mathbf{s}) \tag{5.19}$$

where the $Y_p(\mathbf{s})$ have pairwise uncorrelated increments, so that

$$E\left[Y_p(\mathbf{s}+\mathbf{h}) - Y_p(\mathbf{s})\right]\left[Y_q(\mathbf{s}+\mathbf{h}) - Z_q(\mathbf{s})\right] = 0, \ p \neq q \tag{5.20}$$

and the same variogram, given by

$$E\left[\left(Y_p(\mathbf{s}+\mathbf{h}) - Y_p(\mathbf{s})\right)^2\right] = 2\gamma(\mathbf{h}), \ p = 1,2,\ldots,k \tag{5.21}$$

In this case kriging can be performed separately on the relevant factors of the original set of variables. Two way of checking for intrinsic correlation are described below.

Firstly, if the cross variograms between principal components are not zero at all lags then the principal components are not uncorrelated at all spatial scales and the intrinsic correlation model is not appropriate.

Secondly, consider the codispersion coefficients (Matheron 1965)

$$cc(\mathbf{h}) = \frac{\gamma_{ij}(\mathbf{h})}{\sqrt{\gamma_{ii}(\mathbf{h})\gamma_{jj}(\mathbf{h})}} \qquad (5.22)$$

If these are constant then the correlation of the pairs does not depends on the spatial scale and we can write

$$cc(\mathbf{h}) = \frac{b_{ij}\gamma(\mathbf{h})}{\sqrt{b_{ii}b_{jj}}\gamma(\mathbf{h})} = \frac{b_{ij}}{\sqrt{b_{ii}b_{jj}}} = r_{ij} \qquad (5.23)$$

5.5. Ordinary Cokriging

In ordinary cokriging the secondary variables are normally cross correlated with the primary variables. For example, in mineral industry, core samples may contain not only the grade of copper but also the grade of lead and zinc. We will consider only the case of isotopy here when all the variables are measured at the same sample locations. In this case cokriging is better than separate kriging as it preserves the coherence of the estimators.

This means that if $\hat{W}(\mathbf{s}) = \sum_{i=1}^{k} \hat{Z}_i(\mathbf{s})$ then one always has $\hat{W}(\mathbf{s_0}) = \sum_{i=1}^{k} \hat{Z}_i(\mathbf{s_0})$ under cokriging but, in general, $\hat{W}(\mathbf{s_0}) \neq \sum_{i=1}^{k} \hat{Z}_i(\mathbf{s_0})$ under kriging.

When the joint intrinsic hypothesis holds then ordinary cokriging with unknown means is a simply a minimum variance unbiased prediction of $Z_l(\mathbf{s_0})$ on the basis of observed values $Z_j(\mathbf{s}_i)$ where $i = 1,...,n$, $j = 1,...,k$. We write

$$\hat{Z}_l(\mathbf{s_0}) = p_l(Z; \mathbf{s_0}) = \sum_{i=1}^{n}\sum_{j=1}^{k} \lambda_{ji} Z_j(\mathbf{s_i}) \qquad (5.24)$$

where λ_{ji} is the weight assigned to the jth variable at the ith location.

For unbiased prediction, we need

$$\sum_{i=1}^{n} \lambda_{ij} = \delta_{ij} = \begin{cases} 1, & j = l \\ 0, & j \neq l \end{cases}$$

(5.25)

where δ_{jl} is the Kronecker delta. With this condition

$$E[R(s_0)] = E[p_l(Z;s_0) - Z_l(s_0)]$$

$$= E\left[\sum_{i=1}^{n}\sum_{j=1}^{k} \lambda_{ji} Z_j(s_i) - \sum_{j=1}^{k} \lambda_{li} Z_l(s_0) - \sum_{\substack{j=1 \\ j \neq l}}^{k}\sum_{i=1}^{n} \lambda_{ji} Z_j(s_0)\right]$$

$$= \sum_{j=1}^{k}\sum_{i=1}^{n} \lambda_{ji} E[Z_j(s_i) - Z_j(s_0)] = 0$$

(5.26)

Hence it follows that

$$\mathrm{var}\left(R_l(s_0)\right) = E\left(Z_l(s_0) - \sum_{i=1}^{n}\sum_{j=1}^{k} \lambda_{ji} Z_{j(s_i)}\right)^2$$

(5.27)

and this is to be minimised subject to the constraint in (5.25).

In similar fashion to the details given in section 4.2 for ordinary kriging, $\mathrm{var}\left(R_l(s_0)\right)$ can be written in terms of covariograms (under the assumption of joint second order stationarity) or variograms (under the assumption of joint intrinsic stationarity together with the assumption that all the cross variograms are symmetric). The method of Lagrange multipliers can then be used to carry out the minimization. In terms of the covariogram the resulting ordinary cokriging system of equations is given by

$$\begin{cases} \sum_{i=1}^{n}\sum_{j=1}^{k}\lambda_{ji}C_{jj'}\left(s_i-s_{i'}\right)-m_{j'}=C_{lj'}\left(s_0-s_{i'}\right), i'=1,2,\ldots,n, \ j'=1,2,\ldots,k \\ \sum_{i=1}^{n}\lambda_{ji}=\delta_{jl}, \end{cases}$$
(5.28)

where the $m_{j'}$, $j'=1,2,\ldots,k$ are the Lagrange parameters. The resulting minimum error variance is then given by

$$\sigma_k^2(s_0)=C_{ll}(s_0,s_0)-\sum_{i=1}^{n}\sum_{l=1}^{k}\lambda_{ji}C_{ji}(s_0,s_i)+m_l$$
(5.29)

5.6 Linear Coregionalization Analysis

Let $\{Z_j(s); \ j=1,\ldots k\}$ satisfy the joint second-order stationarity assumptions. Then we decompose $\{Z_i(s)\}$ into sets $\{Z_{ju}(s), u=0,1,\ldots,l\}$ of spatially uncorrelated components so that

$$Z_j(s)=\sum_{u=1}^{l}Z_{ju}(s)+\mu_j, \quad j=1,\ldots,k$$
(5.30)

where for all j, j', u, v

$$E[Z_j(s)]=\mu_j \quad \text{and} \quad E[Z_{ju}(s)]=0$$
(5.31)

and

$$\begin{cases} \text{cov}\big[(Z_{iu}(s),Z_{i'u}(s+h)),(Z_{ju}(s),Z_{j'u}(s+h))\big]=2C_{ij'}(h) \\ \text{cov}\big[(Z_{iu}(s),Z_{i'v}(s+h))\big]=0, \quad u\neq v \end{cases}$$
(5.32)

The cross covariograms $C_{jj'u}(h)$ is composed of real coefficients $b_{jj'u}$ and are proportional to real correlation coefficients $\rho_u(h)$ so that

$$C_{jj'u}(h)=\sum_{u=0}^{l}C_{jj'u}(h)=\sum_{u=0}^{l}b_{jj'u}\rho_u(h)$$
(5.33)

Coregionalzation matrices B_u of order k×k can be set up and the one has a nested covariance function model given by

$$C(\mathbf{h}) = \sum_{u=0}^{l} B_u \rho_u(\mathbf{h}) \tag{5.34}$$

where the B_u are positive semi-definite. Note that this model is equal to the intrinsic correlation model when one can write $B_u = a_u B$. In this case

$$C(\mathbf{h}) = \sum_{u=0}^{l} a_u B \rho_u(\mathbf{h}) = B \sum_{u=0}^{l} a_u \rho_u(\mathbf{h}) \tag{5.35}$$

In the general case we can write

$$C_{jj'u}(\mathbf{h}) = \sum_{u=0}^{l} b_{jj'u} \rho_u(\mathbf{h}), \; jj' = 1,2,...,k \tag{5.36}$$

for some set of correlation functions $\rho_u(\mathbf{h})$.

We then carry out a principal component analysis, as described in section 5.2, on B_u, $u = 1,2,...,l$.

Hence $B = A_u A_u'$ with $A = \left(\sqrt{\lambda_{1u}} \mathbf{e}_1,...,\sqrt{\lambda_{ku}} \mathbf{e}_k \right)$, where $\left(\lambda_{pu}, \mathbf{e}_p \right)$ are the eigenvalue-eigenvector pairs of B_u and $\lambda_{1u} \geq \lambda_{2u} \geq ... \geq \lambda_{ku}$. Thus

$$A_u = (a_{jpu}) = (\sqrt{\lambda_{jpu}} \mathbf{e}_{pj}), \; p, j = 1,2,...,k \tag{5.37}$$

$$b_{jju} = \sum_{p=1}^{k} a_{jpu} a_{j'pu} \; j, \; j' = 1,2,...,k \tag{5.38}$$

The resulting factors are given by $Y_u(\mathbf{s}) = Z(\mathbf{s}) A_u$. Hence we can write

$$Z_{ju}(\mathbf{s}) = \sum_{u=0}^{l} \sum_{p=1}^{k} a_{jpu} Y_{pu}(\mathbf{s}) \tag{5.39}$$

78

and the associated linear model of coregionalization is defined by

$$Z_{ju}(\mathbf{s}) = \sum_{u=0}^{l}\sum_{p=1}^{k} a_{jpu} Y_{pu}(\mathbf{s}) \tag{5.40}$$

where for all j, u, v, p, q

$$\begin{cases} E\big[Y_{ps}(\mathbf{s})\big] = 0 \\ \mathrm{cov}\big(Y_{pu}(\mathbf{s}), Y_{pu}(\mathbf{s}+\mathbf{h})\big) = \rho_u(\mathbf{h}) \\ \mathrm{cov}\big(Y_{pu}(\mathbf{s}), Y_{qv}(\mathbf{s}+\mathbf{h})\big) = 0, u \neq v, or\ p \neq q \end{cases} \tag{5.41}$$

These regionalized factors $Y_{ps}(\mathbf{s})$ can be estimated at location $\mathbf{s}_0$ by ordinary

cokriging with the $Z_j(\mathbf{s})$ as described in section 5.5.

6. Applications

6.1. Overview

In this chapter we shall analyse two original data sets. In doing so we shall use many of the techniques detailed in the preceding chapters. The first data set consists of soil data collected by Bloom (1996) from the Jimperding Brook area of south-west Western Australia. Soil samples were taken at points on two 11×11 grids with grid spacing of 6m. Both grids were located on cleared land but one field had been cropped only once fifteen years ago while the other field had been cropped regularly over the past fifteen years. Several variables were measured in each case but we shall concentrate here on an analysis of two of these variables, Available Phosphate (in ppm) and Potassium (in ppm). These data are given in Appendices A.4, A.5, A.6 and A.7 respectively. The Available Phosphate data will be considered in section 6.2 and the Potassium data will be considered in section 6.3.

The second data set is the distribution of Banksia species in a region of Southwestern Australia. The samples were taken on a 16×22 grid with grid spacing of 50 metres. These data are given in Appendix A.10. These data were sampled and collected by Lamont and Connell (1994) for a study on the biogeography of Banksia distribution in Southwestern Australia. The analysis is given in section 6.4.

MINITAB and GEOEAS were used throughout this chapter to produce the required plots.

6.2. Available Phosphate analysis

A stem-and leaf plot of the cropped field Available Phosphate data is given in figure
6.1. This indicates that these data contain one global outlier. The last row of stem 5
and leaf 1 indicates 51 ppm Phosphate (truncated data).

Stem and leaf plot for cropped field Available Phosphate

4	0	1344
42	0	56666666677777777788888888999999999999
(44)	1	00000001111111222222222222333333333444444444
35	1	5555556666667788888
16	2	00122222
8	2	55679
3	3	01
1	3	
1	4	
1	4	
1	5	1

Figure 6.1

A histogram and box plot of the cropped field Available Phosphate data, along with a
range of descriptive statistics, are given in figure 6.2. From these plots it can be seen
that the data are positively skewed.
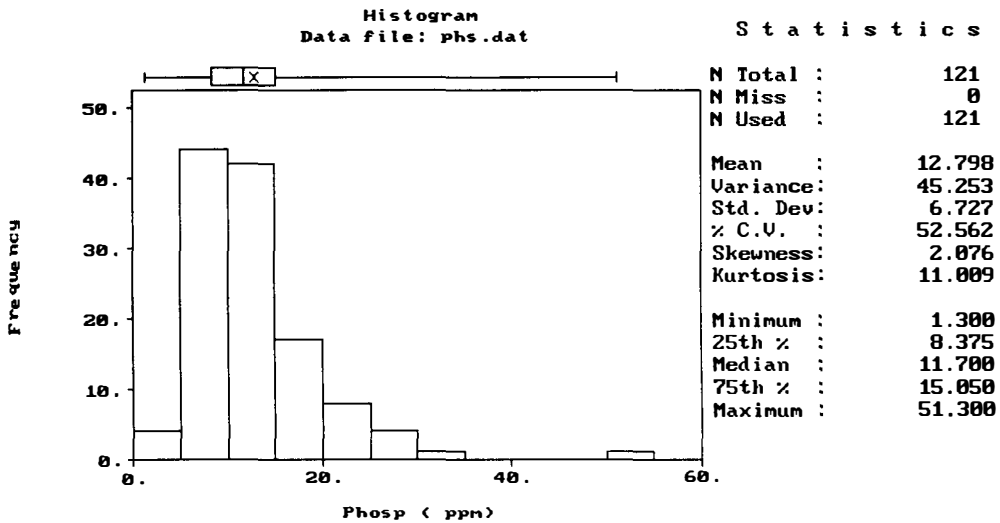
Figure 6.2

81

Post plots and row and column mean and median plots were prepared in order to check for directional trend for the cropped field Available Phosphate data. These post plots are shown in figures 6.3 and 6.4. From these post plots no trend is evident.

Post plot of Phosphate data without values
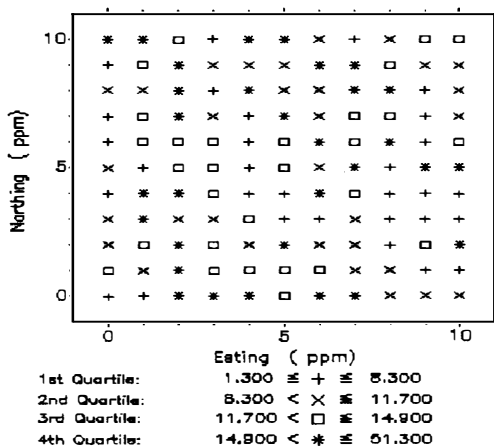
Figure 6.3

Post plot of Phosphate data with values
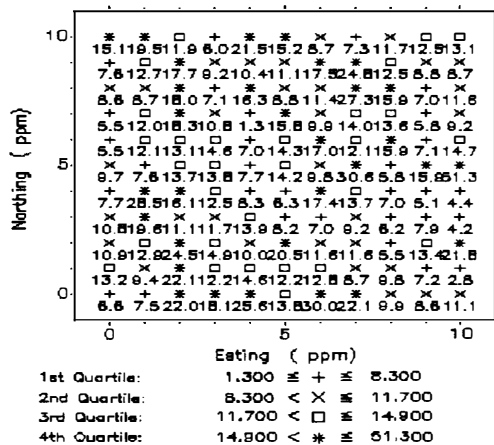
Figure 6.4

The row mean and median plots are given in figure 6.5. These plots indicate that there is no trend in the N-S direction.

Row mean and median plots
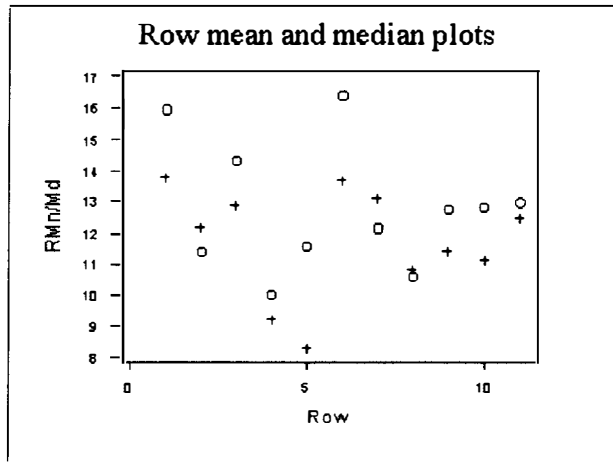
Figure 6.5

The column mean and median plots are shown in figure 6.6. From these plots it can be seen that there is no trend in the E-W direction.

Column mean and median plots
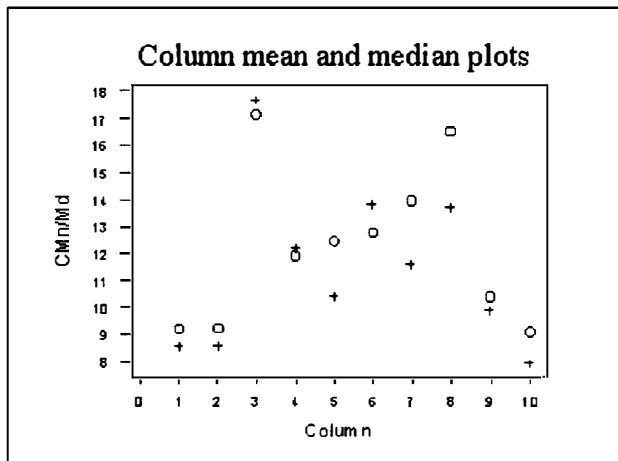
Figure 6.6

The natural log transformation of the original data was used to obtain a more symmetric distribution. A histogram and box plot of the transformed cropped field Available phosphate data are given in figure 6.7.

83
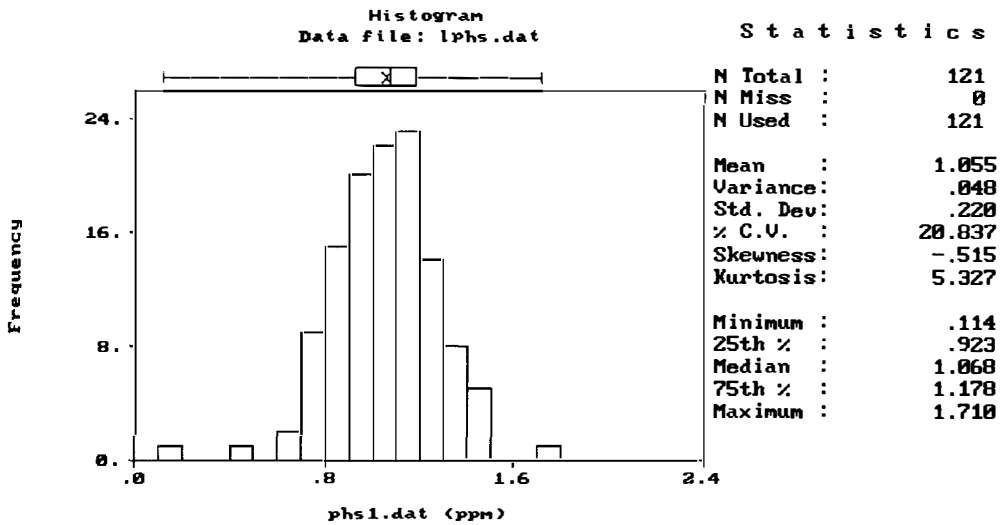
Figure 6.7

From a comparision of figure 6.7 and figure 6.2 it can be seen that the transformed

data has a more symmetric distribution than the original data.

Four directional semivariogram plots were produced to visualise the spatial continuity

of the transformed cropped field Available phosphate data. The four directional

semivariograms were estimated in the E-W, EN-SW, N-S and NE-SW directions and

are given in figures 6.8, 6.9, 6.10 and 6.11 respectively.

Figure 6.8

84
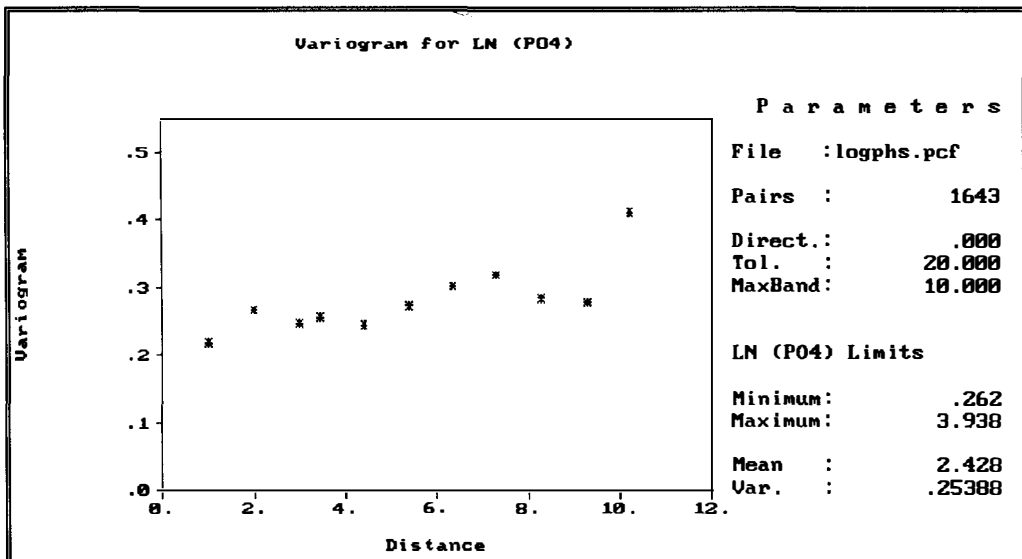
Figure 6.9

Figure 6.10
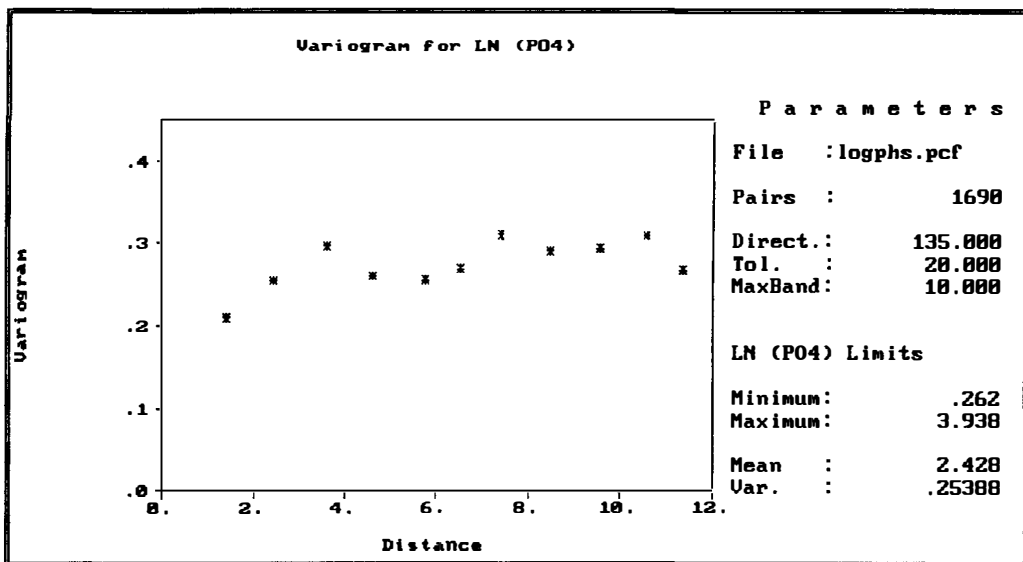
Figure 6.11

The pattern of spatial continuity of the experimental semivariograms in all four directions is essentially the same so an isotropic model can be used and this can be fitted to the omnidirectional semivariogram which is given in figure 6.12. The values of all five experimental semivariograms considered were obtained using GEOEAS and are given in Appendices B.17 to B.21.

Figure 6.12

87
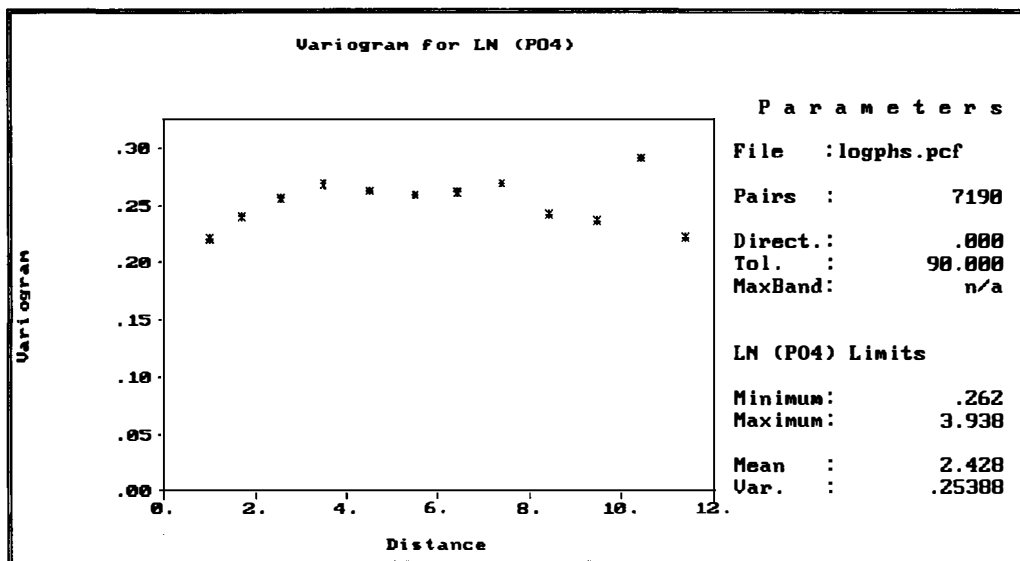
A spherical model with partial sill $c_s = 0.19$, nugget effect $c_0 = 0.06$ and range $a_s = 4$ was fitted to the omnidirectional semivariogram using GEOEAS and this fitted model is shown figure 6.13. Note that the fit is very good near the origin.
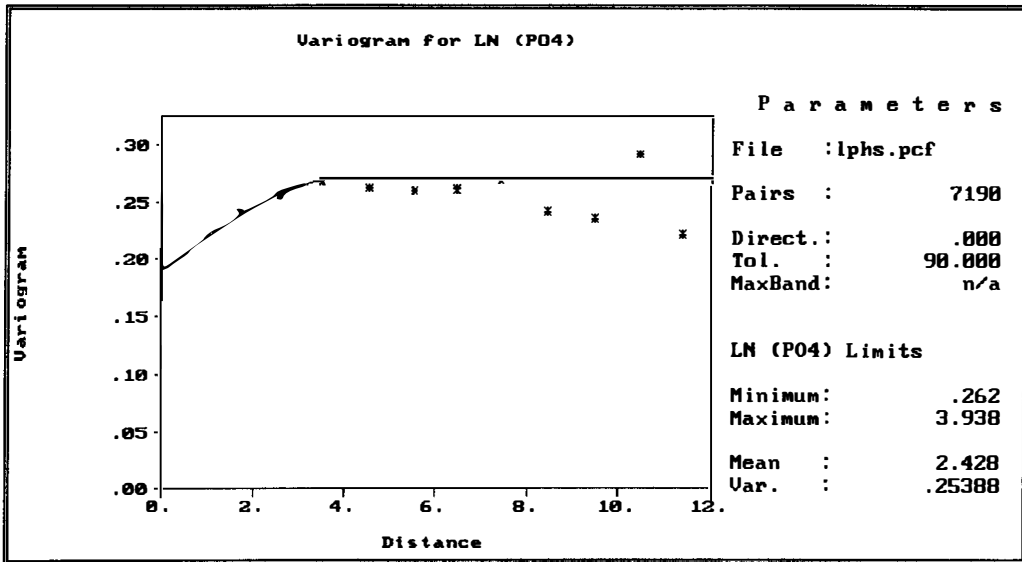
Figure 6.13

The values of the parameters of the spherical model were then estimated by weighted least squares, using the above values as the initial choices. The parameter estimates $c_s = 0.06$, $c_0 = 0.20$ and $a_s = 4$ were obtained and the corresponding fitted semivariogram model is displayed in figure 6.14.

Figure 6.14

88
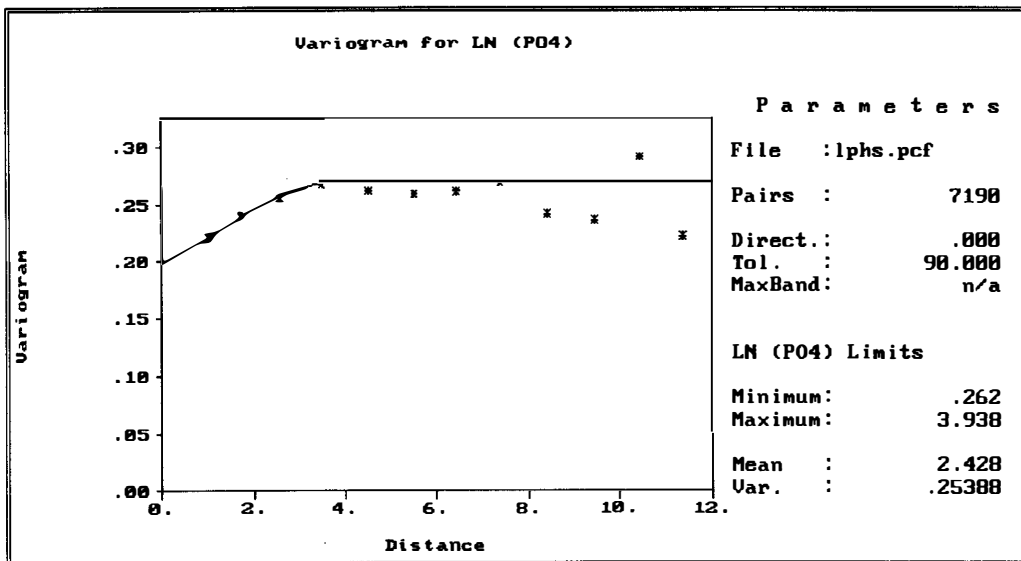
The parameter estimates obtained by weighted least squares were used in kriging. Ordinary kriging was appropriate in this case and was performed on the transformed data. The results of this are given in figure 6.15. The values obtained by GEOEAS are given in Appendix C.3.
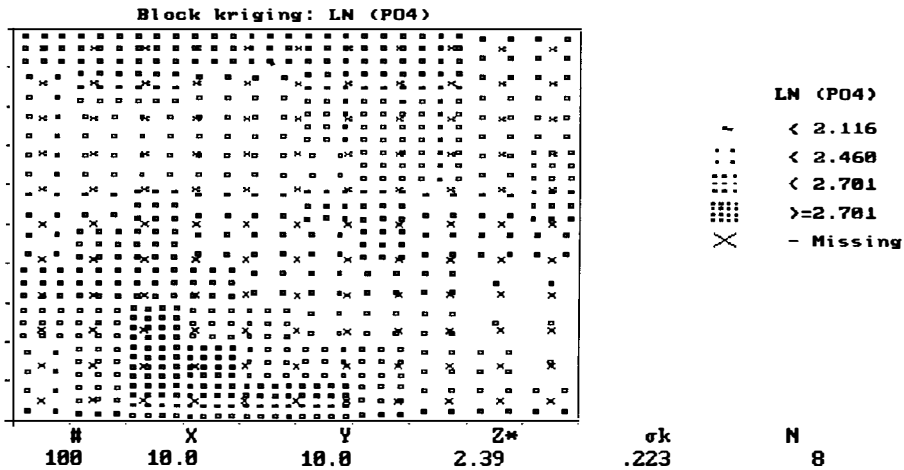
Figure 6.15

For example, the values of the kriging estimate and the kriging variance at location (6.666, 3.333) were obtained as 2.30 and 0.0279 respectively, using the surrounding neighbourhood of cropped field Available phosphate data as shown in figure 6.16.

Figure 6.16

89
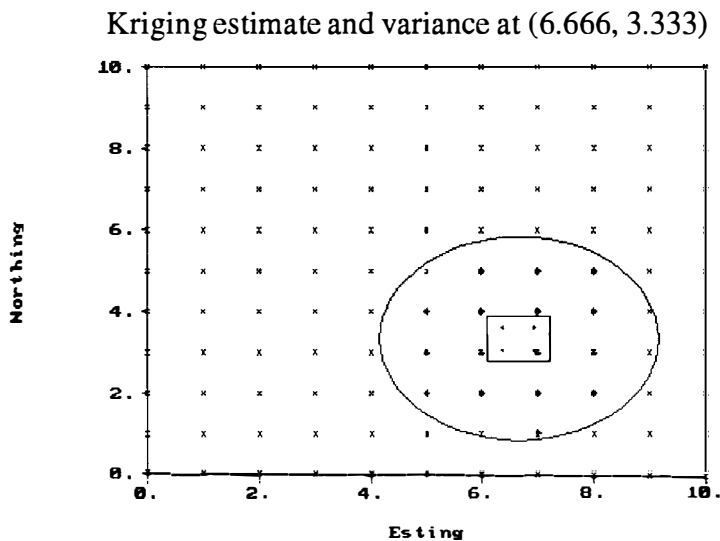
The kriging estimate for the original data can now be obtained, making use of an appropriate correction for the bias involved in using the exponential as a back - transformation here (Cressie 1991).

For the uncropped field Available Phosphate data a stem and leaf plot is given in figure 6.17. We see from figure 6.17 that, in this case, there are two global outliers. For example, the bottom row of stem 3 and leaf 5 indicates 35 ppm Phosphate (truncated data).

Stem and leaf plot for uncropped field Available Phosphate data

```
  9      0  233333333
 42      0  444444444444444445555555555555555
(23)     0  66666666667777777777777
 52      0  8888888889999999999999
 32      1  0000001111
 22      1  222333333
 13      1  4555
  9      1  6677
  5      1  89
  3      2  0
  2      2
  2      2
  2      2
  2      2  9
  1      3
  1      3
  1      3  5
```

Figure 6.17

A histogram and box plot of the uncropped field Available phosphate data were also obtained. These are given in figure 6.18.
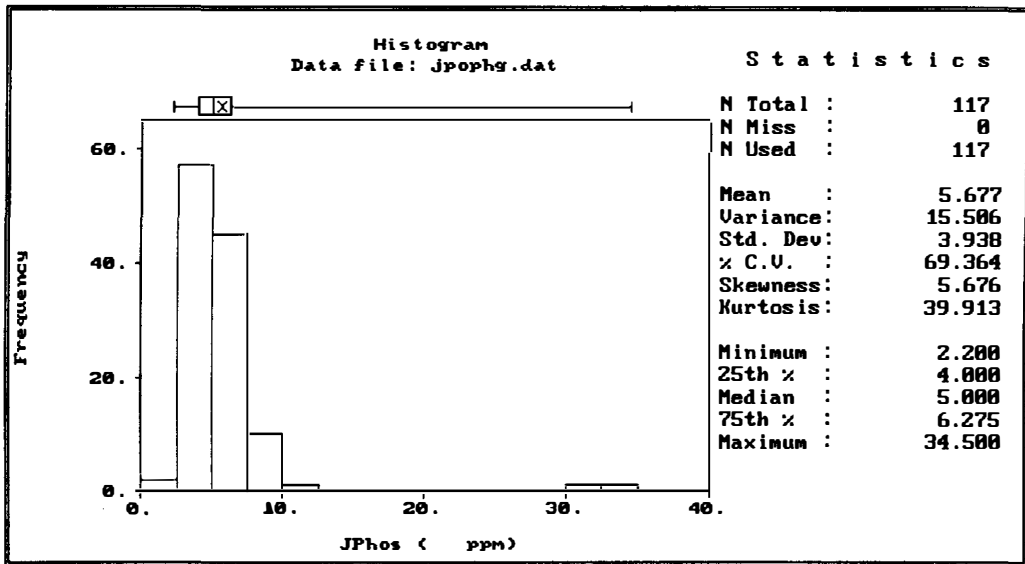
Figure 6.18

From figure 6.18, we can see that the data are positively skewed with two outliers and the natural log transformation will be used in this case. Post plots, row and column mean and median plots were also obtained to check for directional trend. From the postplots in figures 6.19 and 6.20 there is no trend.
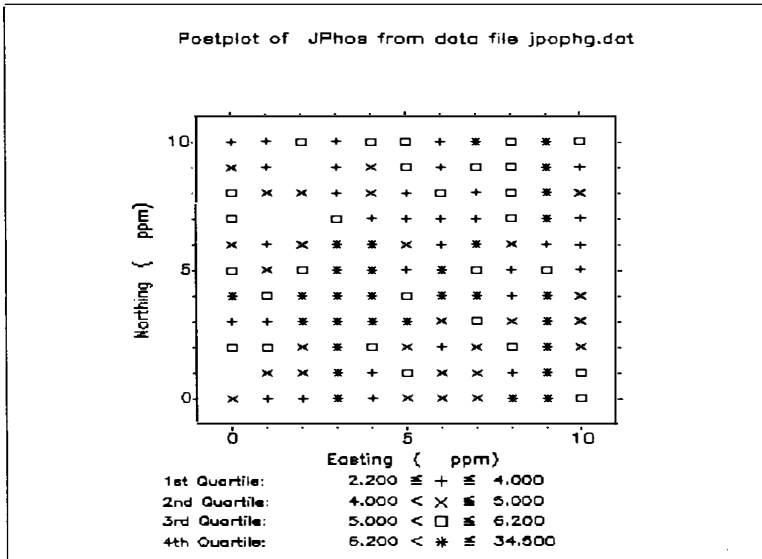
Figure 6.19
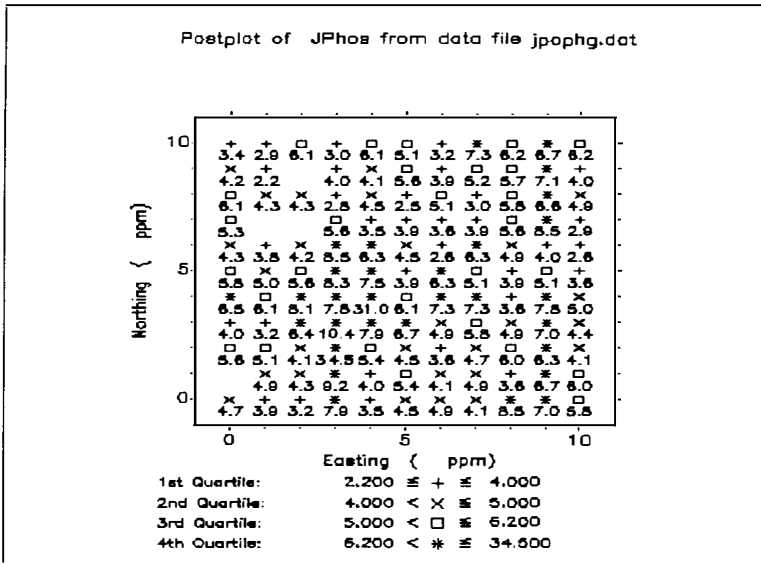
Figure 6.20

The row mean and median plots and the column mean and median plots are given in figures 6.21 and 6.22. These plots confirm the lack of trend in the N-S and E-W directions respectively.
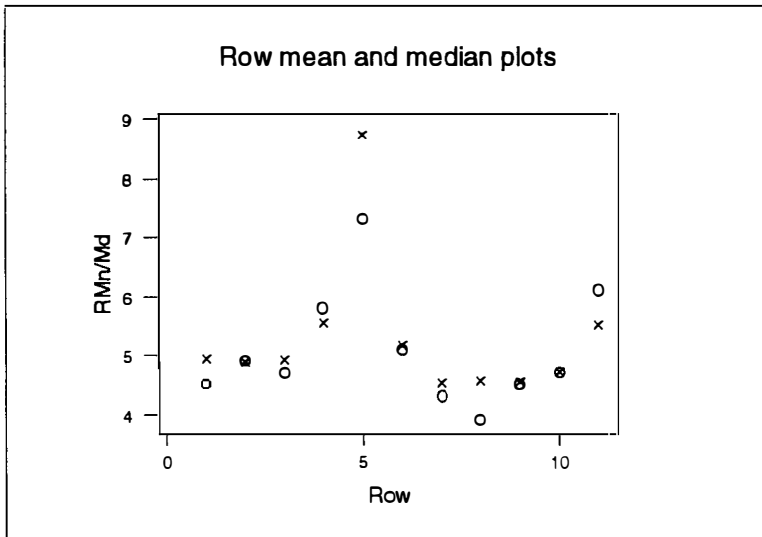
Figure 6.21
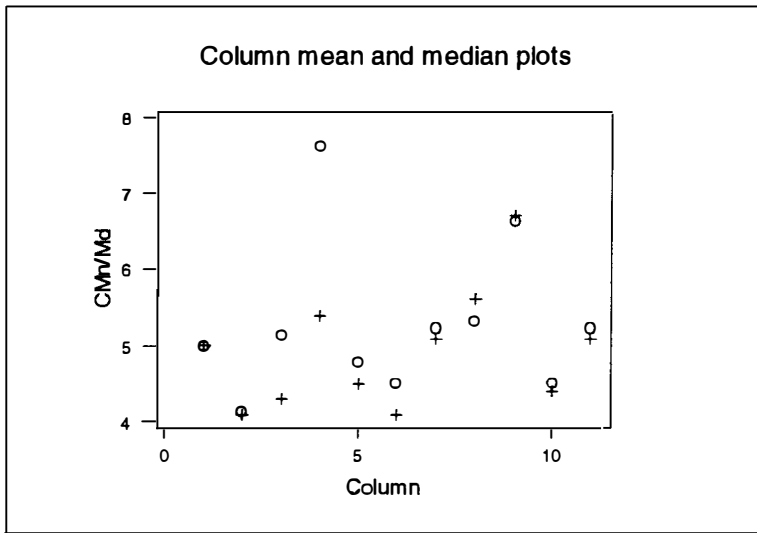
Figure 6.22

Once again four directional semivariograms were obtained to visualise the spatial dependence of the transformed data. The estimated directional semivariograms are in the E-W, N-S, EW-NS and EN-SW directions and are given in figures 6.23. 6.24, 6.25 and 6.26 respectively. The values of all five experimental variograms are also given in Appendices B.31 to B.35 respectively.
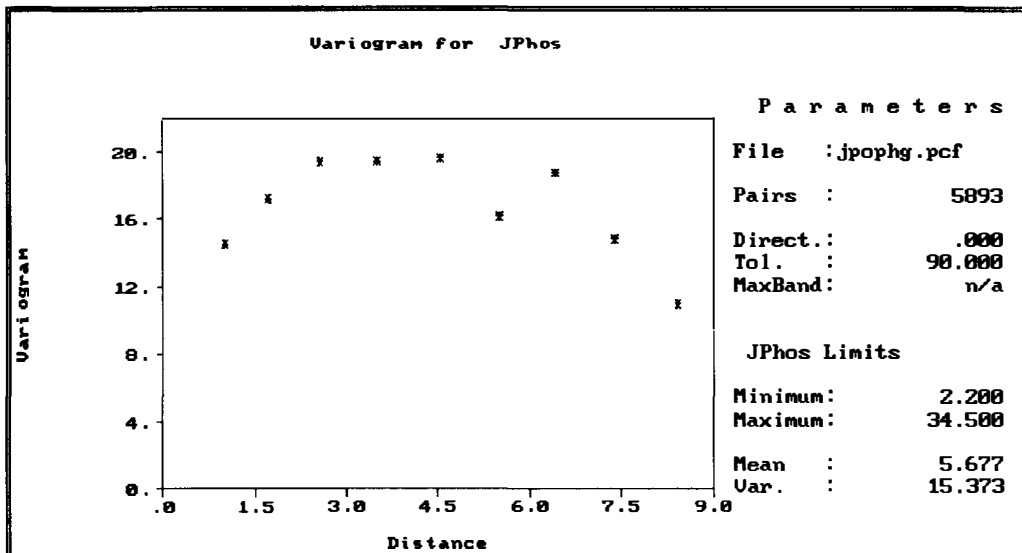
Figure 6.23
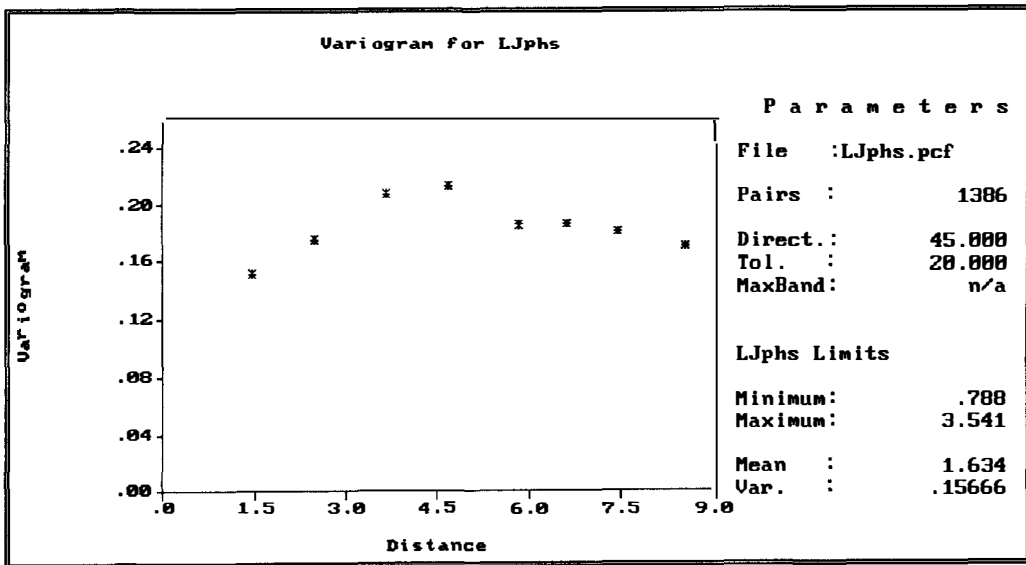
Figure 6.24

Figure 6.25

94
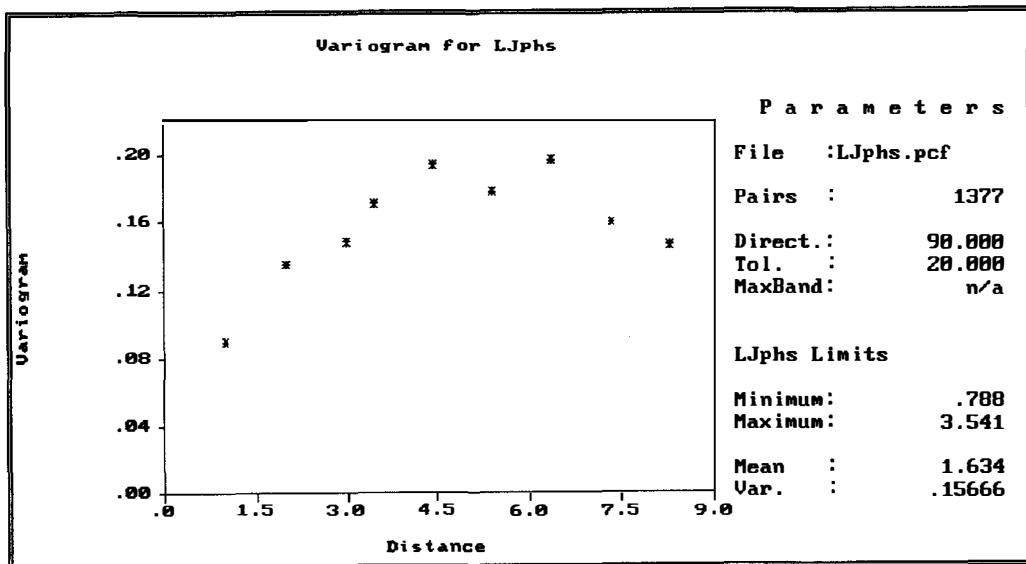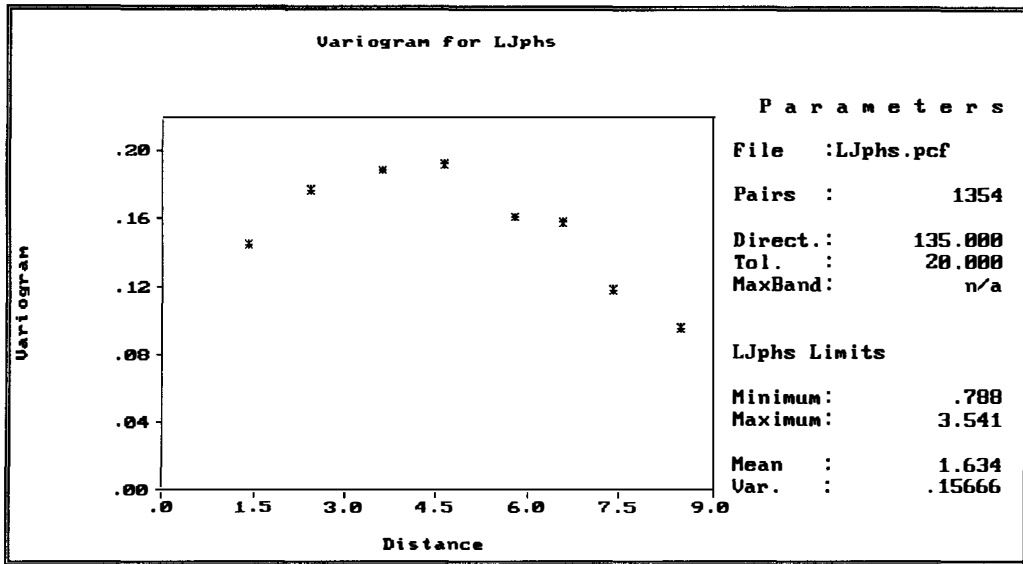
Figure 6.26

The pattern of spatial continuity of the experimental semivariograms in all the four directions is essentially same so an isotropic model can be fitted. This was done for the omnidirectional semivariogram which is shown in figure 6.27.

Figure 6.27

95
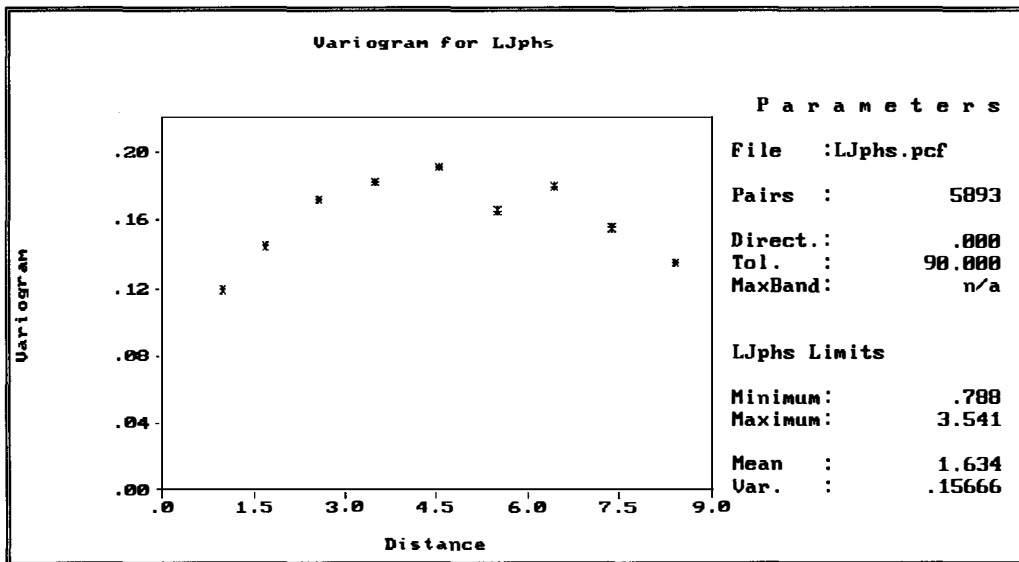
A spherical model was fitted using GEOEAS with partial sill $c_s = 0.11$, nugget effect $c_0 = 0.08$ and range $a_s = 4$ and this fit is shown in figure 6.28.
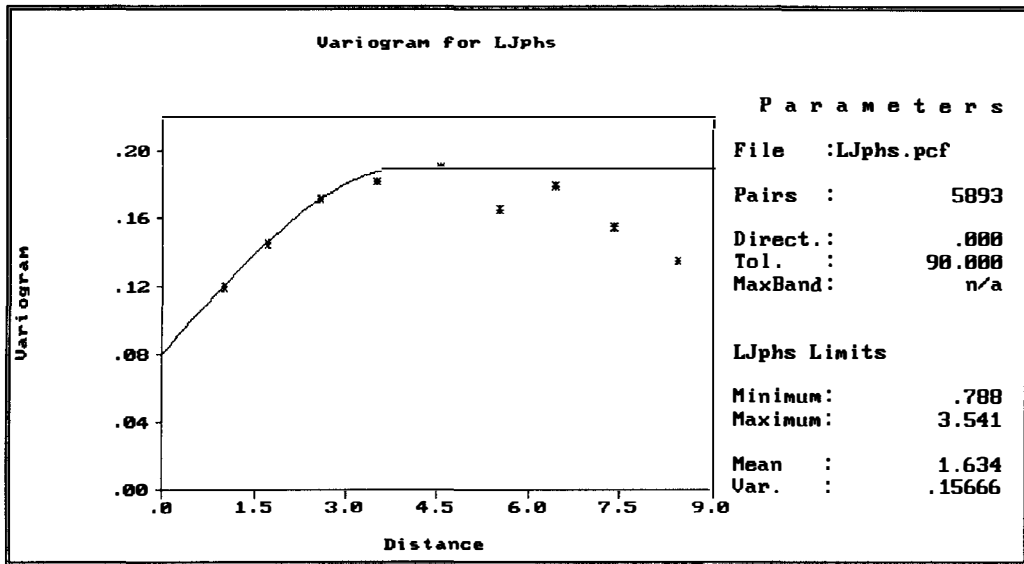
Figure 6.28

The values of parameters for this model were then estimated by using weighted least squares, obtaining the estimates $c_s = 0.09$. $c_0 = 0.08$ and $a_s = 4$. This fitted model is shown in figure 6.29 below.
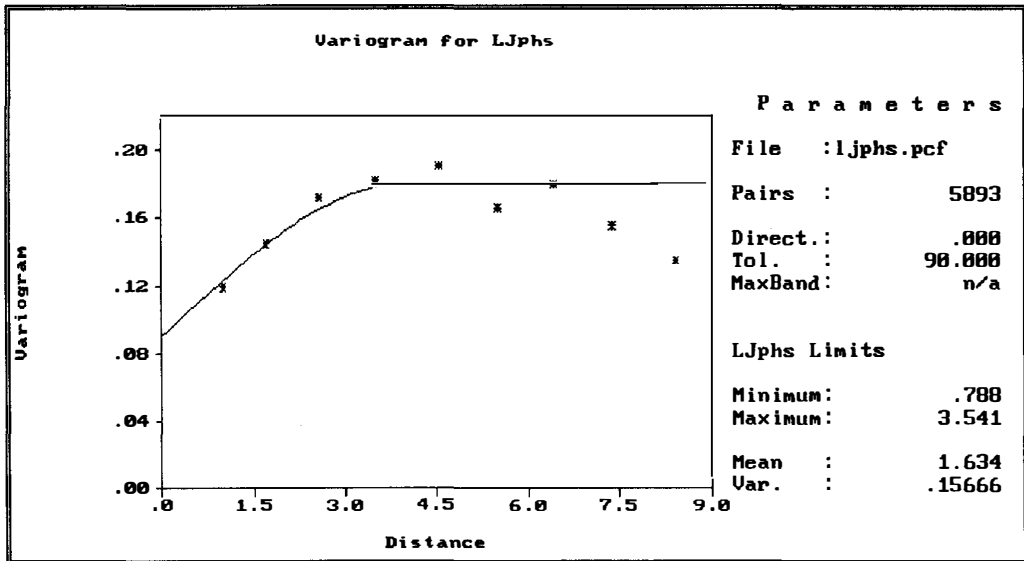
Figure 6.29

The parameter estimates from weighted least squares were then used in the subsequent kriging. Ordinary kriging was appropriate in this case and was performed on the transformed uncropped field Available data and is shown in figure 6.30.
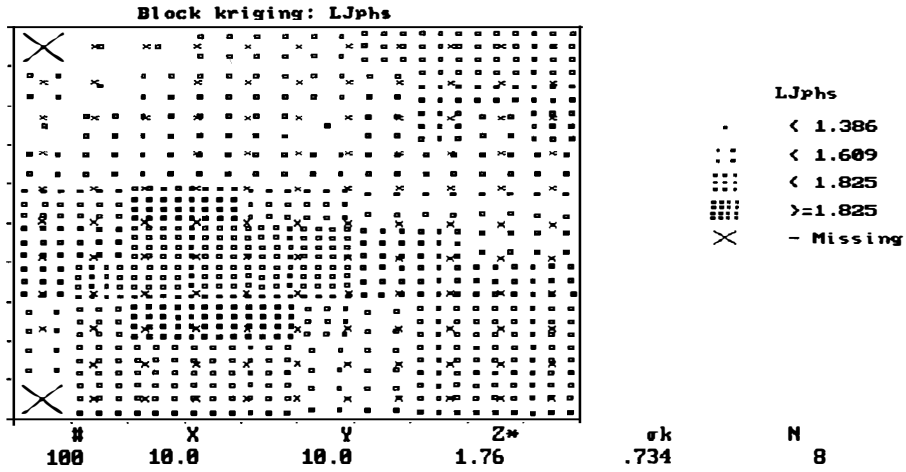
Figure 6.30

The values of the kriging estimate and the kriging variance at location (6.666, 3.333) were obtained here as 1.72 and 0.02 respectively, using the surrounding neighbourhood of uncropped field Available phosphate data shown in figure 6.31 below.

Kriging estimate and kriging variance

Figure 6.31

97
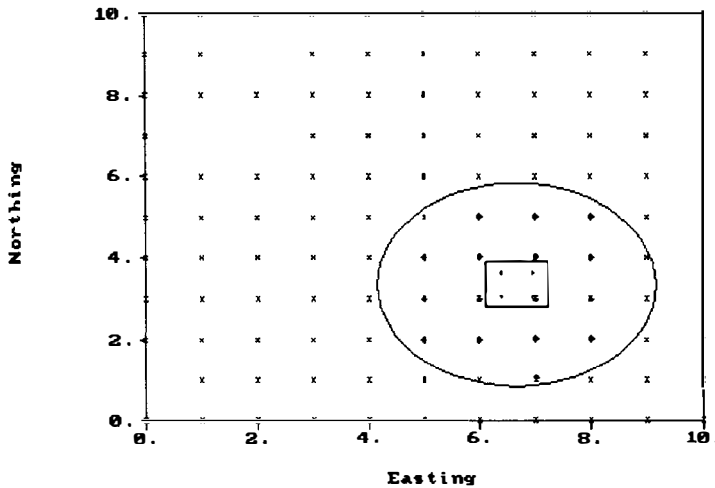
The second and third columns give the stem and leaves respectively. For example in

the last row of 11 stem and 3 leaf indicates 113 ppm cropped field Potassium. The

first column provides a count of the number of leaves.

<p style="text-align:center">Stem and leaf plot for the cropped field Potassium data</p>

```
   1    1    7
  31    2    2344445555566666677788888899999
 (41)   3    00000001111112222223444455566777788889999
  49    4    00001111223334444455789
  26    5    000111245678
  14    6    0112234789
   4    7    4
   3    8
   3    9    29
   1   10
   1   11    3
```

<p style="text-align:center">Figure 6.32</p>
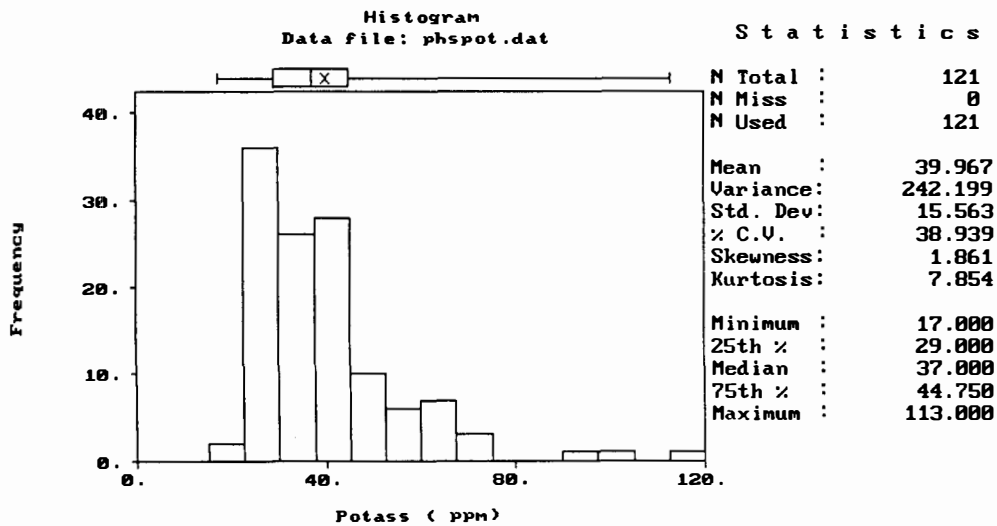
<p style="text-align:center">Figure 6.33</p>

Post plots of these data are given in figures 6.34 and 6.35. From these plots we can

see a trend in the E-W direction, with values increasing from east to west.

The estimates of the original data can then be obtained.

The spatial variation for both cropped and uncropped field Available phosphate data is very clear from figures 6.13 and 6.28. From these figures it is apparant that spatial variation for the uncropped field Available phosphate data is much more rapid than for the cropped field Available phosphate data. For both data the ranges are the same and equal to 24 metres. It means that data pairs for both cropped and uncropped field Available data are correlated at distance up to 24 metres and uncorrelated at distance greater than 24 metres. The nugget effect for the cropped field Available Phosphate data is larger than the uncropped field Available Phosphate data while the sill for the cropped field Available Phosphate data is lower that for the uncropped field Available Phosphate data. This latter result was to be expected given that the actual variance of the cropped field Available Phosphate data is lower than the actual variance for the uncropped field Available Phosphate data.

6.3. Potassium analysis

The stem-and leaf plot of the cropped field Potassium data was obtained and is given in figure 6.32. Two global outliers can be seen. A histogram and box plot, together with summary statistics, for the cropped field Potassium data are given in figure 6.33. From all three plots it can be seen that the data are positively skewed.

Postplot of Potass from data file phspot.dat

Northing (ppm)

10

5

0

0   5   10

Easting (ppm)

1st Quartile:	17.000 ≤ + ≤ 29.000
2nd Quartile:	29.000 < X ≤ 36.000
3rd Quartile:	36.000 < □ ≤ 44.000
4th Quartile:	44.000 < ✳ ≤ 113.000
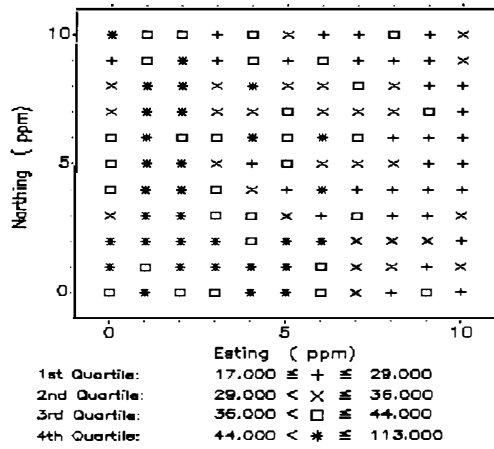
Figure 6.34

Postplot of Potass from data file phspot.dat

Northing (ppm)

10

5

0

0   5   10

Easting (ppm)

1st Quartile:	17.000 ≤ + ≤ 29.000
2nd Quartile:	29.000 < X ≤ 36.000
3rd Quartile:	36.000 < □ ≤ 44.000
4th Quartile:	44.000 < ✳ ≤ 113.000
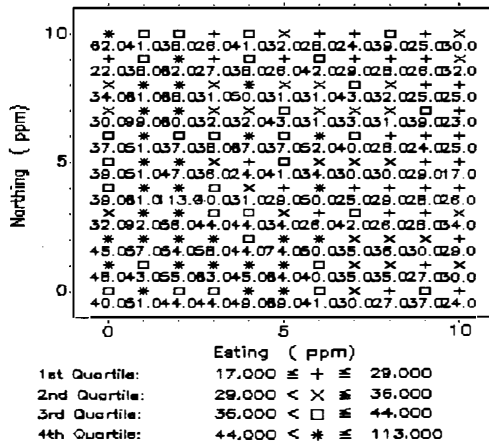
Figure 6.35

The row and column means and medians were also plotted and these plots are given in

figures 6.36 and 6.37 respectively.   Figure 6.36 confirms the lack of trend in the N-S

direction.

100
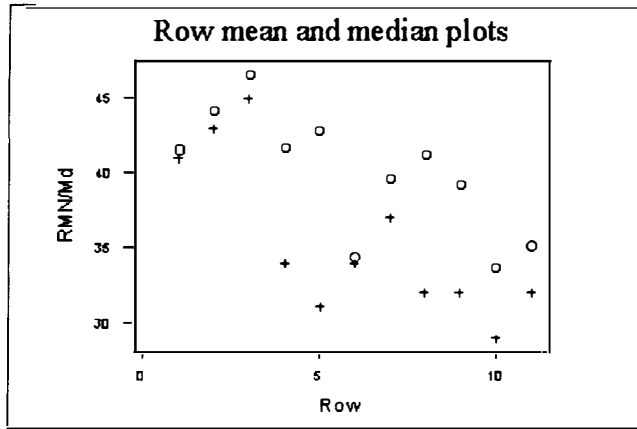
Figure 6.36

From figure 3.37 we can see evidence of the previously mentioned trend in the E-W direction.
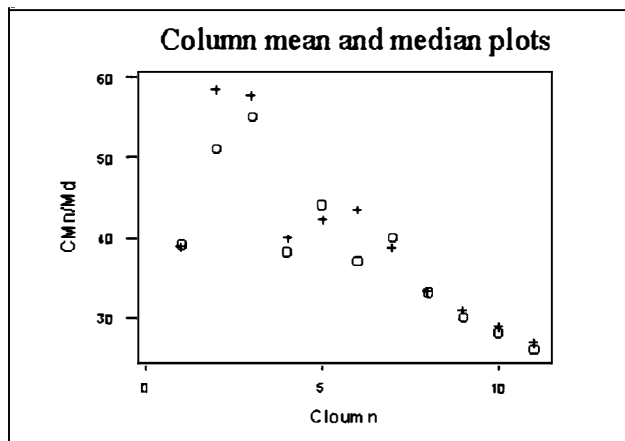
Figure 6.37

As one way of dealing with this trend in the E-W direction, median polish, using six iterations, was performed on the original data and the residuals obtained were treated as a new data set. A stem and leaf plot of the residuals is given in figure 6.38. It can

be seen from this figure that there are two global outliers. The last row of stem 6 and

leaf 0 indicates a residual of 6.0 ppm for the cropped field (truncated) Phosphate data.

Stem and leaf plot for residuals of the cropped Phosphate data

```
   1    -1        6
   7    -1        422100
  24    -0        99987776666655555
  59    -0        444333333322222211111110000000000000
 (27)    0        000000000011111122222222233
  35     0        5556677778888899
  19     1        0011233
  12     1        556
   9     2        3
   8     2        558
   5     3        12
   3     3        9
   2     4
   2     4        7
   1     5
   1     5
   1     6        0
```

Figure 6.38

A histogram and box plot of the residuals, together with descriptive statistics, are

given in figure 6.39. From a comparision of 6.33 and 6.39 it can be seen that the

transformed data has a more symmetric distribution than the original data.
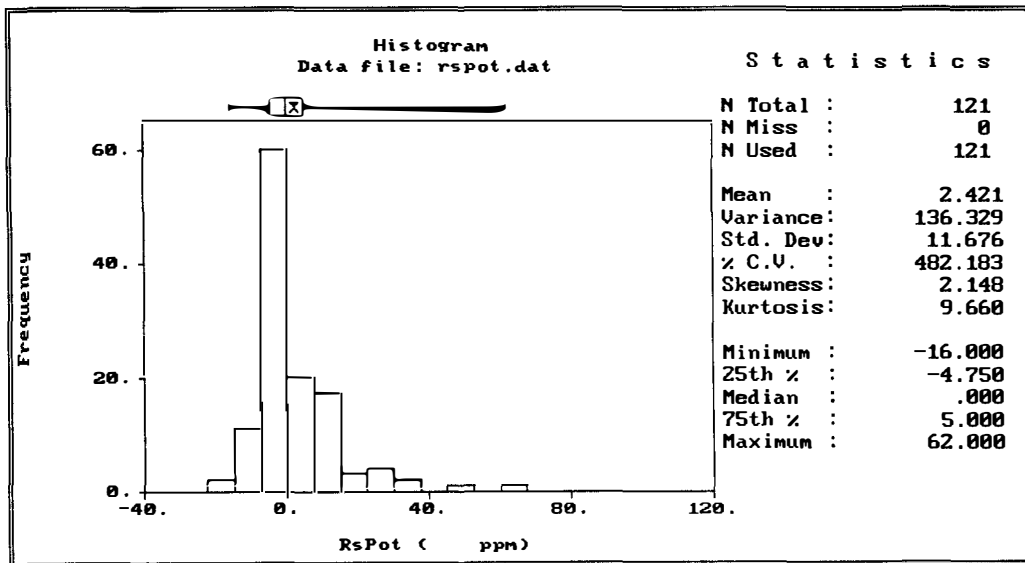
Figure 6.39

Post plots of these residuals are given in figures 6.34 and 6.35. From these plots we can see no trend in any direction.
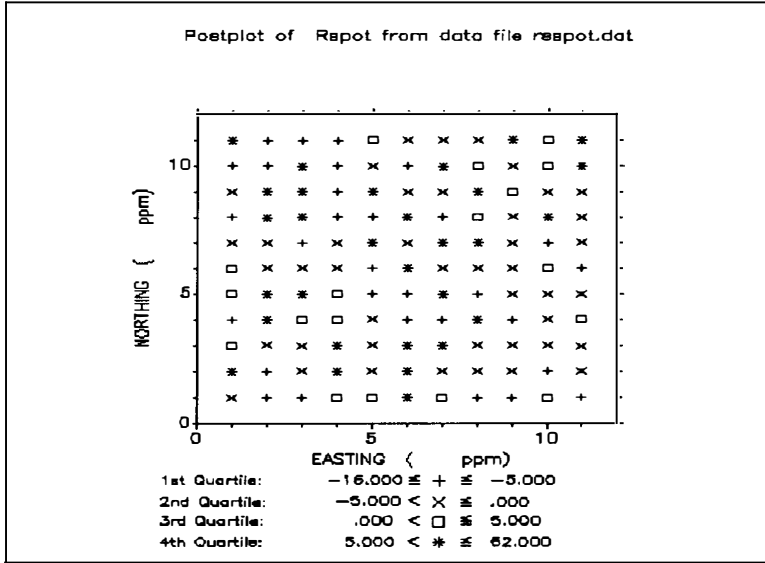
Figure 6.40
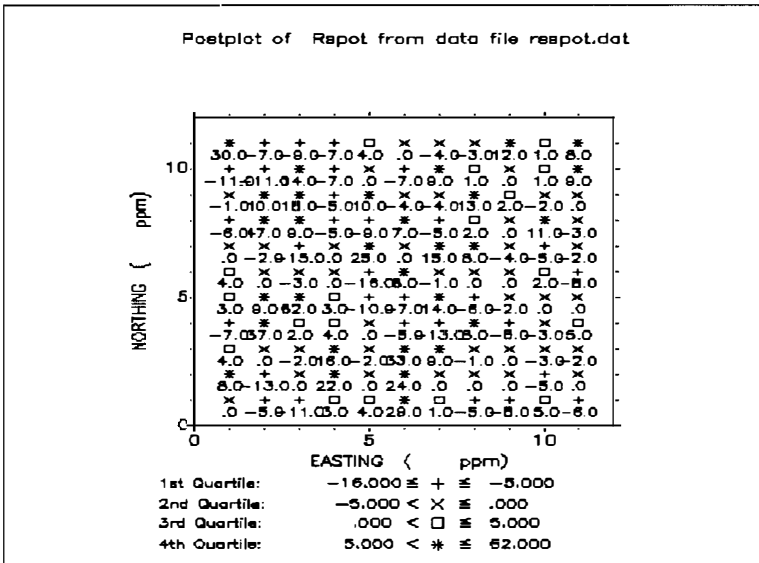
Figure 6.41

The row and column means and medians were also plotted and are given in figures 6.42 and 6.43 respectively. Figure 6.42 confirms the lack of trend in the N-S direction, while figure 6.43 confirms the lack of trend in the E-W direction.

103
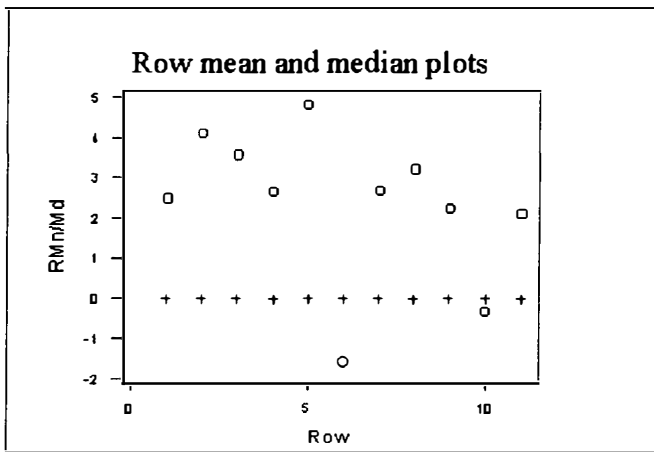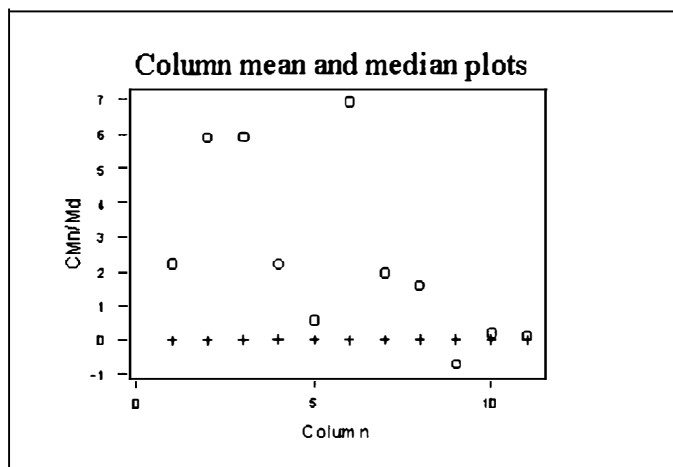
Figure 6.42

Figure 6.43

Four directional semivariogram plots were produced to visualise the spatial continuity

of the cropped field Available Potassium data. These four directional semivariograms

are in the E-W, N-S, EW-NS and NE-SW directions and are given in figures 6.44,

6.45, 6.46 and 6.47 respectively.
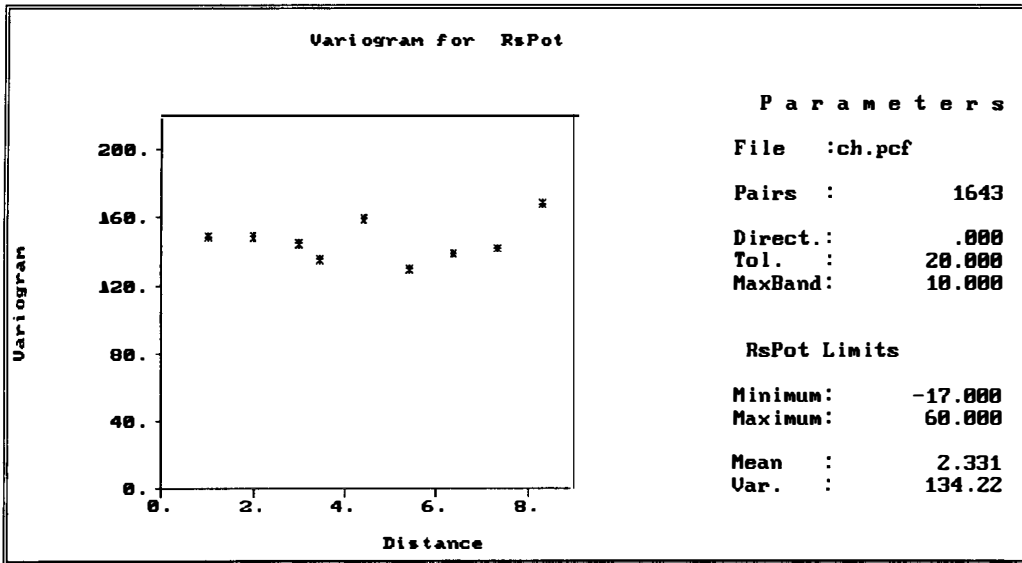
Figure 6.44
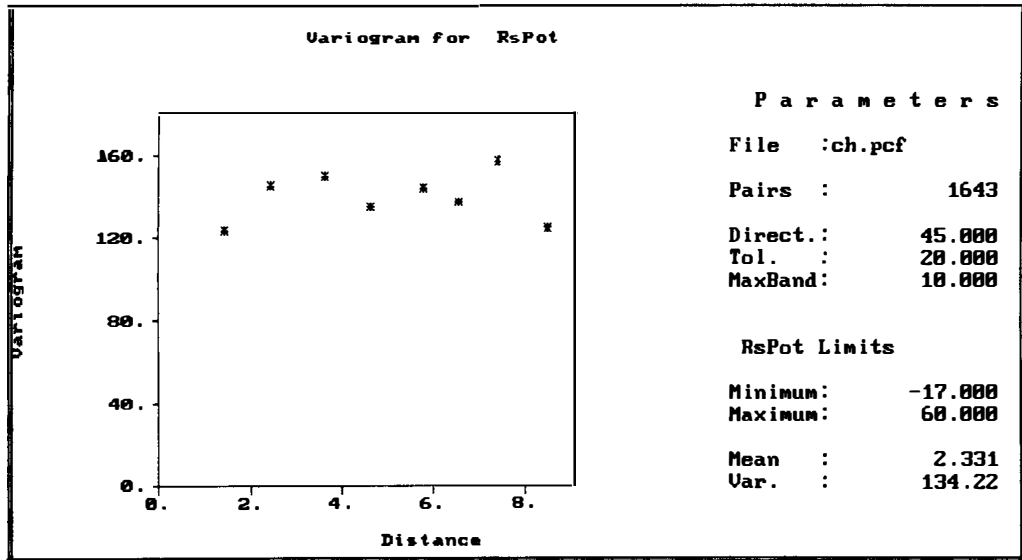
Figure 6.45

105
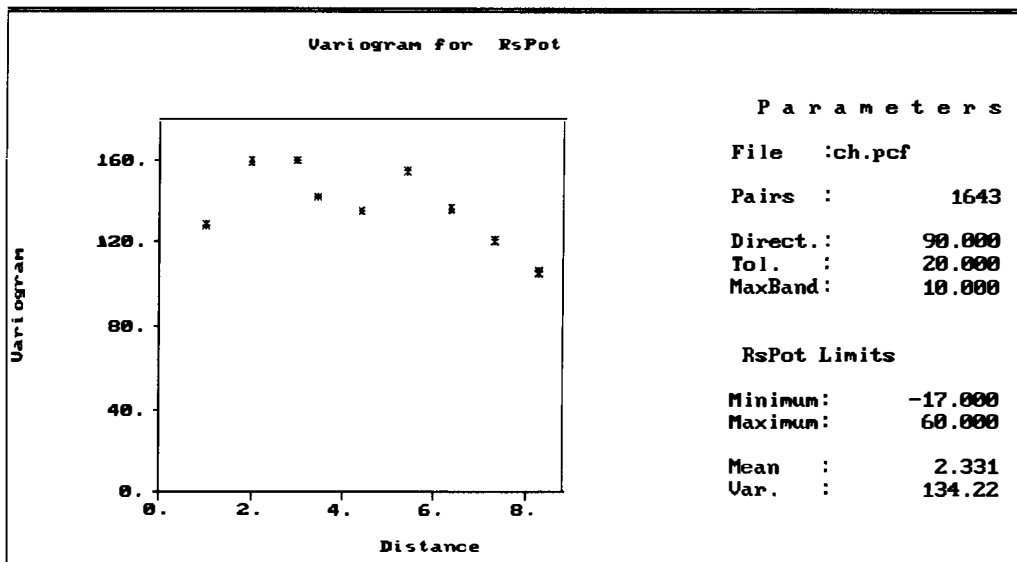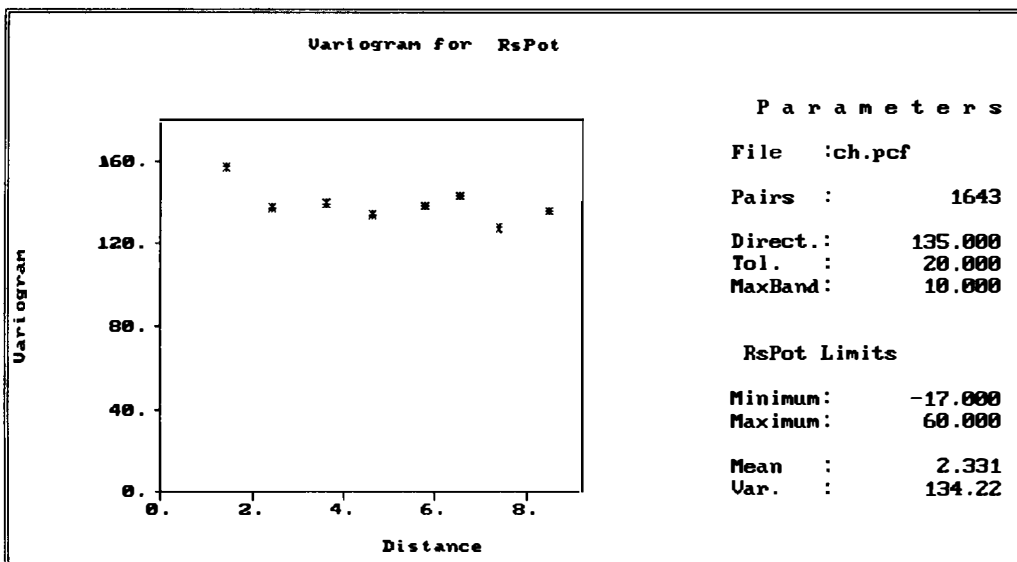
Figure 6.46

Figure 6.47

The pattern of spatial continuity in all the four directions is essentially the same. These semivariograms display a total nugget effect and the residuals can be considered to be spatially uncorrelated. This conclusion can be bring out by consideration of the omnidirectional semivariogram for the residuals which is given in figure 6.48 below.

106
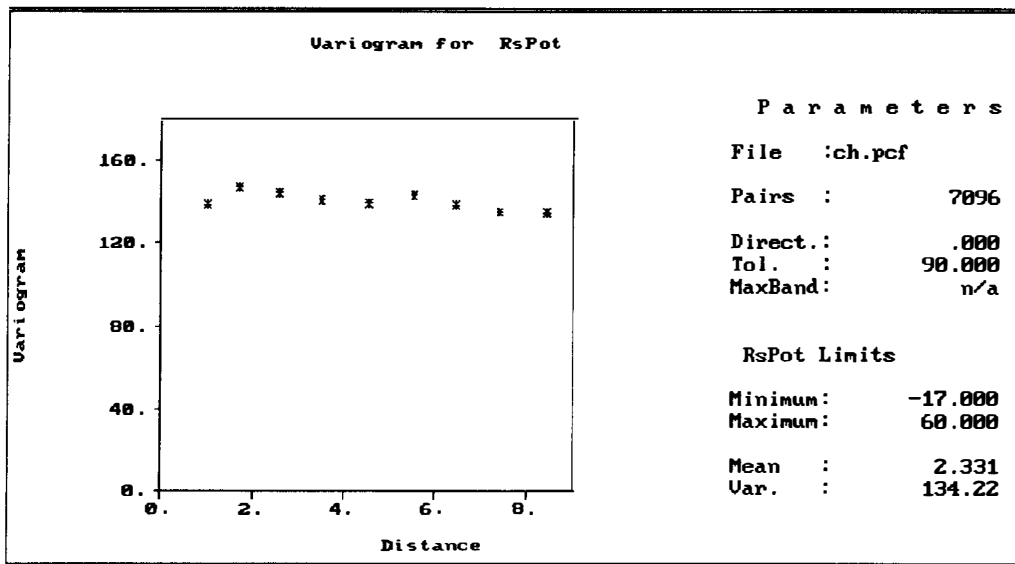
Figure 6.48

This means that all the spatial variation in the data comes from the trend.

The experimental semivariogram values for above directions obtained are given in appendices B.22 to B.26 respectively.

Another approach to the analysis of these data is to consider the semivariogram of the original data in the N-S direction where there is no trend and use this as an overall semivariogram. This semivariogram is given in figure 6.49.

Figure 6.49

107
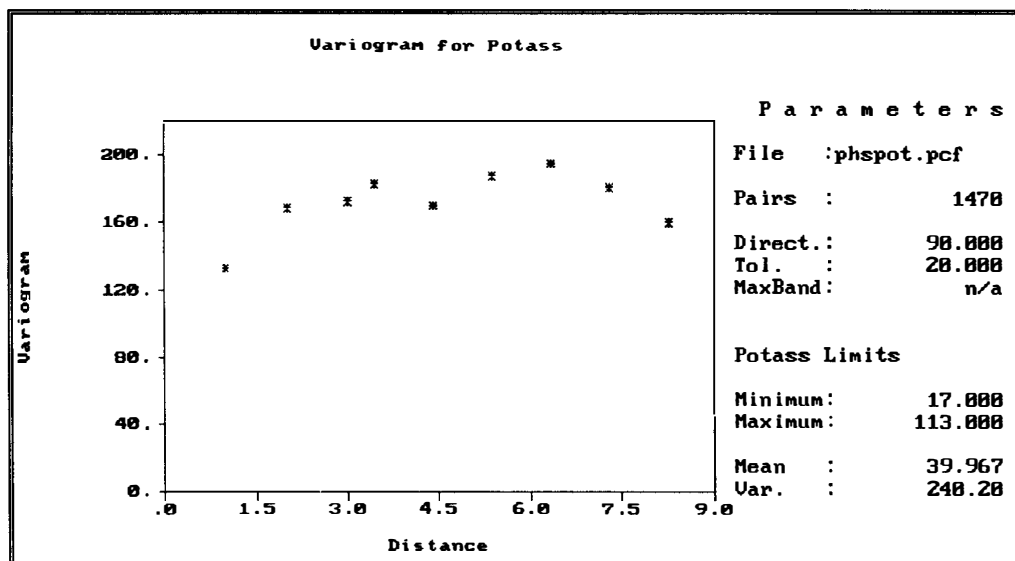
A spherical model with partial sill $c_s = 88$, $c_0 = 100$ and $a_s = 4$, was fitted to the semivariogram estimated in the N-S direction for the cropped field Potassium data using GEOEAS and this fitted model is shown in figure 6.50.
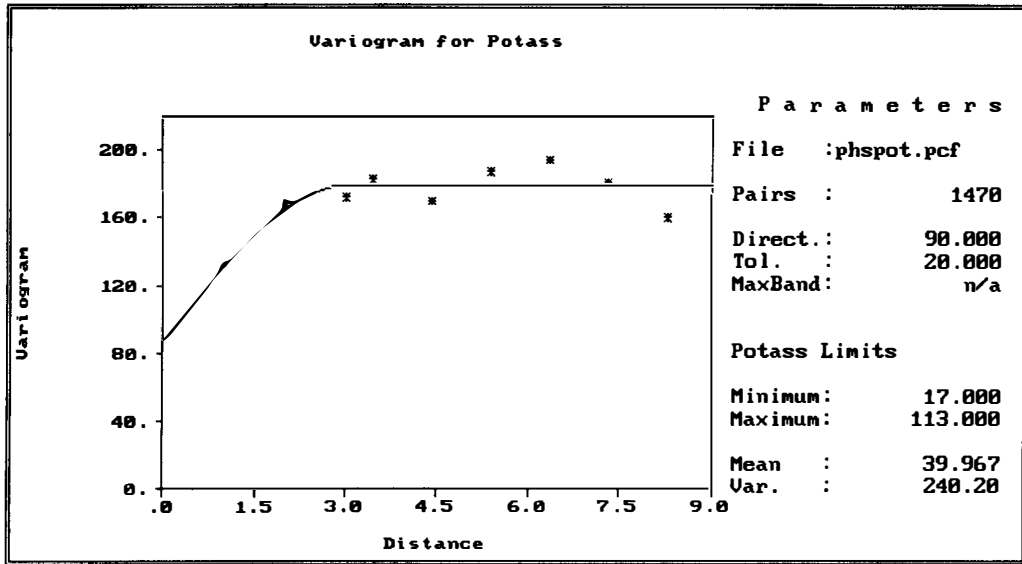
Figure 6.50

Using weighted least squares with these GEOEAS values as the initial values, the parameter estimates $c_s = 93$, $c_0 = 88$ and $a_s = 3$ were obtained in the same direction. This fitted model is given in figure 6.51 below.

Figure 6.51

108
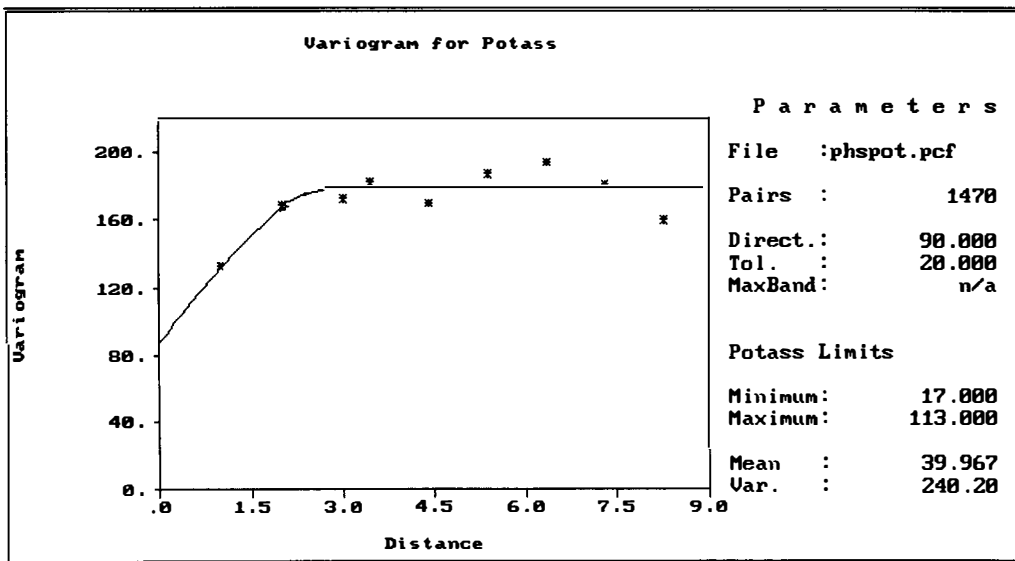
These parameter estimates were then used in the subsequent kriging. Ordinary kriging was appropriate in this case and was performed on the cropped field Potassium data and is shown in figure 6.52.
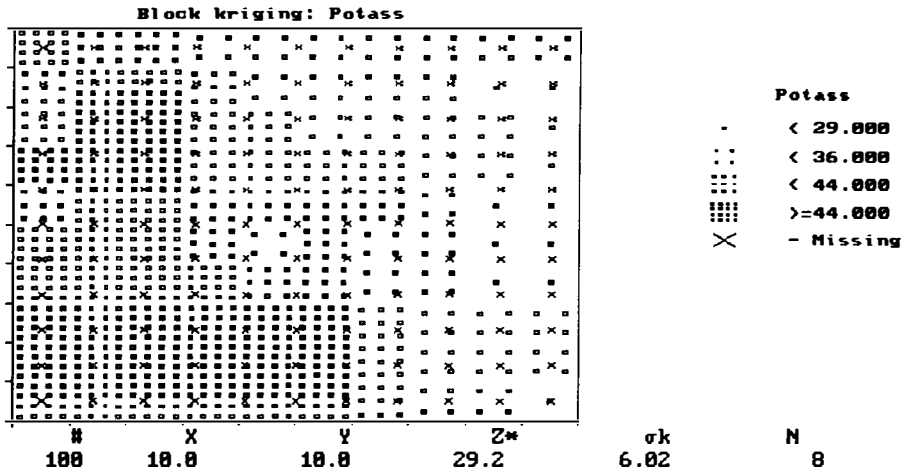
Figure 6.52

The values of the kriging estimate and the kriging variance at location (6.666, 4.444) were obtained as 4.92 and 2.14 respectively, using the surrounding neighbourhood of cropped field Potassium data is shown in figure 6.53.

Kriging estimate and variance at (6.666, 4.444)

Figure 6.53

109
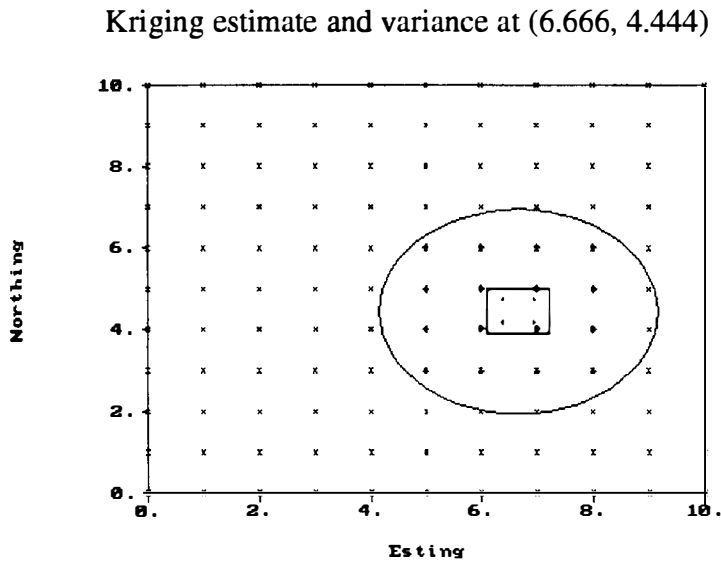
A stem and leaf plot of the uncropped field Potassium data is given in figure 6.54. This indicates that these data contain two global outliers. The last row of stem 3 and leaf 4 indicates 34 ppm uncropped field (truncated) Potassium data. A histogram and box plot are also given in figure 6.55.

Stem and leaf plot for uncropped field Potassium data

```
 27   0   222222233333333333333333333
(51)   0   444444444444444444444444444444455555555555555555555
 39   0   66666666666666666666677777777777
  9   0   888889
  3   1   0
  2   1
  2   1
  2   1
  2   1
  2   2
  2   2
  2   2
  2   2
  2   2
  2   3   1
  1   3
  1   3   4
```

Figure 6.54

Figure 6.55

110
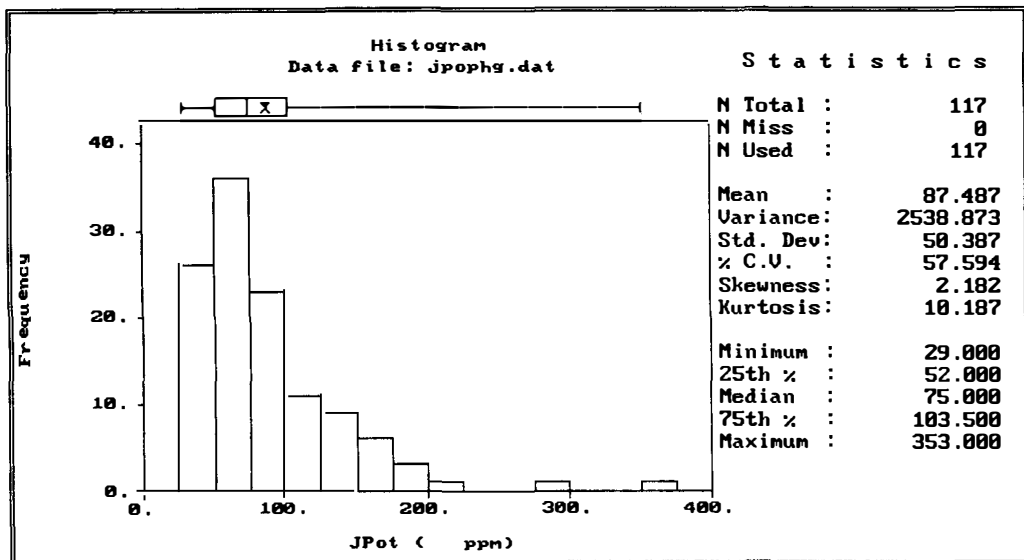
From these figures it can be seen that the data are positively skewed.

Post plot and row and column mean and median plots were also obtained to check for directional trend. These are given in figures 6.56 to 6.59. These plots show trends in the E-W and N-S directions.
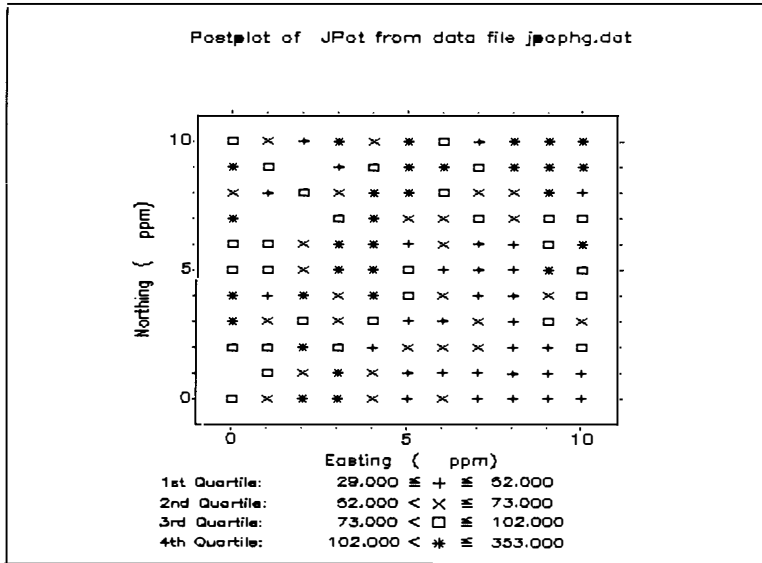
Figure 6.56

Figure 6.57

111
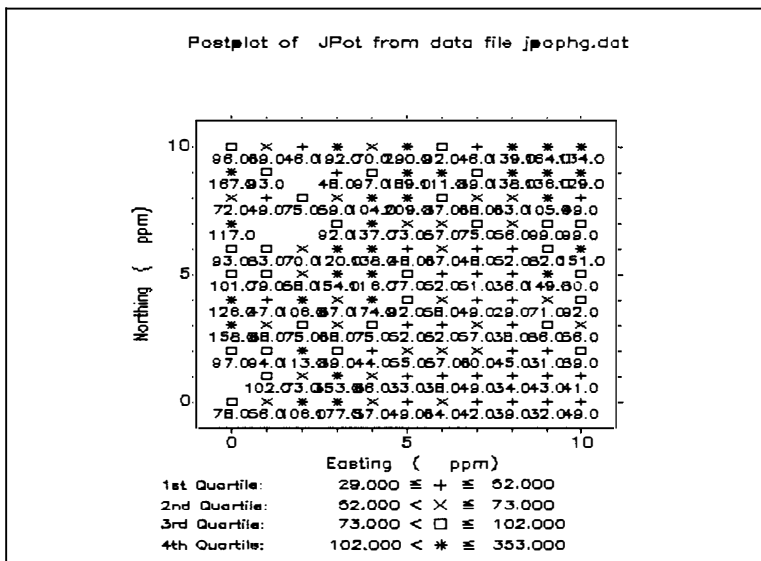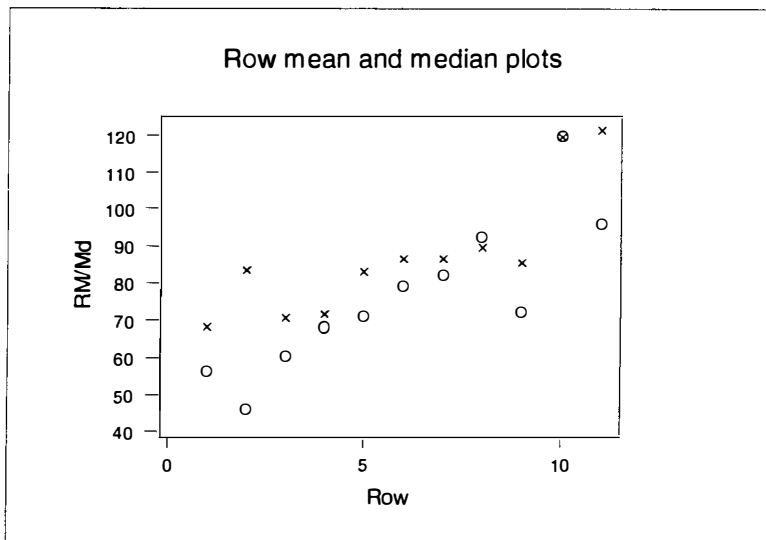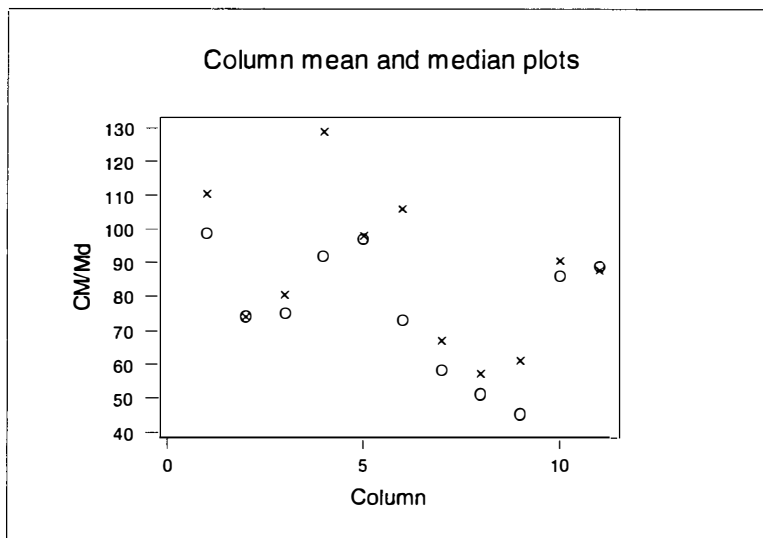
Figure 6.58

Figure 6.59

Median polish will be used here to minimise these trends. Median polish was performed using six iterations on the original cropped field Potassium data and the residuals were treated as a new data set. A stem and leaf plot of the residuals is given in figure 6.60. From this figure we can see three global outliers.

The last row of stem 7 and leaf 3 indicates a (truncated) residual of 7.3 ppm for the

uncropped field Potassium data.

Stem and leaf plot for the residuals of the uncropped field potassium data

```
   3  -1  410
  18  -0  998777766555555
  40  -0  4443333332211111111100
 (31)  0  0001111111222222233333334444444444
  50   0  55555566666666788889
  29   1  01122233
  21   1  555555688
  12   2  022
   9   2  5669
   5   3  00
   3   3
   3   4
   3   4  5
   2   5  3
   1   5
   1   6
   1   6
   1   7  3
```

Figure 6.60

A histogram and box plot are obtained for the residuals are given in figure 6.61.

Figure 6.61

113
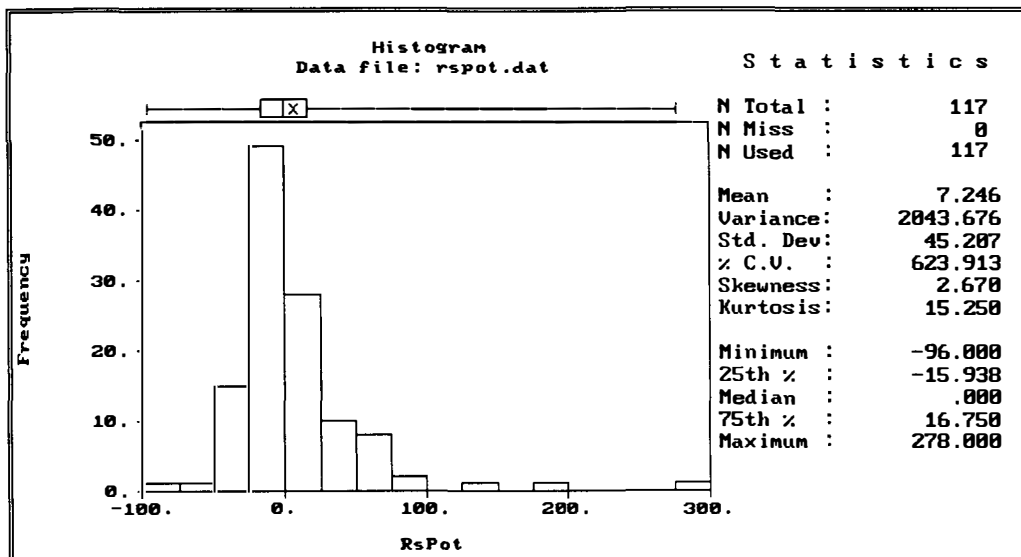
The distribution of the residuals is reasonably symmetric (apart from these global outliers referred to earlier) and so we shall not take any further transformations here. Post plots and row and column mean column mean and median plots were produced for these residuals. These plots are given in figures 6.62 to 6.65.
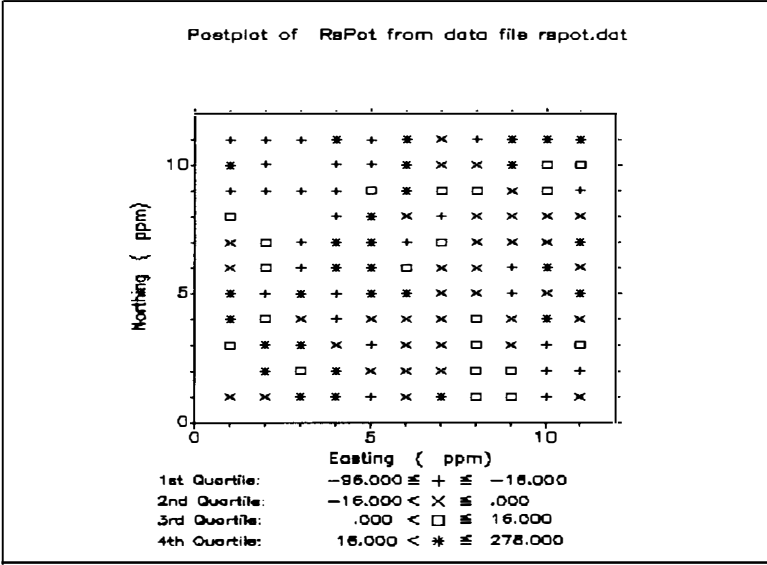
Figure 6.62
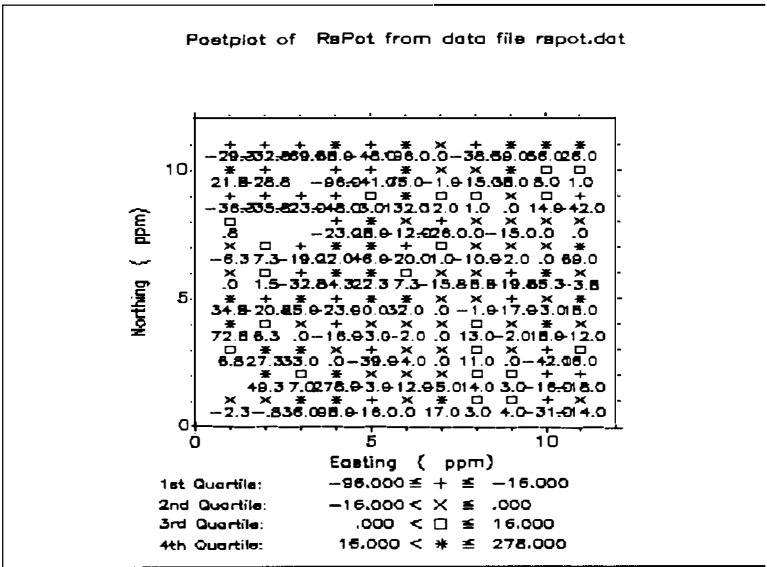
Figure 6.63
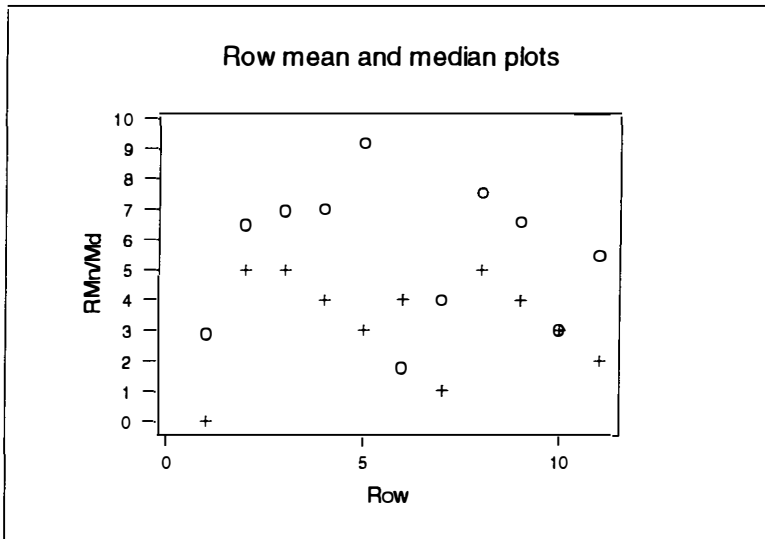
Figure 6.64
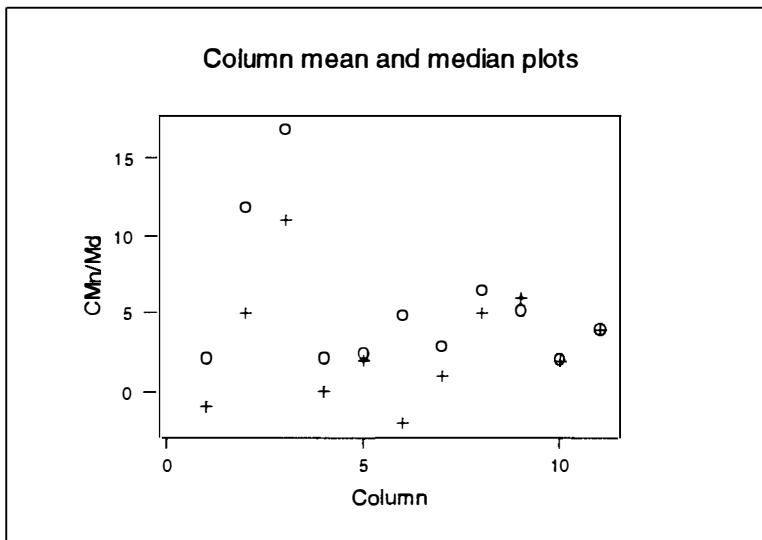
Figure 6.65

An examination of the above plots confirms that the residuals do not exhibit directional trend.

Four directional semivariogram plots were produced to visualise the spatial continuity of the residuals of the uncropped field Potassium data. These four directional semivariograms are in the E-W, N-S, EW-NS and NE-SW directions and are given in figures 6.66, 6.67, 6.68 and 6.69 respectively.
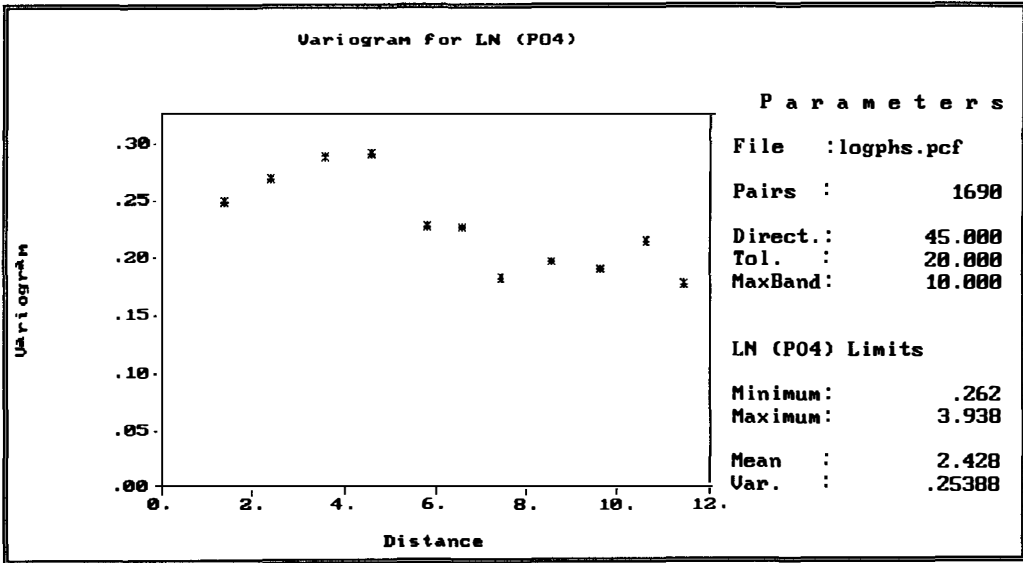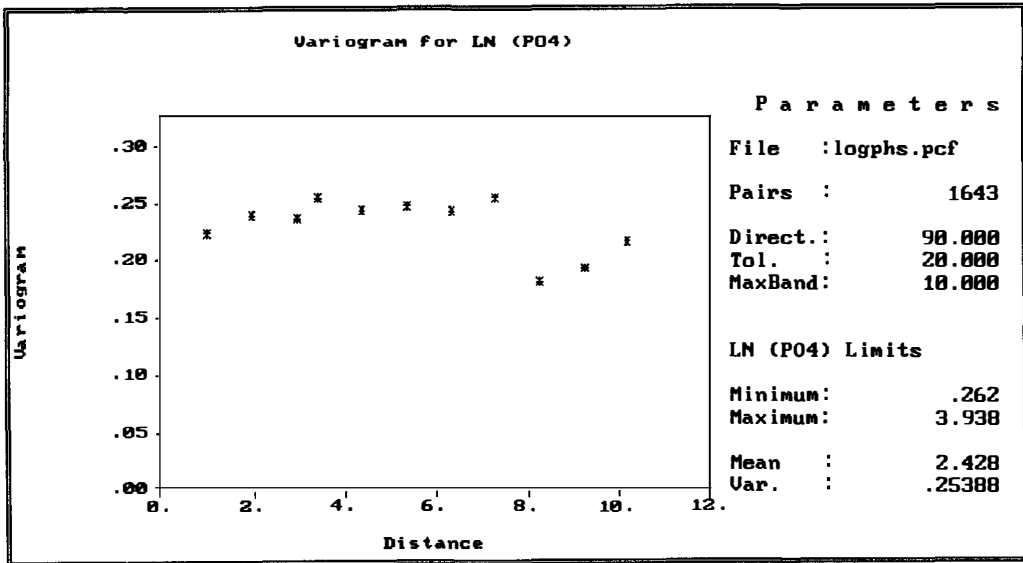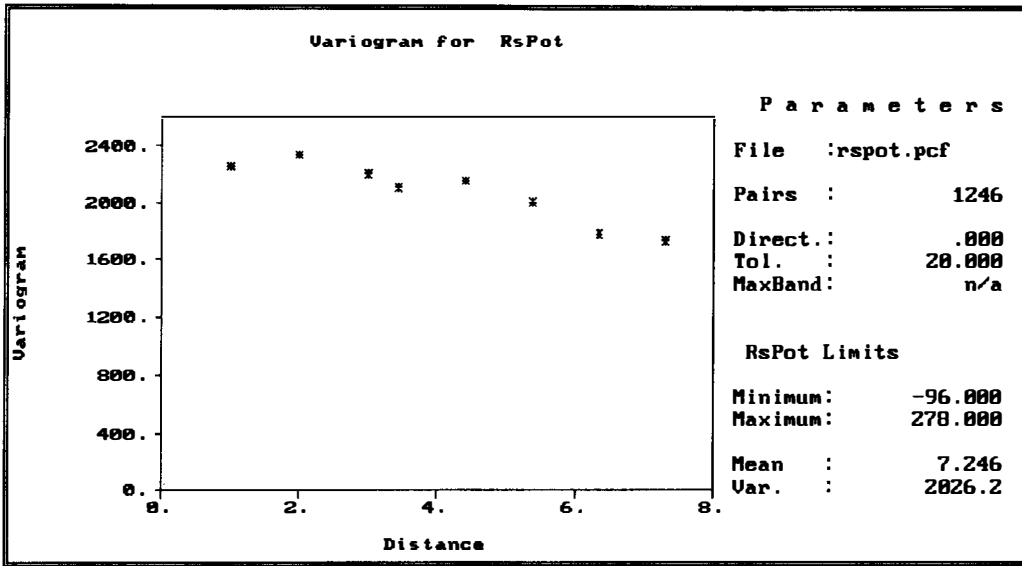
Figure 6.66

Figure 6.67

116
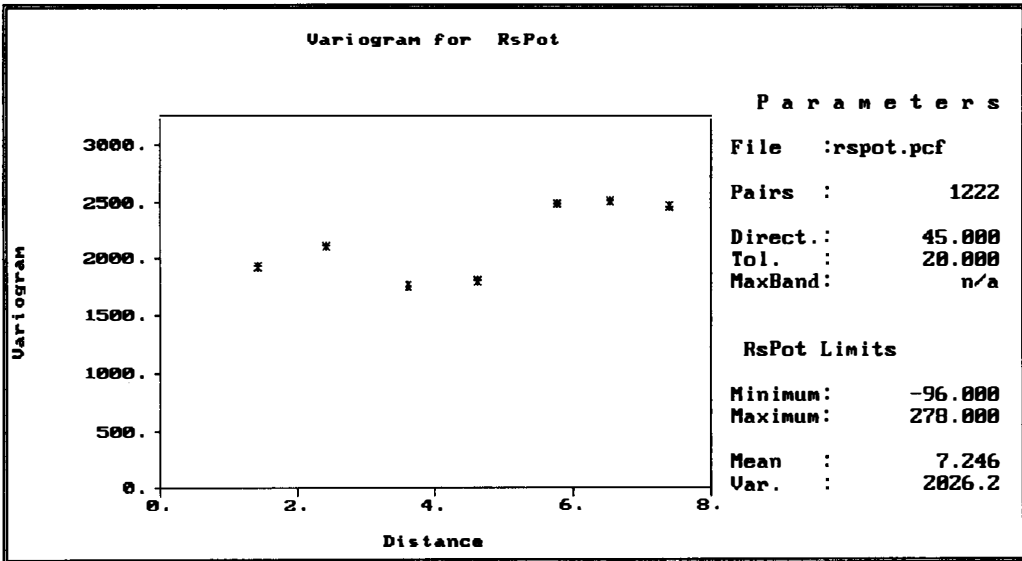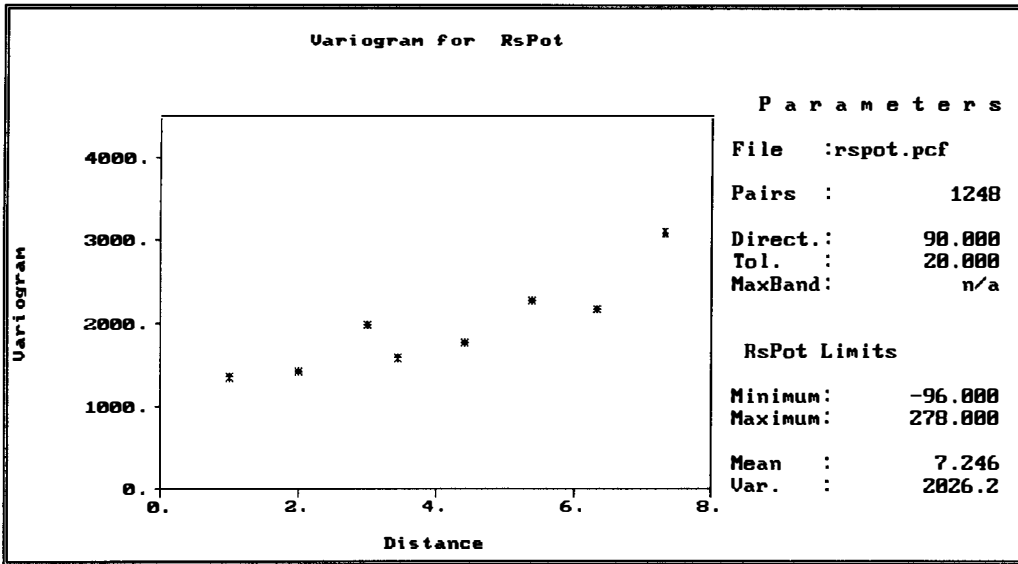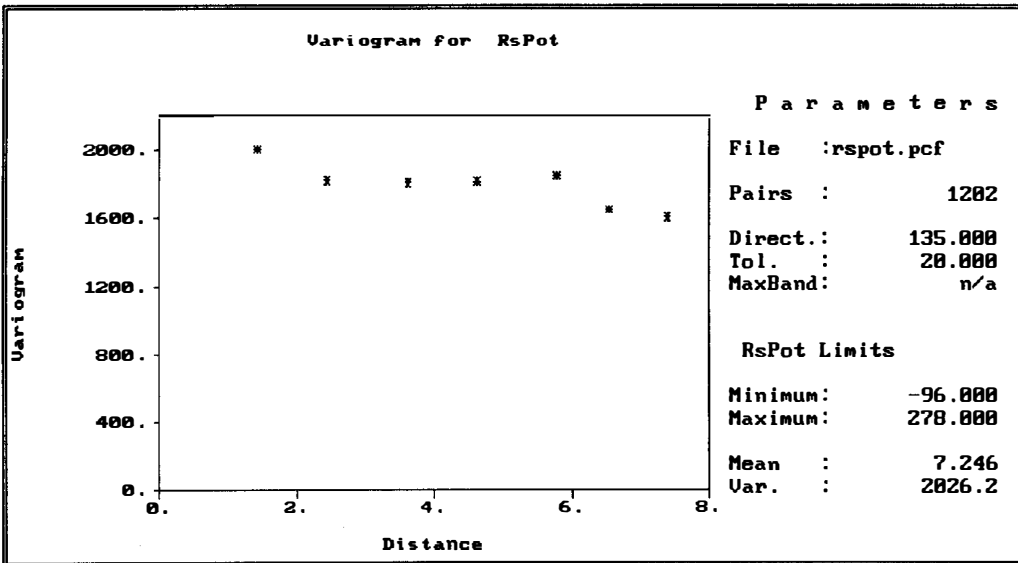
Figure 6.68

Figure 6.69

The pattern of spatial continuity in all directions is essentially the same. These semivariograms display total nugget effect and the residuals are spatially uncorrelated. This means that all the spatial variation in the data comes from the trend. The omnidirectional semivariogram is also obtained to look overall variation in the residuals and is given in figure 6.70.
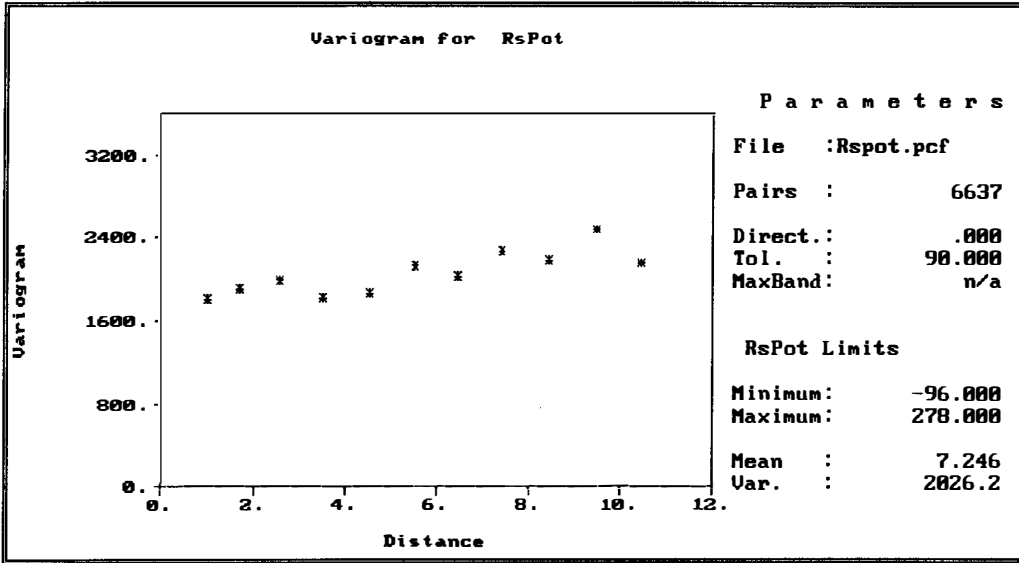
Figure 6.70

All of the experimental semivariogram values for above directions obtained are given in appendices B.35 to B.39.

However, It should be noted here that, although we have taken the Potassium data from the two fields to have the same type of spatial continuity, the situation for the uncropped field is not nearly as clear as that for the cropped field as the semivariograms for the uncropped field are much more erratic than those for the cropped field.

6.4. Banksia Analysis

As mentioned earlier, the Banksia species data are given in Appendix A.10. From the

plots in figures 6.71 and 6.72 it can be seen that, although the distribution of Banksia

species is positively skewed, it contains no global outliers. The last row of stem 2 and

leaf 0 indicates 20 species.

Stem and leaf plot for Banksia data

```
 25   0   111111111111111111111111111
 56   0   2222222222222233333333333333333
(32)  0   4444444444444444455555555555555
 67   0   6666666666666777777777777777777
 37   0   888888888999999
 22   1   0001
 18   1   33333
 13   1   444555
  7   1   67777
  2   1   9
  1   2   0
```

Figure 6.71

Figure 6.72

Post plots and row and column mean and median plots were prepared in order to

check for directional trend for the banksia species. The post plot in figure 6.73

indicates a trend in the N-S direction. The row and column mean and median plots in

figures 6.74 and 6.75 confirm the trend in the N-S direction and the lack of trend in the

E-W direction.

Post plot of Banksia

1st Quartile:	1.000	≤ + ≤	2.000
2nd Quartile:	2.000	< X ≤	5.000
3rd Quartile:	5.000	< □ ≤	7.000
4th Quartile:	7.000	< ＊ ≤	20.000

Figure 6.73

Figure 6.74

120

Figure 6.75

Median polish with four iterations was then performed on the original data of Banksia species to deal with this trend. From the plots of the residuals in figures 6.76 and 6.77 below it can be seen that the distribution of the residuals is reasonably symmetric with no global outliers. The residuals of the Banksia data are given in Appendix A.11. In the last row of stem 1 each leaf 0 indicates a (truncated) residual of 1.0 species.

Stem and leaf plot of residuals for Banksia data

```
  1   -0
  4   -0 544
 13   -0 332222222
 73   1 111000000000000000000000000000000000000000000000000
(36)   0 0000000000000000000000001111111111111
 39    0 2222222222233
 26    0 4444444455555
 13    0 6666
  9    0 88999
  4    1 0000
```

Figure 6.76

121

Figure 6.77

Post plots of these residuals are given in figures 6.78 and 6.79. From these plots we can see no trend in any direction.

Figure 6.78

122
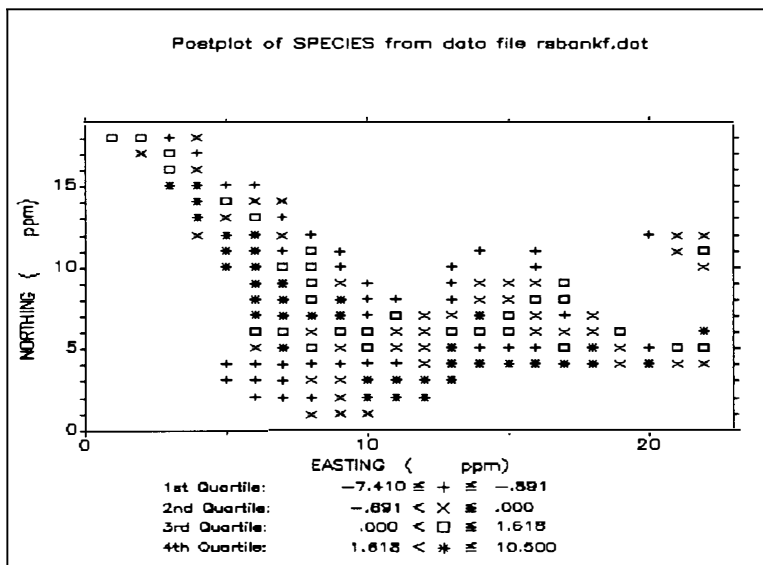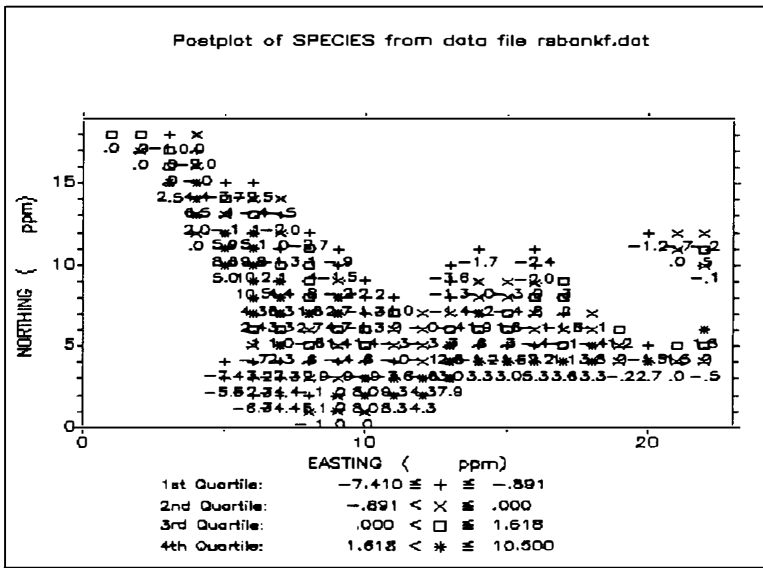
Figure 6.79

The row and column mean and median plots in figures 6.80 and 6.81 confirm the lack

of trend in the E-W and the N-S directions.
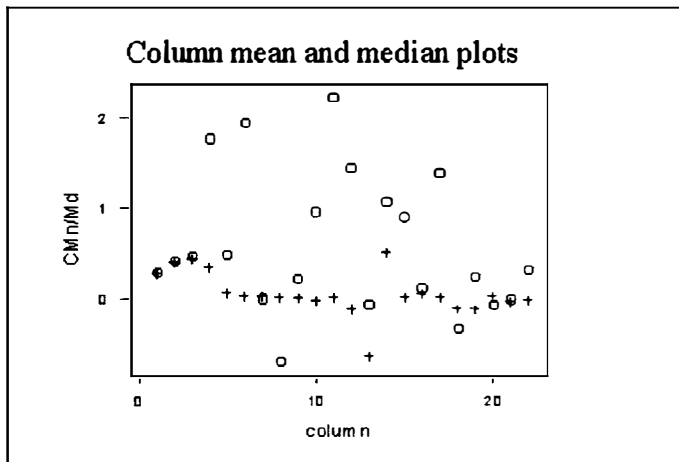
Figure 6.80
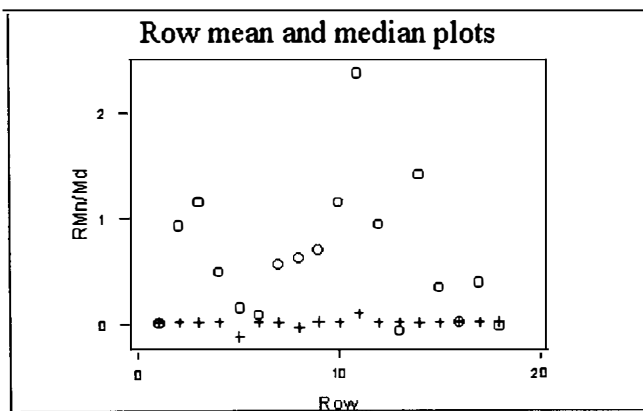
Figure 6.81

123

Four directional semivariogram plots were produced to visualise the spatial continuity of the residuals of the Banksia distribution. These four directional variograms are in the E-W, N-S, EW-NS and NE-SW directions and are given in figures 6.82, 6.83, 6.84 and 6.85 respectively.

The values were obtained using GEOEAS and are given in Appendices B.27 to B30 respectively.
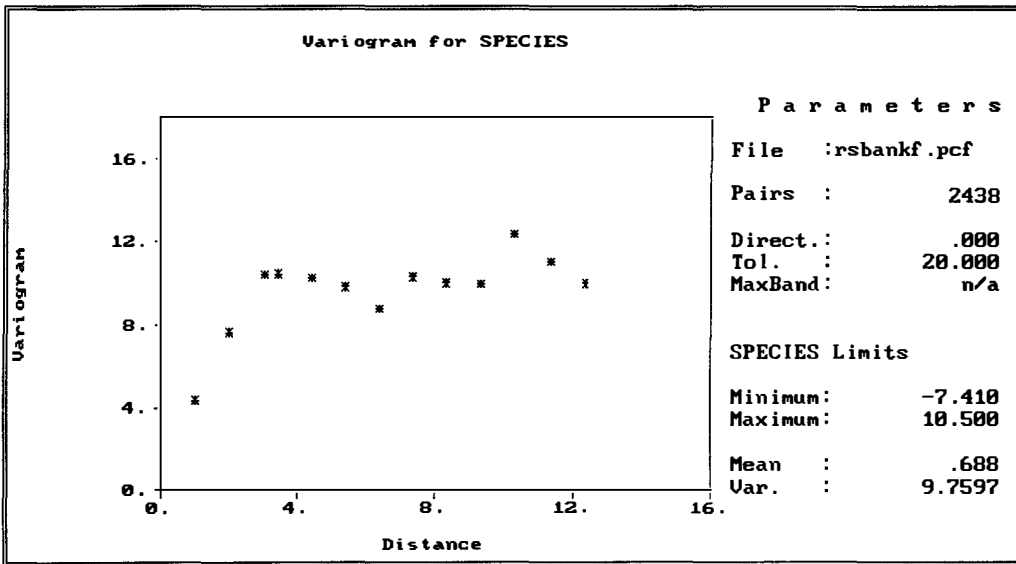
Figure 6.82
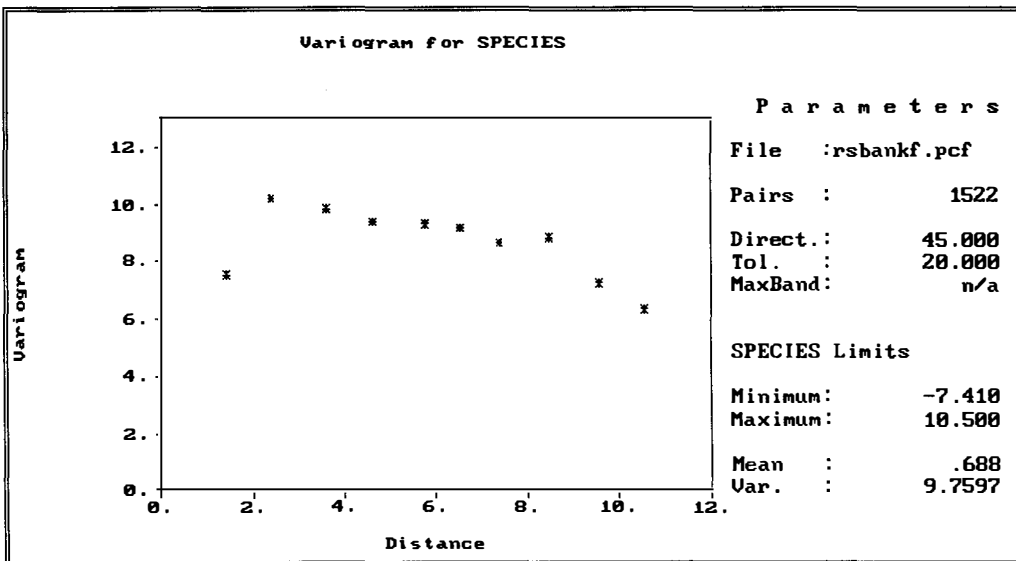
Figure 6.83
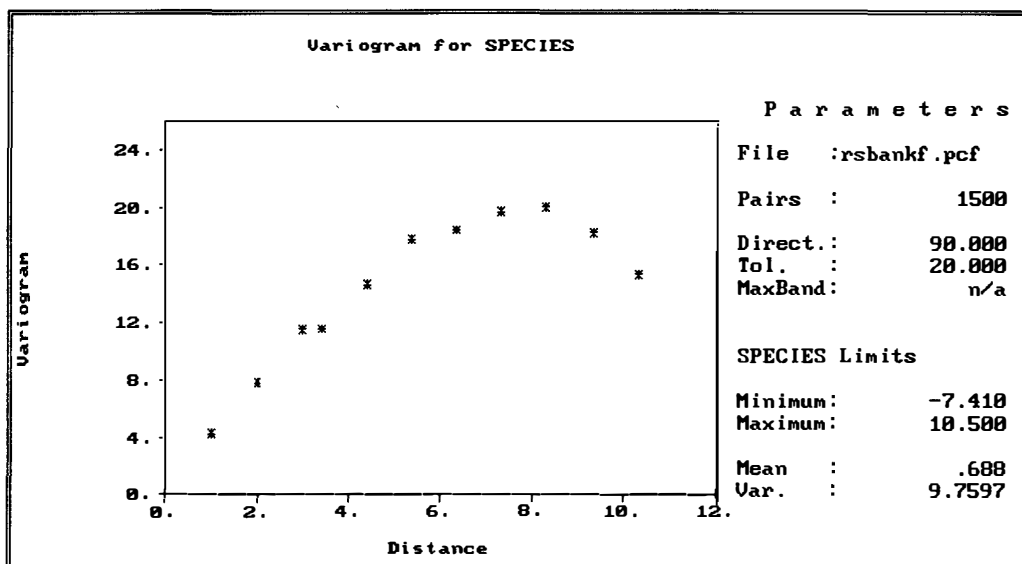
Figure 6.84
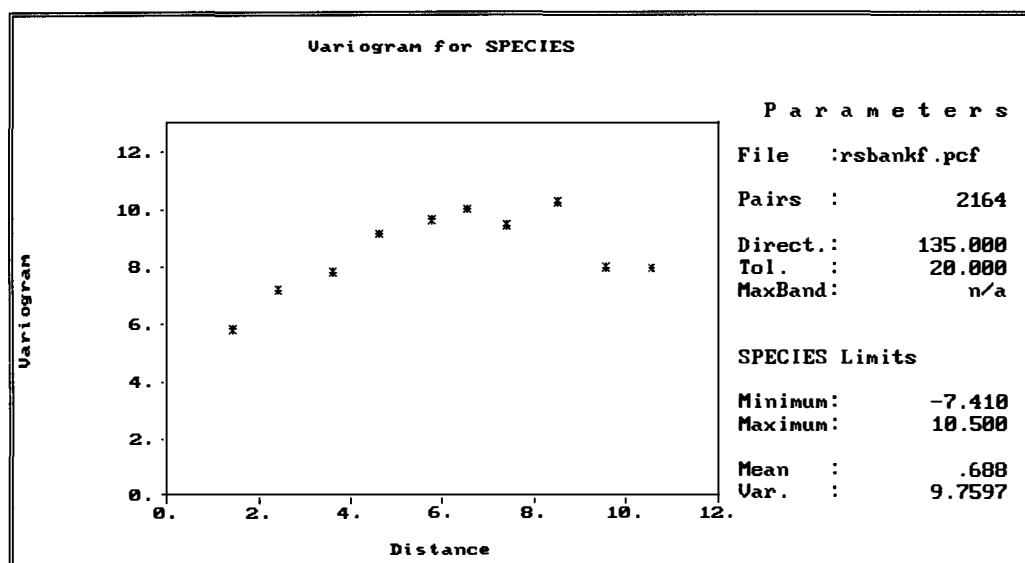
Figure 6.85

From the above figures, we see that, although it would be possible to fit a spherical

model with the same nugget effect and range to all four experimental semivariograms,

and the same sill in the case of the semivariograms in figures 6.82, 6.83, and 6.85. A

different sill is needed for the N-S semivariogram in figure 6.84. This indicates that we

have here a case of sill anisotropy.

The spherical model with $c_s = 9$, $c_0 = 1$ and $a_s = 6$ in the E-W, NE-SW and the SE-NW directions and the spherical model with $c_s = 15$, $c_0 = 1$ and $a_s = 6$ in the N-S direction were fitted to the experimental semivariograms by using GEOEAS. These fitted models are given in figures 6.86, 6.87, 6.88 and 6.89 respectively.
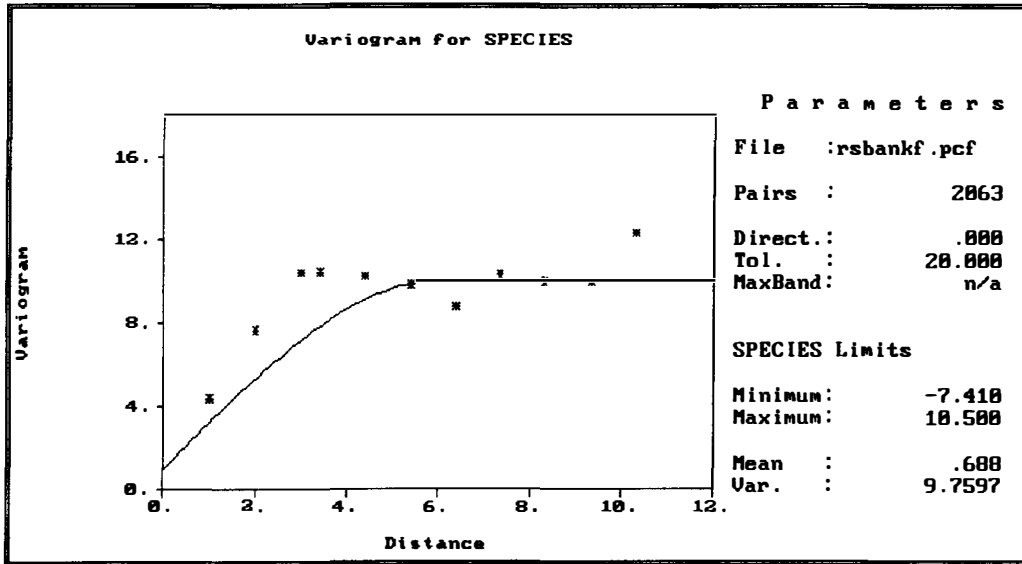
Figure 6.86
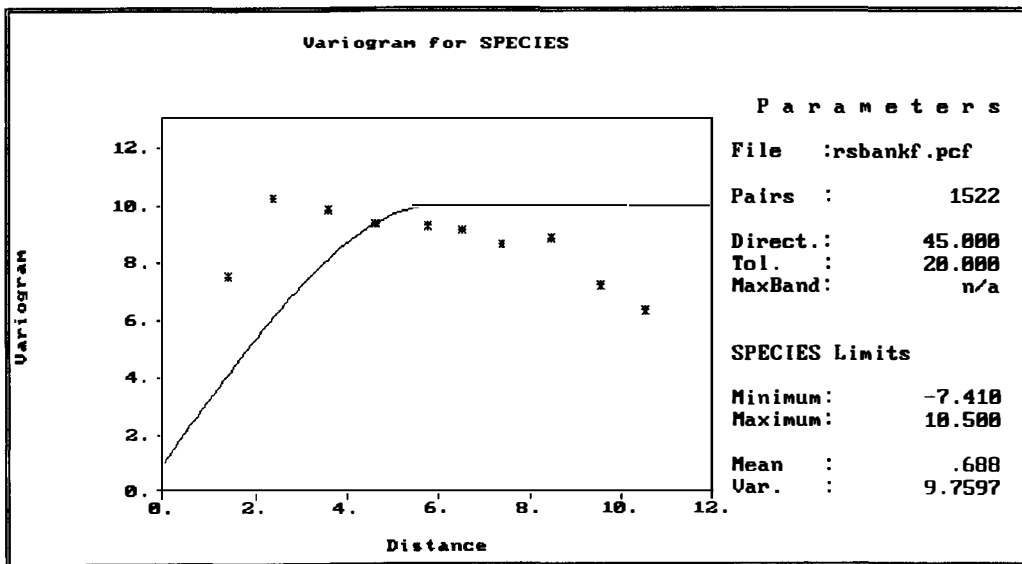
Figure 6.87
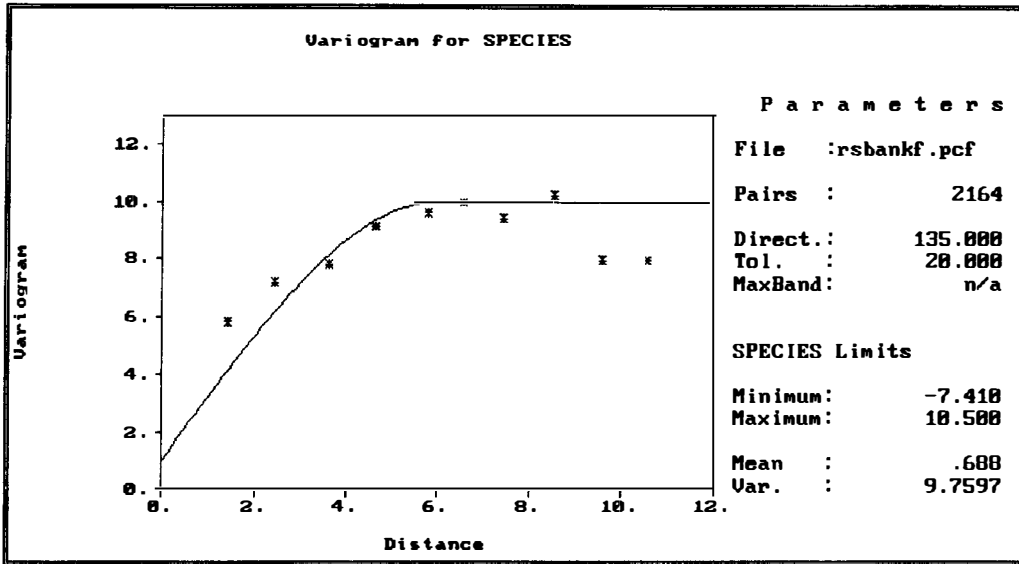
Figure 6.88
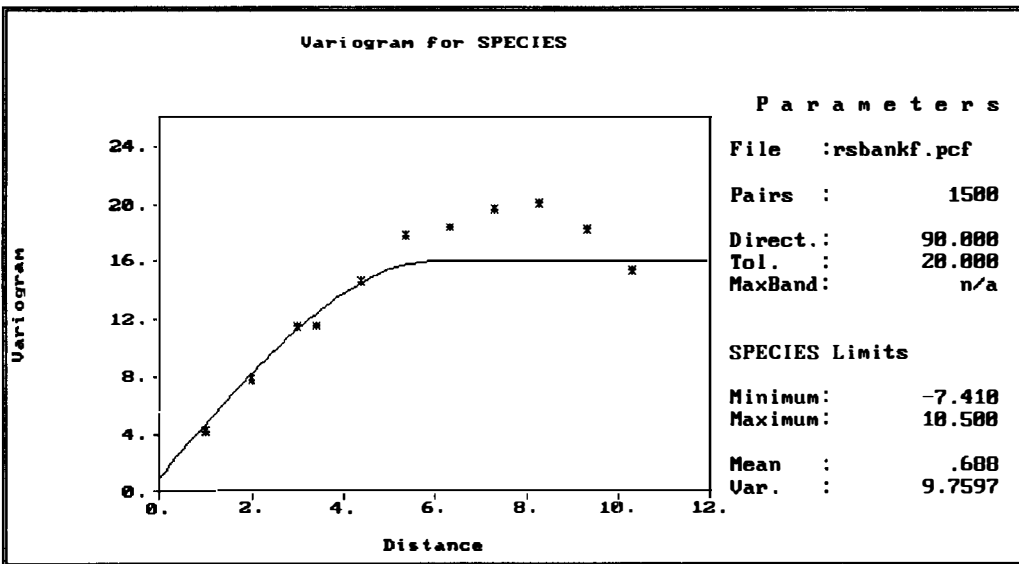
Figure 6.89

To deal with this sill anisotropy we use here a nested semivariogram model as discussed earlier in previous section 3.5. Since the spatial variation in the E-W, SE-NW and NE-SW directions is the same then we can consider one of these direction with the N-S direction.

In this case we will model semivariogram in the E-W diretion with that in N-S the N-S direction to represent the overall spatial continuity of the residuals of the banksia data. The combined (nested) model is given by

$$\gamma(\mathbf{h}) = \begin{cases} 0 & , \ \mathbf{h} = 0 \\ 1 + 9\gamma_s(\|\mathbf{h}\|) + 6\gamma_s(|h_y|) & , \ \|\mathbf{h}\| > 0 \end{cases} \qquad (6.1)$$

where $\mathbf{h} = (h_x, h_y)$ and

$$\gamma_s(h) = \begin{cases} \dfrac{3}{2}\left(\dfrac{h}{6}\right) - \left(\dfrac{1}{2}\right)\left(\dfrac{h}{6}\right)^3 & , \ 0 < h < 6 \\ 1 & , \ h \geq 6 \end{cases} \qquad (6.2)$$

7. Conclusion

In the preceding chapters we have given an account of the Exploratory Data Analysis of spatial data (chapter 2), the essential features of variogram modelling and kriging methods of estimation (chapters 3-5) and an application of these to two new data sets relating to Western Australia (chapter 6).

One advantages of kriging that it provides, via the kriging variance, an idea of the precision of the estimate at each point. However as a measure of the accuracy of the estimates, the kriging variance has disadvantages as well as advantages.

One advantage of kriging itself is that it is an exact interpolator and hence the kriging variance is zero at the data points. Also, the kriging variance depends on the spatial model used. Hence the selection of spatial model is very important and a poor choice of spatial model might lead to imprecise results. On the other hand the kriging variance depends only on the data configuration and is independent of actual data values, producing the same value for two or more locations with same local data configurations. Intuitively the potential for error might be expected to be greater at a location surrounded by unlike data values compared to one with like data values.

What we have given in this thesis constitutes only part (although a very fundamental part) of the vast area of modern geostatistics but should be sufficient to provide the interested reader with a firm foundation for further study in this rapidly expanding area.

For example, we have chosen to restrict our treatment to point kriging. This can be easily extended to block kriging but one then needs to take into account the problem of change of support with subsequent regularisation of the variograms considered. Block kriging is the type used in mining geostatistics.

Another issue of great importance in this area is the problem of conditional bias. Murphy (1988) stated that the conditional bias (regression bias) is zero when the slope of the regression line for predicting values at unknown locations is 1.0 (perfect correlation). In the estimation of climate-scale rainfall for example (Morrissey and Janowiak 1996), the conditional bias produces overestimates of high rainfall and underestimates of low rainfall.

Having noted above that one disadvantage of the kriging variance is that it is not dependent on the actual data values used, one therefore needs some estimate of both local and global uncertainty which is conditioned to the data. One approach to this at the local level is the local confidence interval approach and another is to use conditional simulation (Armstrong and Dowd 1994 ; Cressie 1991).

There is a wide range of simulation algorithms which are currently being used. The method of turning bands was developed by Matheron (1973) for simulating the more general intrinsically stationary processes, and Cressie (1987) shows how to use the Cholesky matrix decomposition method (LU decomposition) for simulation. The LU

decompostion method is also discussed by Dowd and Sarac (1994). Much of the current work using conditional simulation is being done as an application of simulated annealing techniques (Deutsch 1994; Deutsch and Cockerham 1994; Deutsch and Journel 1992).

8. References

Armstrong, M. and Dowd, A. P. (1994). Geostatistical Simulations. Proceedings of the Geostatistical Simulation Workshop, Fontainebleau, France, 27-28 May 1993, Kluwer Academic Publisher.

Bilonick, R. A. (1988). Monthly hydrogen ion deposition maps for the Northeastern U. S. from July 1982 to September 1984. *Atmospheric Environment, 22, 1909-1924.*

Bloom, L. M. (1996). Spatial prediction methods for soil salinity', Zeitschrift fur Angewandte Mathematik und Mechanik Special Issue 3, 381-382.

Boots, B. and Dufournaud, C. (1994). A programming approach to minimizing and maximizing spatial autocorrelation statistics. *Geographical Analysis*, 26(1), 168-184.

Boufassa, A. and Armstrong, M. (1989). Comparision between different kriging estimators. *Mathematical Geology*, 21 (3), 331-345.

Brooker, P. I, Winchester, J. P. and Adams, A. C. (1993). A Geostatistical study of soil data from the Riverland, South Australia. Proceedings: *International congress on Modelling and Simulation*, Vol 4.

Burgess, T. M. and Webster, R. (1980). Optimal Interpretation and isarithmic mapping of soil properties I. The semivariogram and punctual kriging. Journal of Soil Science 31, 135-331.

Chiles, J. P. and Guillen, A. (1984). Variogrammes et krigeage pour la gravimetrie et le magnetisme, *Sciences de la Terre*, Serie Informatique 20, 455-468.

Christensen, R. (1990). The equivalence of predictions from universal kriging and intrinsic random-function kriging. *Mathematical Geology.* 22, 655-664.

Cressie, N. (1985). Fitting Variogram models by weighted least squares. *Mathematical Geology, 17* (5), 563-587.

Cressie, N. (1986). Kriging nonstationary data. *Journal of the American Statistical Association*, 81, 625-634.

Cressie, N. (1987). A nonparametric view of generalized covariance for kriging. *Mathematical Geology*, 19, 425-449.

Cressie, N. (1988). Spatial prediction and ordinary kriging. *Mathematical Geology.*, 20, 405-421.

Cressie, N. (1991). Statistics for spatial data . New York: John Wiley and Sons, INC.

Cressie, N. and Glonek, G. (1984). Median based covariogram estimators reduce bias. *Statistics and probability letters*, 2, 115-125.

Cressie, N. and Hawkins, D.M. (1980). Robust Estimation of variogram. *Journal of the International Association for Mathematical Geology, 12, 115-125.*

Cressie. N and Helterbrand, D. H. (1994). Universal Cokriging Under Intrinsic Coregionalization. *Mathematical Geology*, 26(2), 166-185.

Cressie, N. and Hoef, J. M. V. (1993). Multivariate spatial prediction. *Mathematical Geology*, 25(2), 219-240.

Cressie, N. and Hamlett, J. M. and Horton, R. (1986). Resistant and exploratory techniques for use in semivariogram analysis. *Soil Science Society of Americal Journal*, 50, 868-875.

de Kwaadsteniet, J.W. (1990). On some fundamental weak spots of kriging technique and their consequences., *Journal of Hydrology*, 114(3/4), 277-284.

Deutsch C. (1994). Constrained modelling of histograms and cross-plots with simulated annealing, Report 7, Stanford Centre for Reservoir Forecasting, Stanford, CA.

Deutsch, C. and Journel, A (1992). Annealing techniques applied to the integration of geological and engineering data, Report 5, Stanford Centre for Reservoir Forecasting, Stanford, CA.

Deutsch, C. and Cockerham, P. (1994). Practical considerations in the application of simulated annealing to stochastic simulation. *Mathematical Geology* 26(1), 67-82.

Dowd, P. A. and Sarac, C. (1994). Geostatistical Simulation. An extension of the

LU Decomposition Method of Simulation, in Armstrong, M. and Dowd, P. (eds). Proceedings of the Geostatistical Simulation Workshop, Fontainebleau, France, 27-28 May 1993, Kluwer Academic Publisher.

Englund, E. and Sparks, A. (1991), GEO-EAS 1.2.1 User Guide: US Environmental Protection Agency, Las Vegas.

Furr, J. H. and Hampen W. V. (1986). Geostatistical analysis of potentiometric data in the Wolfcamp Aquifer of the Pale Duro Buston, Texas. Technical report BMI/ ONWI, 587.

Fytas, K., Chaounai, N.-E. and Lavigne, M. (1990). Gold deposits estimation using indicator kriging. *CIM,* v82, No. 934, 77-83.

Galli, A. (1984). Factorial kriging analysis, in Verly et al (eds), *Geostatistic for Natural Resources Characterization, Paty 1, 543-537.*

Galli, A. and Wackernagel, H. (1988). Multivariate geostatistical methods for spatial data analysis. *Data Analysis and Informatics V, Didey ed, 273-280.*

Gandin, L.S. (1963). Objective analysis of meteorological fields. *gidrometeorologicheskoe izdatel'stvo (GIMIZ). Leningard (translated by Israel Program for Scientific Translations, Jerusalem, 1965).*

Goovearts, P. (1991). Etude des relations entre proprietes physico-chimiques du sol par la geostatistique multivariable. In Compte-rendu des Journess de

Geostatistique, Fascicule 1, pp. 247-261, ENSMP.

Goovaerts, P. (1992). Factorial Kriging Analysis: a useful tool for exploring the structure of multivariate spatial soil information. The Journal of Soil Science, 43(4), 597-619.

Goovaerts, P. (1994). On the controversial method for modeling a coregionalization. *Mathematical Geology.* 26(2), 197-204.

Gotway, C. (1992). Fittting semivariogram models by weighted least sqaures. *Computers and Geosciences,* 17, 171-172.

Goulard, M. (1989). Inference in a coregionalization nodel. In Geostatistics, Vol. 1 (ed. M. Armstrong), pp. 397 - 408. Kluwer Academic, Dordrecht.

Goulard, M. and Voltz, M. (1992). Linear Coregionalization Model : Tools for Estimation and choice of the cross variogram matrix. *Mathematical Geology,* 24(3), 269-283.

Hass, T.C (1990). Kriging and automated variogram modelling within a moving window. *Atmospheric Envirnment. part a : General Topi,* 24(7), 1759-1779.

Haining, R. (1990). Spatial data analysis in the social and environmental sciences. Cambridge University Press.

Huijbregts, C. J. and Matheron, G. (1970). Universal kriging (An optimal method for estimating and contouring in trend surface analysis). *In Proceedings of Ninth International Symposium on Techniques for Decision-Making in the Mineral Industry*, J. I. McGerrigle, ed. The Canadian Institute of Mining and Metallurgy, Special Volume 12, 159-169.

Isaaks, E.H and Srivastava, M. (1988). Spatial continuity measure for probabilistic and deterministic geostatistics. *Mathematical Geology, 20 (4), 313-341.*

Isaaks, E.H and Srivastava, M. (1989). An introduction to Applied Geostatistics. New York: Oxford University Press.

Jaquet, O. (1989). Factorial kriging analysis applied to geological data from petroleum exploration. *Mathematical Geology,* 21, 683-691.

James, A. T. and Venables, W. N. (1993). Matrix weighting of several regression coefficient vectors sources. *Annals of Statistics.* 21(2), 1093-1114.

Journel, A. G. (1990). Characteristics behavior and order relation for indicator variograms. *Mathematical Geology,* 22(8), 1011-1032

Journel, A.G and Huijbregts, C. (1978). Mining Geostatistics. Centre de

Geostatistique Fontainebleau, France.

Jowett, G. H. (1952). The accuracy of systematic sampling from conveyer belts. *Applied Statistics*, 1, 50-59.

Kendall, M. G. and Stuart, A. (1969). The Advanced Theory of Statistics, Vol. 1, 3rd ed. Griffin, London.

Kolmorogorov, A.N. (1941). Interpolirovanie I ekstrapolirovanie stationarnykh sluchainykh posledovatel 'nostei' (Interpolated and extrapolated stationary random sequences). Isvestia AN SSSR, seriya matematicheskaya 5(1), 3-14.

Krige, D.G. (1951). A statistical approach to some basic mine valuation problems on the Witwatersrand. *Journal of the Chemical, Metallurgical and Mining Society of South Africa*, 52, 119-139.

Krige, D.G. (1966). Two dimensional weighted moving average trend surfaces for ore-evaluation. *Journal of the South African Institute of Mining and Matallurgy* 66, 13-38.

Lamont, B and Connell, C (1994). The study on the biogeography of Banksia distribution in Southwestern Australia. Department of Environmental Biology, Curtin University of Technology.

Laslett, G. and McBratney, A. (1990). Further comparison of spatial methods for predicting soil pH. *Journal of Soil Science*, 41, 451-471.

Lemmer, I. C. (1984). Estimating local recoverable reserves via indicator kriging. In Geostatistics for Natural Resources Characterization, Part 1, G. Verly, M.

David, A. Journel, and A. Marechal, eds. Reidel, Dordrecht, 365-384.

Ma, Z. Y and Royer, J.J. (1988). Local geostatistical filtering Application to remote sensing. Sciences de la Terre, Serie Informatique 27, 17-36.

Matheron, G (1962). Traite de Geostatistique Appliquee, Tome I. *Memoires du Bureau de Recherche Geologiques et Minieres*, No. 14 Edition Technip, Paris.

Matheron, G. (1963a). Traite de Geostatistique Appliquee, Tome II: Le Krigeage, Memories du Bureau de Recherches Geologiques et Minieres, Editions Bureau de Recherche Geologiques et Minieres, Paris 24.

Matheron, G. (1963b). Principles of Geostatistics. (Economic Geology, 58, 1246-1266.

Matheron, G. (1965). Les variables regionalisees at leur estimation. Masson, Paris.

Matheron, G. (1971a). The theory of regionalized variables and its applications. *Cahiers du Centre de Morphologie Mathematique de Fountainebleau No 5.*

Matheron, G. (1971b). The theory of Regionalized Variables and its applications. *Cahiers du Centre de Morphologie Mathematique, No. 5. Fontainebleau. France.*

Matheron, G. (1973). The intrinsic random functions and their applications, *Ad Appl Prob,* 5, 439-468.

McBratney, A., Webster, R., McLean, R. G. and Spiers, R. B. (1982). Regional variation of extractable copper and cobalt in the topsoil of South East Scotland. *Agronomie* 2, 969-982.

McBratney, A. and Webster, R. (1986). Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates. *Journal of Soil Sciences,* 37, 617-639.

Morrissey M. L. and Janowiak, J. E. (1996). Sampling-induced conditional bias in satellite climate-scale rainfall estimates. *Journal of Applied Meteorology*, 35(4), 541-548.

Murphy, A. H. (1988). Skill scores based on the mean square error and their relationships to the correlation coefficient. *Monthly Weather Review*, 116, 2417-24.

Myers, D. and Journel, A. (1990). Variograms with zonal anisotropies and noninvertible kriging system, *Mathematical Geology*, 22(7), 779-785.

Rouhani, S and Wackernagel, H. (1990). Multivariate geostatistical approach to space-time data analysis. Water Resources Research 26, 585-591.

Royer, J.J. (1984). Proximity analysis: a method for multivariate geodata processing. Sciences de la Terre, Serie Informatique, 20, 223-243.

Ruppert, D. and Wand MP. (1994). Multivariate locally weighted least squares regression. *Annals of Statistics*, 23(3), 1346-1370.

Solow, A. R. (1993). On the efficiency of the Indicator Approach in Geostatistics. *Mathematical Geology.* 25(1), 53-57.

Stein, M.L. (1988). Asymptotic distributions of minimum norm quadratic estimators of the covariance function. *Annals of Statistics*, 16, 55-63

Unlu, K., Nielsen, D. R. and Bigger, J. W. (1990). Statistical Parameters Characterizing the Spatial Variability of Selected Soil Hydralic Properties. *Journal Info Soil Science Society of America Journal,* 54(6), 1537.

Velleman, P. F. and Hoaglin, D. C. (1981). Applications, Basics, and Computing of Exploratory Data Analysis. Duxbury, Boston, MA.

Wackernagel, H. (1988). Geostatistical techniques for interpreting multivarite spatial information, in Chung et al eds. *Quantitative Analysis of Mineral and Energy Resources, 393-409.*

Wackernagel, H. (1989). Overview of methods for analyzing multivarite sequences of soil data, in Bock, H. ed., *Classification and Related Methods of Data Analysis,* 641-650.

Wackernagel, H. (1995). Multivariate Geostatistics. An Introduction with Applications. Springer-Verlag Berlin Heidelberg New York.

Wackernagel, H., Petitgas, P. and Touffait, Y. (1988). Overview of methods for co-regionalization analysis: M. Armstrong (ed.), Geostatistics. *Quantitive Geology and Geostatistics : Kluwer Academic Publishers,* vol. 1, 409-420.

Wartenberg, D., Uchrin, C. and Coogan, P. (1991). Estimating exposure using kriging: a simnlation study., *Environmental Health Perspectives,* 94, 75-82.

Wartenberg, D. (1985). Multivariate spatial correlation: a method for exploratory geographical analysis. Geographical Analysis 17, 263-283.

Webster, R. and McBratney, A.B (1989). On the Akaike Information Criterian

for choosing models for variograms of soil properties. The Journal of Soil Science, 40 (3), 493-520.

Webster, R. and Oliver, M.A. (1990). Statistical Method in soil and land resources survey. New York: Oxford University Place.

Webster, R., Atteia, O. and Dubois, J.P. (1994). Coregionalization of trace metals in the soil in the swiss Jura. *European Journal of soil science*, 45, 205-218.

Weiner, N. (1949). Extrapolation, Interpolation, and Smoothing of stationary time series. MIT Press, Cambridge, MA.

Wold, H. (1938). A study in the analysis of stationary time series. Almqvist and Wiksells, Uppsala.

Yaglom, A.M. (1962). Some classes of random fields in n-dimensional space, related to stationary random processes. *Theory of Probability and its Applications*, 2, 173-320.

Yost, et al. (1982). Geostatistical analysis of soil chemical properties of large landareas. I semi-variogram, II, kriging. *Soil Science Society of America Journal*, 46, 1028-1037.

Zimmerman, L.D. (1993). Another look at anisotropy in geostatistics *Mathematical Geology, 25(4), 453-468.*

Zhang, X. F., Eijkeren, V. J. C. H and Heemink, A. W. (1995). The weighted least square method for fitting a semivariogram model. *Computers and Geoscience*, 21 (4), 605-608.

9. APPENDICES

9.1. Appendix A

A.1 Iron ore data

60.8	54.8	58.9	59.1	57.4	59.1	57.9	54.4								
	57.8	58.1	57.8	56.1	58.7	58.7	57.3	58.1	51.5	54.2					
	55.1	59.6	53.2	58.4	55.9	58.6	56.1	58.5	56.9	56.3	56.6	54.3	52.5	54.3	51.3
	54.3	51.9	52.0	56.7	57.8	50.3	58.2	58.3	59.0	58.9	59.7	52.9	52.8	57.1	
	56.2	55.4	54.3	50.0	58.9	56.5	51.6	57.1	58.4	55.7	59.6	56.6	56.9	59.0	
45.6	46.9	53.5	52.2	53.5	55.7	52.2	47.2	54.1	57.6	58.7	58.6	59.1	56.1	56.7	55.0	56.4
47.1	50.5	51.4	55.8	47.5	51.9	55.0	48.0	45.4	56.9	53.7	57.5	55.1	53.5	47.5	57.6	60.3
	50.4	54.0	51.3	47.8	52.6	50.5	51.2	52.1	53.5	56.7	51.5	56.2	56.8	59.7	
						55.9	55.1	53.1							

Table 2.1

A.2 Residuals of the iron ore data

2.37	-3.86	1.30	0.76	-2.57	.63	-.63	-2.4							
	0.63	.69	1.45	-.98	-0.018	1.49	.018	-.815	.011	-1.159				
0.74	5.00	-0.32	4.18	0.00	4.20	1.63	.333	1.82	1.15	-0.79	-0.76	-1.45	-2.85	-5.43
	-.33	-2.97	-1.80	2.15	1.61	-4.3	3.45	-2.09	.004	-3.2	2.02	-2.48	-1.434	-.333
	0.34	-0.69	-0.73	-5.77	1.49	.60	-4.36	.25	2.14	1.47	.69	-.004	1.44	.34
-1.38	-2.43	-.51	-2.15	0.31	1.77	-3.36	-6.85	-.025	.81	-1.59	1.63	2.03	1.33	3.08	-1.81	.00
1.37	2.42	-1.35	2.80	-4.4	-0.76	.69	-4.79	-7.46	-2.3	-1.07	2.52	-.700	.00	-4.85	2.04	5.16
	-0.66	2.69	1.05	-3.18	-.01	-.609	.018	0.0	-2.57	2.61	4.38	6.12	5.83	
										.004	0.0	-2.54		

A.3 Coalash Data

				8.59	9.00	11.86	8.91	9.99			
		11.62	10.91	8.76	8.89	9.10	7.62	9.65			
	10.39	10.65	10.36	9.58	10.66	8.92	7.80	7.84	9.03	8.60	
	9.79	9.06	10.70	11.21	8.98	9.27	8.19	7.88	7.61	8.20	8.77
	10.74	12.80	10.03	9.36	8.57	9.01	9.04	7.28	9.58	9.69	9.96	9.91
	11.21	9.89	10.34	8.20	9.82	10.06	8.58	8.89	8.64	7.04	8.81	7.95
	9.97	9.70	9.84	10.29	9.84	10.01	9.01	7.68	9.25	7.83	9.14	7.63	9.07
11.17	10.14	9.93	10.27	10.21	11.09	10.63	8.82	10.18	9.34	8.61	
9.92	10.82	11.65	8.96	9.88	8.90	10.18	9.34	10.56	9.06		
10.21	10.73	9.46	9.35	9.78	10.38	9.79	8.91	9.22	11.43		
	12.50	9.63	10.82	10.12	9.40	9.48	10.99	9.92	7.85	8.21	
9.92	11.05	10.11	11.46	10.41	8.45	8.90	8.07	7.96	7.00	7.90	
11.31	9.41	9.37	11.21	9.93	10.70	9.27	9.28	10.13	8.61	8.78	
11.15	9.91	10.17	10.55	11.61	9.16	10.04	11.19	8.10	11.30		
	10.82	11.75	9.78	11.00	9.79	10.19	9.15	8.15	9.20		
10.01	8.23	11.04	10.28	13.07	10.47	11.58	9.46	8.54	10.87		
	10.39	11.11	10.96	10.83	10.09	8.69	11.17	9.39	9.56		
	10.41	10.82	17.61	10.87	*	13.06	11.41	9.96	9.15		
	9.76	11.10	10.80	8.86	9.48	9.22	9.61	8.20			
	10.93	10.94	9.53	10.61	10.27	9.59	9.82	7.81			
		9.64	9.52	10.06	12.65	9.63					
		9.29	8.75	8.96	8.27	8.14					
		10.59	10.43	9.32							

143

A.4 Cropped field Available Phosphate Data (P-PO4)

6.6	7.5	22.0	18.1	25.6	13.8	30.0	22.1	9.9	8.6	11.1
13.2	9.4	22.1	12.2	14.6	12.2	12.8	8.7	9.8	7.2	2.8
10.9	12.9	24.5	14.9	10.0	20.5	11.6	11.6	5.5	13.4	21.8
10.8	19.6	11.1	11.7	13.9	8.2	7.0	9.2	6.2	7.9	4.2
7.7	28.5	16.1	12.5	8.3	6.3	17.4	13.7	7.0	5.1	4.4
9.7	7.6	13.7	13.8	7.7	14.2	9.8	30.6	5.8	15.9	51.3
5.5	12.1	13.1	14.6	7.0	14.3	17.0	12.1	15.9	7.1	14.7
5.5	12.0	18.3	10.8	1.3	15.8	9.9	14.0	13.6	5.8	9.2
8.6	8.7	18.0	7.1	16.3	8.8	11.4	27.3	15.9	7.0	11.6
7.6	12.7	17.7	9.2	10.4	11.1	17.5	24.8	12.5	8.8	8.7
15.1	19.5	11.9	6.0	21.5	15.2	8.7	7.3	11.7	12.5	13.1

A.5 Cropped filed Potassium Data

40	51	44	44	49	69	41	30	27	37	24
48	43	55	63	45	64	40	35	35	27	30
45	57	54	58	44	74	50	35	36	30	29
32	92	56	44	44	34	26	42	26	28	34
39	61	113	40	31	29	50	25	29	28	26
39	51	47	36	24	41	34	30	30	29	17
37	51	37	38	67	37	52	40	28	24	25
30	99	60	32	32	43	31	33	31	39	23
34	61	68	31	50	31	31	43	32	25	25
22	38	62	27	38	26	42	29	28	26	32
62	41	38	26	41	32	28	24	39	25	30

A.6 Uncropped field Available Phophate Data

4.7	3.9	3.2	7.9	3.5	4.5	4.9	4.1	8.5	7.0	5.8
*	4.9	4.3	9.2	4.0	5.4	4.1	4.9	3.6	6.7	6.0
5.6	5.1	4.1	34.5	5.4	4.5	3.6	4.7	6.0	6.3	4.1
4.0	3.2	6.4	10.4	7.9	6.7	4.9	5.8	4.9	7.0	4.4
6.5	6.1	8.1	7.8	31.0	6.1	7.3	7.3	3.6	7.8	5.0
5.8	5.0	5.6	8.3	7.5	3.9	6.3	5.1	3.9	5.1	3.6
4.3	3.8	4.2	8.5	6.3	4.5	2.6	6.3	4.9	4.0	2.6
5.3	*	*	5.6	3.5	3.9	3.6	3.9	5.6	8.5	2.9
6.1	4.3	4.3	2.8	4.5	2.5	5.1	3.0	5.8	6.6	4.9
4.2	2.2	*	4.0	4.1	5.6	3.9	5.2	5.7	7.1	4.0
3.4	2.9	6.1	3.0	6.1	5.1	3.2	7.3	6.2	6.7	6.2

A.7 Uncropped field Potassium Data

78	56	106	177	57	49	64	42	39	32	49
*	102	73	353	66	33	38	49	34	43	41
97	94	113	89	44	55	57	60	45	31	89
158	68	75	68	75	52	52	57	38	86	56
126	47	106	67	174	92	58	49	29	71	92
101	79	58	154	116	77	52	51	36	149	80
93	83	70	120	138	48	67	48	52	82	151
117	*	*	92	137	73	57	75	56	99	99
72	49	75	59	104	209	87	68	63	105	49
167	93	*	48	97	189	111	89	138	136	129
96	69	46	192	70	290	92	46	139	164	134

A.8 Residuals of the Cropped field Potassium Data (P-PO4)

0	-5	-11	3	4	29	1	-5	-8	5	-6
8	-13	0	22	0	24	0	0	0	-5	0
4	0	-2	16	-2	33	9	-1	0	-3	-2
-7	37	2	4	0	-5	-13	8	-8	-3	5
3	9	62	3	-10	-7	14	-6	-2	0	0
4	0	-3	0	-16	6	-1	0	0	2	-8
0	-2	-15	0	25	0	15	8	-4	-5	-2
-6	47	9	-5	-9	7	-5	2	0	11	-3
-1	10	18	-5	10	-4	-4	13	2	-2	0
-11	-11	14	-7	0	-7	9	1	0	1	9
30	-7	-9	-7	4	0	-4	-3	12	1	8

A.9 Residuals of the uncropped field Potassium Data

-1	0	-1	2	5	26	1	-1	-3	6	-3
8	-7	11	22	2	22	1	5	6	-3	4
3	5	8	15	-1	30	9	3	5	-2	1
-5	45	15	6	4	-5	-10	15	0	1	11
3	15	73	3	-8	-9	15	-1	4	2	4
4	6	8	0	-14	4	0	5	6	4	-4
-1	3	-5	-1	26	-3	15	12	1	-4	1
-6	53	20	-5	-7	5	-4	7	6	13	1
-1	16	29	-5	12	-6	-3	18	8	0	4
-11	-5	25	-7	2	-9	10	6	6	3	13
30	-1	2	-7	6	-2	-3	2	18	3	12

(A.10) These are the Banksia data with residuals

Col 1	R1	D1	Rs1	Col 2	R 2	D2	Rs2
1	18	1	0.005	10	4	9	-0.891
2	17	6	0	10	5	7	0.647
2	18	3	0.02	10	6	6	1.413
3	15	7	2.48	10	7	2	-1.28
3	16	7	0.017	10	8	2	-1.335
3	17	6	0.008	10	9	1	-2.185
3	18	2	-0.97	11	2	19	8.3
4	12	4	0	11	3	17	9.3
4	13	5	2.04	11	4	5	-3.58
4	14	8	6.54	11	5	5	-0.045
4	15	7	4.45	11	6	3	-0.282
4	16	5	-0.017	11	7	2	0.02223
4	17	2	-2.026	11	8	1	-1.032
4	18	1	-0.005	12	2	16	4.25
5	3	8	-5.52	12	3	13	4.25
5	4	7	-7.41	12	4	9	-0.632
5	10	14	4.96	12	5	6	-0.093
5	11	17	8.58	12	6	4	-0.326
5	12	15	5.853	12	7	3	.0223
5	13	8	-0.103	12	9	1	7.928
5	14	7	0.396	13	3	17	3.037
5	15	4	-3.694	13	4	13	2.57
6	2	6	-6.314	13	5	9	0.342
6	3	7	-2.314	13	6	5	-0.352
6	4	7	-3.205	13	7	3	-1.406
6	5	6	-0.666	13	8	2	-1.256
6	6	5	0.099	13	9	2	-3.58
6	7	6	2.404	13	10	1	3.264
6	8	8	4.349	14	4	13	-1.196
6	9	14	10.5	14	5	5	0.57
6	10	15	10.17	14	6	5	1.87
6	11	14	9.79	14	7	5	-0.179
6	12	10	5.06	14	8	3	-0.029
6	13	4	0.103	14	9	3	-1.735

6	14	2	-0.396	14	11	2	3.008
6	15	1	-2.486	15	4	15	-1.452
7	2	7	-4.375	15	5	4	0.313
7	3	4	-4.37	15	6	4	1.618
7	4	7	-2.26	15	7	3	-0.436
7	5	8	2.27	15	8	4	-0.285
7	6	5	1.038	15	9	3	5.29
7	7	6	3.34	16	4	13	-2.16
7	8	9	6.289	16	5	7	-0.4
7	9	7	4.439	16	6	4	-0.095
7	10	4	0.1117	16	7	1	0.8496
7	11	2	-1.266	16	8	3	0
7	12	4	0	16	9	3	-2.02
7	13	1	-1.957	16	10	2	-2.4
7	14	1	-0.457	16	11	1	3.594
8	1	8	-0.054	17	4	11	1.133
8	2	6	-5.05	17	5	8	-0.099
8	3	8	-0.054	17	6	2	-1.795
8	4	6	-2.94	17	7	1	0.15
8	5	6	0.593	17	8	1	0.3007
8	6	3	-0.64	17	9	1	3.29
8	7	5	2.66	18	4	10	3.83
8	8	4	1.609	18	5	8	-0.4
8	9	3	0.76	18	6	2	-0.0959
8	10	4	0.43	18	7	1	-0.15
8	11	3	0.054	19	4	9	0
8	12	1	-2.67	19	5	4	1.16
9	1	7	0	19	6	1	2.7
9	2	10	0	20	4	7	-1.532
9	3	7	0	20	5	2	-1.228
9	4	7	-0.891	20	12	1	0
9	5	4	-0.352	21	4	7	1.538
9	6	4	1.41	21	5	5	0
9	7	6	4.71	21	11	1	-0.733
9	8	4	2.66	21	12	1	-0.49
9	9	1	-0.185	22	4	6	0.0488
9	10	1	-1.51	22	5	3	1.815
9	11	1	-0.891	22	6	3	-0.1117

A.11 The data measuring the gravimetric recognition in Vald' Or

-27	[2]	[4]	[8]	[11]	-21	-22	-10	-24	[13]	-15	-7	-6	-35	-34	-42	-63
[1]	[3]	[5]	[9]	[12]	-12	2	13	39	13	0	-14	-18	-26	-40	-60	-89
-27	-24	[6]	[10]	-4	-6	19	25	26	10	10	-6	-16	-32	-68	-91	-122
-11	-4	[7]	14	16	27	21	14	14	5	1	0	-15	-53	-96	-119	-144
-12	5	14	42	57	18	-4	13	18	20	10	-13	038	-97	-133	-139	
-23	-24	18	50	35	9	9	16	12	9	0	-25	055	-80	-94	-133	-155
-12	-8	13	18	9	9	6	2	7	-5	-26	-47	-65	-91	-128	-147	-159
-10	4	-4	-15	-22	-4	-26	-2	-23	-35	-49	-50	-88	-130	-135	-154	-181
-8	-26	-2	-22	-25	-19	-46	-49	-57	-55	-75	-86	-111	-120	-139	-158	-189
-14	-17	1	-21	-41	-45	-67	-64	-74	-91	-97	-120	-138	-165	-171	-199	
-2	-4	-13	-29	-21	-43	-59	-73	-87	-82	-94	-117	-121	-140	-167	-174	-214

9.2 Appendix B

B.1 This macro was written for estimating robust variogram in the E-W direction by using

MINITAB.

```
LET CK3=SQRT(ABSO(Ck2-CK1))
LET K1=K1+1
LET K2=K2+1
LET K3=K3+1
END
```

This macro was written for estimating robust variogram in the E-W direction by using first macro. This is an EXECUTIVE file.

```
LET K1=1
LET K2=K4+1
LET k3=25
execs 'c:\mtbwin\Rvargm2.MAC' K5
LET K6=25+K5-1
STACK C25-CK6 C99
ERASE C25-CK6
LET C98(K4)=N(C99)
LET C100(K4)=(MEAN(C99)**4)/(.457+.494/c98(k4))
LET K4=K4+1
LET K5=K5-1
END
```

This macro was written in the EW-NS direction by using LAG commands.

```
LAG K4 CK1 C95
LET CK3=SQRT(ABSO(C95-CK2))
LET K1=K1+1
LET K2=K2+1
LET K3=K3+1
END
```

```
execs 'c:\mtbwin\vargm2.MAC' K5
LET K6=30+K5-1
STACK C30-CK6 C99
ERASE C30-CK6
LET C98(K4)=N(C99)
LET C100(K4)=MEAN(C99)
LET K4=K4+1
LET K5=K5-1
END
```

This macro was written in EN-ES direction by using Lag commands.

```
LAG K4 CK2 C95
LET CK3=SQRT(ABSO(CK1-C95))
LET K1=K1+1
LET K2=K2+1
LET K3=K3+1
END
LET K1=1
LET K2=K4+1
LET k3=30
execs 'c:\mtbwin\vargm2.MAC' K5
LET K6=30+K5-1
STACK C30-CK6 C99
ERASE C30-CK6
LET C98(K4)=N(C99)
LET C100(K4)=MEAN(C99)
LET K4=K4+1
LET K5=K5-1
END
```

(B.2) The robust and classical variograms for the original iron ore data.

Lag	Count	RobVar	Clvar
1	103	13.110	14.3956
2	94	14.815	18.4433
3	85	14.669	16.9335
4	77	17.318	18.6804
5	69	26.522	23.6839
6	61	19.199	18.5369
7	53	27.329	24.6851
8	45	41.646	33.3302
9	38	35.996	34.2584
10	31	27.162	30.1542
11	25	37.365	39.6672
12	19	39.888	37.0468
13	13	48.084	39.0292
14	7	53.180	48.7786
15	4	165.287	96.2250

B.3 The classical and robust variograms denoted by $2\overline{\gamma}(h)$ and $2\hat{\gamma}(h)$ respectively for the iron ore residuals.

Row	$2\overline{\gamma}(h)$	$2\hat{\gamma}(h)$	Lag
1	11.4178	11.8100	1
2	11.1922	13.3489	2
3	12.0209	12.7622	3
4	13.7444	13.7885	4
5	18.4904	17.2498	5
6	14.7159	15.0603	6
7	21.6566	21.4137	7
8	19.1106	22.8629	8
9	23.1600	20.4777	9
10	14.4020	13.4749	10
11	19.0415	17.6163	11
12	15.7714	16.5716	12

B.4 This is the semivariogram values obtained by using iron ore data in the E-W.

```
+------------------------------ R E S U L T S ------------------------------+
| Variable:  oree      |              | Direction :     .000 |
| Minimum :       45.400 | Estimator  :  Variogram | Tolerance :    1.000 |
| Maximum :       60.800 | Total Pairs :      720  | BandWidth :      n/a |
|-----------------------------------------------------------------------|
|  Pairs  Avg Distance    Estimate |  Pairs  Avg Distance    Estimate  |
|----------------------------------+------------------------------------|
| 1   103     1.000     7.198  | 13  13      13.000     19.515  |
| 2    94     2.000     9.222  | 14   7      14.000     24.389  |
| 3    85     3.000     8.467  | 15                             |
| 4    77     4.000     9.340  | 16                             |
| 5    69     5.000    11.842  | 17                             |
| 6    61     6.000     9.268  | 18                             |
| 7    53     7.000    12.343  | 19                             |
| 8    45     8.000    16.665  | 20                             |
| 9    38     9.000    17.129  | 21                             |
| 10   31    10.000    15.077  | 22                             |
| 11   25    11.000    19.834  | 23                             |
| 12   19    12.000    18.523  | 24                             |
```

B.5 Spherical model is fitted to the iron ore data obtained by GEOEAS.

	Pairs	Avg Distance	Value			
1	197	1.000	7.003			
2	352	1.699	8.126			
3	615	2.560	9.520			
4	659	3.493	10.761			
5	857	4.524	12.725			
6	697	5.506	14.024			
7	614	6.425	13.998			
8	597	7.383	14.231			
9	474	8.415	14.119			
10	367	9.447	12.836	Nugget :	4.500	
11	261	10.410	14.101			
12	202	11.353	15.133	Type	Sill	Range
13	159	12.353	14.642			
14				Spherical	9.500	6.500

B.6 The semivariogram values were fitted using GEOEAS.

Pairs	Avg Distance	Value	Pairs	Avg Distance	Value
1 197	1.000	7.003	18		
2 352	1.699	8.126	19		
3 615	2.560	9.520	20		
4 659	3.493	10.761	21		
5 857	4.524	12.725	22		
6 697	5.506	14.024	23		
7 614	6.425	13.998	24		
8 597	7.383	14.231			
9 474	8.415	14.119	Model		
10 367	9.447	12.836		Nugget :	6.000
11 261	10.410	14.101			
12 202	11.353	15.133	Type	Sill	Range
13 159	12.353	14.642			
14			Gaussian	8.500	6.000

B.7

Variable: Ore			Direction :	.000
Minimum :	-8.207	Estimator : Variogram	Tolerance :	20.000
Maximum :	6.129	Total Pairs : 2380	BandWidth :	10.000

Pairs	Avg Distance	Estimate	Pairs	Avg Distance	Estimate
1 99	1.000	5.863	13 116	12.327	6.710
2 92	2.000	5.932	14 77	13.295	5.912
3 83	3.000	5.976	15 42	14.322	7.294
4 228	3.438	6.607	16 16	15.284	8.519
5 207	4.407	7.654	17 4	16.208	13.451
6 185	5.386	6.953	18		
7 262	6.353	7.842	19		
8 228	7.326	7.344	20		
9 194	8.311	7.469	21		
10 215	9.336	5.758	22		
11 173	10.316	5.895	23		
12 159	11.353	6.151	24		

B.8

```
+----------------------------- R E S U L T S -----------------------------+
| Variable: Ore          |              | Direction :    20.000 |
| Minimum :      -8.207 | Estimator  : Variogram | Tolerance :    20.000 |
| Maximum :       6.129 | Total Pairs :     2124 | BandWidth :    10.000 |
|------------------------------------------------------------------------|
|    Pairs  Avg Distance    Estimate |  Pairs  Avg Distance    Estimate |
|----------------------------------+-------------------------------------|
| 1    99       1.000       5.863 | 13   70      12.401      5.881  |
| 2    92       2.000       5.932 | 14   44      13.377      5.141  |
| 3   165       2.620       5.854 | 15   26      14.355      5.199  |
| 4   214       3.588       6.728 | 16   12      15.359      5.471  |
| 5   240       4.624       8.149 | 17    3      16.267     10.798  |
| 6   213       5.563       7.375 | 18                              |
| 7   217       6.499       7.973 | 19                              |
| 8   202       7.477       7.335 | 20                              |
| 9   185       8.473       6.177 | 21                              |
|10   141       9.489       5.463 | 22                              |
|11   113      10.449       6.127 | 23                              |
|12    88      11.409       5.373 | 24                              |
```

B.9

```
+----------------------------- R E S U L T S -----------------------------+
| Variable: Ore          |              | Direction :    40.000 |
| Minimum :      -8.207 | Estimator  : Variogram | Tolerance :    20.000 |
| Maximum :       6.129 | Total Pairs :     1108 | BandWidth :    10.000 |
|------------------------------------------------------------------------|
|    Pairs  Avg Distance    Estimate |  Pairs  Avg Distance    Estimate |
|----------------------------------+-------------------------------------|
| 1                        |13    8      12.427     10.413  |
| 2    88       1.414       6.781 | 14    2      13.928     17.704  |
| 3   151       2.507       6.469 | 15                              |
| 4   121       3.606       7.113 | 16                              |
| 5   198       4.646       7.325 | 17                              |
| 6   161       5.647       6.204 | 18                              |
| 7    95       6.519       5.103 | 19                              |
| 8   125       7.431       5.410 | 20                              |
| 9    91       8.512       5.290 | 21                              |
|10    39       9.605       9.479 | 22                              |
|11    23      10.589       8.011 | 23                              |
|12     6      11.180      10.026 | 24                              |
```

153

B.10

RESULTS							
Variable: Ore			Direction : 60.000				
Minimum :	-8.207	Estimator : Variogram	Tolerance : 20.000				
Maximum :	6.129	Total Pairs : 809	BandWidth : 10.000				
Pairs	Avg Distance	Estimate	Pairs	Avg Distance	Estimate		
1			13				
2	88	1.414	6.781	14			
3	142	2.524	6.741	15			
4	116	3.380	7.756	16			
5	184	4.438	6.627	17			
6	132	5.491	5.474	18			
7	61	6.460	4.691	19			
8	61	7.317	5.704	20			
9	16	8.401	9.588	21			
10	9	9.295	15.289	22			

B.11

+------ RESULTS ------+							
Variable: Ore			Direction : 80.000				
Minimum :	-8.207	Estimator : Variogram	Tolerance : 20.000				
Maximum :	6.129	Total Pairs : 738	BandWidth : 10.000				
Pairs	Avg Distance	Estimate	Pairs	Avg Distance	Estimate		
1	92	1.000	6.592	13			
2	75	2.000	6.848	14			
3	135	2.587	6.457	15			
4	107	3.538	8.650	16			
5	124	4.483	7.009	17			
6	85	5.419	6.940	18			
7	83	6.366	4.579	19			
8	31	7.236	7.702	20			
9	6	8.062	6.887	21			

B.13

```
+----------------------------- R E S U L T S -----------------------------+
! Variable: Ore          !              ! Direction :   100.000 !
! Minimum :       -8.207 ! Estimator  : Variogram ! Tolerance :   20.000 !
! Maximum :        6.129 ! Total Pairs :       768 ! BandWidth :   10.000 !
!-------------------------------------------------------------------------!
!   Pairs Avg Distance    Estimate ! Pairs Avg Distance    Estimate !
!-------------------------------------+-----------------------------------!
! 1    92      1.000      6.592 ! 13                       !
! 2    75      2.000      6.848 ! 14                       !
! 3   136      2.584      7.064 ! 15                       !
! 4   108      3.535      7.481 ! 16                       !
! 5   123      4.483      8.393 ! 17                       !
! 6    83      5.423      9.016 ! 18                       !
! 7    88      6.377      6.117 ! 19                       !
! 8    41      7.296      5.880 ! 20                       !
! 9    22      8.273      6.880 ! 21                       !
!10                       ! 22                       !
!11                       ! 23                       !
!12                       ! 24                       !
```

B.14

```
+----------------------------- R E S U L T S -----------------------------+
! Variable: Ore          !              ! Direction :   120.000 !
! Minimum :       -8.207 ! Estimator  : Variogram ! Tolerance :   20.000 !
! Maximum :        6.129 ! Total Pairs :       928 ! BandWidth :   10.000 !
!-------------------------------------------------------------------------!
!   Pairs Avg Distance    Estimate ! Pairs Avg Distance    Estimate !
!-------------------------------------+-----------------------------------!
! 1                       ! 13    1      12.042      12.236 !
! 2    88      1.414      5.331 ! 14                       !
! 3   145      2.526      7.042 ! 15                       !
! 4   117      3.378      7.712 ! 16                       !
! 5   186      4.438      8.479 ! 17                       !
! 6   133      5.499      8.810 ! 18                       !
! 7    68      6.465      8.571 ! 19                       !
! 8    88      7.348      8.144 ! 20                       !
! 9    48      8.430      8.046 ! 21                       !
!10    42      9.436      8.331 ! 22                       !
!11    10     10.630      7.451 ! 23                       !
!12     2     11.314      1.831 ! 24                       !
```

B.15

+-------------- R E S U L T S ----------------+						
Variable: Ore			Direction :	140.000		
Minimum :	-8.207	Estimator : Variogram	Tolerance :	20.000		
Maximum :	6.129	Total Pairs : 1401	BandWidth :	10.000		

	Pairs	Avg Distance	Estimate		Pairs	Avg Distance	Estimate
1				13	27	12.432	8.067
2	88	1.414	5.331	14	12	13.606	13.640
3	155	2.507	5.808	15	4	14.430	15.066
4	124	3.606	7.301	16	1	15.232	3.876
5	210	4.647	7.505	17			
6	175	5.643	7.199	18			
7	111	6.524	7.754	19			
8	166	7.441	7.100	20			
9	146	8.519	6.473	21			
10	99	9.607	6.778	22			
11	55	10.607	6.186	23			
12	28	11.367	7.679	24			

B.16

+---------------- R E S U L T S ----------------+						
Variable: Ore			Direction :	160.000		
Minimum :	-8.207	Estimator : Variogram	Tolerance :	20.000		
Maximum :	6.129	Total Pairs : 2438	BandWidth :	10.000		

	Pairs	Avg Distance	Estimate		Pairs	Avg Distance	Estimate
1	99	1.000	5.863	13	93	12.408	8.557
2	92	2.000	5.932	14	54	13.412	9.227
3	167	2.616	5.476	15	24	14.416	9.934
4	220	3.583	6.200	16	7	15.352	11.805
5	253	4.619	6.788	17	1	16.031	21.410
6	230	5.552	6.459	18			
7	246	6.487	7.351	19			
8	245	7.461	6.675	20			
9	243	8.459	6.885	21			
10	191	9.503	6.346	22			
11	152	10.452	6.495	23			
12	121	11.404	7.523	24			

156

B.17

```
-----------------------------  R E S U L T S -----------------------------+       |
| Variable: LN (PO4)       |                  | Direction :    .000 |
| Minimum :        .262 | Estimator  :  Variogram | Tolerance :    90.000 |
| Maximum :       3.938 | Total Pairs :      7258 | BandWidth :      n/a |
|------------------------------------------------------------------------|
|   Pairs  Avg Distance    Estimate |  Pairs  Avg Distance    Estimate |
|----------------------------------+-------------------------------------|
| 1   220      1.000      .220 | 13  60      12.330      .188 |
| 2   398      1.706      .241 | 14   8      13.454      .276 |
| 3   698      2.566      .256 | 15                         |
| 4   762      3.499      .268 | 16                         |
| 5  1016      4.532      .262 | 17                         |
| 6   856      5.515      .259 | 18                         |
| 7   796      6.432      .261 | 19                         |
| 8   830      7.396      .269 | 20                         |
| 9   710      8.421      .241 | 21                         |
|10   538      9.481      .236 | 22                         |
|11   272     10.448      .291 | 23                         |
|12    94     11.384      .222 | 24
```

B.18

```
+-----------------------------  R E S U L T S -----------------------------+
| Variable: LN (PO4)       |                  | Direction :    .000 |
| Minimum :        .262 | Estimator  :  Variogram | Tolerance :    20.000 |
| Maximum :       3.938 | Total Pairs :      1643 | BandWidth :    10.000 |
|------------------------------------------------------------------------|
|   Pairs  Avg Distance    Estimate |  Pairs  Avg Distance    Estimate |
|----------------------------------+-------------------------------------|
| 1   110      1.000      .217 | 13                         |
| 2    99      2.000      .266 | 14                         |
| 3    88      3.000      .246 | 15                         |
| 4   237      3.434      .255 | 16                         |
| 5   206      4.404      .243 | 17                         |
| 6   175      5.382      .272 | 18                         |
| 7   234      6.348      .302 | 19                         |
| 8   185      7.318      .318 | 20                         |
| 9   136      8.287      .284 | 21                         |
|10   119      9.308      .277 | 22                         |
|11    54     10.215      .411 | 23                         |
|12                       | 24
```

157

B.19

```
+----------------------------- R E S U L T S -----------------------------+
| Variable: LN (PO4)      |                  | Direction :   45.000 |
| Minimum :        .262 | Estimator  : Variogram | Tolerance :   20.000 |
| Maximum :       3.938 | Total Pairs :     1720 | BandWidth :   10.000 |
```

Pairs	Avg Distance	Estimate	Pairs	Avg Distance	Estimate
1			13	30 12.330	.075
2 100	1.414	.248	14		
3 261	2.420	.268	15		
4 144	3.606	.288	16		
5 302	4.619	.290	17		
6 145	5.772	.226	18		
7 164	6.552	.226	19		
8 166	7.397	.181	20		
9 171	8.492	.196	21		
10 122	9.564	.189	22		
11 68	10.567	.214	23		
12 47	11.384	.177	24		

B.20

```
+----------------------------- R E S U L T S -----------------------------+
| Variable: LN (PO4)      |                  | Direction :   90.000 |
| Minimum :        .262 | Estimator  : Variogram | Tolerance :   20.000 |
| Maximum :       3.938 | Total Pairs :     1643 | BandWidth :   10.000 |
```

Pairs	Avg Distance	Estimate	Pairs	Avg Distance	Estimate
1 110	1.000	.224	13		
2 99	2.000	.237	14		
3 88	3.000	.234	15		
4 237	3.434	.253	16		
5 206	4.404	.241	17		
6 175	5.382	.246	18		
7 234	6.348	.241	19		
8 185	7.318	.253	20		
9 136	8.287	.180	21		
10 119	9.308	.192	22		
11 54	10.215	.215	23		
12			24		

B.21

	RESULTS			
Variable: LN (PO4)			Direction : 135.000	
Minimum : .262	Estimator :	Variogram	Tolerance : 20.000	
Maximum : 3.938	Total Pairs :	1690	BandWidth : 10.000	
Pairs	Avg Distance	Estimate	Pairs Avg Distance	Estimate
1			13	
2 100	1.414	.211	14	
3 261	2.420	.255	15	
4 144	3.606	.296	16	
5 302	4.619	.260	17	
6 145	5.772	.255	18	
7 164	6.552	.269	19	
8 166	7.397	.310	20	
9 171	8.492	.290	21	
10 122	9.564	.294	22	
11 68	10.567	.309	23	
12 47	11.384	.267	24	

B.22

	RESULTS			
Variable: RsPot			Direction : .000	
Minimum : -17.000	Estimator :	Variogram	Tolerance : 90.000	
Maximum : 60.000	Total Pairs :	7096	BandWidth : n/a	
Pairs	Avg Distance	Estimate	Pairs Avg Distance	Estimate
1 220	1.000	138.300	13	
2 398	1.706	147.085	14	
3 698	2.566	143.897	15	
4 762	3.499	140.785	16	
5 1016	4.532	139.481	17	
6 856	5.515	143.224	18	
7 796	6.432	138.413	19	
8 830	7.396	135.039	20	
9 710	8.421	134.411	21	
10 538	9.481	118.703	22	
11 272	10.448	90.588	23	

B.23

```
+----------------------------- R E S U L T S -----------------------------+
| Variable: RsPot      |                    | Direction :    .000 |
| Minimum :     -17.000 | Estimator  : Variogram | Tolerance :   20.000 |
| Maximum :      60.000 | Total Pairs :     1643 | BandWidth :   10.000 |
|-------------------------------------------------------------------------|
|  Pairs  Avg Distance    Estimate |  Pairs  Avg Distance    Estimate |
|--------------------------------+--------------------------------------|
| 1   110      1.000      148.382 | 13                               |
| 2    99      2.000      148.460 | 14                               |
| 3    88      3.000      144.773 | 15                               |
| 4   237      3.434      135.426 | 16                               |
| 5   206      4.404      158.998 | 17                               |
| 6   175      5.382      129.749 | 18                               |
| 7   234      6.348      138.442 | 19                               |
| 8   185      7.318      141.465 | 20                               |
| 9   136      8.287      168.173 | 21                               |
|10   119      9.308      134.084 | 22                               |
|11    54     10.215       50.269 | 23                               |
```

B.24

```
+----------------------------- R E S U L T S -----------------------------+
| Variable: RsPot      |                    | Direction :   45.000 |
| Minimum :     -17.000 | Estimator  : Variogram | Tolerance :   20.000 |
| Maximum :      60.000 | Total Pairs :     1643 | BandWidth :   10.000 |
|-------------------------------------------------------------------------|
|  Pairs  Avg Distance    Estimate |  Pairs  Avg Distance    Estimate |
|--------------------------------+--------------------------------------|
| 1                      | 13                               |
| 2   100      1.414      123.665 | 14                               |
| 3   261      2.420      144.490 | 15                               |
| 4   144      3.606      149.149 | 16                               |
| 5   302      4.619      135.012 | 17                               |
| 6   145      5.772      143.834 | 18                               |
| 7   164      6.552      137.348 | 19                               |
| 8   166      7.397      156.840 | 20                               |
| 9   171      8.492      124.673 | 21                               |
|10   122      9.564      112.783 | 22                               |
|11    68     10.567       62.949 | 23                               |
|12                      | 24                               |
```

B.25

```
+----------------------------- R E S U L T S -----------------------------+
| Variable: RsPot      |              | Direction :   90.000 |
| Minimum :      -17.000 | Estimator  : Variogram | Tolerance :   20.000 |
| Maximum :       60.000 | Total Pairs :    1643 | BandWidth :   10.000 |
|------------------------------------------------------------------------|
|  Pairs  Avg Distance    Estimate | Pairs  Avg Distance    Estimate |
|------------------------------+----------------------------------|
| 1  110      1.000     128.218 | 13                        |
| 2   99      2.000     159.237 | 14                        |
| 3   88      3.000     159.818 | 15                        |
| 4  237      3.434     142.011 | 16                        |
| 5  206      4.404     134.937 | 17                        |
| 6  175      5.382     154.843 | 18                        |
| 7  234      6.348     136.122 | 19                        |
| 8  185      7.318     120.797 | 20                        |
| 9  136      8.287     105.985 | 21                        |
| 10  119      9.308     101.269 | 22                        |
| 11   54     10.215     121.167 | 23                        |
| 12                    | 24                        |
```

B.26

```
+----------------------------- R E S U L T S -----------------------------+
| Variable: RsPot      |              | Direction :   135.000 |
| Minimum :      -17.000 | Estimator  : Variogram | Tolerance :   20.000 |
| Maximum :       60.000 | Total Pairs :    1643 | BandWidth :   10.000 |
|------------------------------------------------------------------------|
|  Pairs  Avg Distance    Estimate | Pairs  Avg Distance    Estimate |
|------------------------------+----------------------------------|
| 1                     | 13                        |
| 2  100      1.414     157.115 | 14                        |
| 3  261      2.420     137.640 | 15                        |
| 4  144      3.606     139.226 | 16                        |
| 5  302      4.619     133.737 | 17                        |
| 6  145      5.772     138.034 | 18                        |
| 7  164      6.552     142.704 | 19                        |
| 8  166      7.397     127.852 | 20                        |
| 9  171      8.492     135.629 | 21                        |
| 10  122      9.564     127.287 | 22                        |
| 11   68     10.567     128.382 | 23                        |
| 12                    | 24                        |
```

B.27

+------------------------------- R E S U L T S ------------------------------+

Variable: SPECIES	Direction : .000
Minimum : -7.410	Estimator : Variogram	Tolerance : 20.000
Maximum : 10.500	Total Pairs : 2438	BandWidth : n/a

	Pairs	Avg Distance	Estimate		Pairs	Avg Distance	Estimate
1	121	1.000	4.321	13	167	12.372	9.990
2	101	2.000	7.635	14			
3	84	3.000	10.415	15			
4	228	3.416	10.425	16			
5	193	4.396	10.262	17			
6	171	5.394	9.834	18			
7	258	6.359	8.769	19			
8	238	7.328	10.289	20			
9	214	8.299	9.991	21			
10	252	9.337	9.930	22			
11	203	10.317	12.347	23			
12	208	11.385	11.033	24			

B.28

+------------------------------- R E S U L T S ------------------------------+

Variable: SPECIES	Direction : 45.000
Minimum : -7.410	Estimator : Variogram	Tolerance : 20.000
Maximum : 10.500	Total Pairs : 1634	BandWidth : n/a

	Pairs	Avg Distance	Estimate		Pairs	Avg Distance	Estimate
1				13	50	12.397	8.938
2	105	1.414	7.539	14			
3	248	2.413	10.199	15			
4	123	3.606	9.817	16			
5	248	4.622	9.367	17			
6	128	5.774	9.277	18			
7	152	6.556	9.143	19			
8	155	7.398	8.636	20			
9	166	8.494	8.814	21			
10	126	9.572	7.215	22			
11	71	10.560	6.328	23			
12	62	11.385	6.155	24			

B.29

```
+----------------------------- R E S U L T S -----------------------------+
| Variable: SPECIES    |            | Direction :   90.000 |
| Minimum :      -7.410 | Estimator  : Variogram | Tolerance :   20.000 |
| Maximum :      10.500 | Total Pairs :    1500  | BandWidth :     n/a  |
|-----------------------------------------------------------------------|
|   Pairs  Avg Distance    Estimate |  Pairs  Avg Distance    Estimate  |
|--------------------------------+--------------------------------------|
| 1   122      1.000       4.237 | 13                     |
| 2    99      2.000       7.814 | 14                     |
| 3    80      3.000      11.529 | 15                     |
| 4   222      3.408      11.546 | 16                     |
| 5   178      4.379      14.634 | 17                     |
| 6   143      5.357      17.751 | 18                     |
| 7   197      6.342      18.369 | 19                     |
| 8   153      7.304      19.652 | 20                     |
| 9   111      8.275      20.013 | 21                     |
| 10  112      9.340      18.229 | 22                     |
| 11   83     10.308      15.325 | 23                     |
| 12                             | 24                     |
```

B.30

```
+----------------------------- R E S U L T S -----------------------------+
| Variable: SPECIES    |            | Direction :  135.000 |
| Minimum :      -7.410 | Estimator  : Variogram | Tolerance :   20.000 |
| Maximum :      10.500 | Total Pairs :    2164  | BandWidth :     n/a  |
|-----------------------------------------------------------------------|
|   Pairs  Avg Distance    Estimate |  Pairs  Avg Distance    Estimate  |
|--------------------------------+--------------------------------------|
| 1                              | 13                     |
| 2   118      1.414       5.818 | 14                     |
| 3   300      2.424       7.205 | 15                     |
| 4   169      3.606       7.818 | 16                     |
| 5   354      4.624       9.138 | 17                     |
| 6   176      5.772       9.624 | 18                     |
| 7   199      6.549       9.975 | 19                     |
| 8   218      7.404       9.425 | 20                     |
| 9   253      8.516      10.216 | 21                     |
| 10  211      9.582       7.983 | 22                     |
| 11  166     10.569       7.963 | 23                     |
| 12                             | 24                     |
```

B.31

```
+------------------------------ R E S U L T S ------------------------------+
| Variable:  JPhos        |              | Direction :    .000 |
| Minimum :        2.200 | Estimator  : Variogram | Tolerance :   90.000 |
| Maximum :       34.500 | Total Pairs :     5893 | BandWidth :     n/a |
|----------------------------------------------------------------------|
|   Pairs  Avg Distance    Estimate | Pairs  Avg Distance    Estimate |
|--------------------------------+------------------------------------|
| 1  206      1.000      14.535 | 13                          |
| 2  373      1.708      17.201 | 14                          |
| 3  657      2.567      19.392 | 15                          |
| 4  720      3.500      19.459 | 16                          |
| 5  965      4.532      19.677 | 17                          |
| 6  807      5.515      16.169 | 18                          |
| 7  745      6.431      18.777 | 19                          |
| 8  767      7.395      14.853 | 20                          |
| 9  653      8.421      10.983 | 21                          |
```

B.32

```
+------------------------------ R E S U L T S ------------------------------+
| Variable:  JPhos        |              | Direction :    .000 |
| Minimum :        2.200 | Estimator  : Variogram | Tolerance :   20.000 |
| Maximum :       34.500 | Total Pairs :     1366 | BandWidth :     n/a |
|----------------------------------------------------------------------|
|   Pairs  Avg Distance    Estimate | Pairs  Avg Distance    Estimate |
|--------------------------------+------------------------------------|
| 1  104      1.000      16.401 | 13                          |
| 2   93      2.000      17.782 | 14                          |
| 3   84      3.000      20.300 | 15                          |
| 4  225      3.434      19.619 | 16                          |
| 5  194      4.403      17.282 | 17                          |
| 6  163      5.381      16.200 | 18                          |
| 7  216      6.347      18.520 | 19                          |
| 8  167      7.315      13.304 | 20                          |
| 9  120      8.292       2.512 | 21                          |
```

B.33

```
+------------------------------ R E S U L T S ------------------------------+
| Variable:  JPhos        |              | Direction :   45.000 |
| Minimum :        2.200 | Estimator  : Variogram | Tolerance :   20.000 |
| Maximum :       34.500 | Total Pairs :     1386 | BandWidth :     n/a |
|----------------------------------------------------------------------|
|   Pairs  Avg Distance    Estimate | Pairs  Avg Distance    Estimate |
|--------------------------------+------------------------------------|
| 1                             | 13                          |
| 2   93      1.414      17.266 | 14                          |
| 3  245      2.420      18.651 | 15                          |
| 4  136      3.606      22.665 | 16                          |
| 5  289      4.621      21.680 | 17                          |
| 6  140      5.771      21.307 | 18                          |
| 7  158      6.550      22.049 | 19                          |
| 8  161      7.397      25.267 | 20                          |
```

B.34

```
+----------------------------- R E S U L T S ----------------------------+
| Variable: JPhos     |                    | Direction :    90.000 |
| Minimum :        2.200 | Estimator  : Variogram | Tolerance :   20.000 |
| Maximum :       34.500 | Total Pairs :      1377 | BandWidth :      n/a |
|                                                                         |
|    Pairs  Avg Distance     Estimate                                     |
|                                                                         |
| 1   102      1.000       12.632                                         |
| 2    94      2.000       16.299                                         |
| 3    82      3.000       15.325                                         |
| 4   222      3.438       16.904                                         |
| 5   195      4.402       18.563                                         |
| 6   164      5.379       15.069                                         |
| 7   219      6.344       19.931                                         |
| 8   170      7.322       13.710                                         |
```

B.35

```
+----------------------------- R E S U L T S ----------------------------+
| Variable: JPhos     |                    | Direction :   135.000 |
| Minimum :        2.200 | Estimator  : Variogram | Tolerance :   20.000 |
| Maximum :       34.500 | Total Pairs :      1354 | BandWidth :      n/a |
|-----------------------------------------------------------------------|
|    Pairs  Avg Distance     Estimate                                     |
|----------------------- ---- --+--------------------------------------|
| | 2    93      1.414       17.468                                       |
| 3   246      2.421       21.176                                         |
| 4   137      3.606       20.154                                         |
| 5   287      4.618       20.034                                         |
| 6   136      5.772       15.098                                         |
| 7   152      6.552       14.081                                         |
| 8   151      7.395        6.695                                         |
| 9   152      8.493        2.193                                         |
```

B.36

```
+--------------------------- R E S U L T S ---------------------------+
| Variable:  RsPot     |                 | Direction :    .000 |
| Minimum :    -96.000 | Estimator  :  Variogram | Tolerance :   90.000 |
| Maximum :    278.000 | Total Pairs :    6637 | BandWidth :     n/a |
|---------------------------------------------------------------------|
|   Pairs  Avg Distance     Estimate |
```

	Pairs	Avg Distance	Estimate
1	206	1.000	1809.686
2	373	1.708	1917.988
3	657	2.567	1996.186
4	720	3.500	1822.661
5	965	4.532	1869.376
6	807	5.515	2138.142
7	745	6.431	2022.573
8	767	7.395	2269.663
9	653	8.421	2178.811
10	491	9.481	2482.878
11	253	10.443	2157.432

B.37

```
+----------------------------- R E S U L T S ---------------------------+
| Variable:  RsPot     |                 | Direction :    .000 |
| Minimum :    -96.000 | Estimator  :  Variogram | Tolerance :   20.000 |
| Maximum :    278.000 | Total Pairs :    1524 | BandWidth :     n/a |
|-----------------------------------------------------------------------|
|   Pairs  Avg Distance     Estimate |  Pairs  Avg Distance     Estimate |
|-----------------------------------+-----------------------------------|
```

	Pairs	Avg Distance	Estimate		Pairs	Avg Distance	Estimate
1	104	1.000	2251.865	13			
2	93	2.000	2328.331	14			
3	84	3.000	2202.132	15			
4	225	3.434	2101.011	16			
5	194	4.403	2148.984	17			
6	163	5.381	1999.707	18			
7	216	6.347	1776.424	19			
8	167	7.315	1734.729	20			
9	120	8.292	917.217	21			

B.38

```
+---------------------------- R E S U L T S ----------------------------+
| Variable: RsPot    |                    | Direction :   45.000 |
| Minimum :    -96.000 | Estimator  :  Variogram | Tolerance :   20.000 |
| Maximum :    278.000 | Total Pairs :    1563 | BandWidth :    n/a |
|----------------------------------------------------------------------|
|   Pairs  Avg Distance    Estimate |  Pairs  Avg Distance    Estimate |
|----------------------------------+-----------------------------------|
| 1                       | 13                     |
| 2   93    1.414    1930.049 | 14                     |
| 3  245    2.420    2110.486 | 15                     |
| 4  136    3.606    1756.871 | 16                     |
| 5  289    4.621    1802.770 | 17                     |
| 6  140    5.771    2480.683 | 18                     |
| 7  158    6.550    2502.025 | 19                     |
| 8  161    7.397    2454.291 | 20                     |
| 9  164    8.491    2549.628 | 21                     |
|10  115    9.562    2835.673 | 22                     |
|11   62   10.566    2310.900 | 23                     |
```

B.39

```
+---------------------------- R E S U L T S ----------------------------+
| Variable: RsPot    |                    | Direction :   90.000 |
| Minimum :    -96.000 | Estimator  :  Variogram | Tolerance :   20.000 |
| Maximum :    278.000 | Total Pairs :    1542 | BandWidth :    n/a |
|----------------------------------------------------------------------|
|   Pairs  Avg Distance    Estimate |  Pairs  Avg Distance    Estimate |
|----------------------------------+-----------------------------------|
| 1  102    1.000    1358.838 | 13                     |
| 2   94    2.000    1419.922 | 14                     |
| 3   82    3.000    1984.624 | 15                     |
| 4  222    3.438    1590.433 | 16                     |
| 5  195    4.402    1772.581 | 17                     |
| 6  164    5.379    2274.144 | 18                     |
| 7  219    6.344    2177.673 | 19                     |
| 8  170    7.322    3091.332 | 20                     |
| 9  129    8.282    3409.825 | 21                     |
```

B.40

Variable: RsPot			Direction : 135.000
Minimum : -96.000	Estimator : Variogram	Tolerance : 20.000
Maximum : 278.000	Total Pairs : 1519	BandWidth : n/a

Pairs	Avg Distance	Estimate	Pairs	Avg Distance	Estimate
2 93	1.414	1999.014	14		
3 246	2.421	1815.897	15		
4 137	3.606	1807.170	16		
5 287	4.618	1813.238	17		
6 136	5.772	1849.884	18		
7 152	6.552	1650.552	19		
8 151	7.395	1607.547	20		
9 152	8.493	1675.347	21		
10 105	9.562	1241.692	22		
11 60	10.563	1297.836	23		

9.3. Appendix C

C.1

Point kriged is :(6.000, 1.000)					
X(i)	Y(i)	V(i)	Distance	W(i)	B2(i)
6.00	2.00	47.8	1.00	.265	7.32
5.00	2.00	51.3	1.41	.191	6.45
7.00	2.00	52.6	1.41	.191	6.45
6.00	3.00	51.9	2.00	.313E-01	5.25
8.00	2.00	50.5	2.24	.102	4.79
7.00	3.00	55.0	2.24	.194E-01	4.79
4.00	2.00	54.0	2.24	.102	4.79
5.00	3.00	47.5	2.24	.194E-01	4.79
8.00	3.00	48.0	2.83	.237E-02	3.69
4.00	3.00	55.8	2.83	.231E-02	3.69
6.00	4.00	55.7	3.00	-.146E-01	3.39
9.00	2.00	51.2	3.16	.607E-01	3.11
5.00	4.00	53.5	3.16	-.163E-01	3.11
3.00	2.00	50.4	3.16	.606E-01	3.11
7.00	4.00	52.2	3.16	-.163E-01	3.11
3.00	3.00	51.4	3.61	.214E-03	2.41

Neighbors: 16 _ Weights: 1.000
Estimate: 50.7 CVV: 14.000
_k: 2.99 _ WiB2i: 5.976
_²k: 8.96 μ: -.936

C.2

Point kriged is :(4.556, 3.667)					
X(i)	Y(i)	V(i)	Distance	W(i)	B2(i)
5.00	4.00	.313	.555	.143	2.46
4.00	4.00	-2.15	.648	.136	2.42
5.00	3.00	-4.43	.801	.980E-01	2.13
4.00	3.00	2.81	.868	.963E-01	2.12
5.00	5.00	-.731	1.41	.371E-01	1.46
4.00	5.00	-.698	1.44	.353E-01	1.44
6.00	4.00	1.77	1.48	.940E-01	2.04
6.00	3.00	-.769	1.59	.728E-01	1.82
3.00	4.00	-.513	1.59	.844E-01	1.97
3.00	3.00	-1.35	1.69	.692E-01	1.80
5.00	2.00	1.06	1.73	.189E-01	1.13
4.00	2.00	2.69	1.76	.188E-01	1.13
6.00	5.00	-5.77	1.97	.335E-01	1.30
3.00	5.00	.343	2.05	.290E-01	1.24
6.00	2.00	-3.19	2.21	.173E-01	.989
3.00	2.00	-.669	2.28	.168E-01	.999

Neighbors: 16 _ Weights: 1.000
Estimate: -.670 CVV: 2.551
_k: .804 _ WiB2i: 1.969
_²k: .647 μ: -.065

C.3

X(i)	Y(i)	V(i)	Distance	W(i)	B2(i)
, Point kriged is :(6.666, 3.333) for phosphorus data					
7.00	3.00	2.22	.472	.134	.639E-01
6.00	3.00	1.95	.745	.108	.565E-01
7.00	4.00	2.62	.746	.108	.565E-01
6.00	4.00	2.86	.943	.929E-01	.513E-01
7.00	2.00	2.45	1.37	.629E-01	.398E-01
8.00	3.00	1.82	1.37	.656E-01	.398E-01
6.00	2.00	2.45	1.49	.566E-01	.369E-01
8.00	4.00	1.95	1.49	.591E-01	.369E-01
5.00	3.00	2.10	1.70	.474E-01	.318E-01
7.00	5.00	3.42	1.70	.475E-01	.317E-01
5.00	4.00	1.84	1.79	.450E-01	.295E-01
6.00	5.00	2.28	1.80	.448E-01	.295E-01
8.00	2.00	1.70	1.89	.413E-01	.274E-01
5.00	2.00	3.02	2.13	.309E-01	.220E-01
8.00	5.00	1.76	2.14	.318E-01	.220E-01
7.00	1.00	2.16	2.36	.243E-01	.175E-01
Neighbors: 16 _ Weights: 1.000					
Estimate: 2.30 CVV: .066					
_k: .167 _ WiB2i: .044					
_²k: .279E-01 μ: -.005					

C.4

Pairs	Avg Distance	Value for Cropped Phosphate		
1 206	1.000	.119		
2 373	1.708	.145		
3 657	2.567	.172		
4 720	3.500	.182		
5 965	4.532	.190		
6 807	5.515	.165		
7 745	6.431	.180		
8 767	7.395	.155 +----------------------------------		
9 653	8.421	.135	Model	
10			Nugget :	.080
11				
12			Type Sill	Range
13				
14			Spherical .110	4.000

C.5

	Point kriged is :(6.666, 3.333)				
X(i)	Y(i)	V(i)	Distance	W(i)	B2(i)
7.00	3.00	1.76	.472	.207	.751E-01
6.00	3.00	1.59	.745	.150	.664E-01
7.00	4.00	1.99	.746	.150	.664E-01
6.00	4.00	1.99	.943	.118	.602E-01
7.00	2.00	1.55	1.37	.609E-01	.468E-01
8.00	3.00	1.59	1.37	.605E-01	.468E-01
6.00	2.00	1.28	1.49	.508E-01	.434E-01
8.00	4.00	1.28	1.49	.504E-01	.433E-01
5.00	3.00	1.90	1.70	.327E-01	.373E-01
7.00	5.00	1.63	1.70	.327E-01	.373E-01
5.00	4.00	1.81	1.79	.266E-01	.347E-01
6.00	5.00	1.84	1.80	.266E-01	.347E-01
8.00	2.00	1.79	1.89	.208E-01	.322E-01
5.00	2.00	1.50	2.13	.742E-02	.258E-01
8.00	5.00	1.36	2.14	.729E-02	.258E-01
7.00	1.00	1.59	2.36	-.133E-02	.206E-01

Neighbors:	16	_ Weights:	1.000	
Estimate:	1.72	CVV:	.077	
_k:	.139	_ WiB2i:	.058	
_²k:	.194E-01	μ:	.000	

C.6

Pairs	Avg	Distance	Value	for residuals of the Potassium			
1	220	1.000	138.300				
2	398	1.706	147.085				
3	698	2.566	143.897				
4	762	3.499	140.785				
5	1016	4.532	139.481				
6	856	5.515	143.224				
7	796	6.432	138.413				
8	830	7.396	135.039 +-------------------------------------				
9	710	8.421	134.411	Model			
10	538	9.481	118.703		Nugget :	116.000	
11	272	10.448	90.588				
12				Type	Sill	Range	
13							
14				Spherical	30.000	2.000	

Point kriged is :(	2.111,	7.666)			
X(i)	Y(i)	V(i)	Distance	W(i)	B2(i)
2.00	8.00	47.0	.352	.154	19.5
2.00	7.00	.000	.675	.119	14.4
3.00	8.00	7.00	.950	.893E-01	9.93
3.00	7.00	-15.0	1.11	.749E-01	7.54
1.00	8.00	-9.00	1.16	.731E-01	6.82
1.00	7.00	-1.00	1.30	.627E-01	5.12
2.00	9.00	10.0	1.34	.566E-01	4.59
3.00	9.00	16.0	1.60	.446E-01	2.17
2.00	6.00	1.00	1.67	.430E-01	1.72
1.00	9.00	-4.00	1.74	.435E-01	1.34
3.00	6.00	-4.00	1.89	.407E-01	.698
4.00	8.00	-6.00	1.92	.377E-01	.601
1.00	6.00	2.00	2.00	.403E-01	.386
4.00	7.00	1.00	2.00	.394E-01	.360
4.00	9.00	-6.00	2.31	.400E-01	.150E-01
2.00	10.0	-10.0	2.34	.416E-01	.000

Neighbors:	16	_ Weights:	1.000	
Estimate:	6.23	CVV:	19.720	
_k:	4.37	_ WiB2i:	7.550	
^{2}k:	19.1	μ:	-6.905	

C.8 'Estimation of the variogram parameters by MARQUARDT METHOD';

```
Data Variogram;              {Start of the program}
Input Gamastar Count Lag     {define variable}
Cards;                       {Input residuals of iron ore digits}
Proc NLIN Best=30 Maxiter=400  Methods=MARQUARDT;  {Non linear optimization process}
Parms As Cs As                        {Starting value of the parameters}
```

Model Gamastar= $C_0 + C_s \times (3/2) \times (Lag / As) - (1/2) \times (Lag / As)^3$; {Define the model}

Weight= $Count / C_0 + C_s \times (3/2) \times (Lag / As) - (1/2) \times (Lag / As)^3$; {Define Weight}

Der.Co=1

Der.Cs= $((3/2) \times (Lag / As) - (1/2) \times (Lag / As)^3$

Der.As= $Cs \times (3/2) \times (Lag / As)^2 + (3/2) \times (Lag^3 / As^4)$;

```
End;
Else Do;
Model Gamastar=Co + Cs;       {Model of the equation}
```

Weight= $Count / (Co + Cs) \times 2$;

```
End;
If=(_Obs_=1 and _Model_=1) Then Do;
Sill=Co +Cs;
Put Cs=Sill;End
Run;
```

(C.9)

Point kriged is :(9.000, 1.000)					
X(i)	Y(i)	V(i)	Distance	W(i)	B2(i)

9.00	1.00	6.70	.000	.177	5.26
9.00	.000	7.00	1.00	.105	4.80
8.00	1.00	3.60	1.00	.900E-01	4.80
10.0	1.00	6.00	1.00	.105	4.80
9.00	2.00	6.30	1.00	.900E-01	4.80
10.0	.000	5.80	1.41	.836E-01	4.53
10.0	2.00	4.10	1.41	.684E-01	4.53
8.00	.000	8.50	1.41	.684E-01	4.53
8.00	2.00	6.00	1.41	.554E-01	4.53
7.00	1.00	4.90	2.00	.295E-01	4.15
9.00	3.00	7.00	2.00	.295E-01	4.15
7.00	.000	4.10	2.24	.294E-01	3.99
7.00	2.00	4.70	2.24	.195E-01	3.99
8.00	3.00	4.90	2.24	.195E-01	3.99
10.0	3.00	4.40	2.24	.294E-01	3.99
Neighbors: 15 _ Weights: 1.000					
Estimate: 5.91 CVV: 5.211					
_k: .733 _ WiB2i: 4.691					
_²k: .537 μ: -.017					

C.10

+------------------------------ M O D E L ------------------------------+

Pairs	Avg Distance	Value				
1 121	1.000	4.321				
2 101	2.000	7.635				
3 84	3.000	10.415				
4 228	3.416	10.425				
5 193	4.396	10.262				
6 171	5.394	9.834				
7 258	6.359	8.769				
8 238	7.328	10.289				
9 214	8.299	9.991	Model			
10 252	9.337	9.930		Nugget :	1.000	
11 203	10.317	12.347				
12 208	11.385	11.033	Type	Sill	Range	
13 167	12.372	9.990				
14			Spherical	10.000	4.000	