Imperial College London

Department of Environmental Science and Technology

A Spatially-Evaluated Quantitative Risk Assessment for Contaminated Land

by

Jennie Rebecca Gay

A thesis submitted for the degree of
Doctor of Philosophy of the University of London

June 2004

BIBL. LONDIN. UNIV.

Abstract

Contaminated land is a major problem in many developed countries. Most are blighted to some extent by contamination left from past, bad industrial practice and authorities are now enforcing clean up to varying degrees. Recent policy developments are moving towards risk-based contaminated land assessment within a 'fit–for-use' framework, geared around protection of human (and sometimes ecological) health.

Despite potential demand, there is a lack of methodologies for assessing risk to human health from contaminated land. Of those available, most are deterministic, not probabilistic, and do not take account of the spatial distribution of risks. Land contamination is a spatial phenomenon, as are the risks associated with it. The problem at hand lends itself to mapping, and understanding and communicating the risks are improved if the spatial dimension is preserved.

This research combines quantitative probabilistic human health risk assessment and spatial statistical methods (geostatistics) to produce a spatially-evaluated quantitative risk assessment methodology for quantifying risks to human health from exposure to contaminated land.

It is concluded that the interaction between soil contaminant levels, variability in the population characteristics and the regulatory cut offs are complex, making results difficult to predict. Key contributors affecting risk outcomes, introduced as a result of working in a spatially-evaluated manner, include the chosen spatial resolution and location of the population in relation to contaminant hotspots. In addition, variability in population characteristics is important. For Cd intakes, variable soil pH should be accounted for.

The newly developed method provides a depth of information which can be displayed in effective map form whilst conveying uncertainty in results. Examples include maps of soil contaminant levels, probability of soil concentrations exceeding regulatory cut offs and number of people in each location whose intake exceeds the regulatory cut off. Such maps provide valuable information for risk managers.

Acknowledgements

This research was funded by the Environmental Management of Hazardous Mining Wastes and Effluents (ENVIMAN) EU COPERNICUS Programme, Project No: ICA2-1999-10072.

I would like to thank my supervisor, Dr. Anna. Korre, and Professor Sevket Durucan of the Environmental Processes and Systems Research Group for giving me the opportunity to carry out this research. Thanks also to Anna for technical and strategic advice, and the use of her Lavrio data set and maps.

The Russian data was supplied by colleagues on the ENVIMAN project, Professor M. Zilberschmidt and Professor M.Shpirt of Moscow State Mining University, for which I am very grateful. Thanks also to Kate Gorunova for digitising the maps and being my contact between Moscow and London.

Thanks are also extended to the British Geological Survey for providing the data for Stoke-on-Trent.

My colleagues in the research group who have accompanied me through the various enjoyments, trials and tribulations of the past four years are: Dr Sameh Ahmed, Dr. Sungwook Yoon, Dr Edward Lewis, Dr. Laurent Chane Kon; Dr. Ji Quan Shi, Dr Ego Syahrial, Dr Gabriela Munoz Melendes, Dr Claire Imrie, Dr Andrew Richmond, Robert Radloff, Ruben Rubio, Mustafa Ahsan, Dimitrios Papanastasopoulos, Munyindei Masialeti and Vassilis Trikkalis. Best wishes for the future to you all, and good luck to all who are still to write up!

Thanks to Andrew Richmond for his help with geostatistics and to Claire Imrie both for ensuring the computer facilities in general functioned throughout and for sorting out my own computer problems (many operator related!). Also to Dr. Murdoch McAllister for statistical advice when I started designing my intake model.

Thanks to Helen Ferrier for numerous lunch breaks, tea breaks and moral support.

Finally, last, but definitely not least, thanks to my husband Nigel for all his support, encouragement and understanding throughout.

Affirmation

The work submitted in this thesis is my own, and has not been submitted previously for any other degree. The following publications have resulted from this work:

Gay, J.R., Korre, A., and Durucan, S. (2004). Spatially evaluated health risk estimation: the effect of change of support. Proc. 1st International Conference on Advances in Mineral Resources Management and Environmental Geotechnology, 7-9 June 2004, Chania, Crete, Greece, pp437-442.

Gay, J.R., Korre, A. & Durucan, S. (2004). Contaminated Land Health Risk Assessment and Mapping: The Effect of Support on Risk Estimation. Proc. 7th Int. Symp. on Environmental Issues and Waste Management in Energy and Mineral Production (SWEMP 2004), 17-20 May, 2004, Antalya, Turkey, pp103-108.

Korre, A., Gay, J.R., Durucan, S. (2003). Quantitative assessment of the risks associated with high soil heavy metal loads in mining districts. International Conference on Sustainable Development Indicators in the Mineral Industries (SDIMI 2003), 21-23 May 2003, Milos, Greece, pp305-310.

Table of Contents

Abstract.....i
Acknowledgements.....ii
Affirmation.....iii
Table of Contents.....iv
List of Figures.....xi
List of Tables.....xvi

Chapter 1 Introduction.....1

1.1 Motivation for research.....1

1.2 Aims and objectives.....2

1.3 Thesis structure.....3

Chapter 2 Risk Assessment in the context of contaminated land.....6

2.1 Introduction.....6

2.2 Risk: definition.....6

2.3 Source-pathway-receptor paradigm.....7

2.4 Conceptual models for risk assessment.....7

2.4.1 National Research Council (NRC) model.....8

2.4.2 Petts et al. model.....9

2.4.3 Covello-Merkhofer model.....10

2.4.4 Ferguson et al. model.....12

2.5 Uncertainty and variability.....13

2.5.1 Qualitative and quantitative risk assessment.....13

2.5.2 Deterministic versus probabilistic risk assessment.....13

2.5.2.1 Deterministic risk assessment.....14

2.5.2.2 Probabilistic risk assessment.....14

2.5.2.3 Partitioning uncertainty and variability.....16

2.6 Contaminated Land.....16

2.6.1 Definition of contaminated land.....16

2.6.2 Legislation and regulation.....17

2.6.3 Guideline values 18
2.7 Tiered approach to contaminated land risk assessment 21
2.8 Exposure models 23
2.8.1 Detailed example of an exposure model 24
2.9 State-of-the-art versus current practice 25
2.10 Summary 26

Chapter 3 Geostatistics in the context of risk assessment and contaminated land 27

3.1 Introduction 27
3.1.1 Introduction to geostatistics 28
3.2 Fundamental concepts 29
3.2.1 The regionalised variable 29
3.2.2 Stationarity 30
3.2.3 Characterisation of the spatial phenomenon 30
3.2.4 Variography 32
3.2.5 Kriging 34
3.3 Practical considerations 37
3.3.1 Smoothing effect 37
3.3.2 Block estimates 38
3.3.2.1 Theory of block kriging. 39
3.3.3 Non-normality of data 40
3.3.4 Quantifying uncertainty 42
3.3.4.1 Uncertainty in the analytical results 42
3.3.4.2 Uncertainty in the variogram 43
3.3.4.3 Uncertainty in interpolation 43
3.3.5 Simulation 46
3.3.6 Kriging versus simulation 47
3.4 Methods used in this research 48
3.4.1 Indicator kriging: theory 49
3.4.1.1 Problem with block kriging indicators. 50
3.4.2 Sequential indicator simulation: theory 50
3.5 Summary 52

Chapter 4 Conceptual research framework 53

4.1 Introduction 53

4.2 Merging two disciplines 53

4.3 Definitions 56

4.4 Conceptual Framework for this research 56

4.5 Two levels of risk evaluation 58

4.6 Points addressed during development of the methodology 59

4.6.1 Exposure assessment 59

4.6.2 Intake assessment 60

4.6.3 Risk estimation 60

4.6.4 Risk evaluation 60

4.6.5 Risk communication 60

Chapter 5 Exposure assessment using geostatistics 62

5.1 Introduction 62

5.2 Background to the data set 63

5.3 Method 65

5.3.1 Variogram modelling 65

5.3.2 Selection of grid size and moving window neighbourhood 72

5.3.3 Support correction 72

5.3.4 Performing IK 73

5.3.5 Performing SIS 74

5.4 Risk assessment criteria 74

5.5 Analysis of results 76

5.5.1 Uncertainty due to the variogram model 76

5.5.2 Uncertainty due to support size 80

5.5.3 Uncertainty due to method choice: IK versus SIS 84

5.6 Summary 86

Chapter 6 Building the spatial and probabilistic intake assessment methodology 88

6.1 Introduction 88

6.2 Accounting for variability, uncertainty and the spatial dimension 88

6.3 Introduction and background to the CLEA model 90
6.3.1 Definitions 92
6.3.2 Derivation of SGVs 92
6.3.2.1 Threshold contaminants and the Tolerable Daily Soil Intake (TDSI) 93
6.3.2.2 Non-threshold contaminants and the Index Dose (ID) 93
6.3.2.3 Calculating the SGV 93
6.3.2.4 Exposure scenarios and critical receptors 94
6.4 Adaptation of the CLEA model 95
6.4.1 Exposure scenarios 95
6.4.2 Receptor population 95
6.4.3 Basic intake equation 96
6.4.4 Derivation of input variables 97
6.4.4.1 SDR 97
6.4.4.2 CR_{veg} 98
6.4.4.3 BW 98
6.4.4.4 HF_{veg} 100
6.4.4.5 SL_{veg} 101
6.4.4.6 CF_{veg} 101
6.4.4.7 C_{soil} 102
6.4.4.8 EF, ED and AT 102
6.5 The concept of exposure units 103
6.6 Allocating the population to an EU 104
6.6.1 Simple method 104
6.6.2 Advanced method 105
6.6.2.1 Census data 105
6.6.2.2 Allocation procedure 105
6.7 Linking population to soil contamination level. 107
6.8 Calculations 108
6.8.1 Generation of the population 108
6.8.2 Calculating ADI 109
6.9 Results output and evaluation 109
6.9.1 Assessment criteria 110
6.9.2 Risk evaluation 111
6.9.2.1 Two dimensional decision criteria 111

6.9.3 Spatial displays 114
6.9.4 The drawback of thresholds 114
6.10 Summary of methodology 114

Chapter 7 Effect of individual variation on intake 116

7.1 Introduction 116
7.1.1 Between-individual variation 116
7.2 Rearranging the intake algorithm 117
7.3 Deterministic investigation 119
7.4 Probabilistic investigation 121
7.5 Distribution of SE 125
7.6 Proportion of population over-exposed at different soil concentrations 126
7.6.1 As 126
7.6.2 Ni 128
7.6.3 Cd 130
7.7 Relative importance of variability in SE 132
7.8 Summary 135

Chapter 8 Analysis of intake results 137

8.1 Introduction 137
8.1.1 Contaminants of concern 138
8.1.2 Data sets 139
8.2 Interaction of $\dot{SE}$, Csoil and thresholds 139
8.3 Intake results: influence of different factors 142
8.3.1 Change of Support 142
8.3.1.1 Background to the Stoke As data set 142
8.3.1.2 Analysis of results 145
8.3.2 Where in the distribution the threshold lies. 149
8.3.2.1 Background to the Stoke Cd data set 149
8.3.2.2 Analysis of results 150
8.3.3 Location of the population 154
8.3.3.1 Background to the Gai data set 154
8.3.3.2 Analysis of results 156
8.3.4 Location of population combined with change of support 159

8.3.4.1 Analysis of results 159
8.3.5 Loss of information when working with cut offs 161
8.3.5.1 Analysis of results 162
8.3.6 Working with populations or individuals. 164
8.3.6.1 Analysis of results 165
8.3.7 pH dependent Cd intake assessment 165
8.3.7.1 Background to the Vasilevskaya data set 166
8.3.7.2 Effective Cd concentration 168
8.3.7.3 Analysis of results 171
8.4 Spatial displays of results 173
8.4.1 Number of people exceeding the ID/TDSI. 173
8.4.2 Expected number of people and magnitude of excess intake 176
8.5 Summary 177

Chapter 9 General Discussion 180

9.1 Introduction 180
9.2 General applications 181
9.3 Exposure assessment 181
9.4 Intake assessment 184
9.4.1 Major contributions of research 184
9.4.2 Working with a variable population 185
9.4.3 Spatial considerations 187
9.4.4 Populations and individuals 189
9.4.5 Practical considerations 190
9.5 Effective Cd 190
9.6 Limitations of the work 191

Chapter 10 Conclusions and recommendations 193

10.1 Introduction 193
10.2 Conclusions 194
10.2.1 Completion of main objectives 194
10.2.2 Additional novel achievements 195
10.2.3 Insights into probabilistic spatially-evaluated risk assessments 195
10.2.4 Policy implications 196

10.3 Limitations 197
10.4 Recommendations for future work 197

References **198**

Appendix A Geostatistical procedure for Arsenic in Stoke on Trent **208**

Appendix B Geostatistical procedure for Nickel in Gai **216**

Appendix C Geostatistical procedure for Cadmium and pH in Vasilevskaya **225**

Appendix D Geostatistical procedure for Cadmium and pH in Stoke on Trent **241**

List of Figures

Figure 1.1. Two levels for assessment of risks to human health from contaminated land. 2

Figure 2.1. Covello-Merkhofer conceptual risk model. 11

Figure 2.2. Components of a risk assessment (Ferguson et al. 1998). 12

Figure 2.3. Threshold/action concept (Petts et al., 1997). 19

Figure 2.4. Soil and groundwater quality standards and assessment (Swartjes, 1999) 20

Figure 2.5. Tiered approach to risk assessment (DETR et al., 2000). 22

Figure 3.1. Idealised variogram plot. C_0 = nugget, C_1 = sill, a = range 33

Figure 3.2. Four standard variogram models. 34

Figure 4.1. Diagrammatic representation of the conceptual framework for work in this thesis 57

Figure 4.2. Options for evaluating risk 59

Figure 5.1. Two levels for assessment of risks to human health from contaminated land. 62

Figure 5.2. Histogram of soil sample Pb concentrations ($mg.kg^{-1}$). 64

Figure 5.3. Soil sample points with an indication of soil Pb concentration. 65

Figure 5.4. Variogram maps for all cut offs demonstrating an anisotropy, strongest at the highest cut offs 67

Figure 5.5. Sample variograms (symbols) and bi-directional nested model variograms (lines); (BN) 69

Figure 5.6. Sample variograms (symbols) and bi-directional single structure models (lines); (BS) 70

Figure 5.7. Sample variograms (symbols) and omni-directional single structure models (lines); (OS) ... 71

Figure 5.8. Sample variograms (symbols) and model variograms (lines) for Pb sample data. 73

Figure 5.9. Frequency distribution of the probability of exceeding $2000mg.kg^{-1}$ (top) and $500mg.kg^{-1}$ (bottom) according to results of IK on a fine grid. S: percentage of land with probability ≥ 0.2 of exceeding threshold (stringent criterion). W: percentage of land with probability ≥ 0.8 of exceeding threshold (weak criterion). 77

Figure 5.10. Frequency distribution of the probability of exceeding $2000mg.kg^{-1}$ (top) and $500mg.kg^{-1}$ (bottom) according to results of SIS on a fine grid. S: percentage of land with probability ≥ 0.2 of exceeding threshold (stringent criterion). W: percentage of land with probability ≥ 0.8 of exceeding threshold (weak criterion). 77

Figure 5.11. Probability of exceeding 2000mg.kg^{-1} using IK on a fine grid with (a) BN model set (b) BS model set and (c) OS model set. P≥0.8 represents land contaminated under the weak criterion and P≥0.2 represents land contaminated under the stringent criterion..... 78

Figure 5.12. Accuracy plot for IK (top) and SIS (bottom) for the 3 model sets. A, Accuracy; P, Precision; G, Goodness. 80

Figure 5.13. Frequency distribution of the probability of exceeding 2000mg.kg^{-1} (top) and 500mg.kg^{-1} (bottom) according to results of IK on a fine grid (left) and coarse grid (right). S: percentage of land with probability ≥ 0.2 of exceeding threshold (stringent criterion). W: percentage of land with probability ≥ 0.8 of exceeding threshold (weak criterion)..... 82

Figure 5.14. Frequency distribution of the probability of exceeding 2000mg.kg^{-1} (top) and 500mg.kg^{-1} (bottom) according to results of SIS on a fine grid (left) and coarse grid (right). S: percentage of land with probability ≥ 0.2 of exceeding threshold (stringent criterion). W: percentage of land with probability ≥ 0.8 of exceeding threshold (weak criterion)..... 82

Figure 5.15. Soil Pb (mg.kg^{-1}) at (a) 2.5th percentile and (b) 97.5th percentile, being the lower and upper bounds of the 95% confidence interval respectively. Results of performing SIS on a fine grid. 83

Figure 5.16. Soil Pb (mg.kg^{-1}) at (a) 2.5th percentile and (b) 97.5th percentile, being the lower and upper bounds of the 95% confidence interval respectively. Results of upscaling SIS to a coarse grid. 83

Figure 5.17. Soil Pb (mg.kg^{-1}) at (a) 2.5th percentile and (b) 97.5th percentile, being the lower and upper bounds of the 95% confidence interval respectively. Results of performing IK on a fine grid. 84

Figure 6.1. Regression of mean and standard deviation of ln *BW* by age. R^2=0.9984 and 0.9504 for the mean and standard deviation respectively..... 99

Figure 6.2. Representation of GIS layers used to allocate the population of Stoke to a 50x50m grid: 1) ED polygons (different shades indicate differing numbers of people per ED). 2) Points representing position of roofed buildings. 3) Grid cells to which the population must be allocated (50x50m grid used but 200x200m grid displayed for clarity). 4) Final population density (darker shades represent higher density)..... 107

Figure 6.3. Representation of Simulated *ADI* results..... 110

Figure 6.4. Using two dimensional decision criteria: a schematic representation of results. Shading in results table represents ADI estimates in excess of ID/TDSI, where individuals are in rows and simulations are in columns. P: applying criterion by column then row. I: applying criterion by row then column. E: expected proportion of population in excess of ID/TDSI. 5% critical proportion highlighted in histograms..... 113

Figure 6.5. Spatially-evaluated probabilistic risk assessment procedure..... 115

Figure 7.1 Major factors contributing to SE for As, Ni, and Cd at pH7. Left: SE split by four major contributors: eating leafy vegetables; potatoes; and other root vegetables; and by direct ingestion of soil and/or dust. Right: Proportion of SE attributable to contaminant absorbed into consumed vegetables, attached to consumed vegetables and ingested directly from soil and/or dust..... 120

Figure 7.2. Scatter plot of *SDR* versus *SE* for As (left) and home-grown potato consumption versus *SE* for As (right). Correlation coefficient, ρ, shown..... 123

Figure 7.3. Scatter plot of home-grown potato consumption versus *SE* for Cd. Correlation coefficient, ρ, shown..... 124

Figure 7.4. Distribution of *SE* for As and Ni (left) and Cd at pH7 and pH6 (right)..... 125

Figure 7.5. Distribution of standardised *SE* for Cd at pH6 and pH7, As and Ni..... 126

Figure 7.6. Proportion of the population whose intake is in excess of the Index Dose (ID) at different soil concentrations. Solid lines represent the scenario including home-grown vegetable consumption and dashed lines represent the scenario excluding home-grown vegetable intake. SGV is the Soil Guideline Value (see text)..... 127

Figure 7.7. (a) Demonstration of how mean *SE* changes with age; (b) Proportion of the population exceeding the ID for As at the SGV of 20mg.kg^{-1}..... 128

Figure 7.8. Proportion of the population whose intake is in excess of the Tolerable Daily Soil Intake (TDSI) at different soil concentrations. Solid lines represent the scenario including home-grown vegetable consumption and dashed lines represent the scenario excluding home-grown vegetable intake. SGVv is the Soil Guideline Value for the 'with vegetable' scenario, SGVnv is the Soil Guideline Value for the 'no vegetable' scenario (see text). 129

Figure 7.9. Proportion of the population whose intake is in excess of the Tolerable Daily Soil Intake (TDSI) at different soil concentrations: (a) at pH7 (b) at pH6. Solid lines represent the scanario including home-grown vegetable consumption and dashed lines represent the scenario excluding home-grown vegetable intake. SGVv is the Soil Guideline Value for the 'with vegetable' scenario (see text). 130

Figure 7.10. Demonstration of how mean *SE* changes with age (left) and proportion of the population exceeding the TDSI for Cd (right). (a) pH7, SGVv 2mg.kg^{-1}and SGVnv 30mg.kg-1 (b) pH6, SGVv 1mg.kg^{-1}..... 131

Figure 7.11. Interaction of C_{soil} and *SE*: two extreme situations..... 133

Figure 7.12. Proportion of population exceeding the ID/TDSI at different soil concentrations. Plot standardised by the C_{soil}50 (soil concentration at which 50% of the population are expected to exceed the cut off)..... 134

Figure 8.1. Demonstration of how the proportion of the population classed as exceeding the ID/TDSI changes with the distribution of C_{soil} and SE and the relative position of mean C_{soil} to the SGV..... 141

Figure 8.2. Map of Stoke on Trent region showing soil sampling positions..... 143

Figure 8.3. Histogram of distribution of As soil sample concentrations ($mg.kg^{-1}$)..... 144

Figure 8.4. Map of median soil As concentrations after 1,000 simulations on (a) 50m grid and (b) 200m grid. Pale symbols represent population density. 146

Figure 8.5. Map of P95 soil As concentrations after 1,000 simulations on (a) 50m grid and (b) 200m grid. Pale symbols represent population density. 147

Figure 8.6. Distribution of results, from 1,000 simulations, of the estimated fraction of population exceeding the ID for As ($0.0003mg.kg^{-1}_{BW}.d^{-1}$). Left hand column; assuming no home-grown vegetable consumption. Right hand column; assuming home-grown vegetable consumption. 148

Figure 8.7. Histogram of distribution of Cd soil sample concentrations ($mg.kg^{-1}$). One value ($43mg.kg^{-1}$) not shown..... 149

Figure 8.8. Distribution of median *ADI* estimates for As (top) and Cd (bottom) for 1yr olds (left) and 16 yr olds (right). Distributions calculated for three different EU sizes: whole site; 200x200m; and 50x50m. Box and whisker plot lines represent, from top to bottom: Maximum, upper quartile, median, lower quartile, minimum. Triangle represents mean. As ID and Cd TDSI highlighted..... 151

Figure 8.9. Distribution of 1,000 estimates of the proportion of the population exceeding the ID for As (left) and TDSI for Cd (right) on three different supports in Stoke on Trent. (Scenario including home-grown vegetable consumption). 153

Figure 8.10. Gai region showing built up areas in four dispersed settlements. The grid represents the 500x500m EUs and the points the Ni soil sample points, with symbol size in proportion to Ni concentration. 155

Figure 8.11. Distribution of Nickel topsoil sample concentrations. One value of $981mg.kg^{-1}$ not shown. 155

Figure 8.12. Map of median soil Ni concentrations after 1,000 simulations on a 50m grid 157

Figure 8.13. Distribution of soil Ni concentrations in the Gai region. Left; for all 50x50m EUs. Right; distribution of values to which individuals are exposed..... 158

Figure 8.14. Distribution of results, from 1,000 simulations, of the estimated fraction of population exceeding the TDSI for Ni. (a) population randomly distributed in space (b) population correctly allocated to residential EUs..... 159

Figure 8.15. Median Ni soil concentrations in Gai on a 500m support (a) and a 50m support (b)..... 160

Figure 8.16. Comparison of the change in distribution of median Ni concentrations (left) and change in distribution of fraction of population exceeding the TDSI (right) for two different supports; 500m EU (top) and 50m EU (bottom). 161

Figure 8.17. Distribution of 1,000 estimates of the proportion of 16yr olds exceeding the TDSI (0.00077mg.kg$^{-1}_{BW}$.d^{-1}) for Cd in Stoke on Trent. 162

Figure 8.18. Distribution of Cd *ADI*s for each individual 16yr old in Stoke (mg.kg$^{-1}_{BW}$.d^{-1}). Top, ordered by highest intakes first, bottom ordered by lowest intakes first. 163

Figure 8.19. Distribution of median (left) and P95 (right) estimates of Cd *ADI* for 16yr olds in Stoke. 164

Figure 8.20. Vasilevskaya region showing four dispersed settlements. The grid represents the 250x250m EUs and the points the Cd soil sample points, with symbol size in proportion to Cd concentration..... 166

Figure 8.21. Histogram of Cd topsoil sample concentrations (mg.kg^{-1}). (left) and pH values (right)..... 167

Figure 8.22. Map of the median soil Cd concentration after 1,000 simulations on a 25m grid. 169

Figure 8.23. Map of the median soil pH after 1,000 simulations on a 25m grid..... 169

Figure 8.24. Map of the median effective Cd concentration after 1,000 simulations on a 25m grid. 171

Figure 8.25. Distribution of results, from 1,000 simulations, of the estimated fraction of population exceeding the TDSI for Cd, with different soil pH regimes. (a) no spatial variability or uncertainty in pH (b) Spatial variability only in pH (c) Spatial variability and uncertainty in soil pH. 172

Figure 8.26. Number of people in Gai exceeding the TDSI for Ni in each 500m EU. (a) median estimate (b) P95 estimate..... 174

Figure 8.27. Median number of people in Stoke exceeding the TDSI for Cd per 200m EU. (a) calculated on a site wide support (b) calculated on a 200m support (c) calculated on a 50m support. All transformed to a 200m grid for a comparative display. 175

Figure 8.28. Expected number of people per 200m EU who exceed the TDSI by (a) up to 2 times (b) between 2 and 4 times (c) in excess of 4 times..... 177

List of Tables

Table 5.1a. Indicator variogram model structure (BN). 68

Table 5.1b. Indicator variogram model structure (BS). 68

Table 5.1c. Indicator variogram model structure (OS). 68

Table 5.2. Grids used for kriging and simulation. Coordinates represent the central points around which a block is estimated. The coarse and fine grids overlay exactly. 72

Table 5.3. Statistics for Pb sample data (samp) and simulated values (sim) ('000mg.kg^{-1})..... 74

Table 5.4. Percentage of land area classed as contaminated at the two regulatory cut offs using two different probability criteria: S (P≥0.2); W (P≥0.8)..... 85

Table 6.1. Methods to account for variability, uncertainty and spatial dimension in soil concentrations. 89

Table 6.2. Methods to account for variability, uncertainty and spatial dimension in receptor population. 89

Table 6.3. *SDR* values used in calculation of *ADI* 98

Table 6.4. CR_{veg} values used in calculation of *ADI*..... 98

Table 6.5. Parameters for *BW* distributions used in calculation of *ADI*..... 100

Table 6.6. Parameters for HF_{veg} distributions used in calculation of *ADI*..... 100

Table 6.7. SL_{veg} values used in calculation of *ADI*..... 101

Table 6.8. Prescribed CF_{veg} values..... 101

Table 6.9. Dry weight to fresh weight conversion factors for different vegetable types. 102

Table 8.1. Statistics for topsoil (A horizon) As at a depth of circa 15cm (mg.kg^{-1}). 144

Table 8.2. Population numbers in the study area 144

Table 8.3. Statistics for topsoil (A horizon) Cd at a depth of circa 15cm (mg.kg^{-1}). 149

Table 8.4. Statistics of Ni sample distribution (mg.kg^{-1}). 154

Table 8.5. Population in each settlement in the Gai region..... 156

Table 8.6. Risk estimation using three methods. Shaded cells represent at-risk results..... 165

Table 8.7. Simple statistics of the Cd topsoil sample concentrations. 167

Table 8.8. Population in the Vasilevskaya mining region..... 167

Chapter 1 Introduction

1.1 Motivation for research

Contaminated land is a major problem in many developed countries. Most are blighted to some extent by contamination left from past, bad industrial practice. Authorities are enforcing clean up of contaminated sites to varying degrees for reasons including shortage of space for new development and fear of litigation, and also as part of a general move towards more sustainable practices.

Recent policy developments are moving towards regulations which encourage a risk-based approach to contaminated land assessment within a 'fit-for-use' framework, geared around the protection of human (and sometimes ecological) health. A tiered approach to risk assessment is encouraged whereby simple generic methods are used initially, which are replaced by more complex site specific methods should the situation merit it. As a result, earlier qualitative risk assessment methods must be supplemented by more complex quantitative and probabilistic risk estimation methods which take account of variability and uncertainty.

After reviewing the literature, it becomes apparent that there is a lack of methodologies for assessing risk to human health from contaminated land, even though there is a potential demand for them. Of the methods that are available, most are deterministic, as opposed to probabilistic, and they do not take account of the spatial distribution of risks. Land contamination is a spatial phenomenon as are the risks associated with it, and this author feels it is important to preserve the valuable spatial dimension in the assessments. The problem at hand lends itself to mapping, and understanding and communicating the risks are improved if the spatial dimension is preserved.

The process for quantitatively assessing risks from contaminated land starts with collection and analysis of soil samples from the site. Following this, there are two levels at which risk to human health from contaminated land may be evaluated. First, one may quantify the contaminant level in the soil and compare it to a regulatory cut-off level deemed 'safe' or 'acceptable'; level 1 in Figure 1.1. Second, one may calculate the human intakes of soil contaminants and compare them to regulatory 'safe' or 'acceptable' intake levels; level 2 in Figure 1.1. Calculations of the risks at level 1 may be achieved using geostatistical methods, which produce estimates of contaminant concentrations everywhere on the site from sparse sample data, while those at level 2 may be achieved by feeding the contaminant data into intake assessment models, which are generally currently lacking.

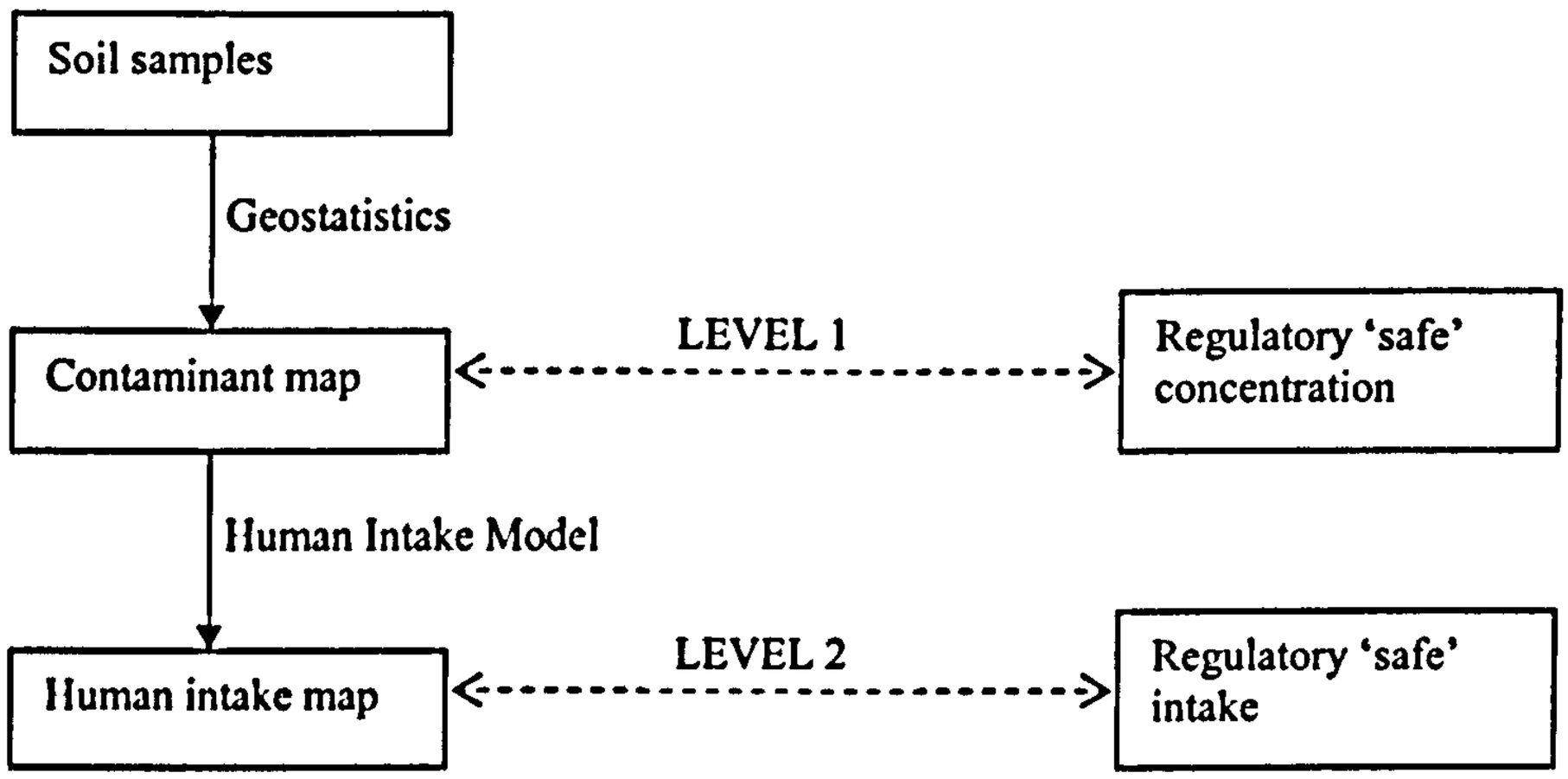


Figure 1.1. Two levels for assessment of risks to human health from contaminated land.

1.2 Aims and objectives

The aim of this research was to produce a methodology for quantifying risks posed to human health by exposure to contaminated land, in a manner which preserves the spatial distribution of the risks and gives a measure of uncertainty in the estimates.

The main objectives were:

- to find the most appropriate geostatistical tools for spatial assessment of land contaminants;

- to identify a suitable intake model and to adapt it for use in a spatially-evaluated, probabilistic methodology;
- to produce meaningful presentations of results for use in risk management and communication.

In addition to achieving these objectives, useful insights into the use of spatial data in risk assessments were produced.

Note: the methodology adapts a current intake model, which, for the purposes of this research, the author accepts as correct (it is not within the remit of this research to verify how good the model is); the methodology assesses risks to human health only; and it assesses direct risks from soil contamination only (it was not within the scope of this research to investigate contaminant effects on groundwater, for example, and intakes from such sources).

1.3 Thesis structure

By means of a literature review, Chapter 2 introduces the general concept of risk assessment before focussing on applications for assessment of risks to human health from exposure to contaminated land. Legislative and regulatory information is introduced to illustrate where such assessments fall within the regulatory framework in general, and particularly in the UK. Uncertainty and variability are discussed in some detail.

Chapter 3 introduces geostatistics, a spatial evaluation tool, in the context of contaminated land and risk assessment. The fundamental concepts of geostatistics are presented along with a discussion of the practical considerations of using geostatistics in this context. Choice of support is an important geostatistical issue which may have a marked effect on results. After reviewing the literature, indicator kriging and sequential indicator simulation are selected for use in this thesis and the theory of these methods is presented.

After having reviewed risk assessment in Chapter 2 and spatial assessment in Chapter 3, Chapter 4 pulls the two disciplines together and presents the research framework for the remainder of the thesis. First, the necessity of this work is justified by means of a short summary of cross disciplinary literature, some of which helped form the seed of the idea for this research. Next, the conceptual framework is presented and two levels at which

risk can be assessed are defined as 'exposure assessment' (level 1 Figure 1.1) and 'intake assessment' (level 2 Figure 1.1).

Chapter 5 presents an exposure assessment case study (level 1 Figure 1.1) using indicator kriging and sequential indicator simulation. It investigates the issues raised in Chapter 3 regarding the uncertainties in geostatistical assessments and looks at the implications for exposure assessments and potential implications for subsequent intake assessments.

Chapter 6 presents the new methodology for probabilistic spatially-evaluated intake assessment, developed especially for this thesis (level 2 Figure 1.1). A probabilistic intake assessment model (Contaminated Land Exposure Assessment model, CLEA) is adapted to incorporate spatial evaluation techniques and to calculate intakes for mixed populations. A description is given of how population data and soil contaminant values from the geostatistical analysis are related spatially during the calculation procedure, along with a description of the results output and how it may be analysed to preserve the depth of information available from the intake assessment.

Having explored, in Chapter 5, issues surrounding estimation of soil contaminant concentrations and the potential effects on risk assessment outcomes, Chapter 7 looks specifically at how variation in the characteristics of the receptor population effect the outcome of the intake assessment. To do this, the CLEA algorithm, is rearranged to separate the soil contaminant concentration term from factors that are specific to the individuals. These factors are summarised into a single variable, the SDR equivalent (*SE*). Key factors contributing to *SE* are investigated, and the distribution of *SE* for a population, which represents the between-individual variation, is studied in detail.

Interactions between soil contaminant concentrations and *SE* are important and change in a complex manner from contaminant to contaminant and support size to support size and depend also on how the components relate in space. The aim of Chapter 8 is to demonstrate the influence of these different factors on intakes through presentation of a varied selection of results from the newly developed intake assessment methodology, which take into account variability, uncertainty and the spatial dimension. The results were produced from the analysis of four diverse example data sets (consisting of varying numbers of exposed people and different elements of concern). In addition, example presentations of intake results are given, which preserve the spatial dimension.

Chapter 9 presents an overall discussion which looks generally at the advantages and disadvantages of the proposed new methodology from both a technical and a policy oriented view. Technically, the practicalities of applying the model are discussed along with an overall assessment of the type of inputs, calculations made and outputs. From a policy point of view, the usefulness of the methodology is assessed in the light of current practices, and possibilities for other useful implementations are proposed.

Finally, Chapter 10 presents the main conclusions and accomplishments and recommends further research to overcome limitations in the new methodology.

Chapter 2 Risk Assessment in the context of contaminated land

2.1 Introduction

This chapter introduces the general concept of risk assessment before focussing on applications for assessment of risks to human health from exposure to contaminated land. Legislative and regulatory information is introduced to illustrate where such assessments fall within the regulatory framework in general, and in that of the UK in particular. Conceptual models provide a basic framework for the steps of the assessment process and various examples are shown. Since the aim of this thesis is to provide a novel, spatially-evaluated quantitative risk assessment which accounts for uncertainty and variability, the two issues of uncertainty and variability are discussed in some detail. The spatial assessment details are discussed in the following chapter. To fulfil the requirements of this thesis a probabilistic exposure assessment model, available in the public domain, is adapted for use in a spatial manner: this model, in its basic form, is introduced here.

2.2 Risk: definition

The New Oxford Dictionary of English (OUP, 1998) defines risk as: "noun: a situation involving exposure to danger" and "verb: expose (someone or something valued) to danger, harm, or loss". Technicians in various fields have proposed more scientific definitions of risk, for example "the likelihood and magnitude of adverse effects under defined exposure conditions" (Gray et al., 1997) and "a combination of the probability and frequency of a harm being realised with an estimate of its magnitude or scale" (Petts et al., 1997). Such technical definitions incorporate probability (likelihood/chance) of an adverse effect and magnitude of the adverse event, should it

occur, both of which are characterised during the risk assessment process. Landis et al. (1998) emphasise the need to keep probability of harm and magnitude of effect separate in risk assessments to distinguish clearly between scenarios that have low probability but catastrophic effects and those with high probability but which have only minor consequences.

Probability is central to risk assessment. A probability of occurrence of zero equates to no risk. Similarly, if we are certain (probability of one) where and when an adverse event will happen (no matter how catastrophic the potential consequences), it is conceivable that there could be no risk, as evasive action could be taken in advance of the event. Usually the probability lies somewhere between zero and certainty, and it is the job of the risk assessor to calculate likely risks in the light of information available.

2.3 Source-pathway-receptor paradigm

Fundamental to the concept of risk assessment is the 'source-pathway-receptor' paradigm, often known as 'source-pathway-target', though this author dislikes the term 'target', it having connotations of intention, which is not usually the case. This paradigm underlines the fact that for a risk to be present there must be a source of potential risk and a pathway to a receptor. If any component is missing, then, by definition, there is no risk.

In terms of contaminated land, it is possible to reduce risks by breaking this chain. For example, take a patch of land which is earmarked for redevelopment, but is known to be contaminated by heavy metals. If the area is designated as residential with gardens, there would be a definite source-pathway-receptor linkage in place: the future householder may choose to grow vegetables in the garden, which could become contaminated with heavy metals and in turn be ingested by the householder, putting his health at risk. Three possible solutions exist. First, one could remove contaminated garden soil and replace it with clean topsoil, thereby removing the source. Alternatively, the pathway could be broken by, for example, laying a hard surface in the garden so that plants cannot be grown. Finally, the receptor could be removed, by re-designating the area for another use, e.g. a car park.

2.4 Conceptual models for risk assessment

The purpose of risk assessment is "to summarise and present scientific knowledge so that it can be used to make sound management decisions" (Landis et al., 1998). Risk

assessment methodology is an attempt to formalise the evaluation of potentially harmful situations in an objective, transparent, systematic and comparable manner. The division between 'risk assessment' and 'risk management' is an important one. Risk assessments should be carried out by a risk assessor using scientific facts in an unbiased manner. Value judgements, perceptions and politics should be left to the risk manager who will judge the scientific results in the light of supplementary subjective information in order to decide what action should be taken; a point supported by various authors (Jimenez-Beltran, 1998; Landis et al., 1998; Vegter, 1998)

Several conceptual models are suggested in the literature, which aim to provide an overarching framework for formalising risk assessments. Some examples are given below to illustrate the similarities and differences between them.

2.4.1 National Research Council (NRC) model

One of the first conceptual models to be published was in a report of the Committee on the Institutional Means for Assessment of Risks to Public Health, in the USA (National Research Council, 1983). One objective of the report was to "assess the merits of separating the analytic functions of developing risk assessments from the regulatory functions of making policy decisions". It was, indeed, recommended that there be a clear distinction between risk assessment and risk management.

The model recommends that a risk assessment consist of some or all of the following steps:

- "*Hazard identification*: The determination of whether a particular chemical is or is not causally linked to particular health effects.
- *Dose-response assessment*: The determination of the relation between the magnitude of exposure and the probability of occurrence of the health effects in question.
- *Exposure assessment*: The determination of the extent of human exposure before or after application of regulatory controls.
- *Risk characterisation*: The description of the nature and often the magnitude of human risk, including attendant uncertainty."

The results of this assessment can be fed subsequently to risk managers who may use them as an aid to decision making.

This model is still quoted widely in a number of different fields (e.g. Commission of the European Communities, 1993; Delignette-Muller & Rosso, 2000).

2.4.2 Petts et al. model

Petts et al. (1997) discuss the issue of risk assessment specifically in the context of contaminated land and propose the following conceptual model:

- "*Hazard identification*: the systematic identification of the hazards that may be associated with a site or a group of sites, considering the existing or proposed use of the site and its environmental setting.
- *Hazard assessment*: aims to understand and describe the plausible and critical pathways by which substances could reach targets, the fate of substances in the environmental media through which they are transported or move and the characteristics of the targets at risk.
- *Risk estimation*: the process of estimating the probability that an unwanted 'event' or outcome will occur under defined conditions.
- *Risk evaluation*: involves identification of appropriate standards/criteria which provide an indicator of acceptability of the risk in human or ecotoxicological terms. It also requires an understanding of the uncertainties inherent in the risk assessment output."

Similar to the NRC recommendations outlined above, these steps feed information into a risk management process which decides how best to tackle the issues raised.

The initial stages of hazard identification and hazard assessment are a sort of screening process whereby one can determine whether the aforementioned source-pathway-receptor linkage exists.

Hazard assessment considers the fate of substances in the environment, alluding to possible chemical speciation and hence alteration of the potential risk. For example, does natural attenuation occur or does speciation lead to non-bioavailability? Petts et al. (based in the UK) consider "both the existing or proposed use of the site and its environmental setting" which may stem from a relation to UK contaminated land legislation which is based on the 'fit-for-use' concept (Section 2.6.2).

Risk estimation involves exposure assessment; "the process of measuring or estimating the intensity, frequency and duration of human or other exposures to the hazards of

concern", and effects assessment; "the process of describing and quantifying the relationship between exposure to the contaminant and health/environmental effects which result". The authors point out that exposure assessment is one of the most difficult steps, being based on fate and transport models of likely movement of contaminants, as actual data is often unavailable.

Evaluation of risk, in this case, takes into account social pressures, costs and benefits, uncertainties in the estimation of the risk, possible use of the sites etc., resulting in transparent statements describing assumptions used, magnitude and effect of the calculated risks and where management of risk should be targeted. The distinction between 'risk assessment' and 'risk management', as described above, is slightly blurred.

2.4.3 Covello-Merkhofer model

Covello & Merkhofer (1993) speak of a process of 'risk analysis' which consists of three stages: hazard identification (identifying risk agents and the conditions and events under which they potentially produce adverse consequences to people or the environment); risk assessment (describing and quantifying the risk); and risk evaluation (comparing and judging the significance of the risk). These three stages provide information necessary for risk management (Figure 2.1).

Covello and Merkhofer's hazard identification essentially determines whether a source-pathway-receptor link exists and proves the necessity, or otherwise, of carrying out a risk assessment.

The Covello-Merkhofer model for risk assessment is described by four steps:

- "*Release assessment*: typically includes types, amounts, timing, probabilities of release of risk agents and description of how these may change as a result of actions or events.
- *Exposure assessment*: relevant conditions and characteristics of human and environmental exposure to risk agents – typically intensity, frequency, duration of exposure through soil, air, water, food; routes of exposure e.g. ingestion, inhalation, absorption through skin; number, nature, characteristics of people/entities that might affect consequences.
- *Consequence assessment*: typically includes specification of human fatalities, illnesses or injuries sustained under given exposure scenarios and/or

specification of ecological damage or adverse effects on natural environment under given exposure conditions.

- *Risk estimation*: integration of results to produce quantitative measures of health and environmental risks. Typically includes estimated number of people experiencing health impacts over time; nature and magnitude of adverse consequences to natural environment; probability distributions, confidence intervals and other means for expressing uncertainties in these estimates."

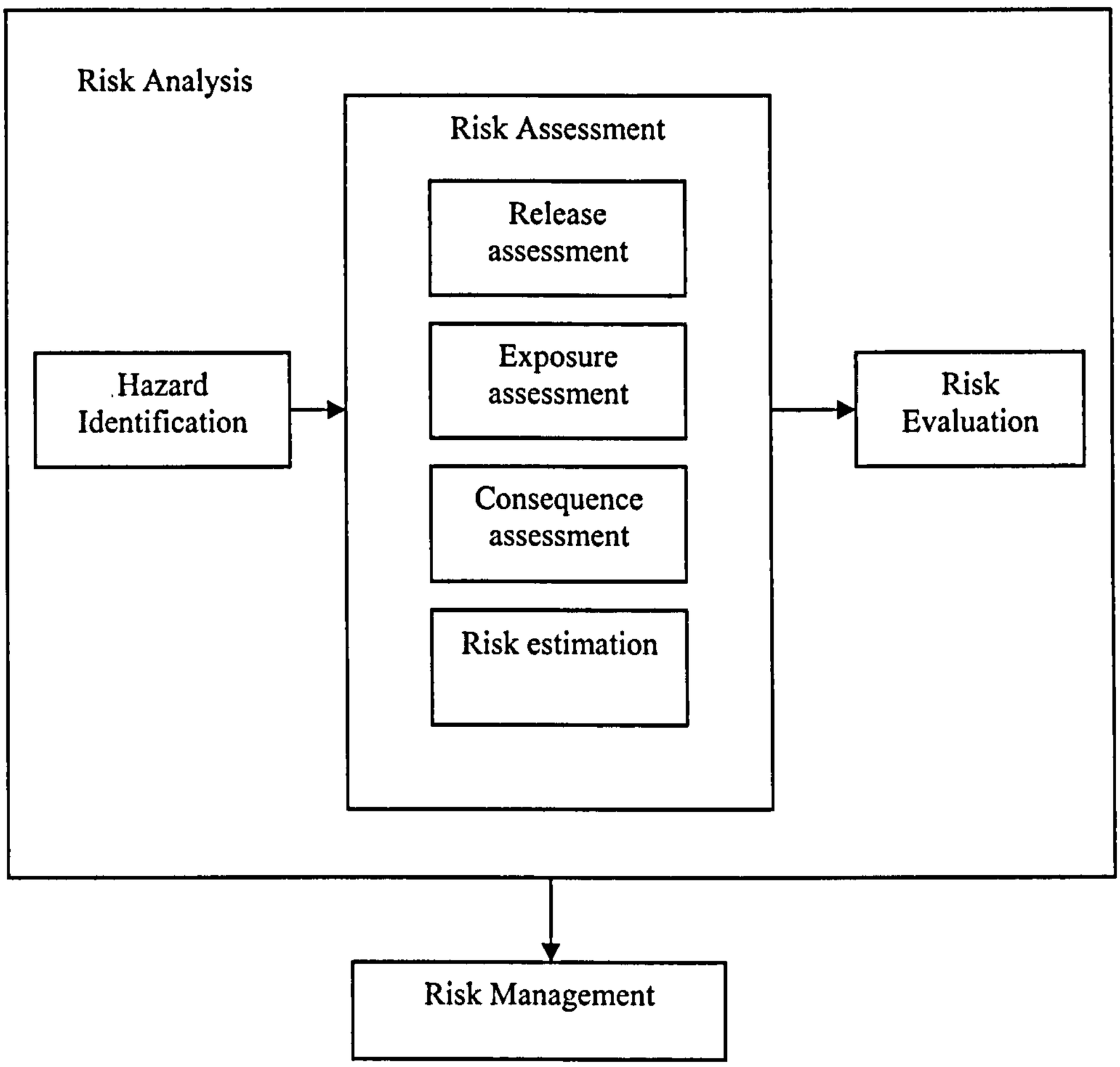


Figure 2.1. Covello-Merkhofer conceptual risk model.

To complete the risk analysis, risk evaluation judges whether the risk is significant. Often at this stage comparisons may be made with regulatory standards of 'acceptable' risk. Finally, the facts are presented to risk managers who must make decisions based on the facts presented during the risk analysis process. Note, that here, risk management ('subjective' decision making) is presented as a separate step outside the risk analysis programme (scientific and 'objective' work).

2.4.4 Ferguson et al. model

Ferguson et al. (1998) question the rationale for separating risk assessment and risk management, and argue for the need for a dialogue between the risk assessor (scientist) and risk manager (decision maker). The management of risk involves the issue of perception of risk and Ferguson et al. acknowledge that whether risk is acceptable or not is not a scientific question. The whole process must be iterative and as new information comes to light, be it hard scientific evidence of links between a chemical and a disease or information on latest public perceptions of a particular risk, it must be taken into account so that conclusions about risk can be modified (Figure 2.2).

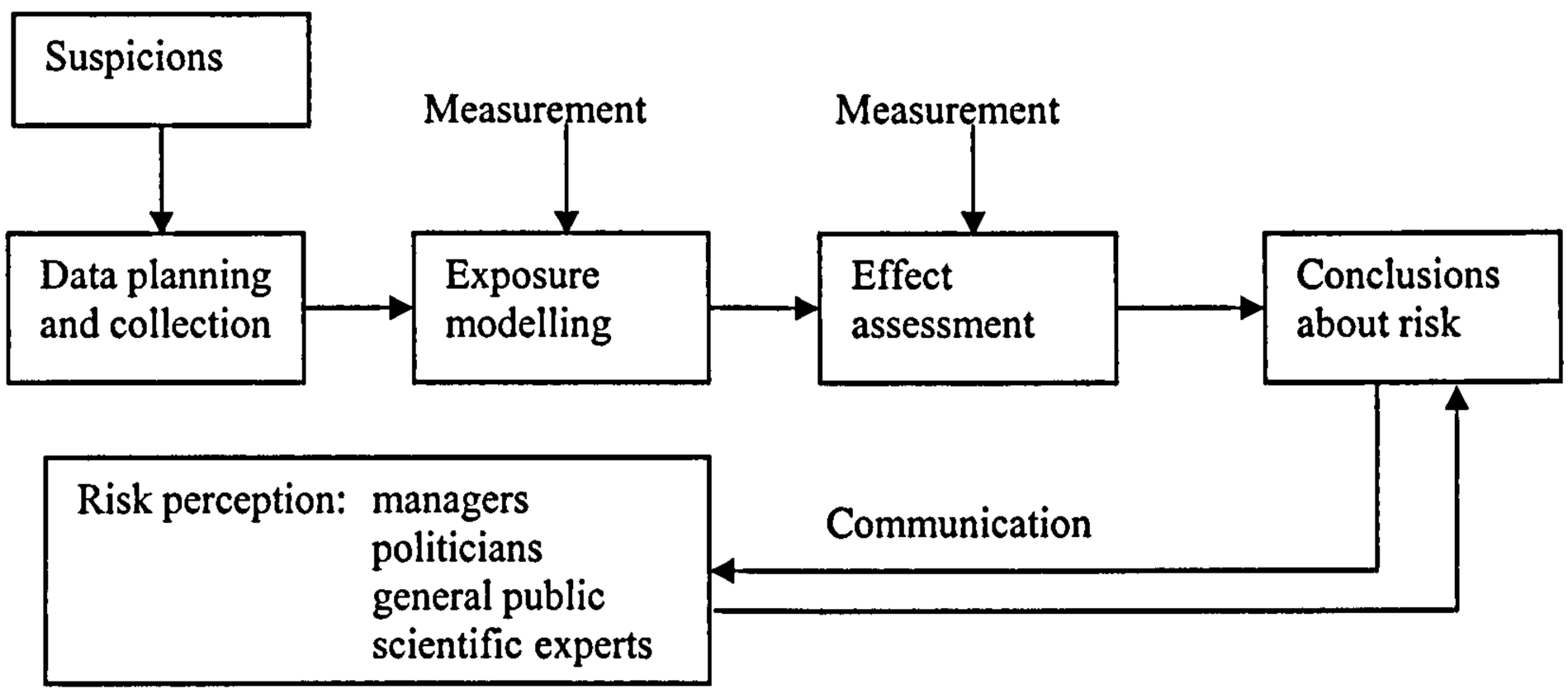


Figure 2.2. Components of a risk assessment (Ferguson et al. 1998).

It is worth emphasising here the importance of communication when practising risk assessment. Like Ferguson et al., Coleman & Marks (1999) for example, acknowledge its importance, seeing an overall risk analysis which links risk assessment with risk communication and management: "The starting point of a risk analysis, however, need not be a risk assessment. Rather in the US, risk communication is receiving increasing attention as the starting point for risk analyses....Dialogue with stakeholders....is increasingly viewed as essential throughout the entire risk analysis process."

The similarities and differences in the conceptual models are subtle. All contain the same key stages of assessment. The main difference appears to be debate about the risk assessment/risk management boundary and the communication of risk.

2.5 *Uncertainty and variability*

Natural phenomena are variable and measurements of their attributes may be uncertain. Uncertainty refers to a state of ignorance about a particular value (e.g. the true body weight of a 30 year old male individual) and variability refers to a measurement that can truly take more than one value (e.g. the body weights of all 30 year old males). Uncertainty can often be reduced by the collection of more information, but variability in a fixed population cannot. Uncertainty (and variability) lead to uncertainty in the results of analyses. This is a long known problem and it is usual for scientists to quote their results with some sort of 'confidence interval', acknowledging the fact that results may be subject to error. Risk assessments should be no different.

2.5.1 Qualitative and quantitative risk assessment

It was usual in the past for risks to be quoted in qualitative terms; high, medium or low risk sites for example. This is still a useful tool for risk ranking exercises, but nowadays more quantitative risk assessments are expected, such as quoting the number of additional cancer deaths per one million population, or the probability of an adverse event. It is at the point of quoting quantitative results that the issue of uncertainty and variability becomes important.

2.5.2 Deterministic versus probabilistic risk assessment

A risk assessment may involve analysis of many different sources and pathways for each receptor and calculations may require the estimation of tens of variables. For example, a contaminant may be inhaled in dust, ingested in food and water, and absorbed through the skin. The amount of exposure to the contaminant depends upon life style, such as time spent in the vicinity of the contaminant sources and media through which they travel. The effect on the receptor could depend upon age, body size and state of health. As the picture is built, so the number of variables to be estimated increases. Each variable will have an estimated distribution of possible values.

Risk assessment methodologies can be classified broadly into two groups; deterministic and probabilistic. Consider a risk which is the function of three independent variables A, B and C such that $Risk = A \times B \times C$. A deterministic assessment uses point estimates, such the mean or 95th percentile, for each variable in the model in order to calculate a point estimate of risk. This results in a single figure, which, for example, may be labelled 'typical' or 'reasonable maximum' with no indication of the range of

risks. In a probabilistic/stochastic assessment, the point estimates are replaced by distributions of possible values for the variables. Results can be presented in the form of a distribution of possible risks within an exposed population, indicating maximal and minimal risks that might be experienced by different individuals, and the relative likelihood of intermediate risks between those two extremes.

2.5.2.1 Deterministic risk assessment

Deterministic assessments can be misleading. For example, in Canada, at the Federal level, risks are calculated for the 'average' or 'typical' individual in a deterministic assessment (Richardson, 1996). In this case it is possible that the average is below the stipulated 'safe dose', leading to a declaration that there is no risk, whilst in actual fact a high percentage of the population actually may be exposed at levels above the safe dose. On the other hand, the USEPA stipulated calculation of risk for the 'reasonable maximum exposure' defined as "the highest exposure that is reasonably expected to occur at a site" (USEPA, 1989) which too can have its problems. For example, taking the 95th percentile value for each variable A, B and C, where all have lognormal distributions with equal variances, will give a risk result above the 99.8th percentile of the entire distribution of risk to the population (Hattis & Burmaster, 1994), potentially leading to a wildly exaggerated impression of the risk. By itself, any single point or summary statistic from a distribution of risk can be misleading.

This is not to say that deterministic assessments are totally inadequate. As an initial screening tool, for example, they are highly valuable (Moore & Elliott, 1996). If a deterministic assessment using the upper 95th percentiles calculates a risk level well below the threshold for action, then there is absolutely no point in pursuing a more complex assessment. It is recommended that deterministic assessments be used as part of a tiered programme which proceeds from simple deterministic assessment through to more complex probabilistic assessment, depending on the outcome of the deterministic assessment e.g. USEPA (1997c; 2001) and see also Section 2.7.

2.5.2.2 Probabilistic risk assessment

Currently, there is a trend towards a more complete characterisation of risk using stochastic modelling techniques which take into account as many sources of uncertainty and variability as possible (Rai et al., 1996). The propagation of uncertainty and variability throughout the assessment produces a more transparent assessment, acknowledging that within a population there will be a range of risks experienced.

Probabilistic assessments are not without their problems. The biggest issue faced by the probabilistic assessor is quantifying and characterising the input distributions. Several authors support the development and use of probabilistic methods, but do so with an attached 'health warning' (Moore & Elliott, 1996; Richardson, 1996; Smith, 1994): Richardson points out that probabilistic assessment does not eliminate the rampant use of assumptions and professional judgment in risk assessment in the face of lack of knowledge, while Smith states that to improve the accuracy of probabilistic risk assessments, risk assessors will need to collect and verify distributions for many currently undescribed input assumptions.

The issue of assumed distributions has been recognised and addressed by some regulatory authorities. The Environmental Health Directorate of Health Canada has initiated the development of standard, accepted probability density functions for exposure factors such as body weight, food consumption, etc. (Richardson, 1996). The Exposure Factors Handbook has been published in the USA (USEPA, 1997a) and in Europe, the Exposure Factors Sourcebook for European Populations (with Focus on UK Data) is now available (ECETOC, 2001), covering all manner of distributions such as bodyweight, age, weekly hours spent away from home, fish and meat consumption etc.

Whilst probabilistic methodology serves to make risk results more transparent, it can also make decisions more complex. It is easier from both a risk management and legal standpoint to make a case for regulating a 'precisely known' hazard than for regulating a hazard with a very uncertain risk (Bailar & Bailer, 1999). There is some concern that without an equally sophisticated framework for risk management and communication, i.e. guidance on how to use these probabilistic results, the information will be used incorrectly, if at all (Griffin et al., 1999; Hope, 1999).

For a fully probabilistic assessment, the 'acceptable risk' levels could also be represented by probability distributions rather than point values. In calculating the point value, several fixed 'uncertainty factors' are used to ensure a conservative acceptable limit. In the probabilistic method, uncertainty factors are replaced by distributions, in a similar way to the risk assessment itself. Several methods for their calculation have been proposed (Baird et al., 1996; Kodell & Gaylor, 1999; Leisenring & Ryan, 1992; Slob & Pieters, 1998; Swartout et al., 1998). Again, problems arise in agreeing which distributions to use and current practice still uses simple thresholds.

2.5.2.3 Partitioning uncertainty and variability

The state-of-the-art in risk assessment research not only seeks to propagate the range of possibilities but seeks to separate 'uncertainty' from 'variability' (Hoffman & Hammonds, 1994; Nauta, 2000; Rai & Krewski, 1998). The purpose is to highlight areas where additional work will serve to reduce uncertainty and increase the reliability of the risk estimates.

The separation of uncertainty and variability further goes to complicate probabilistic risk assessments. It can be difficult to separate and characterise them as often the form of the distribution and the parameters of the distribution are difficult to characterise. There is considerable judgement involved in determining where separation is necessary. Despite USEPA policy stating that "for both input and output distributions, variability and uncertainty are to be differentiated where possible" (USEPA, 1997c), Kelly & Campbell (2000) argue that separation is not imperative. If an assessment gives a result greatly above or below the critical threshold, then why bother to do a probabilistic assessment, let alone separate uncertainty and variability, as this adds no value. The separation of uncertainty and variability should be handled in the appropriate place in a tiered risk assessment (Section 2.7) if necessary.

2.6 Contaminated Land

2.6.1 Definition of contaminated land

Just as there is more than one definition of 'risk', there is more than one definition of 'contaminated land'. One oft-quoted definition of contaminated land, e.g. Harris & Herbert (1994) and Petts et al.(1997), is from the NATO committee of challenges to modern society: "land that contains substances that when present in sufficient quantities or concentrations, are likely to cause harm directly or indirectly, to man, to the environment, or on occasions to other targets".

A definition preferred by this author is one from the UK. This is a legal definition of contaminated land which, aptly, is based on the principle of risk assessment and is found in section 78A(2) of Part IIA of the Environmental Protection Act 1990:

"...any land which appears...to be in such a condition, by reason of substances in, or under the land, that:

a) significant harm is being caused or there is a significant possibility of such harm being caused; or

b) pollution of controlled waters is being or is likely to be caused" (DETR, 1999). Land is not defined in this act, but is defined in schedule 1 to the Interpretation Act 1978 to include "buildings and other structures, land covered with water, and any estate, easement or servitude or right in or over land" (Tromans & Turrall-Clarke, 2000). These authors also state that land would "in principle include sub-soil to any depth; disused mine shafts have sometimes been the recipients of contaminating material in the past."

The 'significant possibility' of harm being caused can be deduced by risk assessment methodology.

2.6.2 Legislation and regulation

Contaminated land is a major problem in many developed countries. Most are blighted to some extent by contaminated land left from past industrial activities. In the spirit of sustainable development in some cases, fear of litigation in others and shortage of space for new development in others, the authorities are enforcing clean up of contaminated sites to one degree or another.

Within the EU, local country legislation is often dictated by EU Directives. However, there are no Directives which specifically cover contaminated land (Busing, 2001). Several impinge on the problem, but are primarily directed towards the protection of water resources, e.g. Water Framework and Landfill Directives (Council of the European Union, 1999; 2000b). Note that the definition of contaminated land used in the UK (Section 2.6.1) includes protection of controlled waters.

Contamination can be 'historic' or 'current'. Historic contamination is an unfortunate legacy, which we cannot, by definition, prevent. Current contamination ought not to occur, since we better understand the effects and are able to put in place preventative measures to avoid it. If new contamination does occur, it should be removed completely. The Integrated Pollution Prevention and Control (IPPC) Directive is a proactive rather than a retrospective directive aimed at prevention of contamination incidents (Council of the European Union, 2000a).

Recently, the significance of soil as a non-renewable resource has been recognised and the degradation of soil, by whatever means, has been acknowledged as a serious problem across the EU. As a result, the European Parliament called on the European Commission to produce a thematic strategy for soil protection by July 2004, to put an end to (among other things) "the accumulation in soil of substances which pose an

environmental and health hazard" and "reversing the alarming trend towards...contamination of soil" (European Parliament, 2002). The UK produced its first soil action plan in March 2004, which includes continuation of the development of soil guideline values for the protection of human health from contaminants in soil (DEFRA, 2004).

In terms of dealing with historic contamination, there are two broad alternatives for cleaning the land: the 'multifunctional' approach; and the 'fit-for-use' approach. In the past, countries such as Denmark and the Netherlands adopted the multifunctional approach for cleaning up contaminated land (Fairman et al., 1998). This involved cleaning all contaminated sites to such a degree that they would be suitable for any chosen use. However, this could prove extremely expensive. For example, by 1995 the Netherlands had identified 300,000 contaminated sites which would have led to a bill of €13 billion if cleaned to multifunctional standards (Ferguson, 1999). Such situations have led to a shift towards more pragmatic 'fit-for-use' policies, and indeed in the Netherlands, currently, 'immobile' soil contaminants are now treated under fit-for-use policies (Lijzen & Swartjes, 2000).

The UK was one of the first countries to adopt pragmatic fit-for-use policies. For instance, planning permission for development of contaminated land is given subject to clean-up of the site to a standard required for that particular development type. For example, land earmarked for a car park may not require remediation work as the hard surface will cap the contamination and act as a barrier, breaking the source-pathway-receptor chain. If at a later stage the land was sold on for redevelopment as housing with gardens, the source-pathway-receptor chain will be recreated, therefore planning permission will only be given subject to a clean up of the contaminated soil which is currently beneath the car park. The contaminant level permitted in a domestic garden is much less than that permitted for non-domestic uses.

'Fit-for-use' is a much more sustainable policy because, in effect, it allows a much greater regeneration and recycling of contaminated land. Costs are kept to a minimum whilst maintaining adequate protection of the population and environment, allowing a greater number of sites to be brought back into use.

2.6.3 Guideline values

Because a full blown risk assessment can be an expensive operation, most countries support the use of generic values (variously defined as screening, threshold, trigger,

guideline or action values) as a tool for use in generic assessments of contaminated land. Care must be taken to understand exactly what these values are, as different authors and governments use different terminology for the same or similar values.

In summary, information about a site can be compared to guidelines of 'safe' contaminant levels or 'acceptable' risks. This provides a method for screening out sites that pose no threat, and a means of ranking potentially harmful sites by degree of likely danger. Further, values are set which indicate when remediation is absolutely required as the risk posed by contaminant levels of this size is deemed too high. Figure 2.3 is a conceptual summary by Petts et al. (1997).

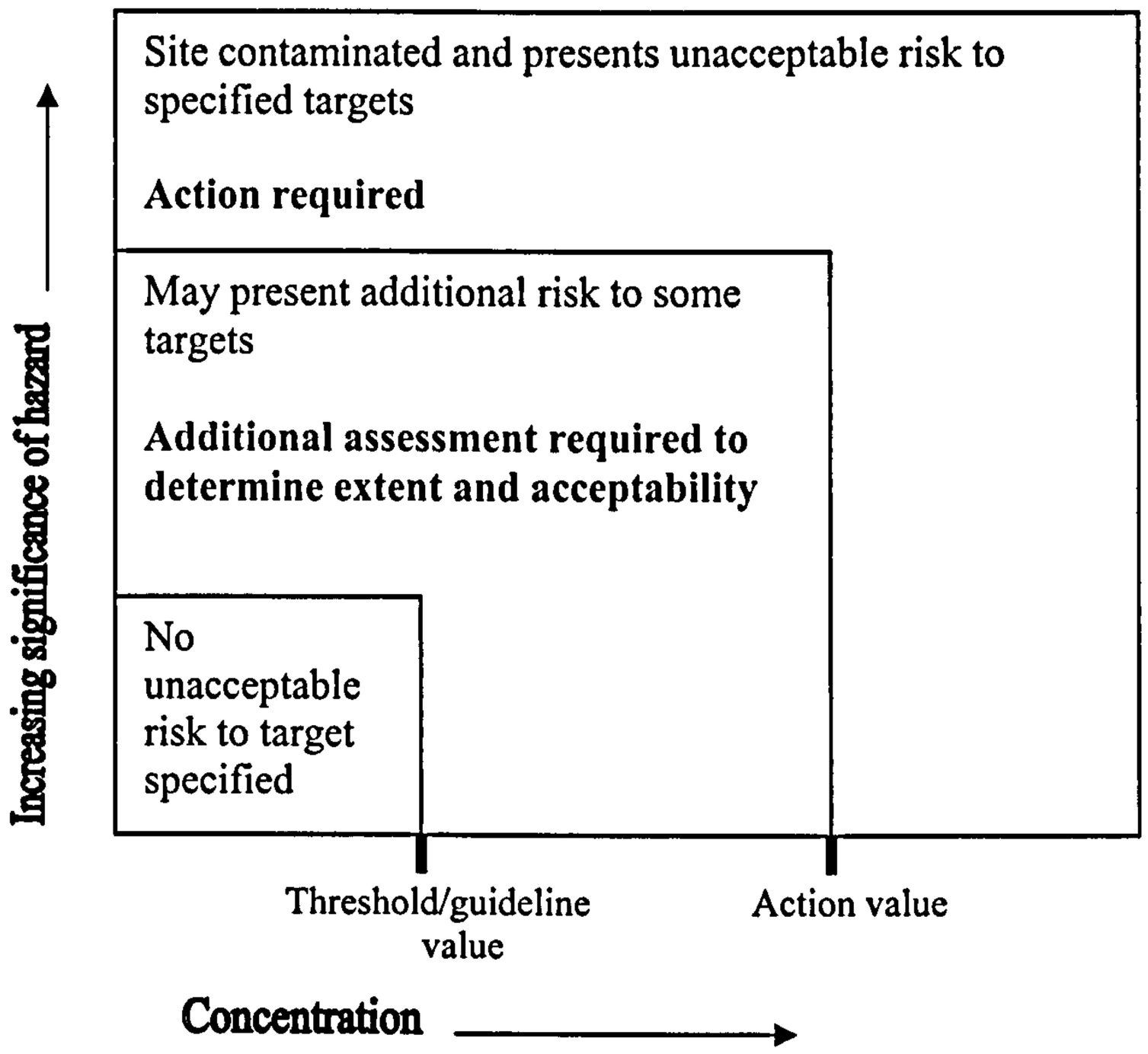


Figure 2.3. Threshold/action concept (Petts et al., 1997).

An example of this is the Dutch system for assessing land contamination. Guideline permissible contaminant levels are determined which are based on assessment of potential risks to ecosystems and humans that would occur under standardised (generic) conditions. The assessment against these guideline levels will determine whether remediation is required to reduce potential risk and if so how urgently it is required. Figure 2.4 illustrates the system (Swartjes, 1999).

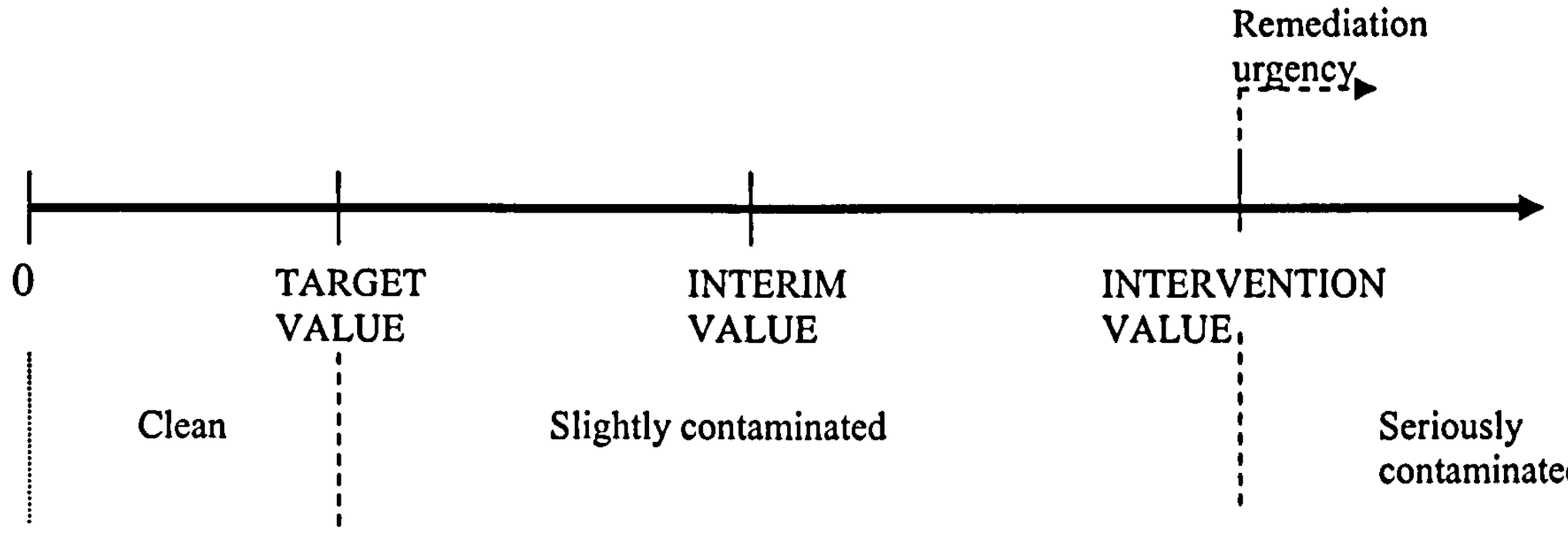


Figure 2.4. Soil and groundwater quality standards and assessment (Swartjes, 1999)

Levels of contaminant below the target value are assessed as clean and require no further action. Values between the target and interim values (the interim value is merely an average of the target and intervention values) also require no further investigation, although some restrictions may be placed on the use of the soil. Between the interim and intervention values, further action is stipulated based on actual risk as determined now by the site specific conditions. After this investigation, if the level is still below the intervention value restrictions may be applied but serious remediation is not necessary. If the contaminant level is above the intervention value, remediation is definitely required as the site is classed as seriously contaminated.

The Dutch intervention values were originally implemented within the framework of the Dutch Soil protection act in 1994. However, in the light of new scientific information, a recent report was commissioned to evaluate and update these values and new Serious Risk Concentrations (SRCs) were derived (Lijzen et al., 2001), to be used as intervention values.

In the past, the guideline values were often based on 'professional judgement', but are increasingly being set using risk estimation models (Ferguson et al., 1998). For example, in the UK, the Interdepartmental Committee on the Redevelopment of Contaminated Land (ICRCL) produced a set of 'tentative' threshold and action values for some contaminants (ICRCL, 1987). However, in 1990, the House of Commons Select Committee on the Environment called for 'more scientifically based guidance' (DEFRA & EA, 2002a) and over the next decade, new guideline 'intervention' values were developed using a probabilistic risk assessment model 'CLEA', which was finally published in 2002 (DEFRA & EA, 2002g). The derivation of the above-mentioned

Dutch SRCs was also made using a risk model. Human toxicological maximum permissible risk levels were used in conjunction with the 'CSOIL' exposure model (Lijzen et al., 2001).

Care must be exercised in employing all such values and effort must be put into understanding their derivation in order to ensure their correct usage. For example, exposure assessment is extremely important and the existence of the source-pathway-receptor chain must be intact for risk to exist. Equally high levels of contaminant may be a threat at one site but not at another, because a pathway does not exist. Such intricacies will not be picked up by using generic guideline levels in isolation without probing deeper into the site specific details. A balance must be struck whereby the use of guideline values is complemented by more site specific studies, bearing in mind the pressures of time and cost, perhaps as part of a tiered risk assessment.

2.7 Tiered approach to contaminated land risk assessment

In the UK, a tiered approach to risk assessment is advocated "where the level of effort put into assessing each risk is proportionate to its priority (in relation to other risks) and its complexity (in relation to understanding of the likely impacts)" (DETR et al., 2000). A risk assessment procedure is carried out at each tier (Figure 2.5), but is simple at the lower tier and becomes more complex at upper tiers. This enables resources to be targeted to where they are required most: if a simple Tier 1 assessment indicates that there is no risk (for example the source-pathway-receptor chain is not intact), there is no point in wasting resources in carrying out a complex, quantitative risk assessment. A similar tiered approach is also recommended in the US (USEPA, 2001).

Though this methodology is advocated in the UK, there is a lack of suitable tools for undertaking such a process (Butler & Petts, 2000). A paper-based, qualitative, UK designed ranking system is available (CLR6: Prioritisation & categorisation procedure for sites which may be contaminated) for a Tier 1 assessment, and the recently released SGVs may be used in generic Tier 2 assessments, but there are no UK based models for a Tier 3 quantitative risk assessment. Therefore, models designed for use in the US (e.g. Risk*Assistant and RBCA) and the Netherlands (Risk-Humaan) are most commonly used in the UK (Butler & Petts, 2000), though their underlying assumptions are not always relevant to the UK situation. These three models are deterministic and use either 'average' (e.g. MLE, most likely exposure) or more conservative maximum

(e.g. RME, reasonable maximum exposure) results as indications of the risk to the population at the contaminated site.

The recently published CLEA model (DEFRA & EA, 2002g) is a probabilistic model, but in its published format is only for use in back-calculating soil guideline values for a specified land use with a critical receptor. It is suitable for a generic Tier 2 assessment but not for 'forward calculating' risks, as would be required in a Tier 3 assessment.

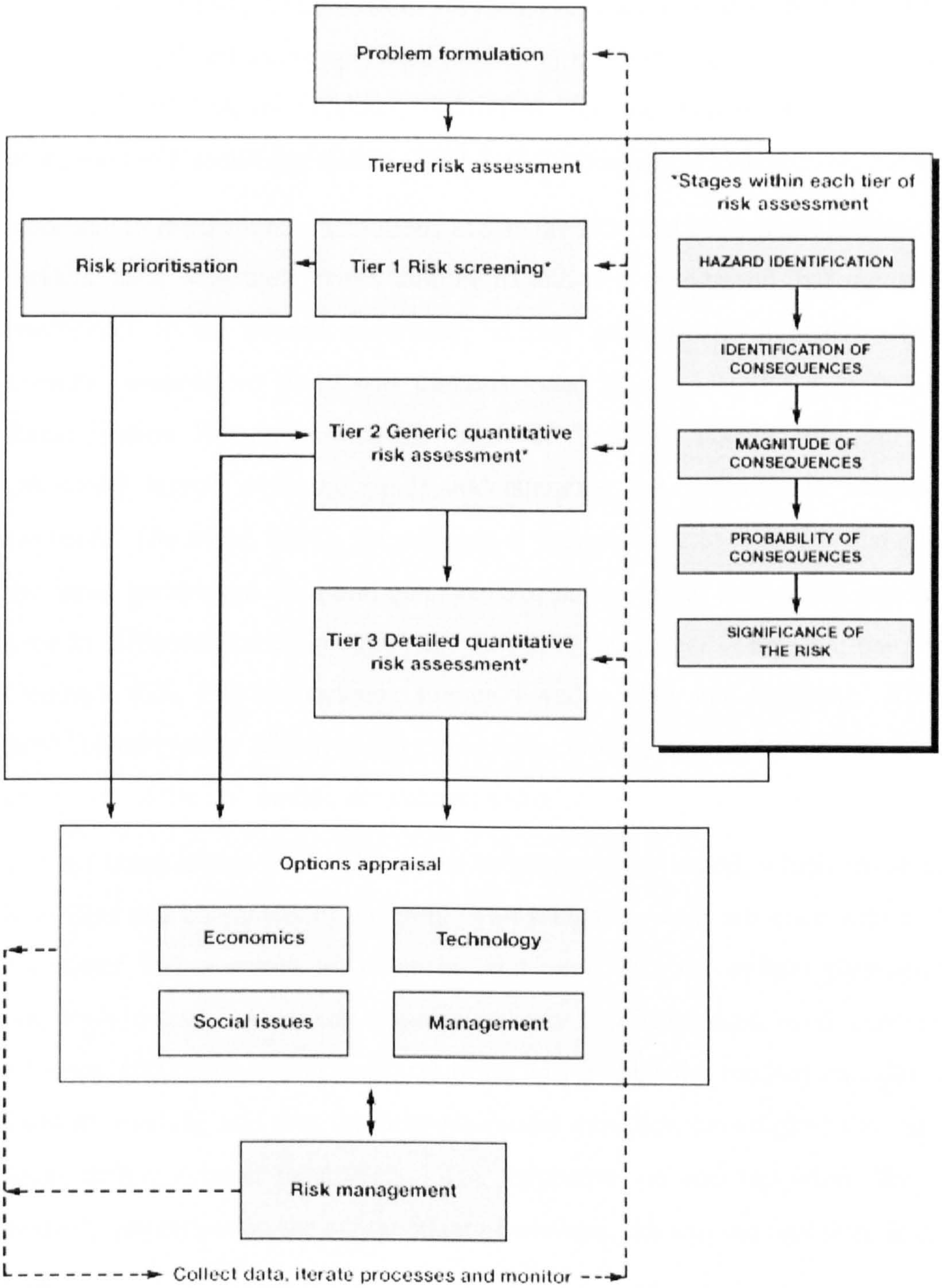


Figure 2.5. Tiered approach to risk assessment (DETR et al., 2000).

2.8 Exposure models

As briefly mentioned in Sections 2.6.3 and 2.7, regulatory standards and actual exposures/risks from contaminated land in European countries and the US are often calculated by means of mathematical models. Swartjes refers to widespread implicit use (i.e. comparison of measured contaminant concentration with soil quality standards which are based on the exposure models) and explicit use (i.e. decision-making based on site-specific exposure calculations) (Swartjes, 2001). Across the EU, policy and practice vary greatly, ranging from very little regulation in some member states, through to highly regulated use of specified models in Belgium (Ferguson, 1999) where the use of VLIER-HUMAAN (Flemish instrument for the evaluation of human risks) is embedded in Flemish legislation (Smit & Goldsborough, 1998).

Research in independent institutions across the EU (and elsewhere) produces individual models, each with their own assumptions and it is recognised that model uncertainty contributes to the overall uncertainty in risk assessment outcomes. An EU cross-country comparative study was commissioned by CLARINET (Contaminated Land Rehabilitation Network) with the aims of "gaining insight into the variation in calculated human exposure [and] understanding the variation in calculated human exposure" (Swartjes, 2002). In addition, it was possible to overview default values for the input parameters used and general information about the human exposure models used in different European countries. Seven models were compared, one from each of Denmark, UK, The Netherlands, France, Sweden, Italy and Belgium. All models are used for derivation of soil quality standards, among other purposes, and all, with the exception of the UK model, are deterministic.

For the comparative study, 20 model scenarios were tested, which involved different land uses and chemicals of concern. The scenarios were run once with a fixed set of parameter values across all countries and once with the default parameters for that country's model. One key conclusion was that calculated total exposures ranged between 100 times higher and 100 times lower than the median calculated exposure from all models, and that the between-model variation outweighed the variation from using different input parameters. For 'exposure via soil ingestion' the variation is scarcely dependent on the contaminant of concern, the soil use and even less on the soil type.

2.8.1 Detailed example of an exposure model

In March 2002, DEFRA and the EA published a series of long-awaited reports and a computer model concerned with the assessment of risks to human health arising from exposure to contaminants in soil in the UK. The series comprised the reports known, for short, as CLR7-10, TOX1-10 and SGV1-10 (more detail is given in Chapter 6). The series of reports focuses around the derivation of Soil Guideline Values (SGVs) using the Contaminated Land Exposure Assessment (CLEA) model (DEFRA & EA, 2002g). SGVs are levels of contaminants in soil ($mg.kg^{-1}$) which are considered 'safe' in particular circumstances. These guideline values are calculated in a probabilistic manner such that they would be protective of 95% of the exposed population, thus protecting the vast majority of people while accepting that there may be some extreme high exposures which will not be prevented. During the process, the Average Daily Exposure (ADE) for 'critical receptors' is calculated. According to the CLEA model, exposure occurs via three routes: inhalation *(inh)*; dermal absorption *(derm)*; and oral intake *(oral)*, and can be summarised by the following equation:

$$ADE = \frac{(IR_{inh} \times EF_{inh} \times ED_{inh})}{BW \times AT} + \frac{(IR_{oral} \times EF_{oral} \times ED_{oral})}{BW \times AT} + \frac{(IR_{derm} \times EF_{derm} \times ED_{derm})}{BW \times AT}$$

where

ADE: average daily exposure to a chemical from soil ($mg.kg^{-1}_{BW}.day^{-1}$)

IR: chemical exposure rate ($mg.day^{-1}$)

EF: exposure frequency ($day.yr^{-1}$)

ED: exposure duration (yr)

BW: bodyweight (kg)

AT: averaging time (day)

In turn, the *IR* terms are made up of sub-equations which model how the exposure occurs in each pathway. For example, the oral pathway is built up from the following equations:

$$IR_{oral} = IR_{dsi} + IR_{isi} + IR_{veg}$$

where

IR_{dsi}: chemical exposure rate from direct (inadvertent) soil ingestion ($mg.day^{-1}$)

IR_{isi}: chemical exposure rate from indirect ingestion of soil attached to vegetables ($mg.day^{-1}$)

IR_{veg}: chemical exposure rate from ingestion of contaminants absorbed into vegetables ($mg.day^{-1}$)

The *IR* sub-equations are made up as follows:

$$IR_{dsi} = C_{soil} \times SDR$$

$$IR_{isi} = C_{soil} \times \sum_{veg\ type}\left(CR_{veg} \times BW \times HF_{veg} \times SL_{veg}\right)$$

$$IR_{veg} = C_{soil} \times \sum_{veg\ type}\left(CR_{veg} \times BW \times HF_{veg} \times CF_{veg}\right)$$

where

C_{soil}: contaminant concentration in soil ($mg.g^{-1}$)

SDR: average daily soil and dust ingestion rate ($g.day^{-1}$)

CR_{veg}: consumption rate per vegetable ($g.FW.day^{-1}$) (FW=Fresh weight)

BW: bodyweight (kg)

HF_{veg}: fraction that is home grown per vegetable (-)

SL_{veg}: amount of attached soil per vegetable ($g.g^{-1}FW$)

CF_{veg}: soil to plant concentration factor ($\mu g.g^{-1}FWplant/\mu g.g^{-1}DWsoil$) (DW=Dry weight)

Of the inputs to the model, SDR (some age categories), BW, CR_{veg} and HF_{veg} are treated probabilistically, while C_{soil}, SL_{veg} and CF_{veg} are treated deterministically.

2.9 State-of-the-art versus current practice

As discussed above, current research into methods of probabilistic risk assessment tend towards separating uncertainty and variability. Several methodologies may be used for propagation of uncertainty and variability through a risk model, for example: Monte Carlo simulation (Hattis & Burmaster, 1994; Hoffman & Hammonds, 1994); probability trees (Sielken & Valdez-Flores, 1999); response surface methods (Isukapalli et al., 1998); and so on. The most popular appears to be Monte Carlo simulation for which USEPA has produced guidance (USEPA, 1997b; 1999), and which is also used in the UK CLEA model. Its popularity may be due to the fact that the method application is relatively simple and has been available for many years. Monte Carlo methods and probability trees are more traditional, with stochastic response surface methods emerging at present.

Dakins et al. (1994) advocate the use of probabilistic methodology and have gone so far as to attempt to quantify, in monetary terms, the 'expected value of including

uncertainty'. They calculated savings made as a result of a reduction in the volume of land to be cleaned up by considering a more realistic probabilistic calculation rather than a deterministic calculation.

Research, naturally, is far ahead of current practice. Roberts (1999) summarises the practical issues with probabilistic risk assessments and highlights the gap between state-of-the-art in science and practice: probabilistic risk assessment is difficult to understand, and difficult to implement in practice in the field. Despite the increased complexity of management decisions under probabilistic risk assessment, public bodies and policy makers are tentatively moving forward on this topic, as demonstrated, for example, by the USEPA's policy statement and code of practice for conducting probabilistic risk assessments for contaminated sites in the USA (USEPA, 1997c; 1999) and the above-mentioned development of acceptable distributions for variables in North America and Europe (ECETOC, 2001; Richardson, 1996; USEPA, 1997a). In the UK, the probabilistic CLEA model has been introduced for developing new guidelines for 'safe' soil pollutant concentrations.

2.10 Summary

The importance of the source-pathway-receptor paradigm and probability in risk assessment has been highlighted. Conceptual models for the procedure are subtly different, though all contain the same basic elements: identifying that the source-pathway-receptor linkage is intact before assessing the size of the risk to the population/environment. Uncertainty and variability have been recognised as important factors in risk assessment and should be quantified for a transparent assessment, but only if the situation merits the additional work. For contaminated land risk assessment, this decision should be part of a tiered risk assessment procedure which starts with a simple assessment and only progresses to more complex assessments should the situation merit it.

A key feature in a contaminated land risk assessment is the concentration of the contaminant in the soil. The next chapter introduces methods for estimating and assessing the contaminant concentration in a way which preserves the spatial distribution of the concentrations and would be of use in a risk assessment.

Chapter 3 Geostatistics in the context of risk assessment and contaminated land

3.1 Introduction

This chapter introduces geostatistics in the context of contaminated land and risk assessment. The fundamental concepts of geostatistics are discussed, followed by a discussion of the practical considerations of using geostatistics for this type of problem. Kriging and simulation are outlined, and the literature is reviewed, and pros and cons of available methods are discussed. Finally, the theory behind the two methods chosen for use in this thesis is laid out.

Typically, in many contaminated land assessments, the assessor is forced to work with limited field samples due to resource limitations in terms of time, manpower and cost. The challenge is to glean as much information as possible from these limited samples as to the amount and distribution of contamination across the site. The development of the methodology in this research means that average values across the area are not of particular interest; rather the interest lies in knowing exactly where the hotspots of contamination occur and how confident we are in these judgements. Thus we need to be able to estimate contaminant concentrations at unsampled locations and attach to them some measure of uncertainty.

There are two levels at which risk from contaminated land may be calculated (Chapter1, Figure 1.1). First, one may quantify the contaminant level in the soil and compare it to a regulatory cut-off level deemed 'safe' or 'acceptable', and second, one may calculate the human intakes of soil contaminants and compare them to regulatory 'safe' or 'acceptable' intake levels. In each case, the contaminant level needs to be quantified

and a measure of the uncertainty in the assessment be preserved. This may be achieved using geostatistics.

3.1.1 Introduction to geostatistics

Spatial assessment of phenomena has been carried out for many years using relatively simple and cheap to implement interpolation methods such as inverse distance weighting or nearest neighbour techniques. However, during the past 50 years a suite of more sophisticated interpolation methods has developed under the umbrella of the now well known subject area of geostatistics.

Geostatistics emerged in the 1950s South African mining industry as a result of trying to solve a very real problem of estimating ore grades from very limited sample data. Mining works on a selective cut-off basis whereby blocks are mined, or not, depending on whether the estimated grade is above or below a critical threshold which is deemed economically viable. Geostatistical techniques were developed to better this estimation and were pioneered by D.G.Krige. Since then, the field has grown phenomenally, both in terms of the suite of methods available and the fields in which they are applied. As well as mining, the techniques have been adapted and developed to solve spatial analysis problems in fields as diverse, for example, as petroleum engineering, forestry, ecology, fisheries, farming and, last but not least, contaminated land investigation.

Earth scientists have long observed that samples in space are more likely to be similar the closer together they are. Geostatistics hinges around the concept that points in space are somehow related to one another and that this relationship can be modelled. The process starts with a traditional statistical analysis and posting of the data, followed by variography, a process whereby the dissimilarity of samples is plotted as a function of their separation distance. If a structured relationship is uncovered, this can be modelled and the model used in a process called kriging, which estimates values at unsampled locations. Kriging is basically a moving window weighted-least-squares estimator, which weights nearby samples according to weights determined from the variogram model, and so exploits the spatial relationship between the samples. Hence, from few original sample data, estimations may be made at all locations.

The theory behind geostatistics, in its many guises, is dealt with rigorously in texts such as Chiles & Delfiner (1999), Goovaerts (1997b), Isaaks & Srivastava (1989) and Webster & Oliver (2001). Here, the fundamental concepts are outlined followed by an

exploration of the areas most pertinent to this research and the justification behind the choice of method used.

3.2 Fundamental concepts

3.2.1 The regionalised variable

The term 'regionalised variable' was coined by Matheron to emphasise two apparently contradictory aspects of these types of variables: a random aspect, which accounts for local irregularities, and a structured aspect, which reflects large scale tendencies (Armstrong, 1998). An ability to model these large scale tendencies allows for spatial interpolation.

Traditional statistical theory is based on the concept of the independent random variable, whereby an outcome of a trial is independent of all others and occurs with some probability. Geostatistics, however, is based on the concept of the regionalised variable, $Z(x)$, whereby for pairs of samples in space their 'outcomes' $z(x)$ are not independent - separation distance h and sample value $z(x)$ are correlated. h is actually a vector quantity, hence has both a distance and direction. Geostatistics quantifies the relationship between separation and sample value using a form of bivariate statistics in which a model is developed to describe this relationship.

The observed value z at each data point x is considered as the outcome $z(x)$ of the random variable $Z(x)$. The family of all these random variables is called a random function. A random function bears the same relation to one of its realisations as a random variable does to one of its outcomes, except that the realisation of a random function is a function whereas the outcome of a random variable is a number. A random function is characterised by its finite dimensional distributions, i.e. by the joint distributions of any set of variables $Z(x_1)$, $Z(x_2)$,....$Z(x_n)$, for all n and for all points x_1, x_2,....x_n (Armstrong, 1998).

The random variables $Z(x)$ are location dependent, thus their inference requires repetitive realisations at each location x, which is never available in practice (Goovaerts, 1997). Hence, we use all pairs of measurements a vector h apart within the study area A, $[z(x_i), z(x_i+h);\ i=1,...,n]$, as a set of repetitions. The implicit assumption is that the corresponding pairs of random variables $[Z(x_i), Z(x_i+h);\ i=1,...,n]$ originate from the same distribution. Such pooling of data pairs, regardless of their locations calls for the

phenomenon under study to be spatially homogeneous within *A* (Goovaerts, 1997); a condition known as stationarity.

3.2.2 Stationarity

It should be emphasised that stationarity is a property of the random function model, and is not a characteristic of the phenomenon under study. Stationarity is a decision made by the user. The user must ask "is a stationary model of the soil realistic?" (Webster, 2000) and if the decision is 'yes' the geostatistical analysis proceeds under this assumption.

The random function is said to be stationary within the study area *A* if the multivariate distribution is invariant under translation (Goovaerts, 1997) i.e. any two vectors of random variables $\{Z(x_1),...,Z(x_n)\}$ and $\{Z(x_1+h),...Z(x_n+h)\}$ have the same multivariate distribution function whatever the translation vector *h*. In other words, regardless of their locations, all pairs of random variables separated by a particular *h* have the same joint probability distribution and their moments are the same everywhere. In practice, we consider different degrees of stationarity.

If all moments are constant, we have full stationarity. If only the first and second moments (mean and variance or covariance) are constant we have second order (or weak) stationarity, whereby the mean is constant and the covariance is constant for a given separation *h* and does not depend on the absolute position of the pairs (Webster & Oliver, 2001). If only the spatial differences (increments) are stationary, then this is covered by the intrinsic hypothesis, wherein one assumes that for every vector *h* the increment $Y_n(x) = Z(x) - Z(x+h)$ is a stationary random function in *x*. Then $Z(x)$ is called an intrinsic random function and its *increments* are second order stationary (Chiles & Delfiner, 1999). Depending upon the stationarity assumption made, the phenomenon can be characterised using different tools.

3.2.3 Characterisation of the spatial phenomenon

Under the assumption of strict or second order stationarity, the covariance function, $C(h)$, may be used. Covariance is a function of *h* and *h* only and tends to decrease as *h* increases. It is the autocovariance function – auto because it represents the covariance of *Z* with itself (Webster & Oliver, 2001). Covariance is calculated as follows:

$$C(h) = \frac{1}{N(h)} \sum_{i=1}^{N(h)} z(x_i) z(x_i + h) - m_{-h} m_{+h}$$

$$\text{with } m_{-h} = \frac{1}{N(h)} \sum_{i=1}^{N(h)} z(x_i) \quad \text{and} \quad m_{+h} = \frac{1}{N(h)} \sum_{i=1}^{N(h)} z(x_i + h)$$

where m_{-h} and m_{+h} are the mean of the tail and head values of the separation vector h and $N(h)$ is the number of pairs separated by vector h.

The covariance may be standardised by the standard deviations of the tail and head values. This converts the measure to the unit-less autocorrelation function, or correlogram, $\rho(h)$:

$$\rho(h) = \frac{C(h)}{\sigma_{-h}\sigma_{+h}}$$

where σ_{-h} and σ_{+h} are the standard deviations of the tail and head values.

The intrinsic hypothesis was proposed by Matheron to circumvent the problem that in many cases the mean appears to change across a region and the variance increases without bound as the area of interest increases. In these circumstances the covariance cannot be defined. Although the mean may vary on the whole, at very short separation distances it would be expected to be equal i.e. E[$Z(x)$-$Z(x+h)$] = 0 for very small h. Under the intrinsic hypothesis, covariances are replaced by variances of differences, var[$Z(x)$-$Z(x+h)$], which also depend not on the absolute position but rather on h. This forms the basis of the intrinsic hypothesis and frees geostatisticians from the constraints of second order stationarity. This variance of differences is called the semivariogram, $\gamma(h)$, (hereafter referred to as the variogram) and is defined as half the average variance of the difference between pairs of tail and head values as a function of h:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i) - z(x_i + h)]^2$$

As h increases the points become less similar and the variogram increases.

For a second order stationary process, the variogram and covariance are equivalent such that a graph of the variogram function is a mirror image of the covariance function. If a process is intrinsic only, there is no equivalence because the covariance function does not exist (Webster & Oliver, 2001). Second order stationarity implies the intrinsic hypothesis but the reverse is not true: an intrinsic random function need not be stationary of order two (Goovaerts, 1997).

The variogram is by far the most common geostatistical descriptor and its salient features are now expanded upon.

3.2.4 Variography

By plotting the average squared difference between the sample values against the separation distance, we can study the general pattern of spatial variability in a spatial phenomenon (Srivastava, 1996). If the phenomenon being studied is very continuous (i.e. highly spatially autocorrelated) over short distances, then the differences between closely spaced data values will be small and will increase gradually as the data pairs become further and further apart. On the other hand, if the phenomenon is completely erratic, then pairs of closely spaced sample values might be as different as pairs of widely spaced data values and in this case the variogram increases rapidly, or may not display a structure at all.

In practice, there are never enough sample data pairs to assess spatial continuity over every separation distance (lag) and direction with any surety. Therefore, a lag and directional tolerance is chosen whereby pairs are pooled into a small number of distance and direction classes and the average sample variogram value is calculated within each distance class for a few key directions. A phenomenon showing spatial continuity that differs according to direction is said to show anisotropy, whereas a phenomenon which displays the same characteristics, whatever the direction, is isotropic. In the isotropic case, all pairs from all directions are placed in the same directional class.

Figure 3.1 shows an idealised variogram plot. There are three main features. The nugget, C_0, is the intercept, which takes a theoretically zero value when $h = 0$. However, due to very short range variability (not picked up in the sampling) and analytical errors (e.g. inaccuracies inherent in the analysis technique), this may take a positive value. The second main feature is the range, a, which is the separation distance at which the plot reaches a plateau. The range delineates the 'zone of influence' such that at separation distances shorter than a, samples are correlated in some way but at greater distances they are not and simply act as independent random variables. The variogram value at range a approximately corresponds to the classical variance of the sample data (Myers, 1997) and is termed the sill, which is the third main feature. The sill, C_1, is present only for second order stationary phenomena. Variograms which reach a sill are termed 'bounded' (Armstrong, 1998), 'transitional' (Isaaks & Srivastava, 1989) or 'transitive' (Myers, 1997). It is possible to have unbounded variograms which never reach a sill, suggesting correlation to an infinite distance due to the presence of underlying drift or trend in the data. Unbounded variograms are associated only with intrinsic variables (Armstrong, 1998).

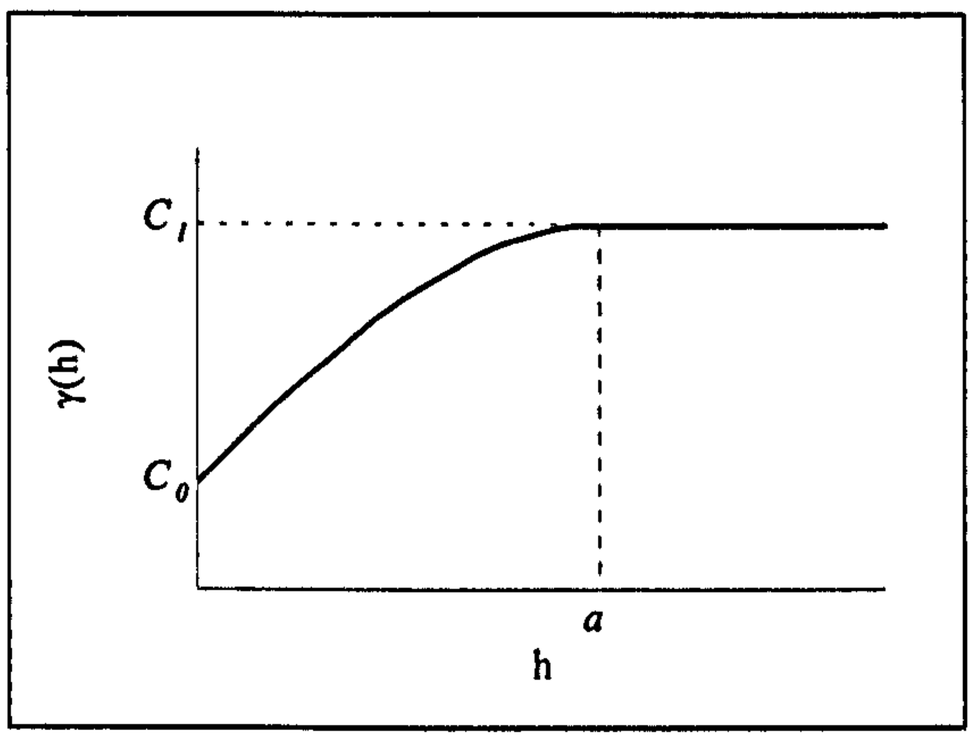


Figure 3.1. Idealised variogram plot. C_0 = nugget, C_1 = sill, a = range

Since variograms are based on sample data, it is not possible to gauge the continuity in every single direction for every possible separation distance. Therefore, in order that the variogram may be used in the kriging process, a model must be fitted so that weights for the local conditioning sample data may be calculated. In order that the model may be used for spatial interpolation it must give a unique solution. The four most commonly used models satisfying this condition are the power, exponential, spherical, and Gaussian. Each model is characterised by a different behaviour near the origin (different manner of modelling the short range continuity) and approach to the sill (for bounded models). Figure 3.2 depicts the four models, which are represented by the following mathematical functions, where C_0 = nugget, C_1 = sill, a = range and h=lag distance (Deutsch & Journel, 1998; Isaaks & Srivastava, 1989; Myers, 1997):

Power $\quad \gamma(h) = C_0 \qquad |h| = 0$

$$\gamma(h) = C_0 + C_1 |h|^{\alpha} \qquad |h| > 0 \qquad 0 < \alpha < 2$$

This model is linear for $\alpha = 1$

Exponential $\quad \gamma(h) = C_0 \qquad |h| = 0$

$$\gamma(h) = C_0 + C_1\left(1 - \exp\left(-\frac{3|h|}{a}\right)\right) \qquad |h| > 0$$

Exponential models reach the sill asymptotically, so a represents the practical range which is the point where the model reaches 95% of the sill value

Spherical $\gamma(h) = C_0$ $|h| = 0$

$\gamma(h) = C_0 + C_1\left(\frac{3}{2}\frac{|h|}{a} - \frac{1}{2}\left(\frac{|h|^3}{a^3}\right)\right)$ $|h| \leq a$

$\gamma(h) = C_0 + C_1$ $|h| > a$

Gaussian $\gamma(h) = C_0$ $|h| = 0$

$\gamma(h) = C_0 + C_1\left(1 - \exp\left(-\frac{3|h|^2}{a^2}\right)\right)$ $|h| > 0$

Gaussian models reach the sill asymptotically, so a represents the practical range which is the point where the model reaches 95% of the sill value.

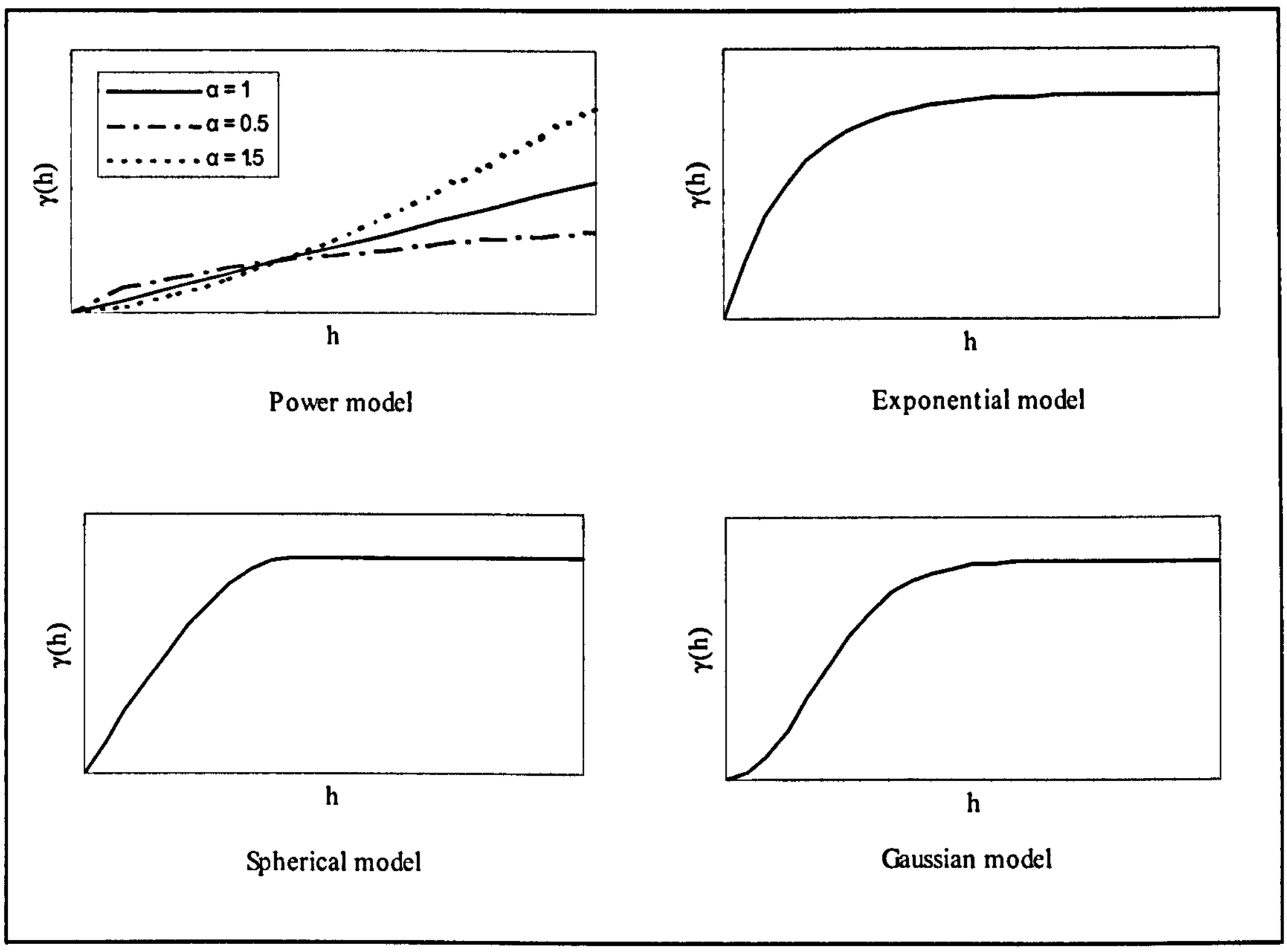


Figure 3.2. Four standard variogram models.

3.2.5 Kriging

Recall that kriging is basically a moving window weighted-least-squares estimator, which weights nearby samples according to weights determined from the variogram model, and so exploits the spatial relationship between the samples. The weights are chosen so as to satisfy two conditions: a non bias condition i.e. the weights must sum to one; and a minimum variance condition, whereby weights are computed by minimising

the estimation variance subject to the non-bias condition. The weights are then used to calculate the interpolated value. If the global mean of the variable is known, the non-bias condition is not required and this leads to simple kriging (Rouhani, 1996). More often than not, the global mean is unknown and ordinary kriging is employed. In this case, the interpolation procedure is free of the global mean value and relies solely on the surrounding sample values.

Ordinary kriging (OK) is often associated with the acronym 'BLUE' - best linear unbiased estimator. It is linear because its estimates are weighted linear combinations of the available sample data, unbiased as it attempts to produce a mean residual error of zero and is best because it minimises the variance of the errors (Isaaks & Srivastava, 1989). A chief feature of kriging, which is ideal for assessment of environmental data, is the automatic declustering of data. Two major problems encountered in spatial interpolation are clustering of samples and distance of samples from the point of interest. As often happens with environmental data, areas of interest (usually areas of high pollutant concentrations) are preferentially sampled, hence over represented in the sample data set. In terms of separation distance, it is intuitive that a sample close by the estimation point should carry more weight than one that is very distant. Kriging makes use of the spatial covariance values to overcome these two issues. In effect, kriging replaces the Euclidean distance by a 'statistical distance' to assign fair weights to all points: physically close, clustered points can be made statistically more remote whilst physically distant, isolated points can be made statistically closer.

Mathematically, the estimate at the unsampled location, $z^*(x_0)$, is given by:

$$z^*(x_0) = \sum_{i=1}^{n} w_i z(x_i) \qquad \text{where} \qquad \sum_{i=1}^{n} w_i = 1$$

and w_i are the weights.

The error variance, σ_r^2, is given by:

$$\sigma_r^2 = \sigma^2 + \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j Cov_{ij} - 2\sum_{i=1}^{n} w_i Cov_{i0}$$

where σ^2 is the overall variance in the area, and Cov is the covariance between two points (Isaaks & Srivastava, 1989). Once the variogram model has been defined, variance and covariance values are calculated and substituted, leaving the error variance as a function of n variables $w_1, ..., w_n$. The function is then minimised with the

constraint that the weights must sum to one (unbiasedness condition). Minimising a function involves setting the partial first derivatives to zero and solving the n simultaneous equations. In this case, however, there is an additional constraint in that the weights must sum to one, giving $n+1$ equations with only n unknowns. To solve this problem, the technique of Lagrange parameters is used whereby an additional unknown parameter, μ, is introduced (Isaaks & Srivastava, 1989). An additional term is added:

$$\sigma_r^2 = \sigma^2 + \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j Cov_{ij} - 2\sum_{i=1}^{n} w_i Cov_{i0} + 2\mu\left(\sum_{i=1}^{n} w_i - 1\right).$$

We know that $\sum_{i=1}^{n} w_i = 1$, therefore $\sum_{i=1}^{n} w_i - 1 = 0$, hence $2\mu\left(\sum_{i=1}^{n} w_i - 1\right) = 0$. In this way the additional variable does not affect the outcome of the equation. The n partial first derivatives can now be set to zero to find the solution that minimises the variance.

The system of equations is known as the 'ordinary kriging system' and can be written in matrix notation (Isaaks & Srivastava, 1989):

$$C \cdot w = D$$

$$\begin{bmatrix} Cov_{11} & \cdots & Cov_{1n} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ Cov_{n1} & \cdots & Cov_{nn} & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ \vdots \\ w_n \\ \mu \end{bmatrix} = \begin{bmatrix} Cov_{10} \\ \vdots \\ Cov_{n0} \\ 1 \end{bmatrix}$$

To solve for the weights, both sides are multiplied by C^{-1}, hence $w = C^{-1} \cdot D$. The minimised error variance, σ_r^2, is commonly called the 'ordinary kriging variance' and is given by $\sigma_r^2 = \sigma^2 - w \cdot D$.

In practice today there are numerous software packages available to solve these kriging systems with relative speed by finding the weights that lead to the best estimate, in the sense of the least squares minimum variance estimator of the value at an unsampled location. Typically, the unsampled locations are the nodes of a regular grid which is laid across the site forming a regular pattern of estimates characterising the whole area.

3.3 Practical considerations

The previous sections have described the fundamental concepts and basic methods behind geostatistical estimation. There are potential drawbacks with these most basic of methods and geostatisticians have adapted the methods to cope with different practical problems and situations that have arisen in the many different disciplines. The focus now turns to practical issues, discussed in the literature, which are of relevance to the research in this thesis.

3.3.1 Smoothing effect

A feature of ordinary kriging, being a least squares estimator, is that although it minimises the error variance at each estimation point in order to produce the 'best' estimate, it can lead to localised biases in the estimates, under-estimating high values and over-estimating low values – the so called smoothing effect (Myers, 1997). While this type of analysis may be satisfactory for some applications, it is not particularly suited to applications which require local assessment of extreme values, such as contaminated land and human health risk assessment. In such studies, high values are of particular interest as they are most likely to pose the biggest risk. Therefore, it is necessary to choose a geostatistical method which preserves the true extremes of the phenomenon and highlights the true extent of the possible risk.

There exists a whole suite of kriging methods adapted to cope with different situations and improve the estimates. However, since all the methods will suffer to some extent from smoothing, it is important that one be chosen that can give a good assessment of the uncertainty in the results. Indicator kriging (IK) is one such method and is discussed in more detail below. Essentially, the emphasis with IK is on estimating the uncertainty rather on producing the 'best' estimate. Disjunctive kriging (DK) can also be used in a similar manner, as explained in the sections below.

Another way of utilising spatial autocorrelation for modelling is to use a simulation algorithm. Instead of producing one map of best local estimates, several maps are generated each of which reasonably match the sample statistics (histogram), variogram model and sample (conditioning) data. The simulated maps are much more erratic and are more likely than kriging to produce extreme estimates. Simulation is therefore appealing for risk assessment and is also discussed in more detail in Section 3.3.5.

3.3.2 Block estimates

Real practical concerns must be considered during a geostatistical investigation. For example, once any hotspots of contamination have been mapped it may be necessary to clean them. The basic geostatistical methods described above concern estimates at unsampled point locations. In practice, every point on the site will not be estimated individually – it is usual to estimate points at the nodes of a grid laid across the site. However, one would not clean only a point of soil, rather one would need to clean the site based on some practical measure, such as the topsoil of an entire domestic garden. It may therefore be necessary to estimate average contaminant concentrations for a whole block instead of a point. Block kriging methods exist, which are able to produce estimates on a block support from point support sample values.

Block kriging, in turn, brings its own difficulties as it can exacerbate the smoothing effect already described. Block kriging is well known to reduce the variance of the values as compared to point kriging (Isaaks & Srivastava, 1989). The 'support effect' manifests itself in a reduction in variance as support size increases. This is due to 'dilution' of extreme high values by the more abundant typical values when the average for the block is taken, thus reducing the maximum estimated value. A similar upward averaging of the low values occurs and hence the distribution of estimates is 'squashed' leading to an overall reduction in variance of the estimates. This is problematic in a risk assessment where extreme high values are of particular interest. Though block values may be desirable, the risk assessor must not lose sight of the fact that, possibly, dangerous hot spots may be 'lost' due to the support effect and will not be cleaned since the block average is below the regulatory threshold.

Block values can also affect the assessment of risk to a receptor. A static receptor e.g. a plant or an animal with a very small foraging range, may live directly on a hotspot of contamination and be subject to very high risk. Assuming a block average value may drastically underestimate the risk for that individual. Alternatively, a hotspot may artificially raise the mean block value, leading to a general over estimation of risk for a mobile receptor who may never come across that hotspot.

The affine and indirect lognormal corrections are two methods which can be applied to the point estimates in order to produce block estimates taking into account the support effect (Isaaks & Srivastava, 1989). The affine correction preserves the mean and the shape of the original distribution while reducing the variance. With the indirect

lognormal correction, the skewness is reduced as well as the variance and the minimum value remains at zero.

Each method requires a variance reduction factor, *f*, which is the ratio of block variance (variation between blocks) to point variance. This can be calculated by knowing that the total (point) variance, $\sigma^2 A$, is equal to the variance within blocks, $\sigma^2 B$, plus the variance between blocks, $\sigma^2 A,B$ (Isaaks & Srivastava, 1989). Typically, one can calculate $\sigma^2 A$ and $\sigma^2 B$. Therefore, block variance can be calculated as $\sigma^2 A,B = \sigma^2 A - \sigma^2 B$. Hence, $f = \frac{\sigma^2 A,B}{\sigma^2 A} = \frac{\sigma^2 A - \sigma^2 B}{\sigma^2 A} = 1 - \frac{\sigma^2 B}{\sigma^2 A}$. Both $\sigma^2 A$ and $\sigma^2 B$ are estimated from the variogram model: $\sigma^2 A$ by finding the average variogram value between all pairs of points within the whole area; and $\sigma^2 B$ from discretising a block and finding the average variogram value between all possible pairs within the block. *f* can then be used to convert the point kriged estimates to representative block kriged estimates using either the affine or indirect lognormal correction method as appropriate.

Another approach is to simulate many point values on a fine grid within the block and use the average of the simulated values to represent better the block value (Goovaerts, 2001). This method is attractive in that it assumes no particular distribution and in addition means that measures other than the block average can be recorded, if desired.

3.3.2.1 Theory of block kriging.

For block estimation, the desire is to estimate the mean value over a given area (block, *V*), rather than at a single point. The mean value of a random function over a local area is simply the average (a linear combination) of all the point random variables contained within that area (Isaaks & Srivastava, 1989) and this mean is also a random variable.

Recall the kriging system for point estimations:

$$C \quad . \quad w \quad = \quad D$$

$$\begin{bmatrix} Cov_{11} & \cdots & Cov_{1n} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ Cov_{n1} & \cdots & Cov_{nn} & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ \vdots \\ w_n \\ \mu \end{bmatrix} = \begin{bmatrix} Cov_{10} \\ \vdots \\ Cov_{n0} \\ 1 \end{bmatrix}$$

The location of the point or block being estimated has no bearing on the calculation of matrix C, therefore C remains unchanged whether points or blocks are being estimated. Matrix D however contains covariances between the random variables at sample locations and the random variable at the point being estimated. While for point estimation, these are point to point estimates, for blocks, these are point to block estimates, therefore matrix D requires modification. The covariance between the *i*th location and the random variable representing the average value of the phenomenon in the block *V* is the same as the average of the covariances between the *i*th location and all the n points within the block *V*. Therefore the D matrix can be rewritten with average covariances, $\overline{Cov}$, between a particular sample location and all the points within the block *V*:

$$\begin{bmatrix} Cov_{11} & \cdots & Cov_{1n} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ Cov_{n1} & \cdots & Cov_{nn} & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ \vdots \\ w_n \\ \mu \end{bmatrix} = \begin{bmatrix} \overline{Cov_{1V}} \\ \vdots \\ \overline{Cov_{nV}} \\ 1 \end{bmatrix}$$

(Isaaks & Srivastava, 1989).

3.3.3 Non-normality of data

Simple geostatistics work best if the underlying distributions are normal (Gaussian). Dealing with contaminated land, more often than not, means dealing with highly positively skewed data sets, e.g. with coefficients of variance of two or more, and this can lead to problems. Since the variogram is a calculated by squaring a difference between pairs of data, it is extremely sensitive to strong positive skewness in the data. A few extreme high data values will contribute to many data pairs and will lead to high variances, which may destabilise the variogram and in some cases render it useless. A common approach to overcome this problem includes trimming off the extreme values while calculating the variogram, but this is rather simplistic, especially in the case of a risk assessment such as this, when high values are of particular importance.

Another option is to transform the data using logs to produce a near normal distribution and then work with the logs of the data and back-transform the final results. This method is known as lognormal kriging (LK) (Webster & Oliver, 2001). It is unfortunate that the back-transforms of the data may be subject to bias since the exponentiation exaggerates any errors associated with the estimates and this is usually worst for the high values which are of interest in a risk assessment (Goovaerts, 1997).

Also, it is not possible to obtain an unbiased back-transformed estimate of the kriging variance, should this be needed, since the mean is unknown (Webster & Oliver, 2001). So, although the variography may be more robust, this may be counteracted in the final results due to sensitivity of the back-transformation.

It is possible to use more robust measures alongside the usual variogram in order to help deduce the key features such as the range and any anisotropy. Using the madogram (mean absolute deviation) and rodogram (square root of the absolute deviation), for example, may help to highlight these key features which are obscured in the variogram plots (Goovaerts, 1997).

Some authors have proposed non-parametric (distribution free) alternatives to the variogram and ordinary kriging, which are more robust to outliers and non-Gaussian data distributions. Walker & Loftis (1997), for example, devised a kriging method based on weighted least absolute deviation regression using weights calculated from a 'medogram' - median absolute difference of pairs versus separation distance. Median estimators are usually more robust to outliers and they found, indeed, that their estimates, in this case, were much less influenced by outliers than those from ordinary kriging.

An alternative, oft-used, non-parametric method is indicator kriging (IK) (Journel, 1983). This method is particularly suited to contaminated land investigations (e.g. Benamghar & Sonnet (1999); Leonte & Schofield (1996); Smith et al. (1993)), since it not only copes well with the skewness of the data, but also gives a risk (uncertainty) measurement, through the conditional cumulative distribution function, ccdf, as well as an estimate of the soil concentration (discussed further below). IK involves a simple transformation of the sample data into binary codes: 1 if the data point is above a critical threshold (e.g. a regulatory threshold) and 0 if below. Variography and kriging are performed on the binary transforms resulting in values which represent the probability of exceeding the threshold value. IK is a general method, in that it can be applied to data sets whatever their distribution shape.

DK (disjunctive kriging) is a parametric method which involves transformation of the non-Gaussian variable, using Hermite polynomials, to a standard normal variable with zero mean and unit variance. The series of Hermite polynomials forming the normal distribution are kriged separately and their results summed to produce an estimate. Back-transformation of results produces contaminant concentration estimates on the

original scale. The method is general, in so much as all data sets can be converted to a standard normal distribution and subsequently back-transformed. Studies have shown, however, that estimates using DK are not significantly different to those achieved by using OK (Webster & Oliver, 1989; Yates & Yates, 1988). It was concluded that since DK is so mathematically complex and time consuming compared to OK, it is not worth the effort for a simple estimate of soil concentration. However, it is pointed out in these same studies that the attractive feature of DK, which could make it worthwhile, is that it is possible to obtain a measure of uncertainty in the estimates directly from the process results, since the ccdf is constructed at each estimation point. This is discussed further below.

3.3.4 Quantifying uncertainty

A major part of this research is quantification of uncertainty. It is vital for a transparent risk assessment to have an assessment of uncertainty in the estimates. For a Level 1 assessment, one can produce a soil concentration estimate with attached confidence intervals or a probability map giving the probability of exceeding a regulatory threshold. For a level 2 assessment, uncertainty in the soil contaminant estimates can be preserved by building a distribution of multiple feasible estimates at each point and feeding these into the exposure equation in a type of Monte Carlo analysis.

Throughout the geostatistical procedure, there are many routes for uncertainty to enter the analysis. First is at the data collection stage where sampling and analytical error may be evident. Next, variography is subject to uncertainty. The structural analysis is usually done 'manually' in that the geostatistician works by trial and error to uncover the structure and any anisotropy in the sample data. Model fitting may be done with the aid of statistical goodness of fit measures, e.g. Pannatier (1995), but is often done by eye – the choice of nugget, range and model type being selected by the practitioner using any knowledge of the phenomenon to help. In practice, a whole range of possible models could look to fit the data equally well and, in the end, the choice is subjective. The interpolation (kriging) procedure is influenced by the measured sample data values, the size of the local search neighbourhood (hence the number of conditioning data used) and the model selection made. All this leads to uncertainty in the interpolated results.

3.3.4.1 Uncertainty in the analytical results

Uncertainty in the analytical results is often assumed to be negligible and sample values are taken at face value. However, some authors have adapted methods to carry the

known analytical error through into the geostatistical analysis. Van Miervenne & Goovaerts (2001), for example, use indicator kriging with 'soft' indicator coding. Instead of hard coding data as 1 if it is above the threshold and 0 if below, they use the coefficient of variation (CV) from the analytical results to represent the uncertainty in the sample value. Each value is represented by a distribution with the measured value as the mean and the CV as the standard deviation (assuming a normal distribution of errors). The soft codes at each cut off are represented by the normal cumulative distribution function, and are indicators valued between 0 and 1, allowing for the fact that, due to analytical errors, we are not confident that the value is definitely above or below the threshold value.

3.3.4.2 Uncertainty in the variogram

Several authors have proposed methods for accounting for variogram model uncertainty. Pilz et al. (1997) recognise that the variogram model used in kriging is only an estimate itself and that uncertainty arising from the model is not fully accounted for. To rectify this, they propose forming a 'belt' including a whole class of plausible variogram models and then selecting the best model by 'minimax' kriging – selecting the model which minimises the maximum kriging variance. Bogaert (1999) uses an analytical approach to calculate confidence intervals for each distance class of the variogram plus distributions on the range and sill parameters. Reliable hypothesis testing can be used to select a suitable variogram model. Safai-Naraghi & Marcotte (1997) account for uncertainty in the variogram model by bootstrapping the original data set to produce 100 plausible sets, fitting a suitable model to each data set, then simulating one map from each of 100 variogram models. This carries the uncertainty from the variogram model into the results. Although these methods are possible, the most common way of accounting for uncertainty in the variogram is to use simulation methods (see Section 3.3.5). Bogaert, (1999) recognises this, but claims simulations are lengthy and approximate, as opposed to the more exact analytical approach. The desired end point of the analysis should be taken into account when deciding if and how the variogram uncertainty should be taken into account.

3.3.4.3 Uncertainty in interpolation

Probably the most common statistical method for reporting uncertainty is the 95% confidence interval. It is usual to report the mean ±1.96 standard deviations, denoting that on repeated estimations 95% of the means will fall within this interval. The one

specific condition when this type of confidence interval may be used in geostatistics is under conditions of strict multi-Gaussianity, when data and errors are normally distributed. In this case, the kriging estimate is the mean of a distribution of possible values with variance equal to the kriging variance (Deutsch & Journel, 1998) and the ±1.96 standard deviations 95% confidence interval is valid.

It is unfortunate in geostatistics that this measure is most often inappropriate for local estimations. Firstly, this approach assumes an underlying normal distribution of errors. It is known that kriging is often locally biased, under-estimating high values and over-estimating low values, so while the global error estimation may be overall normally distributed, the local errors are often highly skewed. Secondly, the kriging (error) variance from the random function model is assumed to be an accurate estimate of the variance of the real errors. The kriging variance is, in fact, independent of the data values (a condition referred to as homoscedasticity), yet the true variance of errors will not be. Intuitively, one would expect the scope for error to be higher for an estimation point surrounded by one very small and one very large value than for a point surrounded by two values of similar magnitude and so the second assumption is rarely met in practice. Hence, it should be with caution that a ±1.96 standard deviations 95% confidence interval is used (Goovaerts, 1997; Isaaks & Srivastava, 1989). Buxton et al. (1997), for example, encounter problems associated with this method, including negative lower confidence intervals which are clearly inappropriate.

Instead of calculating a least squares estimate by kriging and then attaching confidence intervals, an alternative solution is first to assess the uncertainty in the estimate, and then try to pinpoint a likely value for the estimate. To do this, the conditional cumulative distribution function (ccdf) is deduced locally at each estimation point. The various ccdf statistics are dependent on the data values, producing larger uncertainty estimates where there is a large range of data values and smaller uncertainty estimates where the data values are more similar. Hence, uncertainty values deduced using the ccdf approach are likely to be a better reflection of reality. The ccdf, i.e. cumulative distribution function made conditional on the local information available, *(n)*, models this uncertainty by $F(x;z\,|\,(n)) = \Pr\{Z(x) \le z\,|\,(n)\}$. One can derive probability intervals such as

$$\Pr\{Z(x) \in (a,b]\,|\,(n)\} = F(x;b\,|\,(n)) - F(x;a\,|\,(n))$$

or

$$\Pr\{Z(x) > b \mid (n)\} = 1 - F(x; b \mid (n)),$$

which are independent of any particular estimate of *z(x)* and depend only on the locally available information, *(n)* (Goovaerts, 1997).

As has been alluded to in previous sections, both DK and IK are suitable methods for building the ccdf. Of the two, IK is mathematically more simple and flexible and can be a fast and efficient way of obtaining results. For example, it is not necessary, using IK, to always construct the full ccdf. If a regulatory threshold is already known, data can be binary coded at that specific cut off and kriging carried out on those indicators only, immediately resulting in a map of probabilities of exceeding the threshold. One can use IK to assess uncertainty about the results (i.e. probability of exceeding the threshold) without even calculating an actual concentration estimate, e.g. Leonte & Schofield (1996). Though IK is mathematically simpler than DK, it is more cumbersome in the variography should the full ccdf be required, in which case variograms must be fitted at multiple cut offs.

There exists a suite of kriging methods all adapted to suit particular problems (details can be found in texts such as Chiles & Delfiner (1999), Goovaerts (1997b) and Webster & Oliver (2001)), some of which are examined in comparative studies available in the literature. Many comparisons concentrate only on the estimation performances of the method, e.g. Mohammadi et al. (1997) and Saito & Goovaerts (2000) while some, more importantly from the point of view of selecting a suitable risk assessment method, focus on their performance at assessing uncertainty. Chica-Olmo & Luque-Espinar (2002) use DK and IK but merely conclude that the methods are simply different and should be chosen to suit the problem at hand. Lloyd & Atkinson (2001) compared OK, IK and kriging with a trend (KT) with respect to assessing uncertainty in elevation estimates. They concluded that OK and KT were easier to implement, gave better estimates and better global uncertainty estimates, but that IK performed better on the assessment of local uncertainty. Papritz & Dubois (1999) compared OK with three 'non-linear' methods, LK (lognormal kriging), IK and DK on various data sets. They concluded that LK, IK and DK failed to estimate better than OK (even when the Gaussian model was theoretically inappropriate). More importantly, while OK failed on skewed data, LK, IK and DK modelled the uncertainty equally well with no one of these three methods being consistently better than the others. Buxton et al. (1997) used IK, LK and OK to

find the probabilities of exceeding a threshold and found IK uncertainty estimates to be more realistic than either OK or LK.

3.3.5 Simulation

An alternative for deducing the ccdf is to simulate several possible values for *z(x)* using a simulation algorithm. Sequential simulation is a geostatistical Monte Carlo method, whereby instead of producing one map of best local estimates, the emphasis is on producing several maps each of which reasonably match the sample statistics (histogram), variogram model and conditioning (sample) data. It is possible to match the conditioning data exactly (conditional simulation) or re-estimate the data points (non-conditional simulation) e.g. if there is reason to believe they are not accurate. In effect, simulation algorithms automatically account for multiple sources of uncertainty.

The maps produced through simulation are more realistic than those produced by kriging since they do not suffer from smoothing and reproduce the local variation better. Simulation is increasingly preferred to kriging for applications where the spatial variation of the measured field must be preserved, such as delineation of contaminated land (Goovaerts, 1999). Each simulated value, $z^{(l)}(x)$, l = 1,…,L represents a possible outcome for the random variable *Z(x)*. Working with the series of alternative maps (realisations), $\{z^{(l)}(x), x \in A\}, l = 1,...,L$ it is possible to represent the uncertainty in the estimates through summary statistics calculated from the map series: the proportion of simulated outcomes which falls above a regulatory cut off value gives the probability that the point is contaminated; the multiple results produced at each point can be fed into the exposure equation to carry the uncertainty into the exposure analysis.

There are several simulation algorithms available and, like their kriging counterparts, they are suitable in different situations. Since the local uncertainty from the simulations is calculated from the series of alternative maps, the suitability of each simulation algorithm must be judged by the global features reproduced on the maps. For example, sequential Gaussian simulation (SGS) is a maximum entropy method and will likely produce a 'salt and pepper' image with no connectivity between high and low values (Rossi et al., 1993). On the other hand, sequential indicator simulation (SIS) favours the reproduction of connectivity between high values, and as such may be more suitable for a risk assessment where interest lies in patches of contamination (Murray et al., 2002). Other methods are available which produce multiple images, for example LU decomposition, annealing methods, p-field simulation and turning bands, the detailed

methodologies of which can be found in texts including Chiles & Delfiner (1999) and Goovaerts (1997).

Examples in the environmental literature appear, most commonly, to use SGS for a whole range of applications (Faechner et al., 2000; Lapen et al., 2001; McKenna, 1998; Mowrer, 1997; Nathanail et al., 1998; Rossi et al., 1993; Wang et al., 2002). Of these, only two authors justify this choice (Rossi et al., 1993; Wang et al., 2002). A comparative study by Gotway & Rutherford (1994) examined several conditional and non-conditional algorithms on several data sets, and applied a transfer function to the simulated results. They put differences down to many possible factors, including the order in which simulated values are obtained and differences in geostatistical inputs to the algorithms (e.g. variograms). Findings of interest are: conditioning the simulations reduced the differences in results between the different algorithms; conditioning reduced the uncertainty in the response distribution, in particular the tails of the distribution were pulled toward the centre, the median was more accurate and outliers were eliminated; uncertainty distributions from the Gaussian based algorithm were more accurate than those produced by SIS; clustering of similar data is more pronounced in realisations produced by SIS. They found SGS to be flexible and accurate for a variety of data sets, but emphasise that the differences and their impacts must be understood in order that the geostatistician chooses the correct algorithm for the job in hand. Gómez-Hernández (1997) states that multi-Gaussian models should never be used if it is suspected that extreme values may be spatially connected.

3.3.6 Kriging versus simulation

A method should be chosen depending upon the endpoint of the study. Kriging prioritises the local constraint of minimising the estimation error, while simulation prioritises reproduction of global features, such as reproduction of the histogram and variogram (Goovaerts, 2000). If only a local estimate of the soil concentration is needed, without a measure of uncertainty, then simulation is a waste of CPU time and kriging will suffice. If enough simulations are made, the mean of the simulations will tend toward the kriging estimate at that point. On the other hand, simulation does not try to produce the optimal local estimate, but better reproduces the overall spatial continuity, so if the priority is to reproduce global features, simulation may be preferred, at the expense of possible larger mean prediction errors (Goovaerts, 2000).

3.4 Methods used in this research

If only a local estimate of probability of contamination is needed, simulation may be too time consuming. An estimate, albeit possibly smoothed, can be achieved by kriging, and if indicator or disjunctive kriging is used then a suitable assessment of the uncertainty will be available.

On the other hand, if the results are to be fed into another model where the global picture is of importance, then it would seem sensible to carry out a simulation procedure. Multiple values can then be fed into probabilistic exposure models. Another consideration, from a practical point of view, is that public domain software is available for the simulation algorithms (Deutsch & Journel, 1998). Although other methods have been proposed, such as sampling ccdfs created by IK (Van Miervenne & Goovaerts, 2001), software is not readily available. In addition, simulation, importantly, preserves the global patterns of contamination better than kriging, so it is expected that the human distribution patterns combined with simulated soil contaminant patterns would provide good exposure estimates.

For this research project, the theoretical advantages of SIS are appealing. Generally, simulation does not suffer from the effects of smoothing, and lends itself to block estimations by use of the multiple available point simulated values in the block. Simulation implicitly allows for uncertainty in the sample data and variogram model. As a generic method, SIS is suited to assessment of contaminated land and risk assessment for the following reasons: the spatial connectivity of high values of interest may be modelled differently to the low values; it is useful if accurate spatial continuity is required (as opposed to simple global average values); the method is non-parametric and does not rely on Gaussian assumptions (or tranformations), so is suited to skewed data sets; and it has the ability to incorporate secondary data, if necessary (Rossi et al., 1993). Its main drawback is that it is computationally intensive due to multiple kriging systems. Another quoted drawback is that it does not necessarily reproduce the sample histogram accurately (Goovaerts, 1997), but for this research, the correct spatial modelling of high values is more important.

Should a level 1 assessment be required in isolation, IK is the most appealing. DK and IK can be short-listed as suitable candidates, mainly for their ability to assess uncertainty better than other kriging methods. Of the two, IK is probably simpler to

implement and is more flexible, in that quick probability maps may be created for specified regulatory cut off levels.

3.4.1 Indicator kriging: theory

IK is based on the interpretation of the conditional probability as the conditional expectation of an indicator random variable $I(x, z_c)$ given the information *(n)* (Goovaerts, 1997). First of all, a series of cut off values, z_c, is determined which discretise the interval of variability of the continuous attribute *z*. These may be the nine deciles of the sample data values and/or may include values of importance such as regulatory cut off levels. Next, the data values are binary coded for each z_c depending on whether they are above or below that value. These binary codes, or indicators, $I(x_i; z_c)$, are allocated as 1 if $z(x_i) \leq z_c$ and 0 otherwise (ignoring uncertainty). Based on Deutsch & Journel (1992), the ccdf is calculated as

$$F(x; z_c \mid (n)) = \sum_{i=1}^{n} a_i I(x_i; z_c) \qquad \text{with} \qquad a_i \in [0,1] \text{ and } \sum_{i=1}^{n} a_i = 1$$

where a_i is the weight attached to datum location x_i, which is determined through a variogram model. Indicator kriging of a continuous variable is not aimed at estimating the indicator transform (0 or 1), rather it provides a least squares estimate of the ccdf at cut off z_c:

$$[i(x; z_c)] = E\{I(x; z_c \mid (n)\}$$
$$= \Pr\{Z(x) \leq z_c \mid (n)\}$$

where *(n)* represents the conditioning information available in the neighbourhood of location *x* (Deutsch & Journel, 1992). The process is repeated at each cut off, z_c, and since IK is a non-parametric method, one can use a different model for the kriging at each cut off level, modelling high and low values differently, if appropriate. The indicator variogram is a measure of how often a pair of indicators, separated by a vector *h*, falls on opposite sides of a cut off, and only pairs which fall on opposite sides contribute to the variogram at that cut off.

The ccdf built by assembling the *c* indicator kriging estimates represents a probabilistic model for uncertainty about the unsampled value *z(x)* (Deutsch & Journel, 1992). Assumptions must be made for extrapolation beyond the smallest z_c (lower tail), the

BIBL. LONDIN. UNIV.

largest z_c (upper tail) and for the interpolation between two consecutively ranked z_c values in order to construct the full ccdf.

Once the ccdf has been constructed, it is possible to estimate the probability of being above or below any given value at any estimation point. One can choose some appropriate criterion to select the value *z(x)* to represent the unknown data value at the estimation point. For example, one could map the mean value of the ccdf as a representation of the best estimate. Importantly, the uncertainty in the estimate can be represented by a type of confidence interval based on a symmetric probability interval. For example, the 5th and 95th percentile values represent the lower and upper bounds of a 90% confidence interval, whereby there is a 90% probability that the true concentration lies within these bounds.

3.4.1.1 Problem with block kriging indicators.

Block kriging of indicator variables is not straightforward since the block kriging system is only suitable for the evaluation of variables which average linearly. The indicator variable $I(x_i; z_c)$ is a non-linear transform of the original variable $z(x)$, so the block indicator $i_v(x_i; z_c) = 1$ if $z_v(x_i) \geq z_c$, and 0 otherwise, is not a linear average of point indicators.

Goovaerts (1997b) describes the 'block' ccdf and the 'composite' ccdf. The block ccdf is actually the probability that the mean z-value of the block exceeds the threshold z_c. The block ccdf $F_V(x; z_c \mid (n))$ cannot be derived as a linear average of point ccdfs:

$$\left[F_V(x; z_c \mid (n))\right]^* \neq \frac{1}{N}\sum_{i=1}^{N}\left[F(x_i; z_c \mid (n))\right]^* = \left[F_N(x; z_c \mid (n))\right]^*$$

Rather, this linear average, or composite ccdf, $\left[F_N(x; z_c \mid (n))\right]^*$ is an estimate of the proportion of the point values within block V which exceed the threshold z_c. The block kriging system allows one to calculate the composite ccdf as a linear combination of the point indicators $i(x_i; z_c)$.

3.4.2 Sequential indicator simulation: theory

An alternative for deducing the ccdf, is to simulate several possible values for *z(x)* using a simulation algorithm. To recap, each simulated value, $z^{(l)}(x)$, $l = 1,\ldots,L$ represents a possible outcome for the random variable *Z(x)*. Working with the series of alternative

maps (realisations), $\{z^{(l)}(x), x \in A\}$, $l = 1,...,L$, it is possible to represent the uncertainty in the estimates through summary statistics calculated from the map series. For example, the proportion of simulated outcomes which fall above a regulatory cut off value gives the probability that the point is contaminated.

A set of simulated maps, $\{z^{(l)}(x_j), j = 1,...,N\}, l = 1,....L$ can be generated by sampling the N-point ccdf that models the joint uncertainty at N locations x_j (Goovaerts, 1997):

$$F((x_1,...,x_N; z_1,...,z_N \mid (n)) = \Pr\{Z(x_1) \le z_1,...,Z(x_N) \le z_N \mid (n)\}$$

For the sequential simulation algorithm, instead of modelling the N-point ccdf, a one point ccdf is modelled and sampled at each of N points visited along a random sequence. To ensure reproduction of the *z*-covariance model, each one-point ccdf is made conditional to the original *n* data and also all values simulated at previously visited locations. SIS, like IK, is non-parametric and assumes no underlying distribution of data and is robust to outliers.

Consider sequential indicator simulation of a continuous attribute *z* at *N* nodes x_j of a grid, conditional to the data set $\{z(x_i), i = 1,...,n\}$. Based on Goovaerts (1997b), first of all a series of cut off values, z_c, is determined and the data values binary coded for each z_c such that $I(x_i; z_c)$ is 1 if $z(x_i) \le z_c$ and 0 otherwise. A random path is defined such that each grid node x_j is visited only once. At each grid node the ccdf values are calculated at each cut off level using an indicator kriging algorithm and conditioning data consisting of a specified number of both the original data points and the newly simulated values at previously visited nodes. The complete ccdf model $F(x; z \mid (n)), \forall z$ is built using interpolation/extrapolation algorithms. A *z* value is drawn from this ccdf and added to the conditioning data set. The next node is visited and the process repeated, until all nodes have been visited. This procedure is repeated *L* times, using a different random path each time, to produce *L* simulated maps.

The indicator simulation at each grid node can be viewed as a two step procedure. First a simulated class-value is assigned to the grid node, say $x_j \in (z_{c-1}, z_c)$ then a *z* value is drawn from that class using a within class distribution model, e.g. a uniform distribution. Consequently, indicator based simulations reproduce class proportions and corresponding indicator variograms, but not necessarily the overall ccdf and variogram

of the continuous z values. Reproduction of the overall model depends on the number of discretising classes and the interpolation/extrapolation used.

Post processing can be carried out to allow for easy display of the uncertainty surrounding the estimated values at the N grid nodes x_j. Probability maps can be created by counting the proportion of outcomes at each node that exceed a cut off value. Measures of uncertainty at each locality are represented by plotting probability intervals such as the interquartile range of realisations at a grid node or a 90% confidence interval bounded by the 5^{th} and 95^{th} percentile values. The bigger the range, the more uncertain the estimates.

3.5 Summary

A general overview of geostatistical theory has been given and the points most critical to this research have been discussed in more detail. For both practical and theoretical reasons, Indicator methods are selected as being most suitable for use in the spatially-evaluated quantitative risk assessment methodology which is developed in this thesis: public domain software is readily available for their implementation; estimates of uncertainty are achievable; the spatial continuity of contaminant concentrations are preserved in simulations; and these values will provide suitable inputs for a model to assess human health risks.

In the following chapter, risk assessment and spatial assessment are pulled together to show how the two disciplines can be used in conjunction with one another to produce a spatially-evaluated risk assessment. An overall conceptual framework is presented under which the rest of the research is made.

Chapter 4 Conceptual research framework

4.1 Introduction

A framework for the research is laid out in this chapter. To start, the necessity of this work is justified by means of a short summary of literature, some of which helped form the seed of the idea for this research. Next, a conceptual framework is presented and two levels at which risk can be assessed are defined. Finally, the points which need to be addressed during the development of the risk assessment methodology are presented.

4.2 Merging two disciplines

Perusing the specialist risk journals – Risk Analysis and Human and Ecological Risk Assessment – there appear very few papers concerned with the spatial distribution of risk: most are concerned with other methodological aspects such as separation of uncertainty and variability. On the other hand, the geostatistical publications are, obviously, full of papers concerned with spatial data analysis and a number of them cross into the realms of risk assessment. However, the vast majority stop at the assessment of whether a soil may be classified as polluted against a given regulatory threshold. They venture into risk from a probabilistic point of view in so far as they calculate probabilities that soils are truly contaminated, calculate risks of incorrect classification (false positive, false negative) and apply a cost function to value the risks in monetary terms for unnecessary remediation work for instance, e.g. Goovaerts, (1997a; 1997c) and Rautman (1997).

In the more general environmental publications, especially, are some rather more inter-disciplinary studies. Burmaster & Thompson (1997) are critical of the USEPA assumption that every receptor is exposed to the mean (or the upper 95th percentile

confidence limit on the mean) soil concentration in the area of concern. They recommend a methodology for probabilistic human/ecological risk assessments using Bootstrap methods (to provide the uncertainty), random kernels (varying the intensity of exposure) and an interpolation function (for estimating the soil concentrations across the area), which can preserve the spatial nature of the data. However, the method is aimed at providing a more realistic estimate of the 'exposure point concentration' (the concentration to which receptors are exposed), which is still a single (though arguably more realistic) value, but is not a mappable entity.

Hope (2000; 2001) develops a spatial approach (individual based random walk Monte Carlo model) to assess the theoretical exposure of ecological receptors with different population sizes and foraging ranges and demonstrates that the starting point of a receptor, how mobile the species is, and the variability in contaminant levels across the site all have an effect on their calculated exposures. Similar ecological exposure studies have been produced by Marinussen & vanderZee (1996) and Linkov et al. (2001).

Clifford et al. (1995) produced a spatial risk assessment using Geographic Information Systems (GIS) to overlay receptor foraging ranges and soil concentrations in order to calculate overall exposure in terms of hazard quotients (calculated exposure divided by regulatory acceptable exposure). Soil concentrations were assigned by the voronoi polygon technique, whereby the polygon contains a single sample value and every point in the polygon is closer to this sample point than any other. The whole polygon area assumes its sample point concentration. Receptors are randomly placed on the site and their foraging ranges superimposed across the region in order to produce area weighted contaminant concentrations for exposure calculations. Results are produced in the form of a 3-dimensional 'risk surface', where the z axis is the hazard quotient which is plotted across the site (with x-y axis boundaries). A similar study was produced by Kooistra et al. (2001) but rather than using voronoi polygon technique the study region was divided by the researcher into areas whose soil concentrations were deemed the same. Results were mapped with a point symbol representing the centre of a receptors foraging range, with the symbol size in proportion to the risk.

Fewer case studies deal with human risk assessment. In one such study, by Tristan et al. (2000) as well as producing a probabilistic 'spatially resolved hazard assessment' (spatial distribution of soil contamination) a 'spatially resolved exposure assessment' for human receptors is created. Using GIS, soil contamination and soil ingestion maps

were combined to produce a spatial map of 'average daily exposure' which can be compared to a regulatory 'tolerable daily intake' to produce a spatial map of risk. This paper provides mappable results; maps are available to show the soil concentrations on a regular grid across the region and also to indicate exposure levels on the same grid, (assuming a receptor lives in each cell). The method is useful, in principle, but the geostatistical procedures used to derive the soil concentrations are questionable: lognormal kriging was used with an estimate of uncertainty of twice the log of the kriging standard error.

A study by Korre et al. (2002) also did not incorporate a spatially referenced population, and, similarly, produced risk maps to show risks that would occur if receptors lived in each grid cell on the map. Ordinary kriging was used to interpolate the soil contaminant concentrations and no uncertainty was taken into account, though it was included in the population-related variables. This method differs from Tristan's in that it does not rely on combination of data in GIS, but rather incorporates it directly in a risk equation. The spatially referenced results are then plotted in GIS.

Ginevan & Splitstone (1997) devised a methodology based on probability kriging which is capable of incorporating information in addition to the soil contamination levels. They define total soil-related cancer risk, SRC_T as:

$$SRC_T = \sum_{b=1}^{B} SRC_b * U_{b,s}$$

where SRC_b is the soil-related cancer risk attributed to a particular block b and $U_{b,s}$ is the probability of contacting the block b given the sth utilisation scenario. This model is flexible in that, for example, different site-use scenarios can be reflected by different sets of $U_{b,s}$ values. SRC_k values (sample-point specific risks) are calculated first, and these are subject to kriging, to estimate the probability that a block exceeds the risk limit. The results are mappable.

The examples of spatially-evaluated risk assessments above vary greatly. This is partly to do with the aim of the studies, and partly to do with the information available. Some make spatially explicit calculations, but produce results in a non-mappable form. Others produce mappable results, but may still be rather theoretical e.g. assuming one receptor in each location. Some of the geostatistical procedures used could be improved upon. Some take into account uncertainty, others do not.

It is this author's opinion that further integration of methods being developed in the specialist risk and geostatistical fields should be made. This is done by development of a spatially-evaluated quantitative risk assessment method described and demonstrated in the rest of this thesis.

4.3 Definitions

Before explaining in more detail the methodology developed, it is necessary to define the terms as used in the thesis from here onwards. The definitions used by DEFRA and the EA in the UK are adopted as follows (DEFRA & EA, 2002b):

Exposure: the amount of chemical in a medium that is available for intake by the population. Expressed in $mg.kg^{-1}$.

Intake: the amount of chemical entering or contacting the human body at the point of entry. Expressed in $mg.kg^{-1}_{BW}.d^{-1}$.

Uptake: the amount of contaminant that reaches the circulating blood, having being absorbed by the body. Expressed in $mg.l^{-1}$.

4.4 Conceptual Framework for this research

A conceptual framework, depicted in Figure 4.1, was developed for this particular research. It is based on an amalgamation of the conceptual risk models presented in Chapter 2, but has been tailored to help understand the different stages of the new spatially-evaluated risk assessment procedure. The overall structure is quite similar to the Covello-Merkhofer model, since the clear distinction between stages, particularly the risk management stage, is appealing. However, changes have been made both to reflect better the work carried out during this research, and to conform to the definitions adopted for use in this thesis. The main points are highlighted below.

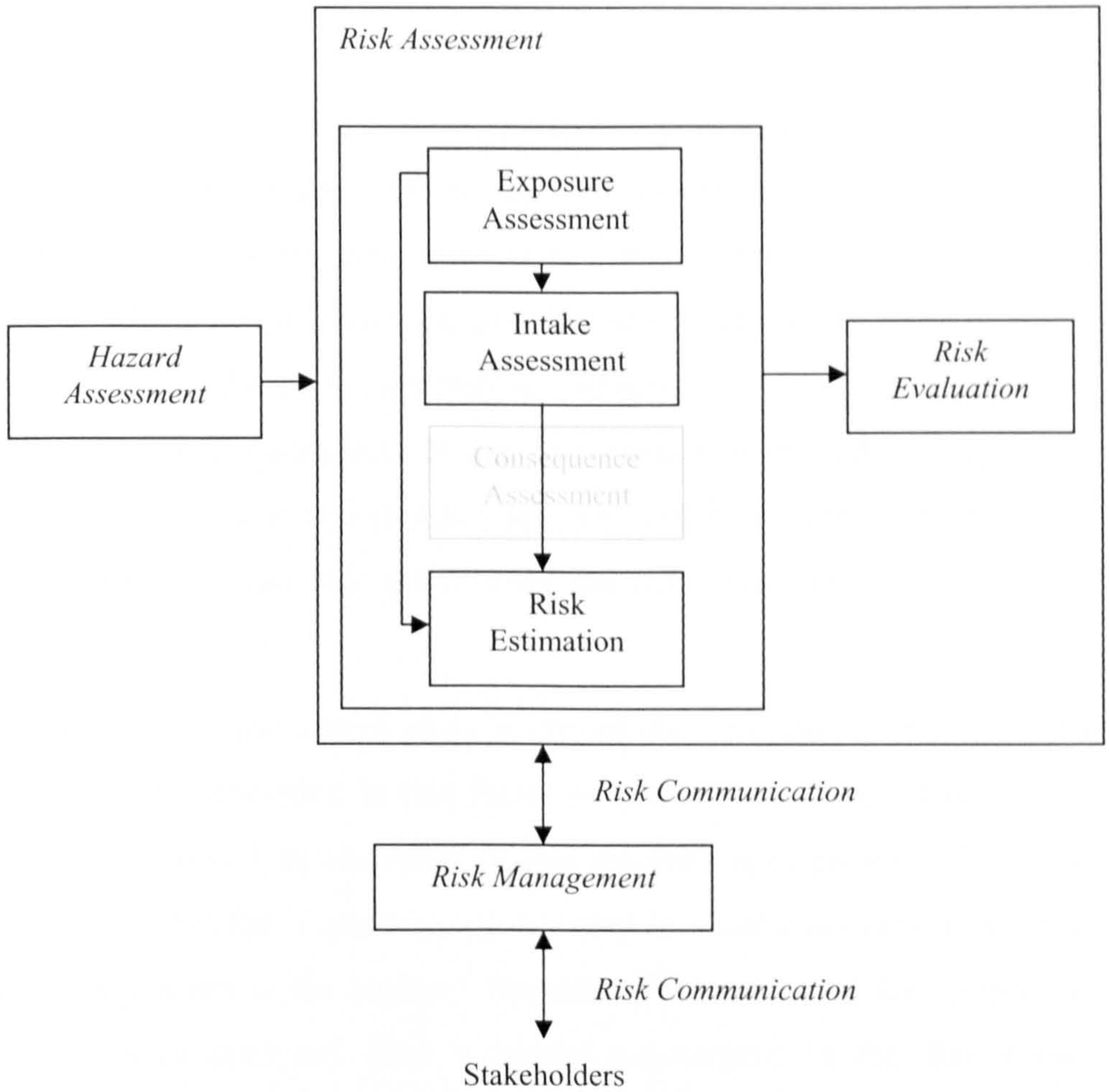


Figure 4.1. Diagrammatic representation of the conceptual framework for work in this thesis

Hazard assessment in this framework is a combination of the Petts et al. 'hazard identification' and 'hazard assessment' stages (Petts et al., 1997). This stage involves the identification of soil contaminants in the vicinity of human receptors and verification that the source-pathway-receptor chain is intact. The overall process called risk assessment consists of a series of calculations as follows:

- *Exposure assessment*: finding the soil contaminant levels to which the receptor is exposed, $mg.kg^{-1}$
- *Intake assessment*: finding the dose received by the individual receptors, $mg.kg^{-1}_{BW}.d^{-1}$
- *Consequence assessment*: deducing the consequences of that intake
- *Risk estimation*: finding the distribution of exposures and intakes and assessing the uncertainties therein and the factors that contribute to those uncertainties

- *Risk evaluation*: comparing the exposure and intake estimates with 'acceptable' standards

Risk management can make use of the results presented to decide on a course of action for the contaminated site. Of paramount importance is communication of the risk. Results of the risk assessment need to be communicated in an effective way to the risk manager, who may communicate his decision criteria in advance of the assessment so that the risk calculation results may be presented to him in an appropriate manner and answer the correct questions. In addition, results of the risk manager's decisions need to be communicated to stakeholders in a way that is relevant to them, and their feedback be communicated back into the process and taken into consideration for refinements of the assessment.

The consequence assessment stage is greyed out. No attempt has been made to do any dose-response assessment in this thesis to quantify the exact medical consequences of the intakes received, as the focus of this research is elsewhere. The stage could have been left out, but the importance of this step in other situations must be acknowledged and the step is left in the outline. Because of the nature of the research for this thesis, data sets were analysed after a hazard assessment in the full knowledge of the consequences that could occur if exposures exceeded 'safe' levels. The case studies were selected on the basis that exposure and intake were likely to exceed safe levels for the population in the region, and have negative consequences which are explained in the literature.

4.5 Two levels of risk evaluation

Figure 4.2 was constructed to illustrate the links between standard setting and the stages of risk assessment and risk evaluation. Risk may be evaluated at more than one level so, in effect, the full chain of assessment need not necessarily be made. For example, exposures may be evaluated by comparing the soil contaminant concentration of arsenic in soil to the regulatory soil guideline value of $20mg.kg^{-1}$. Alternatively, a calculated arsenic intake for a receptor can be compared to a regulatory tolerable intake dose of $3\mu g.kg^{-1}_{BW}.d^{-1}$. Though every step of the process is not undertaken explicitly, it is undertaken implicitly by the fact that the regulatory exposure and intake doses are back calculated from a predefined acceptable risk rate. This research concentrates on developing a new method for undertaking risk assessment at both Levels 1 and 2 (Figure 4.2).

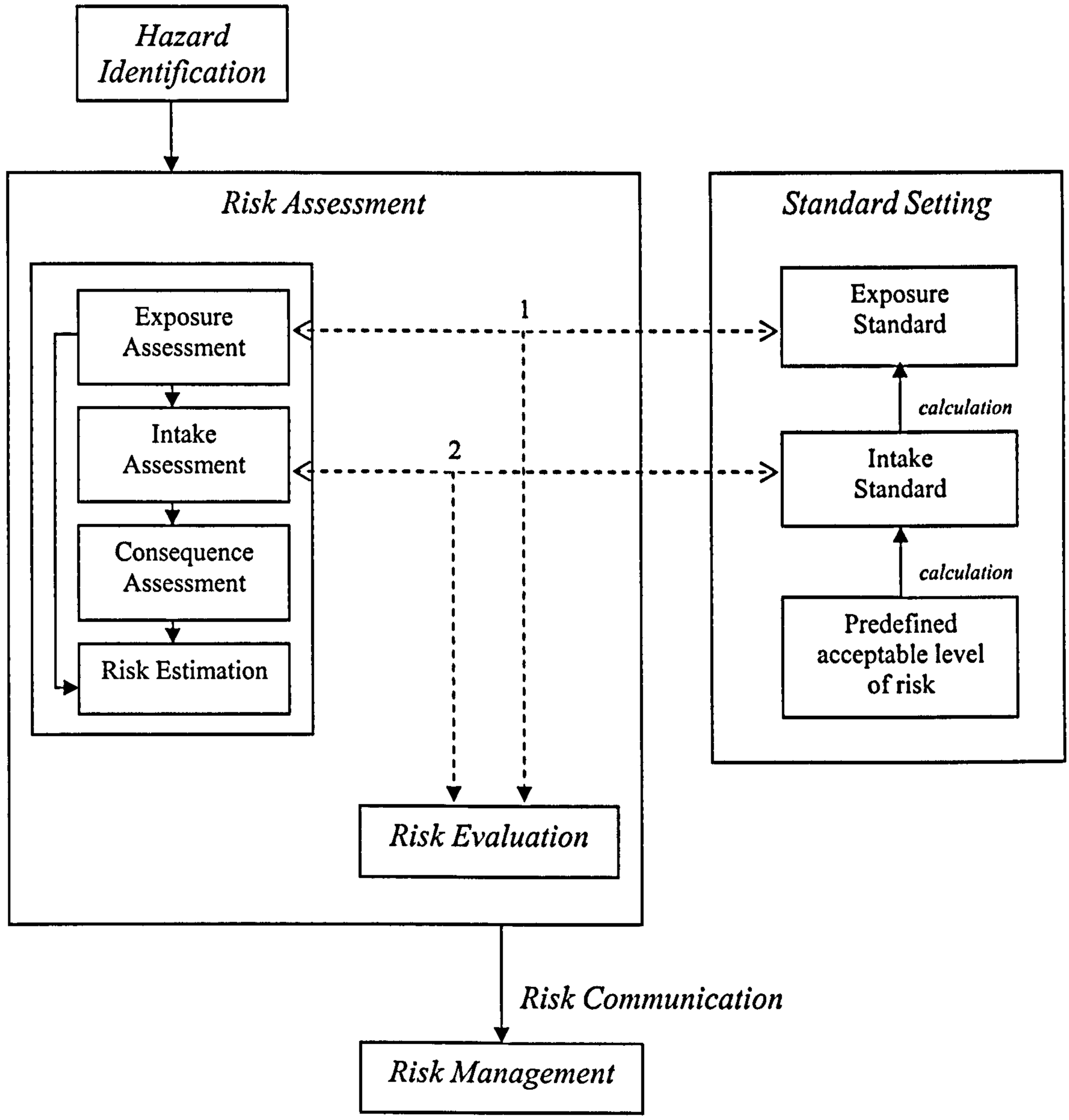


Figure 4.2. Options for evaluating risk.

4.6 Points addressed during development of the methodology

4.6.1 Exposure assessment

Exposure assessment takes on the form of geostatistical analysis of soil contaminant levels using IK and SIS. Issues surrounding the geostatistical methodology are investigated and discussed. Uncertainties due to choice of method (kriging versus simulation), choice of variogram models used and the support effect are presented through a case study.

4.6.2 Intake assessment

A probabilistic, spatially-evaluated intake assessment method is developed using an adaptation of the UK CLEA model. Differences in results due to differences in the treatment of various inputs are studied. For example, spatially-evaluated intake assessment is susceptible to the support effect. In addition, the position of receptors in space with respect to contaminant hotspots affects results as does the variability in the characteristics of the population itself.

4.6.3 Risk estimation

Distribution of risk estimates is looked at as an integral part of the exposure and intake assessments. During the investigation of support effects, for example, distributions of risk are produced and these can be compared to one another to assess the similarities and differences. Risk estimation, in this research, is a means of interrogating the method and discovering the differences in outcome due to the techniques used. A type of sensitivity analysis is carried out on the intake assessment model to discover the main contributors to the distribution of intake estimates.

4.6.4 Risk evaluation

Risk evaluations are performed using regulatory 'safe' levels: UK Soil Guideline Values for exposure assessment; and UK Tolerable Daily Soil Intake and Index Dose for intake assessments.

In addition, due to the probabilistic and spatial nature of this methodology, 'decision criteria' are used for classifying the land or the population as at-risk or not. For example, for an area to be classed as clean, the probability of exceeding the regulatory threshold must not exceed a stipulated critical probability. Similarly, for a population to be classed as not at-risk, the proportion of the population exceeding the tolerable intake must not exceed a stipulated percentage.

4.6.5 Risk communication

An ongoing theme is presentation of the results in a meaningful manner. The probabilistic spatial risk assessments undertaken are very complicated and ways for summarising and filtering out the results in a clear manner that a risk manager can understand is difficult. Histograms and box and whisker plots can be used to present some distributions, while spatial visualisation of results can be presented in map form.

The development of the methodology is described in the course of the following chapters and is demonstrated using case studies of different sizes and with different levels of detail. To start, the following chapter details an exposure assessment (Level 1, Figure 4.2) and associated risk evaluation, and demonstrates the uncertainties involved due to decisions taken during the exposure assessment and the implications of them.

Chapter 5 Exposure assessment using geostatistics

5.1 Introduction

Chapter 4 described two levels at which risk assessment can be carried out. Level 1 entails assessing soil contaminant concentrations (exposure assessment), and level 2 assesses human intake of contaminants (intake assessment). The procedure for both methods begins with collection and chemical analysis of limited soil samples (Figure 5.1) and for the probabilistic, spatially-evaluated risk assessment methodology proposed in this thesis, is followed by a geostatistical analysis to produce estimates of soil contaminant levels across the region of concern.

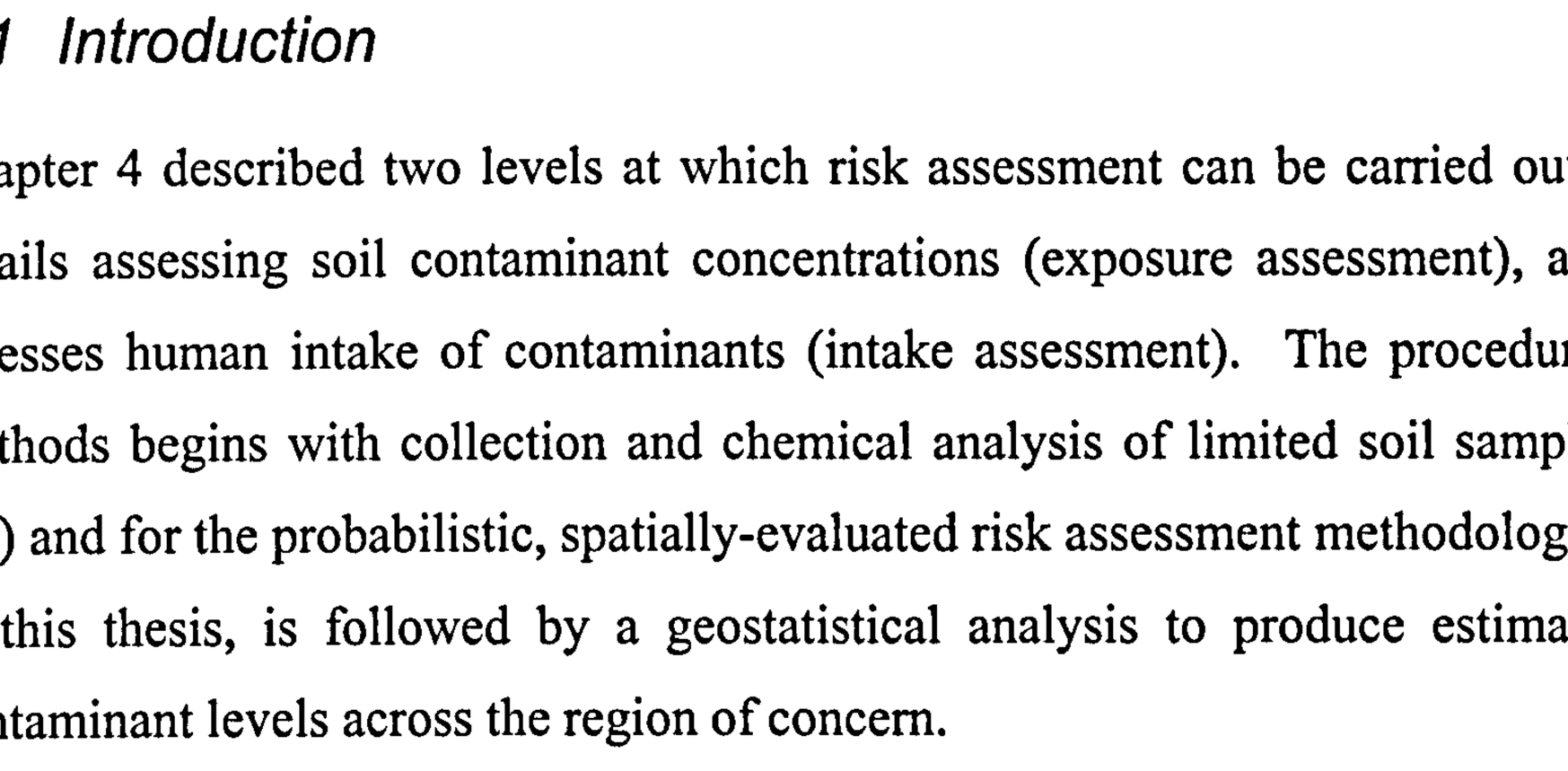


Figure 5.1. Two levels for assessment of risks to human health from contaminated land.

The focus of this chapter is to illustrate, through a case study, the uncertainty that can be introduced through the geostatistical procedure and how it can lead to differences in the results of an exposure assessment. Various sources of uncertainty are addressed: uncertainty due to the variogram modelling is assessed by fitting 3 sets of models to the data which have differing levels of complexity; uncertainty due to the support size is looked into by working with two different support sizes; and uncertainty due to the interpolation method is investigated by comparing results from IK (indicator kriging) and SIS (sequential indicator simulation). Uncertainty due to sampling and chemical analysis cannot be taken into account, since the data are unavailable, though the indicator methods can be adapted to take this into account in other circumstances. Implications for exposure assessment are assessed in two ways: by looking at changes in the probabilities of exceeding regulatory thresholds; and in terms of changes in actual estimated soil concentrations. Since such estimated soil contaminant concentrations may eventually be fed into an intake assessment model, a brief comment is made on the implications for the intake assessment.

5.2 Background to the data set

The soil data set for this case study comes from the South-Eastern part of Attiki peninsula, about 60 km from Athens in Greece. The region has a long history of mining and has been exploited for silver for more than 25 centuries. The main town is Lavrio, around which there are over a thousand ancient mining shafts which were used to explore the subsurface for mixed sulphide ores of lead, zinc and iron. The main ore bodies run in approximately a North North West-South South East direction. Millions of tons of ore have been mined and processed in the region over many centuries and today, as a consequence, enormous heaps of waste remain from the mining and metallurgical work carried out there.

Unfortunately, relevant population data for the region was unavailable, so it was not possible to extend this analysis to an intake assessment as well. However, this comprehensive soil data set was used for the thorough exposure assessment study because it displays good spatial structure in the variograms and leads to a clear demonstration of the exposure assessment method and uncertainties which can affect the results. This is in agreement with the fact that in the Lavrio region both natural geochemical processes and human activities have had a large impact on the soil make-up over many centuries and these processes have produced spatial continuities that can

be discerned in the data during the spatial analysis. This is in contrast to the data sets utilised for the intake assessments, where little structure is found.

420 top soil samples were collected for a previous study on a fairly regular and roughly 500x500m grid over an area of approximately 120km^2 (Korre, 1997). They were not collected with a risk assessment in mind at the time and no re-sampling was possible for validating the current study. Samples were analysed for 24 elements, including heavy metals, and the chemical analysis was one of nearly total decomposition using a combination of perchloric and nitric acids. Full details can be found in Korre (1997).

Lead (Pb) in soil was studied because it is known to have detrimental effects on human health, such as reduction in child IQ, increased risk of miscarriage and effects on the blood and nervous system (DEFRA & EA, 2002e) and is found at high concentrations in topsoil in the Lavrio region. The histogram of Pb sample values is highly positively skewed with some extreme high data points (Figure 5.2). Figure 5.3 illustrates the spread of sample points in space, with symbols in proportion to soil Pb concentration. The unsampled area just to the south east of the centre was inaccessible military land at the time of sampling.

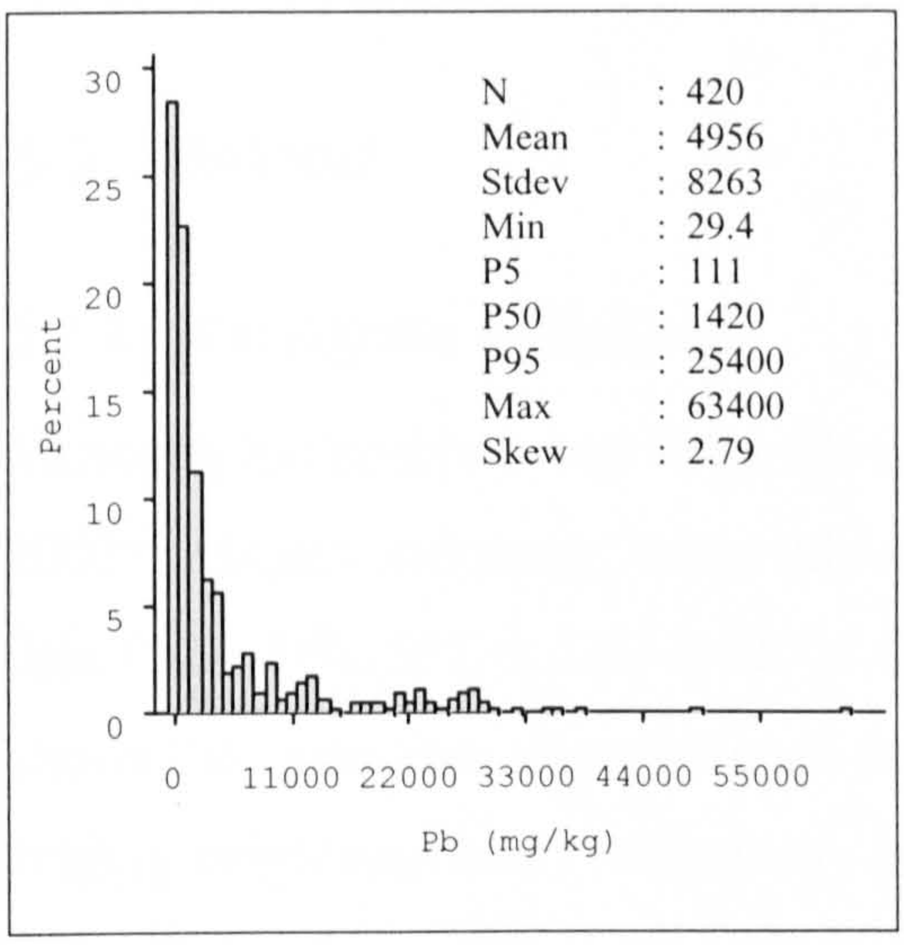


Figure 5.2. Histogram of soil sample Pb concentrations (mg.kg^{-1}).

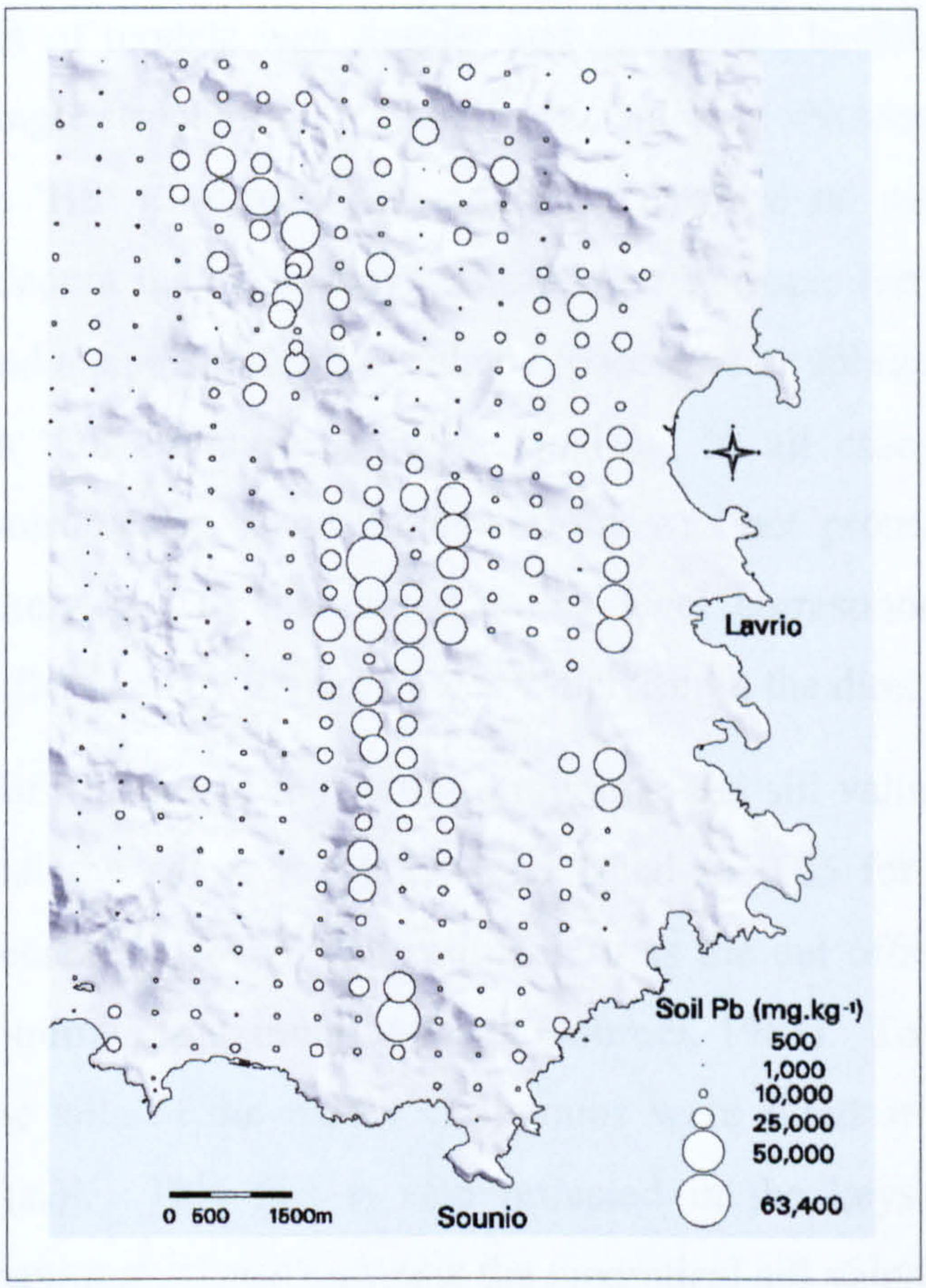


Figure 5.3. Soil sample points with an indication of soil Pb concentration.

5.3 Method

5.3.1 Variogram modelling

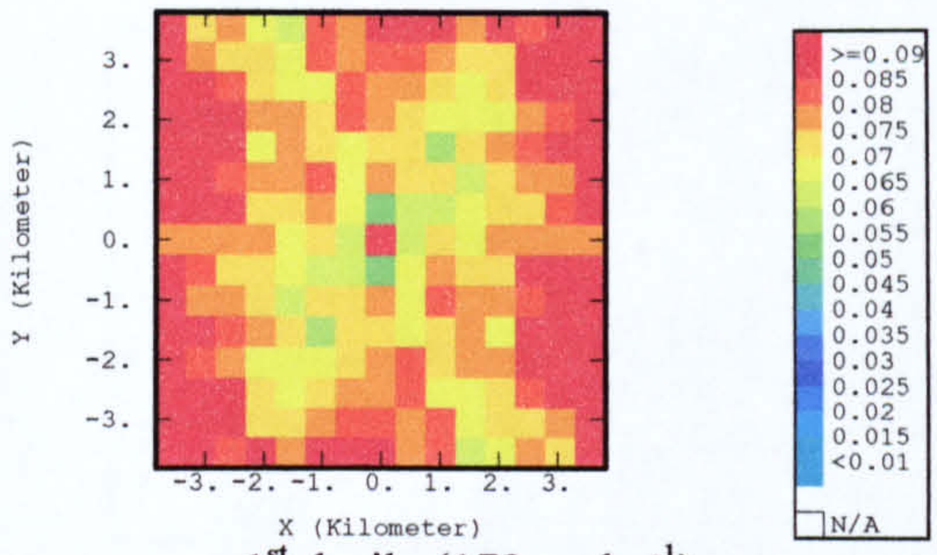
Variographic analysis was undertaken using the Isatis software package (Geovariances, 2002). Binary indicators were calculated at the nine decile values of the raw sample data (178; 319; 567; 1,020; 1,420; 2,045; 3,660; 7,255; and 16,200mg.kg^{-1}). Figure 5.4 shows the variogram maps for each cut off out to a lag distance of 3km, as is used in the kriging neighbourhood. Anisotropy is evident and appeared more pronounced at higher cut off levels. This was interpreted by eye at an angle of approximately 10° anticlockwise from the main North-South and East-West directions.

Three sets of models, with varying complexity, were fitted to the sample variograms in order to investigate uncertainty due to the choice of variogram model. First, a series of bi-directional nested models were fitted, which account for the anisotropy observed. The nested structures are comprised of a nugget, a spherical and an exponential component, and are referred to hereafter as 'BN' (bi-directional, nested). The second

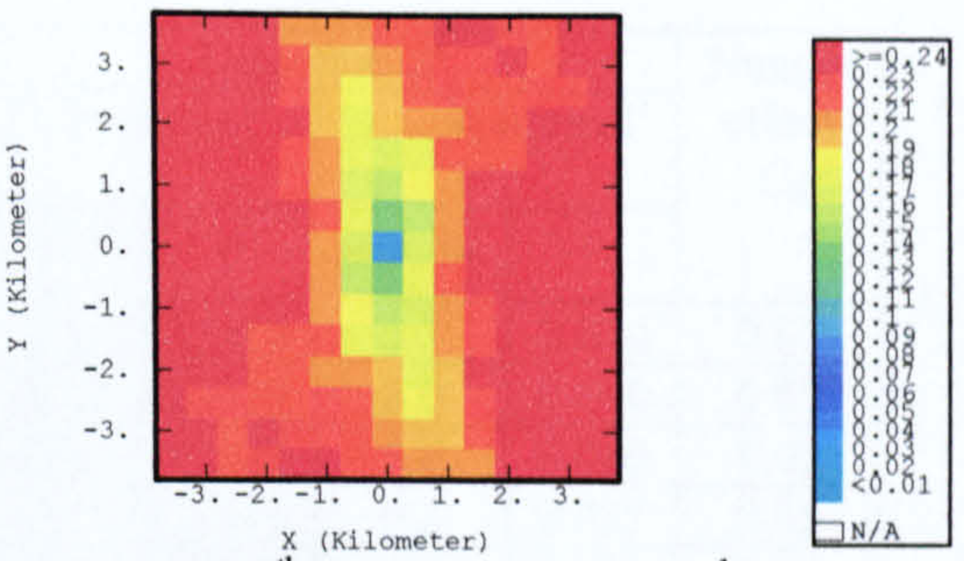
set of models was simpler and comprised bi-directional models with a nugget and a single structure only, either exponential or spherical as appropriate, hereafter referred to as 'BS' (bi-directional, single). The third set was simpler still and did not take into account the anisotropy. These were isotropic (omni-directional) models with a nugget and a single structure, either exponential or spherical as appropriate, and are referred to as 'OS' (omni-directional, single). In all cases, the lowest cut off was modelled isotropically, since the anisotropy was not pronounced at this cut off level. This is likely due to the fact that this level corresponds to background levels and is not influenced by the ore bodies which run in the direction of anisotropy.

For a series of indicator variograms, the sill values should follow order relations such that the sill is maximum and equal to 0.25 for the median cut off (5^{th} decile) and decreases from 0.25 towards zero as the cut offs depart from the median towards the minimum/maximum cut offs (Journel, 1983). To ensure this order relation holds true, the sills of the model variograms were fitted at their theoretical sill values, $F(z_c)[1-F(z_c)]$. This fact is also reflected in the keys of the variogram maps, where the maximum classification is the theoretical sill value for that cut off.

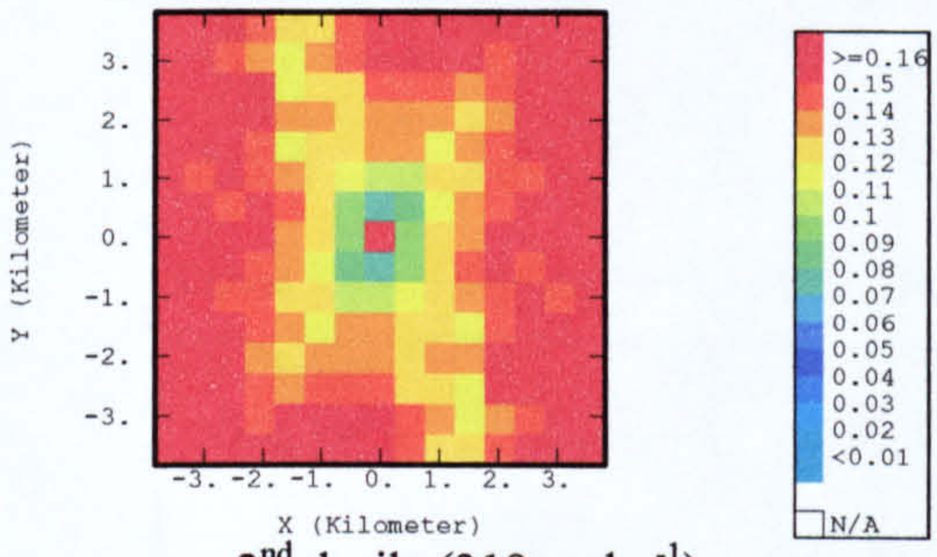
The model parameters are summarised in Table 5.1a (BN), b (BS) and c (OS). Figure 5.5 (BN), Figure 5.6 (BS) and Figure 5.7 (OS) show the sample and fitted model variograms. The same models are used for both IK and SIS processes.



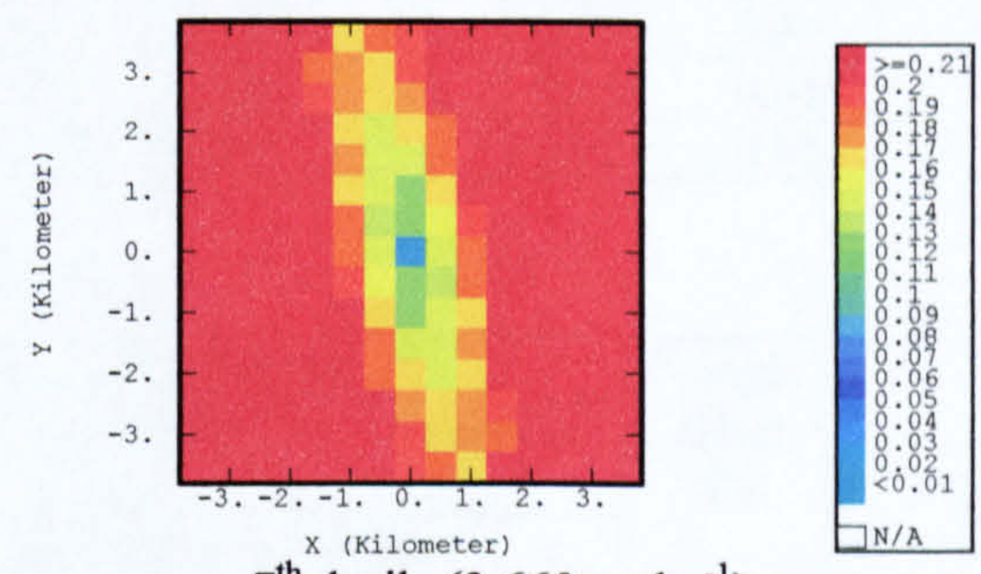
1st decile (178mg.kg^{-1})



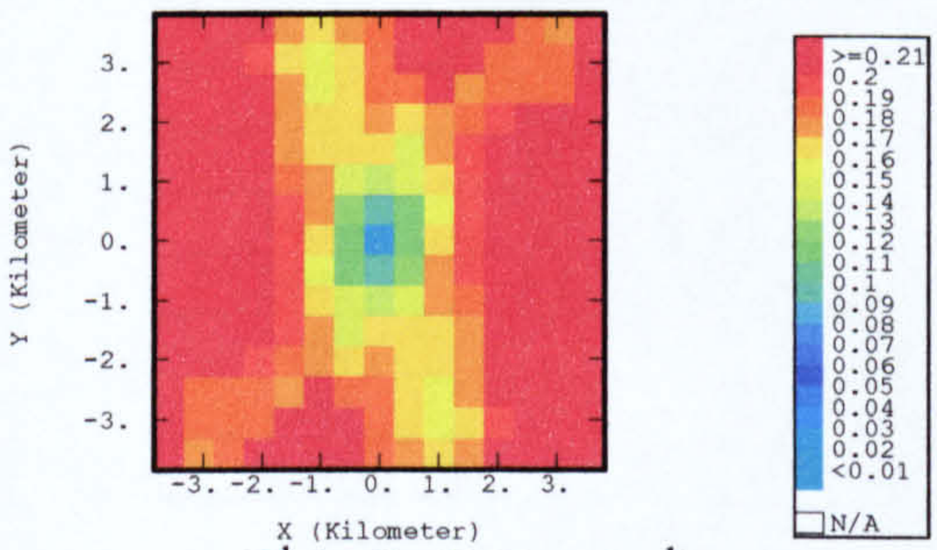
6th decile (2,045mg.kg^{-1})



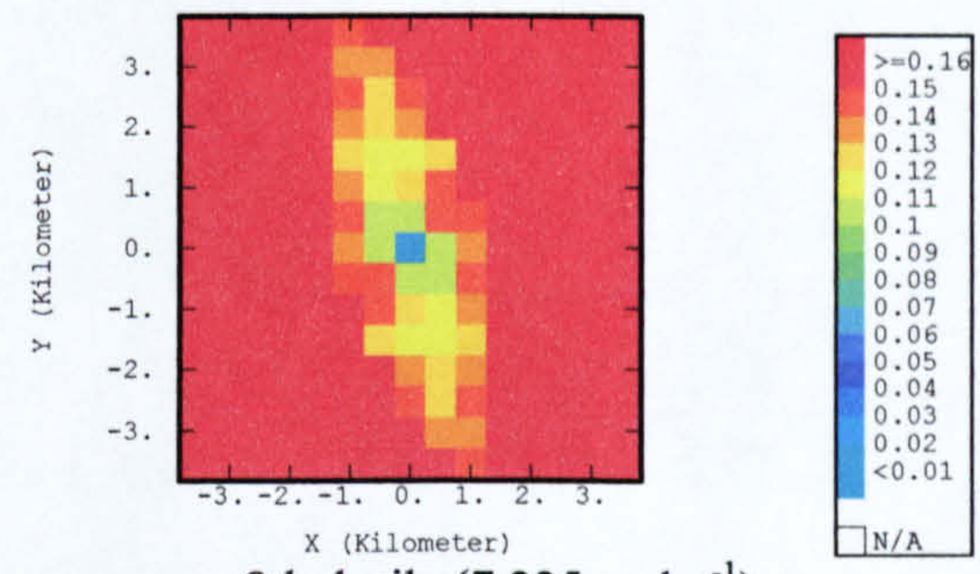
2nd decile (319mg.kg^{-1})



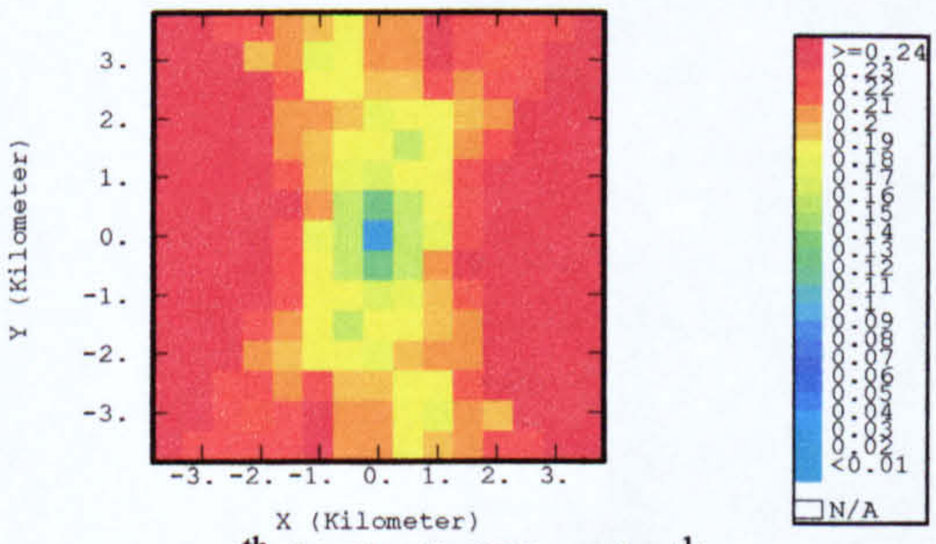
7th decile (3,660mg.kg^{-1})



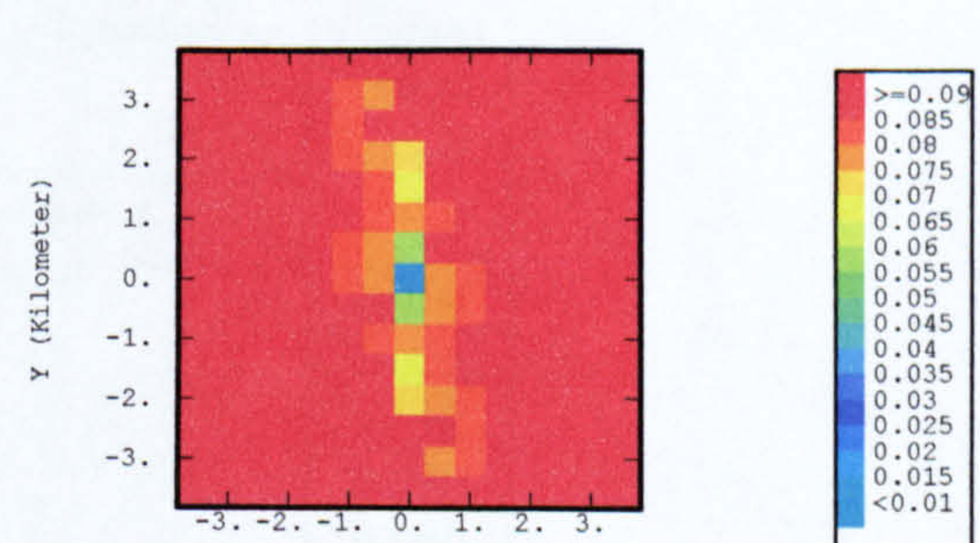
3rd decile (567mg.kg^{-1})



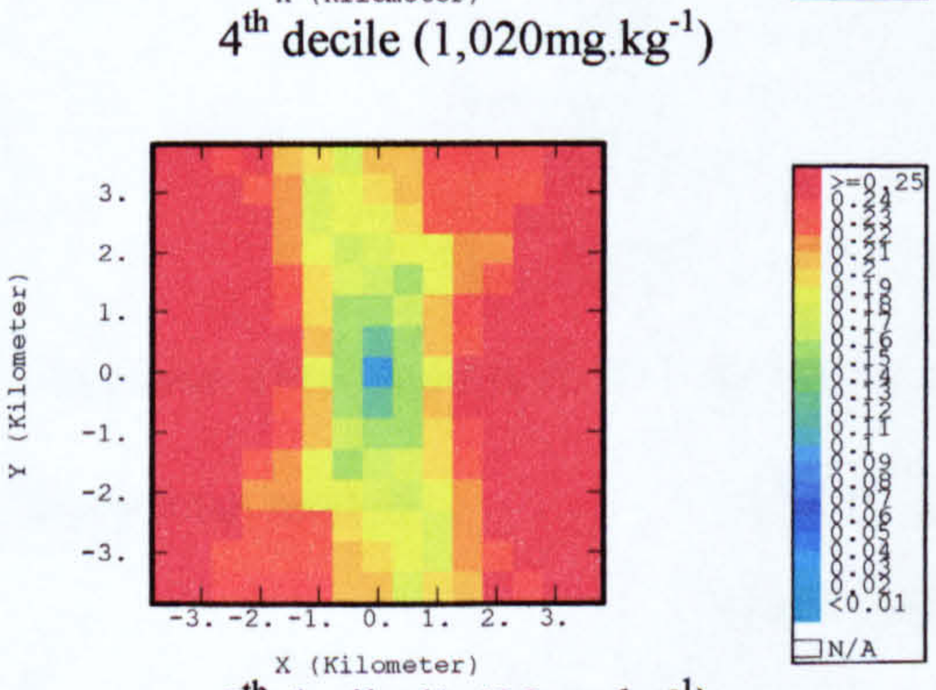
8th decile (7,225mg.kg^{-1})



4th decile (1,020mg.kg^{-1})



9th decile (16,200mg.kg^{-1})



5th decile (1,420mg.kg^{-1})

Figure 5.4. Variogram maps for all cut offs demonstrating an anisotropy, strongest at the highest cut offs.

Table 5.1a. Indicator variogram model structure (BN).

Cutoff ($mg.kg^{-1}$)	Direction 1[a]				Direction 2[a]				Nugget effect C_0
	Exponential[b]		Spherical[c]		Exponential[b]		Spherical[c]		
	a(km)	c	a(km)	c	a(km)	c	a (km)	c	
178	2.5	0.020	3.5	0.020	2.5	0.020	3.5	0.020	0.05
319	3.2	0.070	2.0	0.040	2.8	0.070	6.0	0.040	0.05
567	2.6	0.100	2.0	0.060	3.2	0.100	5.8	0.060	0.05
1020	2.9	0.105	0.9	0.085	1.0	0.105	6.2	0.085	0.05
1420	1.5	0.105	2.0	0.095	1.5	0.105	6.4	0.095	0.05
2045	1.5	0.101	1.7	0.089	1.4	0.101	5.0	0.089	0.05
3660	1.5	0.100	1.4	0.060	2.5	0.100	3.5	0.060	0.05
7255	1.2	0.050	1.4	0.050	2.1	0.050	3.0	0.050	0.06
16200	1.2	0.015	0.6	0.025	2.8	0.015	1.6	0.025	0.05

Table 5.1b. Indicator variogram model structure (BS).

Cutoff ($mg.kg^{-1}$)	Direction 1[a]				Direction 2[a]				Nugget effect C_0
	Exponential[b]		Spherical[c]		Exponential[b]		Spherical[c]		
	a(km)	c	a(km)	c	a(km)	c	a (km)	c	
178	3.8	0.04			3.8	0.04			0.05
319	3.0	0.10			5.0	0.10			0.06
567	2.2	0.15			4.0	0.15			0.06
1020	2.2	0.15			4.8	0.15			0.09
1420	2.4	0.16			5.1	0.16			0.09
2045	1.9	0.14			4.0	0.14			0.10
3660			1.5	0.11			3.5	0.11	0.10
7255			1.5	0.08			3.0	0.08	0.08
16200			1.0	0.03			2.2	0.03	0.06

Table 5.1c. Indicator variogram model structure (OS).

Cutoff ($mg.kg^{-1}$)	Exponential[b]		Spherical[c]		Nugget effect C_0
	a(km)	c	a(km)	c	
178	3.8	0.04			0.050
319	3.7	0.11			0.050
567	3.1	0.15			0.060
1020	3.4	0.15			0.090
1420	3.3	0.17			0.080
2045	2.5	0.16			0.080
3660			2.2	0.110	0.100
7255			2.0	0.075	0.085
16200			1.6	0.025	0.065

[a] Directions: Direction 1 is E10°N (minimum continuity); Direction 2 is N10°W (maximum continuity)

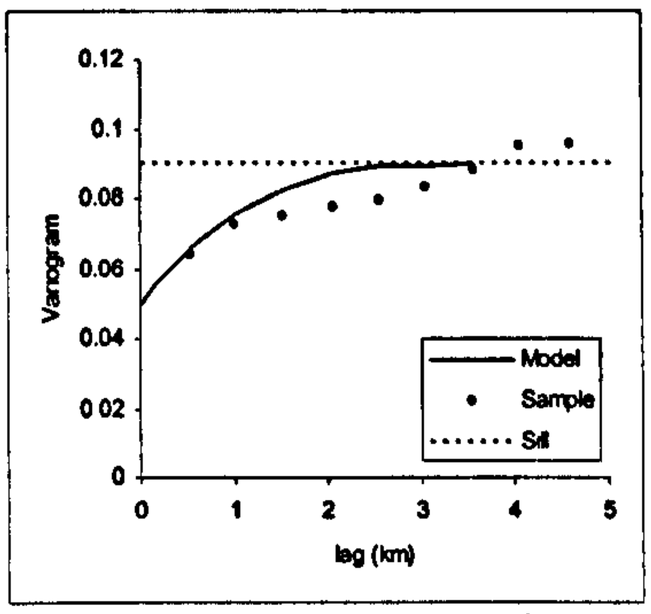
[b] Exponential model: $$\gamma(h) = C_0 + c.\left[1 - \exp\left(-\tfrac{3h}{a}\right)\right]$$

where h is the lag distance

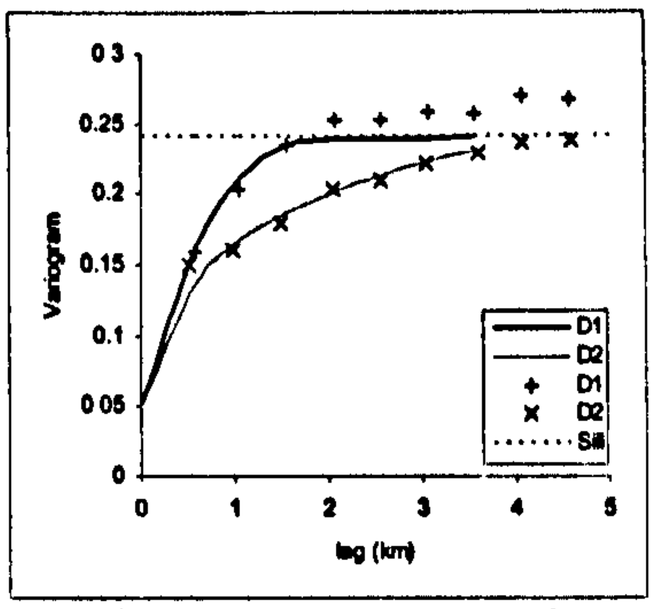
[c] Spherical model: $$\gamma(h) = C_0 + c.\left[1.5\tfrac{h}{a} - 0.5\left(\tfrac{h}{a}\right)^3\right] \quad \text{if } h \le a$$

$$\gamma(h) = C_0 + c \quad \text{if } h > a$$

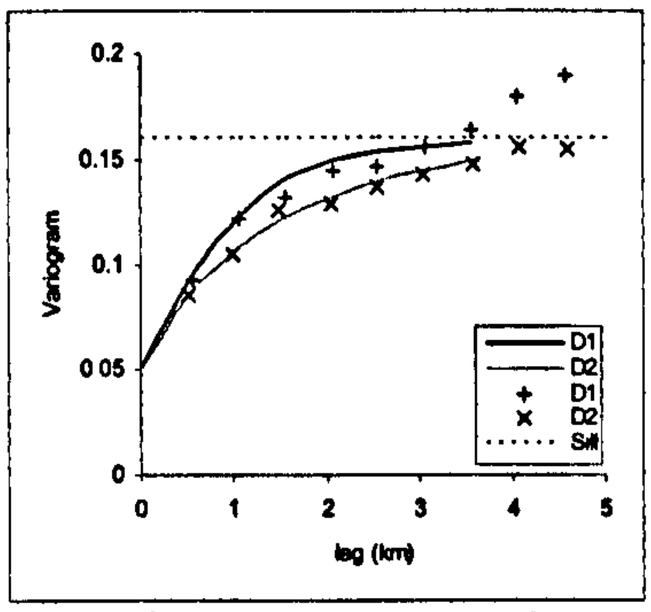
where h is the lag distance



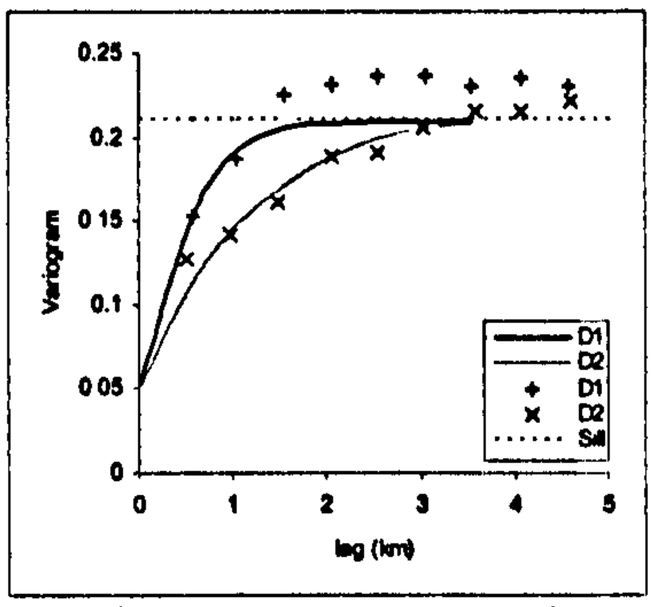
1st decile (178mg.kg-1)



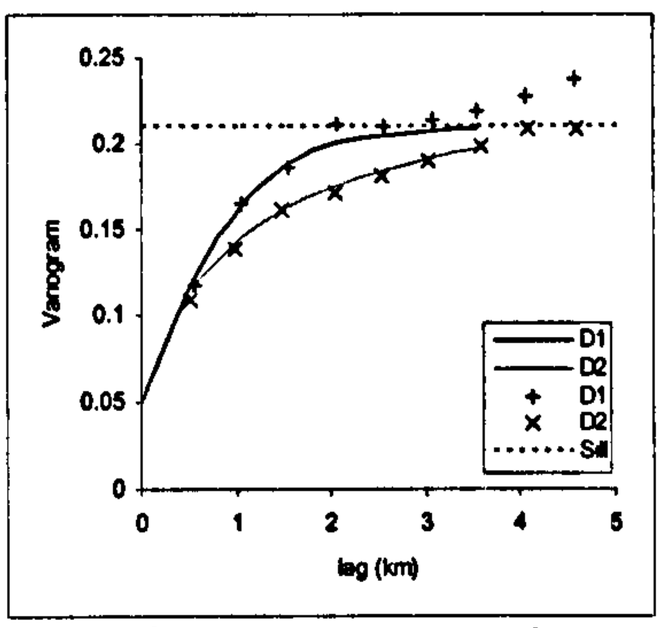
6th decile (2,450mg.kg-1)



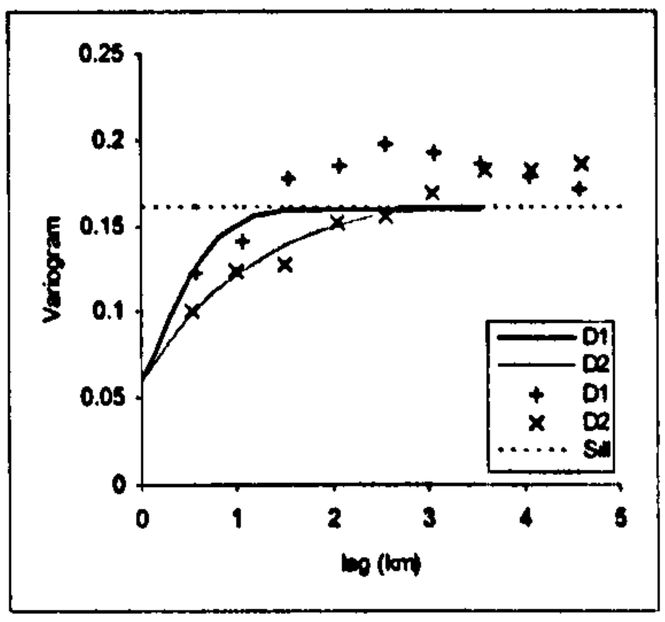
2nd decile (319mg.kg-1)



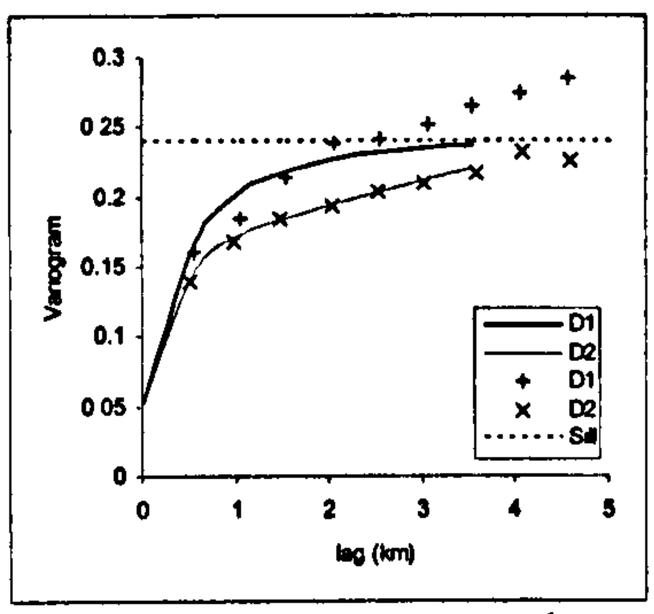
7th decile (3,660mg.kg-1)



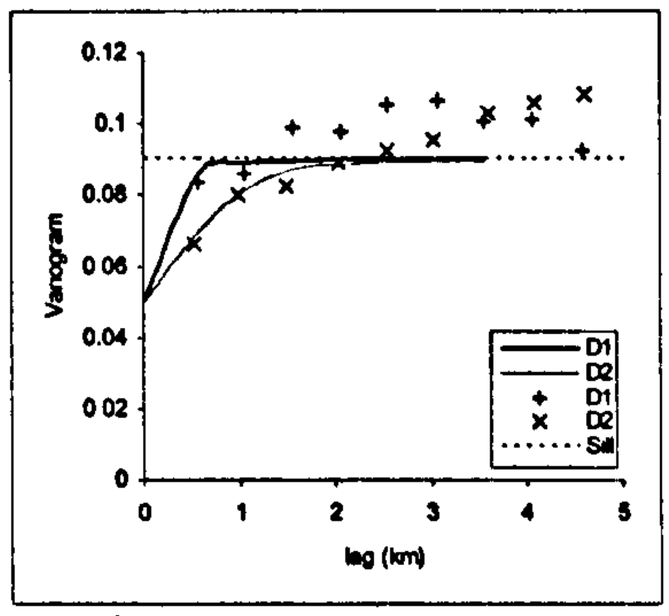
3rd decile (567mg.kg-1)



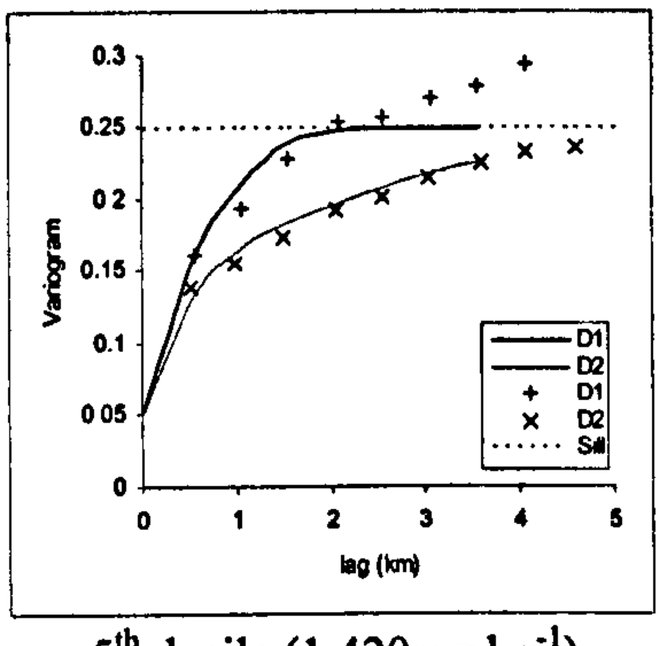
8th decile (7,225mg.kg-1)



4th decile (,1020mg.kg-1)



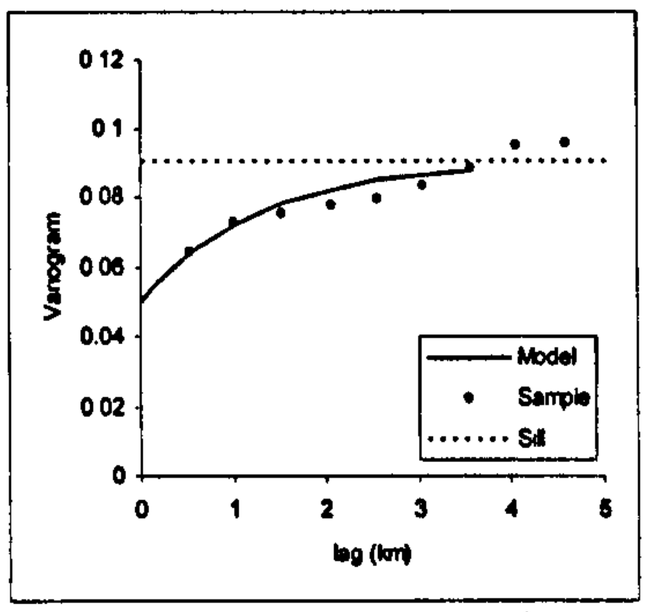
9th decile (16,200mg.kg-1)



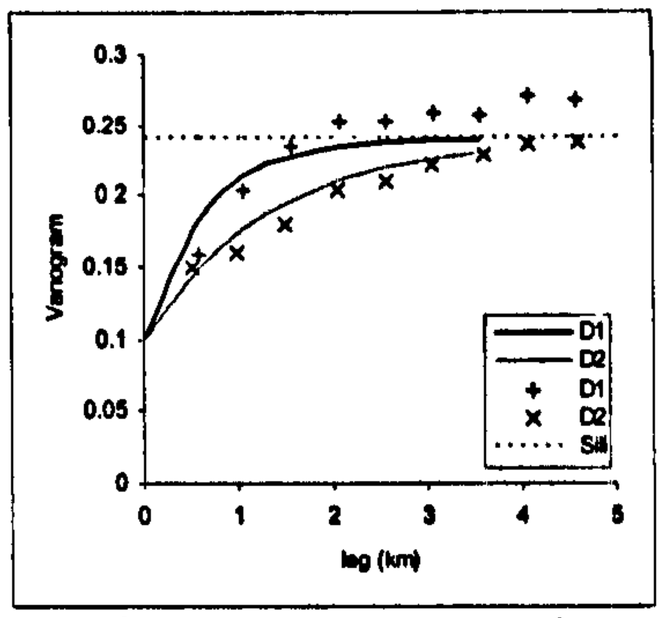
5th decile (1,420mg.kg-1)

D1=Direction 1; E10°N
D2=Direction 2; N10°W

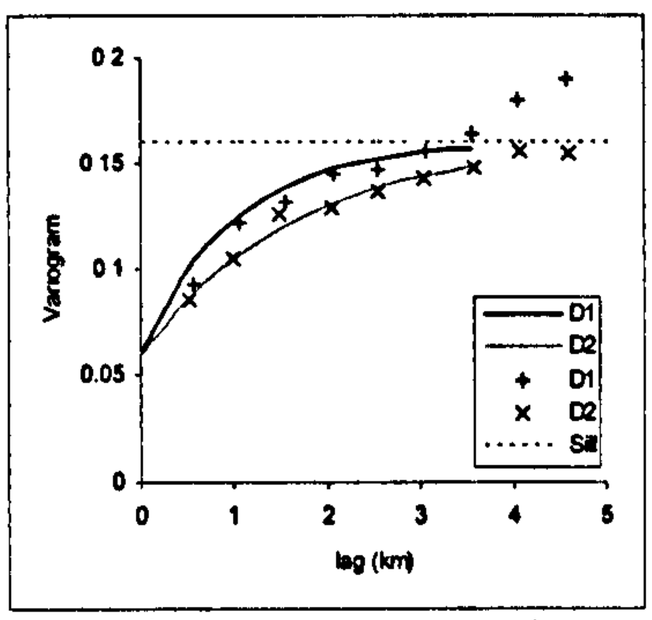
Figure 5.5. Sample variograms (symbols) and bi-directional nested model variograms (lines); (BN)



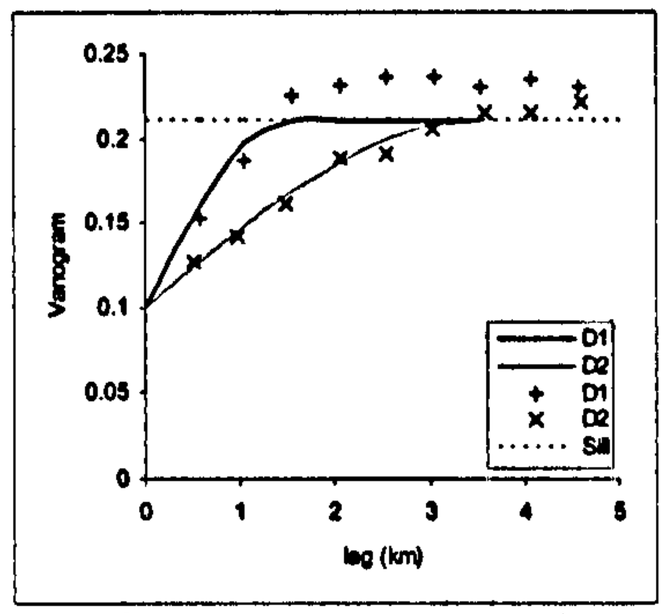
1st decile (178mg.kg^{-1})



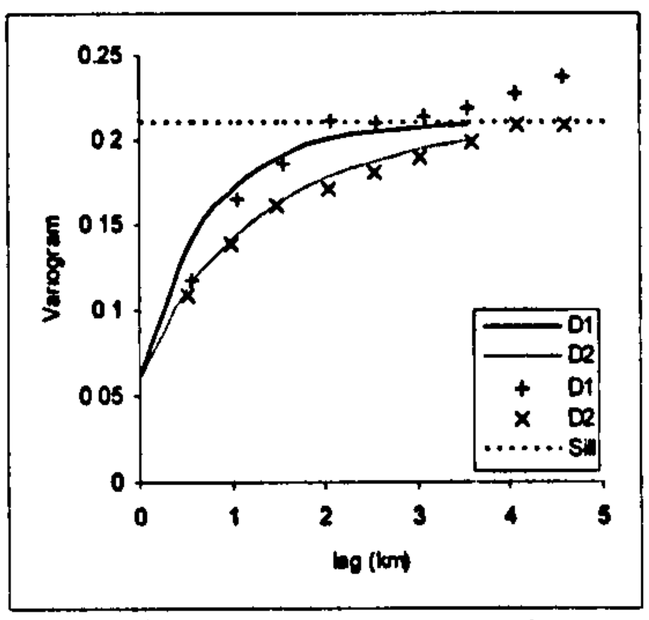
6th decile (2,450mg.kg^{-1})



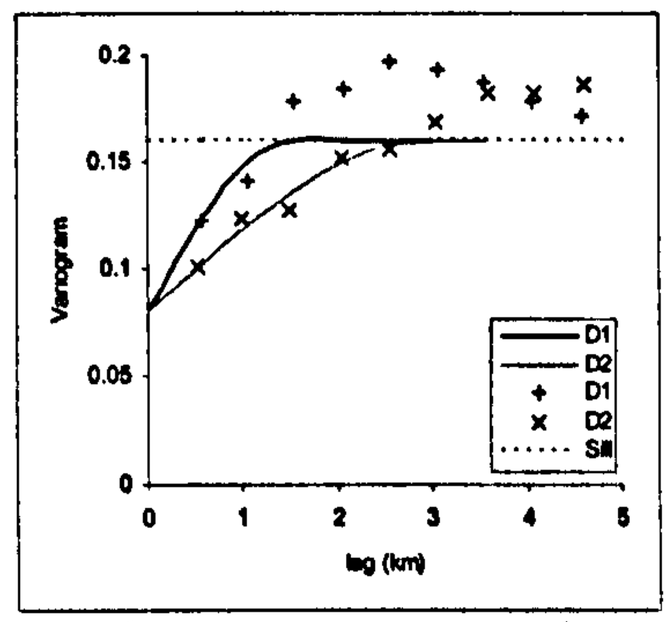
2nd decile (319mg.kg^{-1})



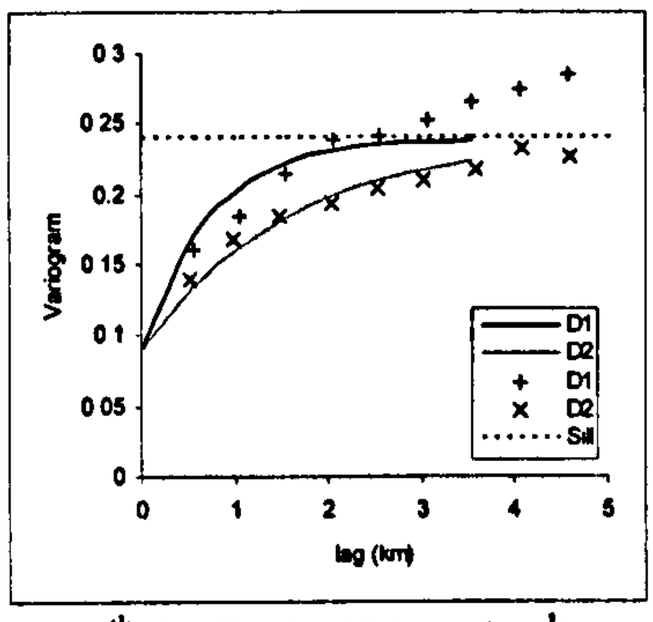
7th decile (3,660mg.kg^{-1})



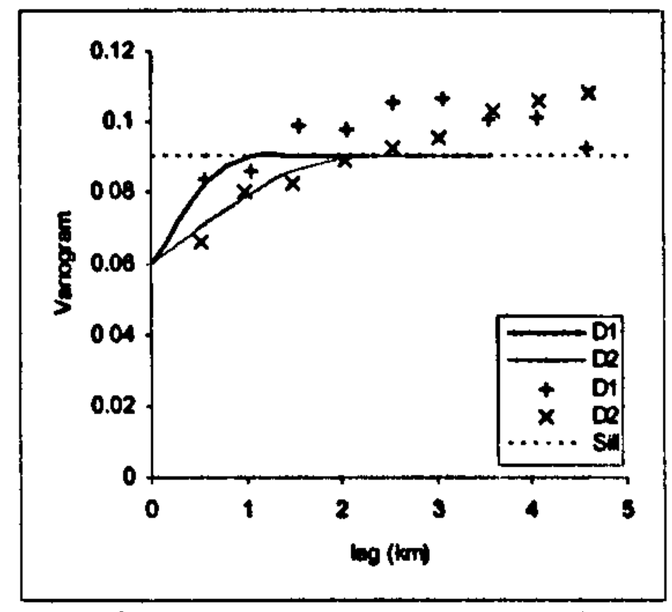
3rd decile (567mg.kg^{-1})



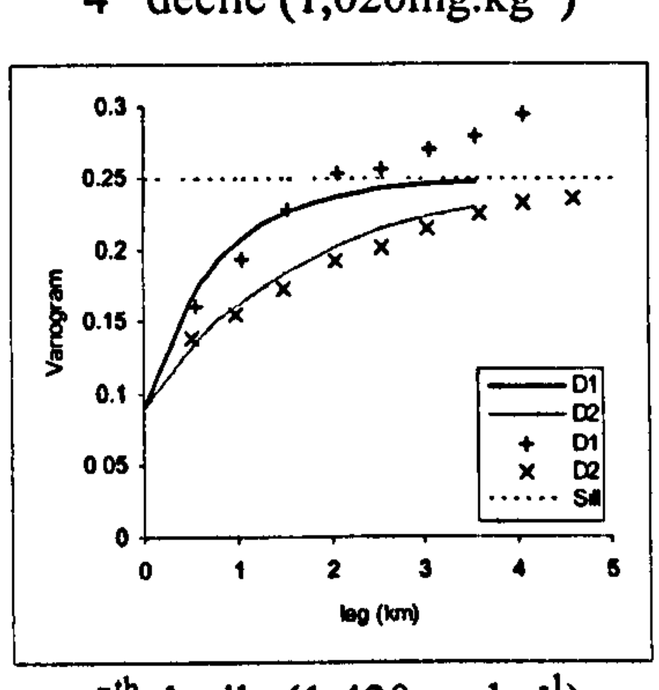
8th decile (7,225mg.kg^{-1})



4th decile (1,020mg.kg^{-1})



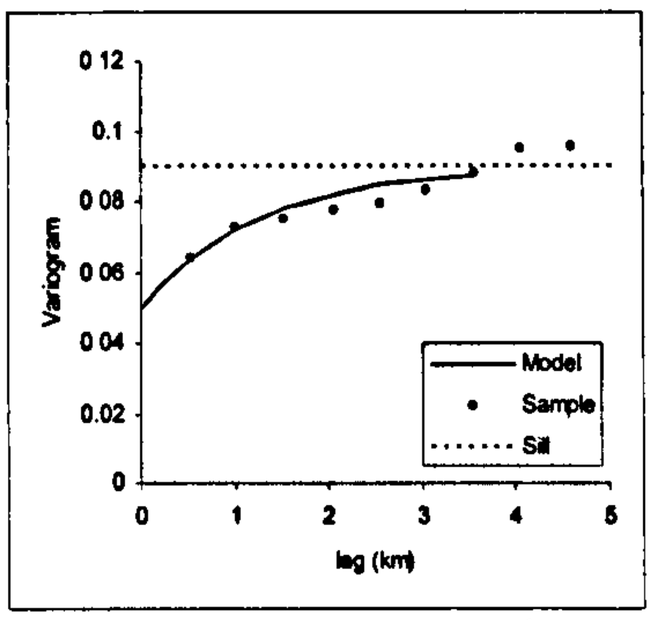
9th decile (16,200mg.kg^{-1})



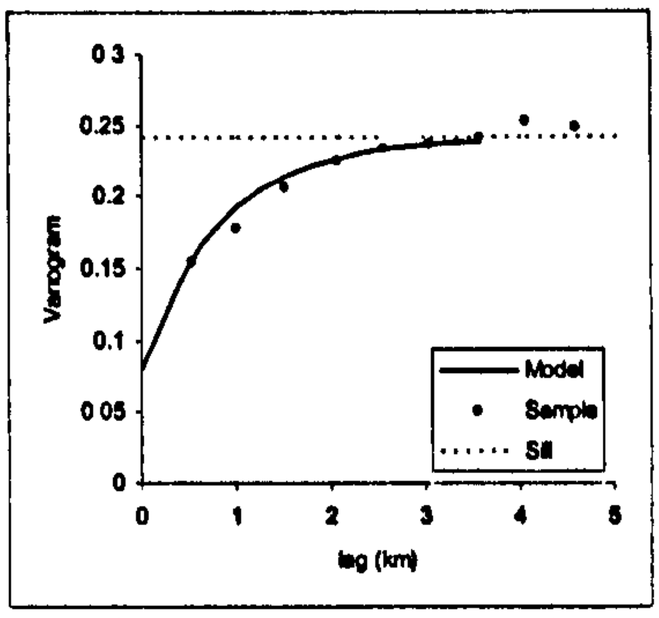
5th decile (1,420mg.kg^{-1})

D1=Direction 1; E10°N
D2=Direction 2; N10°W

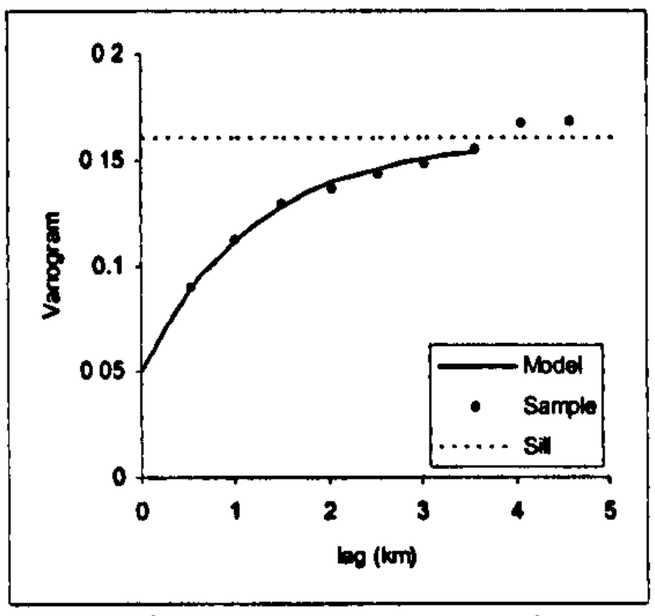
Figure 5.6. Sample variograms (symbols) and bi-directional single structure models (lines); (BS)



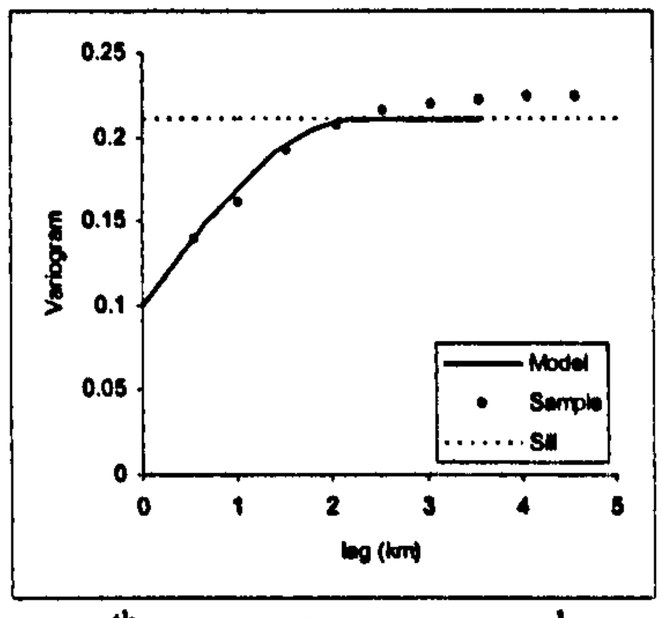
1st decile (178mg.kg^{-1})



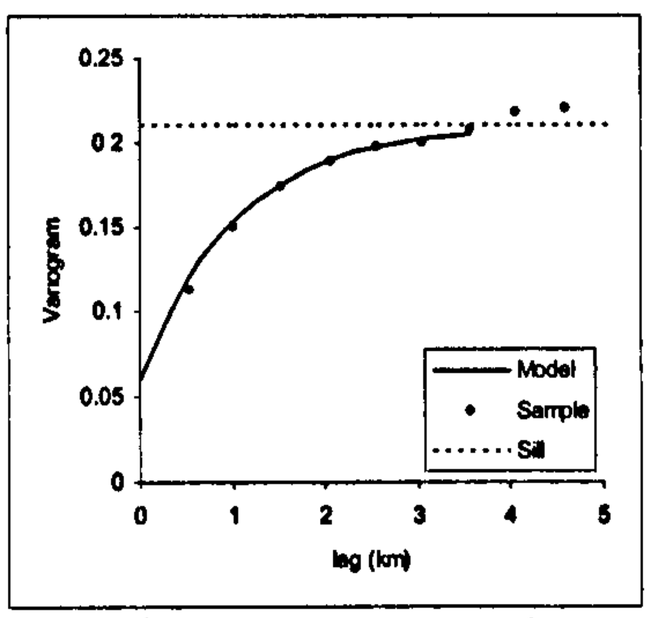
6th decile (2,450mg.kg^{-1})



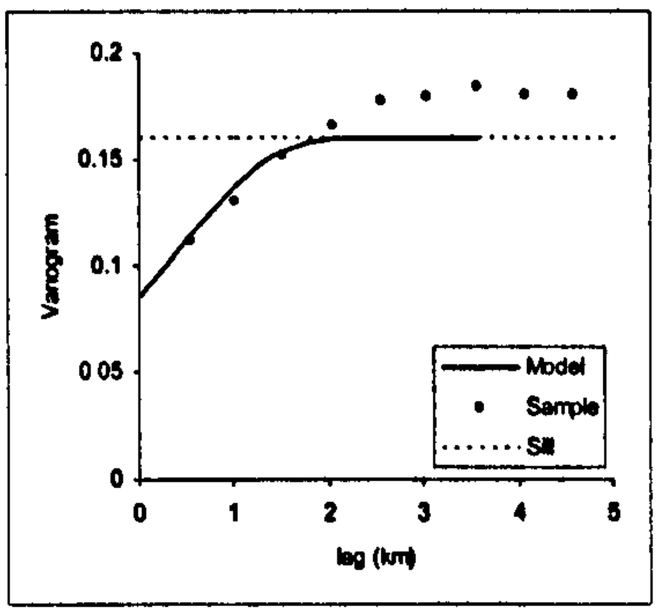
2nd decile (319mg.kg^{-1})



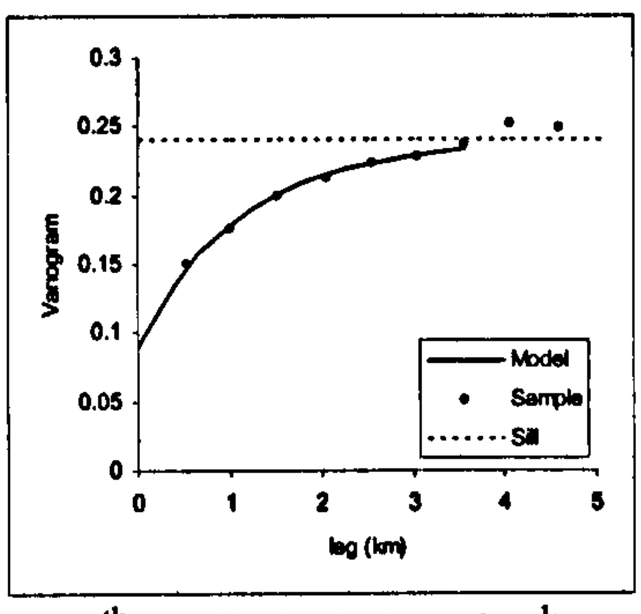
7th decile (3,660mg.kg^{-1})



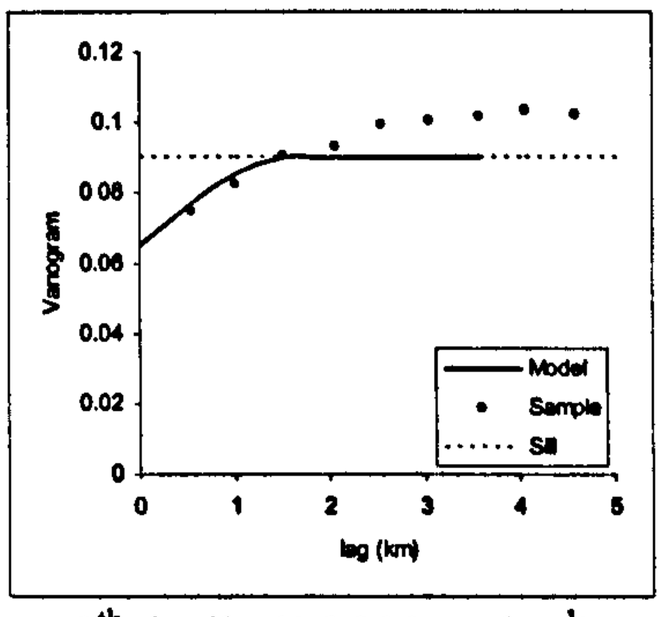
3rd decile (567mg.kg^{-1})



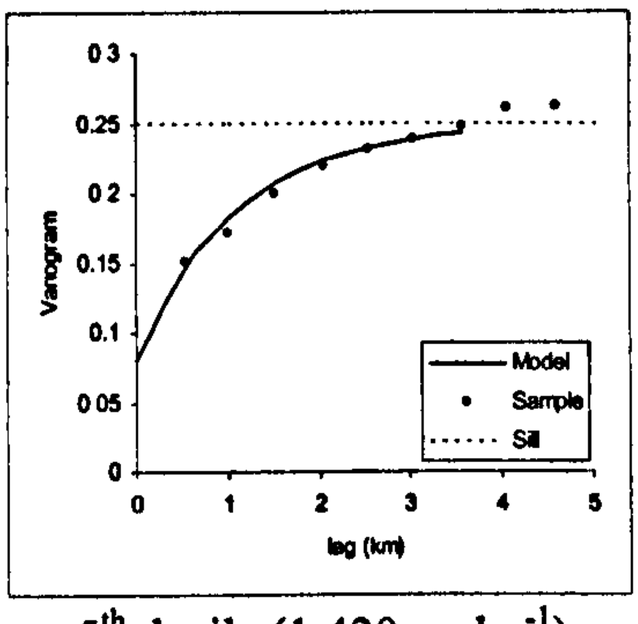
8th decile (7,225mg.kg^{-1})



4th decile (1,020mg.kg^{-1})



9th decile (16,200mg.kg^{-1})



5th decile (1,420mg.kg^{-1})

Figure 5.7. Sample variograms (symbols) and omni-directional single structure models (lines); (OS)

5.3.2 Selection of grid size and moving window neighbourhood

To increase the resolution of the soil Pb concentration information, IK and SIS were performed on two grid sizes: 250x250m and 25x25m. These choices are subjective. Soil samples were collected on a 500x500m grid so the larger grid cell was simply set to half the spacing of the samples. The smaller grid divides the larger spacing by 10 and each large cell contains 100 small cells. Details of the coarse and fine grid are given in Table 5.2.

Table 5.2. Grids used for kriging and simulation. Coordinates represent the central points around which a block is estimated. The coarse and fine grids overlay exactly.

.		number of nodes	coordinates of origin	spacing
Coarse grid	X direction	33	-87.5m	250m
	Y direction	54	-10737.5m	250m
Fine grid	X direction	330	-200m	25m
	Y direction	540	-10850m	25m

For the IK and SIS calculation process, a moving window neighbourhood was used with a radius of 3km, which was of a size sufficient to capture an adequate number of sample points from which to estimate unsampled locations. For IK, the minimum and maximum sample data points used in the estimation process were set to 8 and 27 respectively (a general rule of thumb (Deutsch & Journel, 1998) is to use three to five times the number of fixed cut offs: 3x9=27). For SIS, a mix of sample values and previously simulated nodes is used. A maximum of 10 sample and 17 previously simulated nodes was stipulated.

5.3.3 Support correction

IK and SIS are point estimation methods. Since the size of the 25x25m cell is very small compared to the size of the site (approx. 8x13km) the point estimates are taken to be representative of the 25x25m block. To estimate 250x250m block averages block corrections were performed. The procedure is different for IK and SIS.

Point IK was performed on the 250x250m grid. In order to adjust the point estimate to a block estimate, affine correction was used with a variance reduction factor, f (Isaaks & Srivastava, 1989). f was calculated using a variogram model of the Pb sample data, according to the method of Isaaks & Srivastava (1989) as follows. A sample and model variogram were produced using Isatis (Geovariances, 2002) (Figure 5.8). The model variogram was used to find the average variogram value for the pairs of points in the

region as a whole, and the average variogram value for pairs of points in a 250x250m block. The region 8.25km x 13.5km (A) and a block 250x250m (B) were discretised by more and more points until the average variogram values stabilised. f was calculated by $f = 1 - \frac{\sigma^2 B}{\sigma^2 A}$ and in this case, $f = 0.6351$. Applying the variance reduction factor to the point estimates gives 250x250m block estimates.

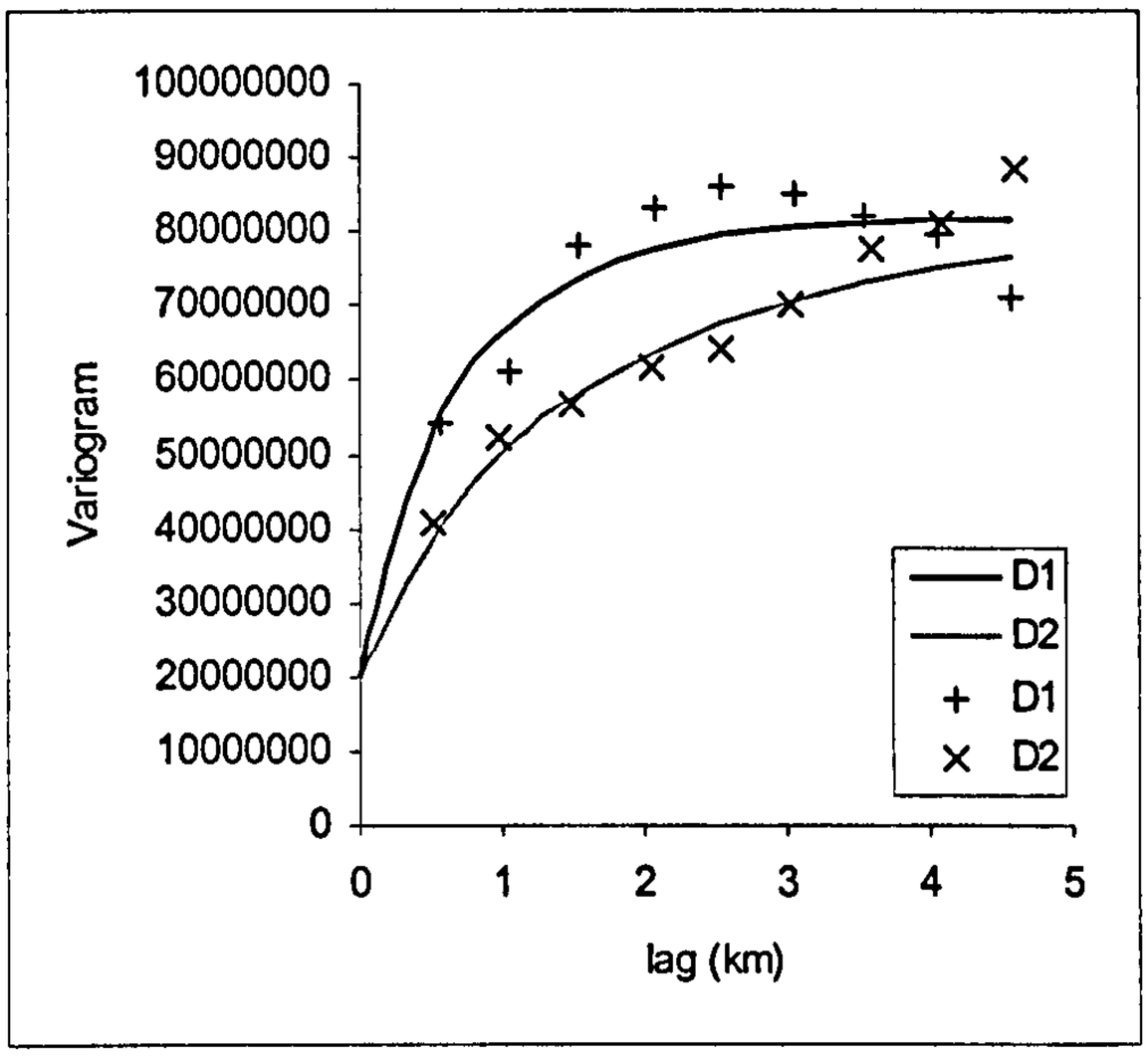


Figure 5.8. Sample variograms (symbols) and model variograms (lines) for Pb sample data.
D1=Direction 1; E10°N and D2=Direction 2; N10°W

For SIS, point simulated results were upscaled to 250x250m blocks by averaging the 100 point estimates in each block and taking the average value as a representative concentration for the whole block.

5.3.4 Performing IK

IK was performed using the GSLIB programme IK3D (Deutsch & Journel, 1998) for each set of models (BN, BS and OS) and both grid sizes (250x250m and 25x25m). IK produces ccdf values at nine discrete points which are the ccdf values at the selected cut off values. The full ccdf was built using the GSLIB POSTIK programme (Deutsch & Journel, 1998) and the minimum possible estimated Pb concentration was set to zero and the maximum to 65,000mg.kg^{-1}, assuming that it was unlikely that the sampling would have picked up the very smallest and very largest values. Linear interpolation was used between the nine discrete ccdf estimates. For extrapolation from the 1st decile ccdf value to zero a power model with shape parameter 2.5 was used, while a

hyperbolic model with shape parameter 3 was used to extrapolate from the 9th decile to the maximum 65,000mg.kg^{-1}. These parameters were made identical to the ones chosen for SIS (see next section for explanation).

Again using POSTIK, the probability of each block being over certain regulatory thresholds was calculated. In addition, a 95% confidence interval on the mean Pb value was calculated, represented by the 2.5th and 97.5th percentile estimates of the soil concentration value in each block.

5.3.5 Performing SIS

SIS was undertaken using the GSLIB programme SISIM (Deutsch & Journel, 1998). 100 simulations were made (based on the fact that cumulative means and variances of the estimates had stabilised) using each set of models (BN, BS and OS) on the 25x25m grid. Again, like IK, SIS produces ccdf values at nine discrete points. The full ccdf is built as an integral part of the SISIM routine. The parameters chosen for the extrapolation of the upper and lower tails of the ccdf are based on matching, as closely as possible, the histogram of simulated values to the histogram of the sample values, and are found by trial and error. The statistics for 100 simulations are averaged and compared to the raw data statistics in Table 5.3. Discrepancies in the tails are due to the minimum and maximum allowable simulated values being lower and higher, respectively, than the sample values.

Table 5.3. Statistics for Pb sample data (samp) and simulated values (sim) ('000mg.kg^{-1}).

	Mean	Stdev	Min	P1	P5	P10	P25	P50	P75	P90	P95	P99	Max
Samp	4.96	8.27	0.03	0.06	0.11	0.18	0.44	1.42	4.91	16.20	25.40	35.20	63.40
Sim	5.02	7.86	0.004	0.08	0.15	0.20	0.50	1.52	5.91	16.22	20.49	35.02	65.00

For results on the 250x250m grid, the 100 25x25m results within each 250x250m block were averaged. Results for both grids were processed using the GSLIB POSTSIM programme (Deutsch & Journel, 1998) in order to calculate the probability of exceeding the regulatory thresholds and to calculate the 95% confidence interval estimates of the Pb concentration for each block.

5.4 Risk assessment criteria

This exposure assessment uses generic risk criteria in the form of regulatory soil guideline values. Any measured concentration of Pb exceeding these guideline values

can be taken as an indication that further more detailed risk assessment should be carried out in these areas. Greece does not have its own regulatory guidelines so, as an initial indication of where excessive risks are, the UK Interdepartmental Committee on the Redevelopment of Contaminated Land (ICRCL) guideline values of 2000mg.kg^{-1} for public open spaces and 500mg.kg^{-1} for domestic gardens and allotments were adopted (ICRCL, 1987). At the time of analysis these threshold values were under review and have since been withdrawn after replacement in the UK by Soil Guideline Values (SGVs) (DEFRA & EA, 2002k). However, for the purposes of illustrating the methodology, this is not important, since the methods are valid whatever the regulatory threshold used and in any case, the estimated Pb levels do not depend on regulatory cut off values.

It is recognised that great care should be exercised in applying regulatory thresholds developed in one country to another. However, the purpose of this chapter is not to produce a comprehensive exposure assessment for the Lavrio region, but is aimed at demonstrating the uncertainties inherent in such an exercise. For this reason, it is sufficient to use the ICRCL guideline values.

Because the method deals with uncertainty, a decision criterion must be used which exploits this uncertainty measure. The results are produced in two forms: the 95% confidence interval Pb estimates are plotted; and the probability of exceeding the regulatory cut offs is plotted. The upper confidence limit (P97.5 Pb concentration) is a conservative upper estimate of the soil concentration which will give a reasonable worst case scenario estimate of the likely soil contaminant levels. This can be compared to the regulatory cut off levels to classify the land into contaminated or clean. The probability values may be plotted, but in order to classify areas as acceptable or unacceptably contaminated, a critical probability must be defined. For illustrative purposes, this assessment uses two critical probability values: 20% (stringent probability criterion) and 80% (weak probability criterion) i.e. if there is a probability of 20% (alternatively 80%) or more that the Pb concentration exceeds the regulatory threshold, land is classed as contaminated.

5.5 Analysis of results

5.5.1 Uncertainty due to the variogram model

Each 25x25m grid cell has a probability of exceeding each of the two regulatory thresholds. Histograms of the distributions of probabilities across the Lavrio region are given in Figure 5.9 for IK and Figure 5.10 for SIS. Each Figure shows the distribution calculated using the BN, BS and OS model sets and gives the percentage of land classified as contaminated under the stringent ($P \geq 0.2$) and weak ($P \geq 0.8$) criteria.

It is clear to see that there is little difference in the distributions between model types. This is especially true at the 500mg.kg^{-1} cut off (bottom row in each Figure) where the biggest range in values is 1.2% using the weak criterion and SIS. For the 2000mg.kg^{-1} cut off, the differences are larger, but still the largest range in values is only 5.2% using the stringent criterion and SIS.

Figure 5.11 shows the IK results in map form as an example. Under the weak criterion, smaller areas are classed as contaminated (dark grey) than under the stringent criterion (dark and pale grey inclusive). The pattern of contamination is very similar in all three cases. In map b (BS), some of the contaminated areas, e.g. the three spots on the left hand side of the map, appear more elongated in the direction of anisotropy, than in map c (OS) illustrating the effect of taking the anisotropy into account. This effect is not as pronounced in map a (BN). However, despite these small differences, for the purpose of a generic regional exposure assessment it appears that whether the model is complex and takes into account the anisotropy (BN) or simple and ignores the anisotropy (OS) or somewhere in between (BS) makes little difference to the outcome. This is true of both IK and SIS.

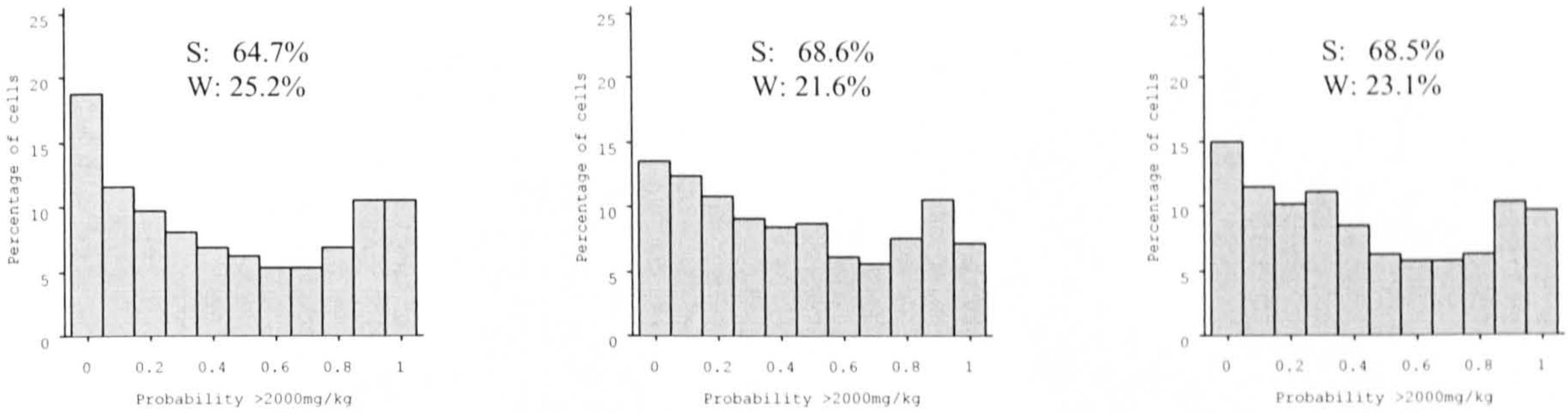


Figure 5.9. Frequency distribution of the probability of exceeding 2000mg.kg^{-1} (top) and 500mg.kg^{-1} (bottom) according to results of IK on a fine grid. S: percentage of land with probability ≥ 0.2 of exceeding threshold (stringent criterion). W: percentage of land with probability ≥ 0.8 of exceeding threshold (weak criterion).

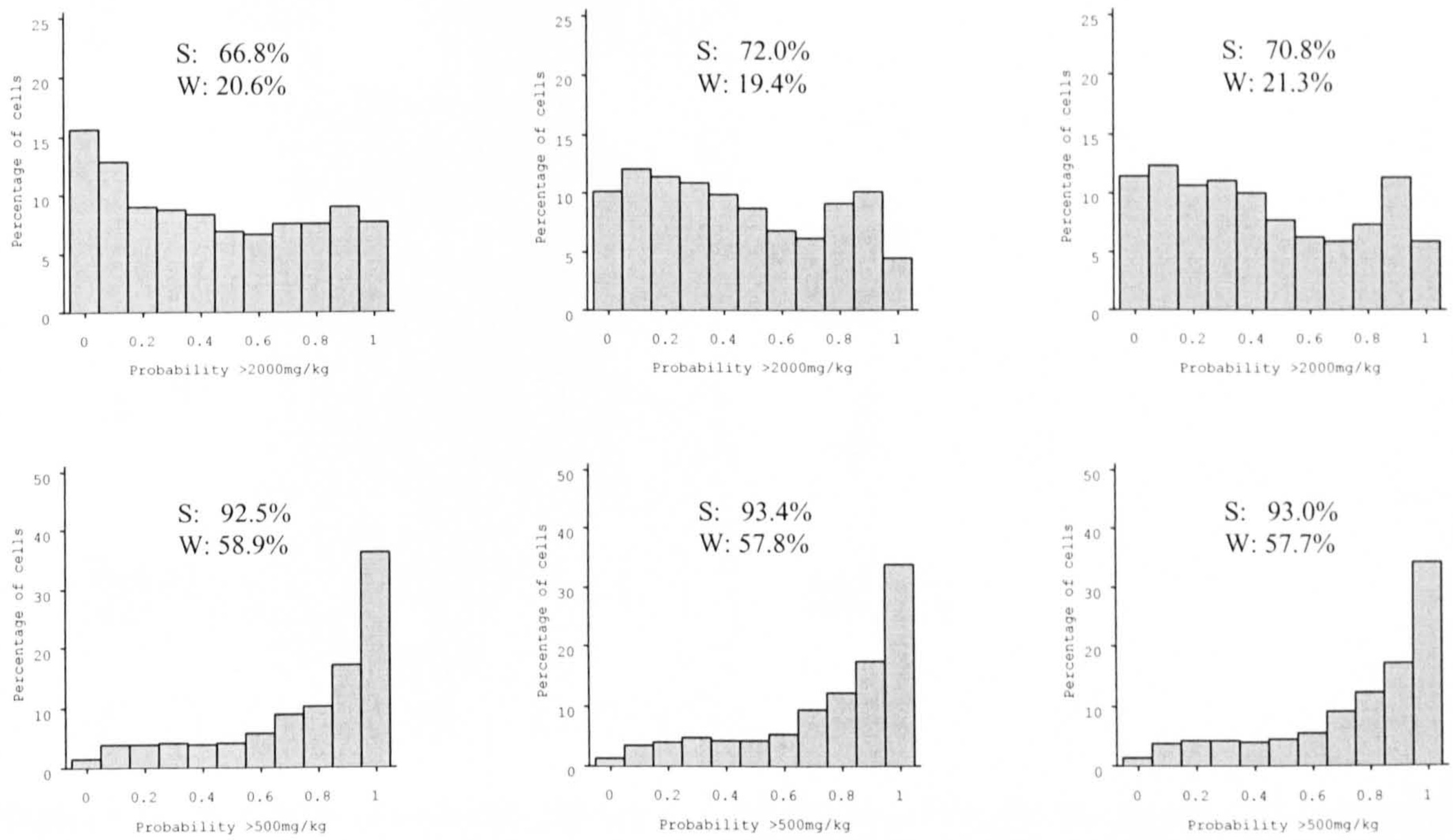


Figure 5.10. Frequency distribution of the probability of exceeding 2000mg.kg^{-1} (top) and 500mg.kg^{-1} (bottom) according to results of SIS on a fine grid. S: percentage of land with probability ≥ 0.2 of exceeding threshold (stringent criterion). W: percentage of land with probability ≥ 0.8 of exceeding threshold (weak criterion).

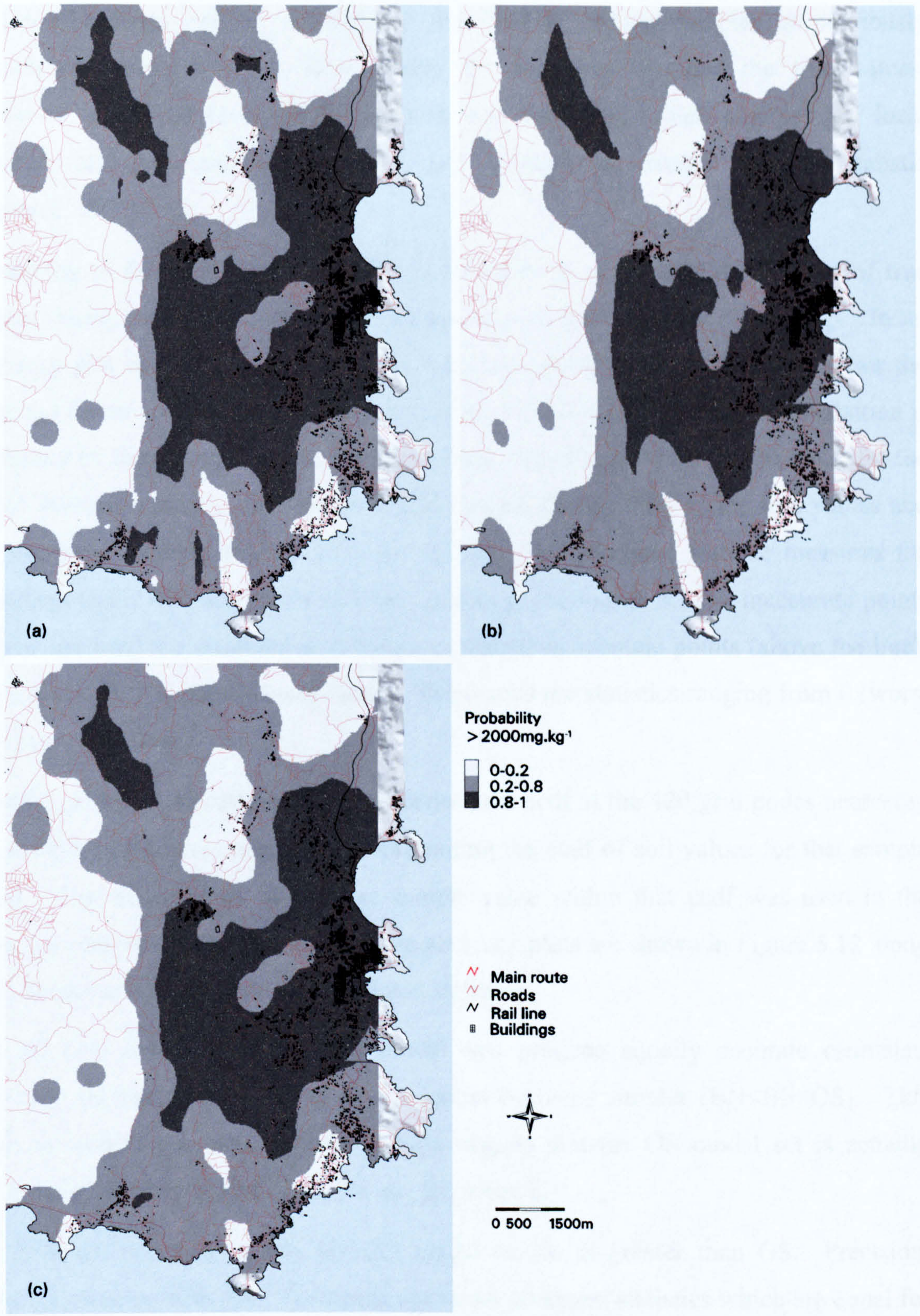


Figure 5.11. Probability of exceeding 2000mg.kg^{-1} using IK on a fine grid with (a) BN model set (b) BS model set and (c) OS model set. $P \geq 0.8$ represents land contaminated under the weak criterion and $P \geq 0.2$ represents land contaminated under the stringent criterion.

The results could not be validated in the field by taking samples at previously unsampled locations to see how closely the estimates matched the true values. However, a method proposed by Deutsch was employed which assesses the local accuracy and precision of the results, and calculates an overall goodness statistic (Deutsch, 1997).

According to this method, a probability distribution is accurate if the fraction of true values falling in the symmetric probability- or p-interval exceeds p for all p. In an accuracy plot (e.g. see Figure 5.12) the calculated points should fall on or above the diagonal for an accurate probability distribution. Precision of an accurate distribution is measured by the closeness of the fraction of true values to p for all p in [0,1] i.e. for the p=0.5 interval a precise distribution would contain exactly 50% of the true values and for the p=0.2 interval exactly 20%, and so on. The goodness statistic measures the closeness to the 45° line on an accuracy plot (e.g. see Figure 5.12). Inaccurate points (below the line) are weighted to be twice as distant as accurate points (above the line), since inaccuracy is more consequential. The results are statistics ranging from 0 (worst case) to 1 (best case).

In practical terms, to carry out this validation, the ccdf at the 420 grid nodes nearest to the sample points were selected as representing the ccdf of soil values for that sample point. The whereabouts of the true sample value within that ccdf was used in the accuracy and precision calculations. The accuracy plots are shown in Figure 5.12 along with the accuracy, precision and goodness statistics.

For IK (top row Figure 5.12) all model sets produce equally accurate estimates, however, the precision increases as the model becomes simpler (BN<BS<OS). This leads to overall goodness statistics which suggest that the OS model set is actually better than the more complicated BN and BS models.

For SIS, the accuracy of BN and BS model results is greater than OS. Precision, however, is lowest for BN. This leads to overall goodness statistics which are equal for BS and OS and both better than BN. Again, similar to IK, this suggests using a simpler model is better than a complicated one, and since omni-directional models are quicker to compute, this would be the best choice.

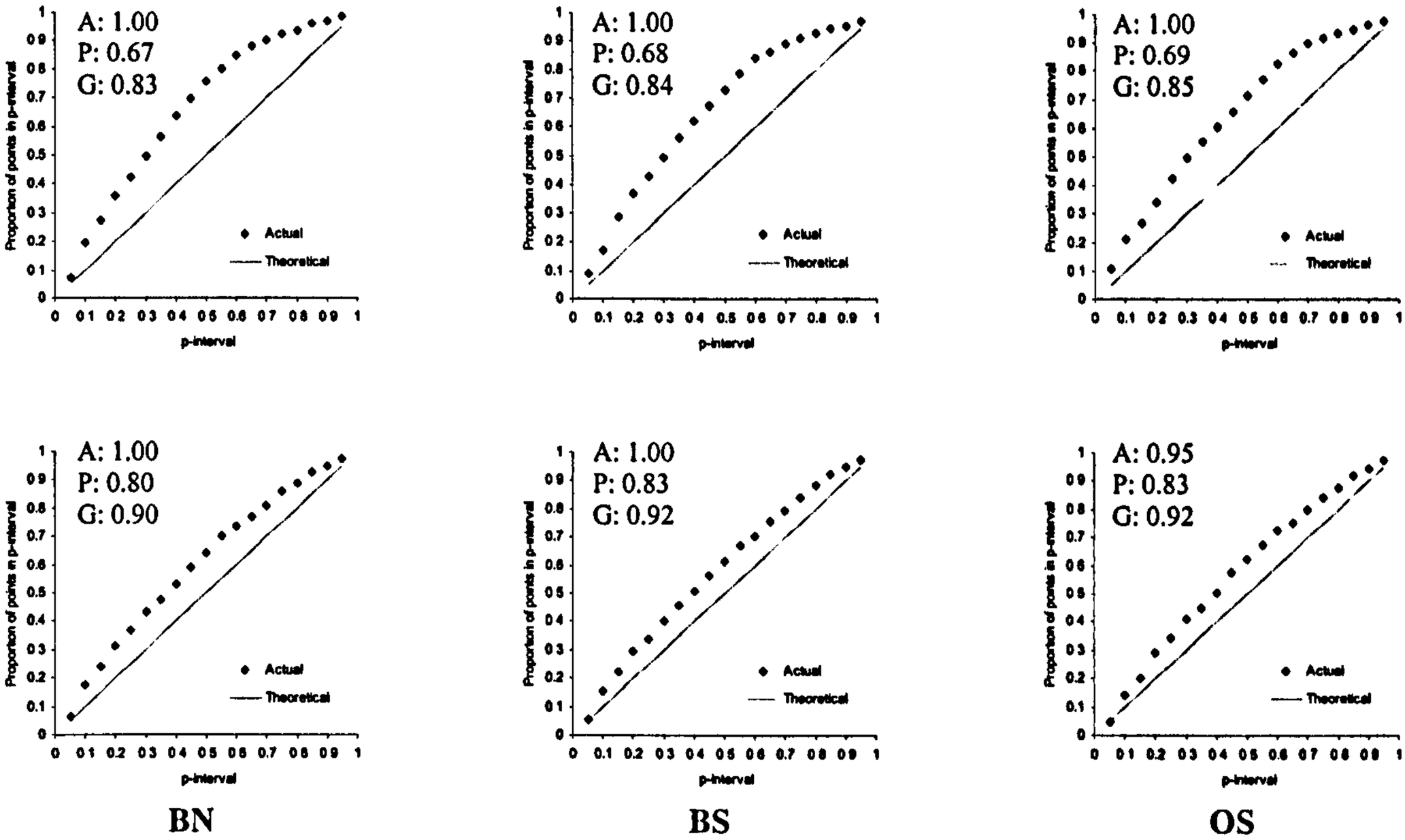


Figure 5.12. Accuracy plot for IK (top) and SIS (bottom) for the 3 model sets. A, Accuracy; P, Precision; G, Goodness.

Overall, it can be seen that for both IK and SIS there appears to be little effect on results output due to whether the BN, BS, or OS model sets are used. This is the case for both the classification of land as contaminated or clean, according to the stringent and weak probability criteria, and in terms of the spatial position of the highest and lowest risk regions. Since using more complex models makes the calculations slower, it seems advisable to fit as simple a model as is possible while, obviously, trying to capture the main shape of the variogram.

This case study bears out recommendations of other practitioners who warn that overly complicated structures are not necessary, since simpler ones often produce results that are just as satisfactory while being much quicker to compute (Deutsch & Journel, 1998; Isaaks & Srivastava, 1989).

5.5.2 Uncertainty due to support size

The design of a generic exposure assessment may be influenced by the end goal. For example, if the purpose is to highlight areas in need of clean up, the assessment may be designed with remediation in mind and the support size used would be chosen according to the most suitable remediation unit size. The purpose of this particular case study is to demonstrate the effect that these decisions may have on the classification of land as

contaminated or clean, and as such the grid cell sizes used for the analysis are rather arbitrary. However the effects demonstrated apply to all studies, to some degree.

Figure 5.13 (IK) and Figure 5.14 (SIS) show the changes in classification of contaminated land when working on the fine grid (25x25m) and coarse grid (250x250m) using the OS model set. It is clear to see that the shape of the distribution of probabilities changes from the fine grid (left) to the coarse grid (right). If using IK (Figure 5.13) with the stringent probability criterion, the classifications are virtually the same on both supports, however using the weak criterion, the contaminant problem looks worse on the coarse grid. For SIS, for both criteria, the picture looks worse when working on the coarse grid. This is because of the change in shape of the Pb concentration distribution on moving from the fine to the coarse grid.

As discussed in Chapter 3, working on a larger support has the effect of reducing the variance of the estimates, since large and small values are smoothed by averaging. Therefore, low values are averaged upwards and high values averaged downwards. This is demonstrated in Figure 5.15 and Figure 5.16 which show the bounds of the 95% confidence interval on the mean Pb concentration for each grid cell. Figure 5.15 shows the soil Pb concentrations as simulated on a fine grid and Figure 5.16 shows them after averaging them to work on a larger support (coarse grid). The differences are pronounced. On the fine grid, most of the lower confidence limit (LCL) values fall below the regulatory cut offs of 500 and 2,000mg.kg^{-1} while the upper confidence limit (UCL) values are mainly well above these and reach into the 40,000mg.kg^{-1} plus category. On the coarse grid, many of the LCLs are higher, falling into the 2-20,000mg.kg^{-1} category, while the UCLs are much smoother overall, mainly falling into the 2-20,000mg.kg^{-1} category. Hence the confidence intervals are much tighter on the coarse grid since the distributions are less varied.

The fact that, generally, the fine grid gives a less contaminated picture than the coarse grid is because the cut offs are both in relatively low percentiles of the Pb distribution (500mg.kg^{-1} $\approx$ 28th percentile, 2,000mg.kg^{-1} $\approx$ 59th percentile of sample data). Working on the larger support, low values are averaged upwards, so pushing more values above the cut offs and giving a worse picture of contamination. Even though higher values are averaged downwards, they are not pulled down sufficiently to pull them below the contaminant thresholds so do not affect the classification.

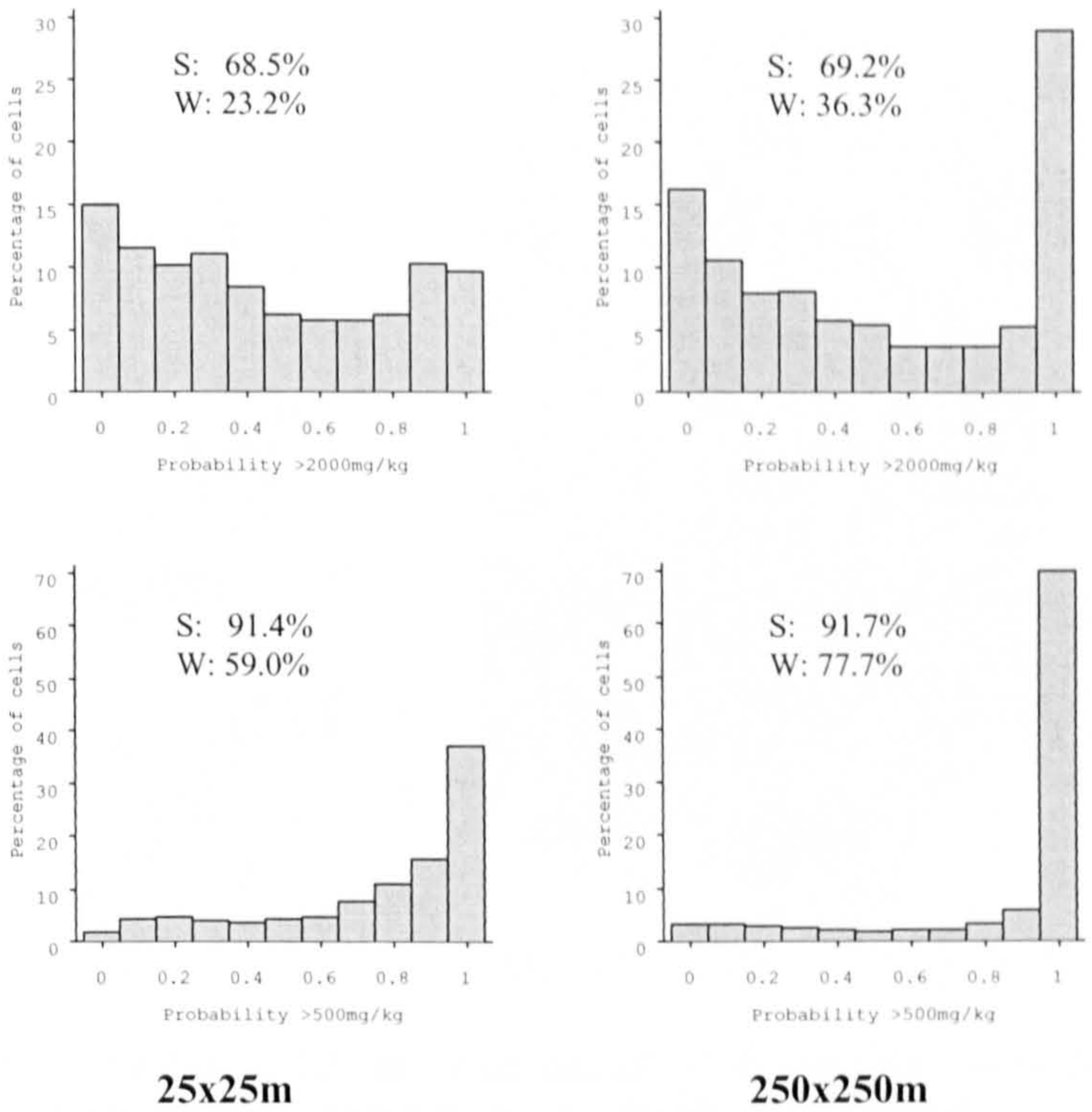


Figure 5.13. Frequency distribution of the probability of exceeding 2000mg.kg^{-1} (top) and 500mg.kg^{-1} (bottom) according to results of IK on a fine grid (left) and coarse grid (right). S: percentage of land with probability ≥ 0.2 of exceeding threshold (stringent criterion). W: percentage of land with probability ≥ 0.8 of exceeding threshold (weak criterion).

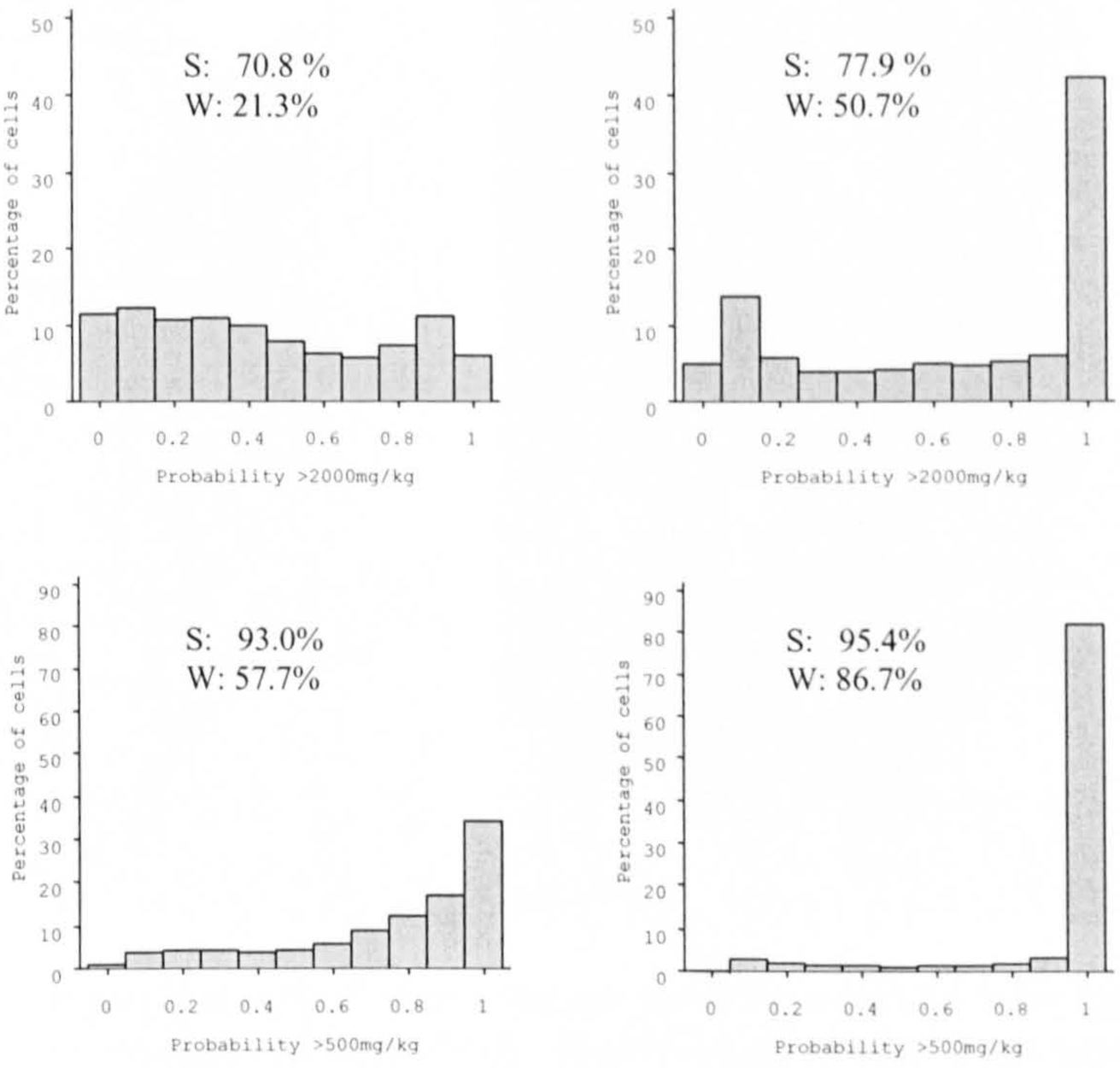


Figure 5.14. Frequency distribution of the probability of exceeding 2000mg.kg^{-1} (top) and 500mg.kg^{-1} (bottom) according to results of SIS on a fine grid (left) and coarse grid (right). S: percentage of land with probability ≥ 0.2 of exceeding threshold (stringent criterion). W: percentage of land with probability ≥ 0.8 of exceeding threshold (weak criterion).

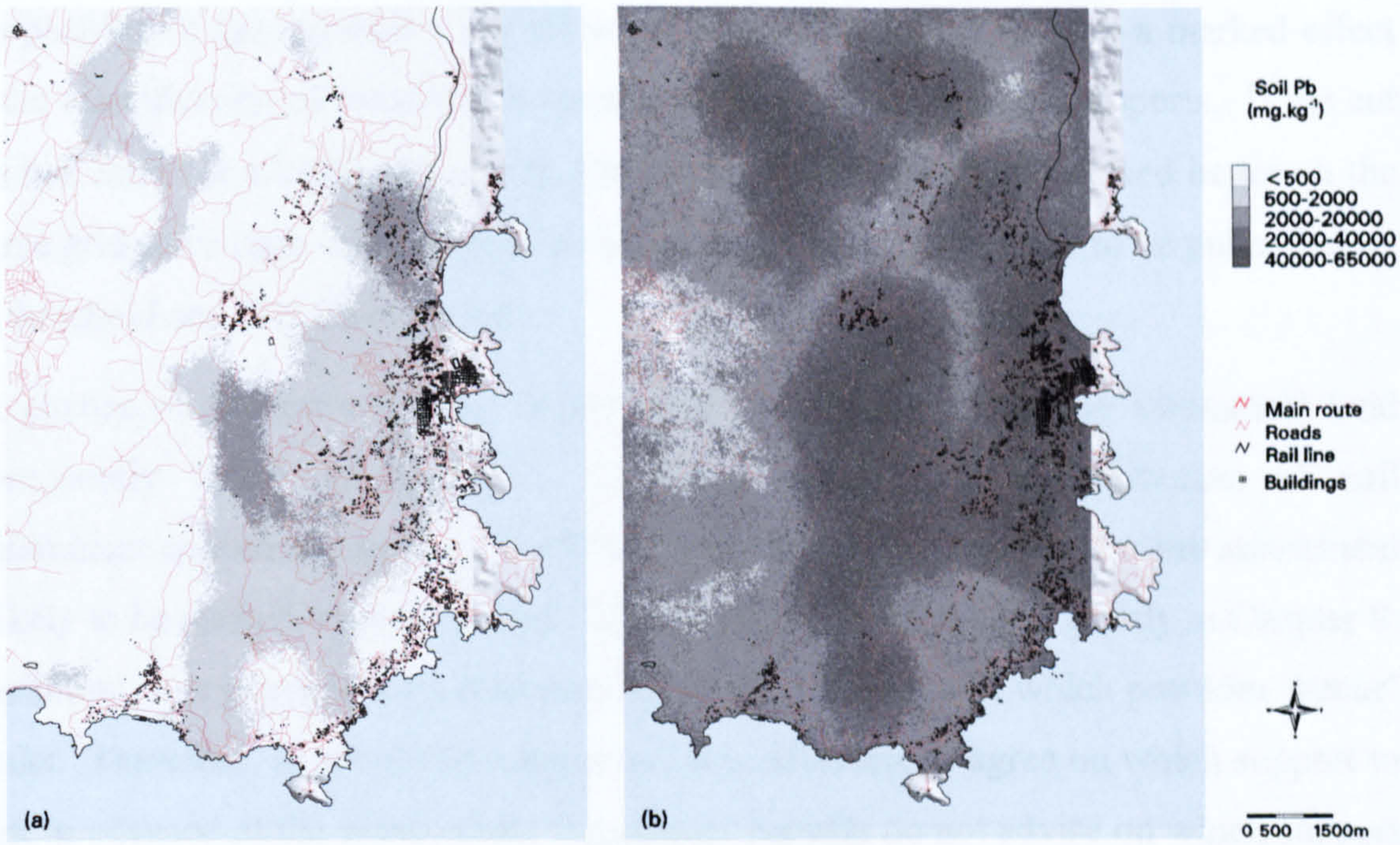


Figure 5.15. Soil Pb ($mg.kg^{-1}$) at (a) 2.5^{th} percentile and (b) 97.5^{th} percentile, being the lower and upper bounds of the 95% confidence interval respectively. Results of performing SIS on a fine grid.

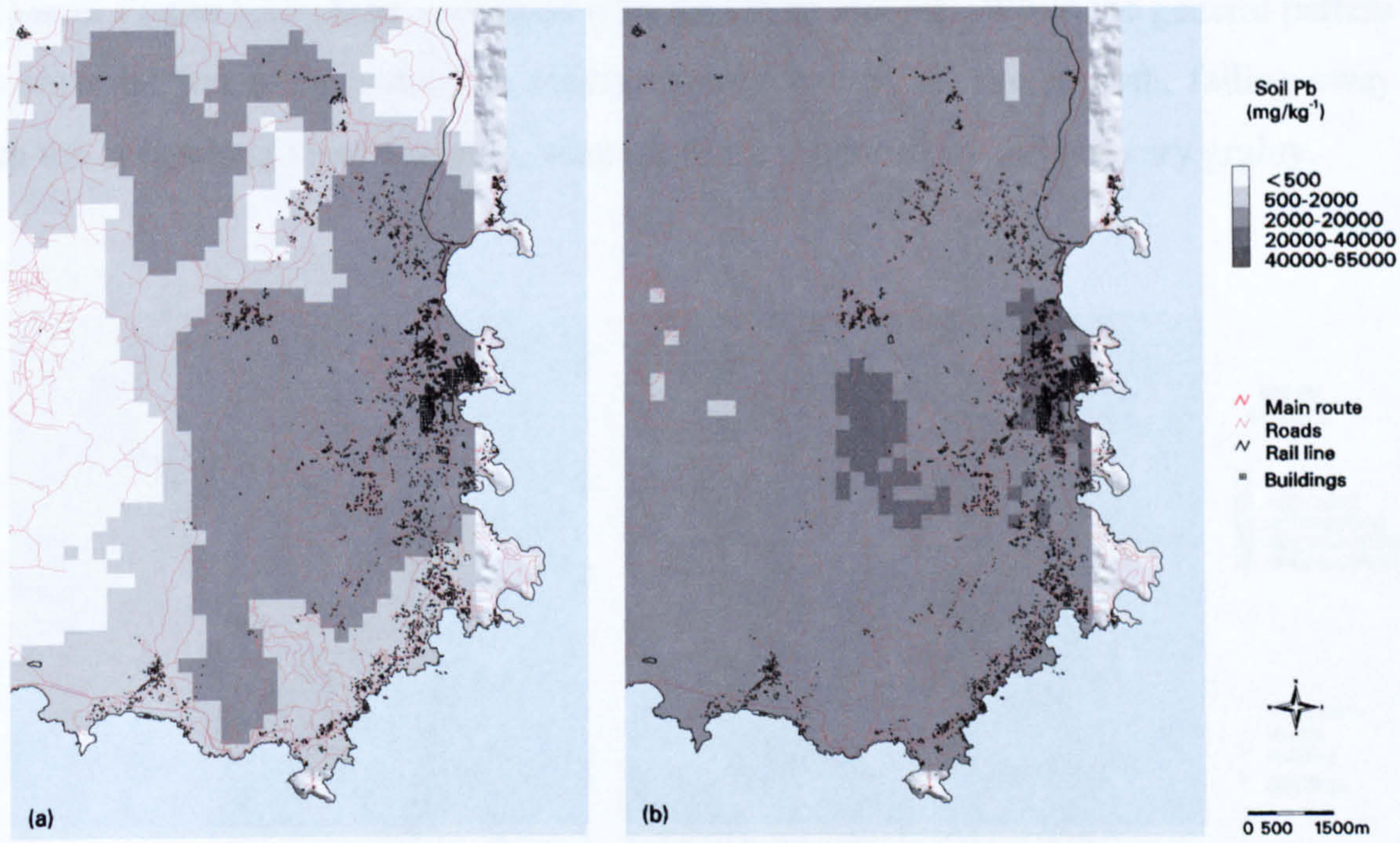


Figure 5.16. Soil Pb ($mg.kg^{-1}$) at (a) 2.5^{th} percentile and (b) 97.5^{th} percentile, being the lower and upper bounds of the 95% confidence interval respectively. Results of upscaling SIS to a coarse grid.

The position of the regulatory cut off within the distribution can have a marked effect on the classification of results with respect to working on different supports. If the cut offs had fallen at a higher percentile, the contamination may have looked better on the coarse grid since high values would be averaged downwards enough to be pulled below the threshold and classified as clean.

The change of support effect has important implications for exposure assessment (and subsequently intake assessment). Checking where in the distribution of soil contaminant concentrations the cut off lies, will help to indicate whether the assessment is likely to be conservative or lenient. This point is expanded upon greatly in Chapter 8. In addition, it is possible for a risk assessor to choose a support which provides 'better' results. Therefore, to avoid this temptation, it is advisable to agree on which support to work in advance of the assessment. Regulatory cut offs do not advise on which support they are to be used.

5.5.3 Uncertainty due to method choice: IK versus SIS

A difference due to using IK as opposed to SIS is clear from comparing Figure 5.17 (IK) with Figure 5.15 (SIS), produced with the OS model set. While the general pattern of contamination is the same, the estimates produced by IK are smooth, falling away from the hotspots in clear contours, whereas those produced by SIS are very grainy.

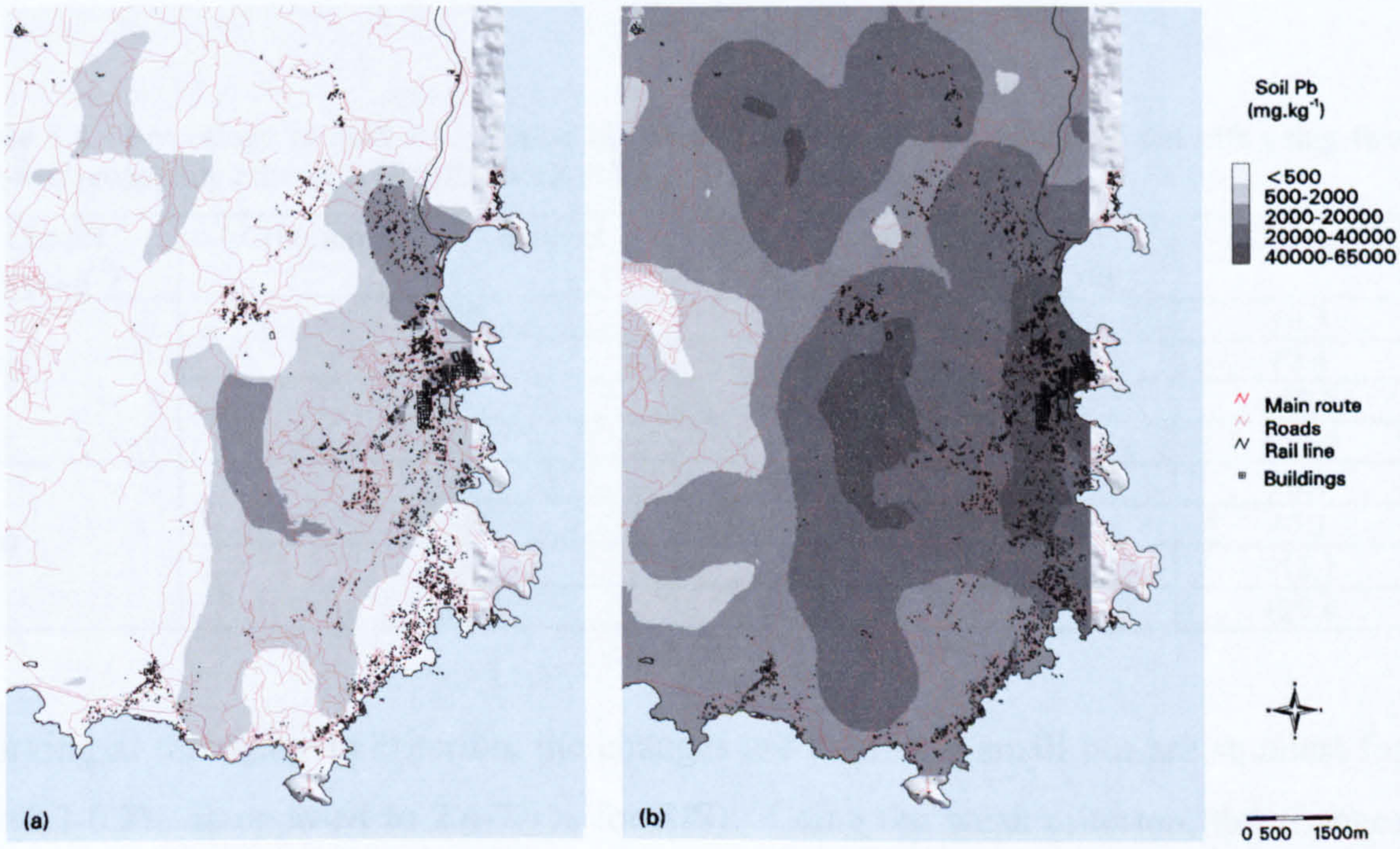


Figure 5.17. Soil Pb ($mg.kg^{-1}$) at (a) 2.5^{th} percentile and (b) 97.5^{th} percentile, being the lower and upper bounds of the 95% confidence interval respectively. Results of performing IK on a fine grid.

This is caused by the way the kriging and simulation procedures work. IK uses the same, sparse sample data for estimating the concentrations in all the unsampled locations. As the location being estimated moves away from one of these samples, the influence of the sample gradually wanes as a new, closer sample begins to exert more influence, and this leads to very smooth maps. SIS, on the other hand, utilises previously simulated values as well as the original sample data, and these too have an influence over the estimated values. Each simulation is made using a different random pathway through the unsampled locations, so each simulated map is different and the results are not smooth.

For an assessment on this scale, whether the map is smooth or grainy is probably not that important: one can get an overall impression of where the contamination exceeds the regulatory threshold. It may be more important, however, when working on a smaller scale and attempting to pinpoint small blocks in need of remediation. Kriging with small blocks leads to over-smooth results (Armstrong & Champigny, 1989).

Another difference between IK and SIS is the effect of change of support. For each regulatory cut off and each probability criterion the amount of land classified as contaminated is different depending upon the support size used (Table 5.4). In this particular case study, working on a fine grid (small support) leads to a smaller percentage of the land being classified as contaminated than working on a coarse grid (large support), though this is not always the case (Section 5.5.2).

Table 5.4. Percentage of land area classed as contaminated at the two regulatory cut offs using two different probability criteria: S ($P \geq 0.2$); W ($P \geq 0.8$).

Cut off ($mg.kg^{-1}$)	Criterion	Method	Support size		Change
			25x25m	250x250m	
500	S	IK	91.4	91.7	+0.3
		SIS	93.0	95.4	+2.4
	W	IK	59.0	77.7	+18.7
		SIS	57.7	86.7	+29.0
2000	S	IK	68.5	69.2	+0.7
		SIS	70.8	77.9	+7.1
	W	IK	23.2	36.3	+13.1
		SIS	21.3	50.7	+29.4

Working at the stringent criterion, the changes are relatively small but are smallest for IK (0.3-0.7% as opposed to 2.4-7.1% for SIS). Using the weak criterion, the changes are more noticeable and are, again, smaller for IK (13.1-18.7% as opposed to 29-29.4% for SIS), suggesting that the support effect is less for IK than SIS. This is a function of

the method of applying the change of support and the difference between smooth IK and grainy SIS results. For IK on a large support, a point kriging was made on a 250x250m grid and a variance reduction factor was applied to the results to adjust them to be representative of the average of the 250x250m block. With SIS, the very variable adjacent point 25x25m results within each block were averaged to give the 250x250m results. It is difficult to compare them directly and it has not been possible to validate the results in the field so it is not possible to say which results are nearer to the truth. The important point to note, however, is that in this case, change of support effects produced in SIS are larger than IK.

A practical consideration of using IK versus SIS is what the results will be used for. If the intention is to carry out a generic exposure assessment only and generally classify the land as contaminated or clean, then IK should suffice. As noted already, Figure 5.17 (IK) and Figure 5.15 (SIS) look broadly similar, but the time taken to produce Figure 5.17 was much less.

If, on the other hand, there is a high likelihood that an intake assessment will be undertaken, then it is better to simulate the results. For the spatially-evaluated intake assessments proposed in this thesis, multiple contaminant maps are required for use in conjunction with spatial population data in order to provide a measure of uncertainty in the intake assessment results. The simulated results may be used map by map, to preserve the spatial correlation of the values, while multiple maps provide a measure of uncertainty. Methods have been proposed recently to sample randomly from the kriged ccdfs produced at each location (Van Miervenne & Goovaerts, 2001), but this author is concerned that the spatial correlation of the soil contaminant concentrations would be lost, and this correlation is important when assessing the risk to a population through a spatially-evaluated intake assessment, as will become clear in future chapters.

5.6 Summary

This chapter has demonstrated the procedure for a spatially-evaluated exposure assessment using a kriging and a simulation method. Uncertainties are introduced into the assessment due to the many subjective choices which must be made. Overall, it is concluded that a simple variogram model which reasonably fits the data should suffice, since fitting a more complicated structure slows the calculation process and has only a minor effect on the results of the exposure assessment. For a generic exposure assessment which should broadly classify land into contaminated and clean classes, the

faster IK should suffice, but if results are to be fed into an intake assessment simulated results are more useful. Choice of support has a large influence on results, especially for SIS, where results are greatly smoothed by the averaging of the highly variable point results.

Having explored spatially-evaluated exposure assessment, and demonstrated the influence that choices can have on the results, the next chapter describes the newly developed spatially-evaluated intake assessment methodology, which utilises results from exposure assessments such as the one just described.

Chapter 6 Building the spatial and probabilistic intake assessment methodology

6.1 Introduction

This chapter presents the new methodology for probabilistic spatially-evaluated intake assessment which has been developed during this research. First of all, an overview of the possible methods for taking into account uncertainty, variability and spatial information is given, and the place of this methodology is highlighted. Next, the background to the CLEA model is given, before an explanation of how the use of the model has been adapted to incorporate spatial evaluation techniques. An account is given of how the populations are allocated in space in the various case studies (to be evaluated subsequently), along with an explanation of how the population and soil values from the geostatistical analysis are related spatially during the calculation procedure. The results output is then described followed by a discussion about how to analyse and evaluate it, so as to preserve the depth of information which is produced with the new model.

6.2 Accounting for variability, uncertainty and the spatial dimension

Using various combinations of data and calculations it is possible to account for one, some or all uncertainty, variability and the spatial dimension. Table 6.1 and Table 6.2 show the differences in methods for dealing with the soil concentrations and receptor populations respectively. These methods can be combined in an intake assessment methodology to provide varying degrees of complexity.

Table 6.1. Methods to account for variability, uncertainty and spatial dimension in soil concentrations.

Method	Variability	Uncertainty	Spatial dimension
Use single site value	-	-	-
Use a distribution fitted to the sample values	Y	-	-
Produce multiple distributions by re-sampling sample values	Y	Y	-
Krige to produce the best estimate per exposure unit (EU) *	Y	-	Y
Multiple site averages produced by simulations	-	Y	-
Multiple values per EU produced by simulations	Y	Y	Y

* Exposure unit (EU) defined in Section 6.5

Table 6.2. Methods to account for variability, uncertainty and spatial dimension in receptor population.

Method	Variability	Uncertainty	Spatial dimension
1 average/critical person	-	-	-
Population of variable people	Y	-	-
1 average/critical person in each exposure unit (EU) *		-	Y
Multiple estimates for average/critical person in EU		Y	Y
1 population located in space	Y	-	Y
Multiple populations located in space	Y	Y	Y

* Exposure unit (EU) defined in Section 6.5

Looking at soil analyses, the first method is really a typical deterministic assessment, taking one value to represent soil values in the area of interest. It does not have to be the mean, but could be the upper 95th percentile value for a more conservative estimate, for example. If several soil samples are available, a distribution can be fitted to them, and sampling from this distribution would give a measure of the variability of soil values in the area. Repeating this sample, say, 1,000 times would give, in addition, a measure of the uncertainty. Still, though, there is no spatial dimension. The spatial dimension can be introduced by kriging to produce a best estimate of the soil concentration at each point on a site. With kriging, variability in space is achieved, but a measure of uncertainty is missing. Using geostatistical simulation, one can introduce an element of uncertainty by producing, say, 1,000 feasible maps. If the average soil concentration were to be used from these, only a measure of uncertainty would be achieved. The most comprehensive method which incorporates variability of the soil values in space and includes a measure of uncertainty is the final one, which uses geostatistical simulations to produce a feasible map of soil values, giving variability in space, and repeats the procedure over again to provide a measure of uncertainty which preserves the spatial dimension. This comprehensive method is used in this research.

For the receptor population, again, using an average or critical person is typical of a deterministic assessment, with no accounting for variability, uncertainty or the spatial dimension. If one was to use a population of people of different ages, weights, sex etc. then variability in the receptors would be accounted for. To account for the spatial dimension, one could place an average or critical receptor in each grid cell on the map of interest, but this would then lose the population variability. The best way to account for population variability and incorporate the spatial dimension properly is to use a real population of people, located in the correct position on the map of interest. This is how the population is treated in this research.

Uncertainty in the population could be taken into account by generating 1,000 populations from the same information. However, this method is purposely not used in this thesis, so as not to over-complicate the analysis and hide effects of interest, such as the effect of support on the intake assessment. One 'real' population is kept for the duration of each intake assessment.

Various parts and/or combinations of the above have been used previously. Examples include: deterministic use of population and soil data (HESP, 1995); uncertainty in receptor data with inferred deterministic soil data (DEFRA & EA, 2002g); variability and uncertainty in the soil data (Burmaster & Thompson, 1997); and spatial variability in soil data by kriging with uncertainty in the receptor (Korre et al., 2002).

Overall, the newly developed method is an improvement on previous methods in as far as it incorporates variability, uncertainty and the spatial dimension in the soil contaminant concentrations and variability and the spatial dimension for the receptor population. The author has not seen another method which tackles the intake assessment in this way.

6.3 Introduction and background to the CLEA model

In 1987 the UK Interdepartmental Committee on the Redevelopment of Contaminated Land (ICRCL) produced a set of 'tentative' threshold and action values for some land contaminants (ICRCL, 1987). However, in 1990, the House of Commons Select Committee on the Environment called for more scientifically based guidance, and the upshot of this was, that after a decade of work, in March 2002 DEFRA and the EA published a series of long-awaited reports and a computer model concerned with the assessment of risks to human health arising from exposure to contaminants in soil. The

series comprised the reports known, for short, as CLR7-10, TOX1-10 and SGV1-10. The whole series was published with the purpose of providing "technical material that can be used in support of the application of the statutory regimes addressing land contamination, particularly PartIIA of the Environmental Protection Act 1990 (the contaminated land regime) and development under the Town and Country Planning Acts" (DEFRA & EA, 2002a). The reports take into account previous guidance on assessing and managing environmental risks (DETR et al., 2000).

The series of reports focuses around the derivation of Soil Guideline Values (SGVs) using the Contaminated Land Exposure Assessment (CLEA) model. SGVs are maximum levels of contaminants in soil ($mg.kg^{-1}$) which are considered 'safe' in particular circumstances. These guideline (or 'intervention') values are calculated in a probabilistic manner such that they would be protective of 95% of the exposed population, thus protecting the vast majority of people while accepting that there may be some extreme high intakes which will not be prevented. First, a Health Criteria Value (HCV) is derived. This is, for example, an intake dose of contaminant which is deemed acceptable or 'safe' (usually measured in $mg.kg^{-1}_{BodyWeight}.d^{-1}$). Next, the level of contaminant in the soil that would lead to an intake dose no greater than the HCV (for 95% of the exposed receptors) is calculated, and this contaminant level is referred to as the SGV. SGVs are to be used as 'intervention values', whereby if contaminants are found in the soil at levels above the SGV, some form of remediation should take place to reduce the soil contaminant level to ensure safety of the site. Having said that, SGVs are not binding – they are advisory guidelines only, and should not be used in isolation. According to DEFRA & EA (2002a), their use is "subject to the provisions of the statutory guidance and other legal requirements".

Contaminated Land Reports (CLR) 7-10 (DEFRA & EA, 2002a; 2002b; 2002g; 2002h) provide the overall framework of, and guidance to, the risk assessment process, including the technical basis for, and use of, the CLEA computer model. The TOX series (e.g. DEFRA & EA, 2002d) provides specific details on the toxicology and uptake of elements (one report per element) in order to derive HCVs. SGV derivations are published in the SGV series (one report per element e.g. DEFRA & EA (2002j)). The original TOX and SGV series are being added to, as and when data for more elements become ready for publication.

6.3.1 Definitions

The terminology used in CLR10 (DEFRA & EA, 2002g) is inconsistent with that already defined in CLR9 (DEFRA & EA, 2002b) which has been adopted for use in this thesis (Chapter 4). CLR10 refers to calculating the 'Average Daily Exposure' in $mg.kg^{-1}_{BW}.d^{-1}$ whereas in CLR 9 the units $mg.kg^{-1}_{BW}.d^{-1}$ refer to 'intake'. For this reason, from here onwards, the 'Average Daily Exposure' will, more correctly, be referred to as the Average Daily Intake (ADI) as defined in CLR9. To confirm:

Exposure: the amount of chemical in a medium that is available for intake by the population. Expressed in $mg.kg^{-1}$.

Intake: the amount of chemical entering or contacting the human body at the point of entry. Expressed in $mg.kg^{-1}_{BW}.d^{-1}$.

Uptake: the amount of contaminant that reaches the circulating blood, having being absorbed by the body. Expressed in $mg.l^{-1}$

6.3.2 Derivation of SGVs

Development of generic SGVs is difficult. As well as sound scientific evidence for toxicological effects due to exposure to contaminants, expert judgement is required to define what 'acceptable risk' is. In addition, contaminant intake from soil is one of many possible routes, and the risk from soil should be kept in perspective with other sources of intake. SGVs must also cover a range of soil types and site conditions. To overcome this, worst-case assumptions can be used but these may lead to overly conservative risk estimates and unnecessary expenditure in some cases. Despite these issues, the use of SGVs can "serve a useful purpose in encouraging a transparent and consistent approach" to risk assessment (DEFRA & EA, 2002a) and can help to focus limited resources to where they are most needed. Details of derivation of the SGVs is given in CLR9 (DEFRA & EA, 2002b) and the most salient points are summarised here.

First, it is important to note that the SGVs refer to chronic intake over an extended period of time, which may be a lifetime. As mentioned above, the starting point for derivation of SGVs is the HCV (health criteria value). There are two types of HCV since there are two classes of contaminant: threshold contaminants; and non-threshold contaminants.

6.3.2.1 Threshold contaminants and the Tolerable Daily Soil Intake (TDSI)

Threshold contaminants are assumed to have no adverse effect on the receptor below a certain threshold level. Therefore, if the receptor is exposed to levels below this threshold, there is deemed to be no risk. This threshold is termed the Tolerable Daily Intake (TDI). The conventional approach for calculating the TDI is to find the 'No Adverse Effect Level' (NOAEL) from animal studies, then divide by 'safety factors' (or 'uncertainty factors') to account for the differences between the animals studied and the human population. Since the TDI refers to chronic intake, a brief period of excess may not have an adverse effect on the receptor.

In terms of deriving the SGV, however, the TDI alone is insufficient, since it applies to intakes form all sources, not just soil contamination. Hence the Tolerable Daily Soil Intake (TDSI) is required i.e. that intake which can be tolerated specifically from contaminated soil. This is calculated by subtracting the Mean Daily Intake (MDI, which is all other 'background' intakes) from the TDI. The difference, TDI-MDI, is put down to the amount of intake that can be tolerated from soil, the TDSI (DEFRA & EA, 2002b).

6.3.2.2 Non-threshold contaminants and the Index Dose (ID)

For non threshold contaminants, the HCV is called the Index Dose (ID). ID is defined as the "dose which can be considered to present a minimal human health risk from exposure to soil contaminants" (DEFRA & EA, 2002b). The ID for deriving the SGVs is set specifically for exposures from soil, and is subject to ALARP (as low as reasonably practicable) conditions. It is assumed that risks from other sources, e.g. drinking water, will also be controlled by ALARP principles. Exceeding the ID even in the very short term would lead to an unacceptable health risk.

6.3.2.3 Calculating the SGV

Once the HCV has been deduced (either as a TDSI or an ID) the SGV is calculated as that concentration in the soil which gives an intake value (Average Daily Intake) equal to the HCV. If ADI=HCV, then soil concentration = SGV. The ADI is calculated through a series of algorithms which form the CLEA model. According to the CLEA model, intake occurs via three routes: inhalation *(inh)*; dermal absorption *(derm)*; and oral intake *(oral)*, and can be summarised by the following equation:

$$ADI = \frac{(IR_{inh} \times EF_{inh} \times ED_{inh})}{BW \times AT} + \frac{(IR_{oral} \times EF_{oral} \times ED_{oral})}{BW \times AT} + \frac{(IR_{derm} \times EF_{derm} \times ED_{derm})}{BW \times AT}$$

where:

ADI: average daily intake of a chemical from soil ($mg.kg^{-1}_{BW}.day^{-1}$)

IR: chemical exposure rate ($mg.day^{-1}$)

EF: exposure frequency ($day.yr^{-1}$)

ED: exposure duration (yr)

BW: bodyweight (kg)

AT: averaging time (day)

In turn, the *IR* terms are made up of sub-equations which model how the intake occurs in each pathway. These sub-equations involve the soil contaminant concentration. Several of the variables in the equations, e.g. bodyweight, are treated probabilistically, recognising that individuals are different. Hence a distribution of ADI values is achieved and when the upper 95^{th} percentile equals the HCV, the soil concentration used in the equations equals the Soil Guideline Value. Hence applying this SGV to the contaminated site should be protective of 95% of the population.

6.3.2.4 Exposure scenarios and critical receptors

The CLEA model, describes four possible scenarios for exposure: residential scenario without home-grown vegetable consumption; residential scenario with home-grown vegetable consumption; allotments scenario and; a commercial/industrial scenario. These conceptual models have been developed in isolation and they should be viewed as such (DEFRA & EA, 2002g). The hours spent in each situation do not 'add up' e.g. the number of hours spent in a residential property do not allow for hours that may be spent in the allotment since the scenarios are worked out completely independently.

Because the model is aimed at back-calculating SGVs, input parameters are focussed around a 'critical receptor' in each scenario, who is thought to be most at risk for behavioural and physiological reasons. If this critical receptor is protected, then it follows that other less at-risk persons will be. For both residential and allotment scenarios, a 6 yr old female child is taken as the critical receptor, whereas in the industrial/commercial scenario a female worker is the critical receptor, thus inputs to the model are tailored accordingly when back-calculating the SGVs.

6.4 Adaptation of the CLEA model

This thesis takes the CLEA model and adapts it for use with a different emphasis. Rather than using it in a regulatory context for back-calculating an SGV, this thesis uses the intake algorithms to calculate actual intakes of an at-risk population due to the variable soil contaminant concentrations in the area in which they reside. Not only is the emphasis changed, but also the model is adapted to take into account the uncertainty in the soil concentrations to which the population is exposed, and is evaluated in a spatial manner, allowing for the interaction of the population and contaminants in space.

Here follows a detailed description of how a new intake model was developed. First, a description of the exposure scenario and categorisation of the population into age classes is given. This is followed by a detailed account of how the basic CLEA equation has been adapted and how the variable inputs were derived. Because the model is evaluated spatially, the population has to be allocated to a point in space and an account of the allocation process is given. Finally, there is a description of how the algorithm is evaluated with the multiple soil values.

6.4.1 Exposure scenarios

Two residential scenarios are studied in detail: one assumes home-grown vegetables are consumed and the other assumes not. Differences in intake depending upon whether home-grown vegetables are consumed or not can then be investigated. Both scenarios assume that the population is exposed to contaminant concentrations in the place where they reside for 24 hours per day over a period of one year. A 'snapshot' of one year allows for a relatively straightforward calculation of intakes: each age-dependent variable can be allocated to a receptor assuming they remain the same age for the duration of the intake period and this means that potentially interesting results will not be clouded by the effect of a constantly changing population.

6.4.2 Receptor population

In this thesis, the focus is on calculating intakes for the entire population. Many of the variables in the intake equation are age dependent, so the age structure of the exposed population is important. Since rapid growth, development and behavioural changes take place during the early years of life, it is necessary to treat these ages differently; age specific values are available for many variables for ages 0 to 15 yrs. All ages from 16 upwards are assumed to have the same characteristics and are treated identically. In the

scenarios studied in this thesis, there is nothing to distinguish between males and females so they are treated as one population.

6.4.3 Basic intake equation

The equation calculates the Average Daily Intake (*ADI*) in $mg.kg^{-1}_{BW}.day^{-1}$ which is the sum of intakes occurring via three routes: inhalation *(inh)*; dermal absorption *(derm)*; and oral intake *(oral)*:

$$ADI = \frac{(IR_{inh} \times EF_{inh} \times ED_{inh})}{BW \times AT} + \frac{(IR_{oral} \times EF_{oral} \times ED_{oral})}{BW \times AT} + \frac{(IR_{derm} \times EF_{derm} \times ED_{derm})}{BW \times AT}$$

where

ADI: average daily intake of a chemical from soil ($mg.kg^{-1}_{BW}.day^{-1}$)

IR: chemical exposure rate ($mg.day^{-1}$)

EF: exposure frequency ($day.yr^{-1}$)

ED: exposure duration (yr)

BW: bodyweight (kg)

AT: averaging time (day)

The *IR* terms are made up of sub-equations which model how the intake occurs in each pathway. For this thesis, the contaminants of concern are metallic/metalloid inorganic contaminants (Cd, As and Ni) which are not volatile. Consequently, contribution to intake from the dermal and inhalation pathways contributes less than 1% to the overall intake (DEFRA & EA, 2002i; 2002j; 2002l) and it is necessary only to include intake from the oral pathway. Thus the intake equation is reduced to:

$$ADI = \frac{(IR_{oral} \times EF_{oral} \times ED_{oral})}{BW \times AT}$$

where

$$IR_{oral} = IR_{dsi} + IR_{isi} + IR_{veg}$$

where

IR_{dsi}: chemical intake rate from direct (inadvertent) soil ingestion ($mg.day^{-1}$)

IR_{isi}: chemical intake rate from indirect ingestion of soil attached to vegetables ($mg.day^{-1}$)

IR_{veg}: chemical intake rate from ingestion of contaminants absorbed into vegetables ($mg.day^{-1}$)

Finally, each *IR* sub-equation is made up as follows:

$$IR_{dsi} = C_{soil} \times SDR$$

$$IR_{isi} = C_{soil} \times \sum_{veg\ type} \left(CR_{veg} \times BW \times HF_{veg} \times SL_{veg} \right)$$

$$IR_{veg} = C_{soil} \times \sum_{veg\ type} \left(CR_{veg} \times BW \times HF_{veg} \times CF_{veg} \right)$$

where

C_{soil}: contaminant concentration in soil ($mg.g^{-1}$)

SDR: average daily soil and dust ingestion rate ($g.day^{-1}$)

CR_{veg}: consumption rate per vegetable ($gFW.day^{-1}$) (FW=Fresh weight)

BW: bodyweight (kg)

HF_{veg}: fraction that is home grown per vegetable (-)

SL_{veg}: amount of attached soil per vegetable ($g.g^{-1}FW$)

CF_{veg}: soil to plant concentration factor ($\mu g.g^{-1}FWplant/\mu g.g^{-1}DWsoil$) (DW=Dry weight)

6.4.4 Derivation of input variables

In its original form, of the inputs to the CLEA model, *SDR* (some age categories), CR_{veg}, *BW*, and HF_{veg} are treated probabilistically, while C_{soil}, SL_{veg} and CF_{veg} are treated deterministically. In this new methodology, C_{soil} is treated spatially and probabilistically, as is CF_{veg} when dealing with Cadmium intake, since it is pH dependent for Cadmium.

The CLR10 report (DEFRA & EA, 2002g) describes in detail inputs used for the original CLEA model. All references to prescribed model values refer to those given in that report. Each variable is now taken in turn and any changes to the prescribed values are noted and justified.

6.4.4.1 SDR

SDR, average daily soil and dust ingestion rate ($g.day^{-1}$), is taken as given in CLR10. Generally, the amount of soil inadvertently ingested decreases with age due to differences in behaviour. Soil ingestion has been studied in more detail in young children and CLR10 suggests using a distribution of possible values for children up to the age of five with point estimates only for children aged 6-15 and persons aged 16+ i.e. *SDR* is treated probabilistically in young children and deterministically in older age groups. The distribution of values used in CLR10 are described through the mean, and

50th and 95th percentile values of the assumed natural lognormal probability density function (PDF) (Table 6.3).

Table 6.3. *SDR* values used in calculation of *ADI*

Age (yrs)	SDR ($g.day^{-1}$)
0-5	0.1* (0.064,0.303)
6-15	0.1
16+	0.06

*Natural lognormal PDF. Mean (50th, 95th percentile) shown.

6.4.4.2 CR_{veg}

Consumption rate of vegetables varies by age and vegetable type and is taken as given in CLR10. Natural lognormal distributions are used to describe the consumption rates: the mean, and 50th and 97.5th percentiles are given in Table 6.4.

Table 6.4. CR_{veg} values used in calculation of *ADI*

Age (yrs)	CR_{veg} ($gFW.kg^{-1}_{BW}.day^{-1}$)*					
	Brussels sprouts	Cabbage	Carrot	Leafy Salad	Onions (shallots and leeks)	Potato
0	0.61 (0.46, 2.01)	0.51 (0.42, 1.37)	1.30 (1.14, 3.05)	0.29 (0.24, 0.82)	0.26 (0.18, 0.95)	3.08 (2.54, 8.63)
1-3	0.44 (0.30, 1.67)	0.46 (0.29, 1.84)	0.62 (0.45, 2.15)	0.32 (0.20, 1.32)	0.32 (0.23, 1.13)	3.63 (3.20, 8.61)
4-15	0.28 (0.22, 0.81)	0.25 (0.20, 0.75)	0.36 (0.26, 1.25)	0.17 (0.11, 0.66)	0.22 (0.16, 0.77)	2.95 (2.66, 6.48)
16+	0.30 (0.25, 0.82)	0.27 (0.21, 0.80)	0.27 (0.22, 0.77)	0.16 (0.11, 0.61)	0.19 (0.15, 0.52)	1.82 (1.66, 3.81)

* Natural lognormal PDF. Mean (50th, 97.5th percentile) shown.

6.4.4.3 BW

BW distributions given in CLR10 were not used in this study. In CLR10, the *BW* distributions are given separately for males and females and for working age adults and retired adults. This is related to the fact that CLEA uses two 'critical receptors', both of which are female, therefore it is necessary to consider the female parameters only. Also, the female critical receptor for their industrial setting is of working age, hence information for retired females must be excluded. However, this study is concerned with calculating *ADI* for receptors of all ages, and from the oral intake pathway only, and since no other variables differ specifically for males and females, it is not necessary

to treat the sexes as two separate populations. Therefore the bodyweight distributions were re-modelled by age group only. Likewise, there is no reason to model retired persons separately from working age persons, since in this model there is no difference in behaviour. All persons aged 16 and over are therefore modelled as one age group.

The same *BW* data set (Department of Health, 1998) as was used in CLR10 was re-worked. The data were split into the age groups and natural lognormal PDFs were fitted to the data. The lowest two age groups (0 and 1yrs) were not included in this data set, so the parameters for these two classes were estimated by linear regression, as illustrated in Figure 6.1. The parameters of the final *BW* distributions are given in Table 6.5.

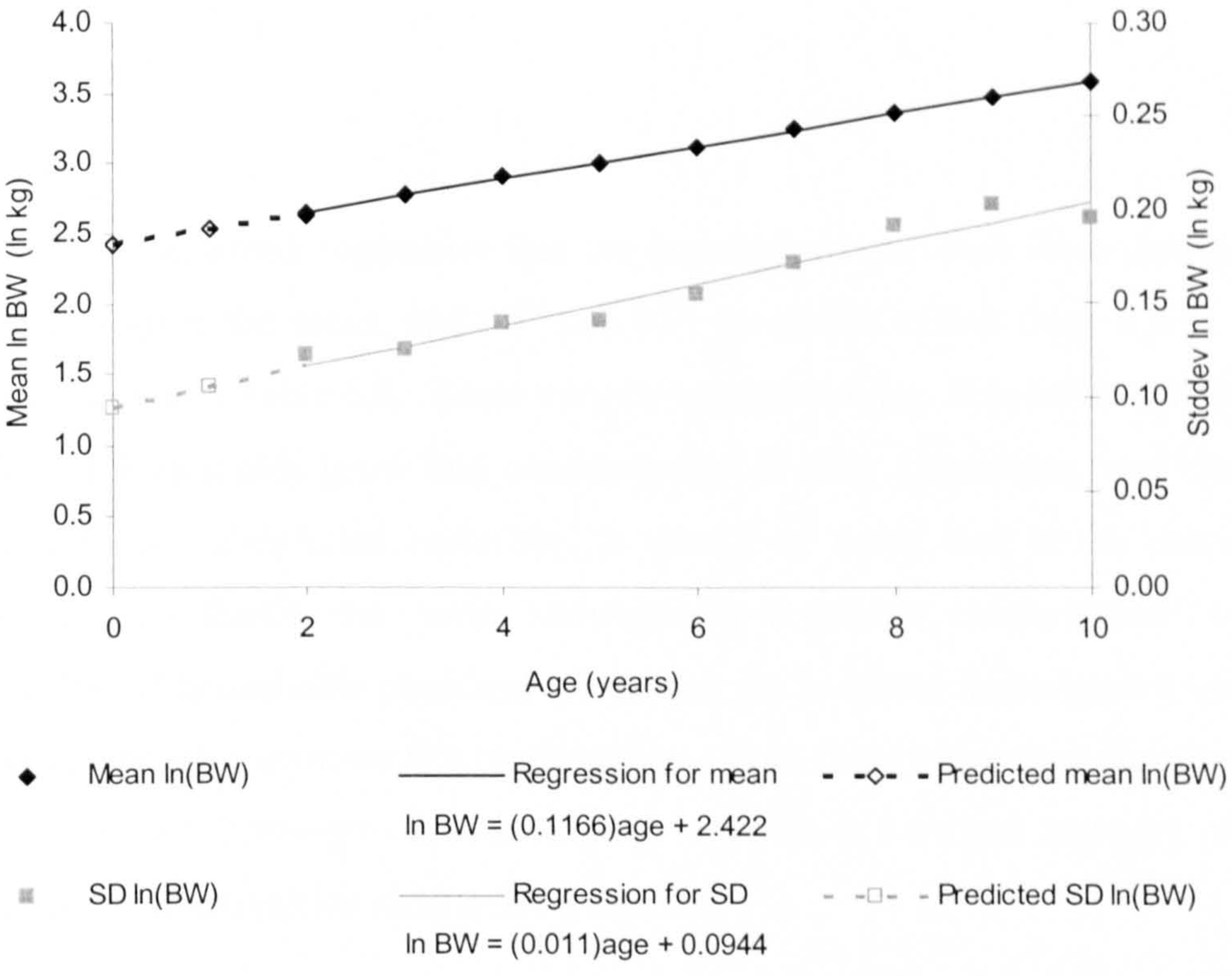


Figure 6.1. Regression of mean and standard deviation of ln *BW* by age. R^2=0.9984 and 0.9504 for the mean and standard deviation respectively.

Table 6.5. Parameters for *BW* distributions used in calculation of *ADI*

Age (yrs)	BW (kg)*	
	Mean	Stddev
0	11.32	1.07
1	12.74	1.35
2	14.01	1.72
3	16.19	2.05
4	18.59	2.60
5	20.41	2.88
6	22.80	3.57
7	26.10	4.51
8	29.09	5.61
9	32.70	6.70
10	36.69	7.24
11	42.19	10.07
12	46.53	10.04
13	52.32	10.74
14	56.45	10.76
15	61.19	12.06
16+	73.86	15.10

* Natural lognormal PDF.

6.4.4.4 HF_{veg}

The fraction of consumed vegetables that are home-grown is taken from data given in CLR10, which gives the mean, and 50th and 95th percentile values from a beta shaped distribution, shown in Table 6.6. These vary by vegetable type. It is estimated that only about 15% of households grow and consume any of their vegetables, and the HF_{veg} values refer to this significant minority. It should be noted that in the calculations performed in this thesis, the 'with home-grown vegetable consumption' scenario assumes 100% of households participate, whereas the 'without home-grown vegetable consumption' scenario assumes 0% participation. This clearly distinguishes the effects of including, or not, home-grown consumption and allows for clear mapping of where the home-grown consumption most affects the *ADI*.

Table 6.6. Parameters for HF_{veg} distributions used in calculation of *ADI*

	Brussels sprouts	Cabbage	Carrot	Leafy Salad	Onions (shallots and leeks)	Potato
HF_{veg}*	0.39, 0.27	0.38, 0.30	0.24, 0.22	0.12, 0.17	0.35, 0.30	0.22, 0.21

* Beta shaped PDF. Mean and standard deviation shown

6.4.4.5 SL_{veg}

Soil loading is taken as given in CLR10. This factor is treated deterministically and varies by vegetable type only. The values are given in Table 6.7.

Table 6.7. SL_{veg} values used in calculation of *ADI*

	Brussels sprouts	Cabbage	Carrot	Leafy Salad	Onions (shallots and leeks)	Potato
SL_{veg} ($g.g^{-1}FW$)	0.001	0.001	0.0001	0.001	0.001	0.0002

6.4.4.6 CF_{veg}

CF_{veg} is treated deterministically and varies by contaminant and also by vegetable class; root vegetables and leafy vegetables are classed separately. Values for Arsenic (As) and Nickel (Ni) are taken as given in CLR10 and are shown in Table 6.8. However, CF_{veg} for Cadmium (Cd) varies according to pH and only the value at pH 7 is given in Table 6.8.

Table 6.8. . Prescribed CF_{veg} values

Contaminant	CF_{leafy} ($\mu g.g^{-1}$DWplant/$\mu g.g^{-1}$DWsoil)	CF_{root} ($\mu g.g^{-1}$DWplant/$\mu g.g^{-1}$DWsoil)
Arsenic	0.009	0.009
Nickel	0.047	0.018
Cadmium	0.793 (at pH7)	0.706 (at pH7)

Instead of assuming an average pH value for a site under investigation, the pH values are allowed to vary across the site according to estimates available from geostatistical simulations. This is a major adaptation of the CLEA model for use in the spatially-evaluated quantitative risk assessments performed in this thesis. These variable pH estimates are used in conjunction with regression equations which link pH to CF_{leafy} and CF_{root} in order to calculate spatially variable CF_{veg} estimates when calculating *ADI* for Cd. This should allow for more realistic estimates of CF_{veg} for Cd. The regression equations are as follows (DEFRA & EA, 2002j):

$$\ln(CF_{root}) = 11.174 - (1.6461 \times \text{pH})$$

$$\ln(CF_{leafy}) = 11.206 - (1.634 \times \text{pH})$$

The regression equations were used separately for analysis of a small data set, but since the regressions are so similar, it was decided to take an average value and use one

equation for all vegetables when dealing with a large data set. This was necessary to reduce, dramatically, the computational burden, but will still be more realistic than assuming a constant pH, and hence constant CF_{veg}, across the site. The average regression equation used is:

$$\ln(CF_{veg}) = 11.19 - (1.64 \times \text{pH})$$

Note that the CF_{veg} values provided in CLR10 and shown in Table 6.8 are in terms of $\mu g.g^{-1}DW$ plant, and the model requires them in $\mu g.g^{-1}FW$ plant. Therefore, a conversion factor is applied to the calculated value to convert it to the correct units. These vegetable specific values are also given in CLR10 and are shown in Table 6.9.

Table 6.9. Dry weight to fresh weight conversion factors for different vegetable types.

	Brussels sprouts	Cabbage	Carrot	Leafy Salad	Onions (shallots and leeks)	Potato
Conversion factor ($gDWplant.g^{-1}FWplant$)	0.095	0.126	0.097	0.04	0.156	0.21

6.4.4.7 C_{soil}

The purpose of the CLEA model is to back-calculate SGVs, i.e. soil concentrations that are deemed safe, so there is no C_{soil} input to the model when it is being used for its designed purpose. C_{soil} 'input' comes at the stage when one is comparing site soil concentrations to the calculated SGV to see whether the soil is above or below this supposed safe level. This is done using a constant average value for the site.

Treatment of C_{soil} marks the major novelty of the method used in this thesis. It is treated in a probabilistic and spatially-evaluated manner, using the outputs of geostatistical simulations as inputs to the risk model. 1,000 maps of soil concentration (simulated by sequential indicator simulation) are used as inputs to the intake model. Each person is allocated to a point in space and linked, therefore, to the 1,000 soil concentrations in that place and the risk equation is solved 1,000 times for each person according to their local soil concentrations.

6.4.4.8 EF, ED and AT

The scenario studied is a 'snapshot' of one year. Therefore, for all ages, except the first year of life, the exposure frequency (*EF*) is 365 days per year, the exposure duration (*ED*) is one year and the averaging time (*AT*) is 365 days. The exception in babies is

that they are assumed only to be exposed to soil contaminants through the oral pathway from the age of 6 months (CLR10). Therefore, their *EF* is taken to be 180 days per year.

6.5 The concept of exposure units

An exposure unit (EU) is defined as "the geographical area in which a receptor moves and contacts contaminated medium during the period of exposure duration" (USEPA, 2001). In practice, for example, an EU for a small child in a residential scenario may be the garden of its house, or an EU could be the entire area of an old industrial complex now designated to be a playground.

This spatially-evaluated quantitative risk assessment uses outputs from geostatistical algorithms as estimates of the local soil contaminant concentration. The geostatistical outputs are produced on a regular grid across the site of interest, so in this case, the EU is designated to coincide with a grid cell of a geostatistical output grid. The EU is not, therefore, as clearly defined as might be, e.g. it is not strictly a residential garden. However, the 'geostatistical EU' is a sufficient unit for the scenarios examined in this thesis.

Having said that the EU sizes were not designed with the population in mind, but rather with the geostatistical process in mind, the soil contaminant data sets were not collected specifically for risk assessment purposes in any case, but were used because they were available to test out this new methodology. If this methodology were to be used in a risk assessment process in the future, soil sample collection could be tailored specifically to allow for EU sizes felt to be appropriate in that instance. For example, it may be tailored to enable classification of EUs appropriately sized to be suitable remediation units.

In this thesis, generally, the size of the large EU was set to about half the sample spacing e.g. if the samples were on a grid with 1km spacing, the large EU would be 500x500m. For two small data sets used later in the thesis (Gai and Vasilevskaya, Russia), the size of the simulation grid was decided simply by dividing the large EU side by 10, so a small EU would be, for example, 50x50m and there would be 100 small EUs in a large EU. For two larger examples (As and Cd from Stoke, UK), the computational burden was just too big to copy this regime. In these cases, the large EUs were divided into 16 smaller EUs (4x4 small EUs per large EU).

As explained in Section 6.4.1, chronic intake is assessed over a period of one year and receptors are assumed to live, work and play in the same place (EU) for that whole year and are exposed to the contaminants in that place only. Simulations produce multiple realisations of the soil concentration for each EU and lead to an estimate of uncertainty on the soil contaminant concentration in that EU.

6.6 Allocating the population to an EU

In order to evaluate the model spatially, each individual in the population needs to be allocated to an EU. Depending upon the amount of information available, this can be carried out with differing degrees of sophistication. This thesis uses populations in three case studies, two of which have simple information, and one which is more comprehensive.

6.6.1 Simple method

For two case studies (Vasilevskaya and Gai, Russia) each study area consists of a region with four settlements. The population numbers by age class are known for each settlement, and the positions of the settlements within the region are known. The age classes are 0-5yrs (pre-school children), 6-17yrs (school-age children) and adult (18+yrs). Overall population numbers are quite small: 6,554 Vasilevskaya; and 34,997 Gai. No other information is available. The allocation procedure was undertaken as follows:

The fine geostatistical grid was laid across the region, and any cell of the grid which contained a building was marked. The population in each settlement was then distributed as uniformly as possible across the marked cells, by age class, ensuring no fractions of people existed. For the pre-school age class, the age of the person was allocated randomly from a uniform distribution between 0 and 5, assuming a uniform birth cohort. This was repeated for the 6-17 age category. Age of adults does not matter, since all ages 16+ are treated identically in the intake model.

Since the risk calculations were to be carried out on different spatial resolutions (in order to investigate the support effect) the population needed to be aggregated up to the larger grid EUs. This was done by simply adding up the populations of each small cell falling within the bounds of the larger cell. This procedure was undertaken by a combination of ArcInfo™ GIS and Microsoft Excel™.

This population distribution was fixed for the duration of the calculations.

6.6.2 Advanced method

For one case study (Stoke on Trent, UK) the information available was much more comprehensive and the data set very much larger than for the Russian case studies (total population 370,440). Detailed UK census data and digital Ordnance Survey (OS) maps are available. These were manipulated in ArcInfo™ GIS to allocate population to EUs. The details of the data sources and allocation procedure are now described.

6.6.2.1 Census data

At the time of analysis, the most recent census data, available in a suitable format, were from 1991. These were downloaded from the CASWEB website ('web interface to census aggregate outputs and digital boundary data') (www.census.ac.uk/casweb). The data are stored in geographic units known as enumeration districts (EDs). EDs were designed with data collection in mind, not data output, and were not scientifically designed, thus the areas differ greatly in geographic extent and shape and outputs are difficult to work with. Within each ED the population is stratified into age classes. For this risk assessment, the age classes used were 0-4yrs (pre-school children), 5-15yrs (school-age children), and 16+yrs (adults).

6.6.2.2 Allocation procedure

A method was required by which to distribute reasonably the population in space. The most accurate method would have been to use Ordnance Survey's (OS) MasterMap™ address layer, Address-Point®. This is a point coverage whereby every address is marked by a spatially referenced point which is designated as either a business or residential address. The ideal solution would be to keep the residential address points and allocate the population among these points. However, the database is not available on an academic licence and is prohibitively expensive. Therefore, an alternative method was sought.

The method devised uses geographic information freely available from EDINA® Digimap®, an on-line mapping data facility which makes OS map data available for academic use (www.edina.ac.uk/digimap). Used in conjunction with the CASWEB population data, a reasonable spatial allocation of population was made as follows.

First a polygon coverage was created from the CASWEB population data which has the population (split by age group) in each polygon, where each polygon is an ED (Frame 1 in Figure 6.2). Next, a point coverage was created from Digimap® Land-Line.plus data.

Land-Line.plus is vector data, and from this it was possible to extract a point coverage for the study region, whereby each point represents a feature on the map and each unique feature type has a unique feature code. The 'buildings' feature code was identified and only points with that feature code were kept. The 'buildings' were described as structures with roofs (Frame 2 in Figure 6.2). A uniform distribution of persons per building was assumed within each ED, so in order to produce an estimate of persons per building, the number of roofed buildings in each ED polygon was counted and the population of that ED was divided by the number of buildings. Each building on the map now has a population allocated to it. Next the geostatistical simulation output grid was laid across the area (Frame 3 in Figure 6.2. Note: a coarse grid is depicted since a fine grid would not display in this schematic diagram). Each cell represents an EU and the population numbers for each building falling inside each EU were summed to give the population per EU (Frame 4 in Figure 6.2). The corresponding population for calculations on a coarse grid was achieved by summing the population in the small EUs falling within each of the larger cells.

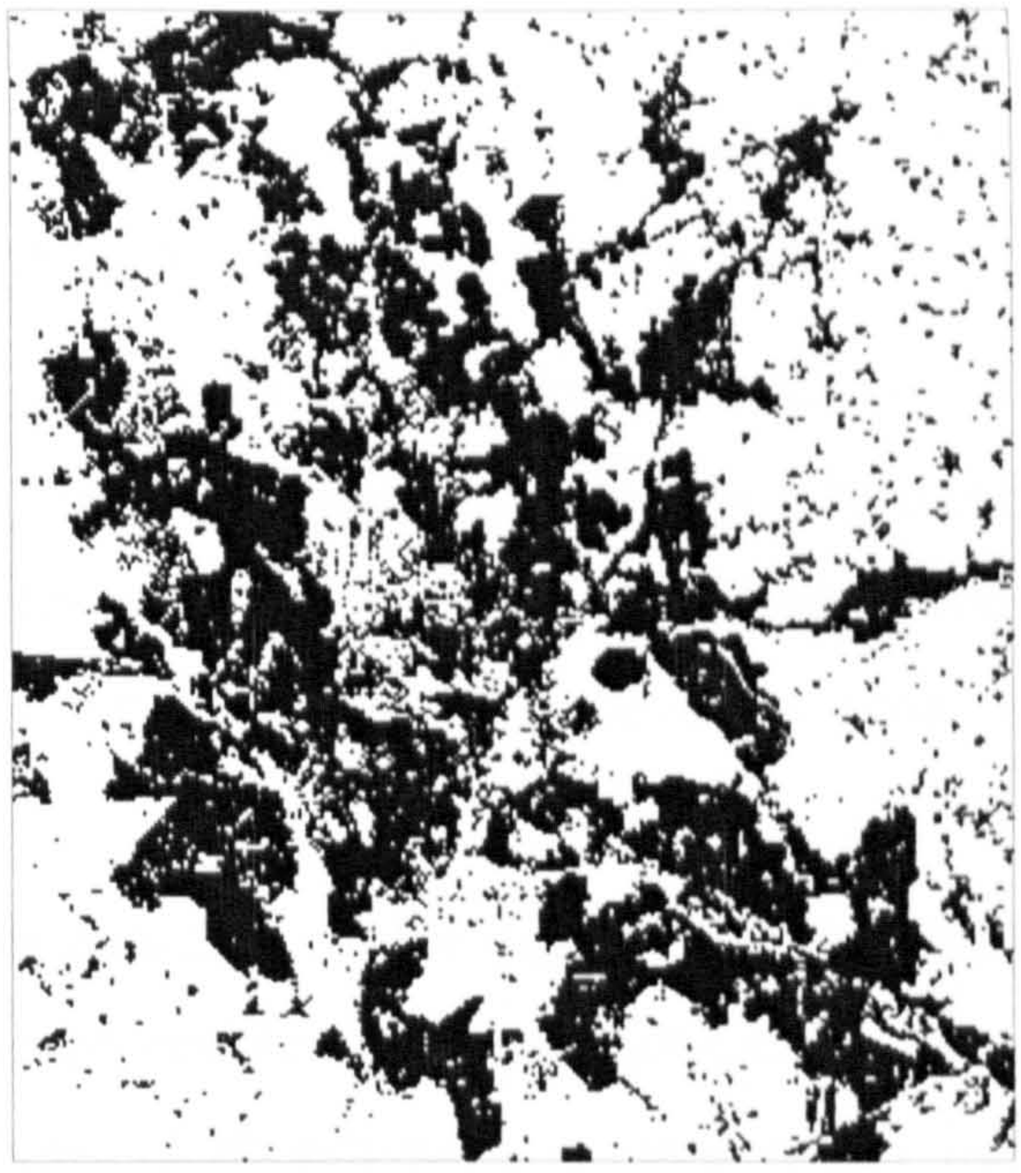
It is recognised that this method is imperfect: some of the buildings to which the population is allocated will be non-residential, e.g. outbuildings, railway stations and industrial units, while some residential buildings will contain multiple residencies, e.g. a block of flats. However, given the density of buildings and the number of people per ED, the overall spatial allocation will be satisfactory.

A national land use database is currently being developed (www.nlud.org.uk) but is still in the pilot stages. A refinement, in future, would be to use a land use layer in conjunction with the buildings point file and keep only those buildings that fall inside residential land use polygons.

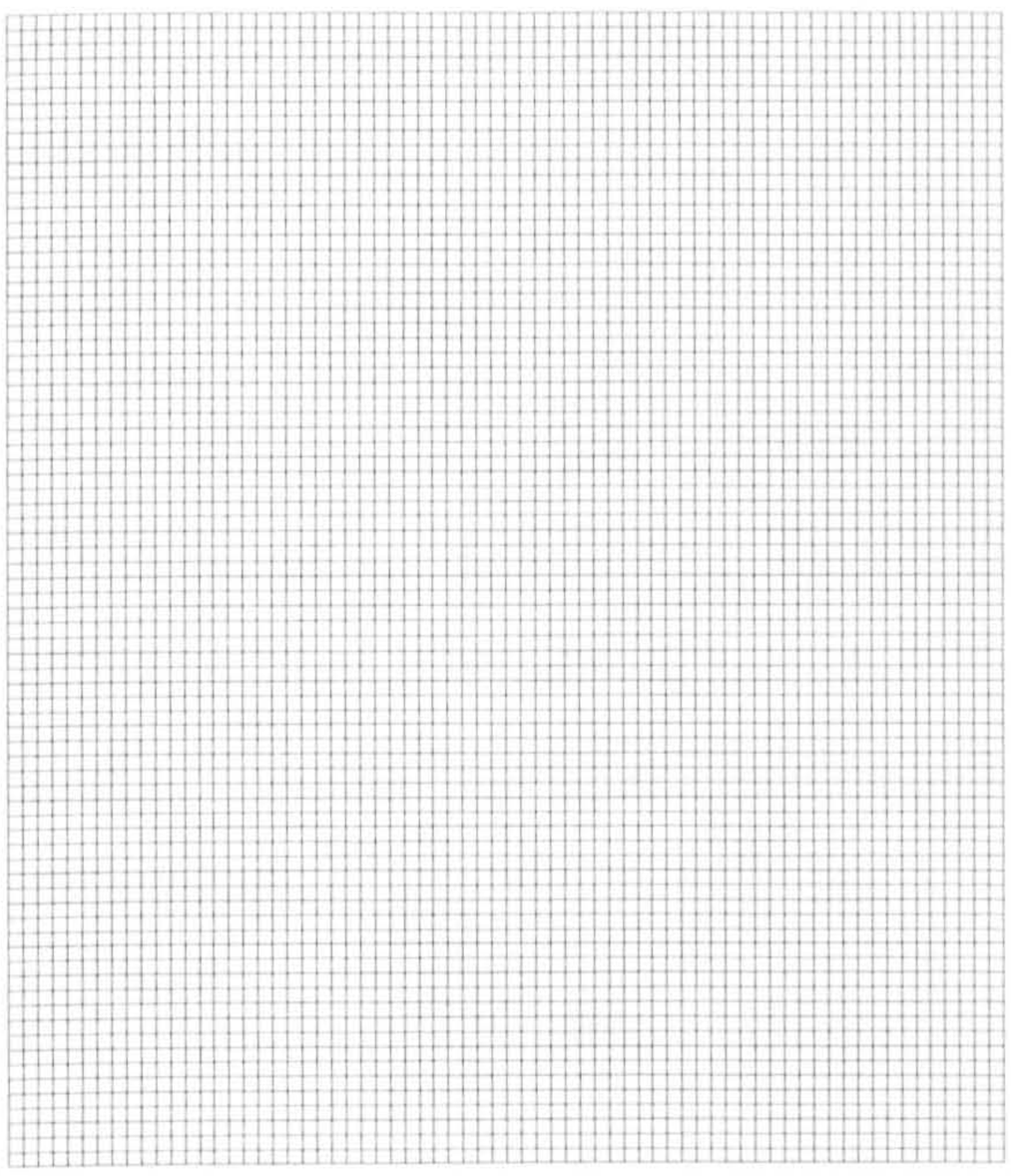
In a similar manner to the simple allocation method, the population ages had to be allocated within the age classes. A uniform birth cohort was assumed and the ages were allocated within each age class by randomly sampling a uniform distribution of 0-4 for the pre-school children, 5-15 for the school-age children. Allocating specific ages to the adults was not necessary since they are not distinguished by age in the risk calculations.



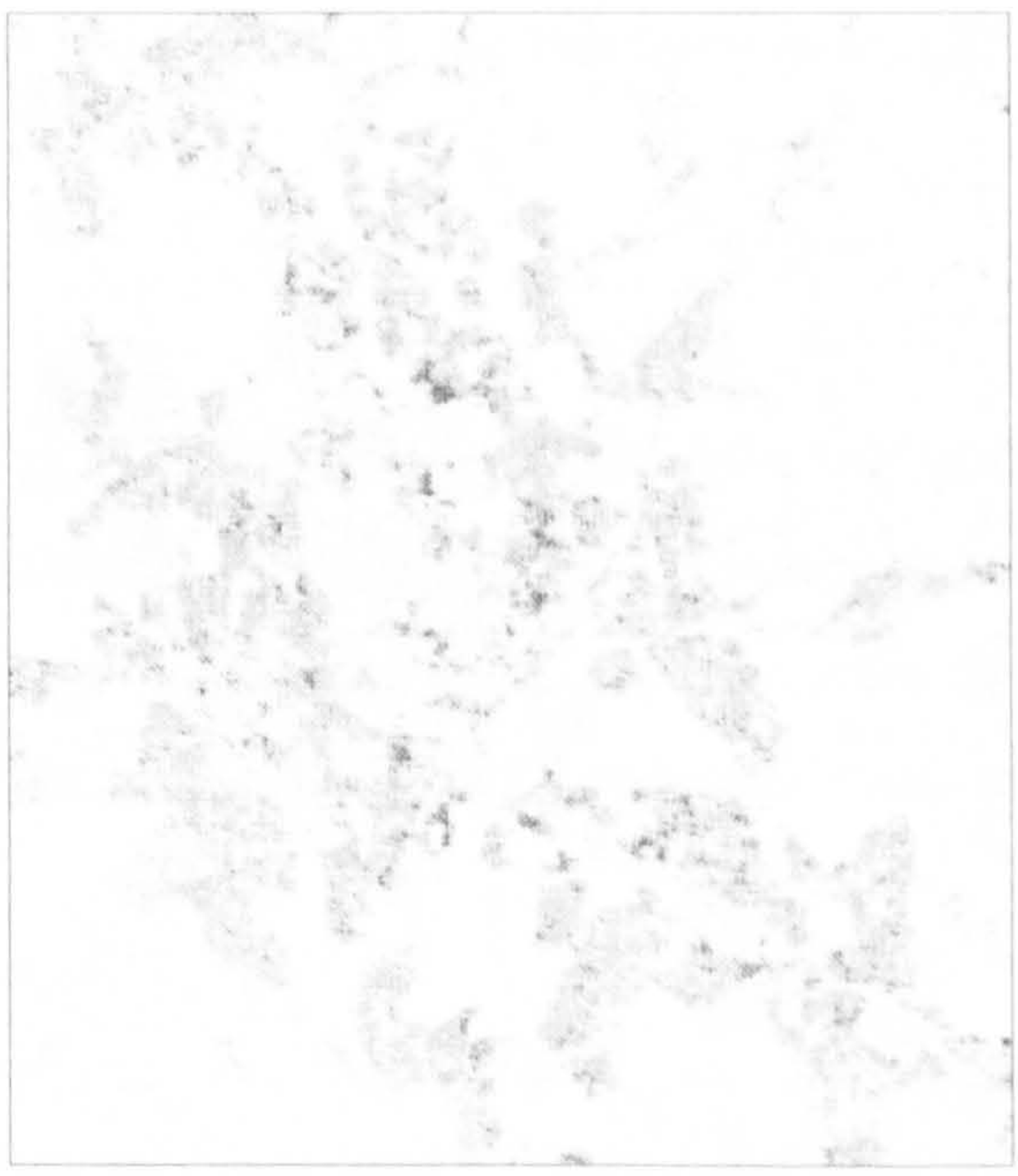
(1)



(2)



(3)



(4)

Figure 6.2. Representation of GIS layers used to allocate the population of Stoke to a 50x50m grid: 1) ED polygons (different shades indicate differing numbers of people per ED). 2) Points representing position of roofed buildings. 3) Grid cells to which the population must be allocated (50x50m grid used but 200x200m grid displayed for clarity). 4) Final population density (darker shades represent higher density).

6.7 Linking population to soil contamination level.

Throughout the calculation procedure, the key issue is ensuring that the link between the population in each EU and the soil contaminant concentration in those EUs is kept intact. This was done by means of xy coordinates and id numbers as follows:

During the geostatistical simulations, each grid cell, or 'geostatistical EU', is identified by an x and a y coordinate. In addition, each EU is given a unique id number. On a single map, there is only one soil concentration per xy coordinate and id number.

When allocating the population to the EUs, each person in the population receives the xy coordinate of the EU and the unique id of the EU. Unlike the soil map, where each EU has only one soil concentration, some EUs can have no people at all, while some can have several. Therefore, it is possible for some people to have the same xy coordinate and EU id number.

During calculation of the intakes, the relevant exposure is obtained by matching the id of the person to the id from the soil map. This procedure is repeated 1,000 times – once for each soil contaminant map – to produce 1,000 estimates of intakes for each individual in the population. By keeping the id numbers and xy coordinates, it is possible to interrogate the intake distributions and summarise the results in meaningful map form.

6.8 Calculations

Programmes were written in SAS® in order to calculate *ADI*. Results from the 1,000 geostatistical simulations (C_{soil} values) were stored as SAS data sets, as were the results of the population allocation procedure (population numbers per EU). These were then used as inputs to the intake calculations. First, a population with characteristics was 'generated'.

6.8.1 Generation of the population

For each case study site, a population had to be 'generated'. Once generated, this population was fixed for the duration of the intake calculations. To facilitate calculation, the intake equation (Section 6.4.3) was rearranged to separate C_{soil} from the other variables as follows:

$$ADI = C_{soil} \times SE$$

where *SE* is referred to as the '$SDR.kg^{-1}{}_{BW}$ equivalent' and is expressed as:

$$SE = \frac{EF_{oral} \times ED_{oral}}{AT}\left(\frac{SDR}{BW} + \sum_{veg\ type}\left(CR_{veg} \times HF_{veg} \times \left(SL_{veg} + CF_{veg}\right)\right)\right)$$

This is discussed in more detail in Chapter 7. The important point to recognise here is that *SE* encompasses all the person specific elements of the intake equation. Hence an

SE can be calculated for each person in the population, by drawing random values of variables from the relevant distributions.

Since many of the distributions are age specific, the first task is to allocate an age to each individual. In the input table, all that is known is the number of people per EU and the age class of those individuals. A specific age was allocated to each person by randomly drawing from a uniform distribution of ages for that age class (i.e. assuming a constant birth cohort), e.g. for 0-4yr olds, a uniform distribution from 0 to 4 inclusive.

Once age is known, for each individual in the population, a value for each variable in the *SE* equation was randomly drawn from the appropriate distribution and the equation solved. Hence each person is represented by an *SE* value. Once generated, the SE values for the population are kept for the duration of the intake analysis, hence there is one population which remains the same for the duration of the intake analysis.

6.8.2 Calculating ADI

To calculate the distribution of *ADI*s for each person, their *SE* was multiplied by the 1,000 different C_{soil} values in the relevant EU to give 1,000 plausible *ADI* values for each person. As mentioned above, each result is tagged with a xy coordinate pair and the matching id number, so results can be interrogated spatially.

6.9 Results output and evaluation

The next task in the development of the spatially-evaluated intake assessment methodology is to find a way to analyse and present the results. After having run the model, we are faced with a great deal of information which must be analysed and presented in a meaningful way.

The format of the results is such that there are 1,000 maps of *ADI* estimates, where each map contains one *ADI* estimate for every member of the population. It can be envisaged as a table (Figure 6.3), whereby each column represents one simulation (map) of *n ADI* estimates (one per person) and each of the *n* rows represents one person with 1,000 *ADI* estimates (one per simulation). The results are variable down each column and uncertain across rows.

Person number	Coordinates x	y	Simulation number 1	2	3	...	1000
1	x_1	y_1	$ADI_{1,1}$	$ADI_{1,2}$	$ADI_{1,3}$	...	$ADI_{1,1000}$
2	x_2	y_2	$ADI_{2,1}$	$ADI_{2,2}$	$ADI_{2,3}$	...	$ADI_{2,1000}$
3	x_3	y_3	$ADI_{3,1}$	$ADI_{3,2}$	$ADI_{3,3}$	...	$ADI_{3,1000}$
...	...	...	...	...	...	...	...
n	x_n	y_n	$ADI_{n,1}$	$ADI_{n,2}$	$ADI_{n,3}$	...	$ADI_{n,1000}$

Figure 6.3. Representation of Simulated *ADI* results.

6.9.1 Assessment criteria

Assume the regulator has determined the ID/TDSI above which intake is deemed unacceptable. For a deterministic assessment, comparison with the threshold is simple. The assessment produces a single intake estimate which is either above (unacceptable) or at/below the threshold (acceptable). This becomes more complicated when the assessment becomes probabilistic.

The next stage would be a probabilistic assessment for one type of critical receptor. In this case, say, 1,000 estimates of intake are produced, for this critical receptor. Now, we have a distribution of 1,000 intake estimates to compare to the regulatory threshold. At this stage a subjective judgement must be made as to what is an acceptable risk. It must be decided whether all estimates must fall at or below the threshold, or whether it is permissible to have a proportion of them over the threshold. This type of assessment forms the basis of the CLEA methodology for deriving SGVs (Soil Guideline Values). The CLEA methodology provides a reasonable worst case scenario by comparing the 95th percentile estimate for a 'critical receptor' to the Health Criteria Value: "By selecting the 95th percentile level, the most unlikely exposure scenarios are excluded but a wide range of more likely situations is included" (DEFRA & EA, 2002g).

USEPA procedures for human health risk assessment from contaminated land do not specify a 'critical receptor' but do recommend using the 'Reasonably Maximum Exposed (RME) individual' for evaluating risk to a population, which is again at the upper 95th percentile of the intake distribution. RME is defined as "the highest exposure that is reasonably expected to occur at a site, and is intended to estimate a conservative exposure case (i.e. well above the average case) that is still within the range of possible exposures" (USEPA, 2001). Their method can take into account variability and uncertainty of the population.

The assessments described above are probabilistic but not spatial. They use a hypothetical population assumed to be exposed to a fixed soil concentration within a

single EU (exposure unit). The intake assessment developed in this thesis adds another dimension. Here we have a whole population of receptors which varies from individual to individual, whose intake varies from simulation to simulation and who are situated in different locations (EUs) in space. This demands a new set of subjective judgements to be made. The subjective judgements can be termed 'decision criteria'.

For the 1,000 *ADI* estimates for the *n* individuals a decision criterion must be designed by which to evaluate the risk to the population. The criterion is two dimensional taking into account variability and uncertainty. The issues surrounding use of two dimensional decision criteria are briefly discussed in the following section.

6.9.2 Risk evaluation

The aim of the intake assessment is to assess if, overall, there is a risk to the population as a whole. The risk may be evaluated in more than one way. A results table, as shown in Figure 6.3, holds information on variability in the population (columns) and uncertainty in the individuals (rows). It is possible to assess the overall risk via population or individual. By population, the fraction of the population per simulation (column) in excess of the ID/TDSI is calculated. The distribution of fractions over 1,000 simulations represents the uncertainty in the fraction of the population that exceeds the ID/TDSI. By individual, the probability that each individual exceeds the ID/TDSI is calculated. This is represented by the fraction of the simulations (across a row) in which their intake exceeds the cut off. The distribution of probabilities for all individuals represents the variability between individuals in the population.

The expected proportion of the population at risk is the proportion of all cells in the table that exceed the ID/TDSI. It is the same whether one calculates rows then columns or columns then rows and is represented by the mean of either distribution described. However, depending on which evaluation method is used (rows then columns or columns then rows), the outcome of an assessment may differ when using seemingly identical two dimensional decision criteria. A demonstration follows in the following section.

6.9.2.1 Two dimensional decision criteria

Working by population (i.e. by column then row: Method P) the percentage of the population allowed to exceed the threshold on each map must be stipulated. In keeping with the criteria described in Section 6.9.1, this is set to 5%. In effect, it is saying that

the 95th percentile individual is at or below the threshold, while accepting that the top 5% of individuals on that map may have excessive intakes. A count of maps in which the percentage of the population above the intake threshold is more than 5% is obtained. The second dimension requires stipulation of the allowable number of maps recording excessive population intakes. For example, if 20 of the 1,000 maps record more than 5% of the population have excessive intakes, is this acceptable? The strictest criterion would stipulate that no maps are allowed to have more than 5% of the population in excess of the threshold. However, this would be considerably stricter than analyses that do not take uncertainty into account. A percentage more in keeping with these would be 50%, and this is adopted for this thesis. It is a purely subjective choice which, in reality, would need to be agreed between the assessor and regulator in advance of the assessment. Thus, in Method P, the decision criterion is passed if in at least 50% of the simulations, no more than 5% of the population has intakes in excess of the ID/TDSI.

Working by individual (row first then column: Method I) one obtains the distribution of individual probabilities of exceeding the ID/TDSI. The decision criterion is passed if no more than 5% of the individuals in the population have a probability of exceeding the ID/TDSI of 50% or more. This is equivalent to saying that the median intake estimate for each individual should not exceed the ID/TDSI in more than 5% of the population.

We now consider some extreme hypothetical examples to illustrate the difference in working with the alternative decision criteria. The schematic in Figure 6.4 represents the results of three intake scenarios. Shading represents the results which exceed the ID/TDSI and is equal to one third in each case. Results table 1 represents a population in which intakes are uncertain but not variable: in 33% of simulations all individuals exceed the ID/TDSI and in the remaining simulations no individuals exceed the ID/TDSI. The results 'pass' whether assessing by Method P (only in 33% of simulations are there more than 5% of the population in excess of the threshold) or by Method I (no individual has a probability greater than 50%), and we conclude that there is no risk to the population. In table 2, the results 'fail' both by Method P (more than 5% of the population exceeds the threshold in every simulation) and by Method I (33% of individuals have a probability in excess of 50%) and we conclude that there is a risk to the population. In table 3, however, Method P fails (33% of the population exceeds the ID/TDSI in 100% of simulations) but Method I passes (each individual has a probability of less than 50%).

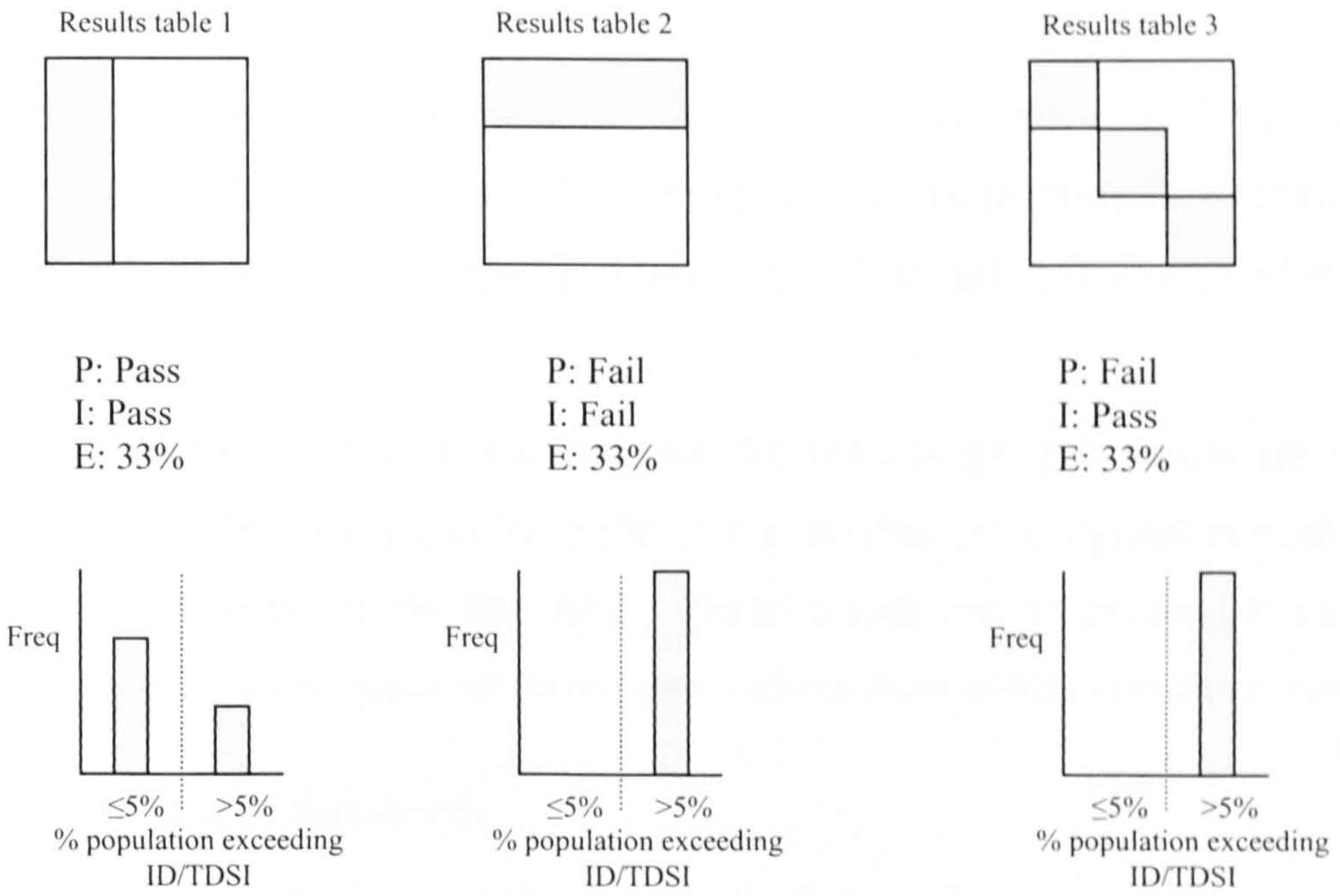


Figure 6.4. Using two dimensional decision criteria: a schematic representation of results. Shading in results table represents ADI estimates in excess of ID/TDSI, where individuals are in rows and simulations are in columns. P: applying criterion by column then row. I: applying criterion by row then column. E: expected proportion of population in excess of ID/TDSI.

In all cases the expected proportion of the population with intakes in excess of the cut off is 33%. It is possible to set a decision criterion for the expected proportion (5% maximum: Method E) in which case all three scenarios 'fail'. The expected number criterion is simple, in that it is one dimensional, but the uncertainty element is lost in presenting the results.

Each method is a valid risk evaluation method. It should be noted that when tackling the problem directly at the population level, risk to the population is assessed a whole map at a time for each of the 1,000 simulations. Therefore, spatial correlation of the soil values to which the population is exposed is preserved. When assessing risk to an individual, in isolation from the rest of the population, spatial correlation of soil values is lost. (What matters in the individual case is how the soil concentration in an individual EU varies over the course of 1,000 simulations.) Because of this, this author prefers the idea of working directly at the population level, since it guarantees that the spatial correlation of the soil values, which the method tries to preserve, is passed on into the intake results. In addition, since the aim is to assess the risk to the population as a whole, this method is more useful since it maintains a measure of the uncertainty in the percentage of the population in excess of the ID/TDSI.

6.9.3 Spatial displays

Having gone to the extent of calculating the results in a spatial manner and preserving the spatial coordinates for each individual in the population, the information is available to be able to map the results. How to summarise the 1,000 individual maps of intake results is also a challenge.

Maps can be produced to show where in space the risks to the population are most prevalent. For example, counts can be made of the number of occupants in each EU that have intakes in excess of the threshold. These counts can be mapped to enable efficient targeting of risk management strategies to where most at-risk receptors reside.

6.9.4 The drawback of thresholds

Thresholds produce easy to interpret results, in as much as, for example, a person is either below the threshold and not at risk or is above the threshold and is at risk. However, this is a very simplistic way of assessing risk, since a person whose intake is ten times in excess of the threshold is classed in the same risk bracket as a person who is only slightly in excess of the limit. Using thresholds can lead to a loss of information and this should be recognised. Targeting risk management strategies to areas where the number of at-risk receptors is highest may help people who are only slightly over the threshold while ignoring an area with few very highly exposed individuals whose health is much more in danger. This problem is countered by bracketing at-risk individuals into intake classes and mapping the number in, for example, highly excessive, medium and slightly excessive categories.

6.10 Summary of methodology

Figure 6.5 summarises the newly developed spatially-evaluated probabilistic risk assessment procedure, the advantages and drawbacks of which will be analysed during the remainder of this thesis.

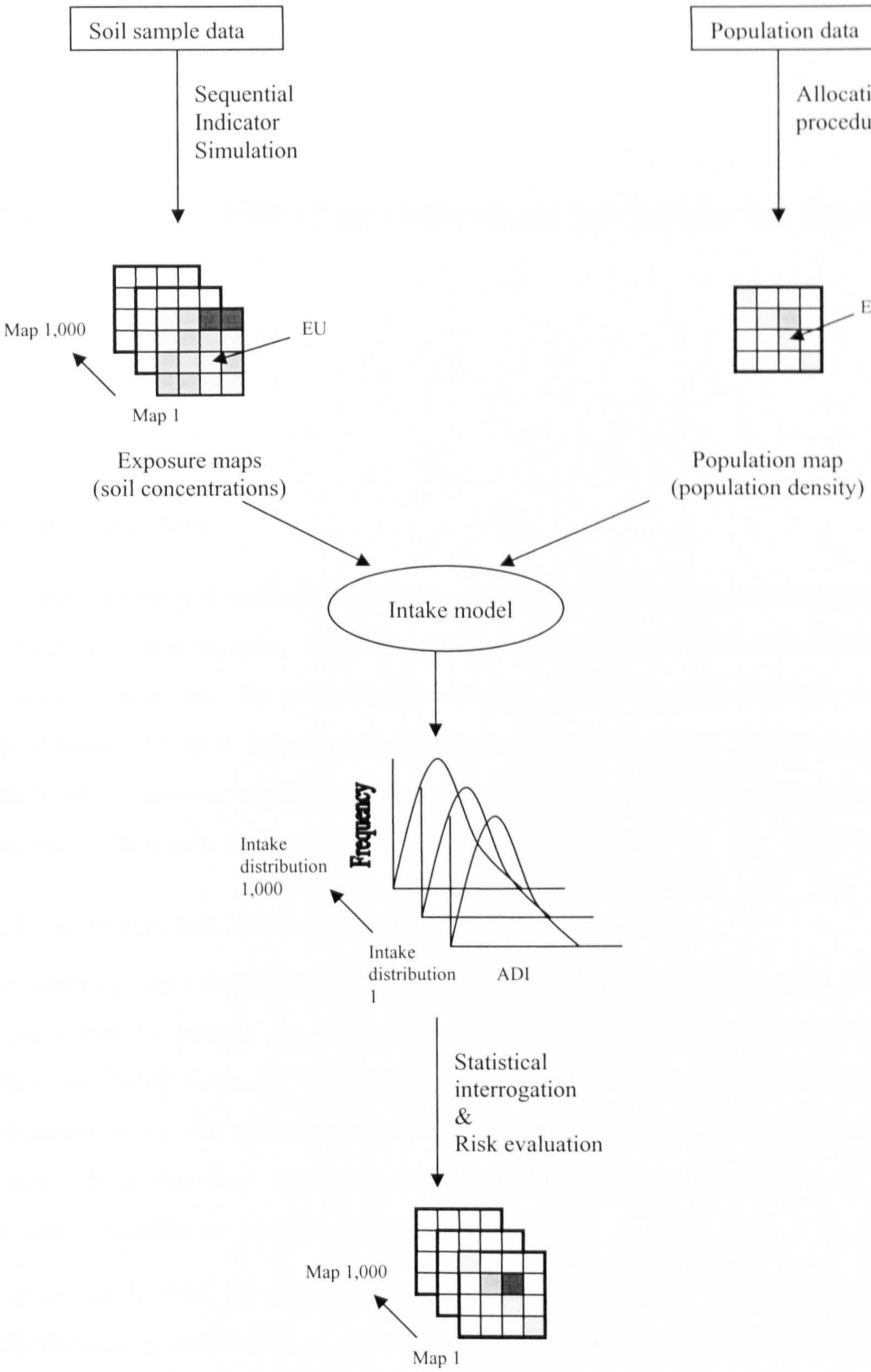


Figure 6.5. Spatially-evaluated probabilistic risk assessment procedure.

Chapter 7 Effect of individual variation on intake

7.1 Introduction

The newly developed methodology for a spatially-evaluated intake assessment was laid out in the previous chapter. There are two sources of information which contribute to the intake assessment: the soil contaminant data produced in the exposure assessment step (Chapter 5); and information about the local population. This chapter looks specifically at how variations in characteristics of the receptor population can affect the outcome of the intake assessment.

7.1.1 Between-individual variation

The methodology developed in this thesis is a probabilistic one that takes into account the fact that all people are different. This variation in the characteristics of the individuals being assessed will have a marked effect on the outcome of the intake assessment: at a given soil concentration, the calculated intake will differ from person to person. It is therefore important and necessary to assess the effect of between-individual variation on predicted intakes.

In order to do this, the CLEA algorithm, used as a basis for the risk assessment methodology, is rearranged to extract the factors that are specific to the individual. These factors are summarised into a single variable, the SDR equivalent (*SE*), which represents the equivalent daily soil ingestion if all the contaminant entered the body directly from soil (as opposed to being absorbed into vegetables, for instance). Thus $ADI = C_{soil} \times SE$. The make-up and distribution of *SE*, representing the between-individual variation can then be studied in detail, independently of the soil concentration component of the algorithm. It is important to recognise that *SE* differs between

contaminants, so will differ for an individual depending upon the contaminant of concern.

To the best of the author's knowledge, such an in depth analysis in terms of '*SE*' has not been presented before.

First of all, the derivation of the *SDR* equivalent, *SE*, is presented. Following this, a deterministic assessment of the factors contributing to *SE* is conducted for different age groups for each element of concern. A population of individuals is then generated probabilistically and the major contributors to the individuals' *SE*s are confirmed and the population *SE* distributions for each contaminant of concern are investigated. Using the distribution of *SE* values for this population, it is possible to assess the proportion that would be over-exposed at different soil concentrations. Finally, the effect of variation in *SE* on the proportion of individuals in excess of the intake threshold is shown.

Example data sets used later in this thesis use Nickel (Ni), Arsenic (As) and Cadmium (Cd), therefore these contaminants will each be investigated in this chapter. In many graphic displays, the data are classed into age groups 0-5yrs, 6-15yrs and 16+yrs. The basis for this is that *SDR* is variable in the 0-5yr age groups, but not in the over-fives, and those aged 16yrs+ are not stratified further by age, so these are natural classes.

7.2 Rearranging the intake algorithm

Recall the basic CLEA intake equation:

$$ADI = \frac{\left(IR_{oral} \times EF_{oral} \times ED_{oral}\right)}{BW \times AT}$$

where

$$IR_{oral} = IR_{dsi} + IR_{isi} + IR_{veg}$$

and

$$IR_{dsi} = C_{soil} \times SDR$$

$$IR_{isi} = C_{soil} \times \sum_{veg\ type}\left(CR_{veg} \times BW \times HF_{veg} \times SL_{veg}\right)$$

$$IR_{veg} = C_{soil} \times \sum_{veg\ type}\left(CR_{veg} \times BW \times HF_{veg} \times CF_{veg}\right)$$

where

ADI: average daily intake of a chemical from soil ($mg.kg^{-1}_{BW}.day^{-1}$)

IR: chemical exposure rate ($mg.day^{-1}$)

EF: exposure frequency ($day.yr^{-1}$)

ED: exposure duration (yr)

BW: bodyweight (kg)

AT: averaging time (day)

IR_{dsi}: chemical exposure rate from direct (inadvertent) soil ingestion ($mg.day^{-1}$)

IR_{isi}: chemical exposure rate from indirect ingestion of soil attached to vegetables ($mg.day^{-1}$)

IR_{veg}: chemical exposure rate from ingestion of contaminants absorbed into vegetables ($mg.day^{-1}$)

C_{soil}: contaminant concentration in soil ($mg.g^{-1}$)

SDR: average daily soil and dust ingestion rate ($g.day^{-1}$)

CR_{veg}: consumption rate per vegetable ($g.FW.day^{-1}$) (FW=Fresh weight)

BW: bodyweight (kg)

HF_{veg}: fraction that is home grown per vegetable (-)

SL_{veg}: amount of attached soil per vegetable ($g.g^{-1}FW$)

CF_{veg}: soil to plant concentration factor ($\mu g.g^{-1}FWplant/\mu g.g^{-1}DWsoil$) (DW=Dry weight)

Rearranging, and making C_{soil} a common factor, we arrive at:

$$IR_{oral} = C_{soil} \times \left(SDR + \sum_{veg\ type} \left(CR_{veg} \times BW \times HF_{veg} \times SL_{veg} \right) + \sum_{veg\ type} \left(CR_{veg} \times BW \times HF_{veg} \times CF_{veg} \right) \right)$$

and therefore;

$$ADI = C_{soil} \times$$

$$\frac{\left(SDR + \sum_{veg\ type} \left(CR_{veg} \times BW \times HF_{veg} \times SL_{veg} \right) + \sum_{veg\ type} \left(CR_{veg} \times BW \times HF_{veg} \times CF_{veg} \right) \right) \times EF_{oral} \times ED_{oral}}{BW \times AT}$$

The second term in this equation has units of ($g.kg^{-1}_{BW}.d^{-1}$). It represents the equivalent daily soil ingestion if all the contaminant entered the body directly from soil (as opposed to being absorbed into vegetables, for instance). It can be thought of in terms

of $SDR.kg^{-1}_{BW}$. This term shall be referred to as the '$SDR.kg^{-1}_{BW}$ equivalent' or *SE* from now on; hence $ADI = C_{soil} \times SE$.

Simplifying further:

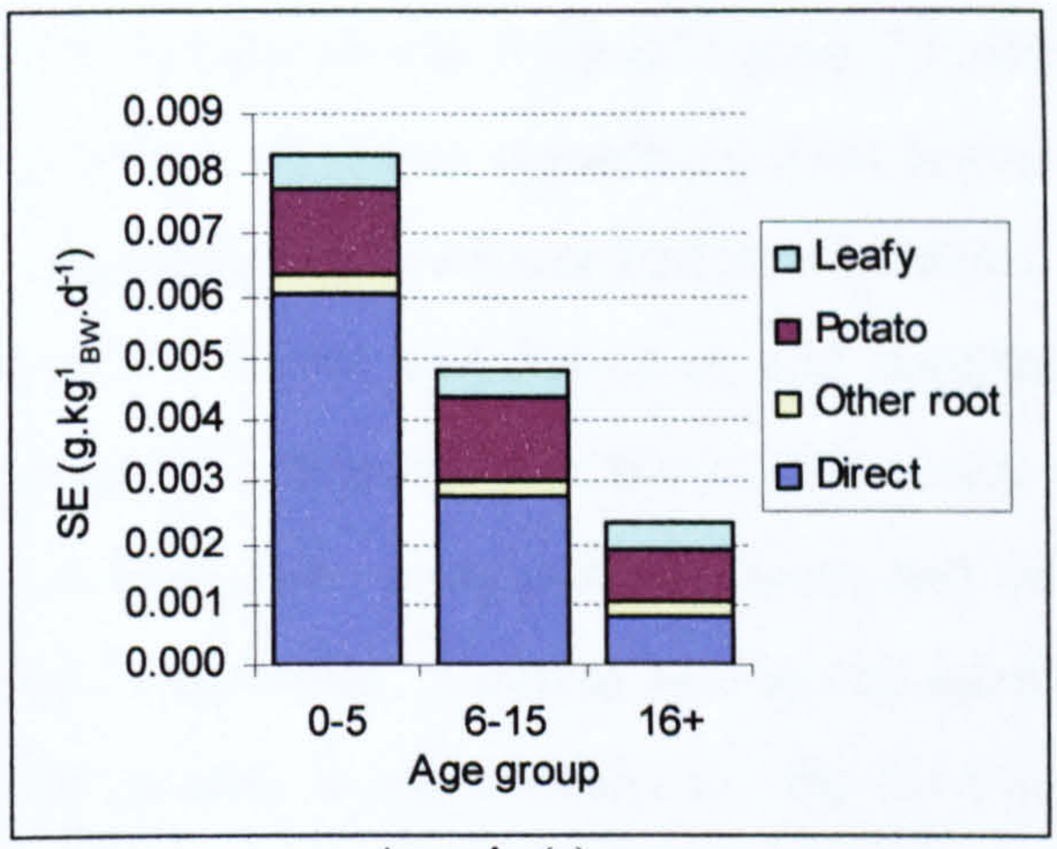
$$SE = \frac{EF_{oral} \times ED_{oral}}{AT}\left(\frac{SDR}{BW} + \sum_{veg\ type}\left(CR_{veg} \times HF_{veg} \times \left(SL_{veg} + CF_{veg}\right)\right)\right)$$

7.3 Deterministic investigation

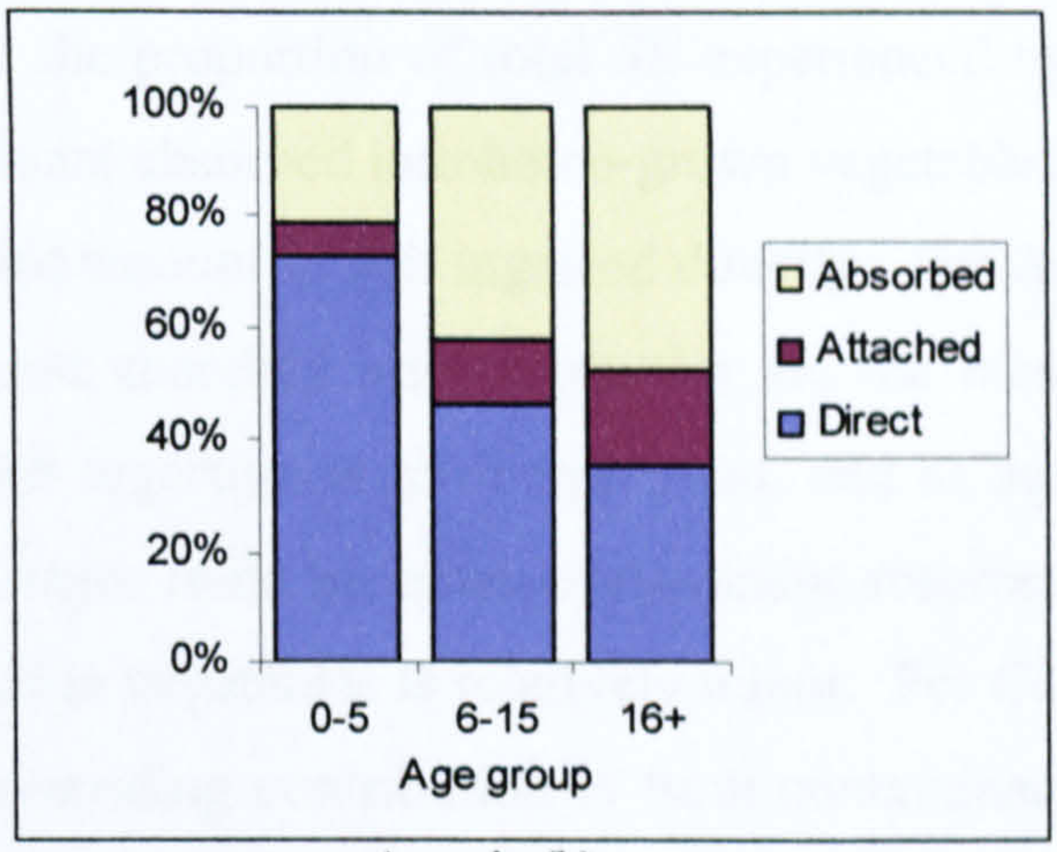
In this section, a deterministic investigation of the influence of variables is undertaken by using the mean values from all the probabilistic variables described in Chapter 6. The *SE* term of the rearranged equation is solved and the proportion of *SE* due to different factors is assessed. The term is solved once for each age class to account for variables that differ by age i.e. separately for each age from 0 to 15 and once for age 16+.

Figure 7.1 shows the major contributors to *SE* for each of the three elements of concern. The left hand column shows the amount due to eating leafy vegetables, potatoes and other root vegetables and due to direct ingestion of soil and/or dust. The right hand column splits the *SE* into the proportion due to contaminant absorbed into home-grown vegetables, attached in soil on the vegetable, and directly ingested.

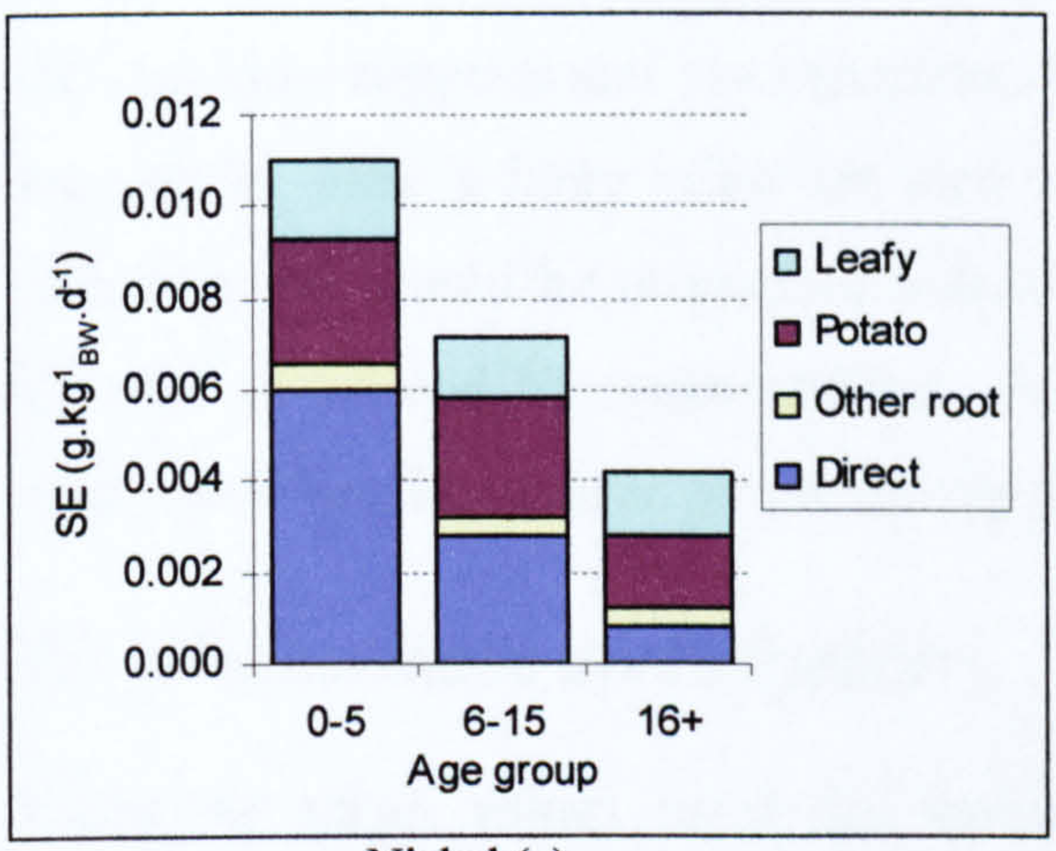
The left hand column in Figure 7.1 clearly demonstrates that, overall for this model and for all three elements considered, *SE* reduces with age. The reason for this reduction differs between elements. In all three cases, direct soil ingestion contribution lessens noticeably with age. This would seem sensible, since young children's behavioural habits (e.g. mouthing of objects, playing on the ground) would lead to higher direct ingestion, and these behaviours are modified as the person grows older. Direct soil ingestion makes a sizeable contribution to *SE* for As, and this falls off for Ni and to such an extent in Cd that this route becomes almost negligible. For Cd, the biggest reduction in *SE* with age is due to lower consumption of potato per kg bodyweight.



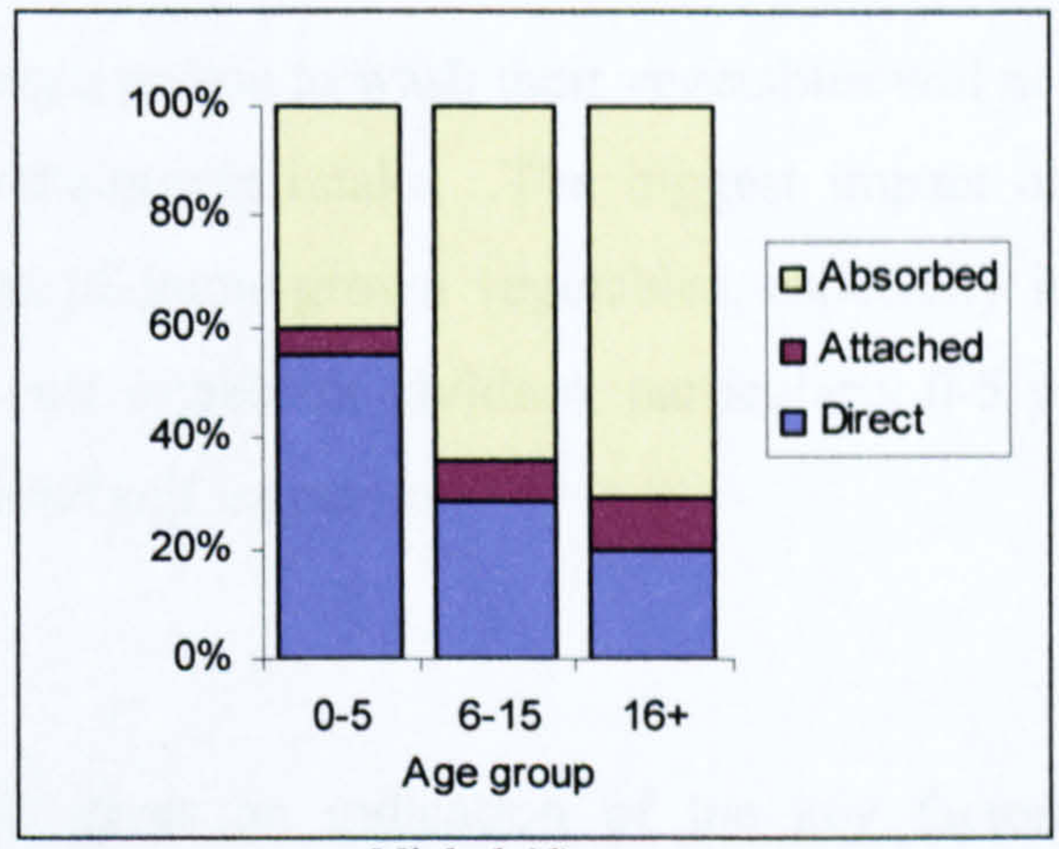
Arsenic (a)



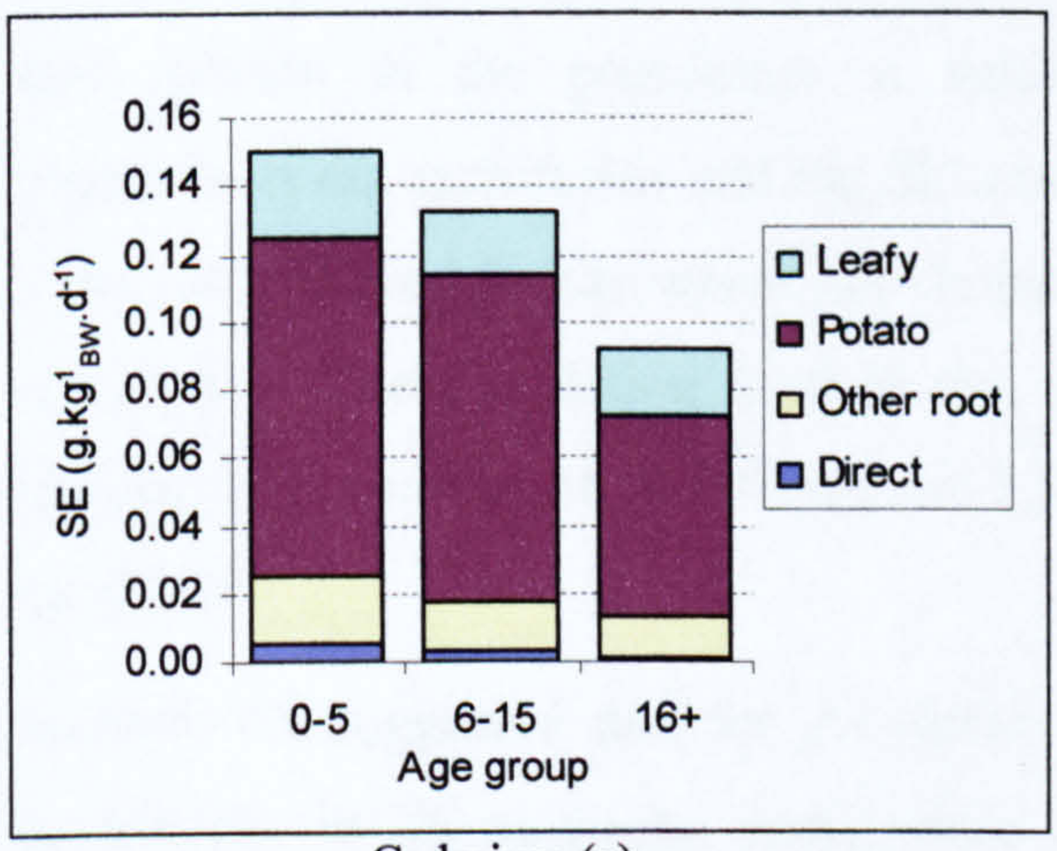
Arsenic (b)



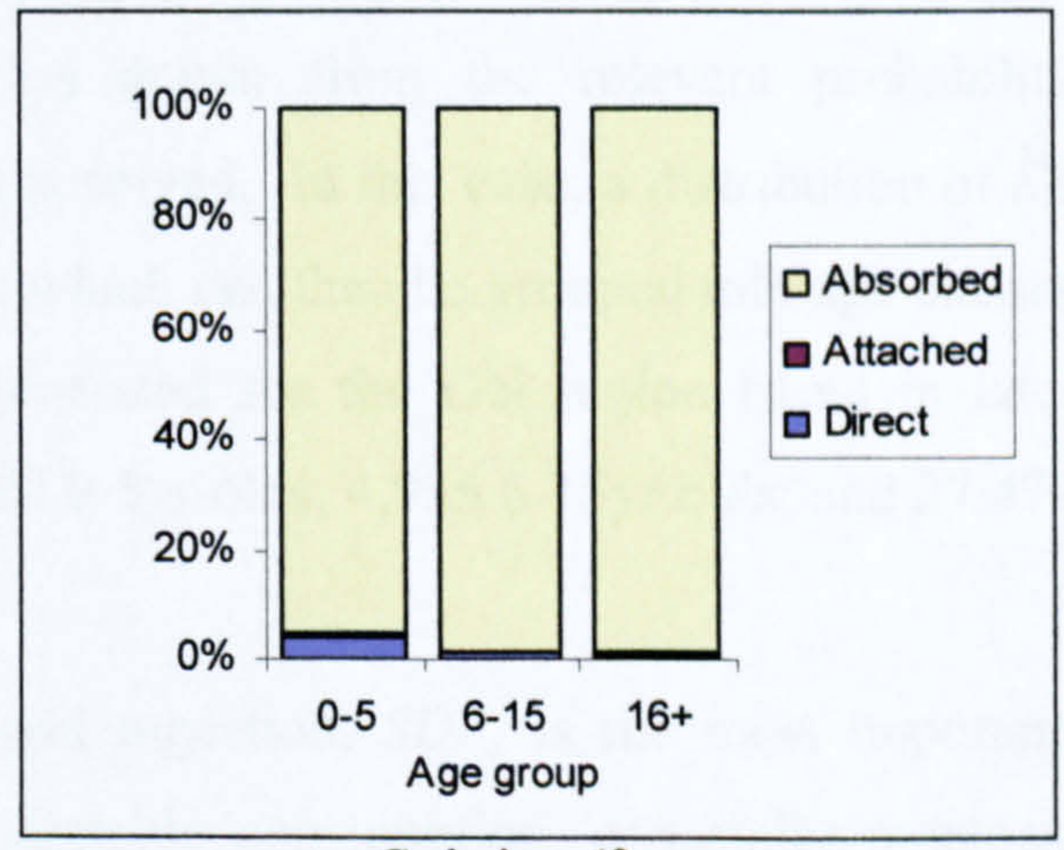
Nickel (c)



Nickel (d)



Cadmium (e)



Cadmium (f)

Figure 7.1 Major factors contributing to *SE* for As, Ni, and Cd at pH7. Left: *SE* split by four major contributors: eating leafy vegetables; potatoes; and other root vegetables; and by direct ingestion of soil and/or dust. Right: Proportion of *SE* attributable to contaminant absorbed into consumed vegetables, attached to consumed vegetables and ingested directly from soil and/or dust.

The graphs on the right of Figure 7.1 show the proportion of total *SE* experienced by each age class that is made up from contaminant absorbed into home-grown vegetables, contaminant in soil attached to vegetables and amount of soil ingested directly. For As and Ni, direct soil ingestion and contaminant absorbed into vegetables are the most important factors. For 0-5yr olds direct soil ingestion is most important, and as age increases, this contribution lessens and the major route becomes contaminant absorbed into vegetables. Contaminant in soil attached to vegetables is relatively minor. For Cd, the picture is quite different. By far the overriding contribution is from contaminant absorbed into vegetables, with the other two routes being negligible.

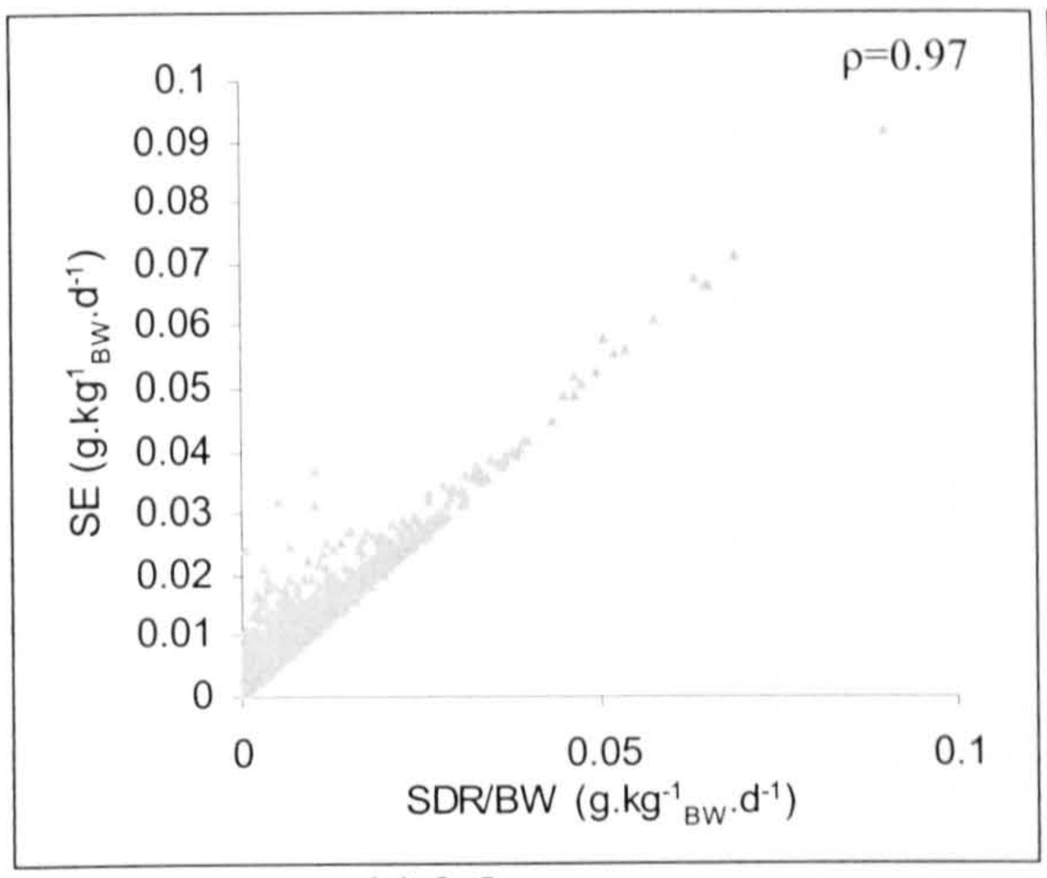
This analysis suggests that just recommending a person to wash their vegetables will not necessarily have a large effect on their contaminant intake. The biggest impact on reducing risk would be to stop consumption of home-grown vegetables, especially in the case of Cd and Ni contamination. Arsenic intake for children, particularly 0-5 yr olds would best be reduced by preventing direct soil ingestion.

7.4 Probabilistic investigation

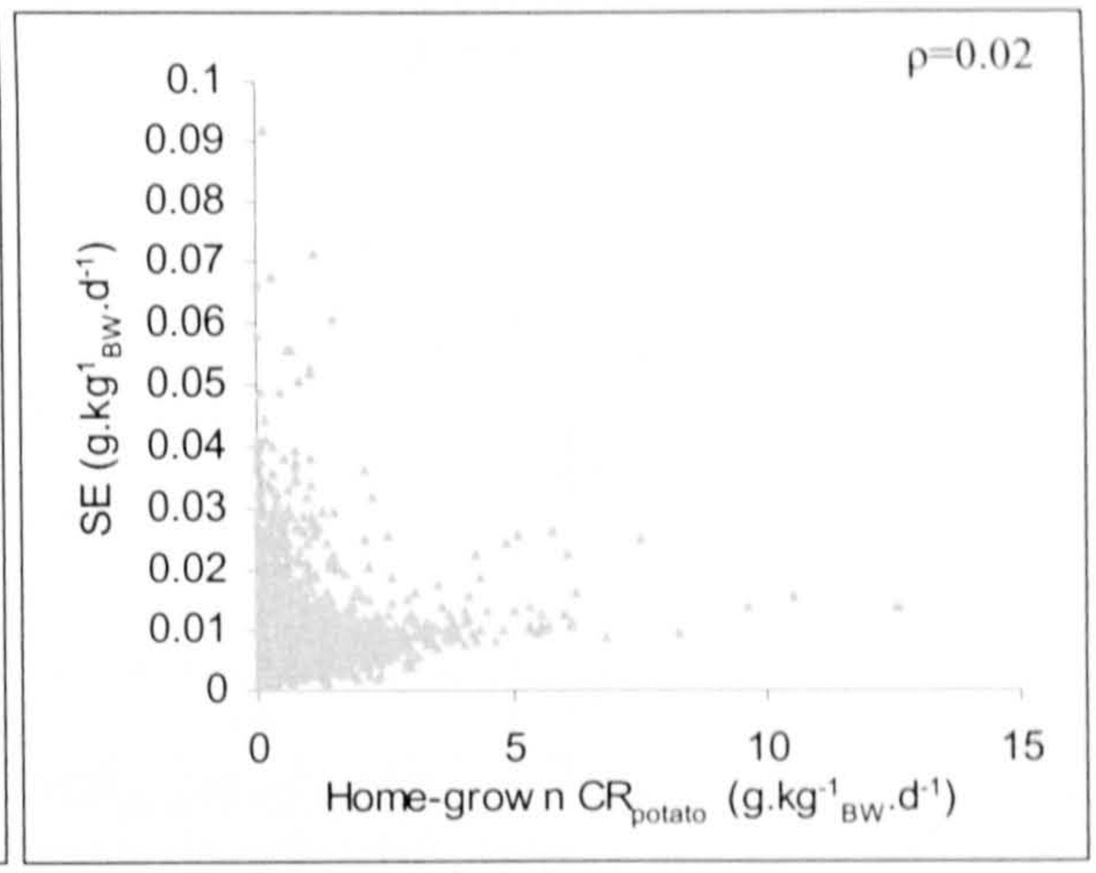
Using the mean values for every variable gives an indication of the key factors contributing to *SE*. Next, a population is investigated in a probabilistic manner. For each person in the population, a value is drawn from the relevant probability distribution for each factor and the *SE* term is solved. In this case, a distribution of *SE* values is achieved for the whole population, which can then be grouped into age classes as before. The population used is that generated for the Gai region (used in later chapters). It consists of 34,997 people: 2,582 0-5yr olds; 4,935 6-15yr olds; and 27,479 aged 16+.

Section 7.3 suggested that for As, direct soil ingestion, *SDR*, is the most important contributor to *SE* at young ages, while vegetable consumption, especially potatoes, becomes most important in the older age groups. To investigate this, *SDR* is plotted against *SE* by age group in Figure 7.2 (left hand column). It is clear that in the lowest age group (plot a), there is a strong correlation between *SDR* and *SE* which is not present for the two older age groups, 6-15yr olds (b) and adults (c). To investigate the importance of home-grown potato consumption ($g.kg^{-1}_{BW}.d^{-1}$) for As, this factor was calculated for each individual by multiplying HF_{potato} (home-grown fraction) with CR_{potato} (total potato consumption rate, $g.kg^{-1}_{BW}.d^{-1}$). Home-grown potato consumption is plotted against *SE*, by age group, in the right hand column of Figure 7.2. Correlation

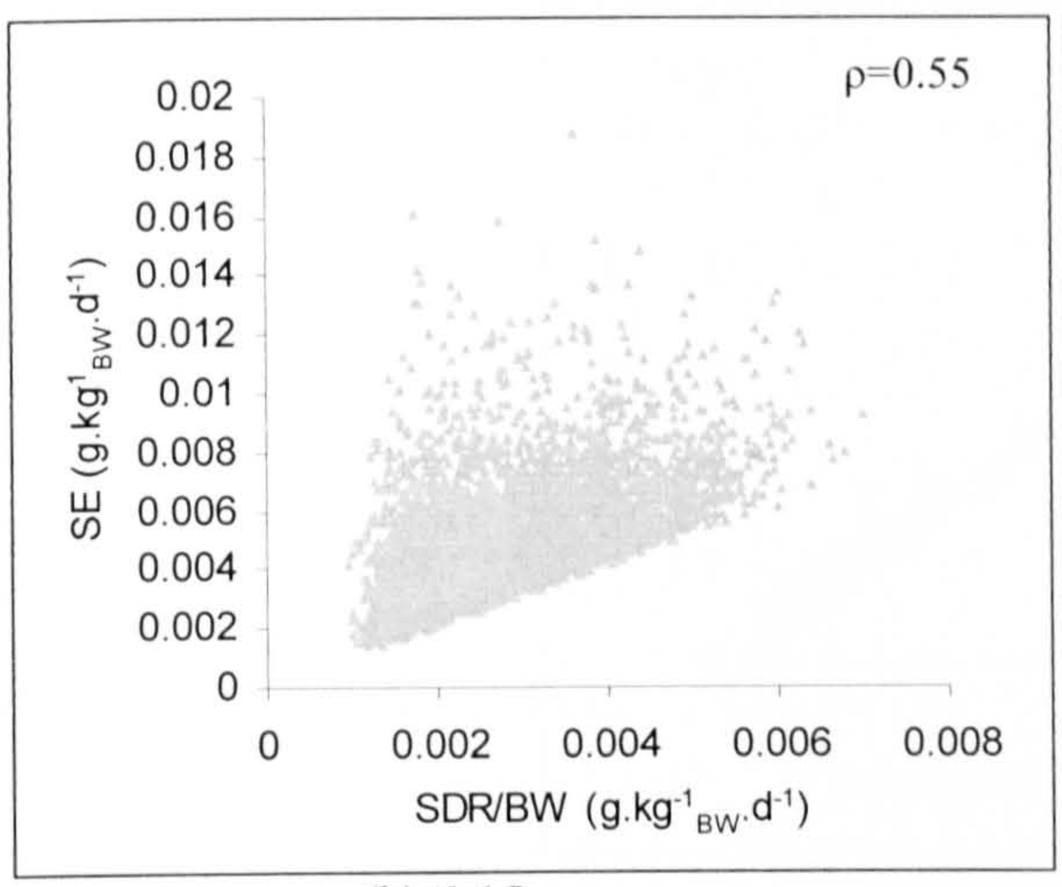
increases with age, confirming the importance of the influence of home-grown potato consumption in older age categories. For Cd, Section 7.3 suggested that potatoes were by far the largest contributor to *SE* for all age groups. Plotting home-grown potato consumption against *SE*, by age group, in Figure 7.3 confirms this: there is a strong relationship at all ages.



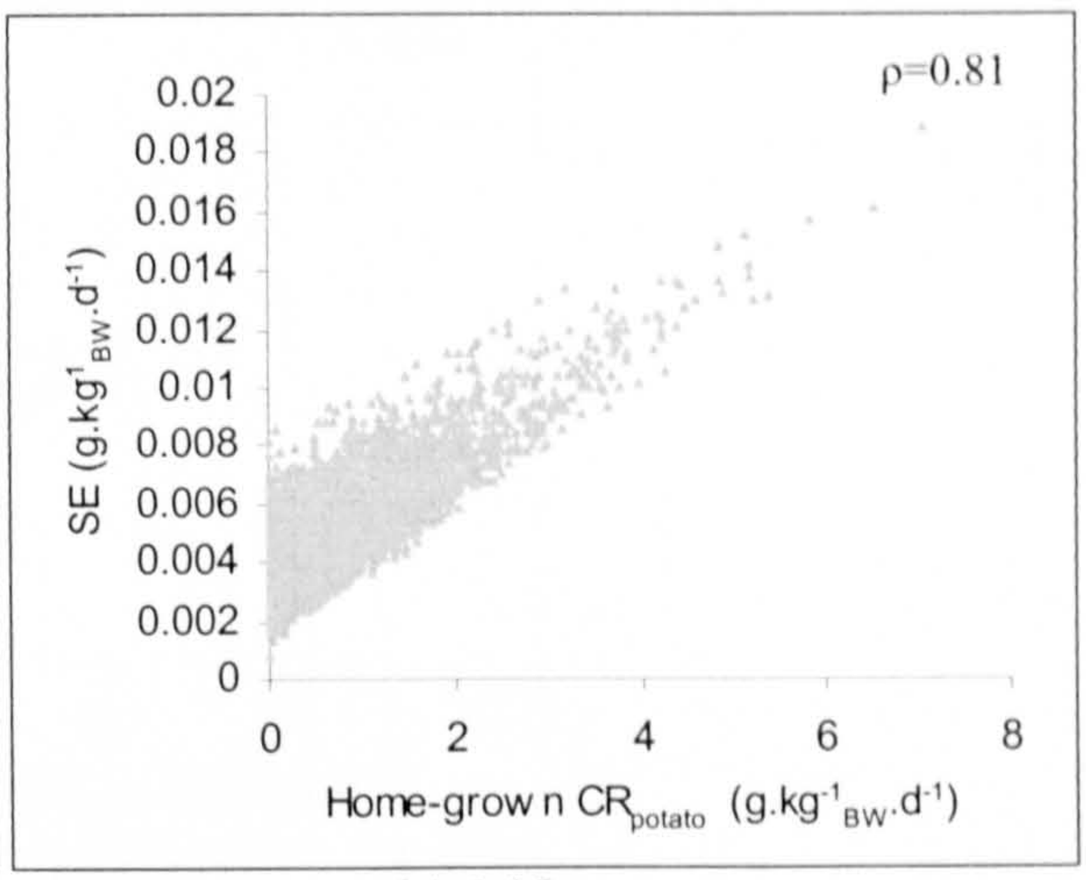
(a) 0-5 yrs



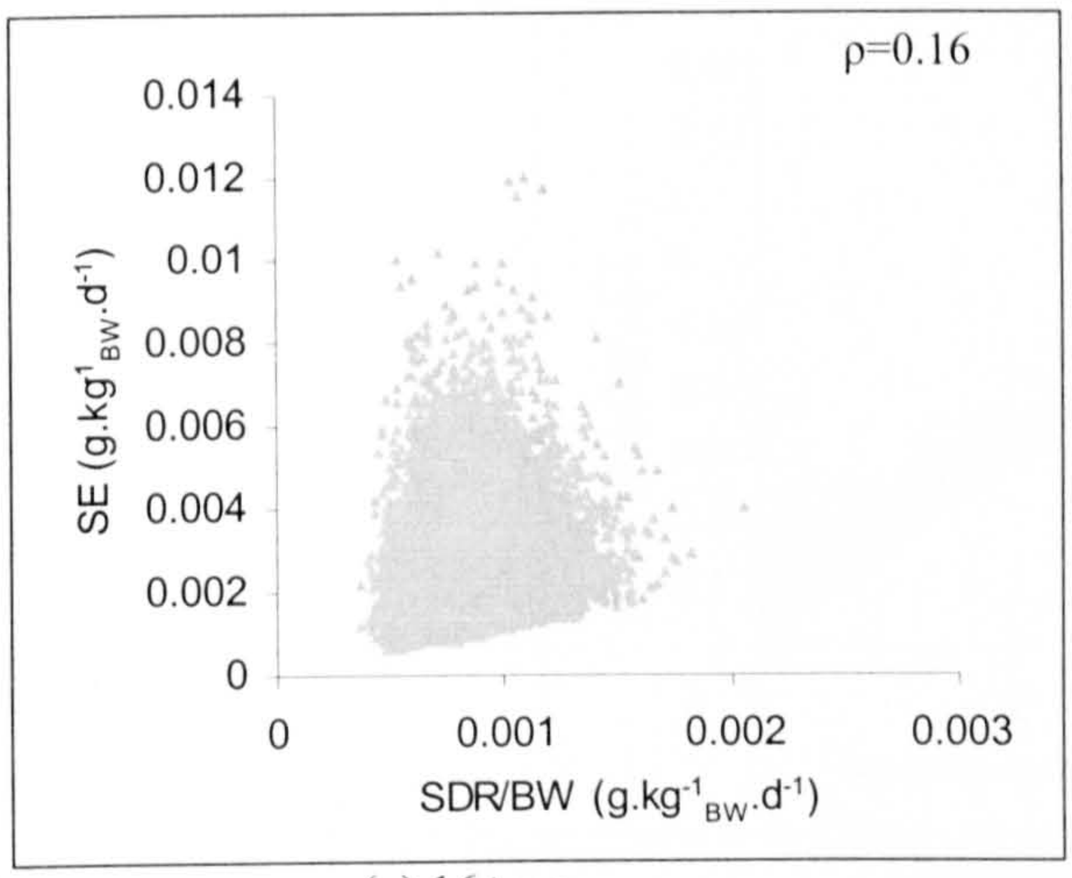
(d) 0-5 yrs



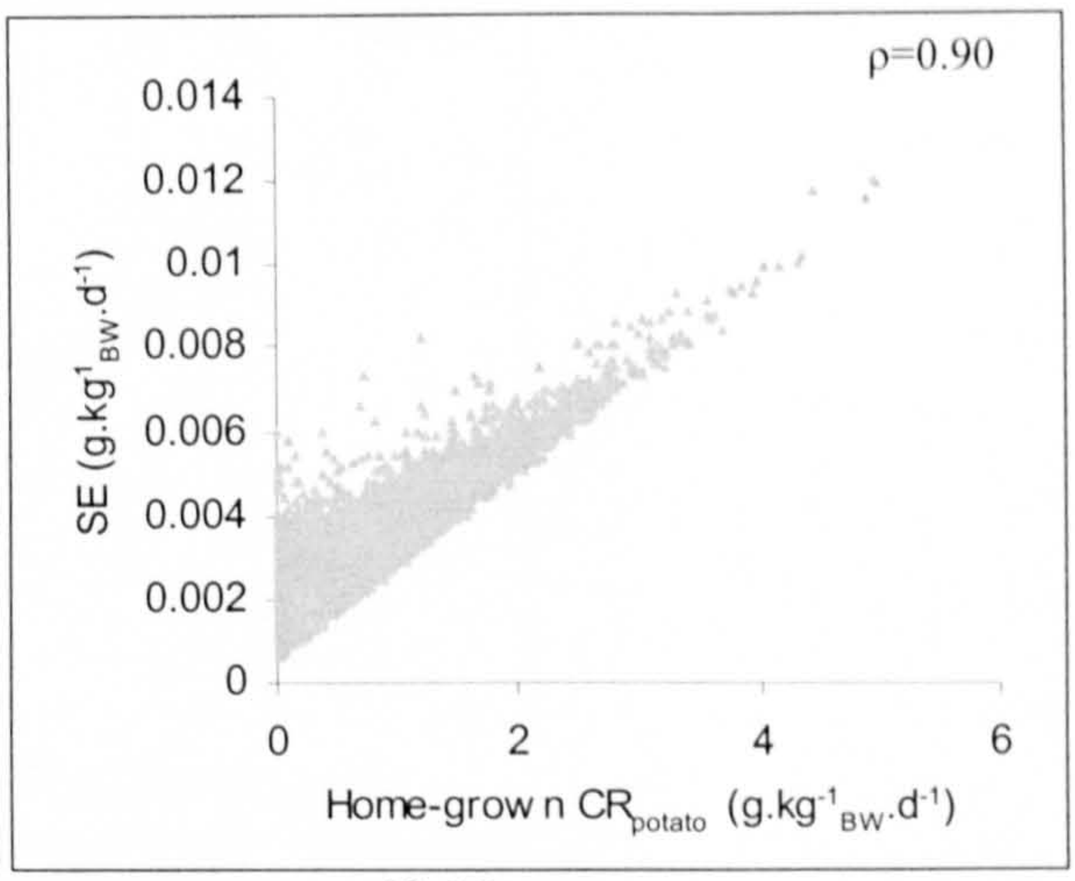
(b) 6-15yrs



(e) 6-15yrs

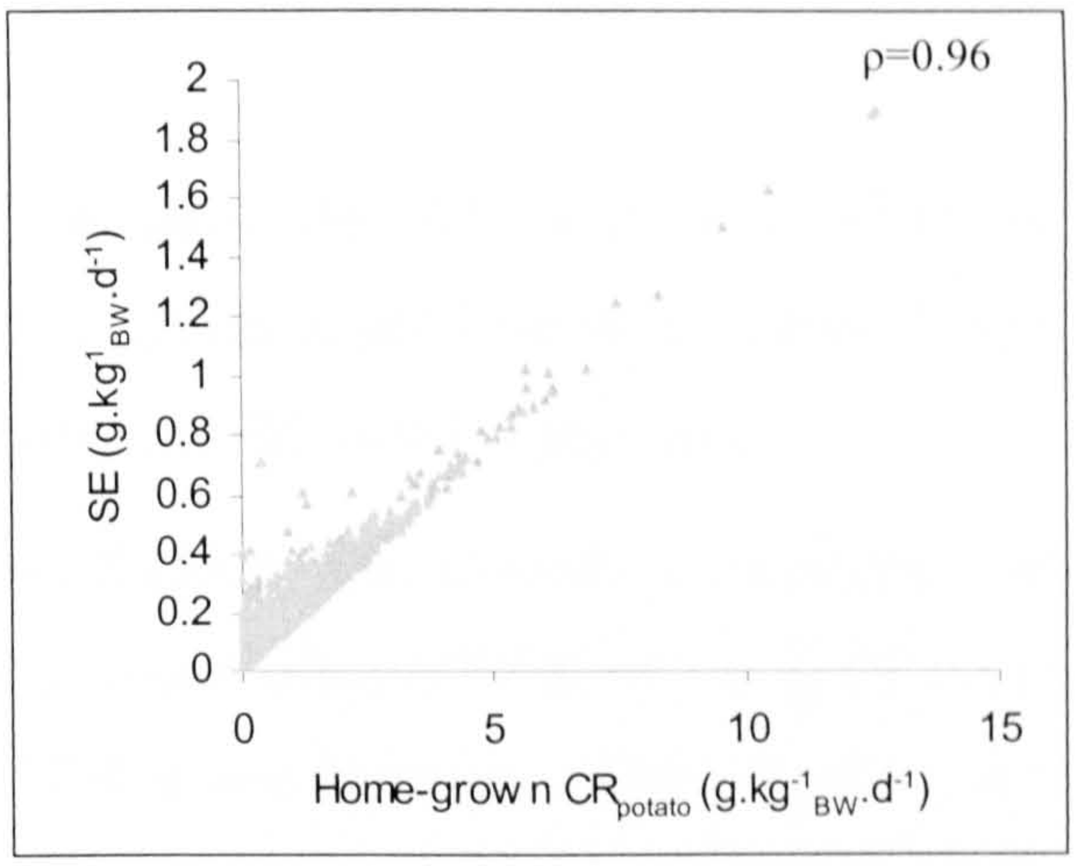


(c) 16+yrs

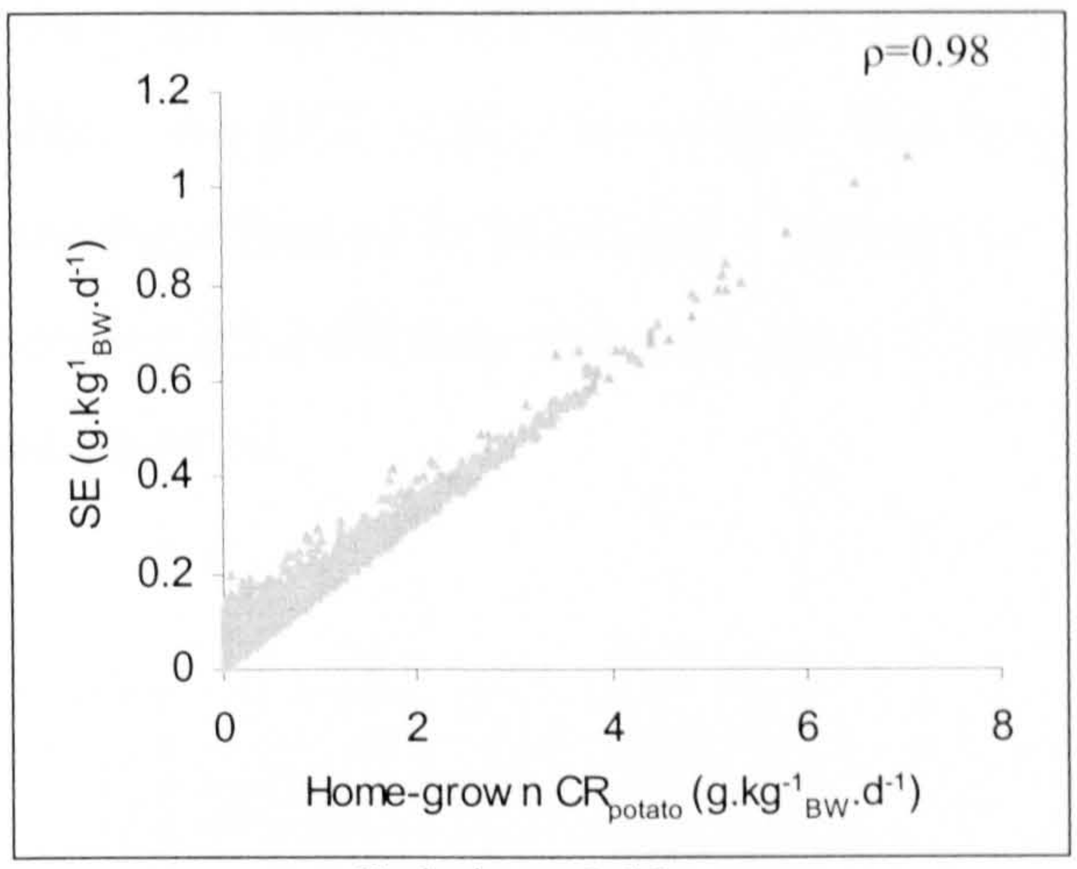


(f) 16+yrs

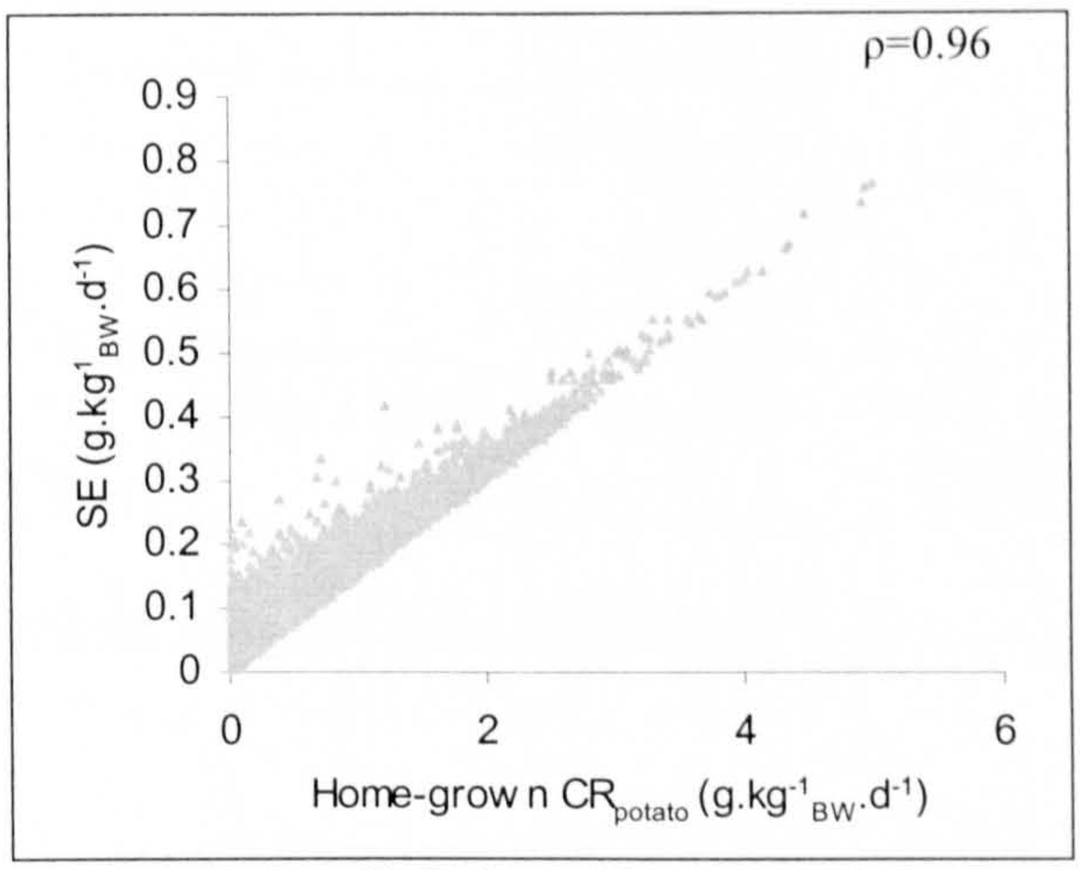
Figure 7.2. Scatter plot of *SDR/BW* versus *SE* for As (left) and home-grown potato consumption versus *SE* for As (right). Correlation coefficient, ρ, shown.



Cadmium; 0-5 yrs



Cadmium; 6-15yrs



Cadmium; 16+yrs

Figure 7.3. Scatter plot of home-grown potato consumption versus *SE* for Cd. Correlation coefficient, ρ, shown.

7.5 Distribution of SE

The plots in Figure 7.4 show the difference in distribution of *SE* for the same population, for the three contaminants of concern. Cd is shown at pH7 and pH6, since CF_{veg}, which is a component of *SE*, is pH dependent.

In the left hand plot, As *SE* distribution is much more peaked than that for Ni. The peak for As is due to the large contribution of *SDR*, which for people aged 6yrs and over is a constant. For Nickel, *SDR* is less important, with a larger contribution from the home-grown vegetable intake route, so the distribution is flatter and more varied than for As. The right hand plots show *SE* for Cd at two pHs. The difference is due to the pH dependent CF_{veg} variable. At pH7, CF_{veg} is smaller than at pH6, so the effect of multiplying by CF_{veg} has the effect of increasing the variance at pH6. Note the x-axis scales: Cd produces a much higher *SE* than either of As or Ni, again because CF_{veg} is so much bigger for Cd than As or Ni.

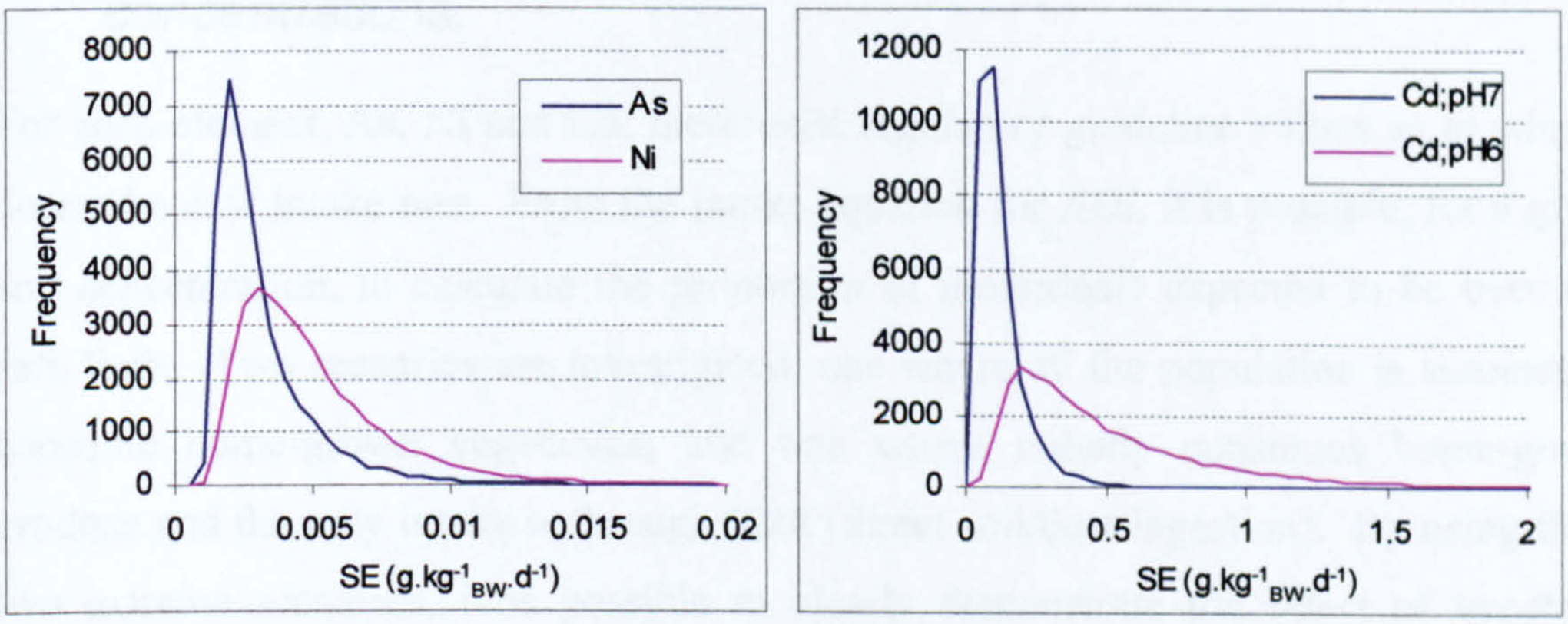


Figure 7.4. Distribution of *SE* for As and Ni (left) and Cd at pH7 and pH6 (right)

The *SE* distributions are standardised by their means, to give a relative measure of spread, depicted in Figure 7.5. The two standardised plots for Cd are practically the same, while Ni and As differ. The x-axis is truncated at $4 g.kg^{-1}_{BW}.d^{-1}$ to increase clarity of the plot. However, it should be noted that As has an extremely long tail (not shown), followed by Ni then by Cd: the skewness statistics are: As 17.63; Ni 9.81; Cd7 and Cd6 3.25. So, although the bulk of the As *SE*s fall in a tight band, it is prone to higher extremes. This can be explained by the high contribution of *SDR* which is a constant for ages over 5yrs (hence the tight peak) but highly variable for ages 0-5 (the long tail).

This effect is less for Ni, and less for Cd as the *SDR* contribution becomes less significant.

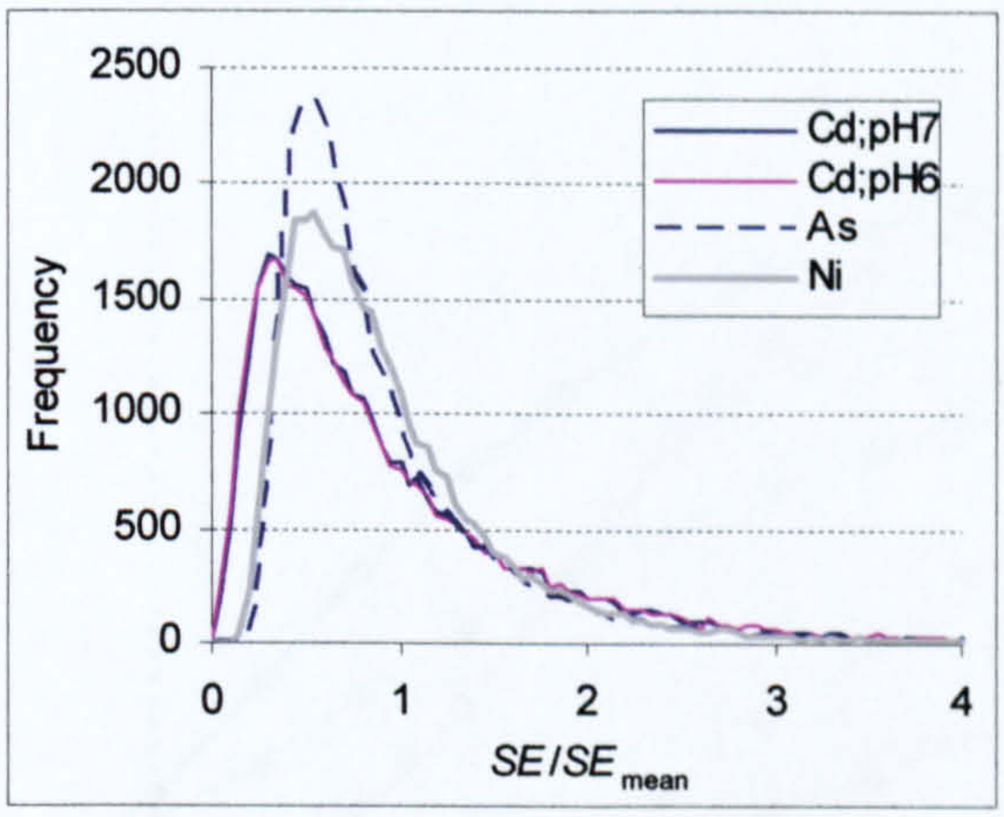


Figure 7.5. Distribution of standardised *SE* for Cd at pH6 and pH7, As and Ni.

7.6 Proportion of population over-exposed at different soil concentrations.

For each element, As, Ni and Cd, there exist regulatory guideline values as to what is deemed a safe intake rate. From the intake equation for *ADI*, it is possible, for a given soil concentration, to calculate the proportion of individuals expected to be over this safe limit. Two scenarios are investigated: one where all the population is assumed to consume home-grown vegetables; and one where nobody consumes home-grown produce and the only intake is through *SDR* (direct soil/dust ingestion). By using these two extreme scenarios, it is possible to clearly demonstrate the effect of vegetable consumption on the intake rate for the population. This analysis also highlights contribution of soil contaminant concentration to *ADI*. Again, the Gai population is used to demonstrate these points.

7.6.1 As

Arsenic is a non-threshold chemical, which means that any intake is potentially harmful, and the intake should be kept as low as is reasonably practicable. This low dose is known as the Index Dose (ID) and is given as $0.3\mu g.kg^{-1}_{BW}.d^{-1}$ in UK guidelines (DEFRA & EA, 2002i). An individual is over-exposed if $SE > ID / Csoil$, thus the proportion of this population expected to be over-exposed at different soil

concentrations is easily calculated. This chart is shown in Figure 7.6, along with an indication of the UK SGV of 20mg.kg^{-1} (DEFRA & EA, 2002i).

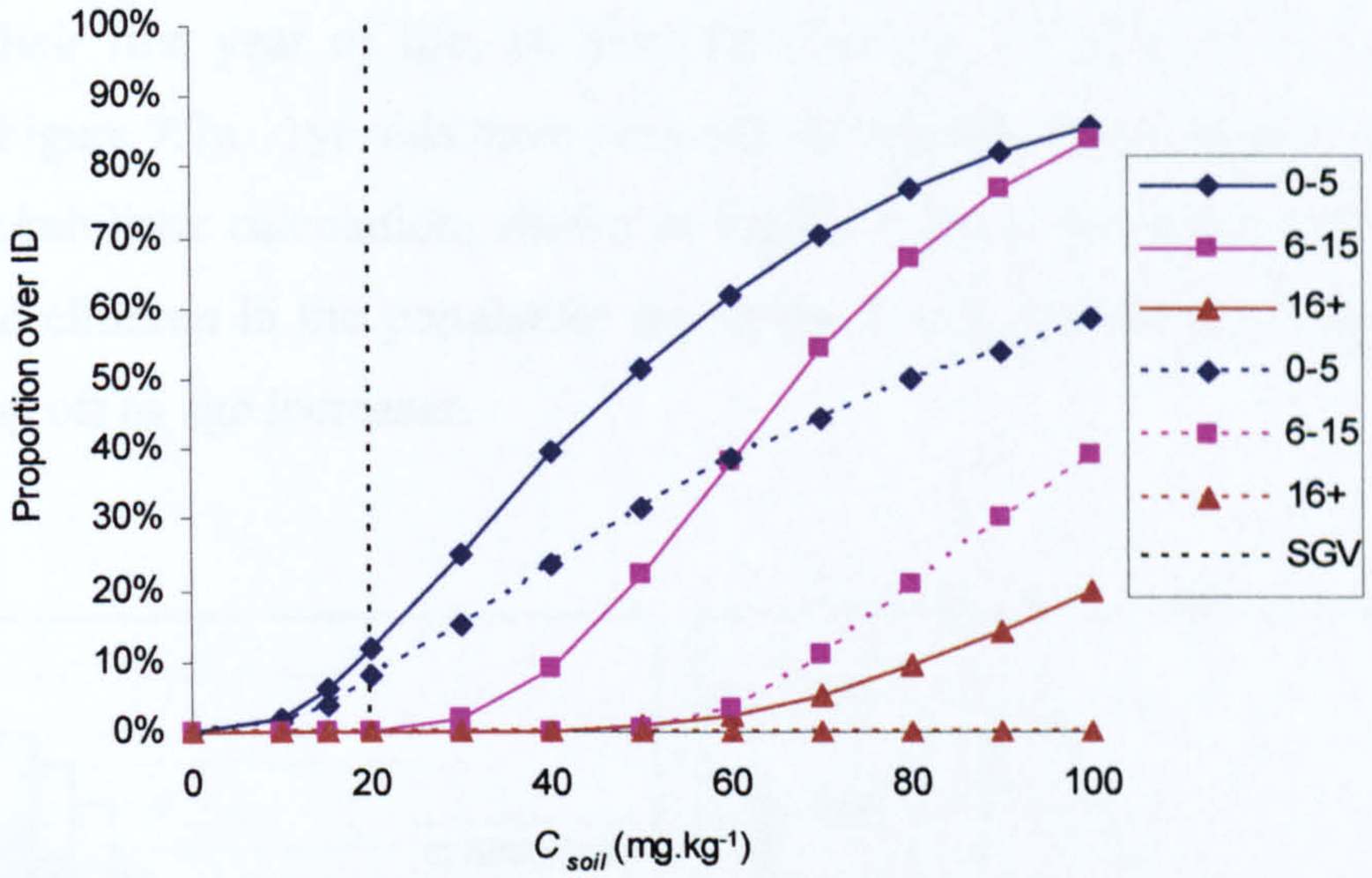


Figure 7.6. Proportion of the population whose intake is in excess of the Index Dose (ID) at different soil concentrations. Solid lines represent the scenario including home-grown vegetable consumption and dashed lines represent the scenario excluding home-grown vegetable intake. SGV is the Soil Guideline Value (see text).

Figure 7.6 shows the differences in intake for each age group: 0-5yr olds fare worse than 6-17yr olds and adults have the least intake, as expected. For both the 'no vegetable' and 'with vegetable' scenarios, at the SGV of 20mg.kg^{-1} the population would appear to be reasonably protected. None of the population aged 6yrs and above is over-exposed to As at this concentration, while approximately 8% and 12% of young children are over-exposed for the 'no' and 'with vegetable' scenarios respectively. One would expect no more than 5%, in as far as the CLEA methodology used to set the SGV is based on protecting a 6yr old critical receptor who falls at the upper 95th percentile of intakes. This apparent over-exposure of youngsters at the SGV is the result of the difference in averaging times used in this thesis and the CLEA methodology. CLEA works out the SGV based on the cumulative lifetime average exposure of a 6yr old critical receptor. This means that, for example, even if as a 2yr old a child has a very high exposure rate, the much lower rates experienced as a 5yr old may pull the lifetime average down below the critical level. Hence, working on lifetime averages, the peaks and troughs are averaged out. The method used in this thesis is for a slightly different scenario, where a 'snapshot' is taken of the exposures of individuals over a one year

period. In this case, the peak exposures at young ages are not averaged out with the lesser levels at older ages, hence the over-exposure level may appear slightly higher and account for the slightly higher than expected rates in the youngest age group.

This effect is illustrated in Figure 7.7. 0yr olds are only exposed for the second 6 months of their first year of life, so have low intakes, as reflected in the mean *SE* numbers in Figure 7.7a. 1yr olds have peak *SE*s which then drop off with age. Results from the probabilistic calculation, shown in Figure 7.7b, demonstrate that most of the over-exposed children in the population are in the 1 and 2yr old age bracket, and the numbers drop off as age increases.

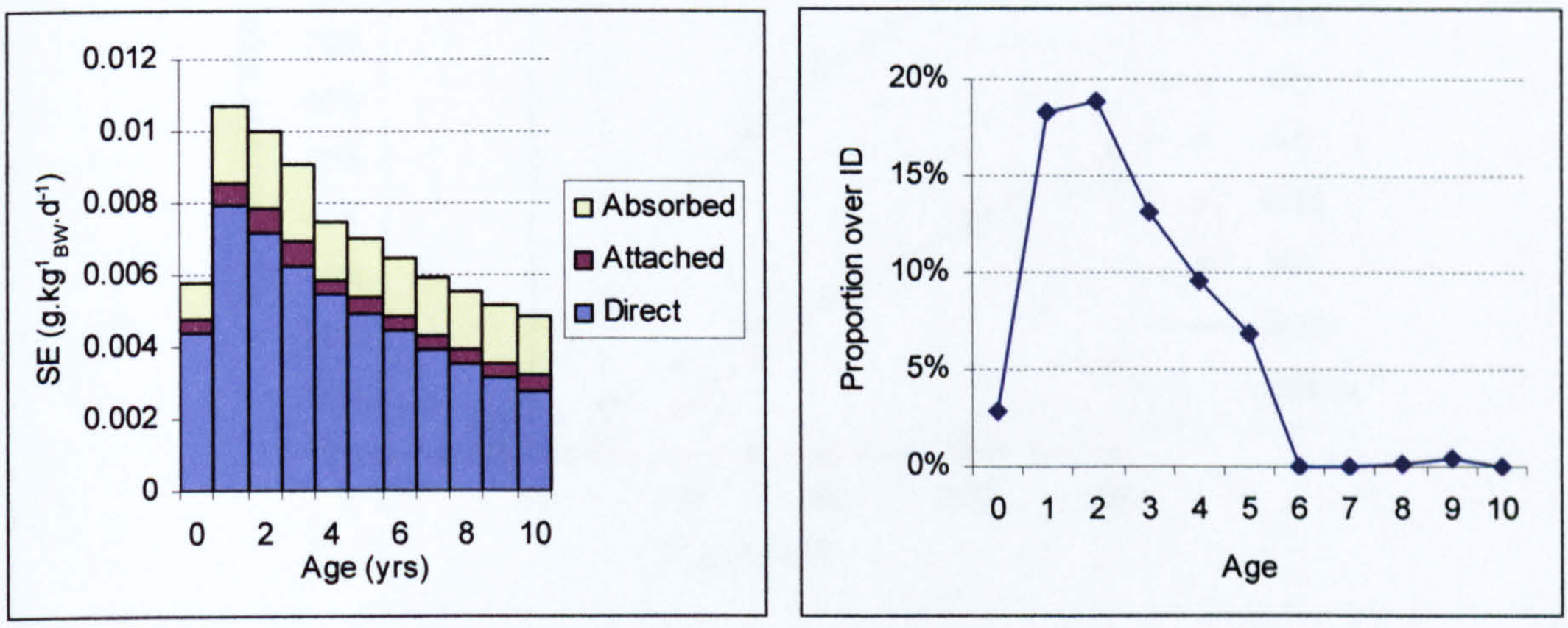


Figure 7.7. (a) Demonstration of how mean *SE* changes with age; (b) Proportion of the population exceeding the ID for As at the SGV of 20mg.kg^{-1}.

7.6.2 Ni

The same exercise is now carried out for Nickel, however the calculation is more complicated. Ni is a threshold chemical, whereby an intake may reach a certain level before it is deemed unsafe. However, the stipulated threshold level (Tolerable Daily Intake, TDI) includes intake from all sources, not just soil, so the background intake (mean daily intake, MDI) per kg_{BW} must be taken from the TDI to find the intake that is acceptable from soil (Tolerable Daily Soil Intake, TDSI).

The TDI is given as 5μg.kg$^{-1}_{BW}$.d^{-1} and the MDI for a 70kg adult as 160μg.d^{-1} (DEFRA & EA, 2002f). However, because adults have a larger MDI than younger age groups, a correction factor must be applied to reduce it for the younger ages. A different correction factor is applied for each age (DEFRA & EA, 2002b). In simple terms:

TDSI ($\mu g.kg^{-1}_{BW}.d^{-1}$) = TDI ($\mu g.kg^{-1}_{BW}.d^{-1}$) – MDI ($\mu g.d^{-1}$) × correction factor/BW (kg).

Finally, if the MDI/BW is calculated to be greater than or equal to 80% of the TDI, then the TDSI is set to 20% of the TDI (DEFRA & EA, 2002b).

Once the TDSI per individual is calculated, the same exercise as for As can be undertaken, whereby the proportion of the population which exceeds the TDSI can be plotted for different soil concentrations (Figure 7.8).

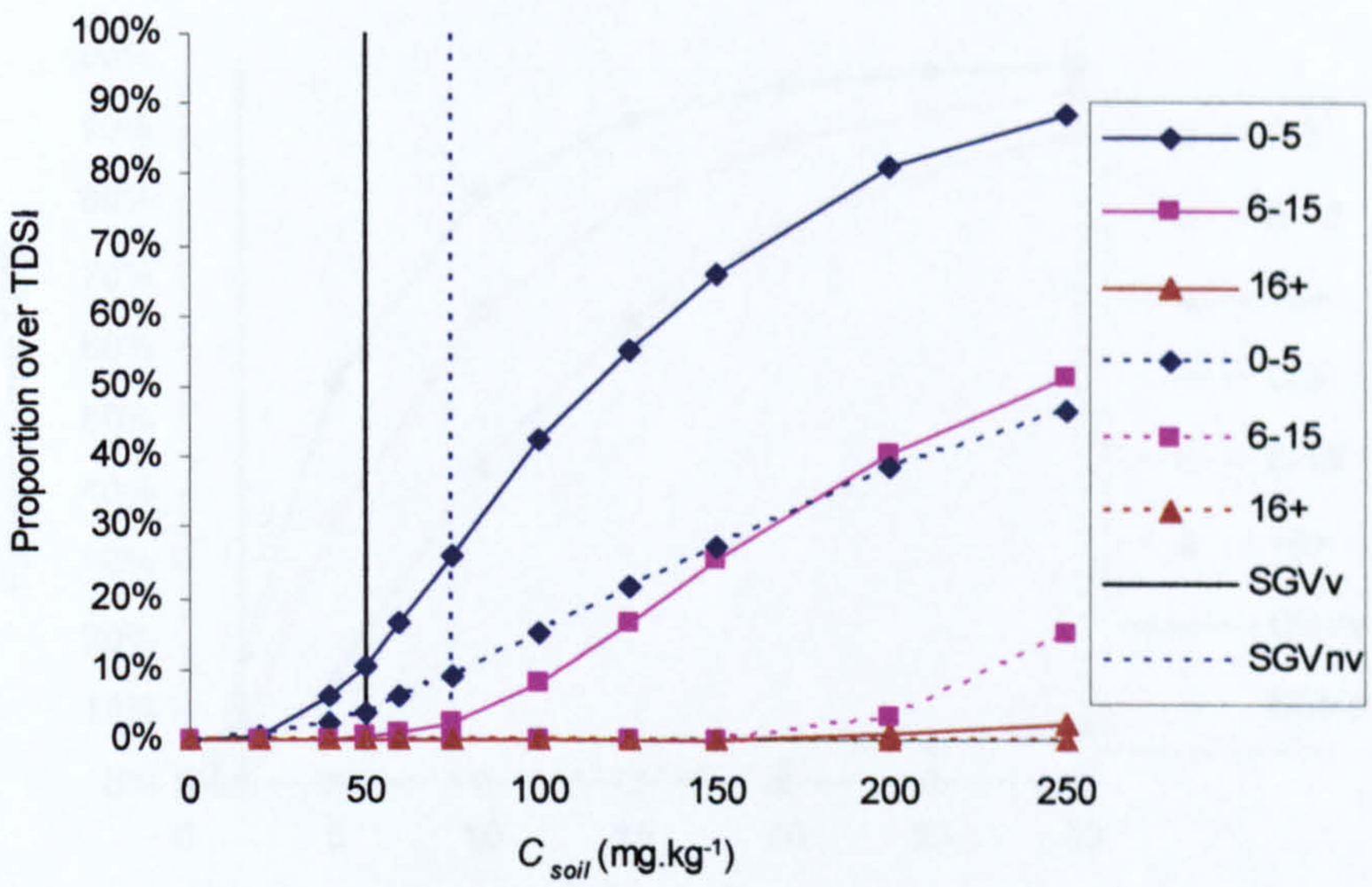


Figure 7.8. Proportion of the population whose intake is in excess of the Tolerable Daily Soil Intake (TDSI) at different soil concentrations. Solid lines represent the scenario including home-grown vegetable consumption and dashed lines represent the scenario excluding home-grown vegetable intake. SGVv is the Soil Guideline Value for the 'with vegetable' scenario, SGVnv is the Soil Guideline Value for the 'no vegetable' scenario (see text).

Two SGVs are given for Nickel (DEFRA & EA, 2002l), one for the 'with vegetable' scenario (SGVv), 50mg.kg^{-1}, and one for the 'no vegetable' scenario (SGVnv), 75mg.kg^{-1}. Similar to the As situation, the older two age groups are adequately protected. No individual exceeds the SGVnv in these two groups and only about 3% of 6-15yr olds exceed the SGVv. Again, the 0-5yr olds appear to be overly-exposed, there being approximately 9% and 10% over-exposure for the 'no vegetable' and 'with vegetable' scenario respectively. Again, this can be explained by methodological differences: SGVs being calculated on cumulative lifetime averages and the method used in this thesis using a snapshot of one year's exposure.

7.6.3 Cd

Cd is a threshold chemical, like Ni, and the TDSI is calculated in the same way. The TDI and MDI for Cd for a 70kg adult are given as $1\mu g.kg^{-1}_{BW}.d^{-1}$ and $16\mu g.d^{-1}$ respectively (DEFRA & EA, 2002d). Cd differs from Ni, however, since the CF_{veg} parameter is pH dependent, increasing as pH decreases. A chart for Cd has been calculated at both pH7 (SGVv $2mg.kg^{-1}$) and pH6 (SGVv $1mg.kg^{-1}$) (Figure 7.9a, b). Since SDR is independent of pH, pH does not affect the 'no vegetable' scenario which is depicted in Figure 7.9a only (SGVnv $30mg.kg^{-1}$).

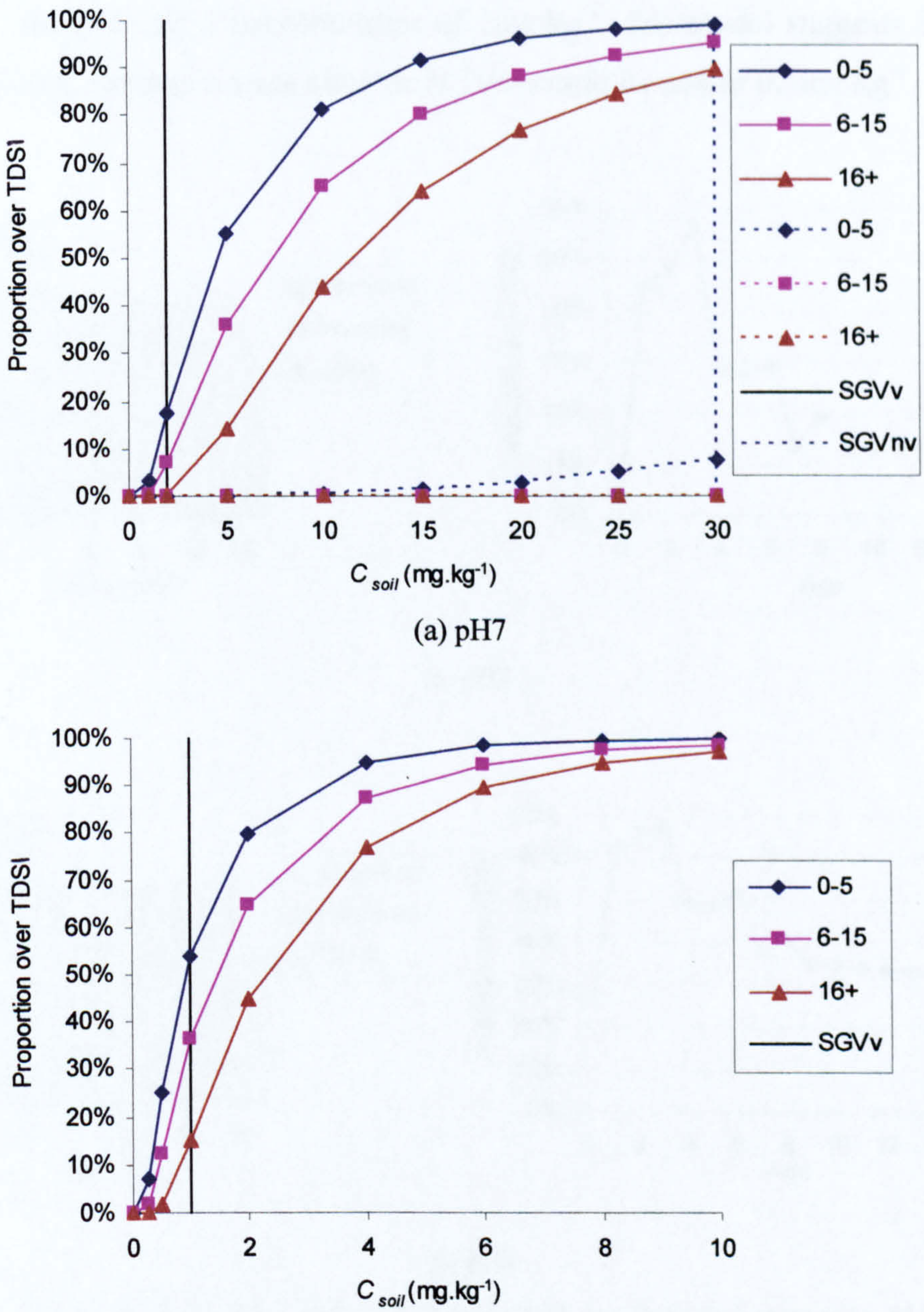
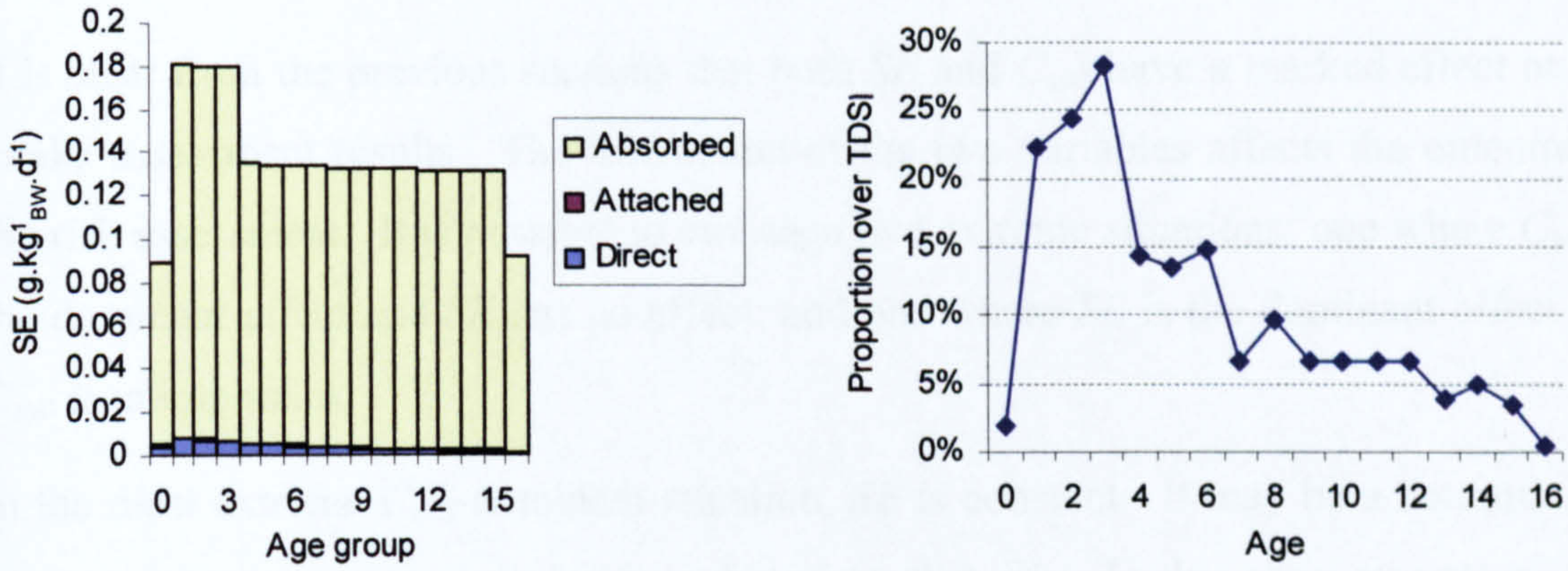
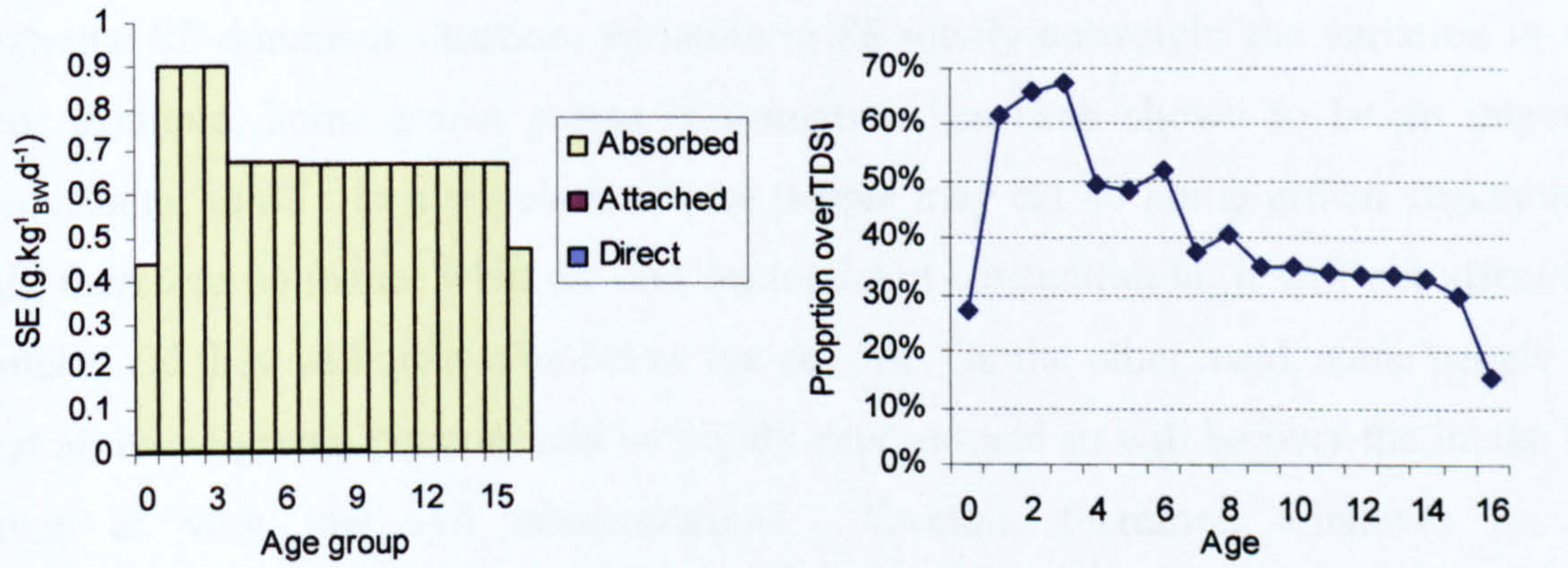


Figure 7.9. Proportion of the population whose intake is in excess of the Tolerable Daily Soil Intake (TDSI) at different soil concentrations: (a) at pH7 (b) at pH6. Solid lines represent the scenario including home-grown vegetable consumption and dashed lines represent the scenario excluding home-grown vegetable intake. SGVv and SGVnv are the Soil Guideline Values for the 'with vegetable' and 'no vegetable' scenarios respectively (see text).

For the 'no vegetable' scenario (Figure 7.9a), the situation is similar to the previous two elements. Only about 7% of the 0-5 year olds are over-exposed and this can again be explained by the difference in methodology used. For the 'with vegetable' scenario, the picture is a little different. At pH7, almost 18% of the 0-5yr olds have an intake in excess of the TDSI, along with 7% of 6-15yr olds. At pH6, the situation is worse, with almost 54% of 0-5yr olds, 36% of 5-16 yr olds and 15% of adults having intakes in excess of the TDSI. Figure 7.10 shows how mean *SE* changes with age, and the percentage of the population in excess of the TDSI at pH7 (a) and pH6 (b). At pH7, with a soil concentration of 2mg.kg^{-1}, lifetime averages will be acceptable on reaching adulthood. At pH6 and a concentration of 1mg.kg^{-1}, this model suggests intakes will still be high, and perhaps a more suitable SGVv would be nearer 0.5mg.kg^{-1}.



(a) pH7



(b) pH6

Figure 7.10. Demonstration of how mean *SE* changes with age (left) and proportion of the population exceeding the TDSI for Cd (right). (a) pH7, SGVv 2mg.kg^{-1}and SGVnv 30mg.kg-1 (b) pH6, SGVv 1mg.kg^{-1}.

The huge differences between the 'with vegetable' and 'no vegetable' scenario reflect the findings from the analysis of the means in Section 7.3. Here the contribution of direct soil ingestion was seen to be negligible for Cd, and depended heavily on home-grown vegetable consumption rate.

The effect of pH on the CF_{veg} variable is quite clear to see (Figure 7.9). At pH6, all age groups are suffering 95%+ over exposure at the 8mg.kg^{-1} contaminant level, whereas at pH7, excesses are only between 33% for adults and 74% for 0-5yr olds at the same contaminant level. This has important implications for the spatially-evaluated risk assessment: using variable pH values in space, along with variable Cd values could make a large difference to the estimated risk.

7.7 Relative importance of variability in SE

It is clear from the previous sections that both *SE* and C_{soil} have a marked effect on the intake assessment results. The interaction of the two variables affects the outcome of the risk assessment. It is possible to envisage two extreme situations: one where C_{soil} is the dominant effect and *SE* has no effect; and one where *SE* is the dominant effect and C_{soil} is unimportant.

In the most extreme C_{soil}-dominant situation, *SE* is constant. It may be a deterministic assessment using an 'average' person found on that site. In this case, as contaminant concentration rises, the population suddenly steps from all below the cut off to all above the cut off as a critical soil concentration is reached (Figure 7.11). Conversely, in the extreme *SE*-dominant situation, variation in *SE* totally outweighs the variation in C_{soil}. For example, home-grown potato consumption has been shown to be an important contributor to *SE*. In a population some people may eat no home-grown vegetables at all, therefore no matter what the soil contaminant concentration, it will not affect their intake and they will always be below the cut off. On the other hand, some people may eat all home-grown potatoes and be highly exposed and so will be over the intake limit even at very low soil concentrations. Overall, therefore, whatever the soil concentration, the proportion of population exposed does not change (Figure 7.11).

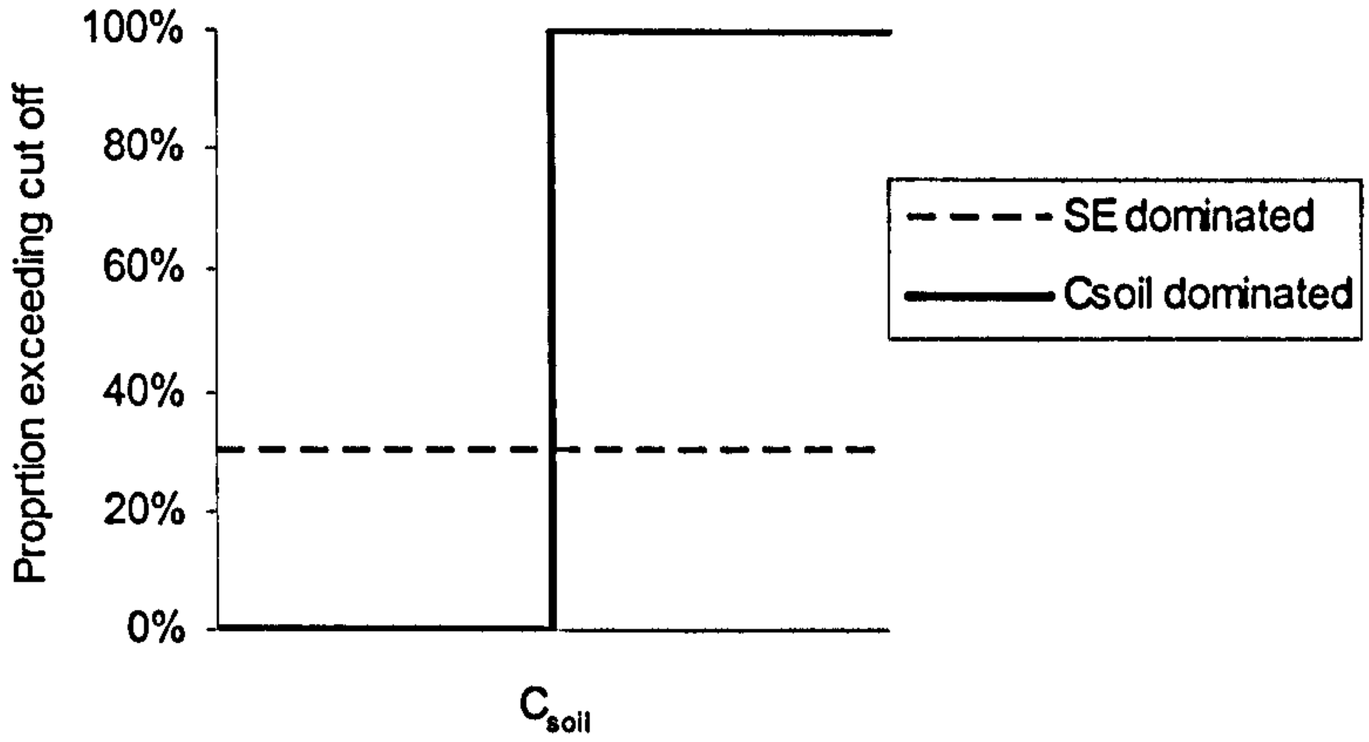
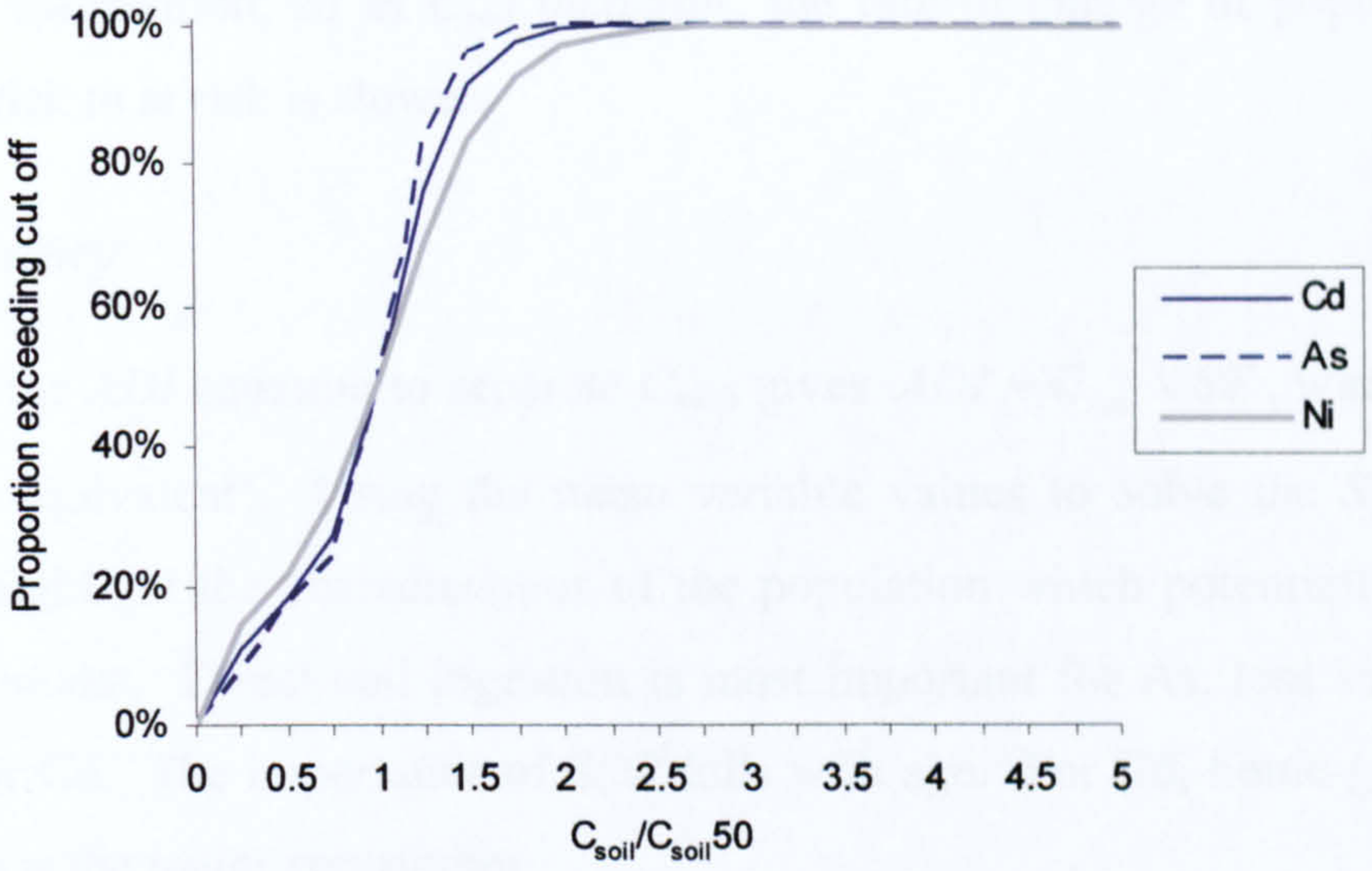


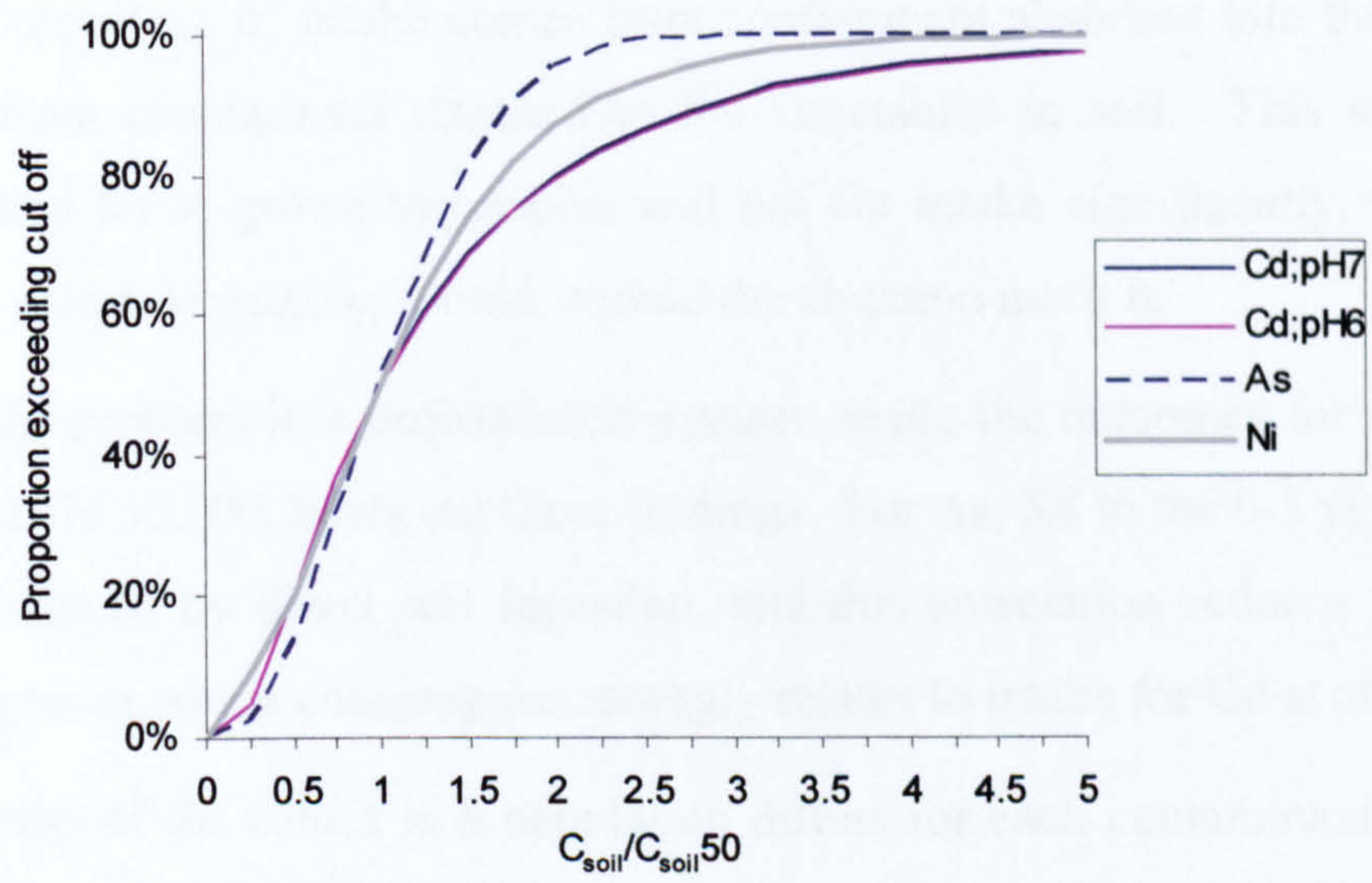
Figure 7.11. Interaction of C_{soil} and *SE*: two extreme situations.

In reality, the situation is likely to be between these two extremes. Figure 7.12 plots the curves of proportion of Gai population exposed at different soil concentrations for As, Ni and Cd for the scenario excluding home-grown vegetable consumption (a) and including home-grown vegetable consumption (b). The soil concentration is standardised by the concentration at which 50% of the population is over-exposed ($C_{soil}50$), so that the plots for the different contaminants can be compared easily. All plots show curves with a slope between the two extremes, suggesting that variation in both C_{soil} and *SE* is important in all cases. The intake assessments are less sensitive to *SE* (between-individual variation) in the 'no vegetable' scenarios (a), indicated by the steeper curves than in (b), the 'with vegetable' scenario. This seems sensible, since there are fewer variables in the 'no vegetable' scenario than the 'with vegetable' scenario.

As is the least *SE*-dependent of the three contaminants, having the steepest curve, regardless of whether or not home-grown vegetable consumption is included. In plot (a), excluding home-grown vegetable consumption, Ni is the most *SE*-dependent contaminant. In plot (b), including home grown vegetable consumption, Cd is the most *SE*-dependent at higher concentrations.



(a) Without home-grown vegetable consumption



(b) With home-grown vegetable consumption

Figure 7.12. Proportion of population exceeding the ID/TDSI at different soil concentrations. Plot standardised by the $C_{soil}50$ (soil concentration at which 50% of the population are expected to exceed the cut off).

This is related to the shape of the SE distributions (Section7.5). For a given population, the *SE* distribution changes according to the contaminant of concern. As, for example, produces a generally narrow distribution, with a few high extremes. This shape of *SE*s (i.e. the bulk of the *SE*s are very tightly distributed) means that a relatively small change in soil contaminant concentration could have a sudden large effect on the number of the population whose intakes are calculated to be in excess of regulatory safe levels. C_{soil} can increase slowly and the proportion of the population exceeding the ID will remain relatively static until the tight peak of the distribution hits the critical mark and the whole of the population suddenly spills over the cut off. Conversely, Ni, for example,

has a flatter distribution, so as C_{soil} increases, the rate of change of population from being not at risk to at risk is slower.

7.8 Summary

Rearranging the *ADI* equation to separate C_{soil}, gives $ADI = C_{soil} \times SE$, where *SE* is the '*SDR*.kg$^{-1}_{BW}$ equivalent'. Using the mean variable values to solve the *SE* term it is possible to highlight the characteristics of the population which potentially contribute the most to intake. Direct soil ingestion is most important for As, less so for Ni and negligible for Cd. The importance of *SDR* falls with age. For Cd, home-grown potato consumption is the major contributor.

In all cases, in terms of the contribution made by home-grown vegetable consumption, the highest percentage of intake comes from contaminant absorbed into the vegetables rather than from contaminant attached to the vegetables in soil. This suggests that simply washing home-grown vegetables will not cut intake significantly, whereas not eating home-grown vegetables would, should the situation merit it.

Looking at the problem in a probabilistic manner, using the outcomes for a population of approximately 35,000, bears out these findings. For As, *SE* in the 0-5 yr olds is most strongly influenced by direct soil ingestion, and this correlation reduces greatly with age. Home-grown potato consumption strongly relates to intake for Cd at all ages.

The distribution of *SE* values in a population differs for each contaminant of concern. The distribution is most skewed for As, with most of the *SE*s tightly distributed, but with some high extremes. Ni has a flatter distribution, with fewer extreme highs and Cd has the least extreme highs, but a slightly flatter more varied distribution over all. This reflects the make-up of *SE*, being most highly influenced by *SDR* for As, less so for Ni and virtually all vegetable intake in Cd.

Taking into account soil contaminant concentration, the proportion of population which exceeds guideline safe intake values increases as concentration increases. The increase is greater when consumption of home-grown produce is taken into account. Regulatory 'safe' soil concentrations, meant to protect the population from over exposure, appear to be adequate for As and Ni. They also appear to be adequate for Cd when there is no home-grown vegetable consumption, and for Cd at pH7 including home-grown vegetable consumption, where lifetime averages appear to be acceptable. When

including home-grown vegetable consumption at pH6, the SGV for Cd seems too lenient.

The variation in *SE* dictates the rate at which the population changes from below to above the ID/TDSI as C_{soil} increases. The less varied/more skewed is *SE*, the quicker the shift from 'safe' to 'unsafe' and vice versa.

Having now looked in detail at soil concentrations (Chapter 5) and between-individual variation (this chapter), the interaction between the two is investigated in the following chapter.

Chapter 8 Analysis of intake results

8.1 Introduction

It has become clear in previous chapters that variations in the component parts of the newly developed model are important and change in a complex manner from contaminant to contaminant and support size to support size. The interaction between these variable components is complicated and the effect on the results of the intake assessment can not be easily predicted. The aim of this chapter is to demonstrate the influence of different factors on intake through presentation of a varied selection of results. In addition, illustrations of the way in which the intake results may be presented so as to preserve the spatial dimension are given.

The newly developed spatially-evaluated intake assessment model has been run on different supports for each of four diverse example data sets, to produce intake predictions for the populations in the study areas. The four example data sets consist of varying numbers of exposed people and different elements of concern (both 'threshold' and 'non-threshold' contaminants).

First of all, general introductory information is given about the elements of concern (and why it might be appropriate to assess intakes of them) and the data sets. Then, as a means of introducing the complexities of the interactions of C_{soil}, SE and a regulatory threshold, a theoretical example is presented. Following this, actual results from the example data sets are presented to illustrate the important features affecting the results of the spatially-evaluated intake assessment. Examples are then given of how to interrogate the results and use the wealth of information available to provide useful map displays for use in risk management.

8.1.1 Contaminants of concern

It was useful to study contaminants with different characteristics, and the three examples used are As, Ni and Cd. As is a carcinogen, and therefore is a non-threshold chemical: any exposure to As should be kept as low as reasonably practicable. The ID (Index Dose) is a single number ($0.0003 mg.kg^{-1}_{BW}.d^{-1}$) and is applicable to all people in the population irrespective of age. Because of this, it is quite straightforward to compare calculated intakes with a single regulatory threshold and display the results in simple graphical form which is easy to interpret, e.g. a distribution of intakes with the ID marked as a reference point.

Ni and Cd are threshold chemicals and the derivation of what is thought to be a 'safe' intake is more complex. The TDI (Tolerable Daily Intake) includes intake from all sources, not just soil contamination, therefore to calculate the TDSI (Tolerable Daily Soil Intake) the background intake from other sources (e.g. drinking water, food) must be subtracted from the TDI. This MDI (Mean Daily Intake from sources other than soil) differs depending upon age, so each age class has a different MDI and hence TDSI. Therefore, when calculating the fraction of the population which has an intake in excess of the TDSI, the intake for each individual must be assessed against the relevant number for the age class. Cd intake estimates are further complicated by the fact that CF_{veg} (soil to plant concentration factor) is pH dependent. In this case, the distribution of pH values is just as important as that of Cd, as the combination of the two contributes to the intake.

In terms of health effects, there is evidence that prolonged low levels of exposure to ingested inorganic As can cause cardiovascular and dermatological problems. Thickening of the inner walls of the blood vessels leads to restricted blood flow and progressive loss of circulation to the extremities. Dermatological problems include skin lesions and skin cancer. It is possible that ingestion of As may also lead to other cancers such as that of the liver and bladder (DEFRA & EA, 2002c). Nickel ingestion may be the cause of dermatitis and, more seriously, is linked to reproductive problems including spontaneous abortion and malformations (DEFRA & EA, 2002f). Ingested Cd is toxic to humans, causing renal damage which in turn can affect vitamin D metabolism. This subsequently affects the calcium balance in the body and can cause weakening of the bones. Evidence suggests that Cd is only carcinogenic through the inhalation route, which is not applicable here (DEFRA & EA, 2002d).

8.1.2 Data sets

The Arsenic data set is from Stoke on Trent in the UK. Both the size of the area (13x15km) and the population (370,000) are very large. The Ni example comes from Gai, Russia, and covers a large area (14x14km) with a relatively small population (35,000) clumped into four settlements. Two Cd examples are used. The first (from Vasilevskaya, Russia) is similar to the Ni example, in that the area is quite large (8x4km) with a population clumped into four small settlements, however the population is very small (6,500). The second Cd example is from the same place as As (Stoke on Trent) so the region and population are both very large. Specific details are given later in the chapter as the data sets are employed.

8.2 Interaction of SE, Csoil and thresholds

Before presenting and analysing the results from the probabilistic, spatially-evaluated intake assessments it is worth demonstrating the effects of combining several factors which may contribute to them. Important aspects are: the distribution of soil concentrations, C_{soil} (depending on support size chosen, Chapter 5); the distribution of individuals' *SE*s (explored in Chapter 7); and use of regulatory cut off values. All three in combination will have varying effects on the assessment.

Interaction of variation in C_{soil} and *SE* has an impact on the results, making the results difficult to predict. Using regulatory thresholds also adds to the difficulties, since, depending on where in the distribution of results the threshold lies, the impacts of change of support (and hence change in C_{soil} distribution) and change in *SE* distribution will be different. A theoretical example is presented in Figure 8.1.

The aim is to calculate the proportion of the population in excess of a threshold. A single regulatory C_{soil} value is stipulated, which, for the sake of this exercise, corresponds to an ID/TDSI which is protective of 100% of a fixed *SE* population i.e. for a fixed *SE*, all the population is 'safe' at or below the SGV and all are over-exposed above the SGV. Two scenarios are presented whereby one set of soil concentrations is low relative to the SGV (left) and one where it is high (right). Soil concentration values are generated on three support sizes, corresponding to a small, large and site EU, which have the same mean but different variances. Each of these soil concentration distributions is then combined with each of three generated lognormal *SE* distributions all with the same mean; one with large variation (SE1), one with small variation (SE2)

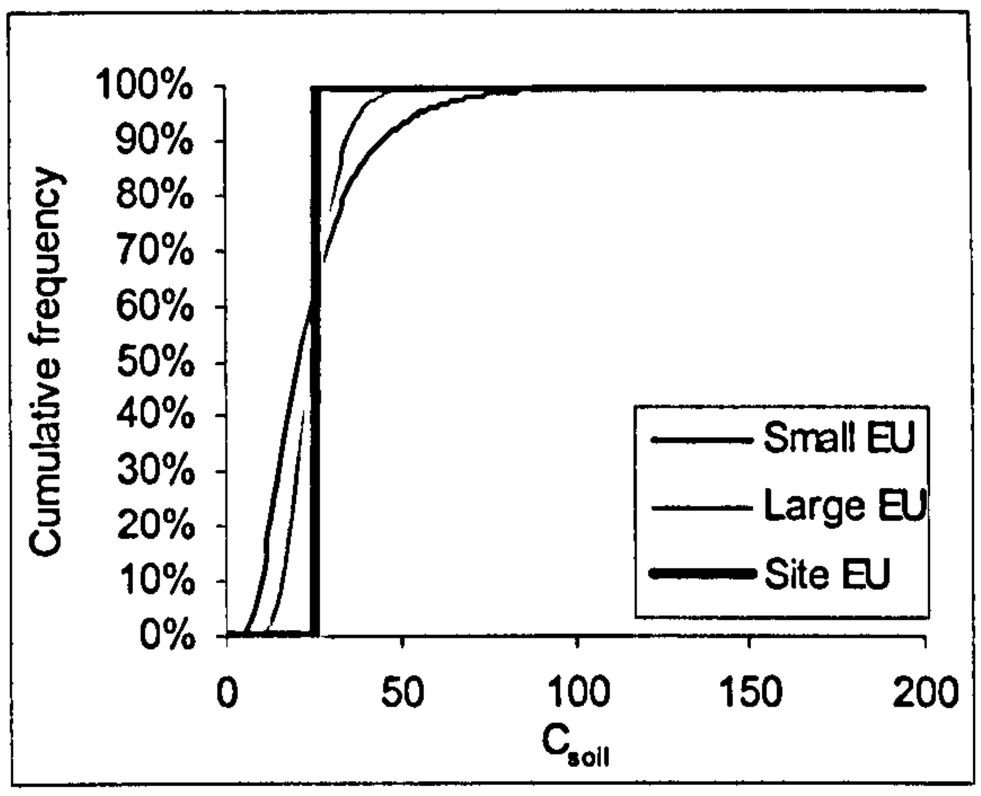
and one with no variation (SE3). For each scenario, nine sets of results are produced. Looking at the results tables presented in the Figure, it is possible to make some general comments about the sort of changes occurring under the different combinations of inputs.

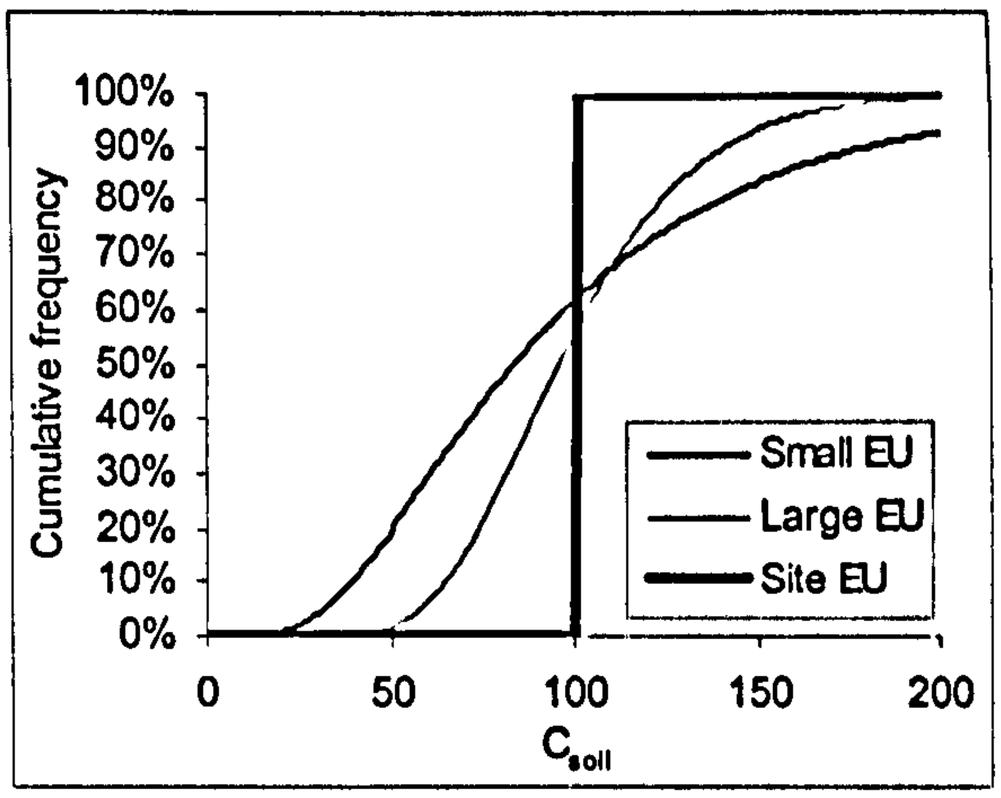
If there is no variation in either C_{soil} or *SE* (Site EU and SE3) then if C_{soil} is low relative to the SGV, none of the population exceeds the ID/TDSI whereas if C_{soil} is high relative to the SGV then all of the population exceeds the ID/TDSI. This is because the assessment is like a deterministic assessment, using an average soil concentration and an 'average' person. Since the cut off represents the C_{soil} at which a person changes from becoming 'safe' to 'unsafe', the person is either above or below the threshold.

With large *SE* variation (SE1), change in EU size (and hence distribution of C_{soil}) has little impact on the proportion in excess of the cut off in either scenario. However, with small *SE* variation (SE2), change in EU size (distribution of C_{soil}) has a bigger impact on the change in proportion in excess of the cut off.

For all *SE* distributions when mean C_{soil} > SGV, the proportion of population in excess of the cut off increases as the variation in C_{soil} decreases. This is directly attributable to the support effect averaging the lower values upwards so that more fall above the cut off.

When C_{soil} < SGV, with the lower *SE* variations, the proportion of population in excess of the cut off decreases as the variation in C_{soil} decreases. Again this can be directly attributed to support effect averaging the high values downwards so that they fall below the threshold. With a large *SE* variation, the opposite is found, and the proportion increases with decreasing soil variation. This could be because the wide range of *SE* values, combined with increased low end C_{soil} values moves the distribution of intakes upwards and over the threshold. Keeping a fixed EU size (and hence fixed C_{soil} distribution) when C_{soil} < SGV results in a decrease in the proportion as *SE* variation decreases, as the results become more focussed around a point below the cut off.

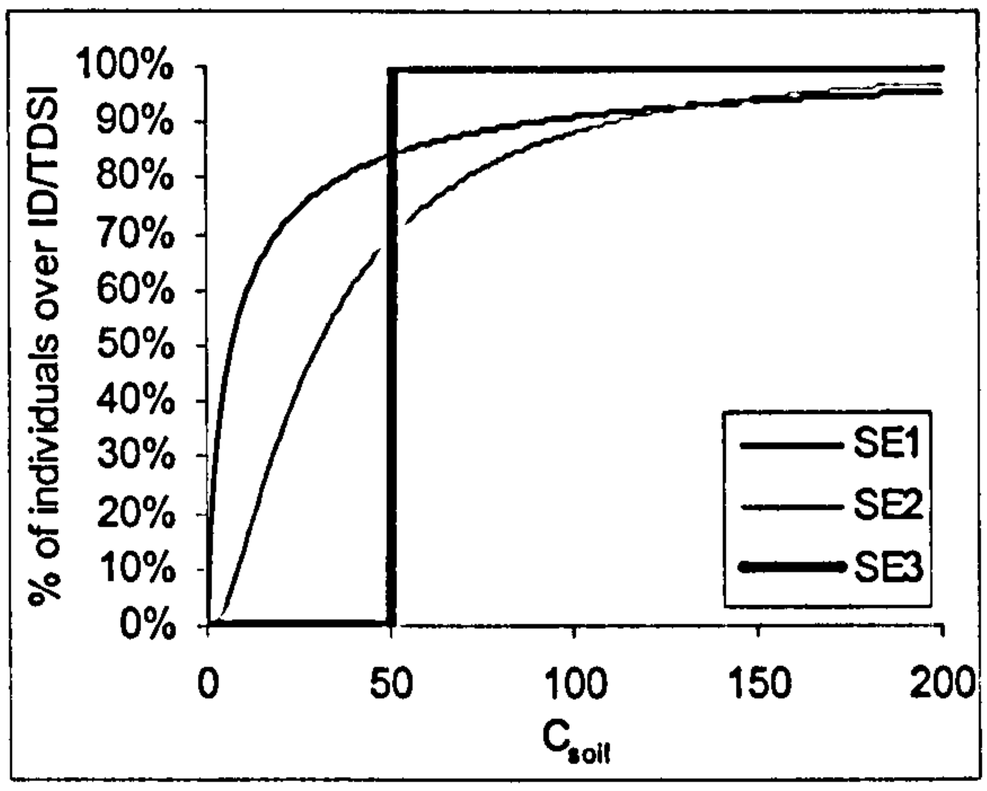




Cumulative distribution of C_{soil} at different EU sizes

Mean C_{soil} (25mg.kg^{-1}) < SGV (50mg.kg^{-1}) Mean C_{soil} (100mg.kg^{-1}) > SGV (50mg.kg^{-1})

COMBINED WITH:



Cumulative distribution of *SE*: SE1 large variation; SE2 small variation; and SE3 no variation

PRODUCES:

NINE SCENARIOS x2

Mean C_{soil} < SGV

Percent of population exceeding ID/TDSI

	SE1	SE2	SE3
Small EU	28.3	8.6	6.8
Large EU	30.4	2.1	0.6
Site EU	32.0	0.0	0.0

Mean C_{soil} > SGV

Percent of population exceeding ID/TDSI

	SE1	SE2	SE3
Small EU	88.4	80.5	79.5
Large EU	90.4	86.3	98.2
Site EU	91.1	88.4	100.0

Figure 8.1. Demonstration of how the proportion of the population classed as exceeding the ID/TDSI changes with the distribution of C_{soil} and *SE* and the relative position of mean C_{soil} to the SGV.

8.3 *Intake results: influence of different factors*

Intake calculations are made for example data sets with varying population sizes, spatial distributions and different elements of concern. Results are analysed and several examples are given here to demonstrate how different factors in the spatially-evaluated intake assessment can affect the results.

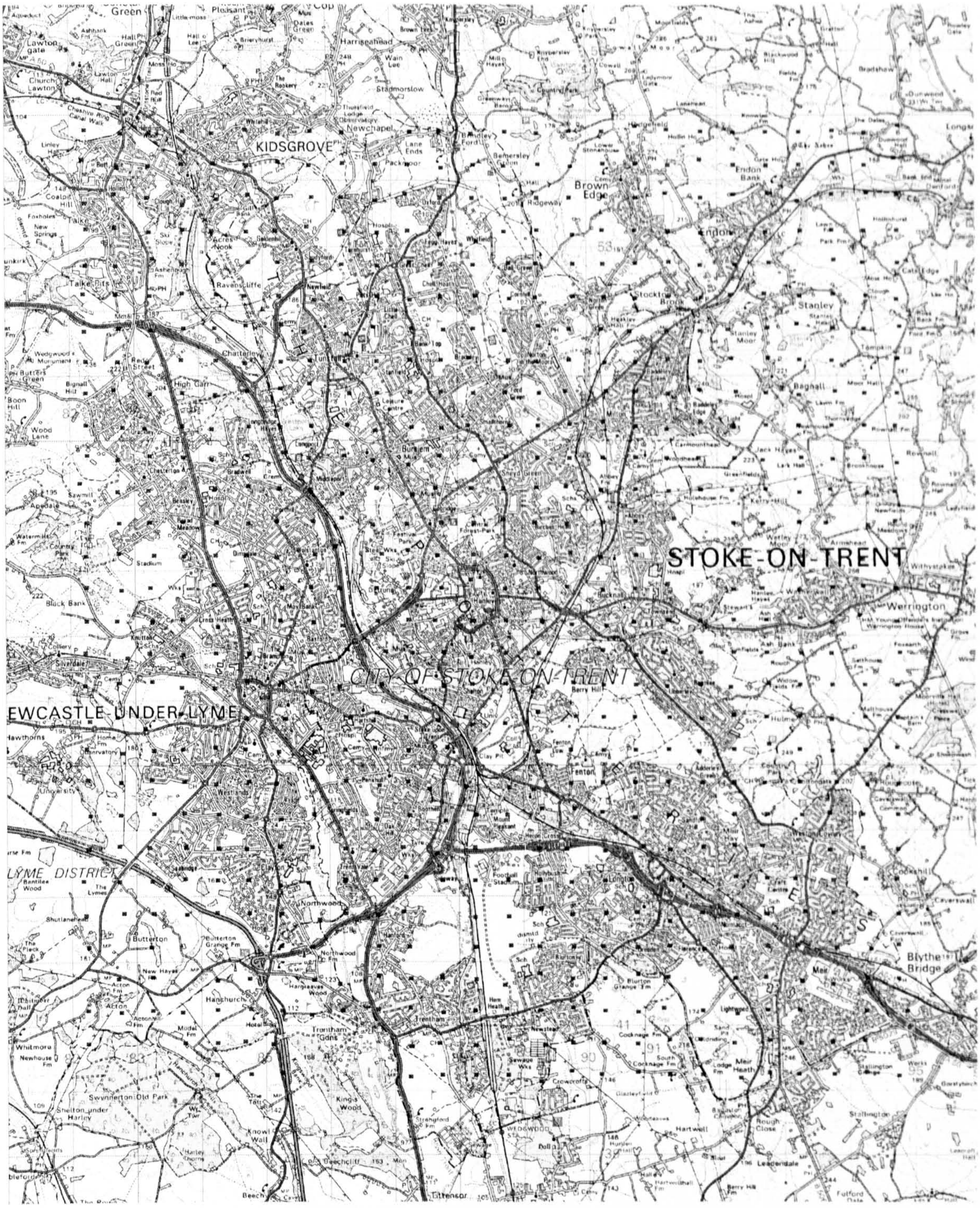
8.3.1 Change of Support

The decision to work at a certain support may affect the results of an intake assessment. Generally speaking, as was discussed in Chapters 3 and 5, as support increases the variation in C_{soil} is reduced because low and high values become averaged. To demonstrate the effect of this phenomenon, an example is given using a data set from Stoke-on-Trent in the UK, where exposure to As in the topsoil is investigated.

8.3.1.1 Background to the Stoke As data set

Stoke is the 15th most populous city in the UK, and lies in the North West Midlands of England in the county of Staffordshire. It has been a heavily industrialised area and is best known for its ceramics industry (and associated colour and glaze chemical works). It has also played host to extensive coal mining (it lies on the North Staffordshire Coalfield), iron and steel works, and electrical and engineering industries. Made-ground, including abundant ceramic and colliery waste, is widely distributed (British Geological Survey, 2000). Because of its industrial heritage, the urban soils of Stoke may contain contaminant levels which are a potential risk to human health.

Soil sampling in the urban area of Stoke on Trent was undertaken by the British Geological Survey (BGS) over 7 summer field seasons between 1988 and 1994 (British Geological Survey, 2000). Sampling was undertaken to correspond as closely as possible to a predefined 500x500m grid, equivalent to a density of four sites per square kilometre, covering an area of approximately 13x15km. Samples were taken from the A horizon, at about 15cm depth. Figure 8.2 shows the extent of the soil sampling in the Stoke on Trent region.



Background image © Crown Copyright Ordnance Survey. An EDINA Digimap/JISC supplied service.

Figure 8.2. Map of Stoke on Trent region showing soil sampling positions.

UK SGVs (soil guideline values) suggest levels of As in the soil of 20mg.kg^{-1} as the tolerable maximum limit. Simple statistics for As soil samples are given in Table 8.1 and a histogram of the sample distribution in Figure 8.3. 20% of the As sample values exceed the SGV. It is possible, therefore, that some of the population of Stoke will be over exposed to As.

Table 8.1. Statistics for topsoil (A horizon) As at a depth of circa 15cm (mg.kg^{-1}).

n	Mean	Std.dev	Min	P1	P5	P10	P25	P50	P75	P90	P95	P99	Max
745	16.15	10.12	2	5	8	9	11	14	18	25	31	60	136

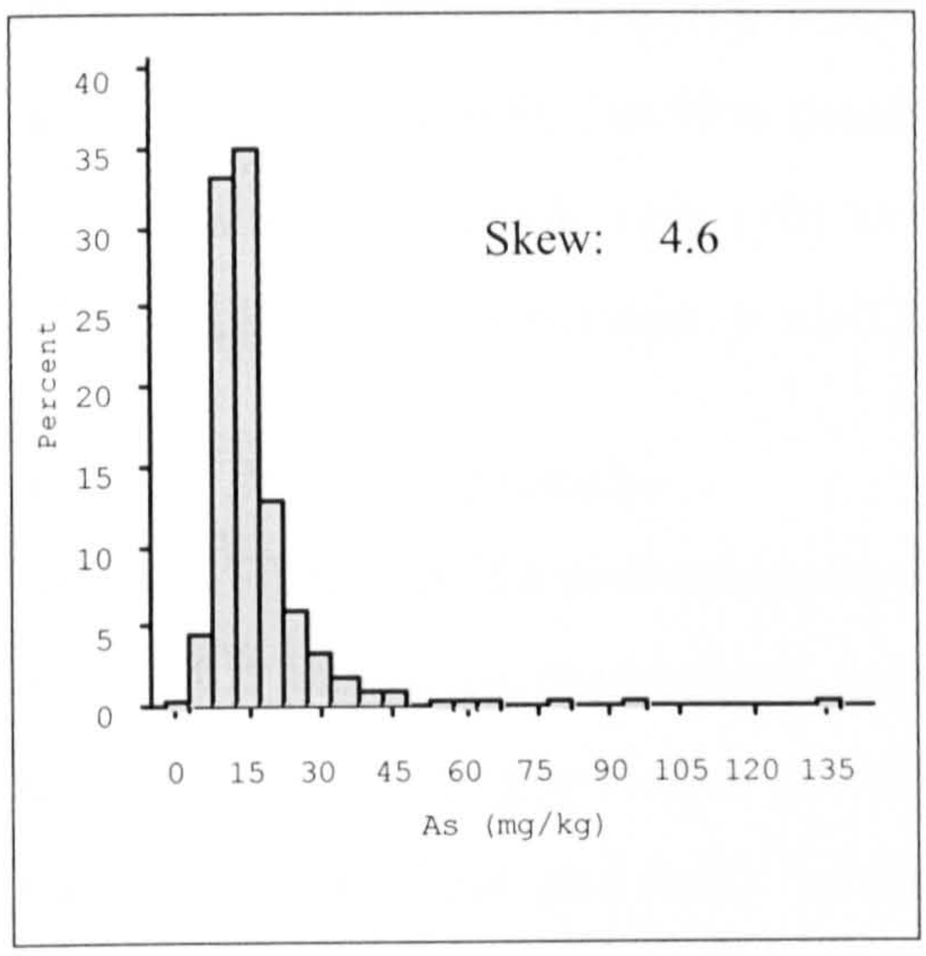


Figure 8.3. Histogram of distribution of As soil sample concentrations (mg.kg^{-1}).

The population breakdown by age class for the study area is given in Table 8.2 and totals some 370,440 people. The advanced method, as described in Chapter 6, was used to allocate the population to the smallest 'geostatistical EUs'. This involved using geographic information on positions of buildings and census data in ArcInfo™ GIS.

Table 8.2. Population numbers in the study area

Age(yrs)	Number
0-4	23,095
5-15	47,517
16+	299,828

Full details of the geostatistical procedure are given in Appendix A and only a brief summary is given here. 1,000 plausible soil concentration maps were produced using SIS (Deutsch & Journel, 1998) on a 50x50m square grid with 272 points in the East

West direction and 312 North South, making a total of 84,864 points. Since each 50x50m cell is so small in comparison to the whole region (13.6x15.6km), the point estimates are taken to be representative of the 50x50m block.

Minimum and maximum As concentrations were set to 10% less and 10% more than the minimum and maximum sample values respectively (i.e to 1.8 and 150mg.kg^{-1}), assuming that the true minimum and maximum values were unlikely to have been found in sampling. The 10% figure is somewhat arbitrary, but may mean more realistic simulations.

For change of support investigation, the 50x50m blocks (50m EUs) were upscaled to 200x200m blocks (200m EUs), whereby the soil concentration of the 200x200m block is the average of the 16 50x50m concentrations contained within. By treating the whole region as a single block, using the average As concentration for the whole region (Site EU), a non-spatial assessment is also produced.

8.3.1.2 Analysis of results

As an indication of the soil concentrations to which the population in Stoke are exposed, Figure 8.4 shows the median soil As concentrations after 1,000 simulations along with an indication of the population density across the city and surrounding area. It can be seen that on the 50m grid (left) the median soil concentrations are generally lower than when working on the 200m grid (right). This is because upward averaging of lower soil concentrations has a bigger effect on the medians than downward averaging of high values, causing a shift upwards in the medians on the 200m grid. Treating the whole region as a single block gives a median value of 15.96mg.kg^{-1} after 1,000 simulations.

In contrast Figure 8.5 shows the P95 values. Working on a 50m support (left) means that the population appear to be exposed to higher extremes than on the 200m support (right), where the extreme highs are averaged downwards. Treating the whole region as a single block leads to a P95 value of 16.11mg.kg^{-1} after 1,000 simulations.

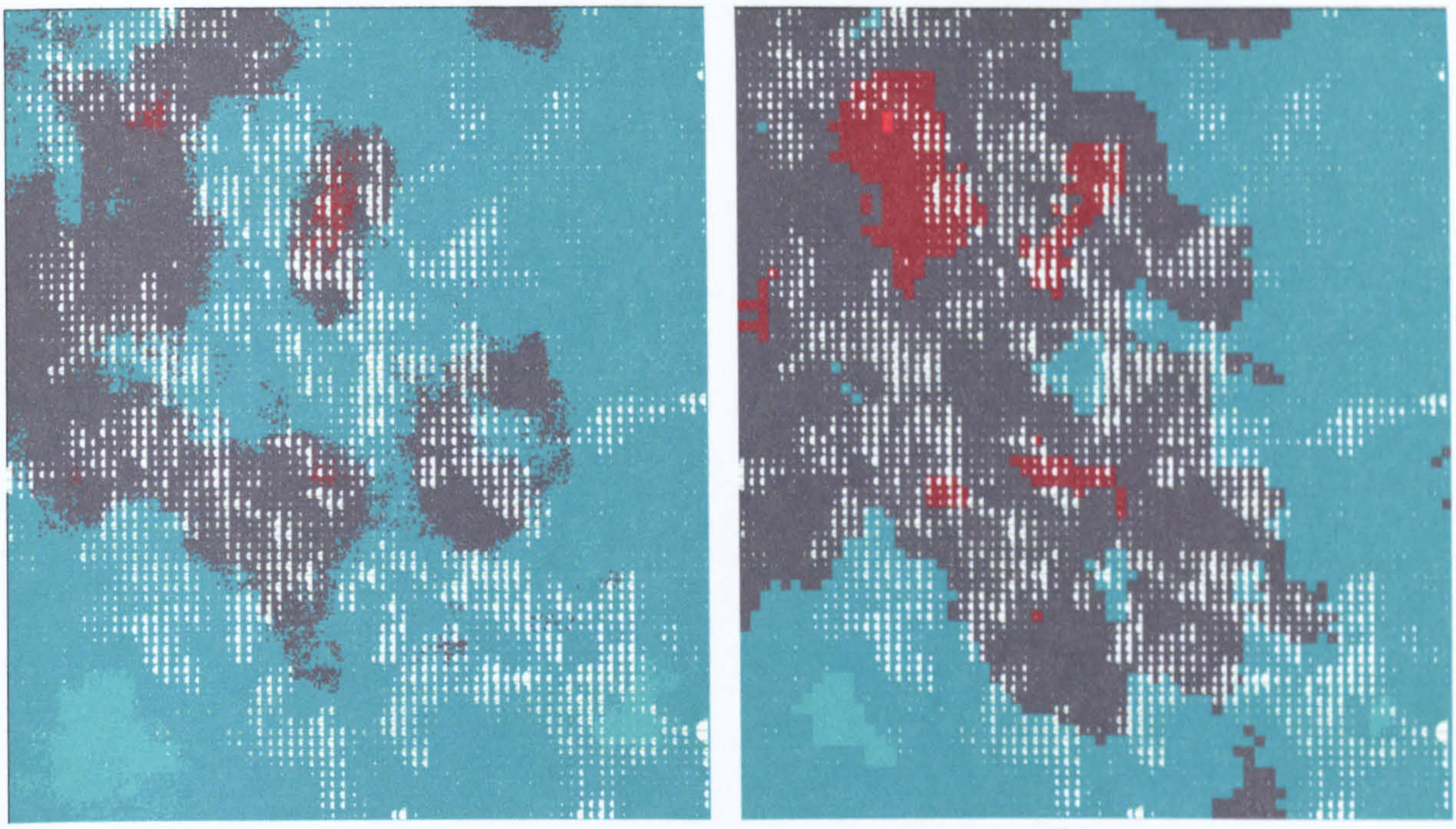


Figure 8.4. Map of median soil As concentrations in the Stoke region after 1,000 simulations on (a) 50m grid and (b) 200m grid. Pale symbols represent population density.

After running the intake algorithm, results are produced for the fraction of the population exposed per simulation, and these are shown in Figure 8.6. The left hand column shows the results for the scenario where the population is assumed not to eat home-grown vegetables and the right hand column shows the scenario including home grown vegetable consumption. In this case, for each scenario, the intakes increase as the support size decreases. This is a direct effect of the change in soil concentrations at the upper end of the distribution. Despite there being a higher median distribution of concentrations at the 200m support, the highest P95s are found at the 50m support, and it is the exposure of the receptors to these extreme highs which is tipping their intake estimates over the ID, thus leading to a higher proportion of the population being in excess of the ID. The increased variance of results at the smaller support is a reflection of the increased variance in soil concentrations to which the population is exposed.

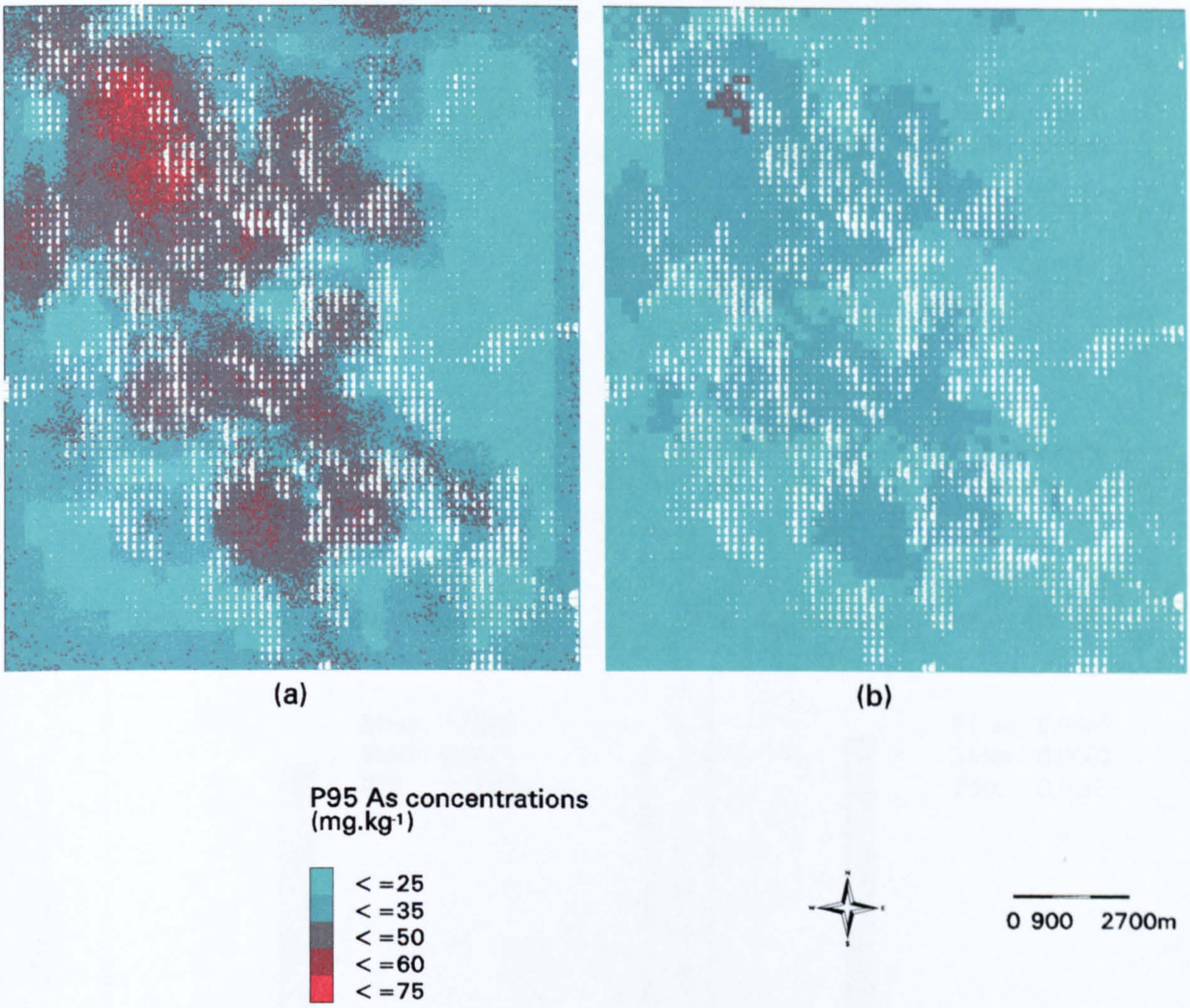
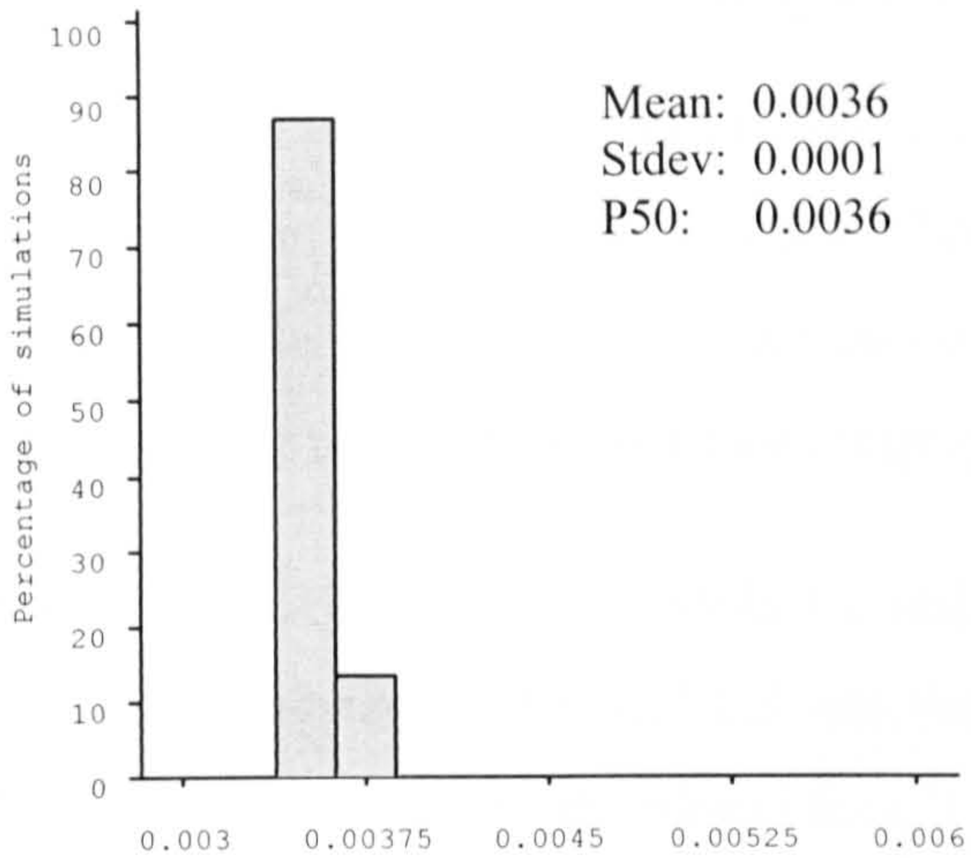


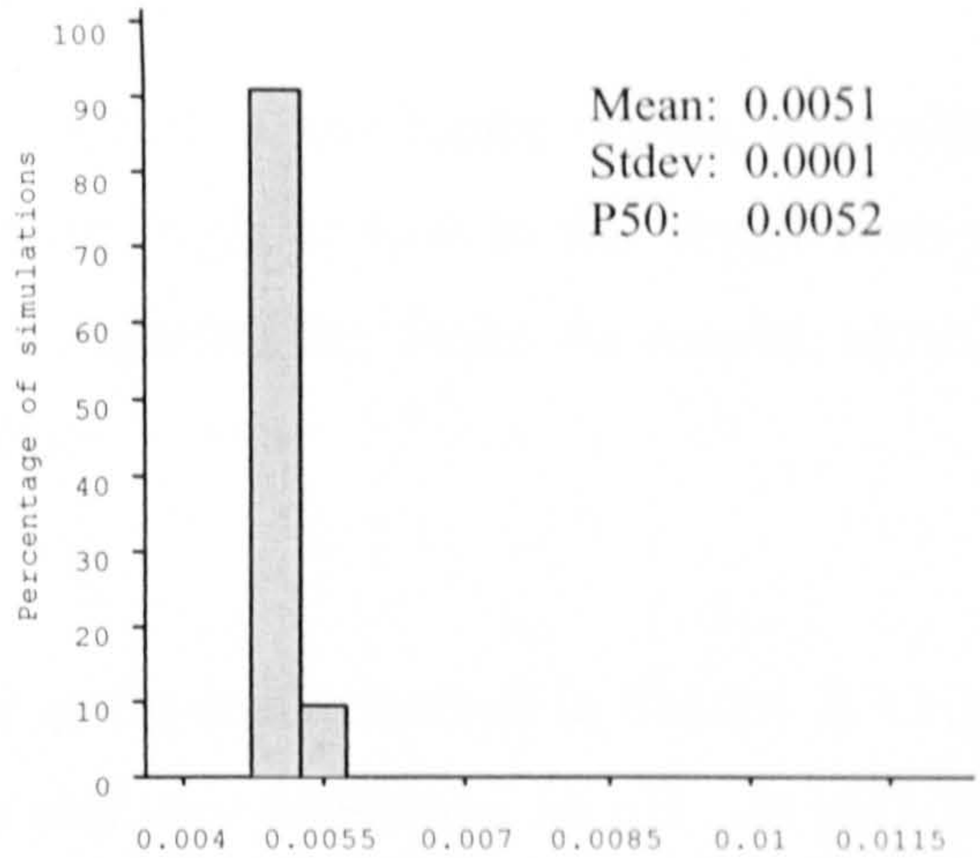
Figure 8.5. Map of P95 soil As concentrations in the Stoke region after 1,000 simulations on (a) 50m grid and (b) 200m grid. Pale symbols represent population density.

This result is as expected (see Section 8.2): the mean As value is less than the SGV, As gives a relatively low variance *SE* and so more people exceed the ID as C_{soil} variation increases (EU becomes smaller).

For an actual risk evaluation, the results are compared to the two dimensional decision criterion stipulated in Chapter 6 i.e. that in at least 50% of the simulations, no more than 5% of the population should have intakes in excess of the ID (Method P). In this case, the result is not affected by whether the assessment is made using a small support (50m EU), large support (200m EU) or using the site averages (Site EU). The intakes are so low that the maximum proportion of the population having an intake in excess of the ID is 1.2%.

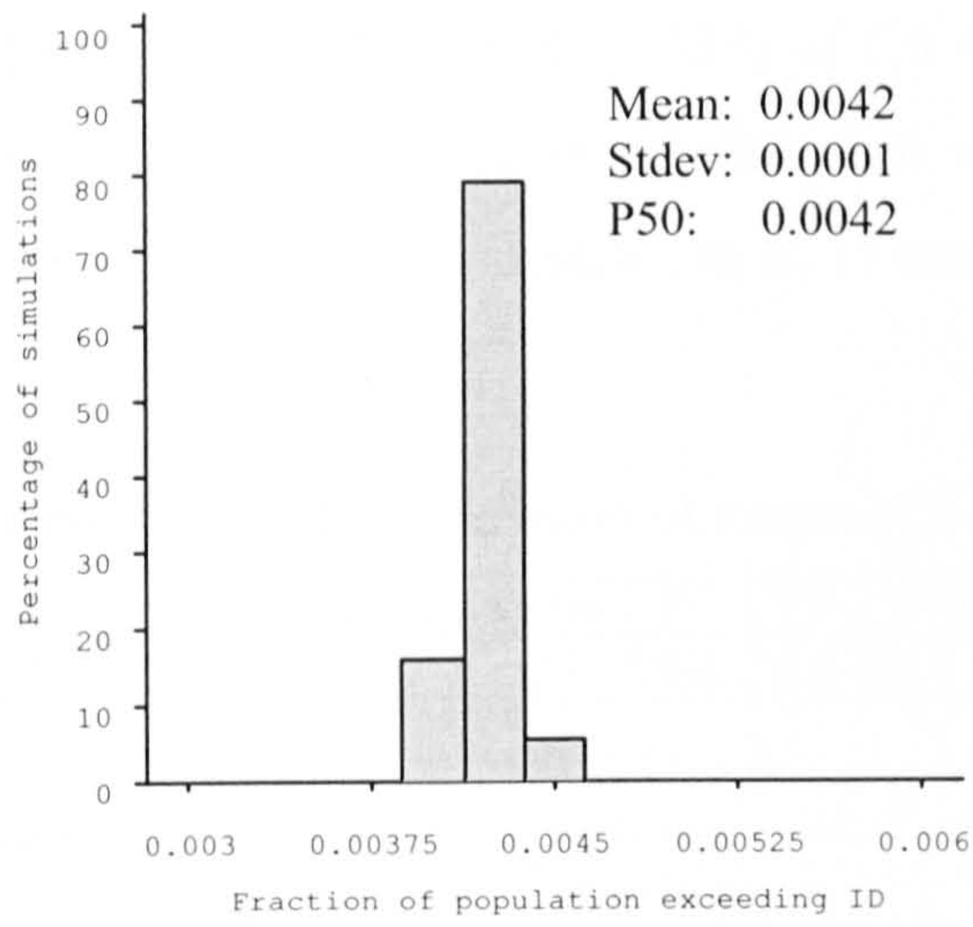
Clearly, there are instances where, with higher overall soil concentrations, working at different supports, and hence EU sizes, could affect whether the population was classed as at-risk or not. This highlights the importance of deciding on a support size at which to work if carrying out a spatially-evaluated assessment.

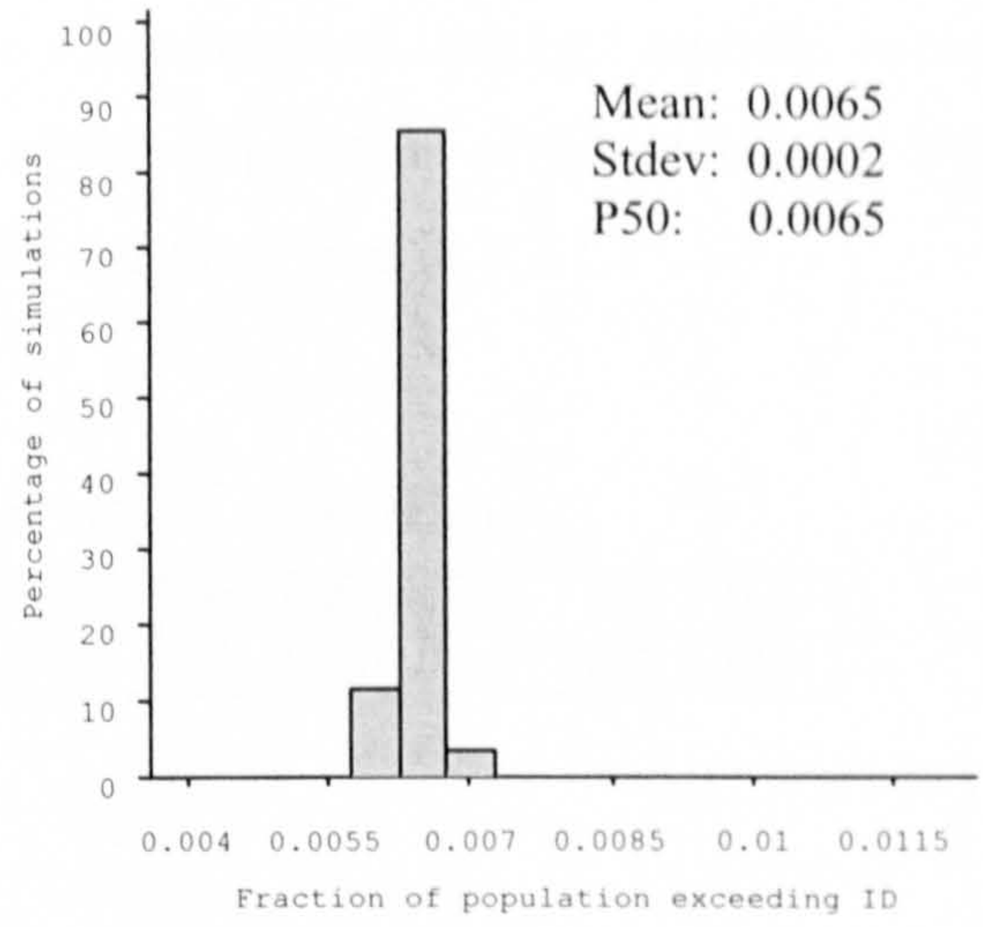




Without vegetable consumption; Site EU

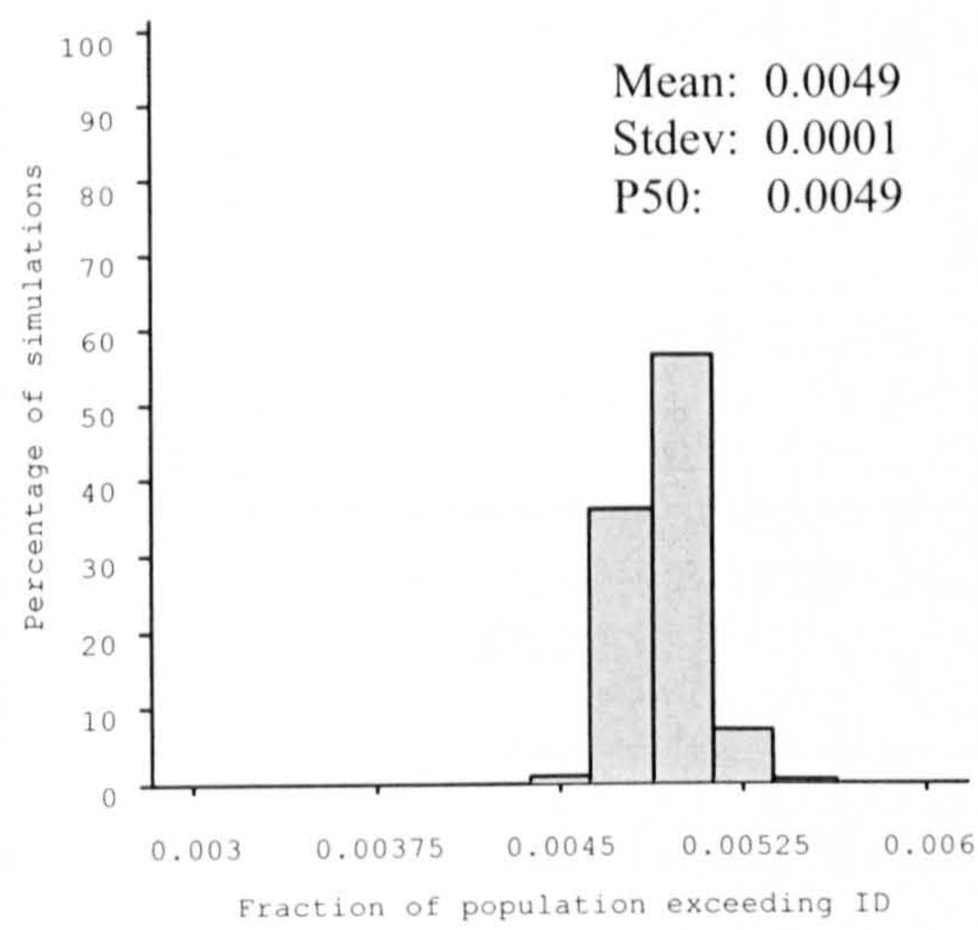
With vegetable consumption; Site EU

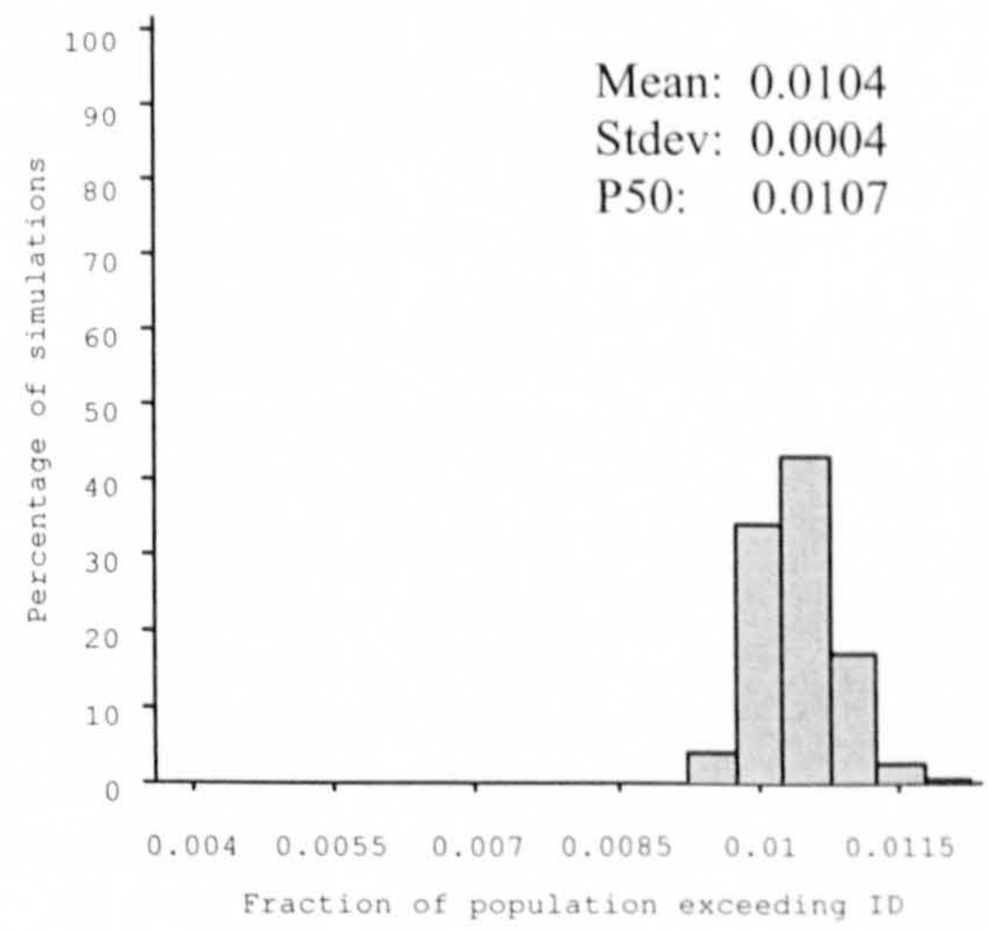




Without vegetable consumption; 200m EU

With vegetable consumption; 200m EU





Without vegetable consumption; 50m EU

With vegetable consumption; 50m EU

Figure 8.6. Distribution of results, from 1,000 simulations, of the estimated fraction of population exceeding the ID for As (0.0003mg.kg$^{-1}_{BW}$.d^{-1}). Left hand column; assuming no home-grown vegetable consumption. Right hand column; assuming home-grown vegetable consumption.

8.3.2 Where in the distribution the threshold lies

The effect of change of support has a different effect on the intake results, depending upon whether the regulatory intake threshold lies high or low in the distribution of intake results. This is demonstrated now by comparing the Stoke As results, already presented, to results from the same region for Cd.

8.3.2.1 Background to the Stoke Cd data set

The sampling regime for soil Cd was the same as for As (described in Section 8.3.1.1). UK SGVs (soil guideline values) for Cd in the soil vary depending on pH. At pH7, the SGV for a situation where home-grown vegetables are consumed is 2mg.kg^{-1}, while at pH6 it is reduced to 1mg.kg^{-1}. The regional average pH for Stoke is 5.4. Simple statistics for Cd soil samples are given in Table 8.3 and a histogram of the sample distribution in Figure 8.7. 53% of Cd sample values exceed 1mg.kg^{-1} and 4% exceed 2mg.kg^{-1}, therefore it is possible that some of the population of Stoke will be over exposed to Cd. The population is, of course, the one described in 8.3.1.1

Table 8.3. Statistics for topsoil (A horizon) Cd at a depth of circa 15cm (mg.kg^{-1}).

n	Mean	Std.dev	Min	P1	P5	P10	P25	P50	P75	P90	P95	P99	Max
744	1.7	2.0	0.4	0.4	0.4	1.0	1.0	2.0	2.0	2.0	2.0	5.5	43.0

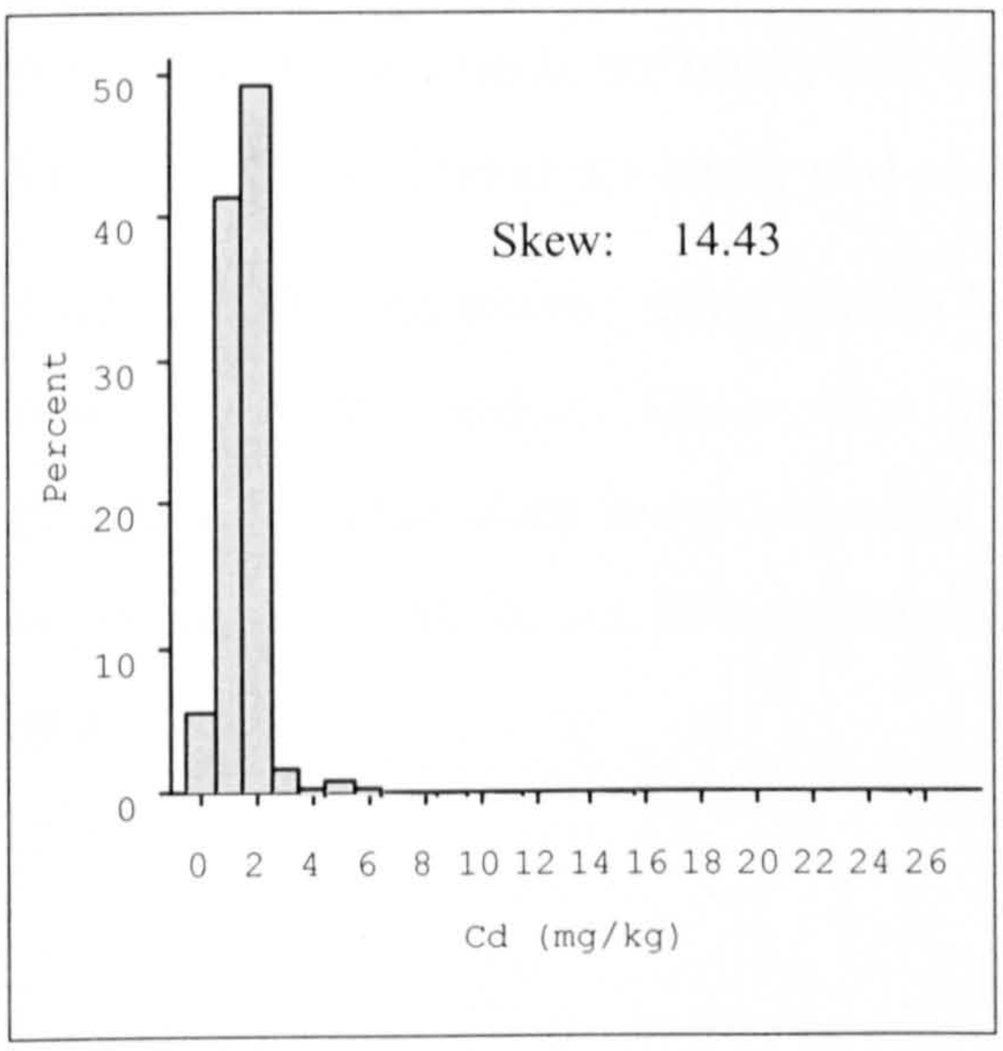


Figure 8.7. Histogram of distribution of Cd soil sample concentrations (mg.kg^{-1}). One value (43mg.kg^{-1}) not shown

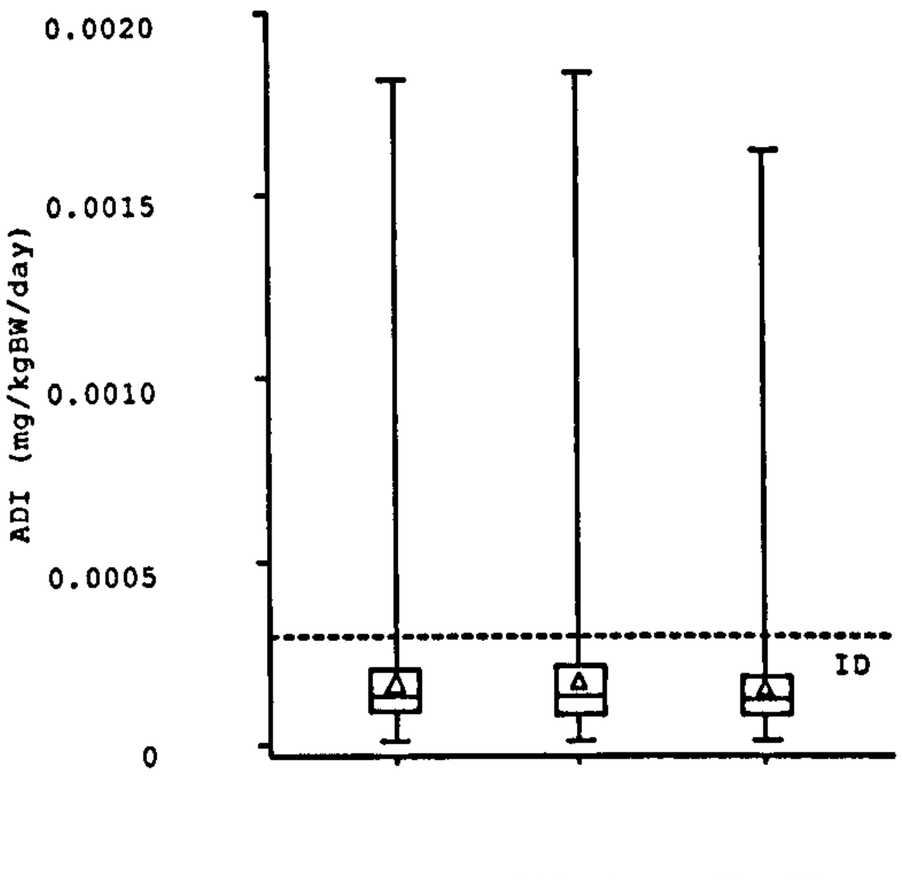
Full details of the geostatistical procedure are given in Appendix D. 1,000 plausible soil concentration maps were produced using SIS on the same supports and grids as for As, above. The minimum Cd concentration was set to 0mg.kg^{-1} (rationale explained in Appendix D) and the maximum to 10% more than the maximum sample value (i.e to 47mg.kg^{-1}), assuming that the true maximum value was unlikely to have been found in sampling. Note; as explained earlier, the new methodology has been designed to incorporate variable pH for Cd intake analyses, however, this example for Cd assumes a fixed pH (mean pH; 5.44), to ensure a clearer illustration of the threshold effect.

The ID for As is a single number (0.0003mg.kg$^{-1}_{BW}$.d^{-1}), so it is easy to display the *ADI* distribution for the population and show where the ID falls in this distribution. The TDSI for Cd varies according to age, so it is not as easy to display the *ADI* distribution with TDSI as a reference point. For this exercise, 1yr olds and 16yr olds are used as examples, so that the position of the TDSI in the *ADI* distribution for Cd can be displayed easily. (The median bodyweight for the age group is used in the TDSI calculation). The scenario including consumption of home-grown vegetables is used.

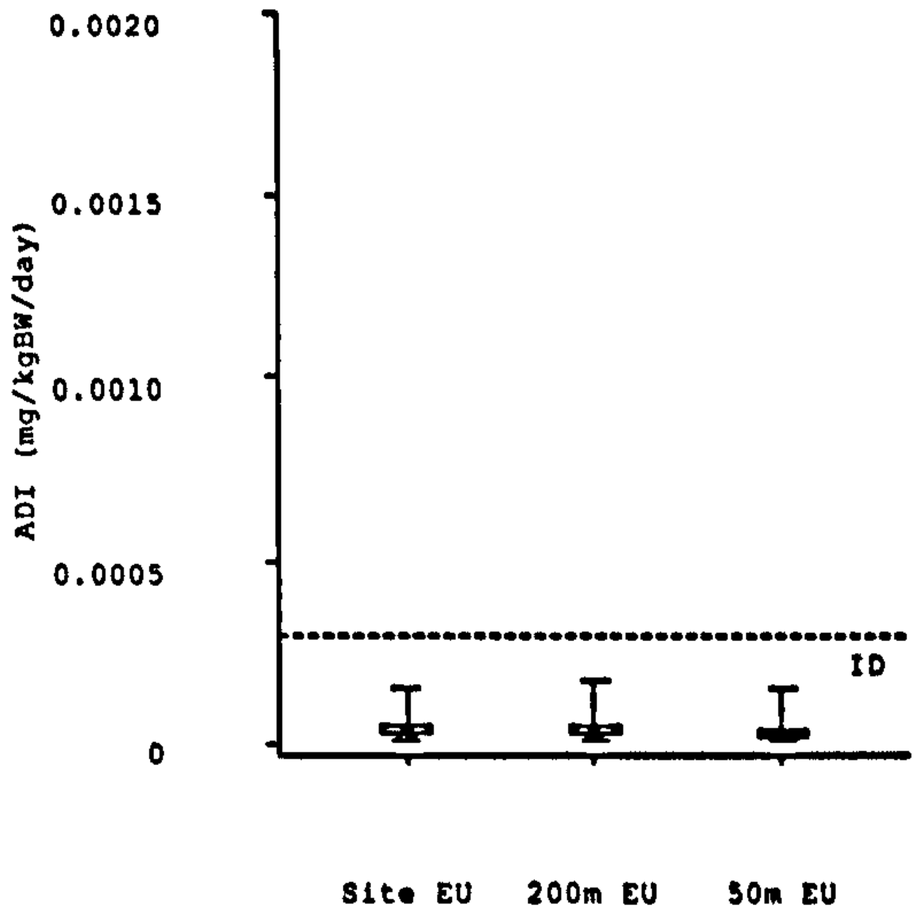
8.3.2.2 Analysis of results

The first case, As, has a mean soil sample contaminant value (16.5mg.kg^{-1}) below the SGV (20mg.kg^{-1}), and the second, Cd, has a mean soil sample contaminant concentration (1.67mg.kg^{-1}) above the SGV (1mg.kg^{-1} at pH6). It is likely, then, that the ID for As will lie at the upper end of the distribution of intake results for As and the TDSI for Cd will lie at the lower end of the distribution of intake results for Cd.

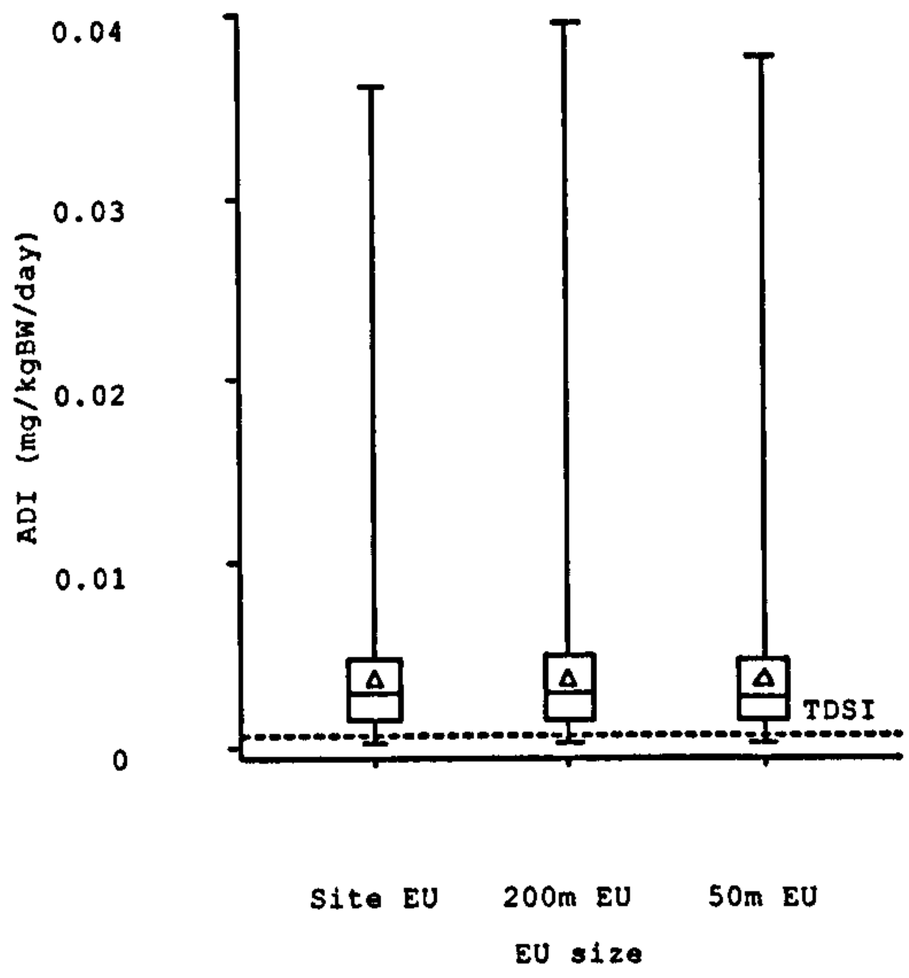
After 1,000 simulations, each person has a distribution of 1,000 *ADI* estimates. A summary of the median *ADI* results for all 1yr olds and all 16yr olds is displayed in Figure 8.8. This does indeed confirm that the regulatory cut-offs fall in the region expected. The ID for As falls at the upper end and the TDSI for Cd falls at the lower end.



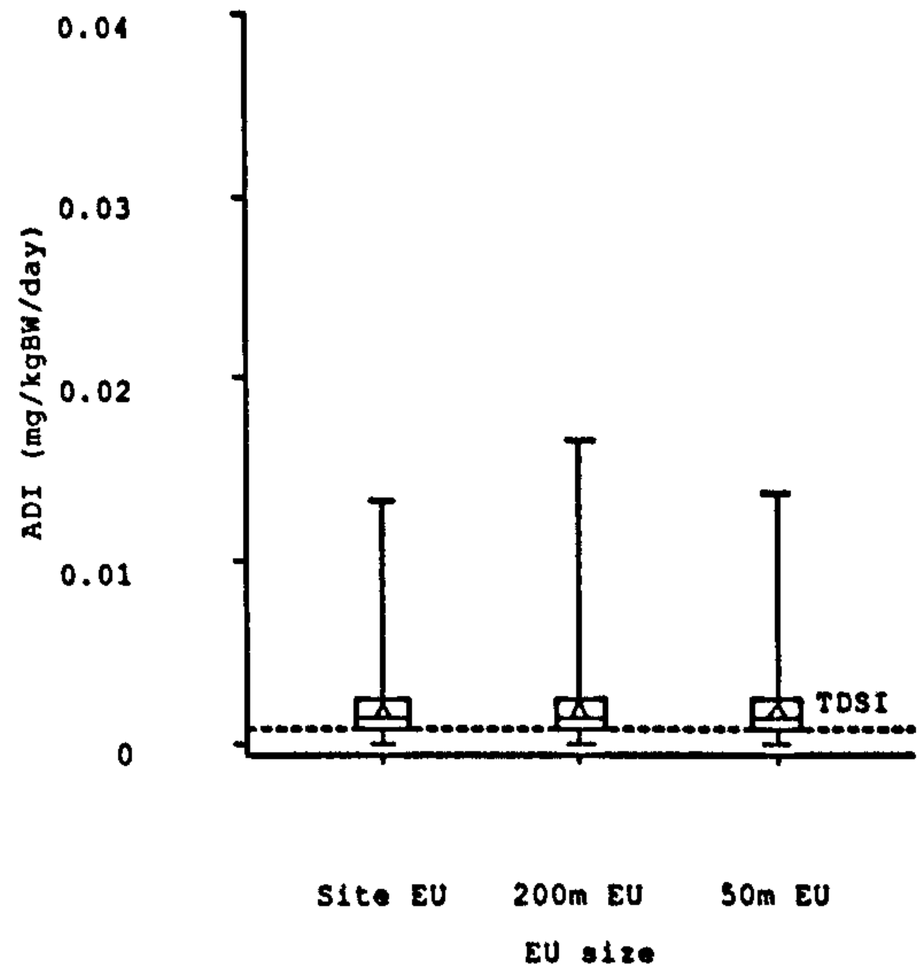
(a) As median *ADI* estimate; 1 yr old



(b) As median *ADI* estimate; 16yr old



(a) Cd median *ADI* estimate; 1yr old



(b) Cd median *ADI* estimate;16yr old

Figure 8.8. Distribution of median *ADI* estimates for As (top) and Cd (bottom) for 1yr olds (left) and 16 yr olds (right). Distributions calculated for three different EU sizes: whole site; 200x200m; and 50x50m. Box and whisker plot lines represent, from top to bottom: Maximum, upper quartile, median, lower quartile, minimum. Triangle represents mean. As ID and Cd TDSI highlighted.

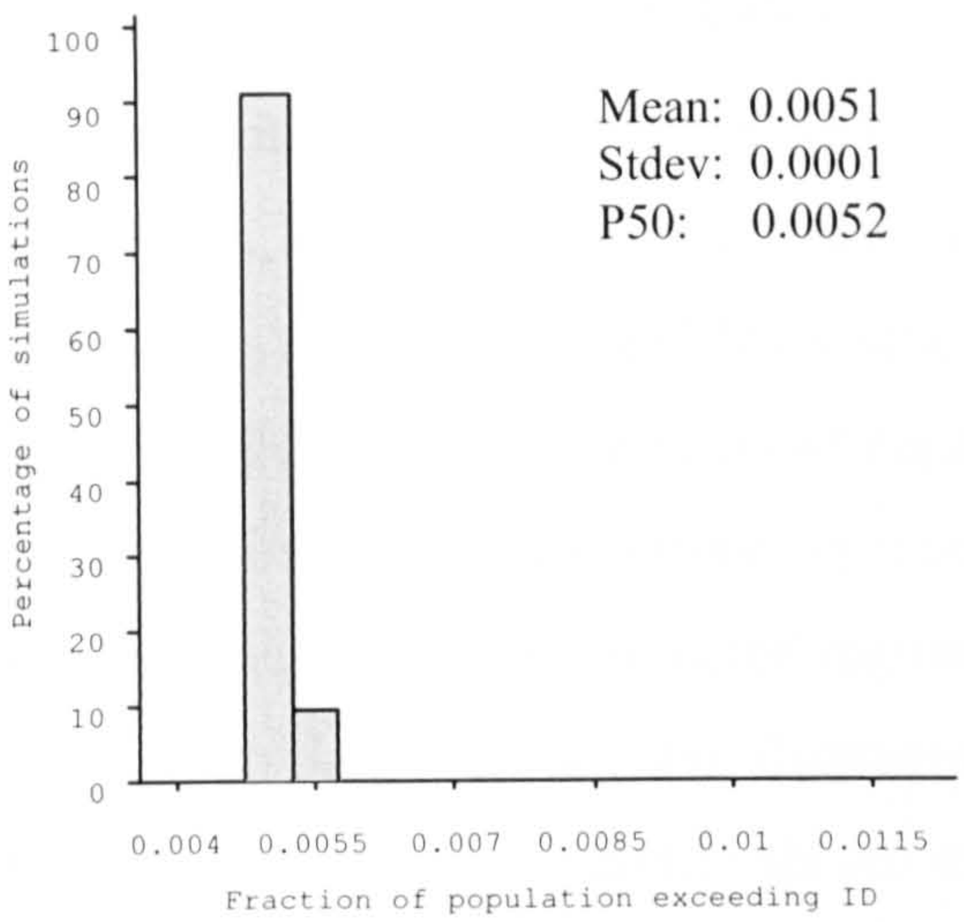
The proportion of the population exceeding the ID/TDSI was calculated for each simulation and the distributions of proportions are shown in Figure 8.9. As seen previously, the proportion of the population with intakes in excess of the ID for As increases as the support size decreases. However, for Cd, the proportion in excess of the TDSI decreases with decreasing support size.

These results are as expected (see Section 8.2): mean As C_{soil} is less than the SGV and mean Cd C_{soil} is greater than the SGV. On moving from a small to a large support, the variance in soil concentrations reduces, resulting in a narrower distribution. This is directly reflected in the *ADI* distributions. In the case of As, where the cut off is at the

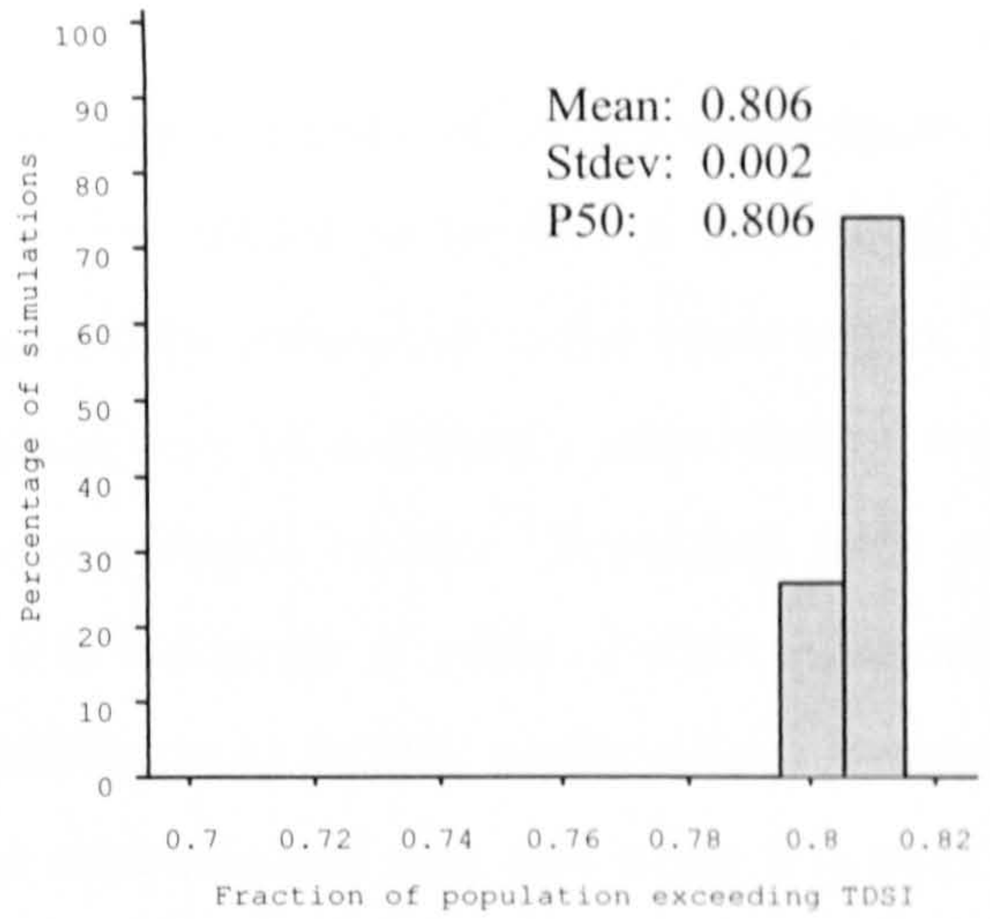
upper end of the *ADI* estimates, the narrow distribution of *ADI* estimates largely falls below the cut off on a large support. At the small support, the distribution becomes wider and some of the *ADI* estimates move in excess of the ID. For Cd, where the TDSI is at the lower end of the *ADI* distribution, at a large support, the narrow *ADI* distribution largely lies above the TDSI. Working at a smaller support, the *ADI* distribution is wider and as a result, some of the *ADI* estimates will move below the TDSI, resulting in a lower proportion of people in excess of the TDSI.

For an actual risk evaluation, judging by the decision criterion (Method P: no more than 5% of the population to have intakes in excess of the ID/TSDI in at least than 50% of simulations) whether this assessment is made using a small support (50m EU), large support (200m EU) or using the site averages (Site EU) does not affect the results. As we have already seen for As, all the intakes are so low that the maximum proportion of the population having an intake in excess of the ID is 1.2%. On the other hand, for Cd, the proportion of the population with excessive intake is always greater than 70%, way in excess of the 5% criterion. Though there is a shift of about 8% on moving from large to small support, it will not make a difference to the results in this instance.

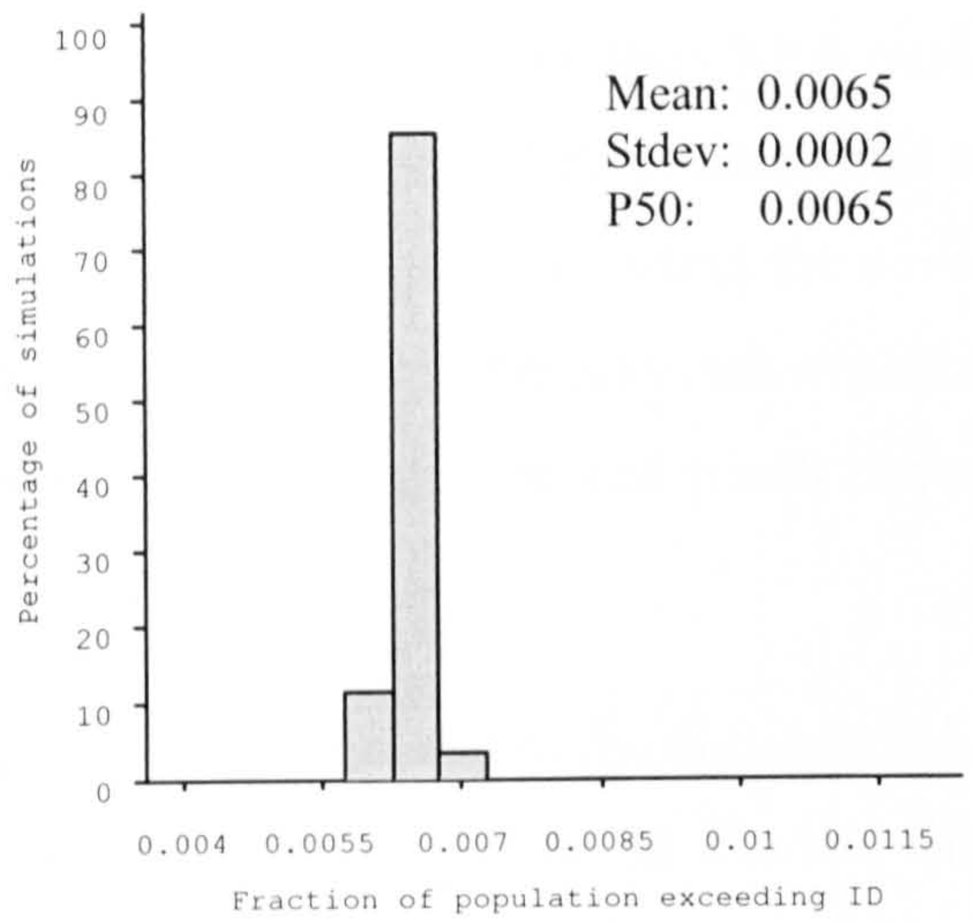
Even though the risk assessment results did not change in this example, the effect of the position of the cut off within the *ADI* distribution has been demonstrated to have an effect on the results. One cannot make a blanket statement which says that working at a small support will lead to higher risk estimates than working on a large support or vice-versa: this is not the case. Rather, it is possible to say that if one is working with a distribution of C_{soil} with a mean below the SGV, working on a small support will likely give the most conservative intake estimates. Whereas if working with a distribution of C_{soil} with a mean in excess of the SGV, working on a large support will likely give the most conservative intake results.



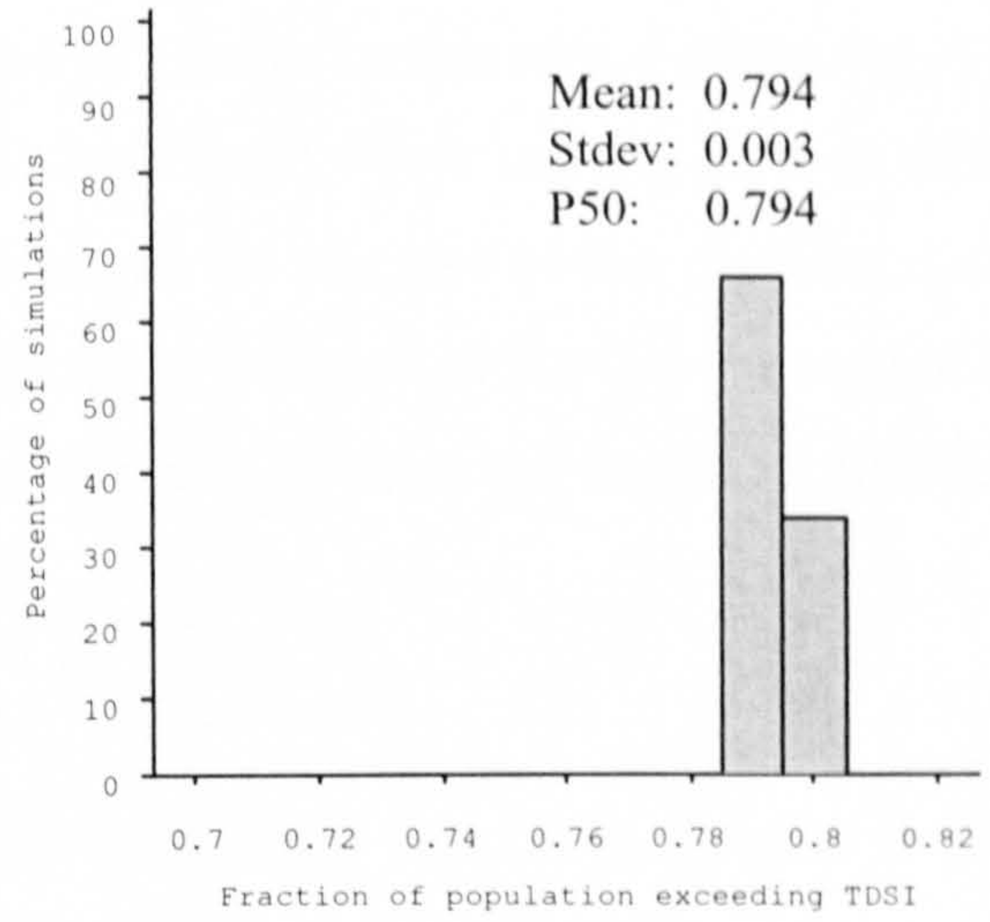
As; Site EU



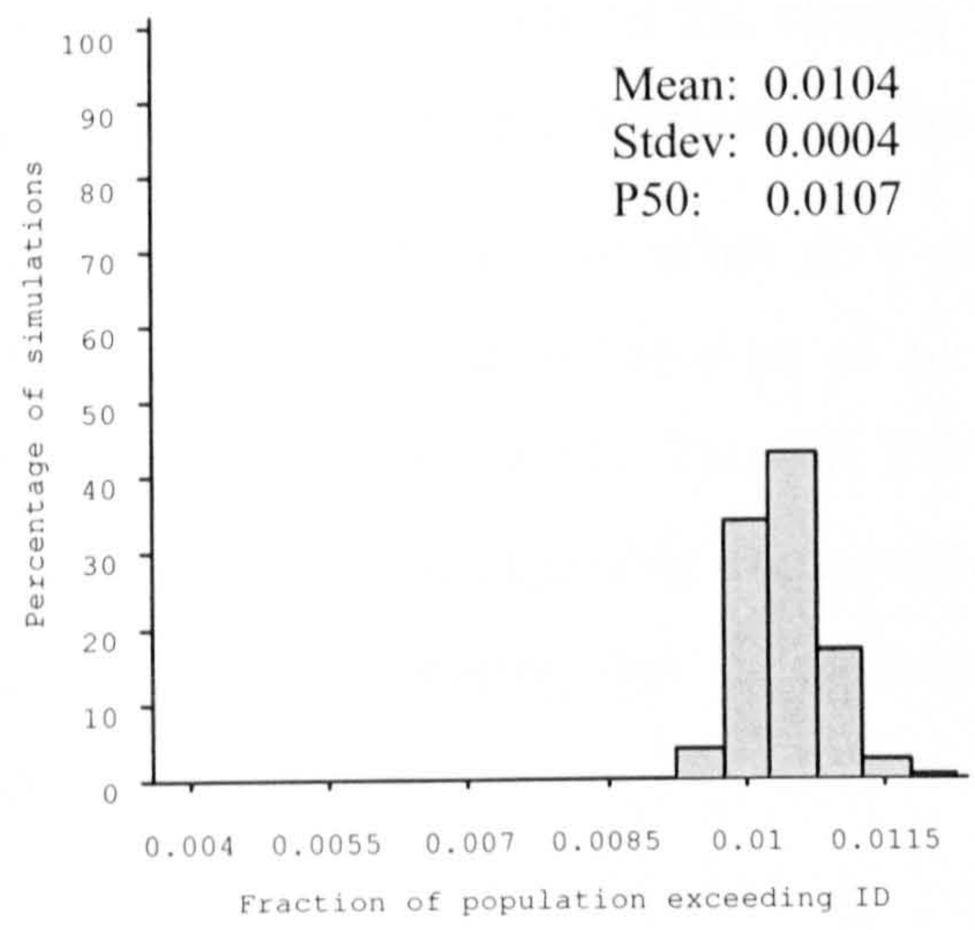
Cd; Site EU



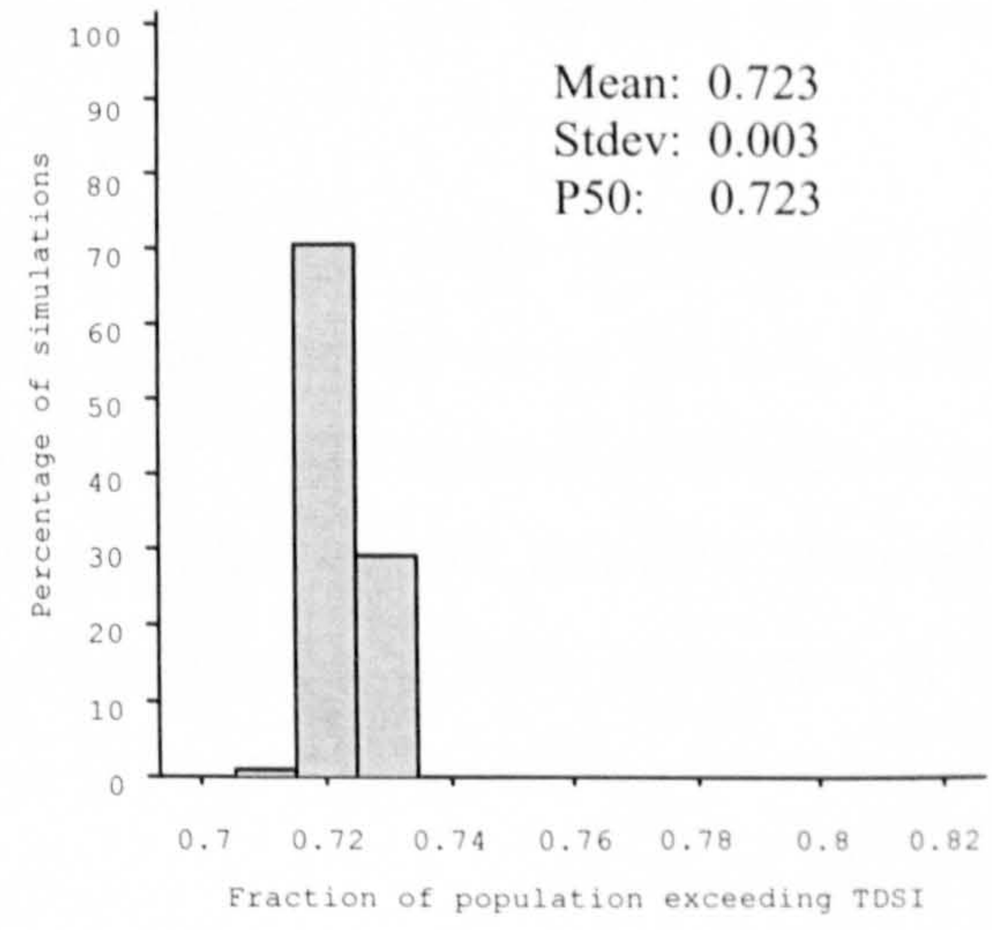
As; 200m EU



Cd; 200m EU



As; 50m EU



Cd; 50m EU

Figure 8.9. Distribution of 1,000 estimates of the proportion of the population exceeding the ID for As (left) and TDSI for Cd (right) on three different supports in Stoke on Trent. (Scenario including home-grown vegetable consumption).

8.3.3 Location of the population

The location of the population with respect to any hotspots of soil contaminants is important. In the methodology developed in this thesis, whereby an individual is assumed to reside in one EU for a whole year and be subjected to the contaminant in that EU only, EUs with hotspots of contamination may be extremely important or may be totally irrelevant, depending on where the population reside. Populated cells are separated out of the contaminated region, and it is the range of values within these cells that is important. This section illustrates the difference in intake results when assuming that the population is randomly spread over the contaminated site, and when it is placed in the correct EU. The point is illustrated with an example data set from the Gai region in Russia.

It is fully recognised that the CLEA model (used as a basis for this new methodology) is designed for use in a UK situation and should not be used where conditions may differ from those assumed in building the model. However, the purpose of this exercise is not to produce a comprehensive risk assessment for the population in the Gai region, but to demonstrate the differences which can occur as a result of using the model in different situations.

8.3.3.1 Background to the Gai data set

Gai is a mining district in Russia: the copper-zinc-pyrite deposit is situated in the Orenburg region of the Southern Urals and has been mined since 1959. Along with mining, dressing and metallurgical processing are also carried out in the region and waste heaps are evident in the vicinity of residential areas. On sampling the topsoil, raised levels of Ni were found in the Gai area.

160 topsoil samples were taken on a grid with a spacing of about 1km (Figure 8.10). The statistics of the distribution of Ni sample values are shown in Table 8.4 and a histogram in Figure 8.11. The UK SGVs, 50mg.kg^{-1} with vegetable consumption and 75mg.kg^{-1} without vegetable consumption, are exceeded by 76% and 28% of samples respectively, indicating that Ni contamination could prove to be a health risk in the region.

Table 8.4. Statistics of Ni sample distribution (mg.kg^{-1}).

Mean	Stdev	Min	P1	P5	P10	P25	P50	P75	P90	P95	P99	Max
74.1	79.3	18	19	31	39	53	63.5	79.5	99	108	280	981

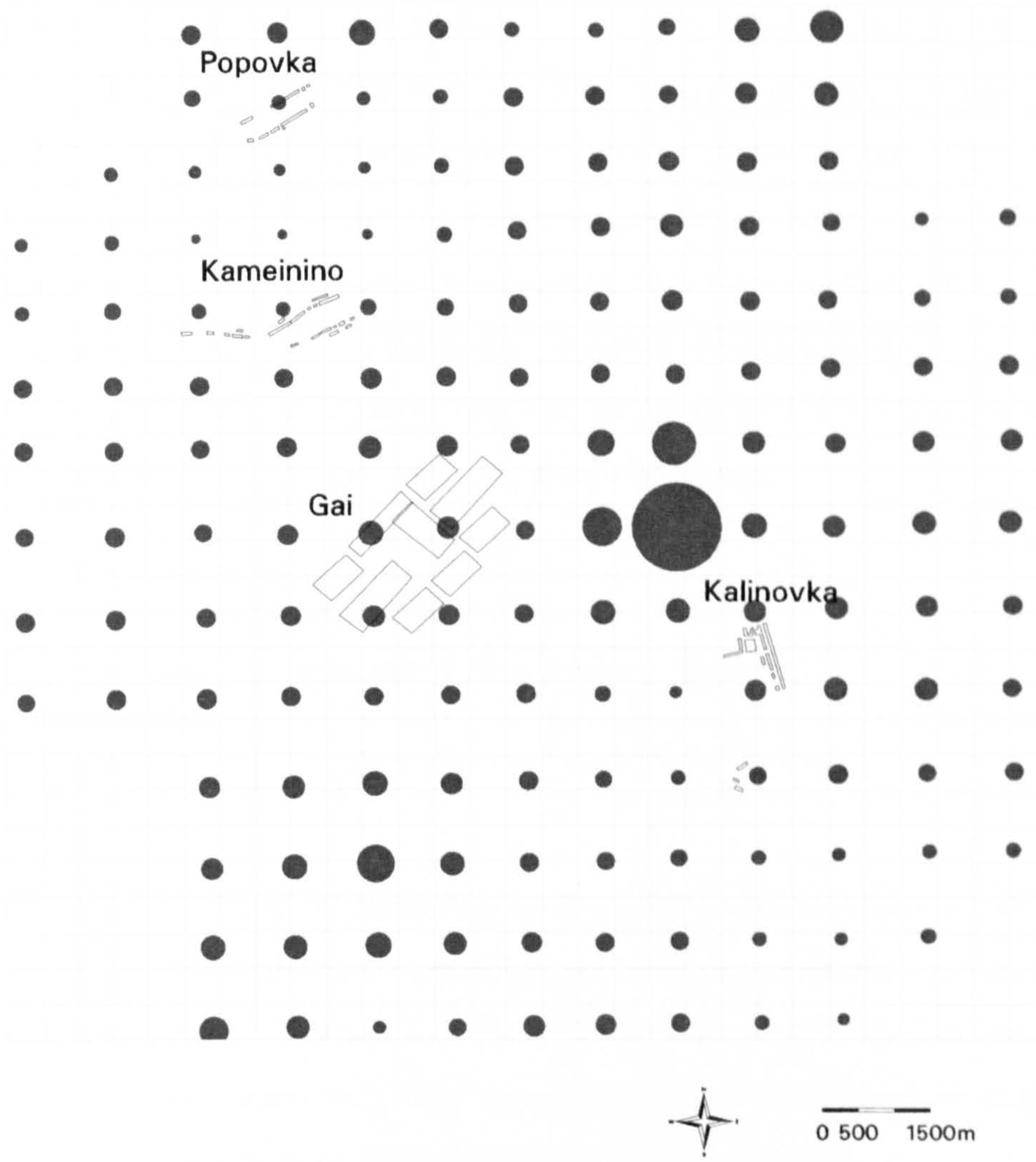


Figure 8.10. Gai region showing built up areas in four dispersed settlements. The grid represents the 500x500m EUs and the points the Ni soil sample points, with symbol size in proportion to Ni concentration.

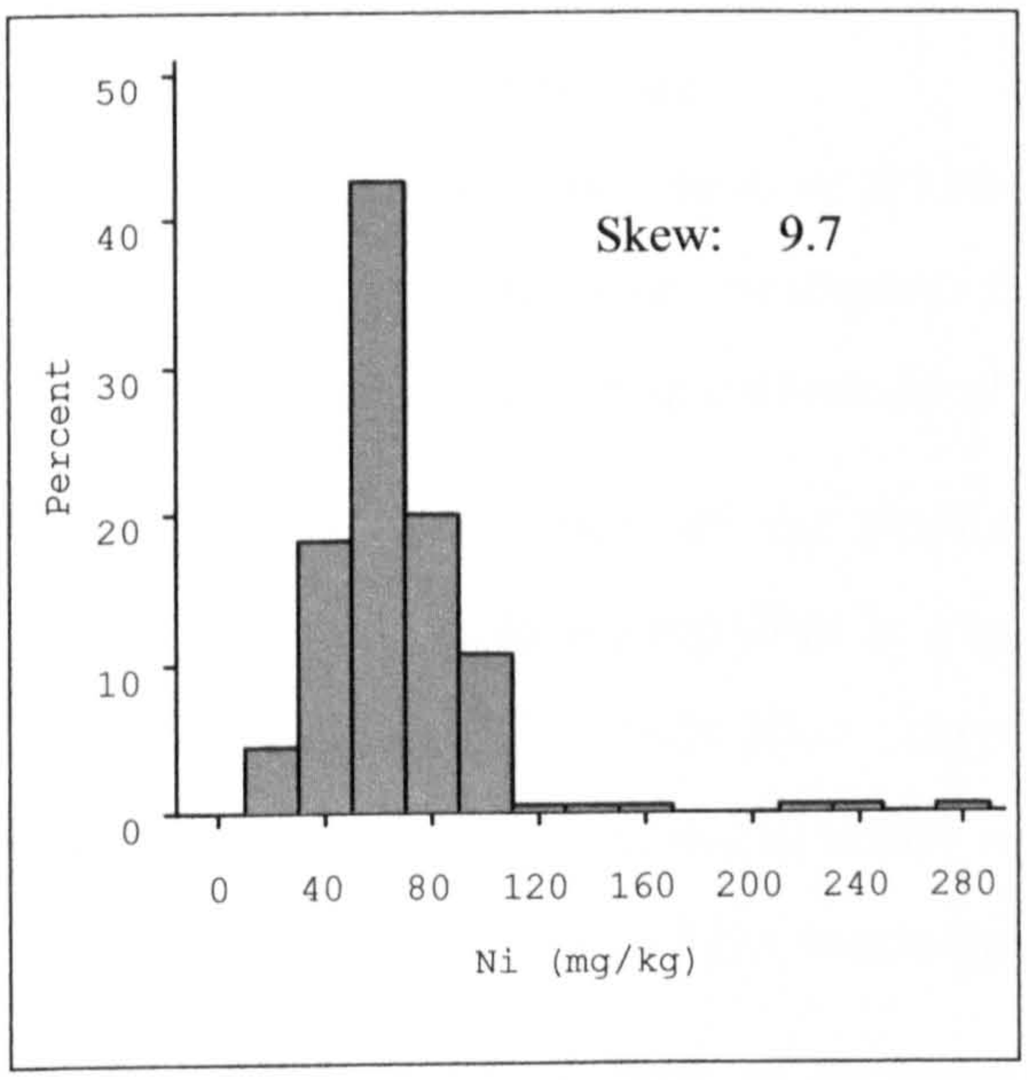


Figure 8.11. Distribution of Nickel topsoil sample concentrations. One value of 981mg.kg^{-1} not shown.

It is clear to see (Figure 8.10) that the population is not evenly spread throughout the region, but is clumped into four settlements: three villages (Kalinovka, Kameinino and Popovka) and a town (Gai). The population, broken down by age class, is given in Table 8.5, and totals 34,997. Ages 0-5 represent pre-school children, 6-17 school-aged children and 18+ adults. Since limited population data were available, the allocation of people to EUs was carried out by the simple method, whereby, within each settlement, individuals are allocated in as uniform a manner as possible to those 'geostatistical EUs' (demarcated by the simulation grid) which contain buildings.

Table 8.5. Population in each settlement in the Gai region

	Gai	Kalinovka	Kameinino	Popovka
0-5yrs	2540	8	22	12
6-17yrs	5760	12	35	26
18+yrs	26076	55	263	188

Full details of the geostatistical procedure used can be found in Appendix B. Suffice it to say that for this example 1,000 plausible maps of soil concentration in the region were produced using SIS (Deutsch & Journel, 1998) on a 50x50m square grid with 290 cells in each direction. Since each 50x50m cell is so small in comparison to the site, the point estimates were taken to be representative of the 50x50m cell.

Minimum and maximum allowable Ni concentrations were set to 10% less and 10% more than the minimum and maximum sample value respectively (i.e. to 16mg.kg^{-1} and 1080mg.kg^{-1}).

8.3.3.2 Analysis of results

To investigate the importance of where on the map the population resides, the intake calculation procedure was undertaken first with the population in the correct EUs and then with the population spread randomly across the site, and the results compared.

The uneven distribution of the population across the region is clear in this case. Contaminant sample information is available across the region, and Ni concentrations can be simulated for each EU. However, using this model for assessing intakes, whereby a person is assumed to reside in their home EU for the full year, means that the simulated C_{soil} results in all the unoccupied EUs do not affect the intake results.

Figure 8.12 shows the median estimated Ni soil concentration after 1,000 simulations on a 50x50m grid. Many of the buildings lie outside the areas with the highest median

Ni concentrations. The corresponding distribution of median Ni values is given in histogram form in Figure 8.13 (left). Also shown is a histogram of the Ni values to which the population is exposed (right). This histogram is achieved by selecting the populated EUs and weighting by the number of residents i.e. so each individual has a median Ni value allocated to them. The difference between the two histograms is striking.

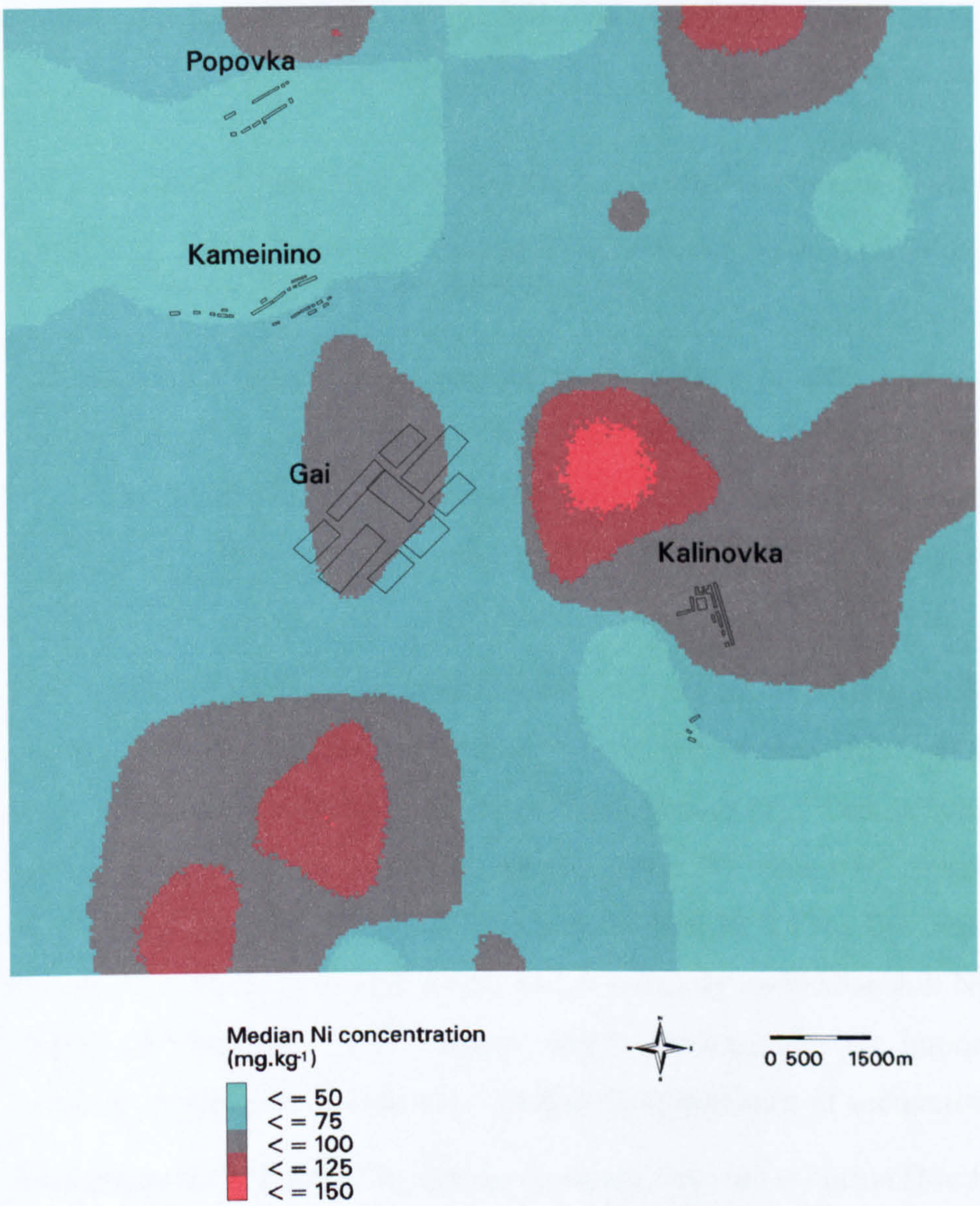
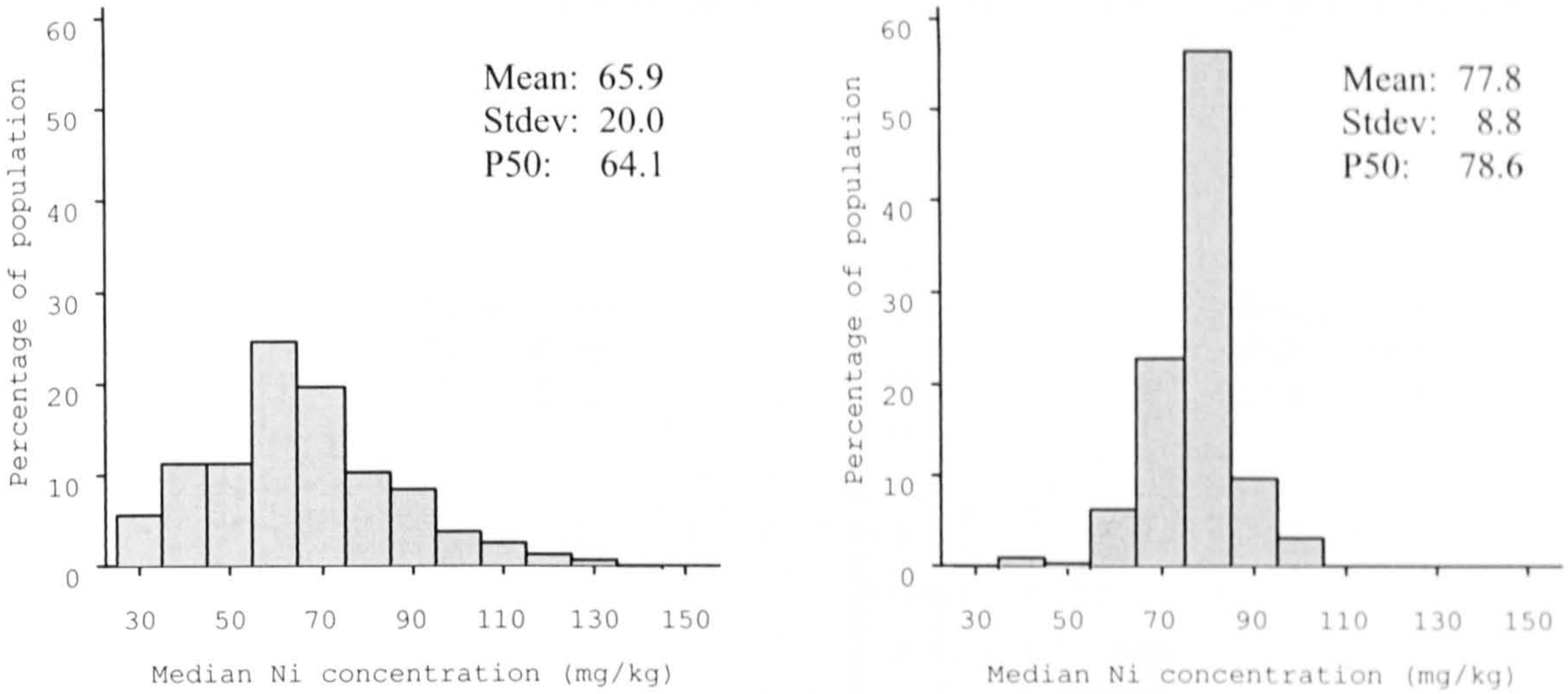


Figure 8.12. Map of median soil Ni concentrations after 1,000 simulations on a 50m grid



(a) Ni concentration for 84,100 50x50m EUs (b) 34,997 person specific Ni concentrations

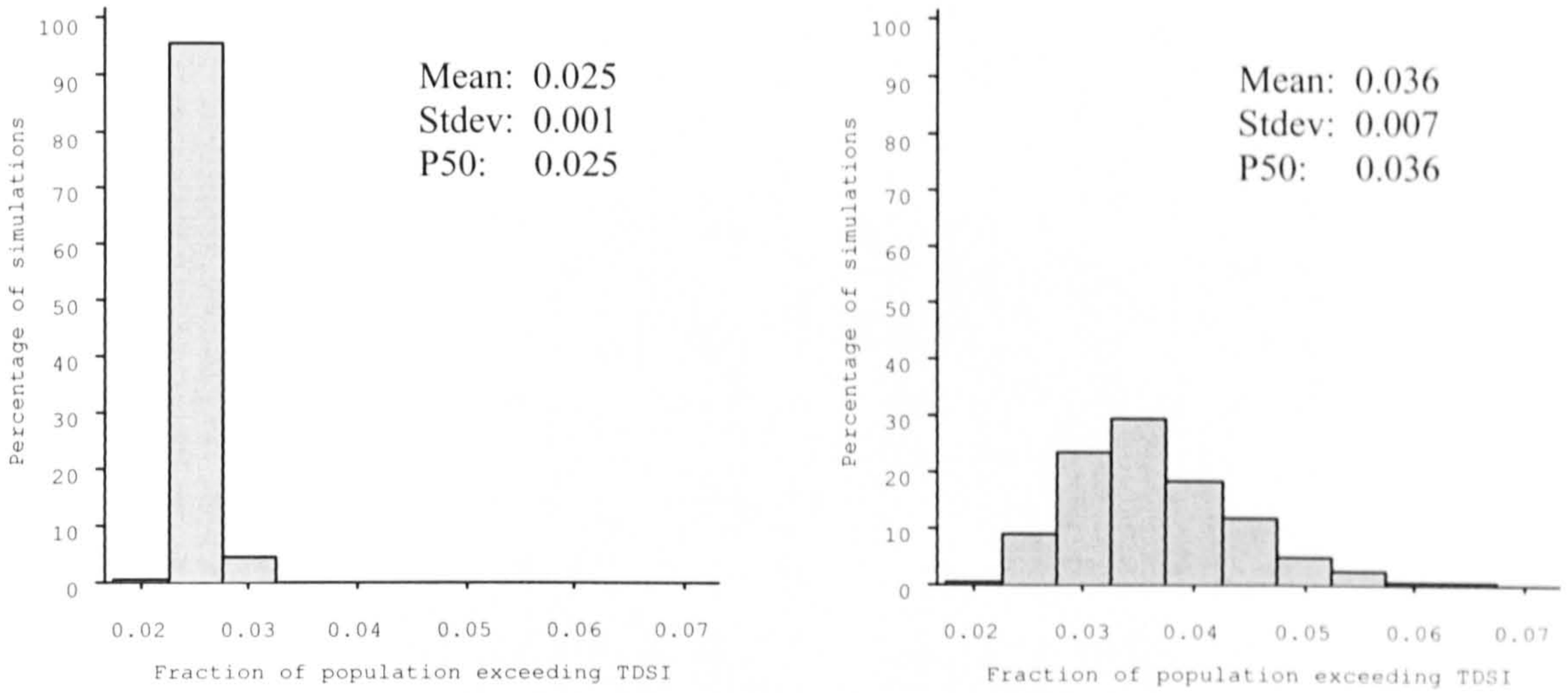
Figure 8.13. Distribution of soil Ni concentrations in the Gai region. Left; for all 50x50m EUs. Right; distribution of values to which individuals are exposed.

The range of median Ni values across the region is quite wide and evenly spread (histogram a) in comparison to the values to which the population are exposed (b). None of the population is exposed to median soil values of 110mg.kg^{-1} or above, and very few at values lower than 60mg.kg^{-1}. The majority are exposed to values around 80mg.kg^{-1}.

The results of running the intake assessment are shown in Figure 8.14. Running the model with a randomly distributed population, in this case, results in a lower estimate of risk (a) than with a correctly distributed population (b). This is because the randomly distributed population is picking up the lower Ni values (<75mg.kg^{-1}) which cover much of the map but are largely unpopulated (Figure 8.12). The bulk of the population lives in the town of Gai (Table 8.5) which is mostly on median soil Ni concentrations in excess of 75mg.kg^{-1}. This example clearly demonstrates the importance of knowing where the population resides in space and the importance of accounting for this fact.

Evaluating the risk using the two dimensional decision criterion (Method P), we see that whether this assessment is made with a randomly or properly distributed population does not affect the result. Although in the correctly distributed assessment a proportion of the simulations have more than 5% of the population with an intake in excess of the TDSI, the proportion is less than the stipulated 50%, so in both cases, the risk is acceptable. Clearly, there may be instances, when the C_{soil}/population-location mix is

different, and the proper allocation of population could affect whether the population is classed as at-risk or not.



(a) Intakes for randomly placed individuals

(b) Intakes for correctly located individuals

Figure 8.14. Distribution of results, from 1,000 simulations, of the estimated fraction of population exceeding the TDSI for Ni. (a) population randomly distributed in space (b) population correctly allocated to residential EUs.

8.3.4 Location of population combined with change of support

Depending upon the location of the population, the decision to work at one support or another may or may not affect the results of the intake assessment. The effect will be more noticeable if the population is located on the more extreme soil values rather than on middle of the range values. Following on from the previous section, the change of support effect is examined for the correctly located Gai population.

8.3.4.1 Analysis of results

Figure 8.15 shows the median Ni soil concentrations after 1,000 simulations on two different supports (500m and 50m). It can be seen that the 500m support produces higher median soil values than the 50m support in some areas. However, it is clear that the population is located outside of the regions showing most change and so may be little affected by the decision to work at one support or another. This is indeed the case and is confirmed in Figure 8.16: despite the marked change in overall soil concentration distribution (left) the fraction of the population exceeding the TDSI is virtually the same (right), since it is located in regions unaffected by change of support.

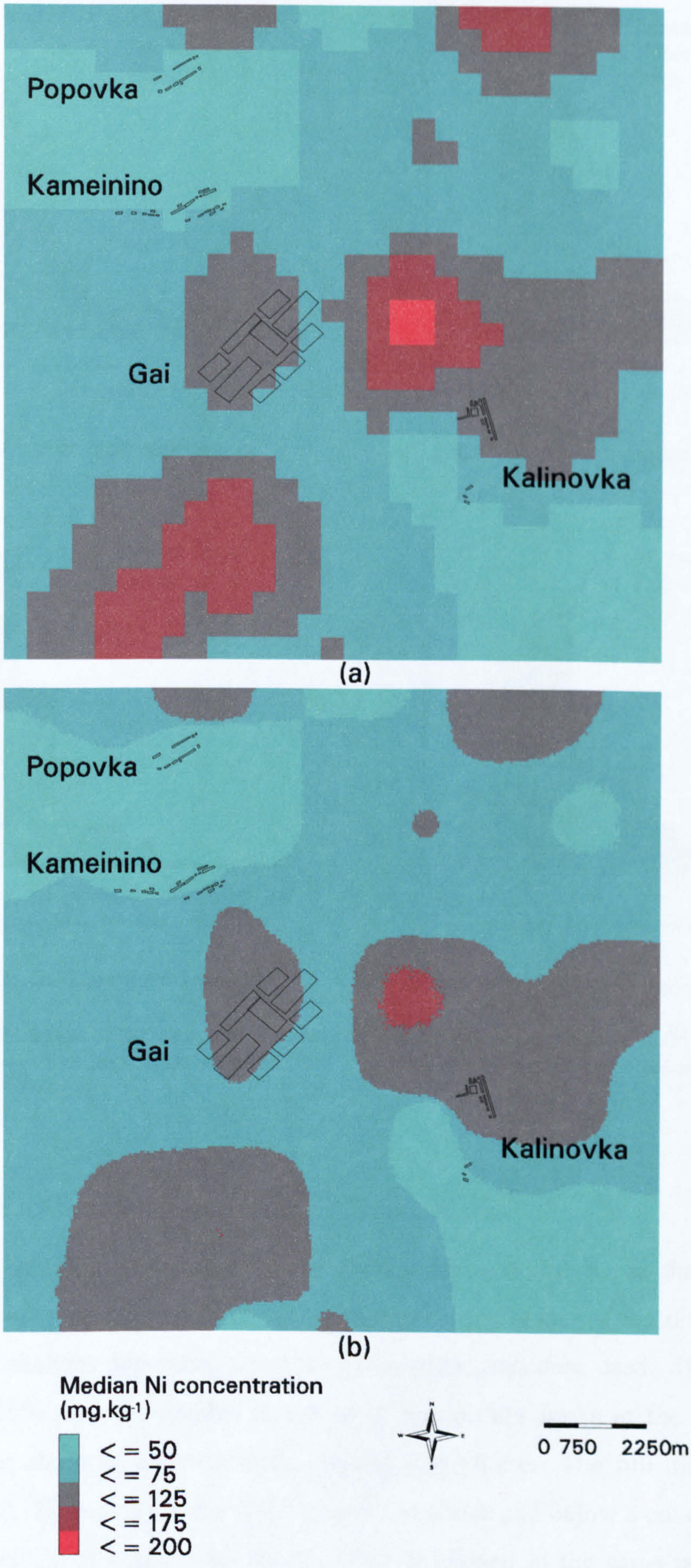
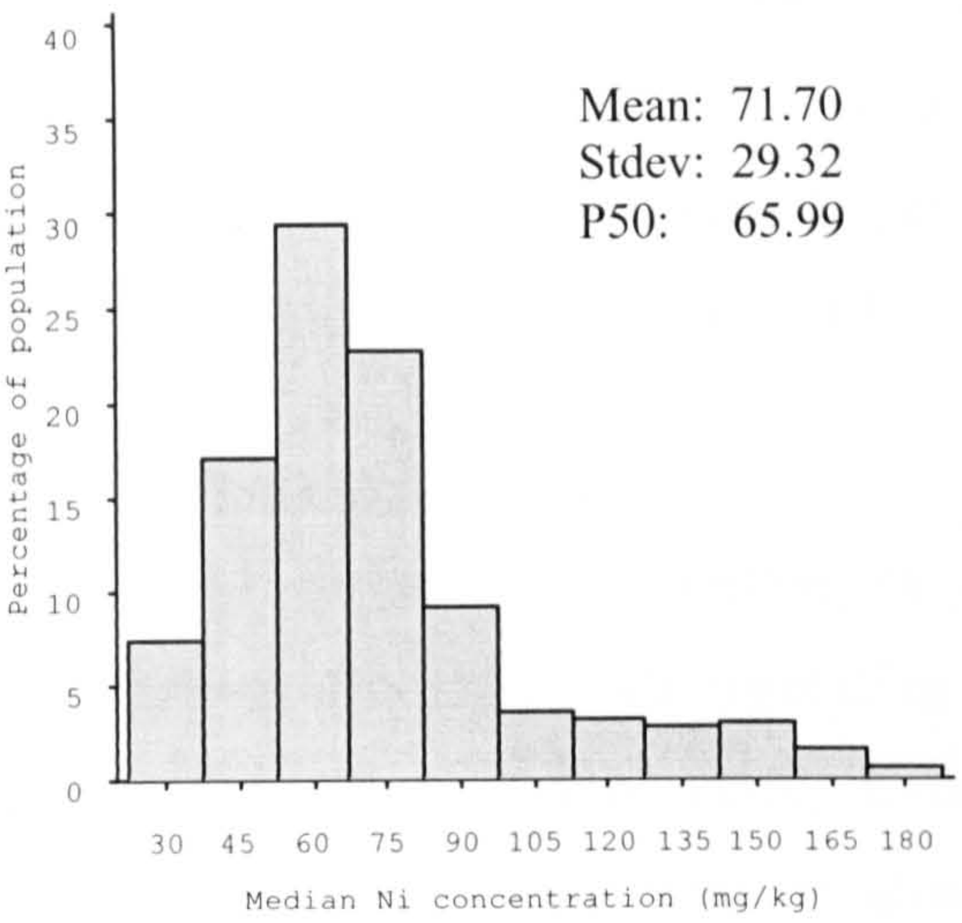
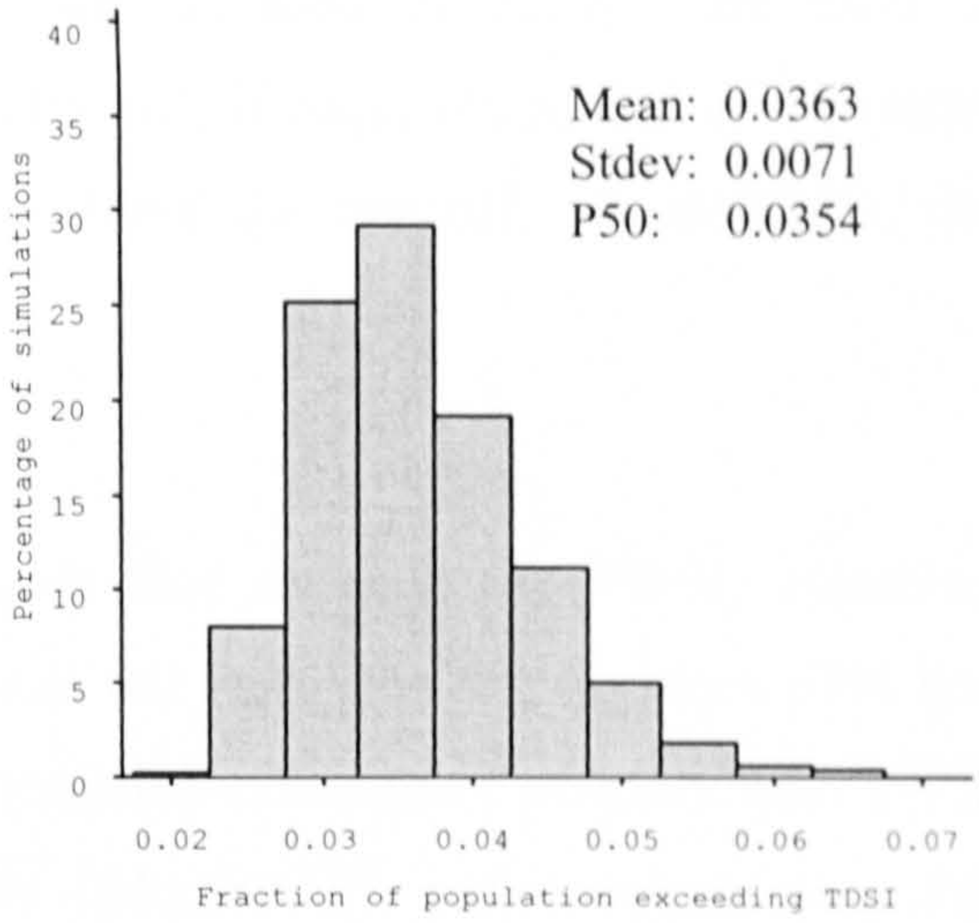


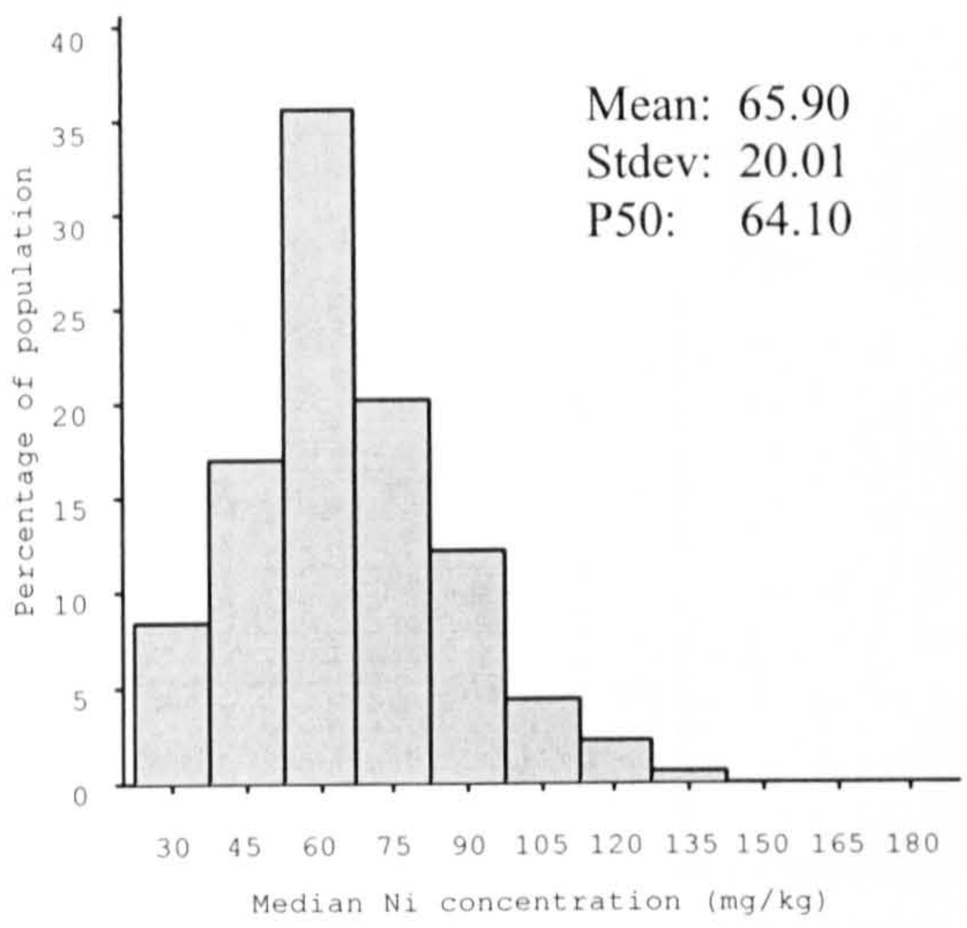
Figure 8.15. Median Ni soil concentrations in Gai on a 500m support (a) and a 50m support (b).



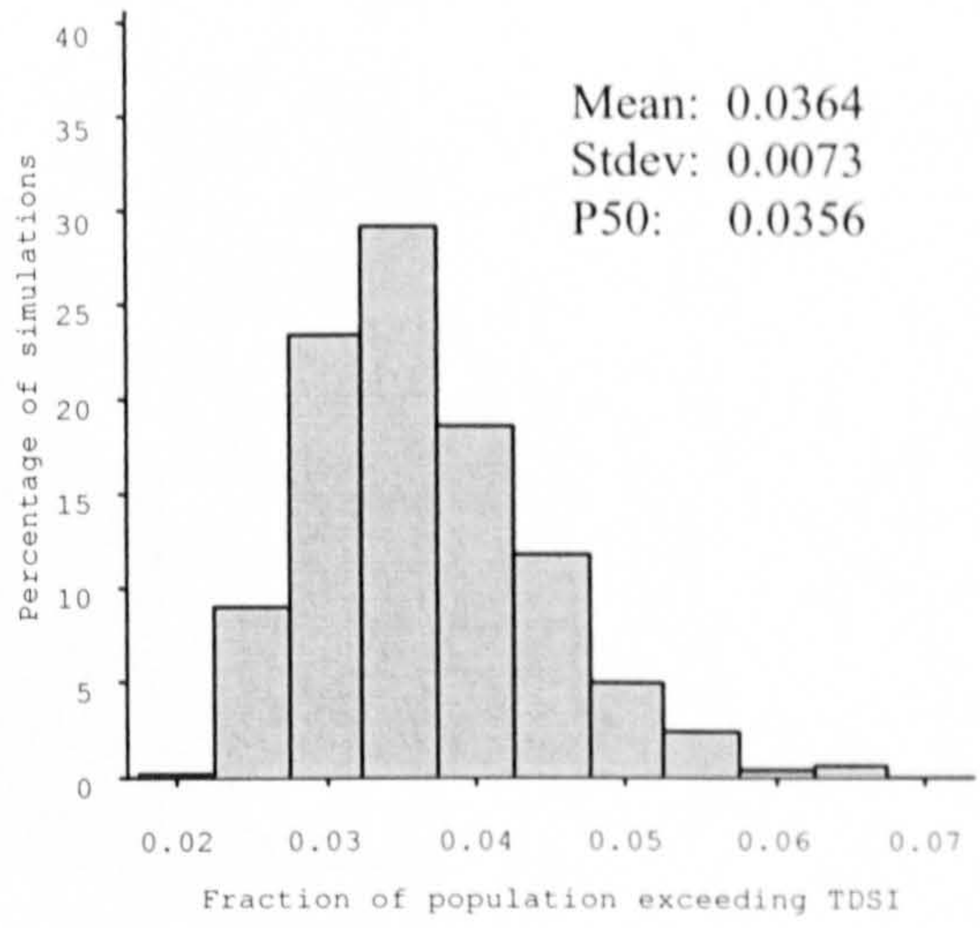
(a) 500m EU; Ni concentration ($mg.kg^{-1}$)



(b) 500m EU; Fraction exceeding TDSI



(a) 50m EU; Ni concentration ($mg.kg^{-1}$)



(b) 50m EU; Fraction exceeding TDSI

Figure 8.16. Comparison of the change in distribution of median Ni concentrations (left) and change in distribution of fraction of population exceeding the TDSI (right) for two different supports; 500m EU (top) and 50m EU (bottom).

8.3.5 Loss of information when working with cut offs

All the above examples use the risk assessment decision criteria in the form of a regulatory cut off. As alluded to in earlier examples, the shape of the distribution of *ADI* estimates changes depending upon the calculation procedure used. However, the full depth of information available is not used if one only looks at the results with regard to being above or below a threshold and stops there. The full distribution of estimates is lost. If one only looks at the proportion above and below a cut off, a person with an *ADI* ten times bigger than the ID/TDSI is classed in the same bracket as an individual who only just exceeds the limit.

The Cd intake estimates for 16yr olds in Stoke, as used in 8.3.2.2, are used to demonstrate the rich information that can be extracted, if required, to add to the rather simplistic, but easy to understand, proportion above the cut off. Specifically, the estimates on the 50m support are used.

8.3.5.1 Analysis of results

Figure 8.17 shows the proportion of 16yr olds that exceeds the TDSI, assuming consumption of home-grown vegetables. We can tell from this that between 67% and 72% of 16yr olds exceed the TDSI. Using the decision criterion of no more than 5% in excess in no more than 50% of simulations (Method P), we conclude that this population of 16yr olds is at risk. However, it would be interesting to find out by how much they are over-exposed. This involves looking at the individuals in the population.

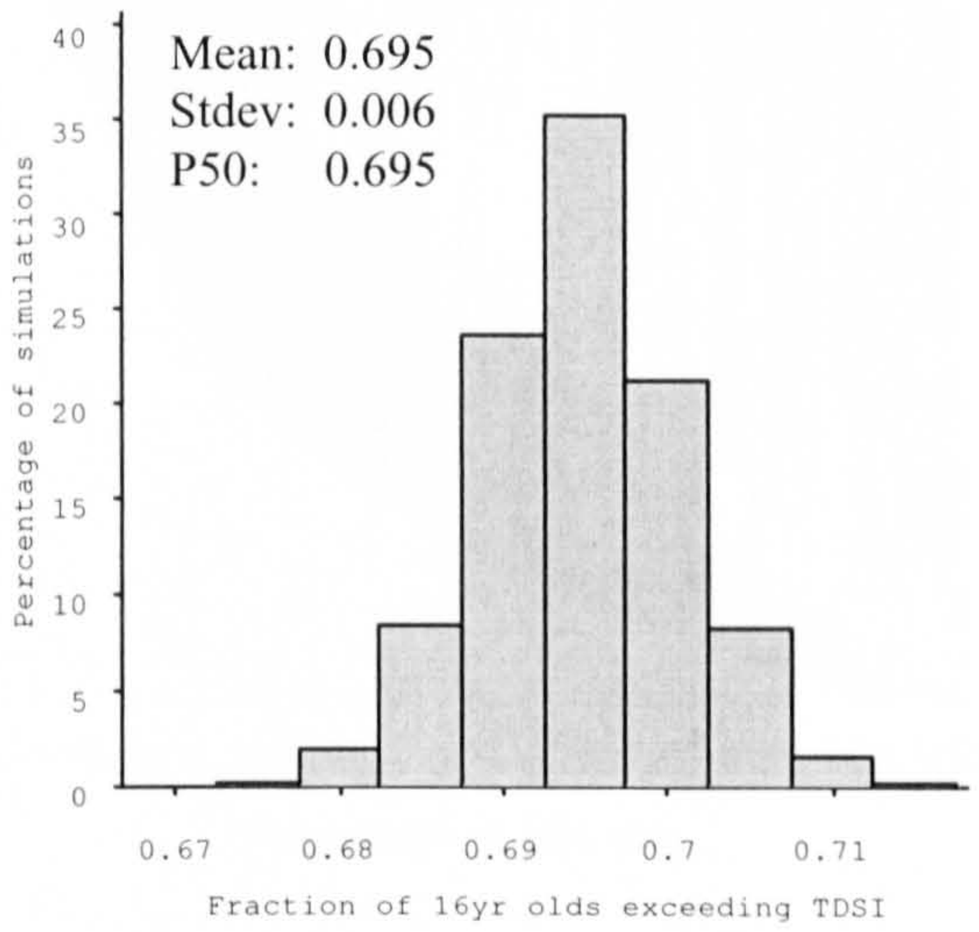


Figure 8.17. Distribution of 1,000 estimates of the proportion of 16yr olds exceeding the TDSI ($0.00077 mg.kg^{-1}_{BW}.d^{-1}$) for Cd in Stoke on Trent.

Figure 8.18 shows the distribution of *ADI* estimates for each individual. The 1,000 estimates are classed into bins according to whether they are below the TDSI, twice, four times, or more than four times over the TDSI limit. The top plot is ordered by decreasing *ADI* and the bottom by increasing *ADI*. Each plot makes it easier to gather information of interest. For example, from the top plot, we can see that 18% of the 16yr olds have at least a 50% probability of exceeding the TDSI by more than 4 times (i.e. having an ADI in excess of $0.00308 mg.kg^{-1}_{BW}.d^{-1}$, for a TDSI of $0.00077 mg.kg^{-1}_{BW}.d^{-1}$). From the bottom plot, we can see that only 8% of 16yr olds have at least a 95% chance of being below the TDSI, and 23% at least a 50% chance.

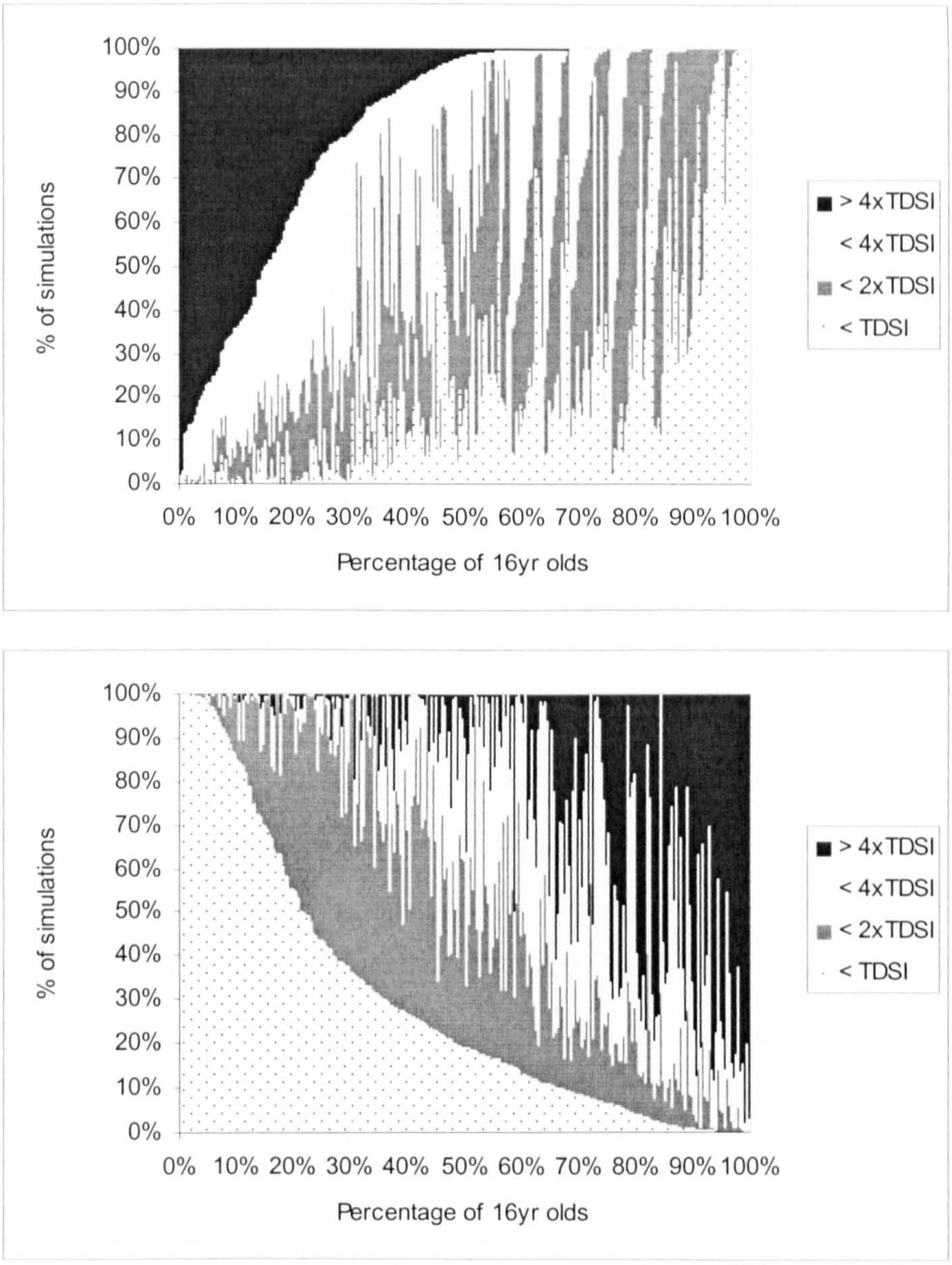


Figure 8.18. Distribution of Cd *ADI*s for each individual 16yr old in Stoke ($mg.kg^{-1}_{BW}.d^{-1}$). Top, ordered by highest intakes first, bottom ordered by lowest intakes first.

The pattern of each vertical line on the plot gives a representation of the uncertainty in the result for each individual. Slicing through the plots in Figure 8.18 at the 50% and 95% level on the y-axis, provides the median and P95 values for each person. These are plotted in Figure 8.19 and illustrate how far above the TDSI the median and P95 estimates for each individual lie.

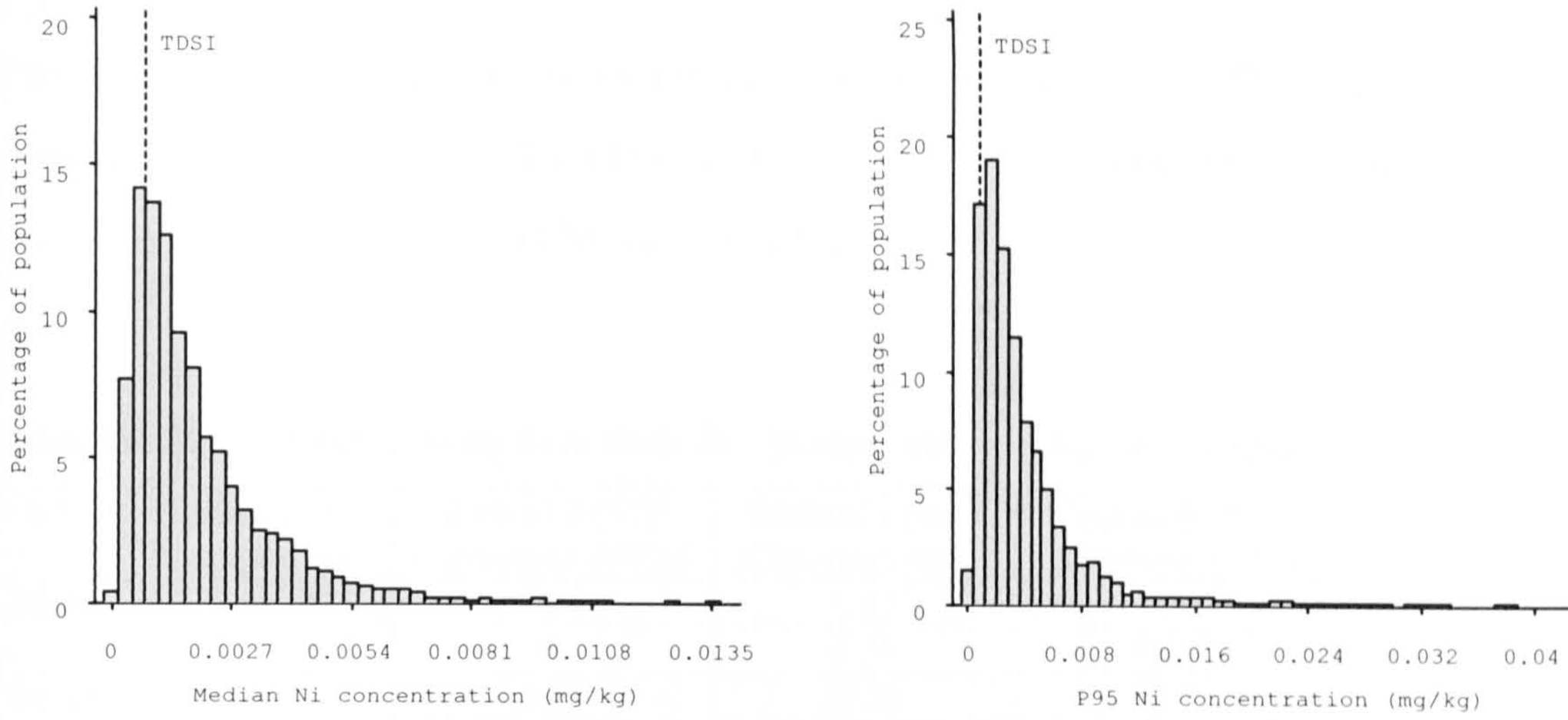


Figure 8.19. Distribution of median (left) and P95 (right) estimates of Cd *ADI* for 16yr olds in Stoke.

This analysis gives a more in depth view of the variability among 16yr olds and uncertainty about the individual results. Though working with cut offs is desirable, from the point of view of decision making, one should not lose sight of the additional information which is available for use as a result of the analysis procedure. Especially in a case like the one demonstrated, where a large percentage of the population is regarded as at risk, it is possible to use the additional information to target those most in need of risk reduction.

8.3.6 Working with populations or individuals

Assume that we, at least initially, work with a threshold and a two dimensional decision criterion. Recalling Chapter 6 Section 6.9.2.1, the evaluation can be performed in more than one way. The preferred method to assess the risk to the population, as has been used so far in this chapter, is directly at the population level (Method P: no more than 5% of the population should exceed the cut off in at least 50% of simulations). The second method is to assess the individual risks first and amalgamate to the population level (Method I: no more than 5% of individuals in the population may have a 50% probability or more of exceeding the cut off). Finally, it is possible to calculate the expected proportion of the population with intakes in excess of the cut off and this should not exceed 5% (Method E).

It was demonstrated with a theoretical example in Chapter 6 that the three assessments may lead to different risk evaluation results. Here, the three evaluation methods are applied to the results of the three example data sets introduced so far.

8.3.6.1 Analysis of results

The results of the intake assessments are shown in Table 8.6. The percentage of the population in excess of the ID/TDSI is given for each calculation method. The results from Method P, Method I and Method E differ.

Table 8.6. Risk estimation using three methods. Shaded cells represent at-risk results.

Site	Contaminant & support size	Method P (%) [a] (Criterion: <50%)	Method I (%) [b] (Criterion: <5%)	Method E (%) [c] (Criterion: <5%)
Stoke	As (50m EU)	0	0.42	1.04
	As (200m EU)	0	0.55	0.65
Stoke	Cd (50m EU)	100	79.21	72.3
	Cd (200m EU)	100	79.96	79.4
Gai	Ni (50m EU)	4.4	2.45	3.64
	Ni (500m EU)	4.5	3.18	3.63

[a] Percentage of simulations with more than 5% of population exceeding ID/TDSI
[b] Percentage of individuals with probability of exceeding ID/TDSI ≥ 50%
[c] Expected proportion of population exceeding ID/TDSI

The differences are generally quite small, except for the difference between the population method and others for Cd in Stoke. In these examples, the differences do not affect the classification of the populations as at-risk or not. Both the Stoke As and Gai Ni do not pose a threat to the receptors, whereas the receptors have excessive intakes of Cd in Stoke. In the no-risk situations (Stoke As and Gai Ni) the population method gives the healthiest result for As, whereas for Ni, the individual method paints the healthiest picture. In the at-risk situation (Stoke Cd), the expected method produces the least risk result.

It should be noted that when tackling the problem directly at the population level, risk to the population is assessed a whole map at a time for each of the 1,000 simulations. Therefore, spatial correlation of the soil values to which the population is exposed is preserved. Because of this, this author prefers the idea of working directly at the population level, since it guarantees that the spatial correlation of the soil values, which the method tries to preserve, is passed on into the intake results. In addition, the uncertainty in the proportion of the population exceeding the cut off is preserved.

8.3.7 pH dependent Cd intake assessment

Cd intake is sensitive to soil pH, since the soil to plant concentration factor, CF_{veg}, is pH dependent: the lower the pH, the higher the CF_{veg}. Therefore, for a more realistic

spatially-evaluated intake assessment, it is necessary to take variable pH into account. This is demonstrated first with a small data set form Vasilevskaya in Russia.

Again, it is fully recognised that the CLEA model (used as a basis for this new methodology) is designed for use in a UK situation and should not be used where conditions may differ from those assumed in building the model. But again, the purpose of this exercise is to demonstrate the differences which can occur as a result of using the model in different situations, rather than provide a comprehensive risk assessment for the population in Vasilevskaya.

8.3.7.1 Background to the Vasilevskaya data set

The Vasilevskaya coal mines lie in the Tula region of Russia, approximately 200km south of Moscow. Opencast mining methods are used to extract Lignite which is processed at the sites and, consequently, large waste dumps are found in the area. 135 topsoil samples were taken across the region on a square grid with an approximate spacing of 500m (Figure 8.20).

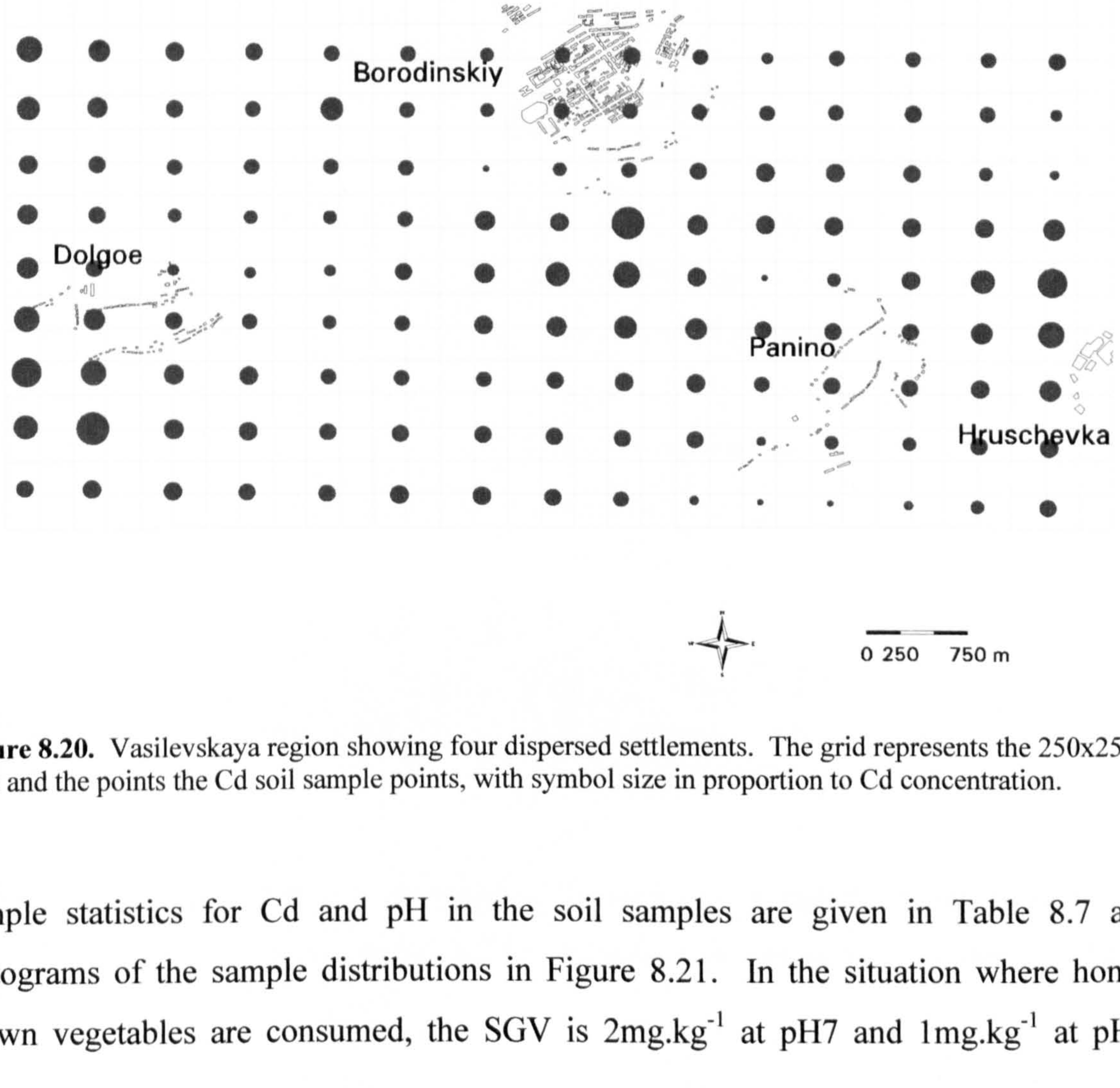


Figure 8.20. Vasilevskaya region showing four dispersed settlements. The grid represents the 250x250m EUs and the points the Cd soil sample points, with symbol size in proportion to Cd concentration.

Simple statistics for Cd and pH in the soil samples are given in Table 8.7 and histograms of the sample distributions in Figure 8.21. In the situation where home-grown vegetables are consumed, the SGV is 2mg.kg^{-1} at pH7 and 1mg.kg^{-1} at pH6.

There is a possibility that some of the population may be overly exposed to Cd, depending on the combination of pH and Cd found in each EU.

Table 8.7. Simple statistics of the Cd topsoil sample concentrations ($mg.kg^{-1}$) and pH values.

	Mean	Stdev	Min	P1	P5	P10	P25	P50	P75	P90	P95	P99	Max
Cd	0.65	0.30	0.1	0.1	0.2	0.4	0.5	0.6	0.7	1.0	1.2	1.9	1.9
pH	6.86	0.81	4.2	4.8	5.2	5.7	6.4	7.1	7.5	7.7	7.7	7.8	7.8

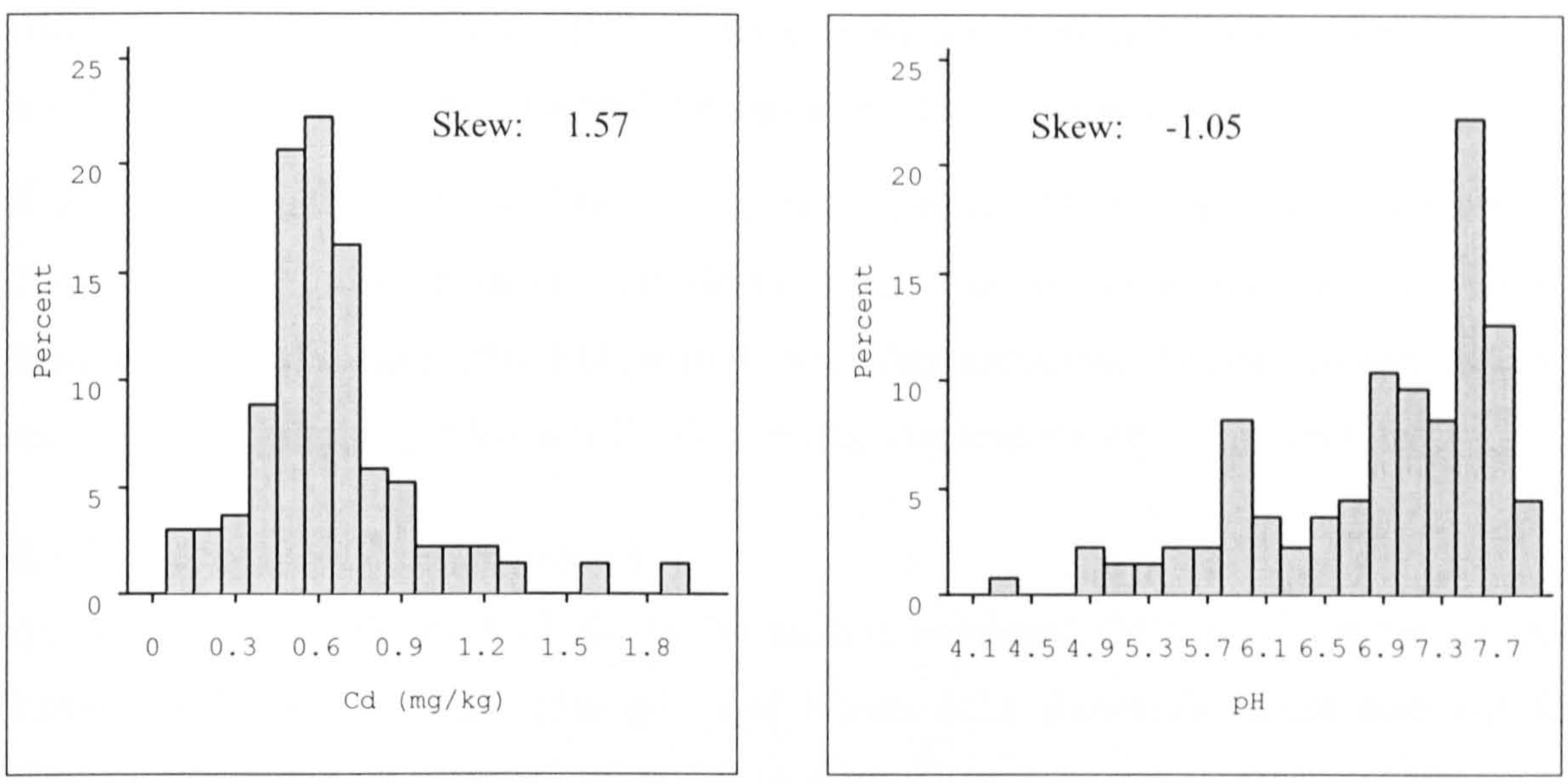


Figure 8.21. Histogram of Cd topsoil sample concentrations ($mg.kg^{-1}$) (left) and pH values (right).

In the study area are four settlements (Figure 8.20), relatively small in size, with a total population of just 6,554 people (Table 8.8). Ages 0-5 represent pre-school children, 6-17 school-aged children and 18+ adults. Since limited population data were available, the allocation of people to EUs was carried out by the simple method, whereby population is allocated in as uniform a manner as possible to those 'geostatistical EUs' (demarcated by the simulation grid) which contain buildings.

Table 8.8. Population in the Vasilevskaya mining region.

	Panino	Borodinskiy	Dolgoe	Hruschevka
0-5yrs	7	402	8	0
6-17yrs	12	969	17	15
18+yrs	46	4917	104	57

In this case, pH and Cd must be assessed for each EU, so the simulation procedure is carried out for both. Full details of the procedure are given in Appendix C. 1,000

plausible soil concentration maps were produced using SIS (Deutsch & Journel, 1998) on a 25x25m square grid with 330 points in the East West direction and 160 North South, making a total of 52,800 points. Since each 25x25m cell is so small in comparison to the whole region, the point estimates are taken to be representative of the 25x25m block.

Minimum and maximum Cd concentrations were set to 0.1mg.kg^{-1} (see Appendix C) and 10% more than the maximum sample value (i.e to 2.1mg.kg^{-1}), assuming that the true maximum value was unlikely to have been found in sampling. Minimum and maximum pHs were set to 4.1 and 7.9 respectively (see Appendix C).

Results were calculated for three scenarios: a single pH for the whole site for the duration of the calculations (no variability or uncertainty accounted for); one with a constant pH within each 25m EU (spatial variability accounted for, but no uncertainty); and one with variable pH in each EU (variability and uncertainty accounted for).

8.3.7.2 Effective Cd concentration

As an illustration, Figure 8.22 shows the median estimated Cd soil concentration after 1,000 simulations on a 25x25m grid and Figure 8.23 shows the same map but for median pH values. A simple correlation of the medians gives ρ=0.29, suggesting there is not a very strong spatial relationship between median Cd and median pH values. The weak correlation is positive i.e. high median Cd tends to be found in the same place as high median pH.

In order to map the regions with the highest likely exposures, a function is developed to approximate the spatial distribution of 'effective Cd' concentrations i.e. a function which can quantify the combination of the effects of Cd and pH. High Cd/low pH leads to higher intakes, while low Cd/high pH leads to lower intakes.

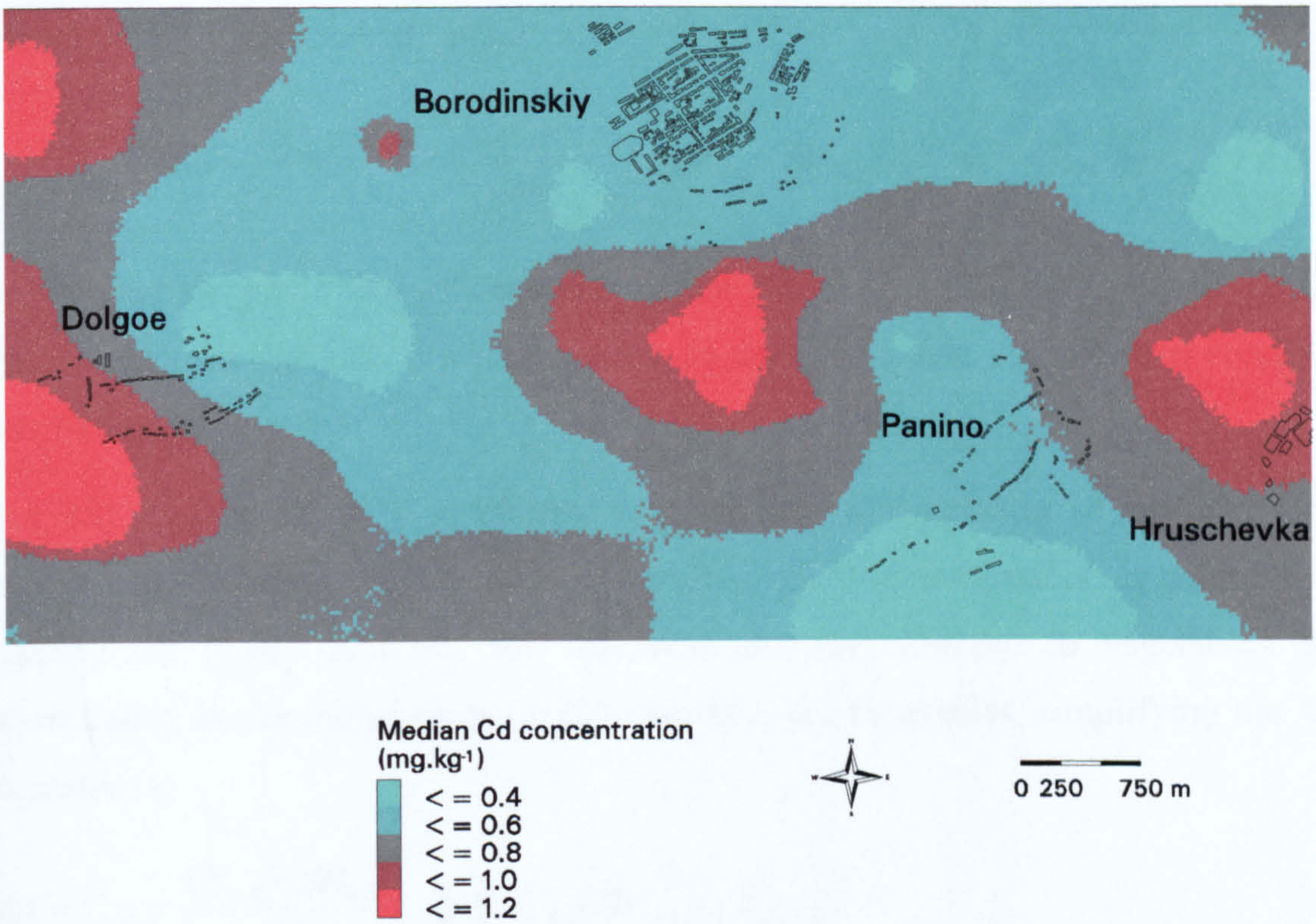


Figure 8.22. Map of the median soil Cd concentration after 1,000 simulations on a 25m grid.

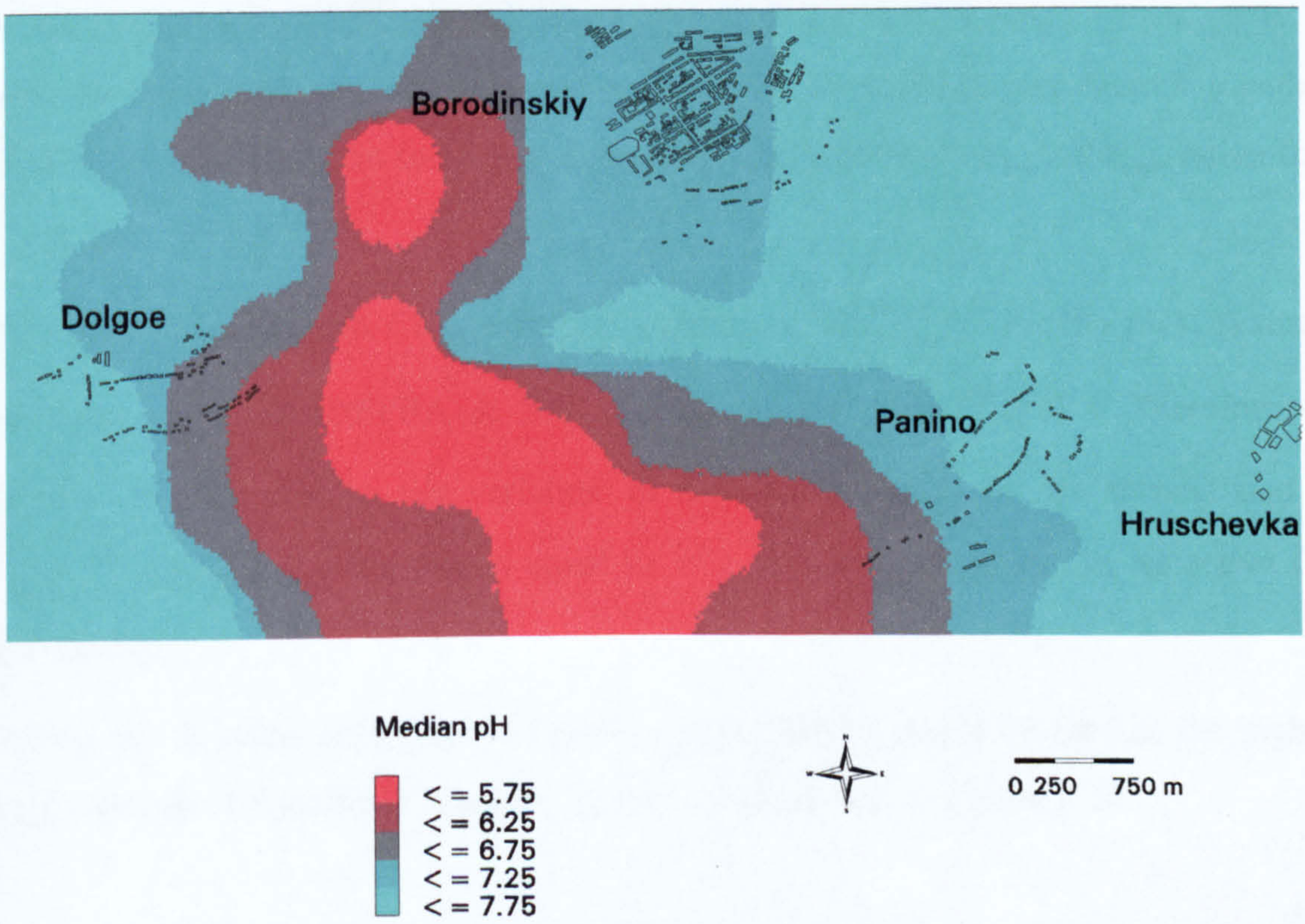


Figure 8.23. Map of the median soil pH after 1,000 simulations on a 25m grid.

Going back to the intake equation, $ADI = C_{soil} \times SE$, variable pH affects the *ADI* through the CF_{veg} term which is part of *SE*:

$$SE = \frac{EF_{oral} \times ED_{oral}}{AT}\left(\frac{SDR}{BW} + \sum_{veg\ type}\left(CR_{veg} \times HF_{veg} \times \left(SL_{veg} + CF_{veg}\right)\right)\right)$$

In effect, the soil part of the *ADI* equation is now made up of C_{soil} and CF_{veg}, as CF_{veg} is dependent on the soil pH. The aim is now to derive a single variable that combines these two effects.

Recalling the analysis of *SE* in Chapter 7, it was seen that virtually all intake of Cd is through that absorbed into vegetables. Hence, a good approximation is achieved by dropping the routes of direct soil ingestion and soil attached to vegetables, and concentrating on the intake route for Cd absorbed into vegetables, simplifying the *ADI* calculation to:

$$ADI = C_{soil} \times \frac{EF_{oral} \times ED_{oral}}{AT}\left(\sum_{veg\ type}\left(CR_{veg} \times HF_{veg} \times CF_{veg}\right)\right)$$

Using average values for CF_{veg} across all vegetable types (see Chapter 6),

$$ADI = C_{soil} \times CF_{veg}\left(\frac{EF_{oral} \times ED_{oral}}{AT}\sum_{veg\ type}\left(CR_{veg} \times HF_{veg}\right)\right)$$

Thus, the rearrangement of the equation brings both variable soil factors together: variation in *ADI* due to the soil component is now dependent on $C_{soil} \times CF_{veg}$, rather than C_{soil} alone.

The relationship between CF_{veg} and pH is given as $\ln(CF_{veg}) = a - (b \times \mathrm{pH})$, where a and b are constants (DEFRA & EA, 2002g), so $CF_{veg} = e^{-bpH} \times e^{a}$. Therefore the variation in *ADI* due to the variation in soil components can be summarised as $kC_{soil}e^{-bpH}$ (where k is a constant equal to e^{a}) and shall be called the 'effective Cd' concentration.

Plotting this function on a map will give an indication of where on the site the highest likely risk from Cd intake will occur. This is demonstrated in Figure 8.24.

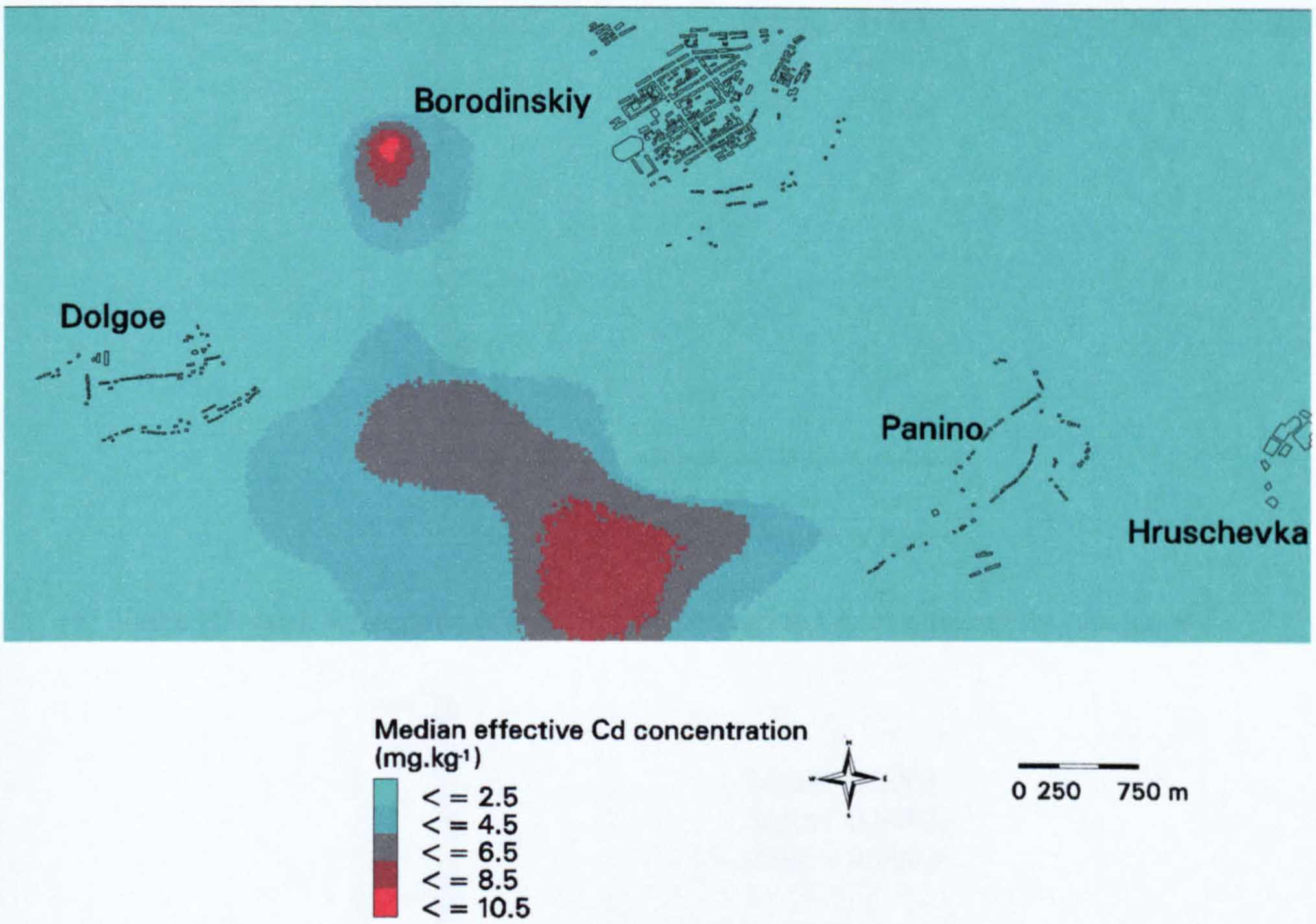
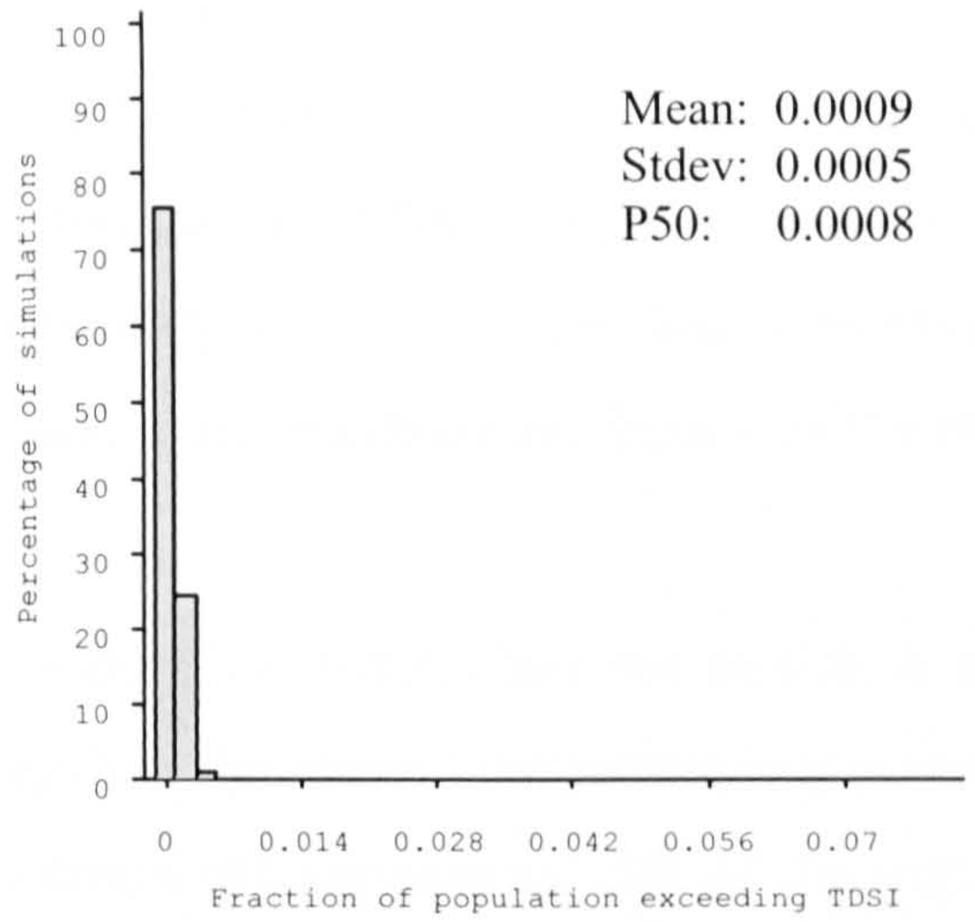


Figure 8.24. Map of the median effective Cd concentration after 1,000 simulations on a 25m grid.

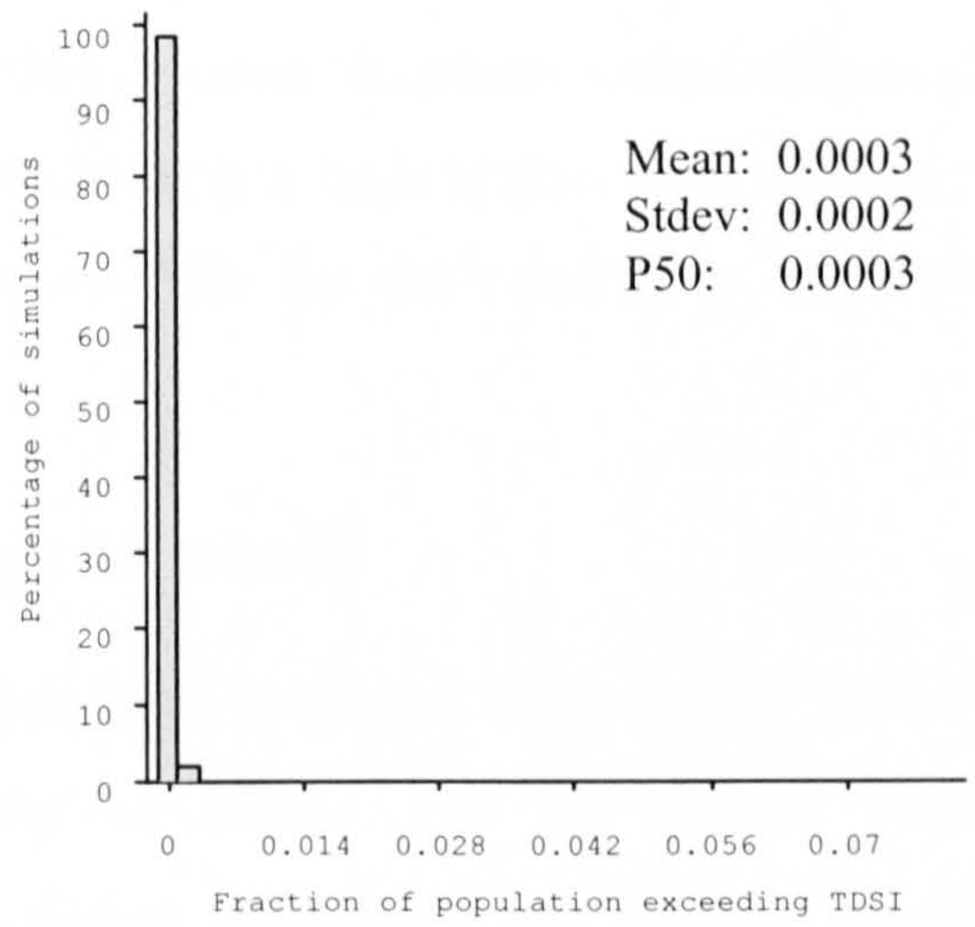
The hotspot to the East of Borodinskiy in Figure 8.24 is where high Cd concentrations and low pHs coincide (see Figure 8.22 and Figure 8.23). The highs and lows of the effective Cd concentrations follow a similar pattern to the pH spatial distributions, suggesting that pH has a stronger influence than Cd concentrations on the effective Cd, and proving the importance of including variable pH.

8.3.7.3 Analysis of results

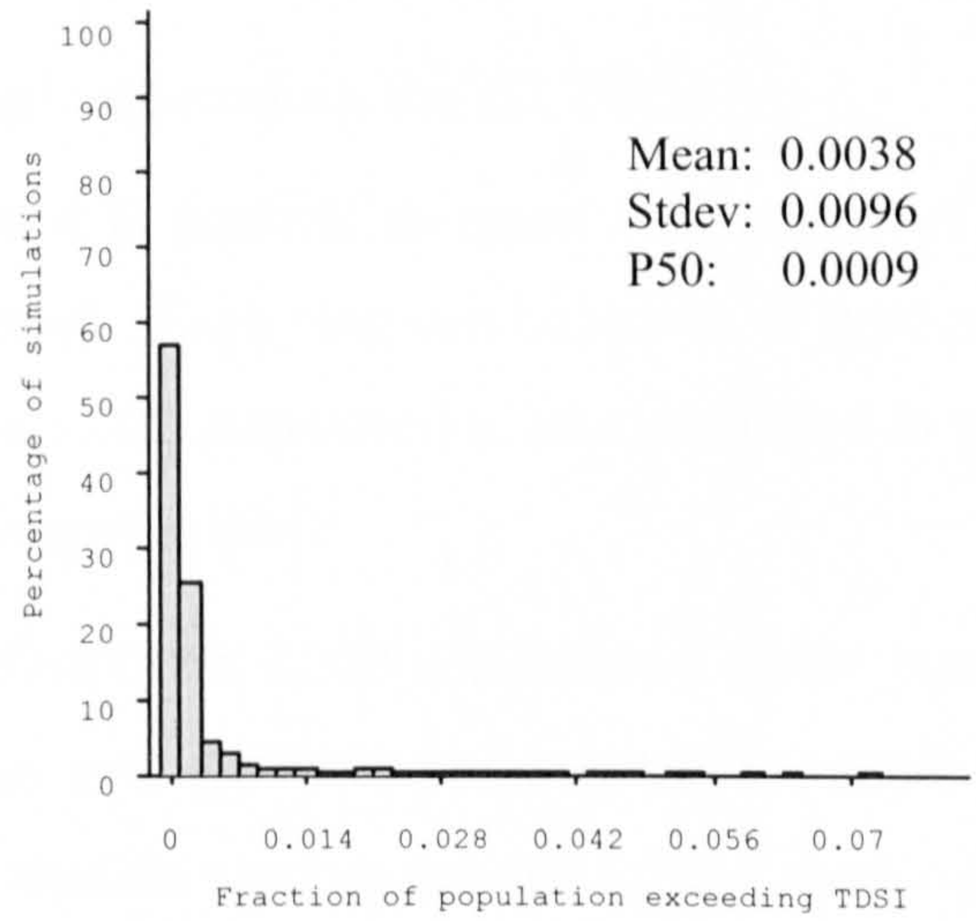
Figure 8.25 shows the results for the proportion of population with intakes in excess of the TDSI for the three scenarios described above: a single pH (mean pH: 6.86) for the whole site for the duration of the calculations (no variability or uncertainty accounted for); a constant pH within each 25m EU (spatial variability accounted for, but no uncertainty); and variable pH in each EU (variability and uncertainty accounted for).



(a) Single pH value throughout calculation (no pH variability or uncertainty accounted for)



(b) Single pH value per EU (pH variability accounted for)



(c) pH changes in each EU in each Simulation (pH variability and uncertainty accounted for)
(one fraction, 0.12, not shown)

Figure 8.25. Distribution of results, from 1,000 simulations, of the estimated fraction of population exceeding the TDSI for Cd, with different soil pH regimes. (a) No spatial variability or uncertainty in pH (b) Spatial variability only in pH (c) Spatial variability and uncertainty in soil pH.

The healthiest picture is painted when variability only in pH is taken into account (b). The worst predictions of *ADI* are produced, in this case, when full spatial variability and uncertainty in the pH is considered in the intake calculations (c). Taking account of both variability and uncertainty in the soil pH has a marked effect on the results, producing a much wider and more uncertain prediction of the fraction of the population which is over-exposed.

Judging the results by the decision criteria laid out earlier, it is evident that no matter how pH is treated, the result is the same – the population is not at risk from Cd intake. For the variable and uncertain pH scenario (c) 5% of the population does exceed the TDSI but in only 1% of simulations. The mean fraction exposed after 1,000 simulations is 12 times higher in the variable and uncertain scenario (c) than the variable only scenario (b). It is possible in some situations with different pH/Cd combinations, the difference in approach could have a large impact on whether a population is declared as at-risk or not. To account fully for the uncertainty and variability is the preferred method for this author.

8.4 Spatial displays of results

Up until now, results have been displayed in histogram form. The histograms give an idea of the uncertainty in results, but the spatial variability element is not communicated. Since, during calculation, the x and y coordinates are preserved for each individual it is possible to analyse and plot results in map form.

8.4.1 Number of people exceeding the ID/TDSI

For each simulated map it is possible to count the number of people in each EU that exceed the intake threshold. Each map can be drawn to give one possible outcome for the result. However, the 1,000 maps need to be assimilated to produce an overall useful measure for the risk manager to use.

Each EU has a distribution (over 1,000 simulations) of the number of people in excess of the ID/TDSI. For an average measure, one could summarise the 1,000 results by displaying the mean or median number per cell. For a more conservative measure, one could display the upper 95th percentile result. An example is shown in Figure 8.26. This shows a spatial display of the number of people per 500m EU who exceed the TDSI for Ni in Gai. The left hand map uses the median estimate from the 1,000 simulated maps and the right the P95 estimate. The P95 estimates are slightly higher, as

expected. Using the median results, no people in Popovka are classed as at risk, for example, whereas at the P95 level two EUs have up to 50 people with intakes in excess of the TDSI.

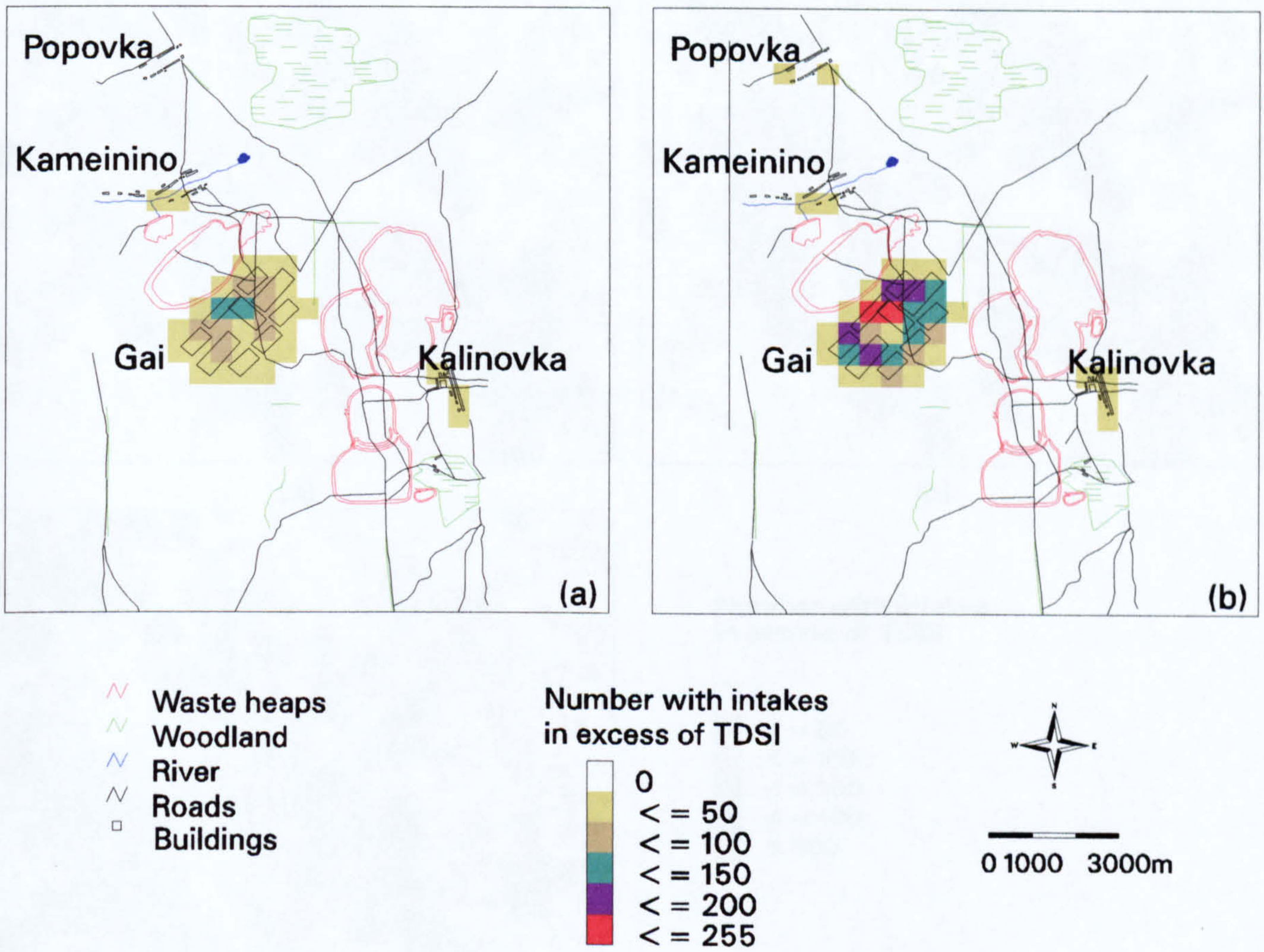


Figure 8.26. Number of people in Gai exceeding the TDSI for Ni in each 500m EU. (a) median estimate (b) P95 estimate.

A similar map is shown in Figure 8.27. This time a comparison is made of calculating exposure to Cd in Stoke using the Site EU, 200m EU and 50m EU. The results are disaggregated/aggregated in order to plot on a 200m grid so that they may be compared. One can see that the Site EU calculation produces a higher median estimate of the number of people per 200m cell who exceed the TDSI for Cd. This is consistent with the non-spatial presentation of results in Section 8.3.2.2, Figure 8.9. This comparison is made out of interest to reinforce the change of support effect, however, this would not be the main purpose of producing such a map.

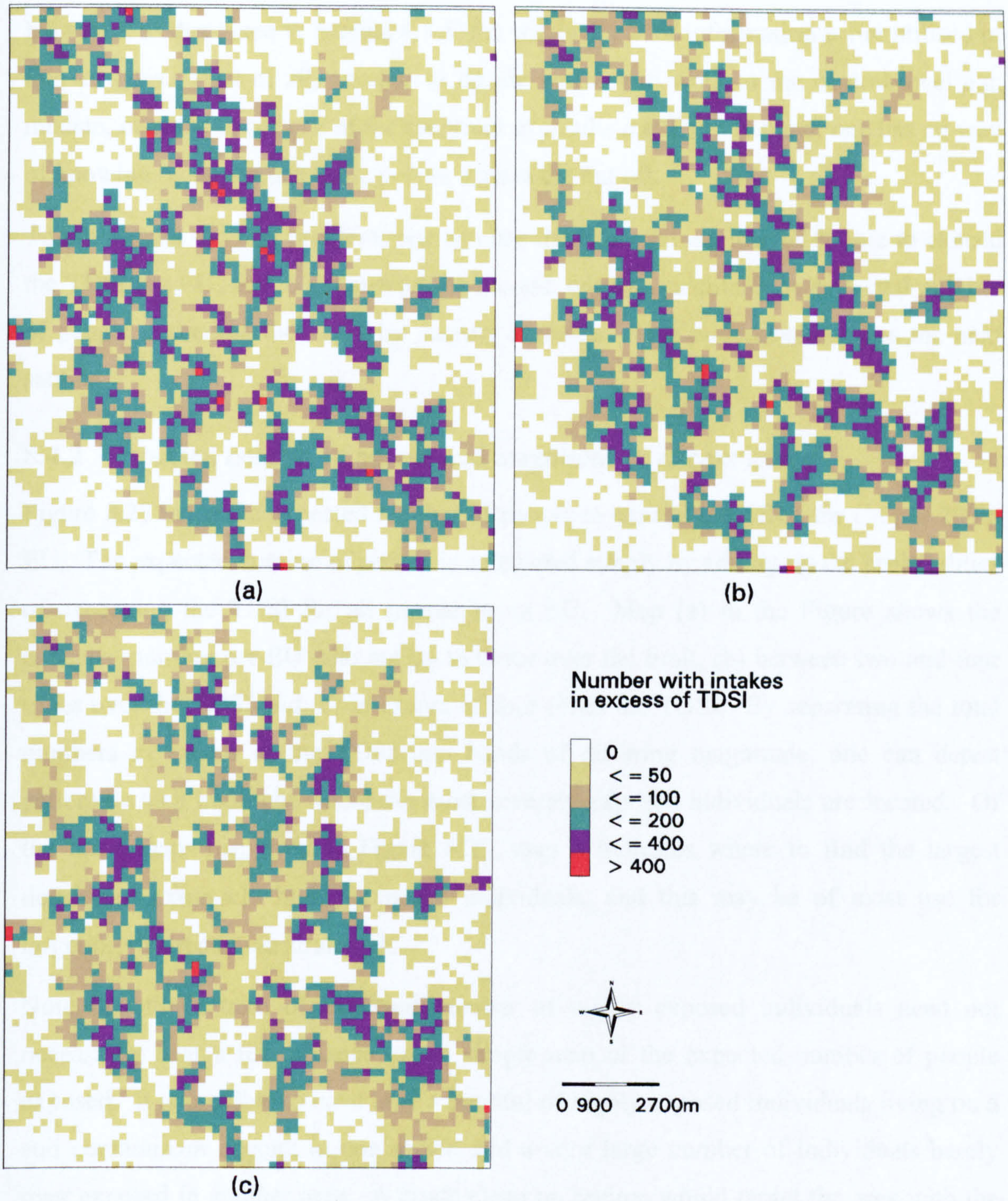


Figure 8.27. Median number of people in Stoke exceeding the TDSI for Cd per 200m cell. (a) calculated on a site wide support (Site EU) (b) calculated on a 200m support (200m EU) (c) calculated on a 50m support (50m EU). All transformed to a 200m grid for a comparative display.

Using maps such as these (Figure 8.26 and Figure 8.27) enables the risk manager to target remedial measures to the correct place. In order to reduce the risk for the most people in an efficient manner, the highest valued EUs on the map should be targeted first.

However, as discussed in Section 8.3.5, the whole depth of information is not utilised in this analysis. First of all, one has to decide whether to work at the central level e.g. median, or at the upper end of the spectrum of results e.g. P95. Second, one has no way of knowing by how much these people exceed the cut off.

To get around the first problem, one can use the expected number of people to exceed the TDSI per EU, and to get around the second, one can calculate the expected number of people exceeding the TDSI by varying amounts. This is demonstrated in the next section.

8.4.2 Expected number of people and magnitude of excess intake

Figure 8.28 maps the expected number of people to exceed the TDSI for Cd per 200m EU. The expected number of people is calculated simply by adding up the probabilities of exceeding the TDSI for all people in an EU. Map (a) in the Figure shows the expected number per EU who are up to twice over the limit, (b) between two and four times over the TDSI and (c) in excess of four times the TDSI. By separating the total numbers in excess of the TDSI into bands of differing magnitude, one can detect patterns which highlight where the most severely exposed individuals are located. Of the three maps displayed in Figure 8.28, map c indicates where to find the largest density of extremely highly exposed individuals, and this may be of most use for targeting risk reduction strategies.

Note that the pattern of expected number of highly exposed individuals need not necessarily bear a resemblance to the simple map of the expected number of people exposed. It is possible to have only a handful of highly exposed individuals living on a soil contaminant hotspot in one region and a very large number of individuals barely over exposed in another area. A crude clean up regime would target the area with the highest density of people exceeding the TDSI. A more sophisticated and ethically sound regime would target the areas where the individuals are most highly exposed.

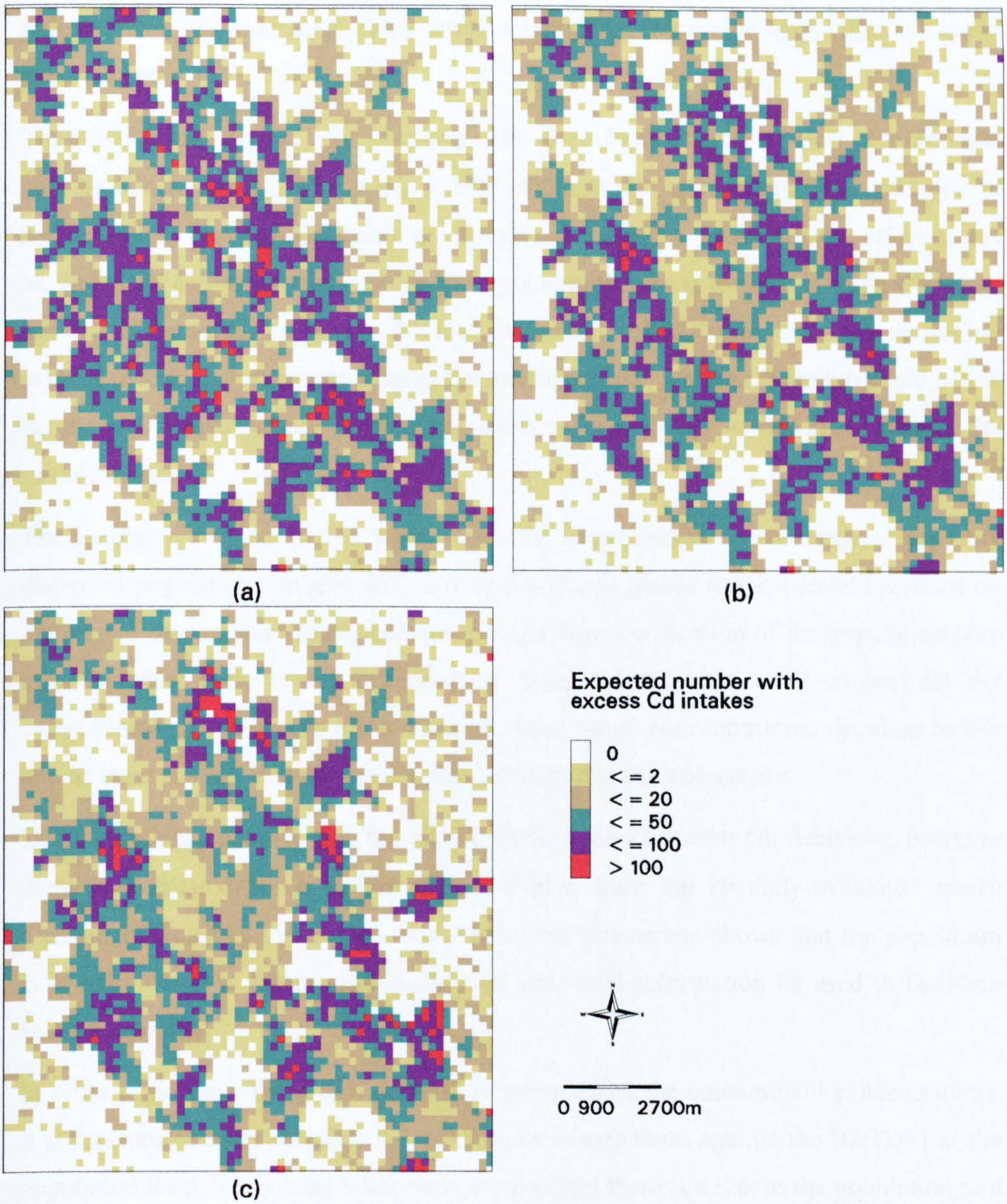


Figure 8.28. Expected number of people per 200m EU who exceed the TDSI by (a) up to 2 times (b) between 2 and 4 times (c) in excess of 4 times.

8.5 Summary

Interaction between the key components of the intake assessment model (*SE* and C_{soil}) and regulatory thresholds makes predicting the outcome of the spatially-evaluated intake assessment model difficult to predict. However, the results, once achieved, hold

a rich amount of information which can be interrogated to provide useful displays to aid risk managers.

The choice of support for working out the distributions of soil contaminant concentrations in space affects the results of the intake assessment. The variation in concentrations generally decreases as support size increases. However, the effect of this on the intake assessment depends on the regulatory threshold the risk is being evaluated against. If the cut off happens to fall high in the range of ADI estimates, then increasing support (decreasing variance) reduces the calculated risk, whereas if the threshold is low in the ADI distribution, increasing support (decreasing variance) increases the calculated risk.

The assumed location of the population can affect results. Assuming a randomly dispersed population can give different results to one placed in the correct locations on the map of contaminant concentrations. In addition, the location of the population may affect whether there is any difference in using a large or a small support for the calculations. If the people are located on mean range concentrations, deciding to use one or other support will not make much difference to the assessment.

Using regulatory cut offs, like the ID and TDSI makes for clear cut decisions, however there is a great depth of information available from the spatially-evaluated intake assessment which can be tapped into. If the risk assessment shows that the population is at risk, then it is recommended that this additional information be used to facilitate risk management.

In order to preserve the spatial correlation provided by the contaminant concentrations, it is recommended to calculate the *ADI*s and evaluate them against the ID/TDSI at the population level. Only after it has been decided that there is a risk to the population as a whole should the individual results be interrogated for additional risk management purposes.

The model is able to handle variable pH and Cd in conjunction with one other. A better indication of where in space the most likely risk from Cd will come would be gained by plotting the function $kC_{soil}e^{-bpH}$. Incorporation of variable pH affects the results of the intake assessment, and is recommended.

Spatial displays of the individual results can be produced in terms of the numbers of individuals exceeding the ID/TDSI per EU. Better still, the expected number of people

per EU can be displayed and this can be broken down into displays of slightly, moderately and highly exposed people, to facilitate an efficient, ethical risk management strategy.

A general discussion of the findings presented in the thesis follows in the next chapter.

Chapter 9 General Discussion

9.1 Introduction

After reviewing the literature, it becomes apparent that methodologies for assessing risks to human receptors from contaminated soil in a manner that preserves the spatial distribution of risks are lacking. While methodologies for incorporating spatial information are being developed, they are largely in the ecological risk field and those that assess risk to human receptors are far from complete (Chapter 4). This is not to say that they are not useful: each proposed methodology makes a contribution to building a better model for assessing human health risks, and this, too, is the aim of the research carried out in this thesis. It is far from complete in itself, but brings a new piece of information to add to the jigsaw that is risk assessment.

The proposed methodology is for a spatially-evaluated quantitative risk assessment, which preserves the spatial dimension of the results and provides a measure of uncertainty in them. Details of the procedure have been discussed throughout the preceding chapters and the discussion that follows looks generally at the advantages and disadvantages of the proposed new methodology from both a technical and a policy oriented view. Technically, the practicalities of applying the model are discussed along with an overall assessment of the type of inputs, calculations made and outputs. From a policy point of view, the usefulness of the methodology is assessed in the light of current practices, and possibilities for other useful implementations are proposed. Since the author is based in the UK, the discussion mainly concentrates on where this methodology fits into the regulatory regime in the UK. However, likely generic uses applicable to situations anywhere are also highlighted.

9.2 General applications

When the idea for the methodology was first conceived it was envisaged on a scale that related to a small development scheme, for example, a tool that could be used for assessing risks related to the redevelopment of a contaminated industrial site as a housing development. However, the data sets available for the research were on a much larger scale than anticipated, covering large areas several kilometres by several kilometres in size. This led to the realisation that a methodology could be developed for use with a much more strategic bent, to assist in regional planning and health protection. It is proposed that this methodology is flexible enough to be used both for small scale development and regional planning applications.

Land contamination, as examined in this thesis, is historic. Most of the work undertaken with regard to contaminated land is related to historic contamination, and, for example, both the UK PartIIA regulations and US Superfund regulations deal with historically contaminated land. Whether risk assessments carried out on historically contaminated sites are 'proactive' or 'reactive' is a matter of definition. On the one hand, an assessment which looks at an old industrial site, assesses the potential risks to human receptors and removes the risks prior to redevelopment could be seen as proactive. On the other hand, an assessment which calculates the risks to a population already dwelling in an area with high contaminant levels can be seen as reactive. It is proposed that this new methodology can be used in both proactive and reactive risk assessments.

9.3 Exposure assessment

Sources of uncertainty in the soil contaminant estimates were examined thoroughly in the case study presented in Chapter 5. Geostatistics provides a suite of methods suitable for assessing the concentration of a soil contaminant in unsampled locations. This is useful, since often in contaminated land assessments, samples are relatively sparse and the concentrations in large tracts of land are unknown. For the spatially-evaluated risk assessments presented in this thesis, concentrations are required all across the regions in question and geostatistics is proposed as a suitable way of producing these estimates. Though it is possible to employ simpler interpolation methods for estimating concentrations in unsampled areas, such as inverse distance weighting, geostatistics appeals because it not only interpolates, but is able to provide measurements of uncertainty as well, and this is an essential part of the new methodology.

The geostatistical procedure itself introduces uncertainties, such as through the choice of variogram model, the choice of support at which to work, and adjusting point estimates to be representative of a block size of interest. The main conclusions from the case study were as follows. The choice of variogram model is not a major source of uncertainty, so the simplest model ought to be chosen, as this will allow for speedier calculations. Kriging is quicker than simulation and ought to be used if working on a fine grid with point estimates. However, the adjustment of point kriged estimates to block estimates involves deciding which type of block correction factor to select. In this case the affine correction factor seemed more appropriate than the indirect lognormal one, in the sense that the change in results from point to block seemed most similar the changes produced in averaging small scale simulated results to block results. In the past, it has been recommended to use small scale simulations and then aggregate them to larger block sizes as necessary, to avoid needing to 'guess' the best correction factor (Deutsch & Journel, 1998).

After debating the pros and cons of various approaches, indicator methods were selected for use in this methodology. The reason for this is that they are versatile and may be adapted for use in many situations (some more complex than have been tackled in this thesis): indicator methods do not rely on dubious assumptions regarding underlying data distributions; they are suitable for use with highly skewed data; connectivity of high/low values is preserved; secondary data may be incorporated, as may uncertainty in the analytical results; uncertainty is easily displayed for results maps in terms of probabilities of exceeding regulatory thresholds; and theoretically correct confidence intervals can be plotted. The drawback of indicator methods is that they are computationally intensive and sometimes tedious to work with because of having to model multiple variograms for each assessment.

One drawback of the exposure assessment study was that it was not possible to validate the results. The pros and cons of the methods could be discussed, but the decisions made to recommend one above the other could not be validated in the field, since re-sampling of the area was not possible.

However, despite this, geostatistical exposure assessment results are potentially extremely useful in the context of risk assessment. Mapped information is a very good risk communication tool. Maps showing high and low contaminant concentrations in the region make visualisation of the problem much easier. Displayed in GIS with

infrastructure information it is possible to relate the areas estimated to be above/below regulatory guidelines to where the population of receptors lives. From these maps it is possible to perform a type of simple hazard assessment. One can say that if there is no receptor in an area then there is no risk since the source-pathway-receptor linkage is broken. This is true even if the contaminant level is extremely high. In one sense, one could argue that it is a retrospective hazard assessment. Alternatively, such maps could be of use in the strategic planning process enforced as part of the Town and Country Planning Regulations (HMSO, 1999). They could be used in a type of proactive hazard assessment whereby regional planners could use such information to avoid planning new housing projects on highly contaminated land. The planning process can be made to be as sustainable as possible through the 'fit-for-use' ethos encouraged in UK policy, which is facilitated by the use of such exposure maps. Since all estimates are subject to uncertainty, decision makers can be helped by maps of 'average' or 'high end' soil concentrations, for example, or by probability maps showing the probability of the land exceeding the SGV, all of which are relatively straightforward to produce with the recommended indicator methods.

A tiered approach to contaminated land risk assessment is recommended in the UK (DETR et al., 2000) and exposure assessment maps could be used to assist the Tier 2 generic risk assessment process. The maps are generic in the sense that they can be related to regulatory SGVs which are derived by ensuring a critical receptor would be adequately protected at that soil concentration. Plotting colour coded maps showing soils above and below the SGV will highlight the areas which pose most risk and which should be targeted for clean up so that the site may be made safe.

Much uncertainty is produced by the decision to work on one support or another. This is because working with blocks, instead of points, leads to smoothing of the soil contaminant concentrations i.e. a reduction in variance meaning that the distribution of values becomes narrower. When working with a regulatory cut off, this can have a large effect on the results of the exposure assessment, as values are pushed over or pulled below the cut off, simply by the effect of averaging. Regulatory cut offs do not, as a rule, specify a support size, and so are open to abuse: a risk assessor could choose a support size either to 'reduce' or 'increase' contaminant levels to suit his own end goal. Hence, specifying a support on which to work, in advance of the assessment, is advisable. This may be based on a suitable remediation unit size, for example.

Use of geostatistics for exposure assessment requires a competent geostatistician. Although, on the one hand, we are saying that the exposure assessment is generic, that is only in terms of the comparison of concentrations to a generic regulatory threshold. Fitting the variogram models and deciding on the moving window neighbourhood etc. must be done on a site by site basis and requires some level of competence: it is not possible to apply these methods without some degree of understanding of the subject. In practice then, a local authority wishing to use these techniques must hire the expertise to undertake them, if the skills are not already in house. In addition, users of the results should be aware of limitations when utilising them.

To recap, generic exposure assessment provides an overall picture of where the risk to the population is likely to be: soil concentration maps may be produced to highlight areas with concentrations in excess of the SGV. As already alluded to, these are most useful if displayed in GIS along with an indication of where the population is (to be) located. Following the tiered approach to risk assessment, if there are no areas on the exposure map with excessive concentrations, then since the generic SGV is designed to be protective of the most vulnerable receptor, one can assume that the area produces no risk. However, should areas show up with concentrations in excess of the SGV, a more in-depth study should be instigated. This can take the form of an intake assessment which can utilise the new spatially-evaluated probabilistic methodology.

9.4 Intake assessment

9.4.1 Major contributions of research

Since there is an apparent lack of intake models available in the UK (Butler & Petts, 2000), adaptation of the CLEA model for use as a stand alone intake assessment model for use in a Tier 3 risk assessment seems sensible. The assumptions in the model are geared towards the UK population already, so justification for its use in preference to foreign models should not be required. In addition, the model is probabilistic, unlike any of the other models mentioned in the European comparative study (Swartjes, 2002) or those apparently commonly in use in the UK (Butler & Petts, 2000), so is already an advancement on these. In addition, it can be adapted to cater for 'real' populations of mixed ages (as opposed to one critical receptor) and can be used within a spatial framework.

One of the major achievements of this research has been to highlight the importance of spatial information. The spatial dimension is important for two reasons: first,

preserving the spatial positions of contaminant concentrations in relation to the spatial positions of the receptors during intake calculations enables a better assessment of the likely contact of the receptors with the contaminants and hence their risks; and second, being able to see the information in map form aids the decision process enormously by allowing efficient targeting of resources to where they are needed most.

Another major achievement is the ability of the methodology to deal with uncertainty, variability and the spatial dimension at once. State of the art in risk assessment research is the separation of uncertainty and variability and this has been done whilst at the same time preserving the spatial nature of the data. Spatial variability and uncertainty was incorporated for the soil contaminant concentrations, but only variability and spatial location was incorporated for the receptor population. This was simply a matter of having limited time for the research. It was important not to overcomplicate the assessment so that a thorough analysis of results could be made in the time available. The model, if necessary, could easily be extended to incorporate uncertainty in the population data in the future.

A further valuable contribution made by the newly developed intake assessment model is its ability to deal with 'real' populations. This makes the model flexible, in that it can be used for a reactive assessment (i.e. calculating risks as experienced by a population already living on a contaminated site) as well as proactive (assuming a population will live on a site in the future). The reactive assessment can assess risks using the correct population densities and age structure found in the area of concern, to provide a good assessment of risks as they stand at the moment. For the proactive assessment, it is still possible to generate a theoretical population and assess the theoretical risks to them, in a similar manner as has been done in previous spatial assessments, e.g. Korre et al. (2002).

9.4.2 Working with a variable population

In Chapter 7, the CLEA intake algorithm, which was adapted for use in this new methodology, was rearranged so as to separate the soil component from the population specific components of the equation. A new variable, *SE* (SDR equivalent, $mg.kg^{-1}_{BW}.d^{-1}$), was derived which summarised the population specific components. This variable could be interrogated to discover the main features affecting its make up. The percentage of intake via each pathway for the CLEA 'critical receptor' only, is given in the relevant SGV reports. The work carried out in this thesis splits the *SE* by pathway

for each receptor age group, and is important for interpreting intake results for a mixed population.

For the three contaminants studied, As, Ni and Cd, their *SE* compositions were seen to be quite different. *SE* for As, was found to be much more dependent on intake via direct, inadvertent soil and dust ingestion. Intake through vegetable consumption was less significant, especially in the youngest age groups. Cd, on the other hand, has an *SE* which is made up almost entirely of intake through contaminant absorbed into vegetables and the contribution of direct soil ingestion is negligible. Ni falls between these two extremes. Of the vegetables consumed, potatoes were the largest contributor to *SE*.

These findings have important implications. It should be possible to recommend different risk management strategies depending upon the contaminant of concern. This is especially true in the case of a reactive risk assessment i.e. discovering a contaminant problem once the population is already installed. Uncovering a problem of high soil contaminant levels in an area already populated can be difficult to resolve. From both a cost and disruption point of view, undertaking a soil clean-up programme, by removing contaminated topsoil and replacing it with clean soil, is undesirable. However, it may be possible, armed with the knowledge of the make up of the *SE* for the contaminant of concern, to recommend a non-intrusive risk management plan, even if it is on a temporary basis until a more definitive plan is implemented. For example, if Cd is the problem, recommending (enforcing) no home-grown vegetable consumption on a housing estate would stand a good chance of reducing the risk to receptors to acceptable levels, since intake through other routes is negligible in comparison. For As, this recommendation may be ineffective, since most of the intake is through inadvertent soil and dust ingestion, so the only workable scenario may be to remove the offending top soil.

Also of note (and demonstrated in Chapter 7) is the shape of the distribution of *SE*s for a population of mixed ages. For a given population, the *SE* distribution changes according to the contaminant of concern. As, for example, produces an *SE* distribution where generally the bulk of the *SE*s are very tightly distributed. This means that a relatively small change in soil contaminant concentration could have a sudden large effect on the number of the population whose intakes are calculated to be in excess of regulatory safe levels. As C_{soil} increases slowly, the proportion of the population

exceeding the cut off will remain relatively static until the tight peak of the distribution hits the critical mark and the whole of the population suddenly spills over the cut off. Conversely, Cd, for example, has a flatter distribution, so as C_{soil} gradually increases, the rate of change of population from being not at risk to at risk is slower.

Classification of receptors' intakes as 'safe' or 'unsafe' can be performed, as in the exposure assessment, by comparing them to strict cut offs. The ID and TDSI are derived by scientists to represent intakes above which the population would be deemed to be at risk of damaging their long term health. The problem with using cut offs with a strict 'above' or 'below' classification is that the depth of available information is lost as we have no idea of knowing by how much a receptor exceeds the cut off. Thinking about the width of the *SE* distribution could prove useful. For a contaminant with a narrow *SE* distribution, as mentioned above, the population may move suddenly from 'safe' to 'unsafe' with a relatively small increase in concentration. So, a risk manager may be faced with bleak picture in the form of a map showing vast regions of population with intakes above the 'safe' level. However, it may be that all these receptors, unbeknown to him, are only marginally above the cut off. A flatter *SE* distribution gives a slower rate of change with increase in soil concentration. A risk manager may now be faced with a map showing small regions of population with intakes above the 'safe' level. However, in this case, though the situation appears to be relatively good, many of these people may have intakes which greatly exceed the cut off and whose health is in great danger. Even though there are fewer of them, because of the flatter more varied *SE* distribution, of the few that are over, many may be very much over. For this reason, it is recommended to look at the actual size of the excess, not simply at whether the receptors are over or under the threshold. This should, of course, be carried out within the tiered risk assessment approach as and when it is deemed necessary.

9.4.3 Spatial considerations

Chapter 8 presented results of intake assessments made with various example data sets. The purpose was not to produce comprehensive risk assessments for several case studies, but to highlight the differences which can result from using available data in a spatial and probabilistic manner. In this model, a 'snapshot' of intakes is calculated: the receptors are assumed to live and be exposed to the soil contaminant concentration in

their home EU for a 12month period. Their chronic intake in $mg.kg^{-1}_{BW}.d^{-1}$, *ADI*, is assessed.

A major issue which has become evident is that on conducting spatial assessments as recommended in this research, the support size on which one looks at the results becomes very important. Changes in the apparent intakes can come about simply by looking at the results with a different resolution. Similar problems have been documented before. Sexton et al. (2002) present an excellent account of the effects of spatial scale (geographic extent) and spatial resolution (level of aggregation for statistical data) on the outcome of environmental health assessments. They particularly look from the point of view of the environmental epidemiologist but some of the arguments are pertinent to the type of spatial risk assessment being proposed in this thesis. The resolution chosen for the analysis can have important consequences on the results. Looking in too much detail with too fine a resolution, one may lose the bigger picture and miss key large scale trends at work in the population. Conversely, working with a coarse resolution it is possible to blur results and miss important peaks and troughs which may be visible if one looked at a smaller resolution. This is commonly known as the 'modifiable areal unit problem'. These issues of scale and resolution must be confronted and resolved for the problem at hand.

In this research, the differences in intake results when working with different resolutions are brought about by a change in the soil contaminant concentrations. In order to work with a larger EU size, the point soil concentrations are simply averaged over the larger area and the receptor, by default, is assumed to be exposed to the average concentration across this larger EU. This is, in effect, falling into the 'trap' of not considering all the spatial information available, because we know the point where the receptor lives, which could be on a hotspot or not and make a large difference to the intake compared to working with an average value. The question of scale is then raised. If the intake assessment is at the scale of a housing estate, then it is likely that point simulations with a small resolution could be used without the need to aggregate to higher EU sizes, so none of the detailed spatial information will be lost. If working at the regional scale, then fine details will not necessarily be required as a more general picture is sought. In this case, aggregating some of the finer details and achieving a more 'averaged' result will not necessarily be problematic. After highlighting areas of concern, using the larger EUs, it may be decided to invest in additional sampling in a key region and produce a smaller resolution assessment in this area. Again, it is

emphasised that thought should be made in advance of the intake assessment as to what is the aim of the assessment and therefore what details one is prepared to 'lose' by working at a coarser resolution.

It was found (Chapter 8) that how high the soil contaminant concentrations are in relation to the regulatory soil guideline values can be used as an indication of the likely conservativeness of an intake assessment, where intakes are measured against an ID/TDSI. A quick look at the soil sample information and the SGV in advance of the assessment could contribute towards choosing a suitable support size to use for the assessment. It was found that if the mean C_{soil} < SGV, working on a larger support size generally decreased the intake estimates. Hence, if the mean of the available soil samples tends to fall below the SGV, working on a smaller support will give a more conservative assessment of the risks. If the mean C_{soil} > SGV, the pattern is less clear, but generally a good indication would be that working on a larger support would tend to give a more conservative estimate of intakes, measured in terms of the fraction of the population whose intakes are in excess of the ID/TDSI. With respect to the discussion about site and regional scale assessments, this point should be borne in mind.

The location of the population matters a great deal in this model. Because the receptors are allocated to an EU and stay there, the soil contaminant concentration in that cell only dictates their intakes. Recalling the source-pathway-receptor paradigm, if there are no receptors in an EU, there is no risk. This means that, potentially, in some areas, large swathes of land can be declared to pose no risk for the simple reason that no one lives (or will live) there. On the other hand, hotspots become very important. If receptors live on top of hotspots their intakes will potentially be very large. Hence, knowing the whereabouts of the population in space is extremely important and assuming that each receptor is exposed to the average soil contaminant concentration can have large effects on the results as demonstrated in the example in Chapter 8. Whether the effects will be to increase or decrease receptor intakes will not be known until the population location has been looked at in detail with respect to the contaminant concentrations.

9.4.4 Populations and individuals

Since we are able to quantify both variability and uncertainty in this methodology, the decision to declare a population at risk or not involves stipulating a two-dimensional decision criterion to cope with the variability and the uncertainty. Also raised in

Chapter 8, was the fact that it is possible to quantify variability and uncertainty for individuals and/or populations. In order to maintain the integrity of the spatial information, it was recommended that the results be calculated at the level of the population i.e. directly calculating, simulation by simulation (map by map), the fraction of the population with intakes in excess of the ID/TDSI. Only once it has been decided that there is a risk to the population as a whole, should the intakes of the individual receptors be interrogated further. This recommendation, again, fits comfortably within a tiered approach to risk assessment.

Having the opportunity to interrogate the results at the level of the individual opens up a rich source of information. As mentioned earlier (and demonstrated in Chapter 8) maps can be produced which highlight where the individuals with the highest excess intakes live. These areas can be prioritised in terms of clean up. Merely looking at the numbers of individuals per EU who are in excess of the ID/TDSI can result in an 'unfair' risk management plan. In this situation the problem is presented as a function of population density. Targeting the highest density areas may lead to the most efficient reduction in numbers of overly exposed individuals, and hence reduce the fraction of the population below the critical 5% mark, but it is questionable from an ethical point of view. Better, if possible, to reduce the risk for receptors with the highest intakes.

9.4.5 Practical considerations

The methodology, as it stands at the moment, is not a stand alone system. It is a collection of calculations and procedures made in a variety of software packages, which only the author has carried out. It is difficult to envisage the development of a user friendly package which relatively novice users could use, due to the nature of the calculations required e.g. it is necessary to make subjective judgements in so many stages (geostatistics, EU size, decision criteria etc.). However, it is possible to envisage guidance for the use of such methods, as has been done in the US, e.g. USEPA (2001). The various parts can be used within the framework of a tiered risk assessment and, should future validation suggest that they are useful methods, guidance for their consistent use would be recommended.

9.5 Effective Cd

A novel approach to assessing the intakes of Cd is proposed in this thesis. A new variable called 'effective Cd' concentration was derived. This was in recognition of the fact that Cd intake is sensitive to changes in pH: intake increases with increasing Cd and

decreasing pH. Merely looking at the pattern of Cd concentrations may be insufficient. Using the SGVs produced by the CLEA model, for example, one has to assume a uniform pH for the area of concern. This is not realistic, and since pH has been demonstrated (in Chapter 8) to have a big influence on the intake assessment results, even as part of a generic risk assessment, it is felt unacceptable to look purely at Cd concentrations. The 'effective Cd' variable should better reflect the areas of concern. To enable a spatially-evaluated tiered risk assessment to be undertaken for Cd, it is recommended that an SGV for 'effective Cd' be derived, and that this be used to judge risks from Cd contamination.

9.6 Limitations of the work

Because the spatial element can make such a difference to the results, the author feels that it is important that it be included in analyses in the future. However, the author also recognises the limitations of the methodology.

From a methodological point of view, the assumption that the receptor never moves from its home EU is simplistic. Time did not allow investigation of the effects of receptors spending time in both their 'home EU' and their 'work EU', for example. It is recommended that this research be carried out as a next step in method development, by updating the model to include receptor movements. Histograms of the fraction of the population exceeding the ID/TDSI could be produced easily, but mapping the results of a mobile population for risk communication would be more complex. One option would be to map the intake for the individual to its home EU, no matter where the individual picked up the contaminant. Alternatively, one could map the intakes to the source cell by summing the amount of intake received by all the individuals visiting that cell. This would be useful for prioritising areas for clean up.

A further, possible, limitation of the method is that it considers a 'snapshot' of intakes expected over a one year period only to assess the chronic intakes. In this snapshot, the population remains in place and the concentration of the soil contaminant is constant. It is possible that the population would change and the contaminant concentrations would change, by natural attenuation for example. However, boundaries for the study must be set, and the one year snapshot was the chosen boundary: the author still feels that this snapshot is a useful indicative tool. In the future, the effect of changing soil concentrations could be looked into if deemed appropriate.

Another limitation is that the new methodology deals with the risks from one contaminant at a time, and for metallic/metalloid non-volatile substances only. It may be possible to amalgamate the results for more than one contaminant by deriving a type of 'equivalents' system (e.g. express the intakes in metal equivalents) but this is difficult. As discussed in DEFRA & EA (2002b) combining the risks calculated from different elements is tricky, since effects are not necessarily additive and do not necessarily affect the same bodily functions.

The focus of this thesis has been on method development, and though differences in intakes have been demonstrated as a result of including definitive population locations and maps of soil contaminant levels etc., it has not been possible to validate the results. This is a major gap. It is recommended that future work should address this issue by making further such comparative studies with the express intention of validating them to discover whether this method is an improvement over less complex methods.

Finally, it is recognised that all the effort has been channelled into calculating exposures and intakes without considering in depth the original source of the soil contaminant. In reality, this should be considered for remediation purposes and possible prevention of future contamination. It may be that the 'contamination' is a result of natural high background concentrations due to underlying geology. Some authors have undertaken work to 'fingerprint' contaminants and discover their origins (Carlon et al., 2001; Facchinelli et al., 2001; Korre, 1999a; 1999b) and it may be possible to link such work with exposure and intake assessments in the future to refine the information provided.

Chapter 10 Conclusions and recommendations

10.1 Introduction

After reviewing the literature, it becomes apparent that there is a lack of methodologies for assessing risk to human health from contaminated land, even though there is a potential demand for them. The demand comes as a result of regulations which encourage a risk-based approach to contaminated land assessment within a 'fit-for-use' ethos. A tiered approach to risk assessment encourages initial simple generic methods which should be replaced by more complex site specific methods should the situation merit it. Of the methods that are available, most are deterministic and do not take account of the spatial distribution of the risks.

The aim of this research, therefore, was to produce a methodology for quantifying risks posed to human health, by exposure to contaminated land, in a manner which preserves the spatial distribution of the risks and gives a measure of uncertainty in the estimates.

The main objectives were:

- to find the most appropriate geostatistical tools for spatial assessment of land contaminants;
- to identify a suitable intake model and adapt this model for use in a spatially-evaluated, probabilistic methodology;
- and to produce meaningful presentations of results for use in risk communication.

These main objectives were accomplished, and during the process spawned other useful achievements, described below. In addition, useful insights into the use of spatial data

in risk assessments were produced. Despite the accomplishments, limitations are recognised and recommendations are made for future work to overcome them.

10.2 Conclusions

10.2.1 Completion of main objectives

- Indicator methods are useful for characterising soil contaminant levels since they are versatile and may be adapted for use in many situations: indicator methods do not rely on dubious assumptions regarding underlying data distributions; they are suitable for use with highly skewed data; connectivity of high/low values is preserved; secondary data may be incorporated; uncertainty in the analytical results may be incorporated; uncertainty is easily displayed for results maps in terms of probabilities of exceeding regulatory thresholds; and theoretically correct confidence intervals can be plotted. The drawback of indicator methods is that they are computationally intensive and sometimes tedious to work with because of having to model multiple variograms for each assessment.

- The CLEA model was identified as suitable for use in the risk assessment process: the model is already probabilistic so has many input distributions already assigned; the model is UK based (as is the author) and its use is already embedded in UK regulations. It was successfully adapted from its intended use (for back-calculating generic Soil Guideline Values based on lifetime average intakes of a critical receptor) for use in calculating probabilistic intakes for a mixed population located in space: intakes were calculated per person depending on age – this gives a measure of variability in the population; chronic intakes were calculated for a 'snapshot' period of one year; intakes for individuals were calculated with respect to the soil contaminant concentration in their home location; 1,000 simulations of soil contaminant concentration were produced to give 1,000 estimates of intake for each individual – this gives a measure of uncertainty in the results.

- Useful graphic summaries of results from the 1,000 spatially-evaluated simulations were produced: box and whisker plots show variation in intakes for a given (selection of) population; histograms display uncertainty in results for the fraction of the population exposed in excess of regulatory 'safe' levels; maps show areas of high/low soil contaminant concentration with respect to the

location of the population; maps show regions where the greatest number of people exceed the 'safe' intake level; maps show where the individuals with the greatest excess intakes live.

The new methodology is versatile in that: it can be used in both a proactive and reactive manner; it may be used on local or regional scales; it may be used with real populations or hypothetical ones; use of indicator methods allows incorporation of additional information, if available, to aid the exposure assessment process; and uncertainty for populations may easily be incorporated if required.

10.2.2 Additional novel achievements

- As a result of developing the methodology and interrogating the results, a new variable, *SE*, (SDR equivalent, $mg.kg^{-1}_{BW}.d^{-1}$) was derived which represents an amalgamation of all the population specific variables in the intake equation. *SE* can be a useful tool: composition of *SE* is different for each contaminant and knowledge of the composition can help inform risk management plans; shape of the *SE* distribution for a population can help indicate the likely rate at which the population moves from being below to above a regulatory 'safe' intake standard.
- Intake of Cd is pH dependent. It was felt unacceptable to assume constant pH, therefore the intake model was adapted to take into account both variable Cd and variable pH levels in soil. A new variable 'effective Cd' was derived to be indicative of regions of likely high Cd intake: effective Cd = $kC_{soil}e^{-bpH}$ (where k and b are constants). Effective Cd should be used in maps to highlight likely regions of high Cd intake.

10.2.3 Insights into probabilistic spatially-evaluated risk assessments

- Interaction of C_{soil} distribution, *SE* and regulatory threshold is complex:
  - if mean C_{soil} < SGV, generally, as C_{soil} variation increases, the proportion of the population in excess of the ID/TDSI increases. This indicates that the most conservative support size would be a small one;
  - if mean C_{soil} > SGV, generally, as C_{soil} variation increases, the proportion of the population in excess of the TDSI decreases. This indicates that the most conservative support size would be a large one;

- if ID/TDSI falls towards the bottom end of the intake distribution, as C_{soil} variation increases, the proportion of the population in excess of the ID/TDSI decreases;
- if ID/TDSI falls towards the top end of the intake distribution, as C_{soil} variation increases, the proportion of the population in excess of the ID/TDSI increases.

- Location of the population affects the intake assessment results:
  - a randomly located population may have a different intake distribution to a properly located one, especially if there are hotspots of contamination;
  - the decision to work on one support or another (and hence change the C_{soil} distribution) may or may not effect the intake results, depending on whether the population is located in hotspot areas or on average range values.
- Working with strict regulatory cut offs leads to loss of information:
  - an exposure assessment shows where the soil exceeds the SGV, but not by how much;
  - an intake assessment shows how many people exceed the ID/TDSI, but not by how much.
- Using intake results from individuals reveals a depth of information otherwise hidden:
  - One can assess by how much each individual exceeds the ID/TDSI;
  - One can map the expected number of individuals in each EU who exceed the ID/TDSI;
  - One can stratify and map the results according to by how much the individuals exceed the ID/TDSI.

10.2.4 Policy implications

- The spatially-evaluated exposure assessment would:
  - sit in UK policy as Tier 2 generic risk assessment procedure;
  - be a useful tool for regional planning;
  - be applicable as a US tiered risk assessment tool.

- The spatially-evaluated intake assessment would:
  - sit in UK policy as a Tier 3 site-specific risk assessment tool;
  - be applicable as a US tiered risk assessment tool.

10.3 Limitations

- The results have not been validated.
- The methodology is still open to improvement for the following reasons:
  - only one contaminant is addressed at a time i.e. mixtures of contaminants are not catered for;
  - it is only applicable to non-volatile metals/metalloids;
  - the population is assumed to be static in one place for the duration of the exposure period;
  - soil contaminant concentration is assumed constant for the duration of the exposure period.

10.4 Recommendations for future work

The most important piece of work would be a validation study. This thesis has clearly demonstrated the differences which occur when treating data in a spatial way, however, it has not been possible to validate the results and conclude whether this methodology is a significant improvement over previous ones. Studies should be carried out with the express intention of validating the results.

Should the validation studies suggest that the results are acceptable, it is recommended that a continuation of this research would be useful. In terms of method development, possibilities for improvement include: population movements; mixtures of contaminants; changing concentrations over time; and non-metallic contaminants.

References

Armstrong, A. (1998). *Basic linear geostatistics*. Berlin Heidelberg: Springer-Verlag.

Armstrong, M., & Champigny, N. (1989). A Study on Kriging Small Blocks. *Cim Bulletin, 82*(923), 128-133.

Bailar, J. C., & Bailer, A. J. (1999). Risk assessment - the mother of all uncertainties - Disciplinary perspectives on uncertainty in risk assessment. *Annals of the New York Academy of Sciences, 895*, 273-285.

Baird, S. J. S., Cohen, J. T., Graham, J. D., Shlyakhter, A. I., & Evans, J. S. (1996). Noncancer risk assessment: A probabilistic alternative to current practice. *Human and Ecological Risk Assessment, 2*(1), 79-102.

Benamghar, A., & Sonnet, P. (1999). Performance comparison of cumulative and class indicator approaches for pollution risk assessment. In J. Gomez-Hernandez & A. Soares & R. Froidevaux (Eds.), *geoENV II - Geostatistics for environmental applications* (pp. 357-368). Dordrecht: Kluwer Academic Publishers.

British Geological Survey. (2000). *Regional geochemistry of Wales and part of west-central England: stream sediment and soil*. Keyworth, Nottingham: British Geological Survey.

Bogaert, P. (1999). Assessing the variability of the variogram estimator. In J. Gomez-Hernandez & A. Soares & R. Froidevaux (Eds.), *geoENV II - Geostatistics for environmental applications* (pp. 479-490). Dordrecht: Kluwer Academic Publishers.

Burmaster, D. E., & Thompson, K. M. (1997). Estimating exposure point concentrations for surface soils for use in deterministic and probabilistic risk assessments. *Human and Ecological Risk Assessment, 3*(3), 363-384.

Busing, J. (2001). Contaminated land research under the EU RTD programme 'Environment and Sustainable Development'. *Land Contamination and Reclamation, 9*(1), 55-70.

Butler, B., & Petts, J. (2000). *Land contamination risk assessment tools. An evaluation of some of the commonly used methods*. (R&D technical report P260). Bristol: Environment Agency.

Buxton, B. E., Wells, D. E., & Diluise, G. (1997). Comparison of three kriging methods for making soil remediation decisions. In E. Y. Baafi & N. A. Schofield (Eds.), *Geostatistics Wollongong '96* (Vol. 2, pp. 984-995). Dordrecht: Kluwer Academic Publishers.

Carlon, C., Critto, A., Marcomini, A., & Nathanail, P. (2001). Risk based characterisation of contaminated industrial site using multivariate and geostatistical tools. *Environmental Pollution, 111*(3), 417-427.

Chica-Olmo, M., & Luque-Espinar, J. A. (2002). Applications of the local estimation of the probability distribution function in environmental sciences by kriging methods. *Inverse Problems, 18*(1), 25-36.

Chiles, J.-P., & Delfiner, P. (1999). *Geostatistics: Modelling spatial uncertainty*. New York: John Wiley and Sons, Inc.

Clifford, P. A., Barchers, D. E., Ludwig, D. F., Sielken, R. L., Klingensmith, J. S., Graham, R. V., & Banton, M. I. (1995). An Approach to Quantifying Spatial Components of Exposure for Ecological Risk Assessment. *Environmental Toxicology and Chemistry, 14*(5), 895-906.

Coleman, M. E., & Marks, H. M. (1999). Qualitative and quantitative risk assessment. *Food Control, 10*(4-5), 289-297.

Commission of the European Communities. (1993). Commission Directive 93/67/EEC of 20 July 1993 laying down the principles for assessment of risks to man and the environment of substances in accordance with Council Directive 67/548/EEC. *Official Journal, L227*(08/09/1993), 9-18.

Council of the European Union. (1999). Council Directive 1999/31/EC of 26 April 1999 on the landfill of waste. *Official Journal, L182*(16/07/1999), 1-19.

Council of the European Union. (2000a). Council Directive 96/61/EC of 24 September 1996 concerning integrated pollution prevention and control [Official Journal L 257 of 10.10.1996]. *Official Journal, L257*(10/10/1996).

Council of the European Union. (2000b). Directive 2000/60/EC of the European Parliament and the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. *Official Journal, L327*(22/12/2000), 1-72.

Covello, V. T., & Merkhofer, M. W. (1993). *Risk assessment methods. Approaches for assessing health and environmental risks*. New York: Plenum Press.

Dakins, M. E., Toll, J. E., & Small, M. J. (1994). Risk-Based Environmental Remediation - Decision Framework and Role of Uncertainty. *Environmental Toxicology and Chemistry, 13*(12), 1907-1915.

DEFRA. (2004). *The first soil action plan for England: 2004-2006* (PB9411). London: Department for Environment, Food and Rural Affairs.

DEFRA & EA. (2002a). *Assessment of risks to human health from land contamination: An overview of the development of soil guideline values and related research* (R&D publication CLR7). Bristol: Department for Environment, Food & Rural Affairs and the Environment Agency.

DEFRA & EA. (2002b). *Contaminants in soil: Collation of toxicological data and intake values for humans* (R&D publication CLR9). Bristol: Department for Environment, Food & Rural Affairs and the Environment Agency.

DEFRA & EA. (2002c). *Contaminants in soil: Collation of toxicological data and intake values for humans. Arsenic* (R&D publication TOX1). Bristol: Department for Environment, Food & Rural Affairs and the Environment Agency.

DEFRA & EA. (2002d). *Contaminants in soil: Collation of toxicological data and intake values for humans. Cadmium* (R&D publication TOX3). Bristol: Department for Environment, Food & Rural Affairs and the Environment Agency.

DEFRA & EA. (2002e). *Contaminants in soil: Collation of toxicological data and intake values for humans. Lead* (R&D publication TOX6). Bristol: Department for Environment, Food & Rural Affairs and the Environment Agency.

DEFRA & EA. (2002f). *Contaminants in soil: Collation of toxicological data and intake values for humans. Nickel* (R&D publication TOX8). Bristol: Department for Environment, Food & Rural Affairs and the Environment Agency.

DEFRA & EA. (2002g). *The contaminated land exposure assessment model (CLEA): Technical basis and algorithms* (R&D publication CLR10). Bristol: Department for Environment, Food & Rural Affairs and the Environment Agency.

DEFRA & EA. (2002h). *Potential contaminants for the assessment of land* (R&D publication CLR8). Bristol: Department for Environment, Food & Rural Affairs and the Environment Agency.

DEFRA & EA. (2002i). *Soil guideline values for arsenic contamination* (R&D publication SGV1). Bristol: Department for Environment, Food & Rural Affairs and the Environment Agency.

DEFRA & EA. (2002j). *Soil guideline values for cadmium contamination* (R&D publication SGV3). Bristol: Department for Environment, Food & Rural Affairs and the Environment Agency.

DEFRA & EA. (2002k). *Soil guideline values for lead contamination* (R&D publication SGV10). Bristol: Department for Environment, Food & Rural Affairs and the Environment Agency.

DEFRA & EA. (2002l). *Soil guideline values for nickel contamination* (R&D publication SGV7). Bristol: Department for Environment, Food & Rural Affairs and the Environment Agency.

Delignette-Muller, M. L., & Rosso, L. (2000). Biological variability and exposure assessment. *International Journal of Food Microbiology, 58*(3), 203-212.

Department of Health. (1998). *Health survey for England, 1998* (Data set 4150). UK Data Archive, University of Essex, www.data-archive.ac.uk.

DETR. (1999). *Contaminated land: Implementation of Part IIA of the Environmental Protection Act 1990. Draft DETR circular*. London: Department of the Environment Transport and the Regions.

DETR, Environment Agency, & Institute for Environment and Health. (2000). *Guidelines for environmental risk assessment and management*. London: The Stationery Office.

Deutsch, C. V. (1997). Direct assessment of local accuracy and precision. In E. Y. Baafi & N. A. Schofield (Eds.), *Geostatistics Wollongong '96* (Vol. 1, pp. 115-125). Dordrecht: Kluwer Academic Publishers.

Deutsch, C. V., & Journel, A. G. (1992). *GSLIB Geostatistical software library and user's guide*. New York: Oxford University Press, Inc.

Deutsch, C. V., & Journel, A. G. (1998). *GSLIB Geostatistical software library and user's guide* (2nd ed.). New York: Oxford University Press, Inc.

ECETOC. (2001). *Exposure Factors Sourcebook for European Populations (with Focus on UK Data)* (Technical Report No.79). Brussels: European Centre for Ecotoxicology and Toxicology of Chemicals.

European Parliament. (2002). *European Parliament resolution on the Commission communication 'Towards a Thematic Strategy for Soil Protection' (COM(2002) 179 - C5-0328/2002 - 2002/2172(COS))*.

Facchinelli, A., Sacchi, E., & Mallen, L. (2001). Multivariate statistical and GIS-based approach to identify heavy metal sources in soils. *Environmental Pollution, 114*, 313-324.

Faechner, T., Pyrcz, M., & Deutsch, C. V. (2000). Soil remediation decision making in presence of uncertainty in crop yield response. *Geoderma, 97*(1-2), 21-38.

Fairman, R., Mead, C. D., & Williams, W. P. (Eds.). (1998). *Environmental risk assessment. Approaches, experiences and information sources*. Copenhagan: European Environment Agency.

Ferguson, C., Darmendrail, D., Freier, K., Jensen, B. K., Jensen, J., Kasamas, H., Urzelai, A., & Vegter, J. (Eds.). (1998). *Risk assessment for contaminated sites in Europe; Volume 1. Scientific Basis*. Nottingham: LQM Press.

Ferguson, C. C. (1999). Assessing risks from contaminated sites: policy and practice in 16 European countries. *Land Contamination & Reclamation, 7*(2), 33-54.

Geovariances. (2002). Isatis, Version 4.0.6 (Version 4.0.6). 77212 Avon Cedex, France: Geovariances.

Ginevan, M. E., & Splitstone, D. E. (1997). Improving remediation decisions at hazardous waste sites with risk-based geostatistical analysis. *Environmental Science & Technology, 31*(2), A92-A96.

Gómez-Hernández, J. J. (1997). Issues on environmental risk assessment. In E. Y. Baafi & N. A. Schofield (Eds.), *Geostatistics Wollongong '96* (Vol. 1, pp. 15-26). Dordrecht: Kluwer Academic Publishers.

Goovaerts, P. (1997a). Accounting for local uncertainty in environmental decision-making processes. In E. Y. Baafi & N. A. Schofield (Eds.), *Geostatistics Wollongong '96* (Vol. 2, pp. 929-940). Dordrecht: Kluwer Academic Publishers.

Goovaerts, P. (1997b). *Geostatistics for natural resources evaluation*. New York: Oxford University Press.

Goovaerts, P. (1997c). Kriging vs stochastic simulation for risk analysis in soil contamination. In A. Soares & J. Gomez-Hernandez & R. Froidevaux (Eds.), *geoENV I - Geostatistics for environmental applications* (pp. 247-258). Dordrecht: Kluwer Academic Publishers.

Goovaerts, P. (1999). Geostatistics in soil science: state-of-the-art and perspectives. *Geoderma, 89*(1-2), 1-45.

Goovaerts, P. (2000). Estimation or simulation of soil properties? An optimization problem with conflicting criteria. *Geoderma, 97*(3-4), 165-186.

Goovaerts, P. (2001). Geostatistical modelling of uncertainty in soil science. *Geoderma, 103*(1-2), 3-26.

Gotway, C. A., & Rutherford, B. M. (1994). Stochastic simulation for imaging spatial uncertainty: comparison and evaluation of available algorithms. In A. Armstrong & P. A. Dowd (Eds.), *Geostatistical Simulation* (pp. 1-21): Kluwer Academic Publishers.

Gray, G. M., Jeffery, W. G., & Merchant, G. E. (1997). *Risk assessment and risk management of non-ferrous metals. Realizing the benefits and managing the risks*. Ottawa: International Council on Metals and the Environment.

Griffin, S., Goodrum, P. E., Diamond, G. L., Meylan, W., Brattin, W. J., & Hassett, J. M. (1999). Application of a probabilistic risk assessment methodology to a lead smelter site. *Human and Ecological Risk Assessment, 5*(4), 845-868.

Harris, M., & Herbert, S. (1994). *ICE design and practice guide. Contaminated land. Investigation, assessment and remediation*. London: Thomas Telford.

Hattis, D., & Burmaster, D. E. (1994). Assessment of variability and uncertainty distributions for practical risk analysis. *Risk Analysis, 14*(5), 713-730.

HESP. (1995). Human Exposure to Soil Pollutants. Version 2.10b (Version 2.10b). The Hague: Shell Internationale Petroleum Maatschappij B.V.

HMSO. (1999). The Town and Country Planning (Development Plan) (England) Regulations 1999, *Statutory Instrument 1999 No. 3280*: Her Majesty's Stationary Office.

Hoffman, F. O., & Hammonds, J. S. (1994). Propagation of uncertainty in risk assessments: The need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability. *Risk Analysis, 14*(5), 707-712.

Hope, B. K. (1999). Assessment of risk to terrestrial receptors using uncertain analysis - A case study. *Human and Ecological Risk Assessment, 5*(1), 145-170.

Hope, B. K. (2000). Generating probabilistic spatially-explicit individual and population exposure estimates for ecological risk assessments. *Risk Analysis, 20*(5), 573-589.

Hope, B. K. (2001). A case study comparing static and spatially explicit ecological exposure analysis methods. *Risk Analysis, 21*(6), 1001-1010.

ICRCL. (1987). *Guidance on the assessment and redevelopment of contaminated land* (ICRCL guidance note 59/83). London: Interdepartmental Committee on the Redevelopment of Contaminated Land, Department of the Environment.

Isaaks, E. H., & Srivastava, R. M. (1989). *Applied geostatistics*. New York: Oxford University Press.

Isukapalli, S. S., Roy, A., & Georgopoulos, P. G. (1998). Stochastic Response Surface Methods (SRSMs) for uncertainty propagation: Application to environmental and biological systems. *Risk Analysis, 18*(3), 351-363.

Jimenez-Beltran, D. (1998). Preface. In R. Fairman & C. D. Mead & W. P. Williams (Eds.), *Environmental risk assessment. Approaches, experiences and information sources*. Copenhagan: European Environment Agency.

Journel, A. G. (1983). Nonparametric-Estimation of Spatial Distributions. *Journal of the International Association for Mathematical Geology, 15*(3), 445-468.

Kelly, E. J., & Campbell, K. (2000). Separating variability and uncertainty in environmental risk assessment - Making choices. *Human and Ecological Risk Assessment, 6*(1), 1-13.

Kodell, R. L., & Gaylor, D. W. (1999). Combining uncertainty factors in deriving human exposure levels of noncarcinogenic toxicants, *Uncertainty in the Risk Assessment of Environmental and Occupational Hazards* (Vol. 895, pp. 188-195). New York: New York Acad Sciences.

Kooistra, L., Leuven, R., Nienhuis, P. H., Wehrens, R., & Buydens, L. M. C. (2001). A procedure for incorporating spatial variability in ecological risk assessment of Dutch River floodplains. *Environmental Management, 28*(3), 359-373.

Korre, A. (1997). *A methodology for the statistical and spatial analysis of soil contamination in GIS.* Imperial College London.

Korre, A. (1999a). Statistical and spatial assessment of soil heavy metal contamination in areas of poorly recorded, complex sources of pollution Part 1: factor analysis for contamination assessment. *Stochastic Environmental Research and Risk Assessment, 13*(4), 260-287.

Korre, A. (1999b). Statistical and spatial assessment of soil heavy metal contamination in areas of poorly recorded, complex sources of pollution Part 2: Canonical correlation analysis and GIS for the assessment of contamination sources. *Stochastic Environmental Research and Risk Assessment, 13*(4), 288-316.

Korre, A., Durucan, S., & Koutroumani, A. (2002). Quantitative-spatial assessment of the risks associated with high Pb loads in soils around Lavrio, Greece. *Applied Geochemistry, 17*(8), 1029-1045.

Landis, W. G., Moore, D. R. J., & Norton, S. B. (1998). Ecological risk assessment: looking in, looking out. In P. E. T. Douben (Ed.), *Pollution risk assessment and management* (pp. 464). Chichester: John Wiley & Sons.

Lapen, D. R., Topp, G. C., Hayhoe, H. N., Gregorich, E. G., & Curnoe, W. E. (2001). Stochastic simulation of soil strength/compaction and assessment of corn yield risk using threshold probability patterns. *Geoderma, 104*(3-4), 325-343.

Leisenring, W., & Ryan, L. (1992). Statistical Properties of the Noael. *Regulatory Toxicology and Pharmacology, 15*(2), 161-171.

Leonte, D., & Schofield, N. (1996). Evaluation of a soil contaminated site and clean-up criteria: a geostatistical approach. In R. M. Srivastava & S. Rouhani & M. V. Cromer & A. J. Desbarats & A. I. Johnson (Eds.), *Geostatistics for Environmental and Geotechnical Applications, ASTM STP 1283* (pp. 133-145): American Society for Testing and Materials.

Lijzen, J. P. A., Baars, A. J., Otte, P. F., Rikken, M. G. J., Swartjes, F. A., Verbruggen, E. M. J., & van Wezel, A. P. (2001). *Technical evaluation of the intervention values for soil/sediment and groundwater* (RIVM 711701023). Bilthoven: RIVM.

Lijzen, J. P. A., & Swartjes, F. A. (2000). *Soil-use specific remediation objectives.* Paper presented at the Proceedings of the seventh international FZK/TNO conference on contaminated soil, 18-22 September, 2000, Leipzig, Germany.

Linkov, I., Grebenkov, A., & Baitchorov, V. M. (2001). Spatially explicit exposure models: application to military sites. *Toxicology and Industrial Health, 17*(5-10), 230-235.

Lloyd, C. D., & Atkinson, P. M. (2001). Assessing uncertainty in estimates with ordinary and indicator kriging. *Computers & Geosciences, 27*(8), 929-937.

Marinussen, M., & vanderZee, S. (1996). Conceptual approach to estimating the effect of home-range size on the exposure of organisms to spatially variable soil contamination. *Ecological Modelling, 87*(1-3), 83-89.

McKenna, S. A. (1998). Geostatistical approach for managing uncertainty in environmental remediation of contaminated soils: Case study. *Environmental & Engineering Geoscience, 4*(2), 175-184.

Mohammadi, J., Van Miervenne, M., & Goovaerts, P. (1997). Mapping Cadmium concentration and the risk of exceeding a local sanitation threshold using indicator geostastistics. In A. Soares & J. Gomez-Hernandez & R. Froidevaux (Eds.), *geoENV I - Geostatistics for environmental applications* (pp. 327-337). Dordrecht: Kluwer Academic Publishers.

Moore, D. R. J., & Elliott, B. (1996). Should uncertainty be quantified in human and ecological risk assessments used for Decision-Making? *Human and Ecological Risk Assessment, 2*(1), 11-24.

Mowrer, H. T. (1997). Propagating uncertainty through spatial estimation processes for old-growth subalpine forests using sequential Gaussian simulation in GIS. *Ecological Modelling, 98*(1), 73-86.

Murray, C. J., Lee, H. J., & Hampton, M. A. (2002). Geostatistical mapping of effluent-affected sediment distribution on the Palos Verdes shelf. *Continental Shelf Research, 22*(6-7), 881-897.

Myers, J. C. (1997). *Geostatistical error management:quantifying uncertainty for environmental sampling and mapping*. New York: Van Nostrand Reinhold.

Nathanail, C. P., Furguson, C. C., Brown, M. J., & Hooker, P. J. (1998). *A geostatistical approach to spatial risk assessment of lead in urban soils to assist planners.* Paper presented at the Proceedings of the 8th international congress, International Association for Engineering Geology and the Environment, Rotterdam.

National Research Council. (1983). *Risk assessment in the Federal Government: Managing the process*. Washington DC: National Academy Press.

Nauta, M. J. (2000). Separation of uncertainty and variability in quantitative microbial risk assessment models. *International Journal of Food Microbiology, 57*(1-2), 9-18.

OUP. (1998). *The new Oxford dictionary of English*. Oxford: Oxford University Press.

Pannatier, Y. (1995). *Variowin. Software for spatial data analysis in 2D*. New York: Springer-Verlag.

Papritz, A., & Dubois, J. P. (1999). Mapping heavy metals in soil by (non-)linear kriging: an empirical validation. In J. Gomez-Hernandez & A. Soares & R. Froidevaux (Eds.), *geoENV II - Geostatistics for environmental applications* (pp. 429-440). Dordrecht: Kluwer Academic Publishers.

Petts, J., Cairney, T., & Smith, M. (1997). *Risk-based contaminated land investigation and assessment*. Chichester: John Wiley and Sons.

Pilz, J., Spoeck, G., & Schimek, M. G. (1997). Taking account of uncertainty in spatial covariance estimation. In E. Y. Baafi & N. A. Schofield (Eds.), *Geostatistics Wollongong '96* (Vol. 1, pp. 302-313). Dordrecht: Kluwer Academic Publishers.

Rai, S. N., & Krewski, D. (1998). Uncertainty and variability analysis in multiplicative risk models. *Risk Analysis, 18*(1), 37-45.

Rai, S. N., Krewski, D., & Bartlett, S. (1996). A general framework for the analysis of uncertainty and variability in risk assessment. *Human and Ecological Risk Assessment, 2*(4), 972-989.

Rautman, C. A. (1997). Geostatistics and cost-effective environmental remediation. In E. Y. Baafi & N. A. Schofield (Eds.), *Geostatistics Wollongong '96* (Vol. 2, pp. 941-950). Dordrecht: Kluwer Academic Publishers.

Richardson, G. M. (1996). Deterministic versus probabilistic risk assessment: Strengths and weaknesses in a regulatory context. *Human and Ecological Risk Assessment, 2*(1), 44-54.

Roberts, S. M. (1999). Practical issues in the use of probabilistic risk assessment. *Human and Ecological Risk Assessment, 5*(4), 729-736.

Rossi, R. E., Borth, P. W., & Tollefson, J. J. (1993). Stochastic Simulation for Characterizing Ecological Spatial Patterns and Appraising Risk. *Ecological Applications, 3*(4), 719-735.

Rouhani, S. (1996). Geostatistical estimation: Kriging. In R. M. Srivastava & S. Rouhani & M. V. Cromer & A. J. Desbarats & A. I. Johnson (Eds.), *Geostatistics for Environmental and Geotechnical Applications, ASTM STP 1283* (pp. 20-31): American Society for Testing and Materials.

Safai-Naraghi, K., & Marcotte, D. (1997). Bootstrapping variograms. In E. Y. Baafi & N. A. Schofield (Eds.), *Geostatistics Wollongong '96* (Vol. 1, pp. 188-199). Dordrecht: Kluwer Academic Publishers.

Saito, H., & Goovaerts, P. (2000). Geostatistical interpolation of positively skewed and censored data in a dioxin-contaminated site. *Environmental Science & Technology, 34*(19), 4228-4235.

Sexton, K., Waller, L. A., McMaster, R. B., Maldonado, G., & Adgate, J. L. (2002). The importance of spatial effects for environmental health policy and research. *Human and Ecological Risk Assessment, 8*(1), 109-125.

Sielken, R. L., & Valdez-Flores, C. (1999). Probabilistic risk assessment's use of trees and distributions to reflect uncertainty and variability and to overcome the limitations of default assumptions. *Environment International, 25*(6-7), 755-772.

Slob, W., & Pieters, M. N. (1998). A probabilistic approach for deriving acceptable human intake limits and human health risks from toxicological studies: General framework. *Risk Analysis, 18*(6), 787-798.

Smit, P. J., & Goldsborough, D. G. (1998). VLIER-HUMAAN: A human exposure model embedded in Flemish Legislation, *Contaminated Soil '98* (pp. 1045-1046). London: Thomas Telford.

Smith, J. L., Halvorson, J. J., & Papendick, R. I. (1993). Using Multiple-Variable Indicator Kriging for Evaluating Soil Quality. *Soil Science Society of America Journal, 57*(3), 743-749.

Smith, R. L. (1994). Use of a Monte-Carlo Simulation for Human Exposure Assessment at the Superfund Site. *Risk Analysis, 14*(4), 433-439.

Srivastava, R. M. (1996). Describing spatial variability using geostatistical analysis. In R. M. Srivastava & S. Rouhani & M. V. Cromer & A. J. Desbarats & A. I. Johnson (Eds.), *Geostatistics for Environmental and Geotechnical Applications, ASTM STP 1283* (pp. 13-19): American Society for Testing and Materials.

Swartjes, F. A. (1999). Risk-based assessment of soil and groundwater quality in the Netherlands: Standards and remediation urgency. *Risk Analysis, 19*(6), 1235-1249.

Swartjes, F. A. (2001). Human exposure model comparison study: state of play. *Land Contamination and Reclamation, 9*(1), 101-106.

Swartjes, F. A. (2002). *Variation in calculated human exposure. Comparison of calculations with seven European human exposure models* (RIVM 711701030). Bilthoven: RIVM.

Swartout, J. C., Price, P. S., Dourson, M. L., Carlson-Lynch, H. L., & Keenan, R. E. (1998). A probabilistic framework for the reference dose (probabilistic RfD). *Risk Analysis, 18*(3), 271-282.

Tristan, E., Demetriades, A., Ramsey, M. H., Rosenbaum, M. S., Stavrakis, P., Thornton, I., Vassiliades, E., & Vergou, K. (2000). Spatially resolved hazard and exposure assessments: An example of lead in soil at Lavrion, Greece. *Environmental Research, 82*(1), 33-45.

Tromans, S., & Turrall-Clarke, R. (2000). *Contaminated land. The new regime.* London: Sweet & Maxwell.

USEPA. (1989). *Risk Assessment Guidance for Superfund (RAGS), Volume I -- Human Health Evaluation Manual, Part A*: United States Environmental Protection Agency.

USEPA. (1997a). *Exposure Factors Handbook.* (EPA/600/P-95/002 Fa.). Washington DC, USA.: United States Environmental Protection Agency.

USEPA. (1997b). *Guiding principles for Monte Carlo analysis* (EPA/603/R-97/001). Washington, DC, USA: Risk Assessment Forum, United States Environmental Protection Agency.

USEPA. (1997c). *Policy for use of probabilistic analysis in risk assessment at the US Environmental Protection Agency*: United States Environmental Protection Agency.

USEPA. (1999). *Risk assessment guidance for Superfund volume 3 part A: Process for conducting probabilistic risk assessment (RAGS 3A) (DRAFT)*: United States Environmental Protection Agency.

USEPA. (2001). *Risk Assessment Guidance for Superfund (RAGS) Volume 3 Part A: Process for conducting probabilistic risk assessment*: United States Environmental Protection Agency.

Van Miervenne, M., & Goovaerts, P. (2001). Evaluating the probability of exceeding a site-specific soil cadmium contamination threshold. *Geoderma, 102*, 75-100.

Vegter, J. J. (1998). Developments of risk assessment and risk management, *Contaminated Soil '98* (pp. 291-299). London: Thomas Telford.

Walker, D. D., & Loftis, J. C. (1997). Alternative spatial estimators for ground-water and soil measurements. *Ground Water, 35*(4), 593-601.

Wang, G. X., Gertner, G., Singh, V., Shinkareva, S., Parysow, P., & Anderson, A. (2002). Spatial and temporal prediction and uncertainty of soil loss using the revised universal soil loss equation: a case study of the rainfall-runoff erosivity R factor. *Ecological Modelling, 153*(1-2), 143-155.

Webster, R. (2000). Is soil variation random? *Geoderma, 97*(3-4), 149-163.

Webster, R., & Oliver, M. A. (1989). Optimal Interpolation and Isarithmic Mapping of Soil Properties .6. Disjunctive Kriging and Mapping the Conditional-Probability. *Journal of Soil Science, 40*(3), 497-512.

Webster, R., & Oliver, M. A. (2001). *Geostatistics for Environmental Scientists*. Chichester: John Wiley & Sons, Ltd.

Yates, S. R., & Yates, M. V. (1988). Disjunctive Kriging as an Approach to Management Decision- Making. *Soil Science Society of America Journal, 52*(6), 1554-1558.

Appendix A Geostatistical procedure for Arsenic in Stoke on Trent

A.1 Simple statistics

Simple statistics of the 745 topsoil samples are given in Table A 1 and a histogram of the distribution in Figure A 1.

Table A 1. Simple statistics of the As topsoil sample concentrations at a depth of circa 15cm ($mg.kg^{-1}$).

Mean	Std.dev	Min	P1	P5	P10	P25	P50	P75	P90	P95	P99	Max
16.15	10.12	2	5	8	9	11	14	18	25	31	60	136

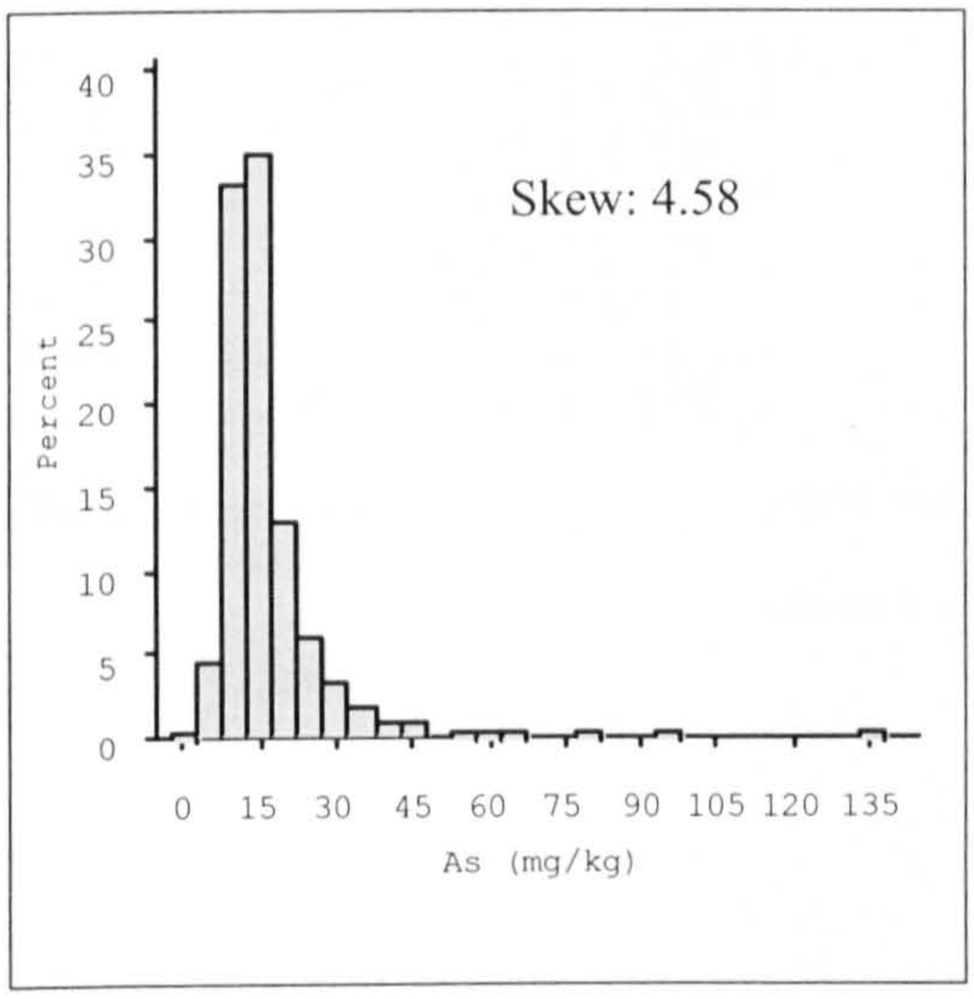


Figure A 1. Histogram of the As soil sample values.

A.2 Variogram modelling

Exploratory data analysis was undertaken using Isatis (Geovariances, 2002). Eight indicator cut offs were used for modelling the As data. In this case, setting the cut offs to the exact decile values of the sample data was difficult since the As values are

recorded to the nearest $mg.kg^{-1}$ value, are very repetitive and so difficult to rank. Therefore, the eight cut offs were spread as best as possible which was dictated by the recorded values, in some cases only $1mg.kg^{-1}$ apart. The following cut offs were modelled: 9, 11, 13, 14, 15, 17, 20 and $25mg.kg^{-1}$, which represent the 10^{th}, 22^{nd}, 38^{th}, 46^{th}, 54^{th}, 69^{th}, 80^{th} and 89^{th} percentiles respectively. To avoid order relation problems, model sills were fixed at the theoretical variances, p(1-p), where p is the percentile at that cut off.

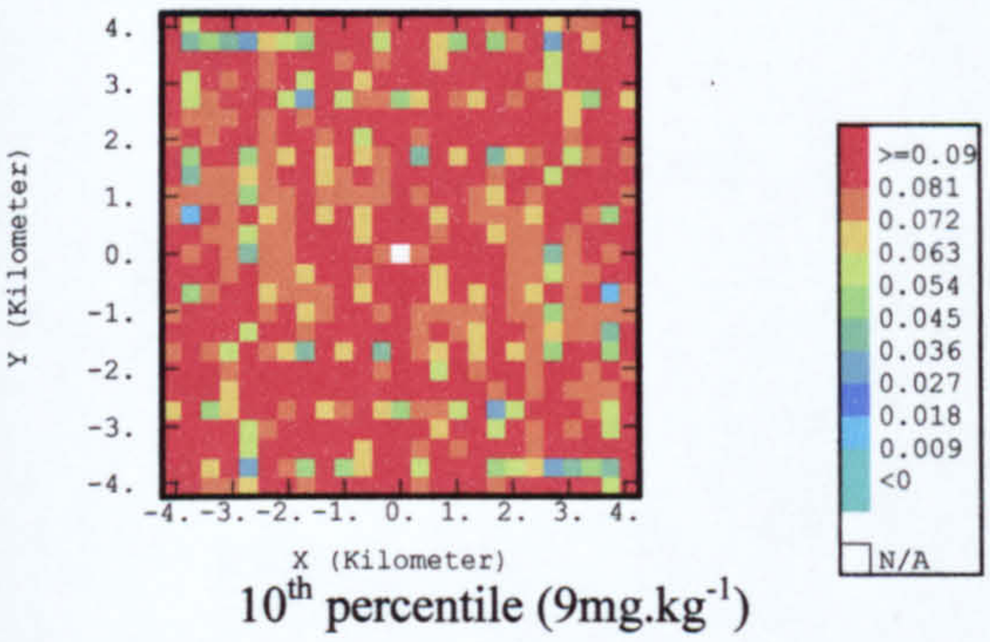
Variogram maps were examined for signs of anisotropy Figure A 2 but none was found, so the variograms were modelled isotropically. Little structure could be found; the most was discerned using a lag spacing of 340m and a lag tolerance of 50%. This lack of structure is probably because of many different human influences on the soil in a very mixed urban environment which includes much made-ground. The sample and model variograms are shown in Figure A 3 and parameters of the models in Table A 2. Exponential models with a relatively high nugget were fitted.

Table A 2. Parameters of the indicator variogram models for As.

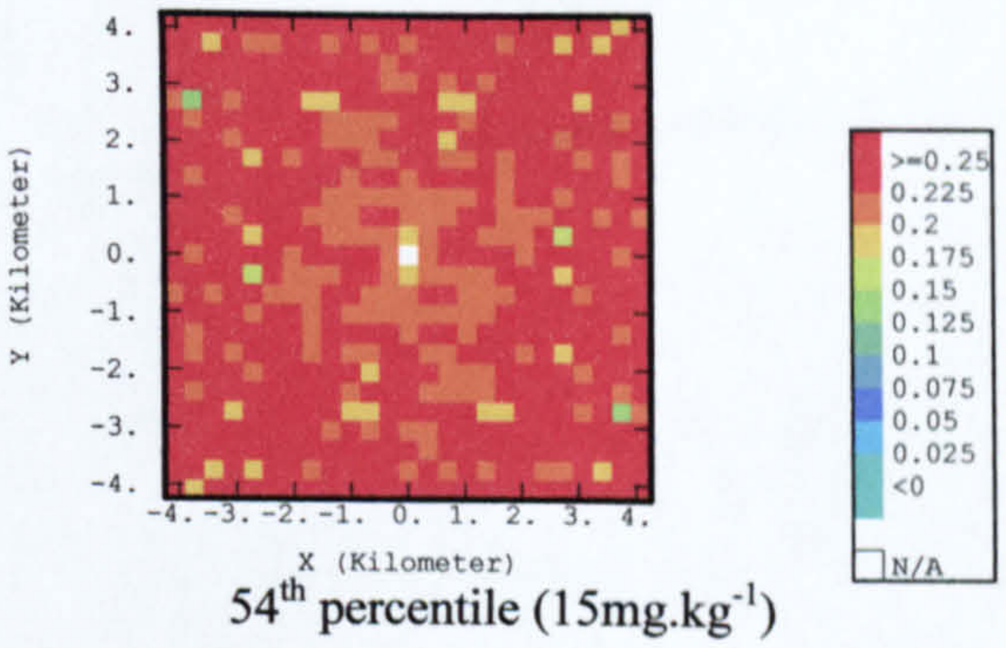
Cut off ($mg.kg^{-1}$)	Exponential*		nugget C_0
	a (m)	c	
9	6000	0.015	0.075
11	4000	0.025	0.145
13	5000	0.050	0.190
14	3000	0.050	0.190
15	3500	0.050	0.190
17	2500	0.035	0.175
20	1300	0.030	0.130
25	1000	0.010	0.090

*Exponential model: $$\gamma(h) = C_0 + c.\left[1 - \exp\left(-\frac{3h}{a}\right)\right]$$

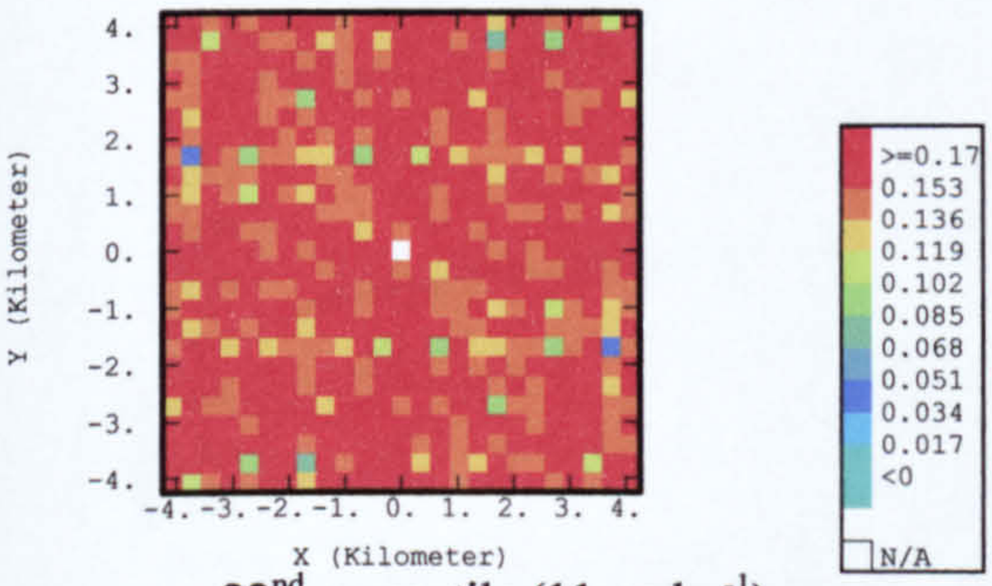
where h is the lag distance



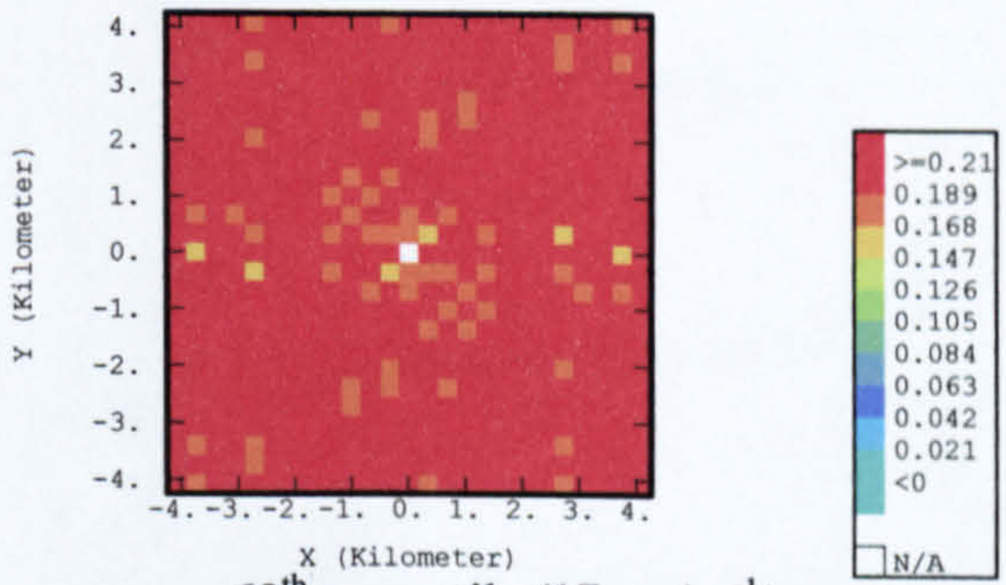
10th percentile (9mg.kg-1)



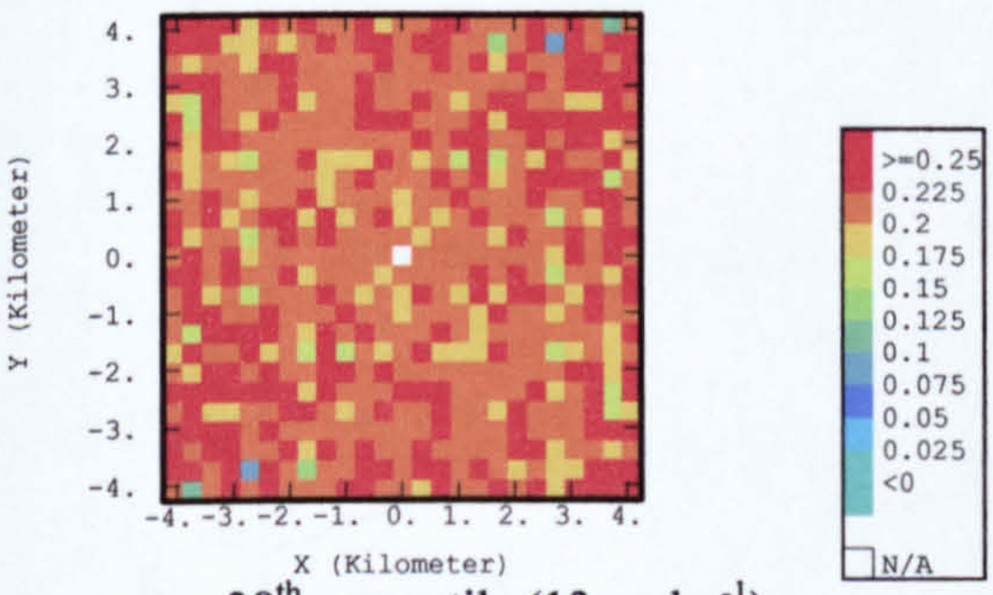
54th percentile (15mg.kg-1)



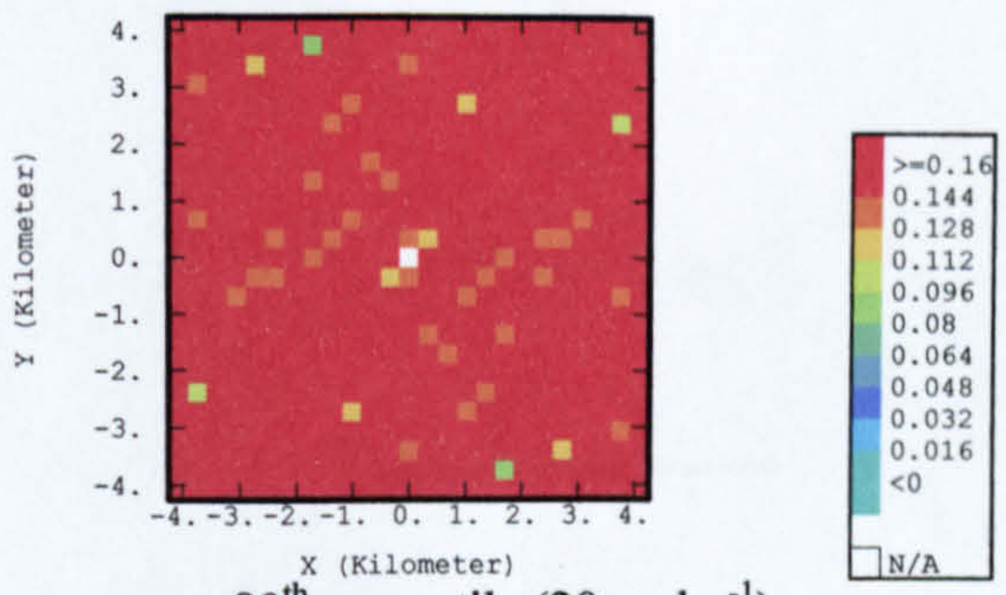
22nd percentile (11mg.kg-1)



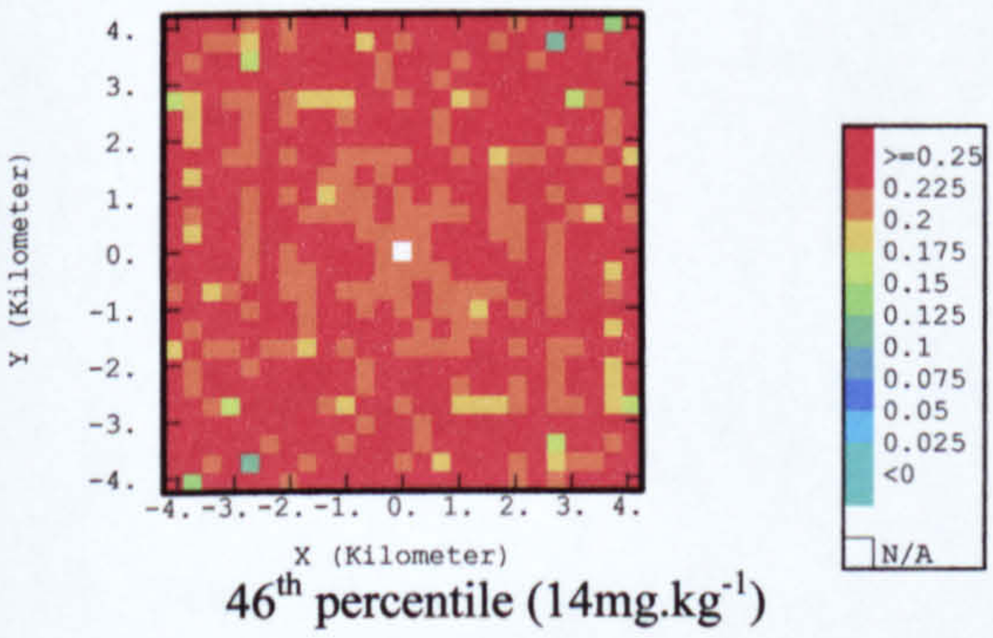
69th percentile (17mg.kg-1)



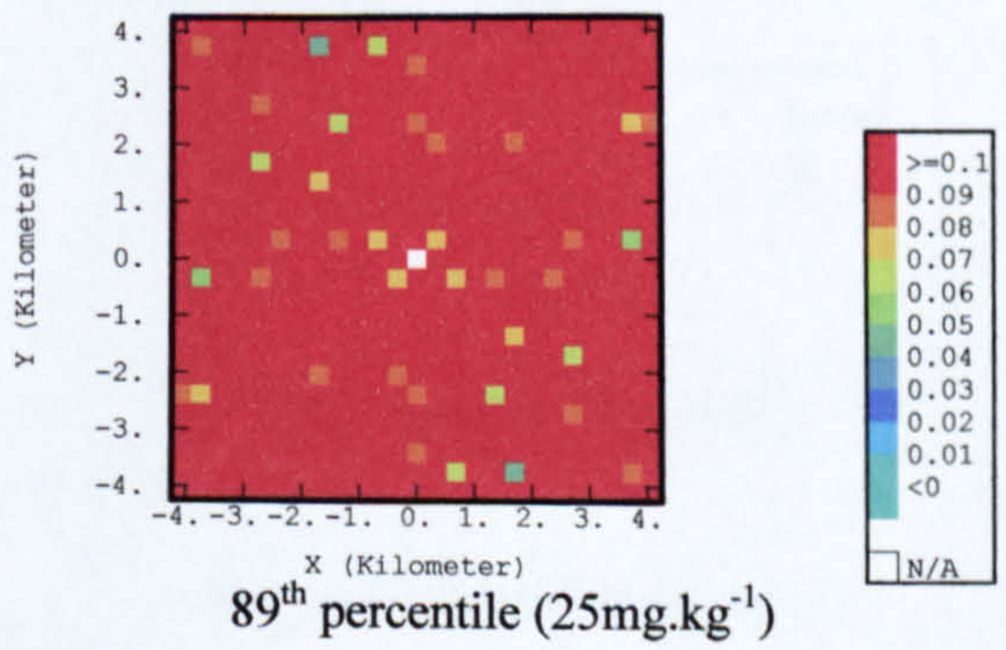
38th percentile (13mg.kg-1)



80th percentile (20mg.kg-1)

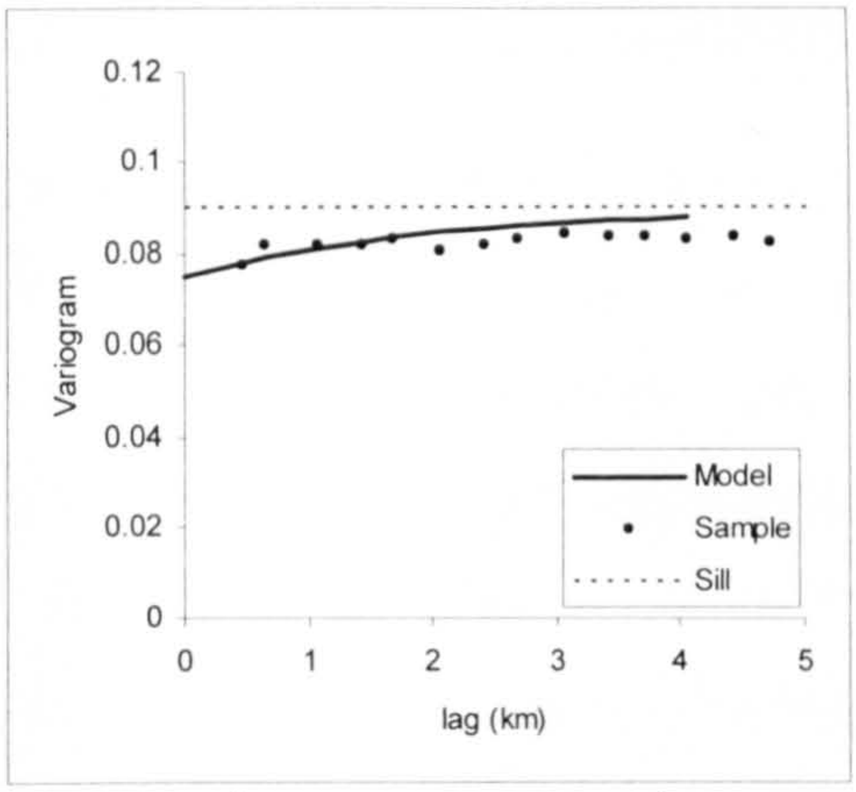


46th percentile (14mg.kg-1)

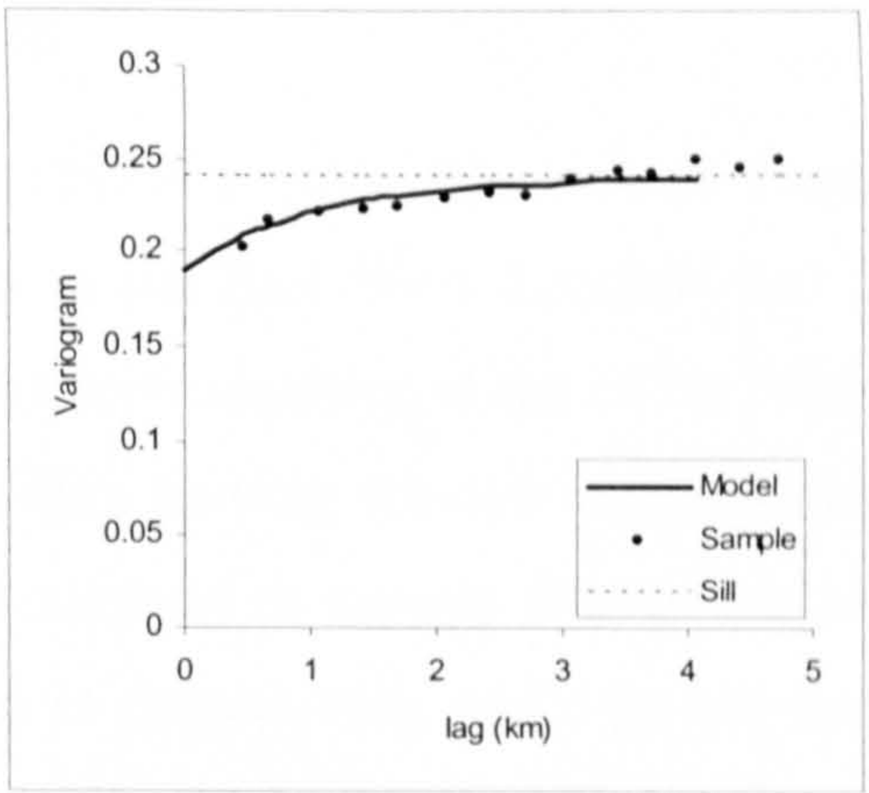


89th percentile (25mg.kg-1)

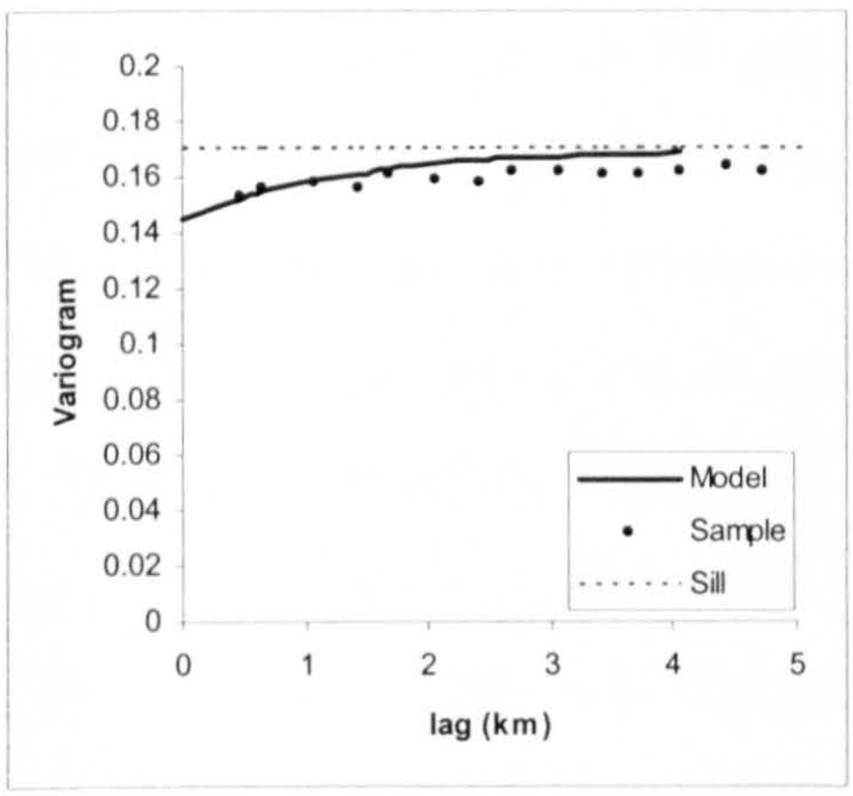
Figure A 2. Indicator variogram maps for As at eight cut offs.



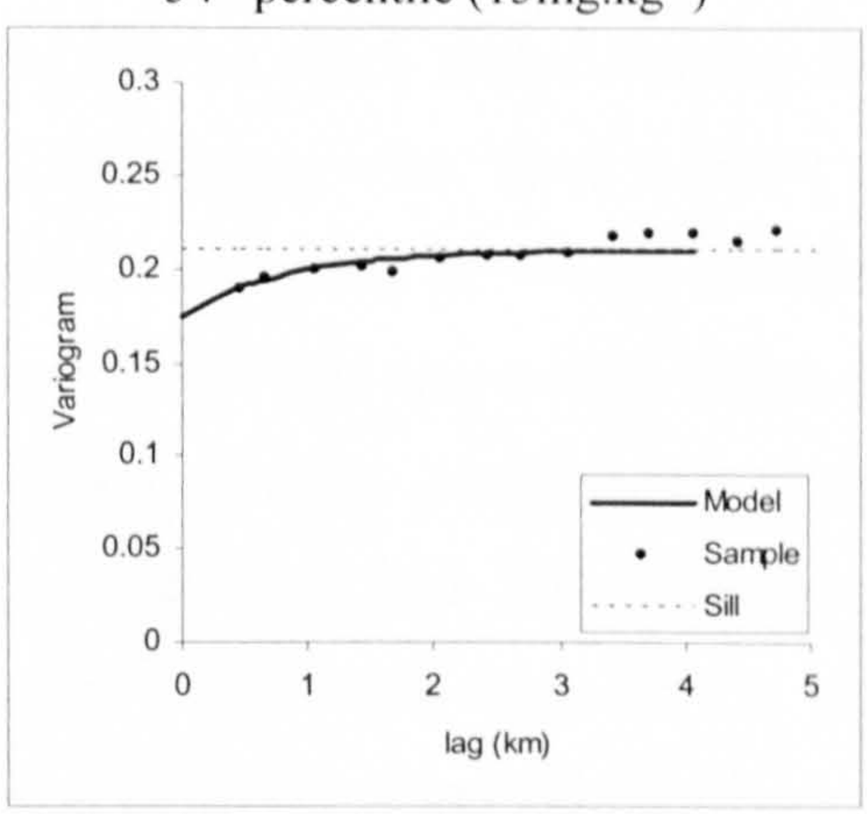
10th percentile (9mg.kg-1)



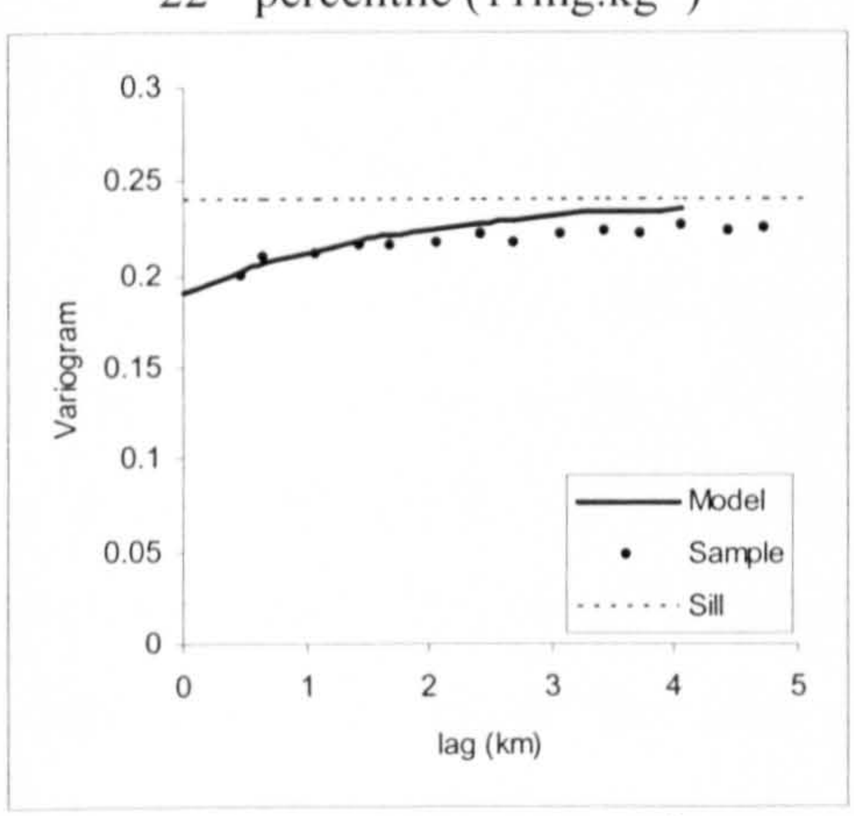
54th percentile (15mg.kg-1)



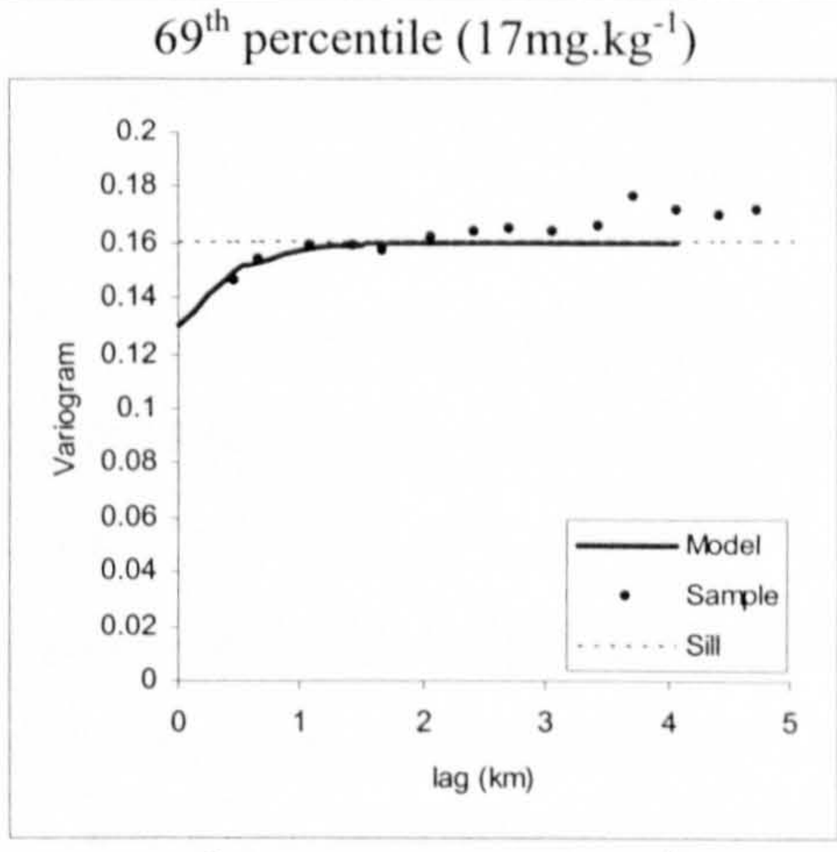
22nd percentile (11mg.kg-1)



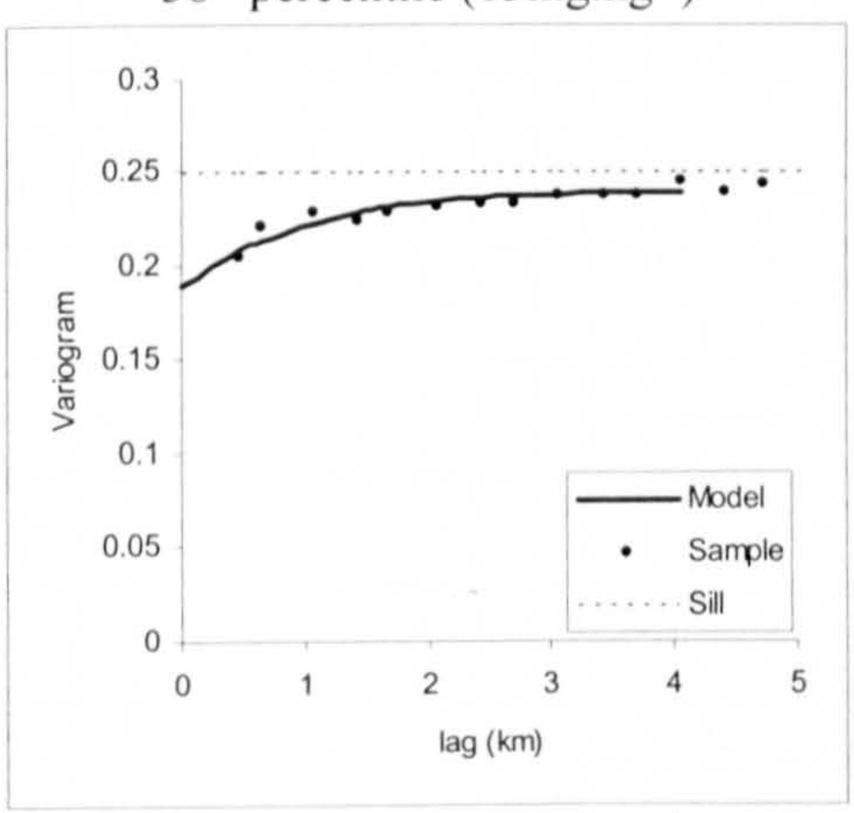
69th percentile (17mg.kg-1)



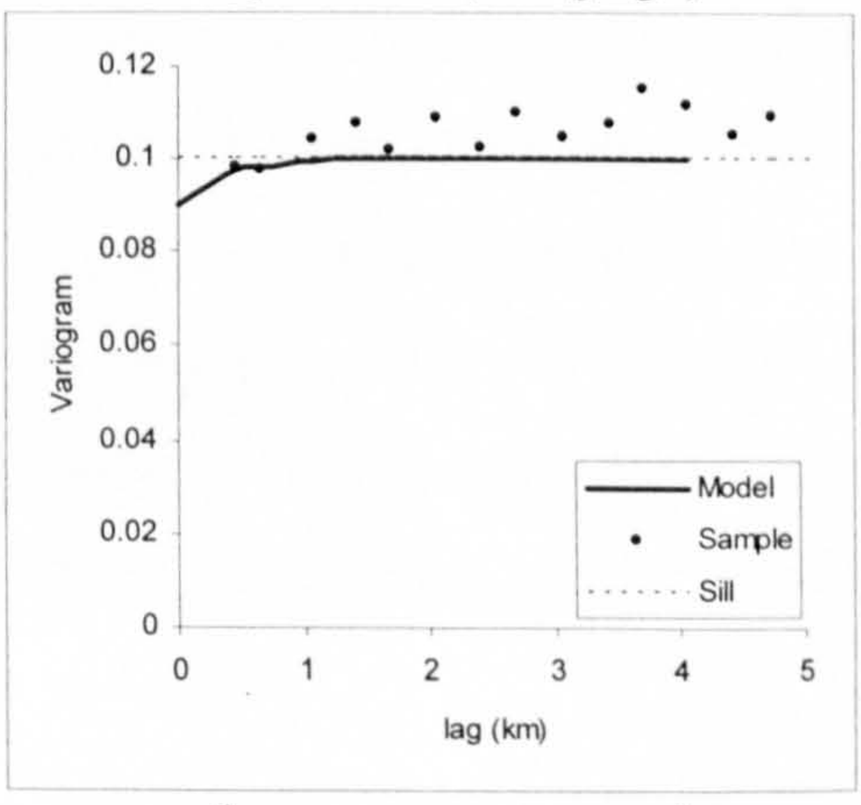
38th percentile (13mg.kg-1)



80th percentile (20mg.kg-1)



46th percentile (14mg.kg-1)



89th percentile (25mg.kg-1)

Figure A 3. Sample and model indicator variograms for As.

A.3 Running and validating simulations

Simulations were carried out using the GSLIB SISIM routine (Deutsch & Journel, 1998) on a 50mx50m square grid with 272 cells in the East-West direction and 312 North-South. Each simulated value was taken to be representative of the entire 50x50m square. The ordinary kriging routine was used with a moving window neighbourhood of radius 4,000m. The minimum and maximum numbers of sample data points to be used in simulation of a grid point were set to 0 and 16 respectively, while the number of previously simulated nodes was set to 16. These parameters ensure that before any nodes are simulated, up to 16 sample data points will be used, and thereafter, once simulated nodes are available, they will be taken into account also, up to a 50:50 mix between sample data and previously simulated nodes. Using the previously simulated nodes ensures the joint ccdf is reproduced and using sample data assists reproduction of the sample histogram. To speed up the simulation process, sample data can be optionally allocated to grid nodes prior to simulation, and kept fixed throughout. However, this was deemed inappropriate since every simulated value at that point would be the same with no uncertainty and could lead to 'holes' in probability maps at a later stage i.e. that grid point will always have a probability of 0 or 1 of being over a threshold, no matter what the results around it.

On building the ccdfs for each cell, the minimum As value was set to 10% less and the maximum to 10% more than the minimum and maximum sample values respectively (i.e. to 1.8 and 150mg.kg^{-1}), assuming that it was unlikely that the maximum value would have been found in sampling. The upper and lower tails of the ccdf must be extrapolated from the highest cut off to the maximum and the lowest cut off to the minimum using models. Initially, small batches of sequential indicator simulations (SIS) were carried out with different extrapolation parameters for the upper and lower tails of the ccdfs. Selection of the final models was based on the best reproduction of the sample histogram (i.e. by comparing the statistics of the samples with those of the simulated values). This was found with a lower tail consisting of a power model with parameter 3.5 and the upper tail with a hyperbolic model with parameter 2.5. Figure A 4 shows a plot of the sample value statistics versus the average of the simulated statistics, and Table A 3 summarises the statistics.

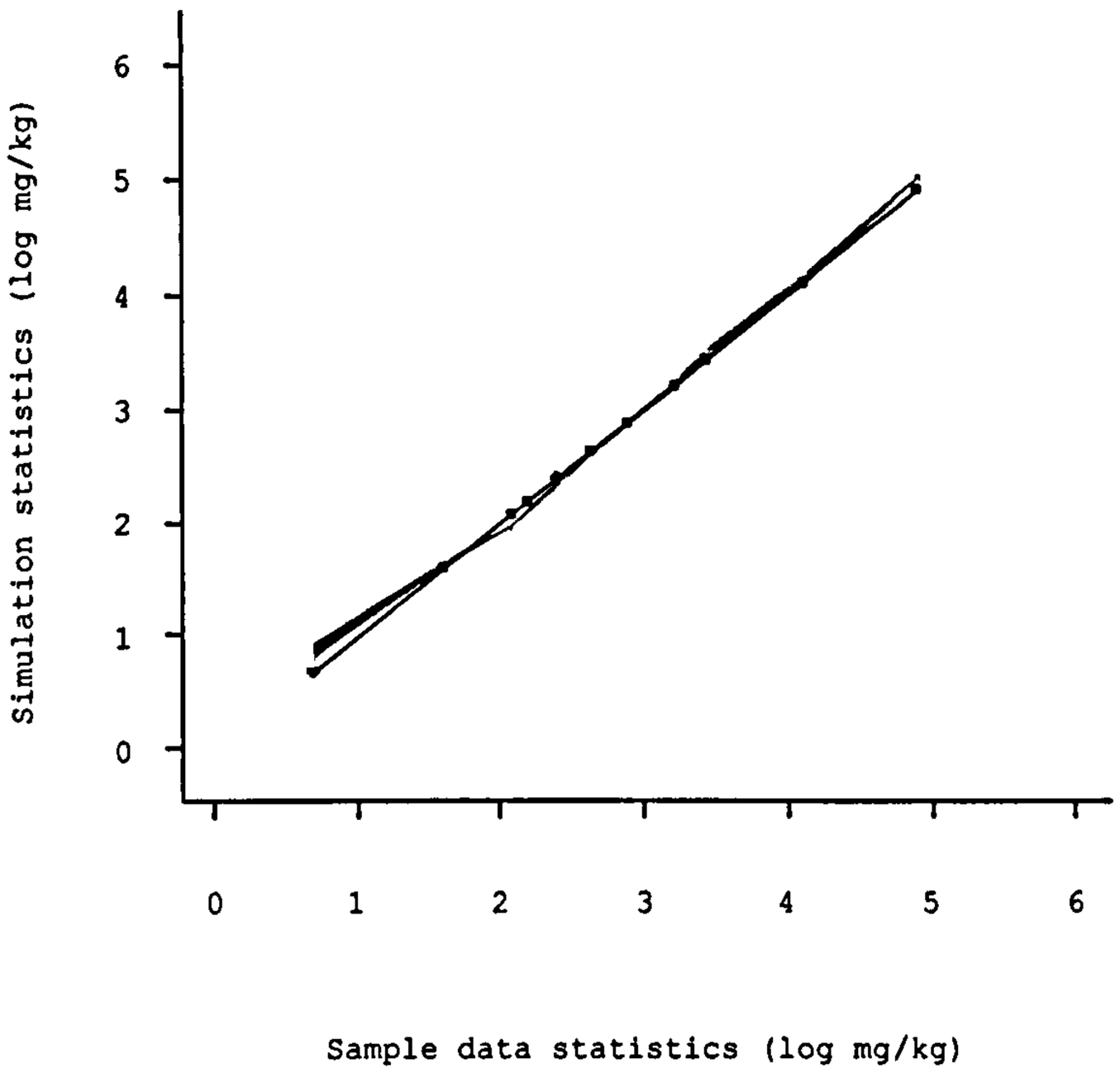


Figure A 4. Sample As statistics plotted against simulation statistics for first 50 simulations, as an example. Black diagonal represents perfect match. Statistics plotted are; min, P1, P5, P10, P25, P50, P75, P90, P95, P99 and max.

Table A 3. Statistics of the sample data and the average statistics from 100 simulations ($mg.kg^{-1}$).

	mean	stdev	min	p1	p5	p10	p25	p50	p75	p90	p95	p99	max
Samples	16.1	10.1	2	5	8	9	11	14	18	25	31	61	136
Simulations 50m grid	16.0	11.3	2.3	5.1	7.1	8.3	10.4	13.4	17.6	28.8	32.6	62.0	150
Simulations 200m grid	16.0	4.5	7.6	9.2	10.4	11.2	12.7	15.0	18.2	22.0	24.7	30.5	43.6

Indicator procedures at multiple cut offs with different indicator variogram models can result in order relation problems, whereby $(P > z_c) < (P > z_{c+1})$, which is clearly impossible. Order relation problems are corrected within the SISIM programme and are low for these simulations: in the order of 8% of ccdfs requiring corrections, with an average magnitude of correction in the order of 0.002.

In order to assess the goodness of the simulations the method devised by Deutsch was followed (Deutsch, 1997). This method produces Accuracy, Precision and Goodness statistics between 0 (worst case) and 1 (best case) by comparing the sample values with the simulated values at that point. An accurate probability distribution is one where the fraction of true values falling in the p-interval exceeds p for all p in [0,1]. In an accuracy plot, the calculated points should fall on or above the line for an accurate

probability distribution. A precise (and accurate) distribution is measured by the closeness of the fraction of true values to p for all p in [0,1] e.g.. for the p=0.5 interval a precise distribution would contain exactly 50% of the true values. Goodness is a measure of departure from the 45degree line on the accuracy plot and weights inaccurate results twice, as they are more consequential than accurate ones. The results gave A=0.89, P=0.93 and G=0.96; an accuracy plot is shown in Figure A 5. In addition, a histogram is plotted of where in the distribution of simulations, the true value lies. First of all, the 745 grid nodes nearest to the sample points were selected as representing the simulated distribution of soil values for that sample point. The plot is shown in Figure A 6. A chi^2 test was not significant, so it can be concluded that the results do not differ significantly from the expected. On the basis of these tests, the simulations were accepted as being satisfactory.

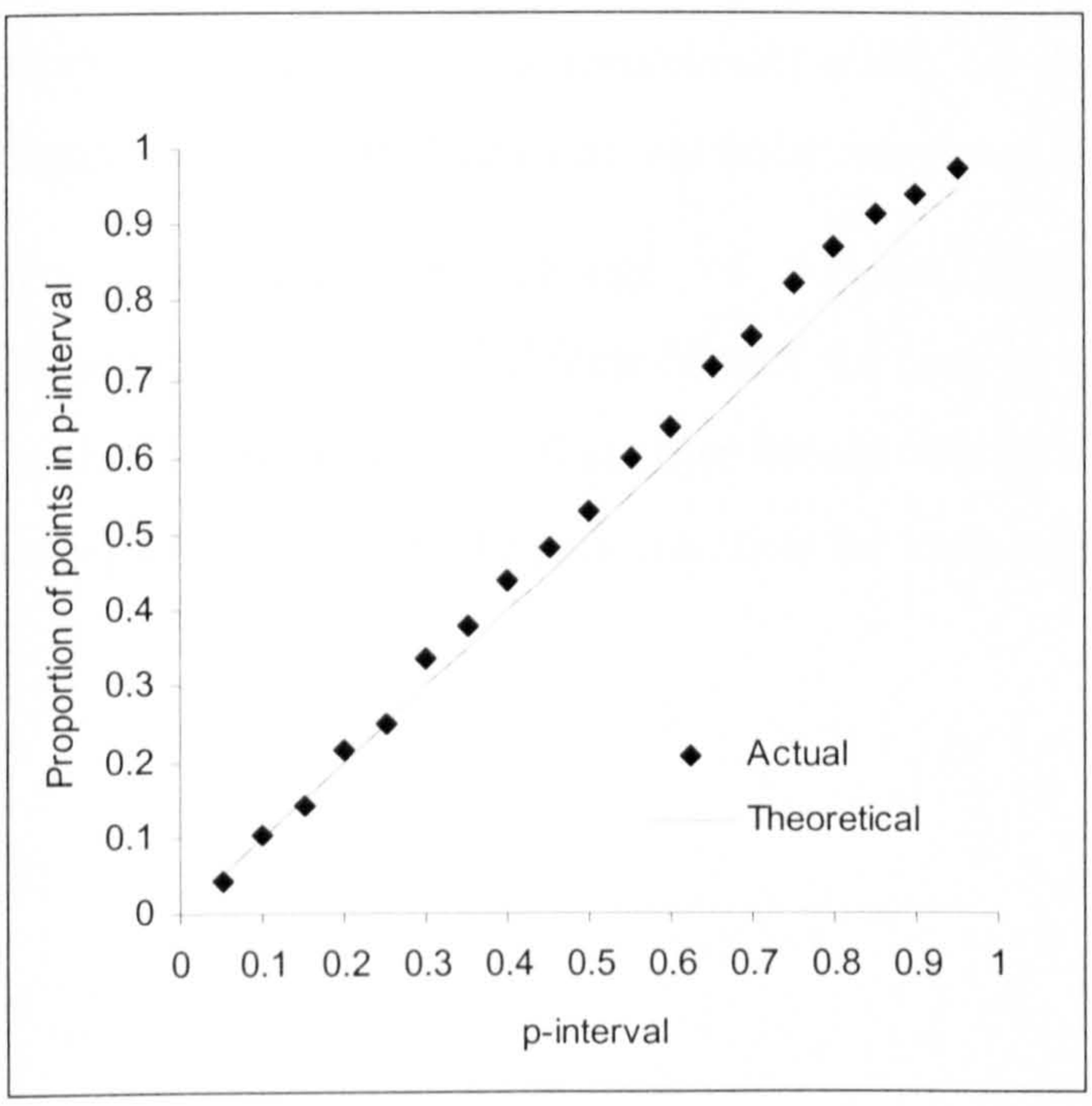


Figure A 5 Accuracy plot for As simulations.

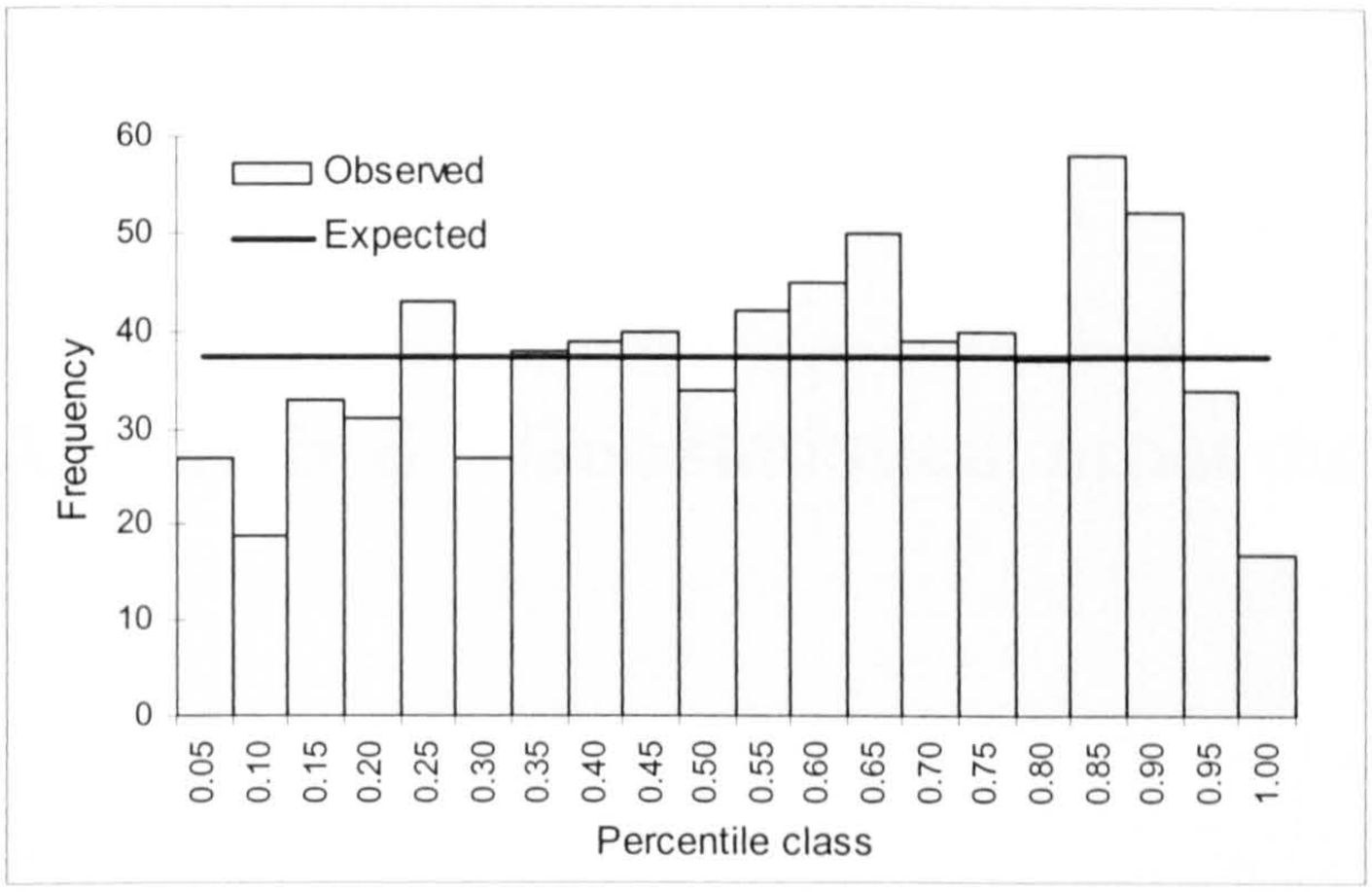


Figure A 6. Distribution of where in the simulated As values the true As value lies.

1,000 simulations were produced in all based on the fact that cumulative map statistics (mean, median and variance) had stabilised and at individual grid points the 'space of uncertainty' had reached a maximum width i.e. the interval between the minimum and maximum simulated value at that point increased no more.

To demonstrate the change of support effect, the 50x50m simulations were amalgamated into 200x200m blocks for use in the exposure analysis, by taking the arithmetic mean of the 16 smaller blocks within each large block. For the non-spatial analysis, the average As concentration for each map was also found.

Appendix B Geostatistical procedure for Nickel in Gai

B.1 Simple statistics

Simple statistics of the Ni concentrations for the 160 topsoil samples are given in Table B 1 and a histogram of the distribution in Figure B 1.

Table B 1. Simple statistics of the Ni topsoil sample concentrations ($mg.kg^{-1}$).

Mean	Stdev	Min	P1	P5	P10	P25	P50	P75	P90	P95	P99	Max
74.1	79.3	18	19	31	39	53	63.5	79.5	99	108	280	981

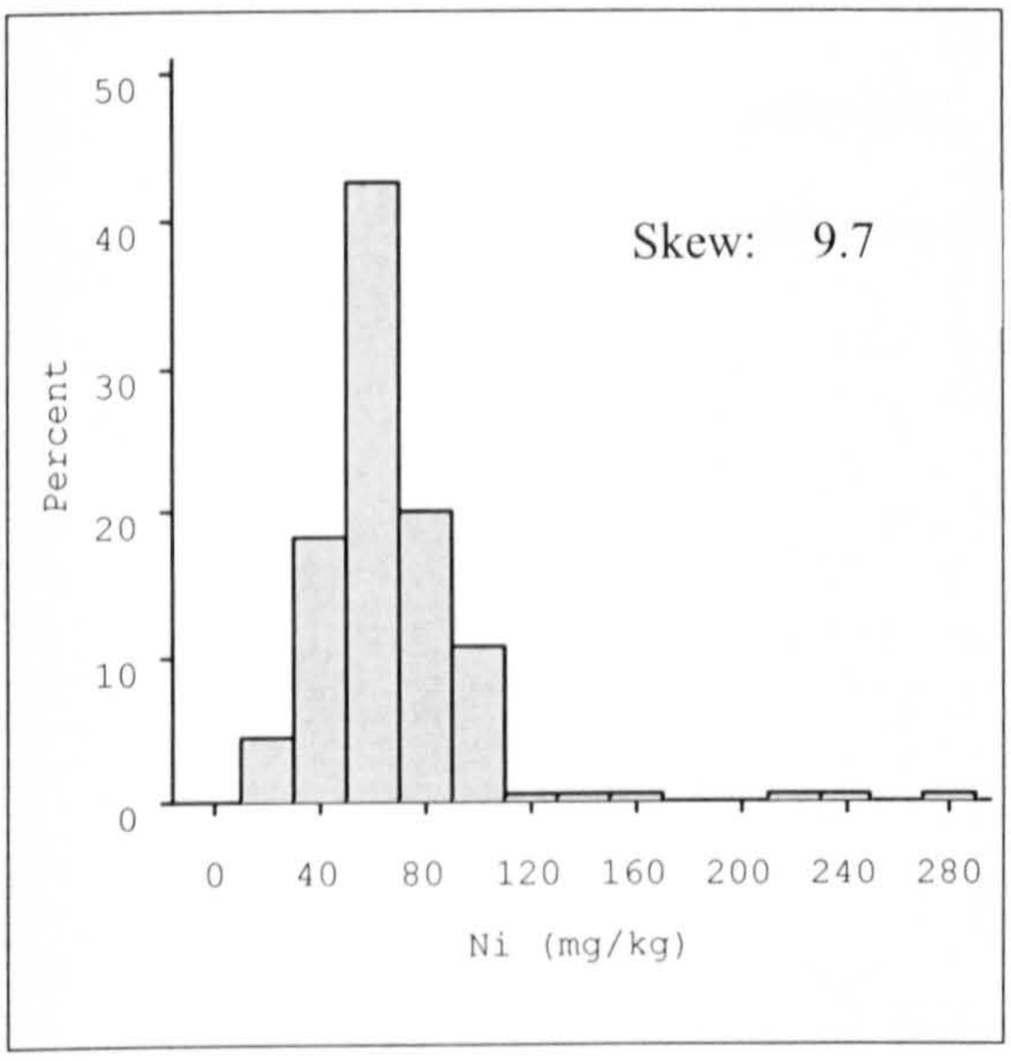


Figure B 1 Histogram of Ni topsoil sample concentrations ($mg.kg^{-1}$).

B.2 Variogram modelling

Variogram maps are shown in Figure B 2. A small hint of anisotropy was seen at some cut offs, but it was deemed not strong enough to warrant modelling. Most structure was apparent at a lag of 900m. Omni-directional indicator variograms were fitted at the nine decile values of the sample data: 39, 46.5, 56.5, 61.5, 63.5, 67, 74, 85 and $99mg.kg^{-1}$.

The variogram models, fitted using Isatis (Geovariances, 2002), were composed of nugget, exponential and/or spherical components. Sample and model variograms are shown in Figure B 3 and the model parameters are given in Table B 2. To help avoid order relation problems, the sills of the variogram models were fitted to the theoretical variances, p(1-p), where p is the percentile of the cut off.

Table B 2. Indicator variogram model parameters.

Cut off (ppm)	Exponential[1]		Spherical[2]		Nugget C_0
	a	c	a	c	
39	5000	0.06	2000	0.02	0.01
46.5	6000	0.14	3500	0.01	0.01
56.5	4000	0.13	5500	0.07	0.01
61.5	3000	0.15	5000	0.08	0.01
63.5	3000	0.1	4500	0.14	0.01
67	-	-	3600	0.22	0.02
74	-	-	3500	0.19	0.02
85	-	-	3600	0.14	0.02
99	-	-	3400	0.06	0.03

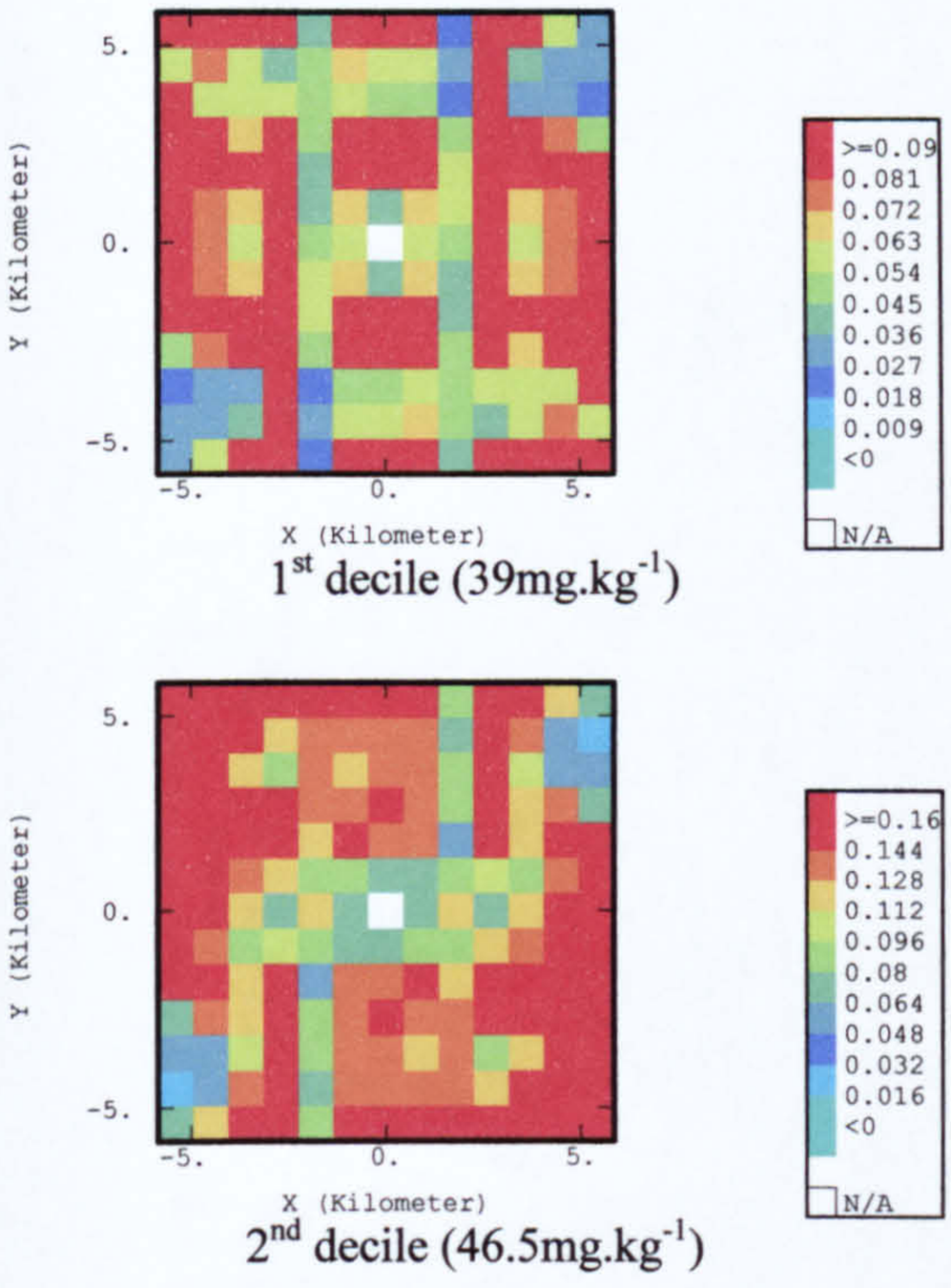
[1] Exponential model: $\gamma(h) = C_0 + c.\left[1 - \exp\left(-\frac{3h}{a}\right)\right]$

where h is the lag distance

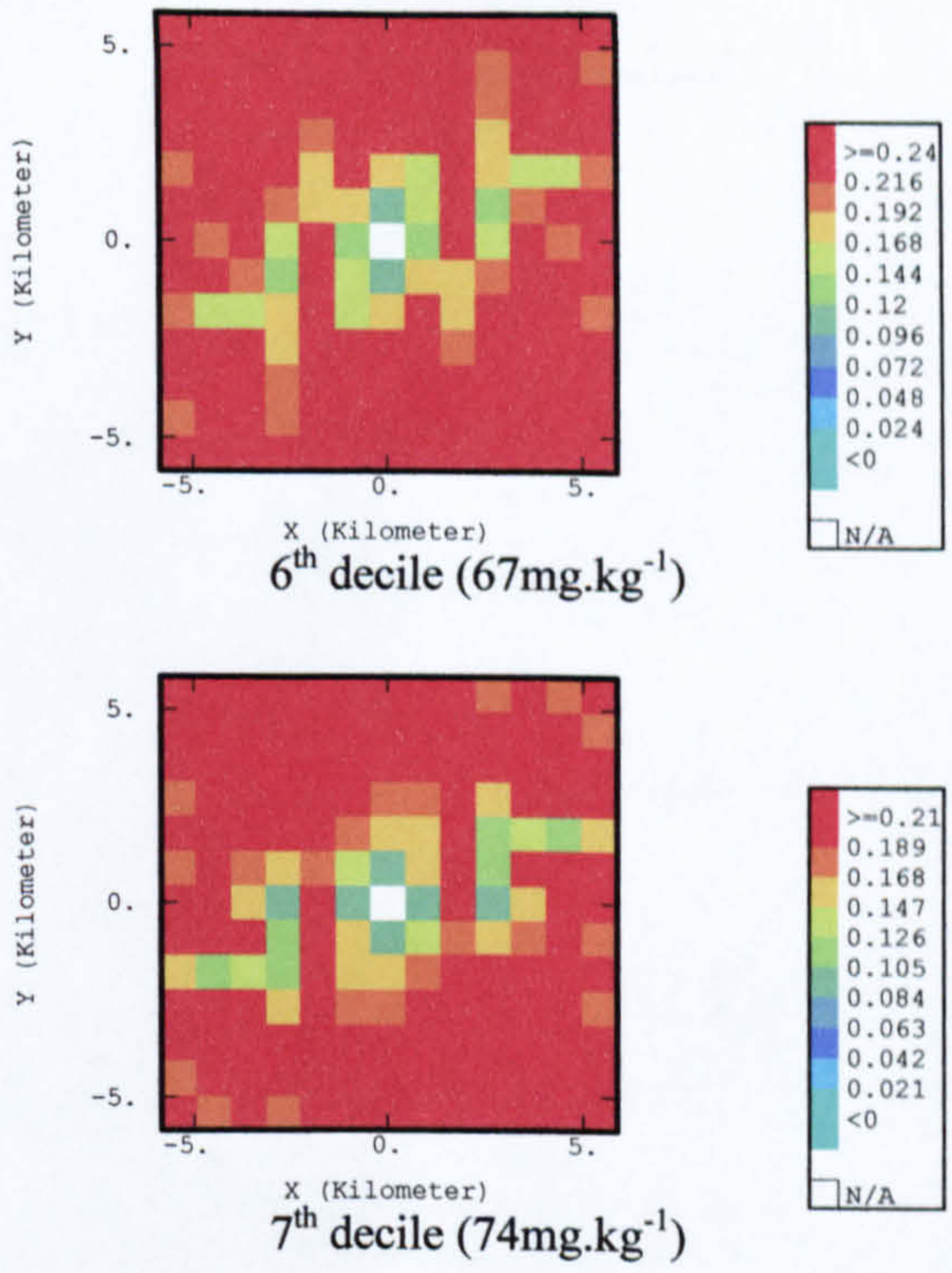
[2] Spherical model: $\gamma(h) = C_0 + c.\left[1.5\frac{h}{a} - 0.5\left(\frac{h}{a}\right)^3\right]$ if $h \le a$

$\gamma(h) = C_0 + c$ if $h > a$

where h is the lag distance



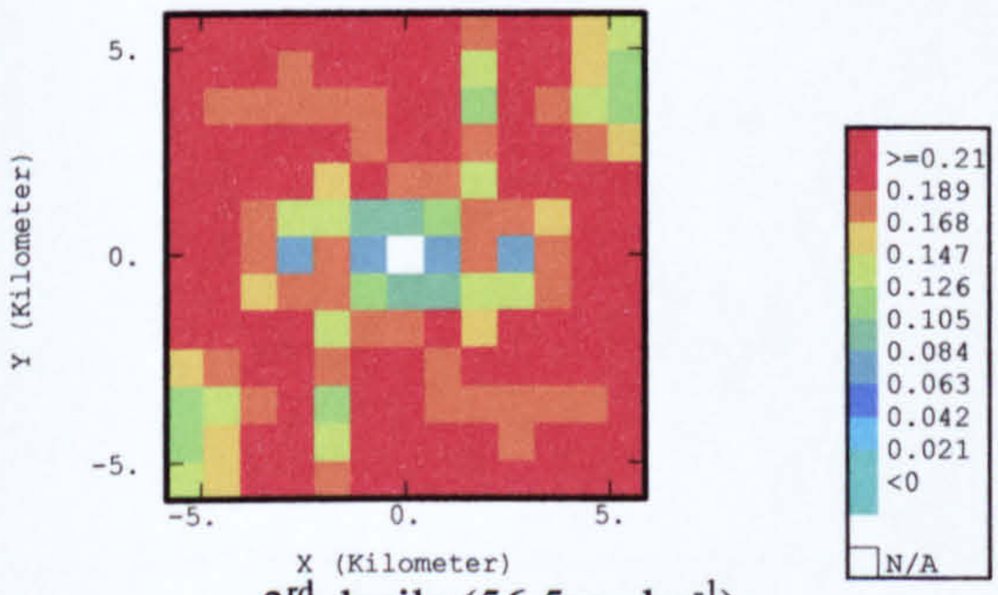
1st decile (39mg.kg^{-1})



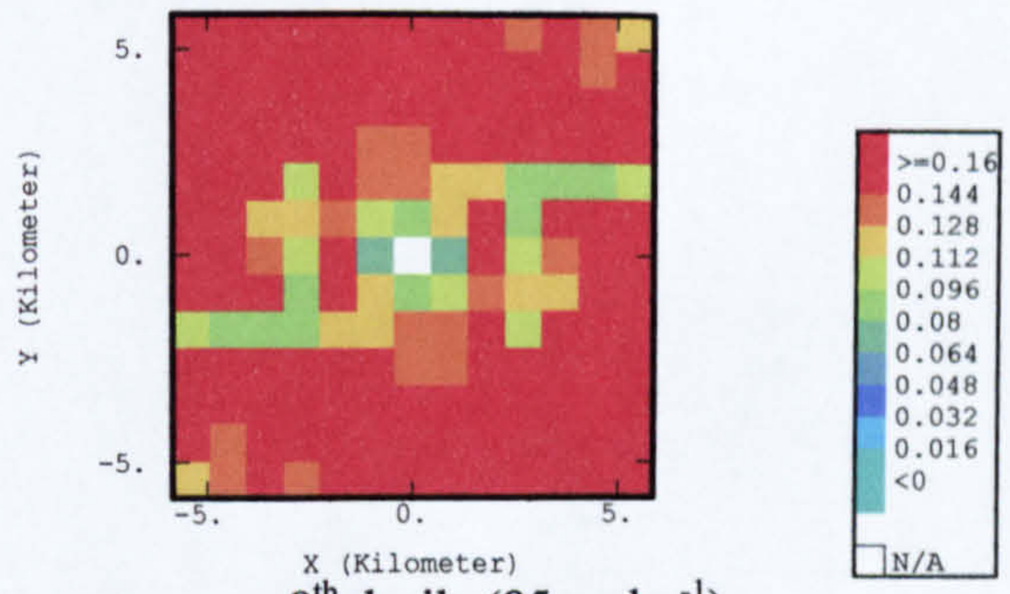
6th decile (67mg.kg^{-1})

2nd decile (46.5mg.kg^{-1})

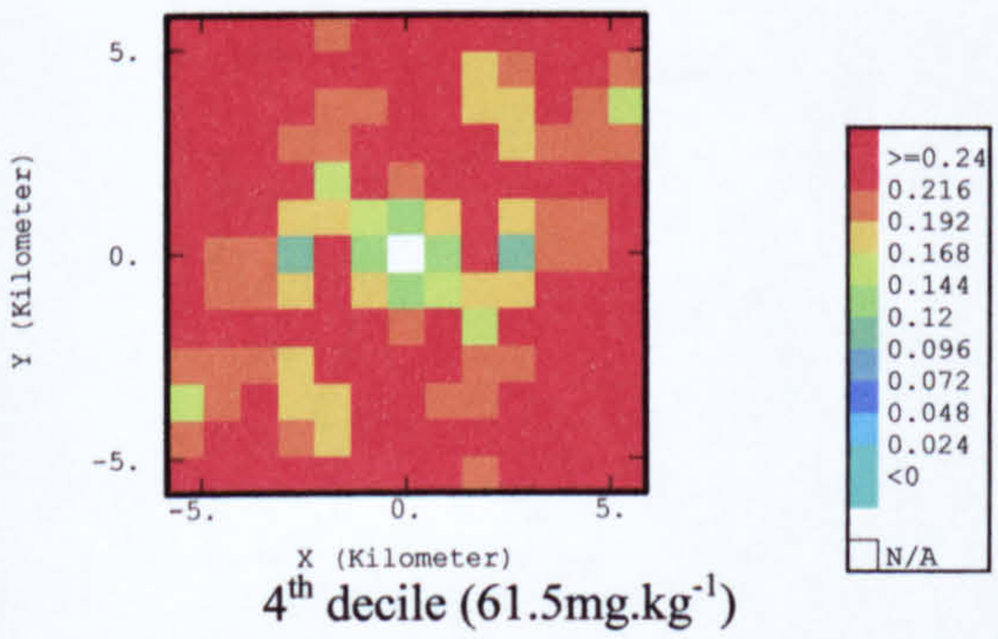
7th decile (74mg.kg^{-1})



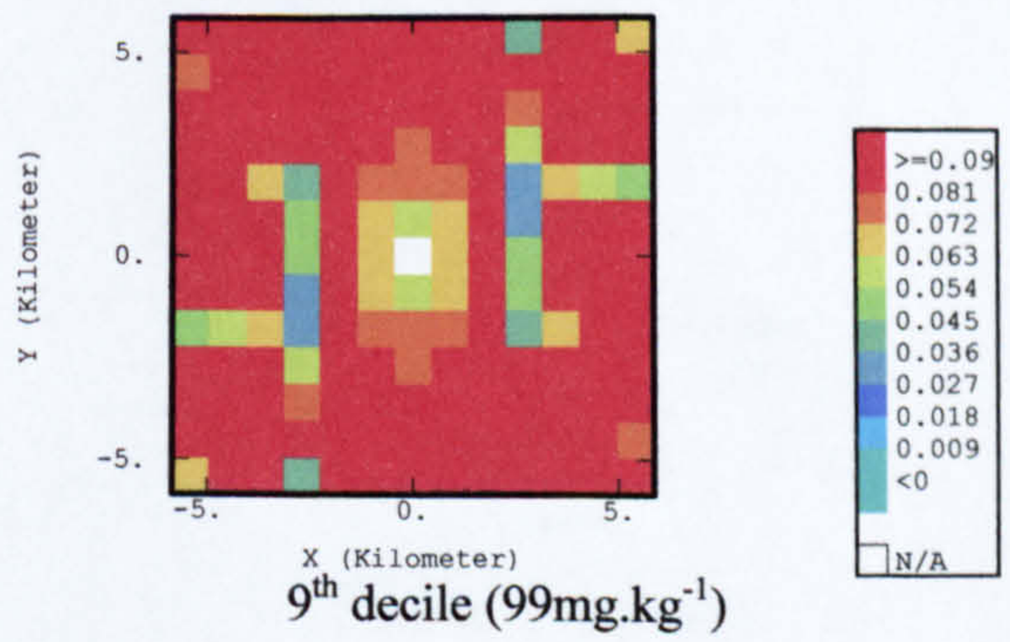
3rd decile (56.5mg.kg^{-1})



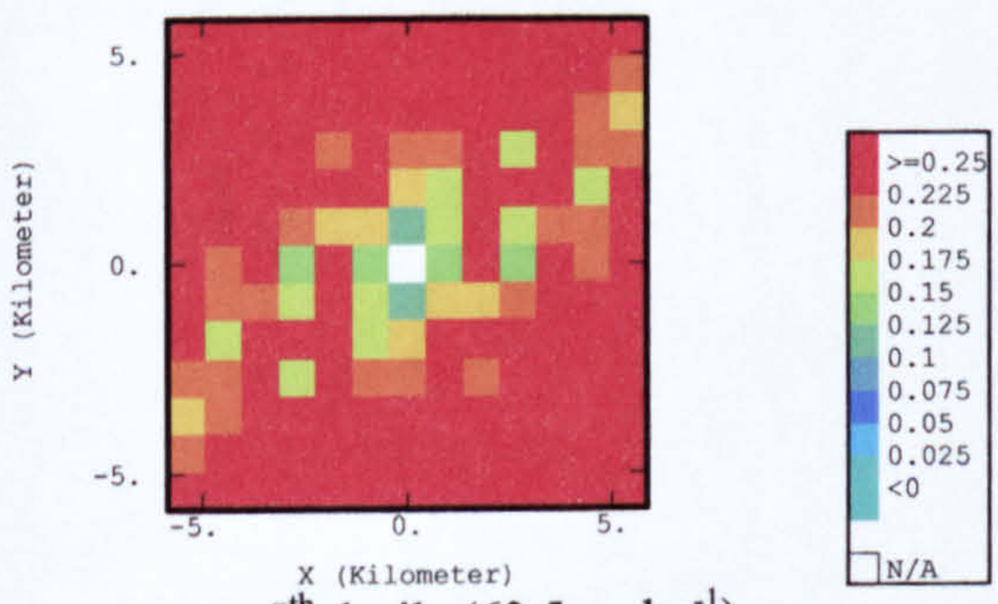
8th decile (85mg.kg^{-1})



4th decile (61.5mg.kg^{-1})

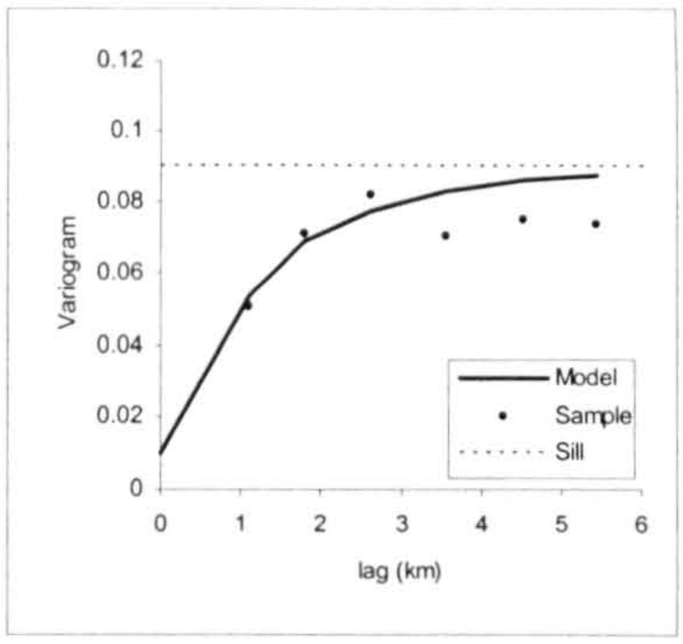


9th decile (99mg.kg^{-1})

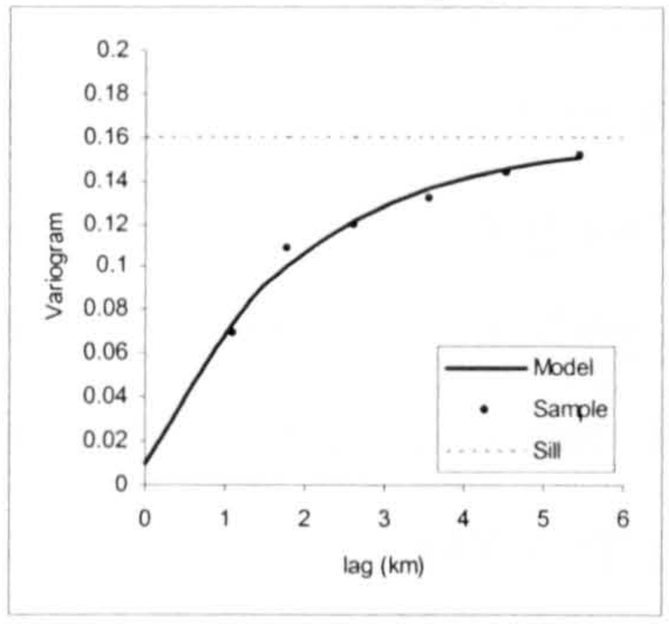


5th decile (63.5mg.kg^{-1})

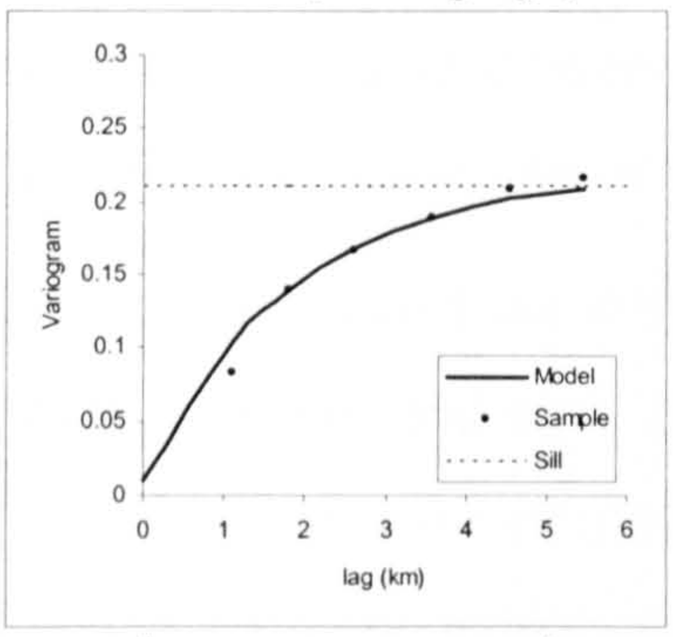
Figure B 2. Indicator variogram maps for Ni at nine cut offs.



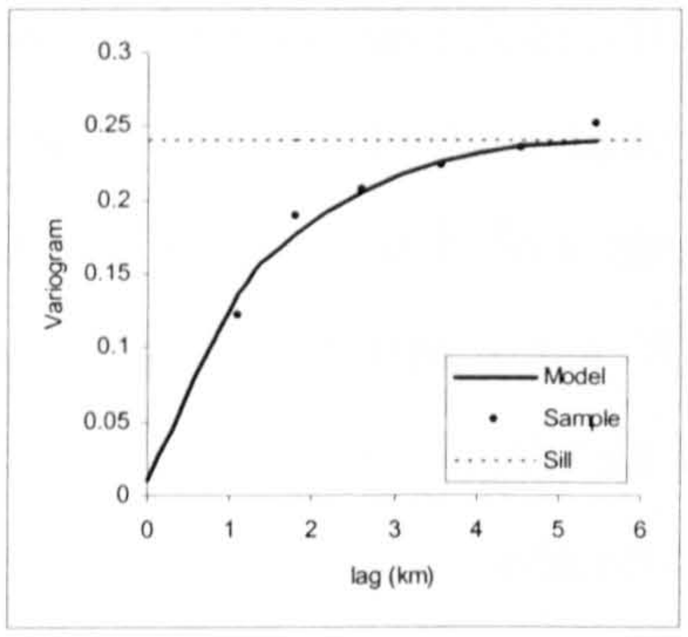
1st decile (39mg.kg^{-1})



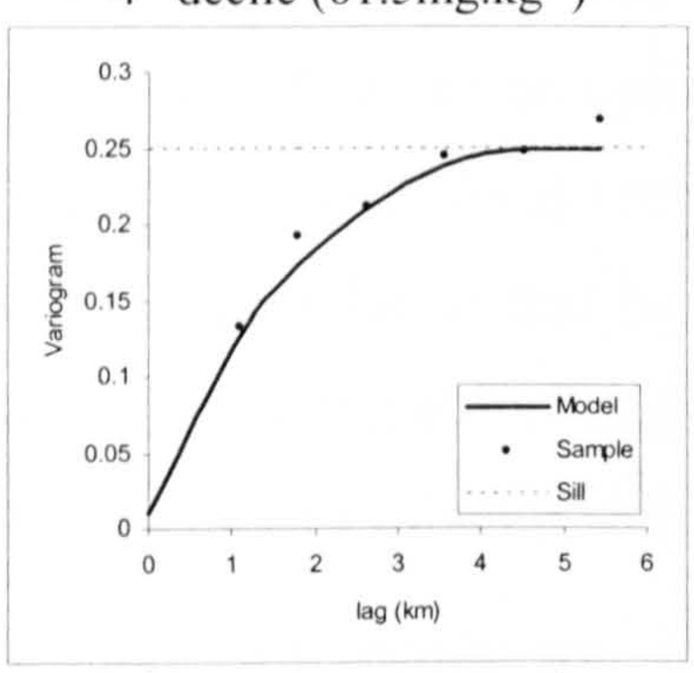
2nd decile (46.5mg.kg^{-1})



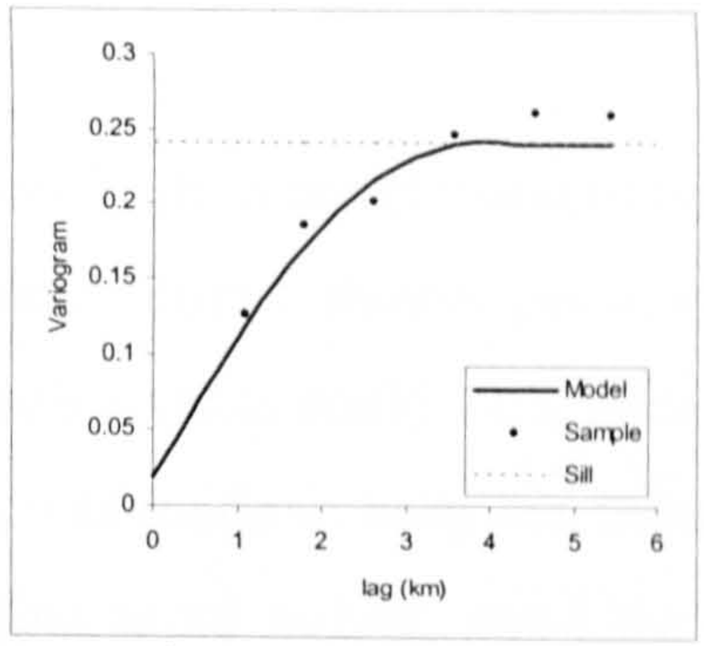
3rd decile (56.5mg.kg^{-1})



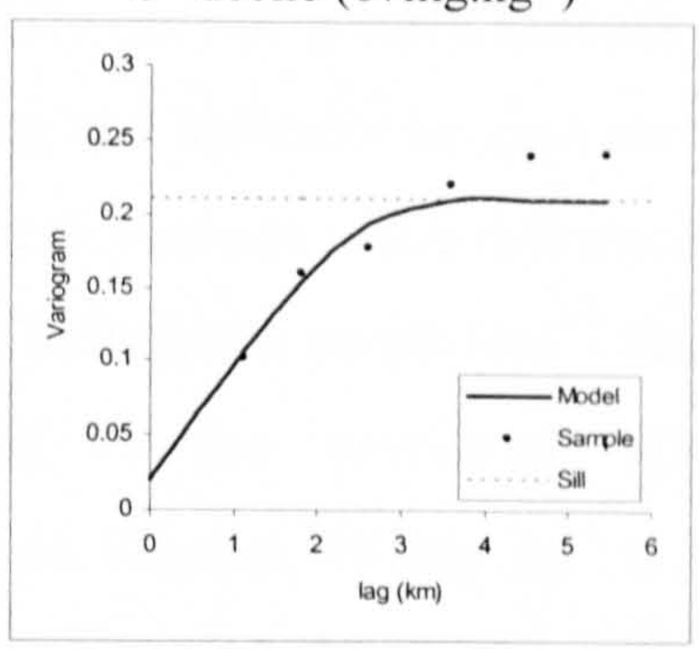
4th decile (61.5mg.kg^{-1})



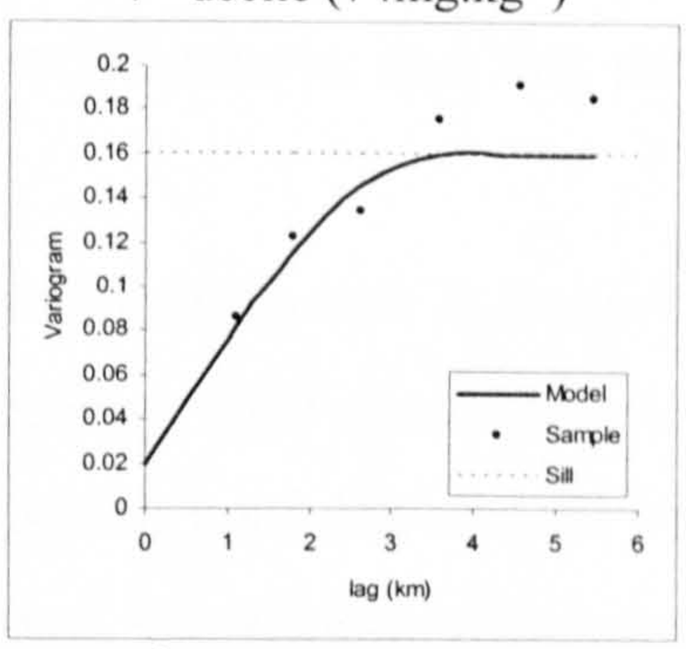
5th decile (63.5mg.kg^{-1})



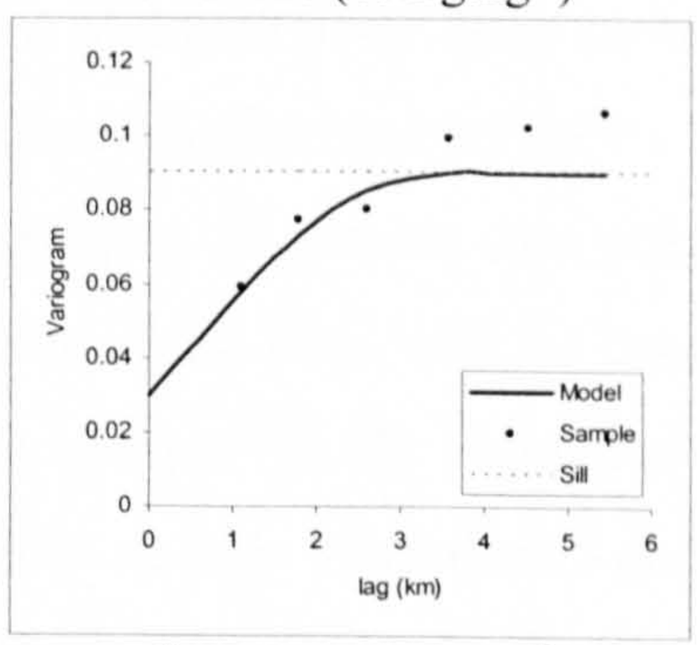
6th decile (67mg.kg^{-1})



7th decile (74mg.kg^{-1})



8th decile (85mg.kg^{-1})



9th decile (99mg.kg^{-1})

Figure B 3. Sample and model indicator variograms for Ni.

B.3 Running and validating simulations

The maximum sample value of 981mg.kg^{-1} appears very high in comparison to the other samples (the next highest value is 280mg.kg^{-1}). It comes from a sample point which is in the vicinity of a spoil heap, so it is possible that such a value could have been taken. In the absence of any other information, the decision was made to keep the value as is. Using indicators, it is not necessary to delete the value as an outlier, since the binary code will be the same whether the sample is very slightly over or greatly exceeds the cut off of interest so will not affect the indicator modelling and indicator kriging procedure. The biggest effect of leaving such a high value in the calculations is the influence it will have when it comes to building the ccdfs during the simulation procedure. Since this involves extrapolating from the highest cut off to the maximum allowable concentration, results will vary greatly depending upon whether 981mg.kg^{-1} is kept as the maximum or 281mg.kg^{-1} is taken as the maximum. Keeping the 981mg.kg^{-1} sample will lead to much more conservative exposure and intake results, and this should be borne in mind during analysis of the results.

Simulations were carried out with the GSLIB SISIM routine (Deutsch & Journel, 1998) on a 50x50m square grid with 290 cells in each direction. Each simulated value was taken to be representative of a 50x50m square. The ordinary kriging routine was used with a moving window neighbourhood of radius 5,000m. The minimum and maximum numbers of sample data points to be used in simulation of a grid point were set to 0 and 16 respectively, while the number of previously simulated nodes was set to 16. These parameters ensure that before any nodes are simulated, up to 16 sample data points will be used, and thereafter, once simulated nodes are available, they will be taken into account also, up to a 50:50 mix between sample data and previously simulated nodes. Using the previously simulated nodes ensures the joint ccdf is reproduced and using sample data assists reproduction of the sample histogram. To speed up the simulation process, sample data can be optionally allocated to grid nodes prior to simulation, and kept fixed throughout. As in previous case studies, this was deemed inappropriate since every simulated value at that point would be the same with no uncertainty and could lead to 'holes' in probability maps at a later stage i.e. that grid point will always have a probability of 0 or 1 of being over a threshold, no matter what the results around it.

On building the ccdfs for each cell, minimum and maximum allowable Ni concentrations were set to 10% less and 10% more than the minimum and maximum

sample value respectively (i.e. to 16mg.kg^{-1} and 1080mg.kg^{-1}), assuming that it was unlikely that the minimum and maximum values would have been found in sampling. The figure of 10% is arbitrary, but may well lead to more realistic simulations. As mentioned above, the tails of the ccdf must be extrapolated from the highest cut off to the maximum and the lowest cut off to the minimum using models. After much trial and error, the upper tail model was selected as a hyperbolic with parameter 2 and the lower a power model with parameter 1.5. The selection was based on comparing the statistics of the sample and simulated distributions of Ni values. The closer the statistics, the better the model parameters. Figure B 4 shows a plot of the sample value statistics versus the average of the simulated statistics and Table B 3 summarises the statistics.

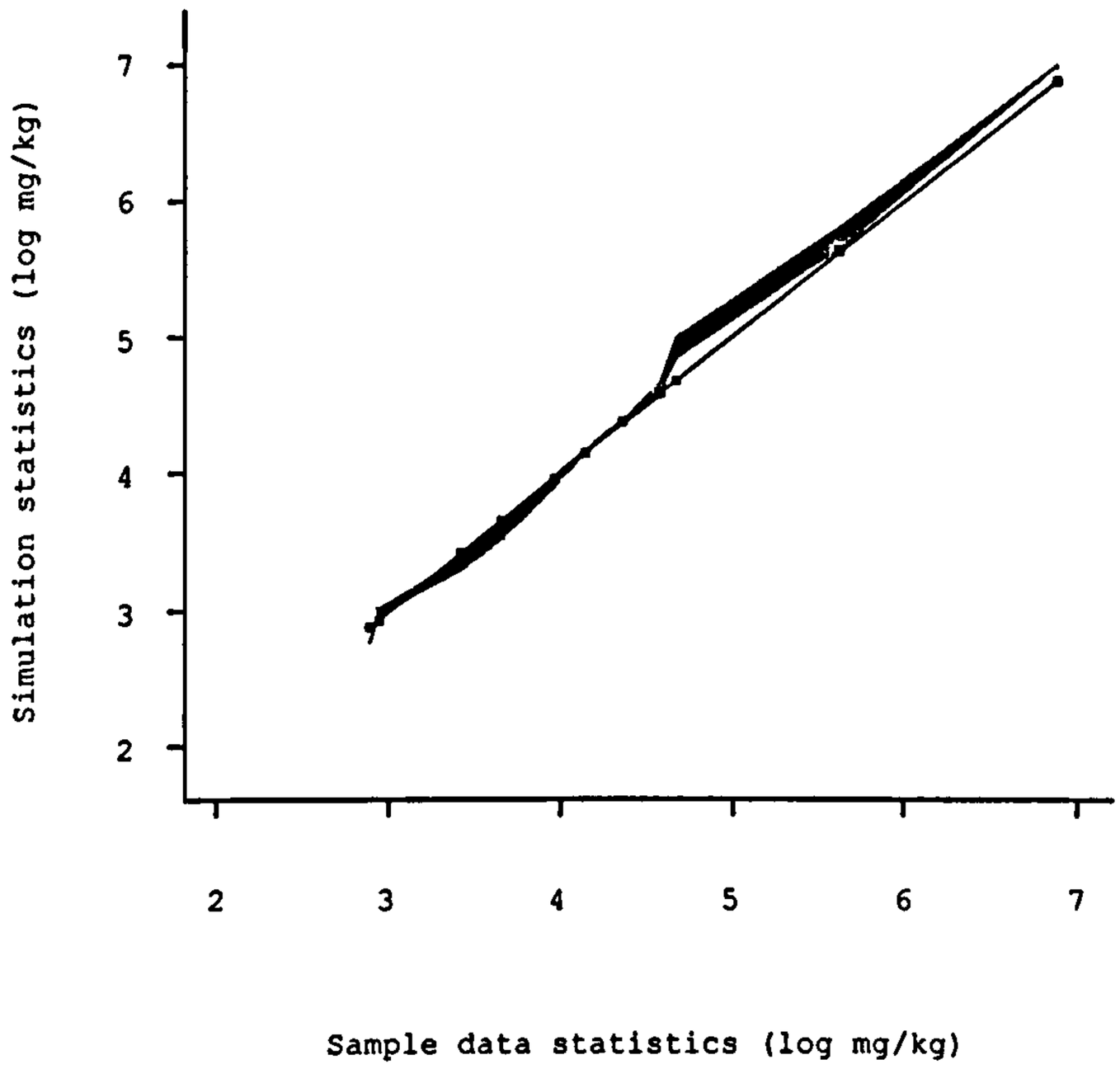


Figure B 4. Sample Ni statistics plotted against simulation statistics for first 50 simulations, as an example. Black diagonal represents perfect match. Statistics plotted are; min, P1, P5, P10, P25, P50, P75, P90, P95, P99 and max.

Table B 3. Statistics of the Ni sample data and the average statistics from 100 simulations (mg.kg^{-1}).

	mean	stdev	min	p1	p5	p10	p25	p50	p75	p90	p95	p99	max
Samples	74.1	79.3	18	19	31	39	53	64	80	99	108	280	981
Simulations 50m grid	74.0	61.7	16	20	29	36	51	64	80	100	139	310	1080
Simulations 500m grid	74.0	33.8	29	30	33	39	53	66	84	124	151	183	209

Indicator procedures at multiple cut offs with different indicator variogram models can result in order relation problems, whereby $(P > z_c) < (P > z_{c+1})$, which is clearly impossible. Order relation problems are corrected within the SISIM programme. In the order of 72% of ccdfs require corrections, with an average magnitude of correction in the order of 0.0006.

In order to assess the goodness of the simulations the method devised by Deutsch was followed (Deutsch, 1997). This method produces Accuracy, Precision and Goodness statistics between 0 (worst case) and 1 (best case) by comparing the sample values with the simulated values at that point. An accurate probability distribution is one where the fraction of true values falling in the p-interval exceeds p for all p in [0,1] and a precise (and accurate) distribution is measured by the closeness of the fraction of true values to p for all p in [0,1]. Goodness is a measure of departure from the 45degree line on the accuracy plot and weights inaccurate results twice, as they are more consequential than accurate ones. The results gave A=0.95, P=0.78 and G=0.89. An accuracy plot is shown in Figure B 5. In addition, a histogram is plotted of where in the distribution of simulations, the true value lies. First of all, the 160 grid nodes nearest to the sample points were selected as representing the simulated distribution of soil values for that sample point. The plot is shown in Figure B 6. The highest and lowest percentile classes are a little under-represented, and one reason for this may be due to the fact that the simulations were allowed to vary to 10% higher and lower than the highest and lowest sample values respectively. In any case, a chi^2 test was not significant, so it can be concluded that the results do not differ significantly from the expected. On the basis of these tests, the simulations were accepted as being satisfactory.

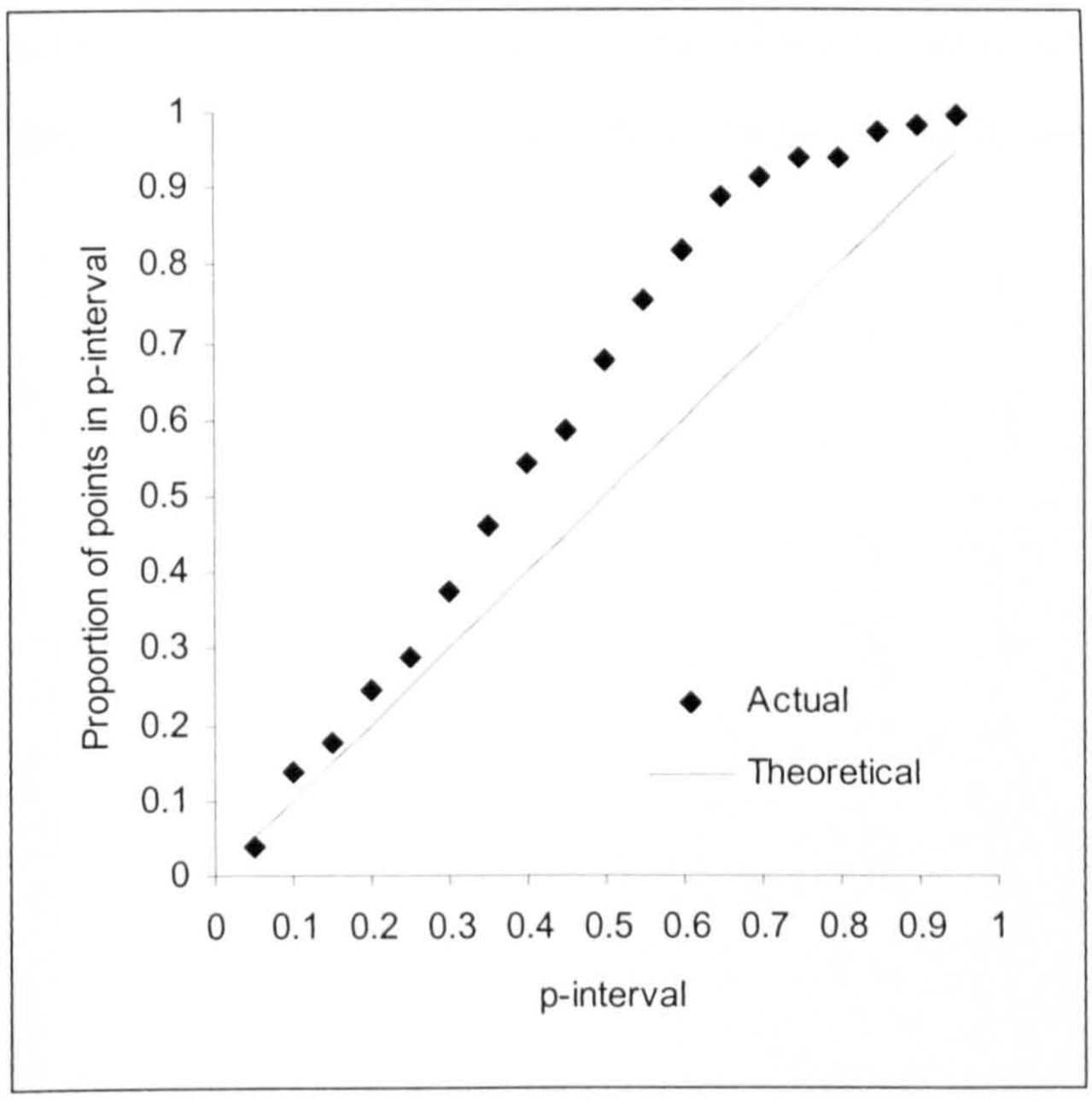


Figure B 5. Accuracy plot for Ni simulations.

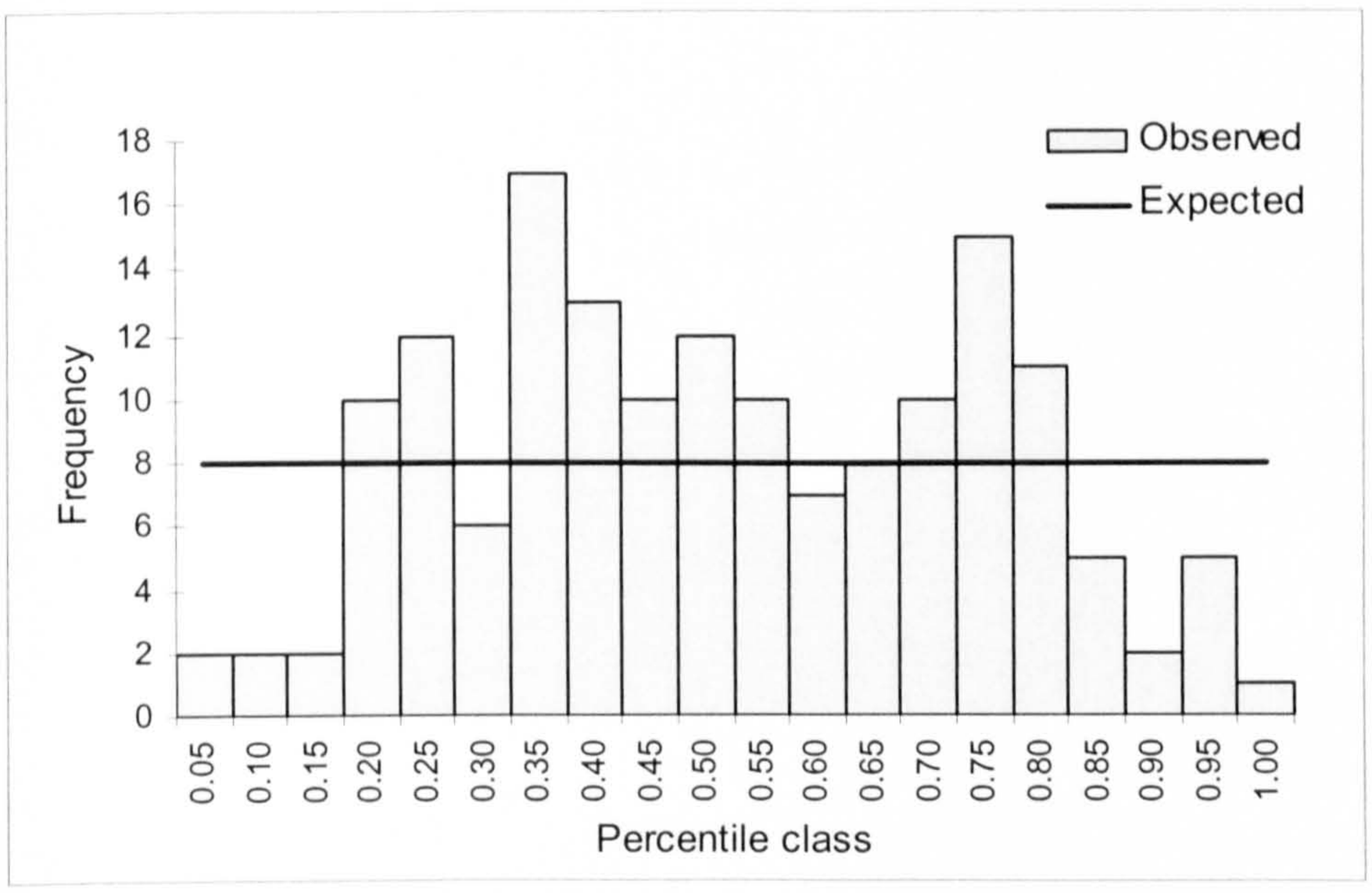


Figure B 6. Distribution of where in the simulated Ni values the true value falls.

1,000 simulations were produced in all based on the fact that cumulative map statistics (mean, median and variance) had stabilised and at individual grid points the 'space of uncertainty' had reached a maximum width i.e. the interval between the minimum and maximum simulated value at that point increased no more.

To demonstrate the change of support effect, the 50x50m simulations were amalgamated into 500x500m blocks for use in the exposure analysis, by taking the

arithmetic mean of the 100 small blocks contained within each large one. For the non-spatial assessment, the overall average for each map was found.

Appendix C Geostatistical procedure for Cadmium and pH in Vasilevskaya

C.1 Simple statistics; Cd

Simple statistics of the 135 topsoil samples are given in Table C 1 and a histogram of the distribution in Figure C 1. The minimum and maximum Cd sample values were 0 and 1.9mg.kg^{-1} respectively. Values recorded as zero were not detectable by the instrumentation used and were classed as 'non-detects'. Three such values were recorded and were changed to be equal to the detection limit of the analytical test, given as 0.1mg.kg^{-1}. Therefore the range of Cd values becomes 0.1 to 1.9mg.kg^{-1}.

Table C 1. Simple statistics of the Cd topsoil sample concentrations (mg.kg^{-1}).

Mean	Stdev	Min	P1	P5	P10	P25	P50	P75	P90	P95	P99	Max
0.65	0.30	0.1	0.1	0.2	0.4	0.5	0.6	0.7	1.0	1.2	1.9	1.9

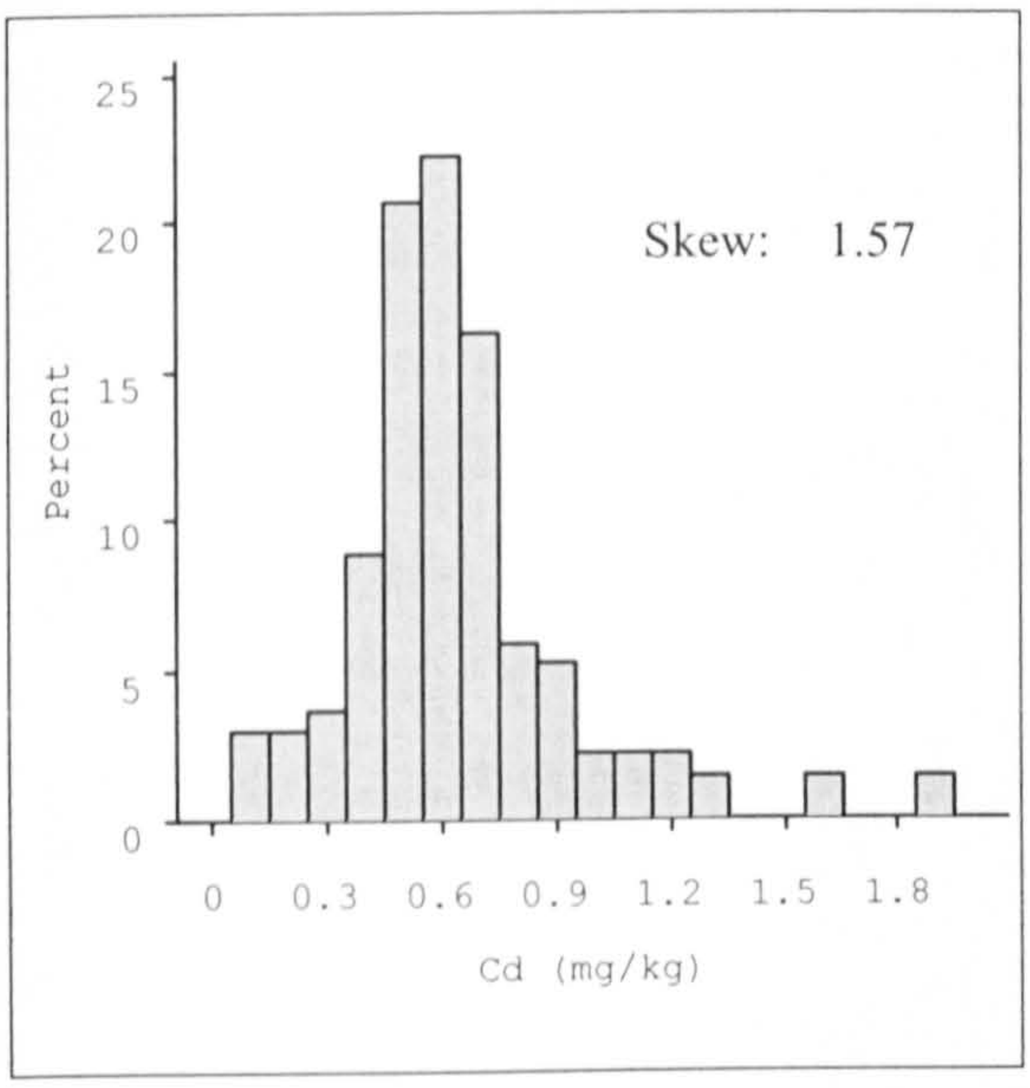


Figure C 1. Histogram of Cd topsoil sample concentrations (mg.kg^{-1}).

C.2 Variogram modelling; Cd

Since the analytical method is only accurate to one decimal place, this led to many repetitive sample values and as such it is not practical to define the cdf through 10 deciles in this case. Indicators were modelled using Isatis (Geovariances, 2002) at the 10^{th}, 19^{th}, 39^{th}, 61^{st}, 78^{th} and 89^{th} percentiles being 0.35, 0.45, 0.55, 0.65, 0.75 and 0.95mg.kg^{-1} respectively. To avoid order relation problems, model sills were fixed at the theoretical variances, p(1-p), where p is the percentile at that cut off. The variogram maps are shown in Figure C 2: a lag of 475m was used. The variograms were modelled isotropically using spherical and nugget components. The sample and model variograms are shown in Figure C 3 and the model parameters given in Table C 2.

Table C 2. Indicator variogram model parameters for Cd.

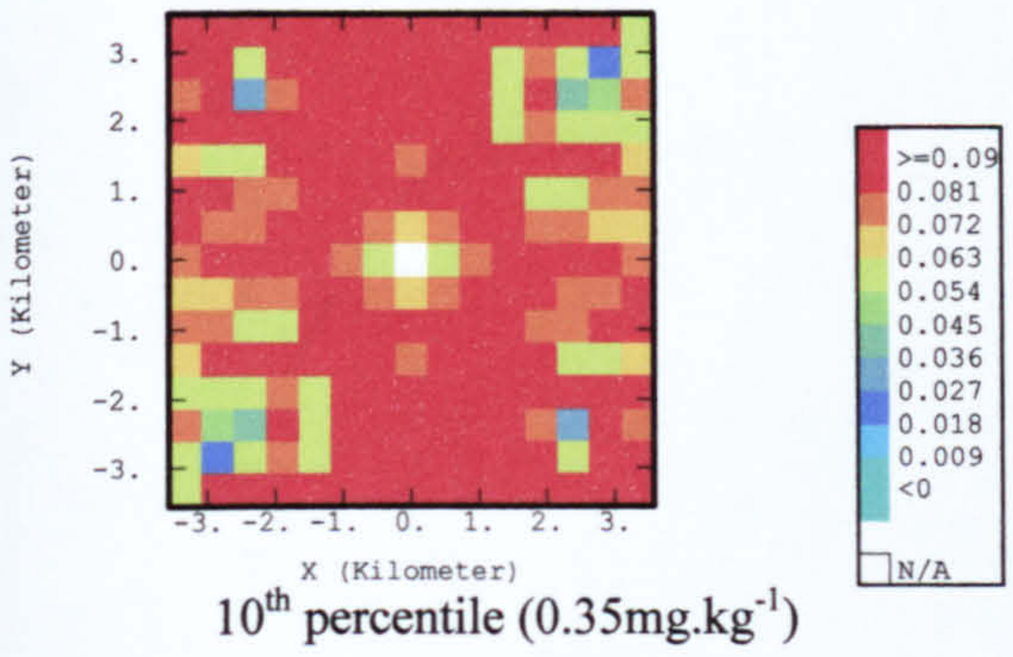
Cut off (mg.kg^{-1})	Spherical[1]		Nugget
	a	c	C_0
0.35	1250	0.05	0.04
0.45	1100	0.11	0.04
0.55	1450	0.21	0.03
0.65	1400	0.21	0.03
0.75	1400	0.13	0.04
0.95	1300	0.07	0.03

[1]Spherical model:

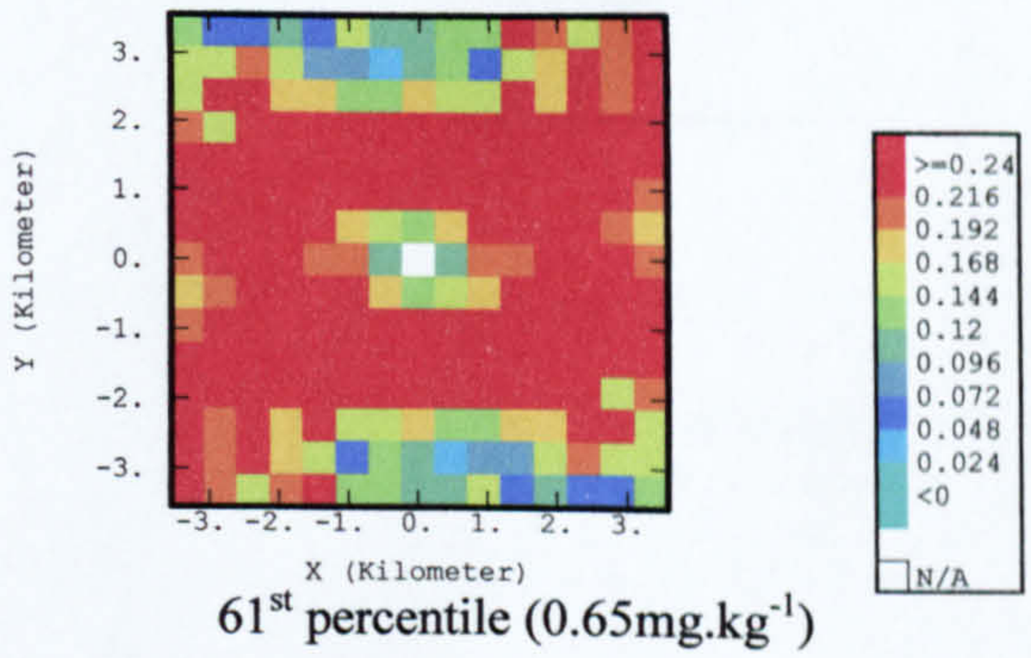
$$\gamma(h) = C_0 + c.\left[1.5\frac{h}{a} - 0.5\left(\frac{h}{a}\right)^3\right] \quad \text{if } h \le a$$

$$\gamma(h) = C_0 + c \quad \text{if } h > a$$

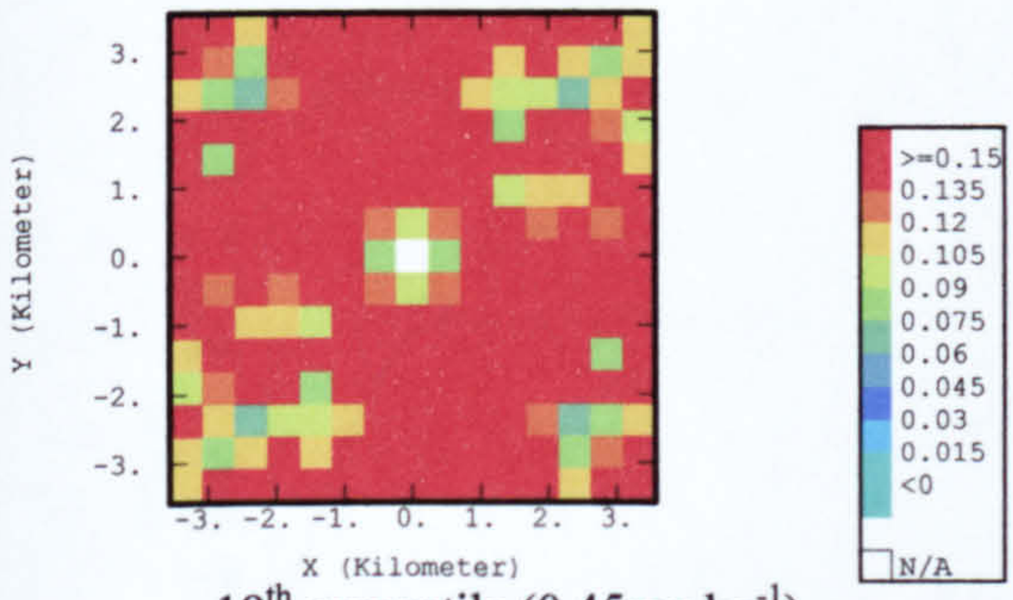
where h is the lag distance



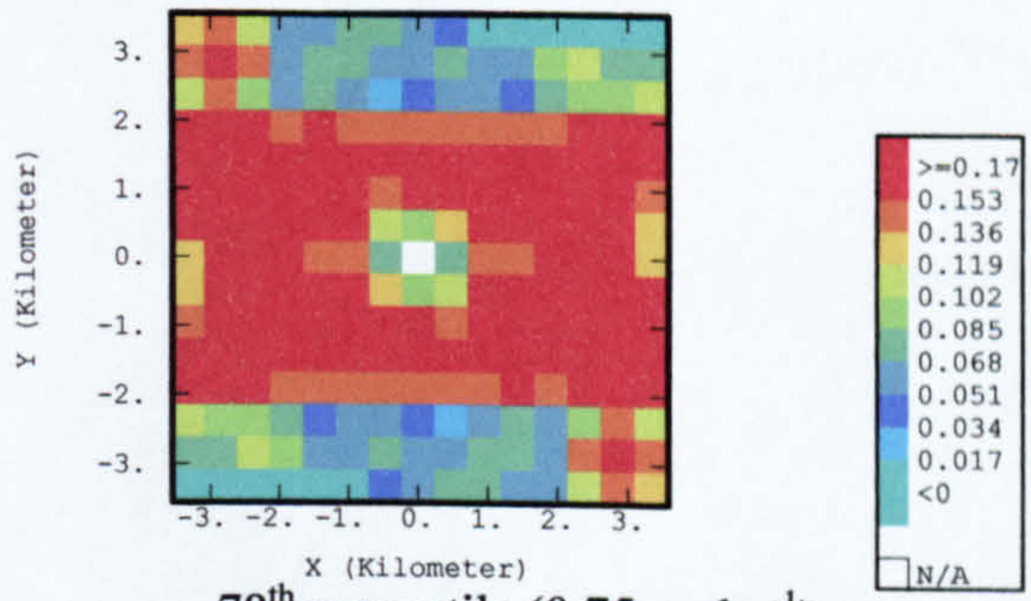
10th percentile (0.35mg.kg^{-1})



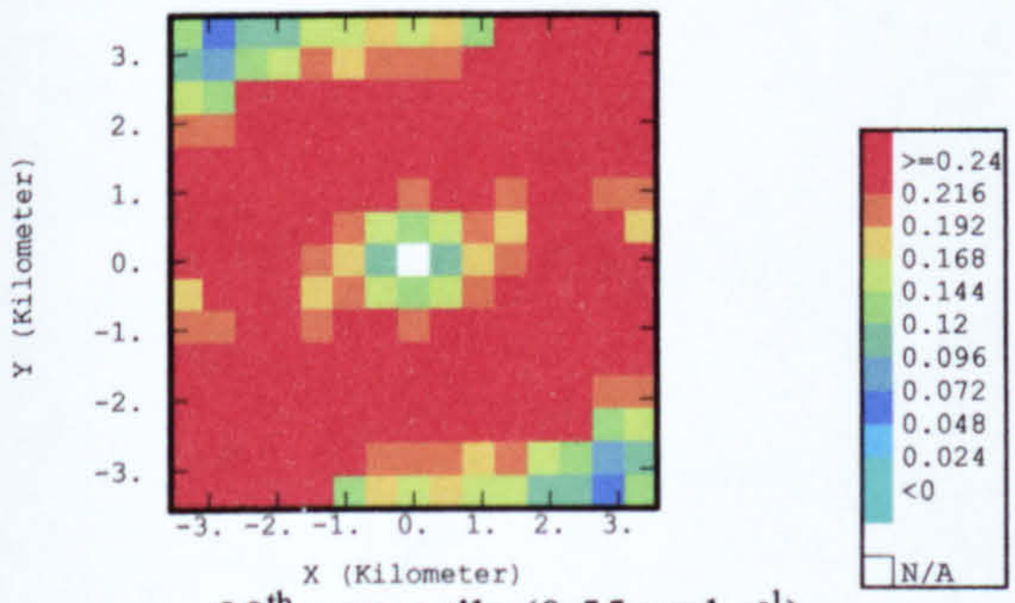
61st percentile (0.65mg.kg^{-1})



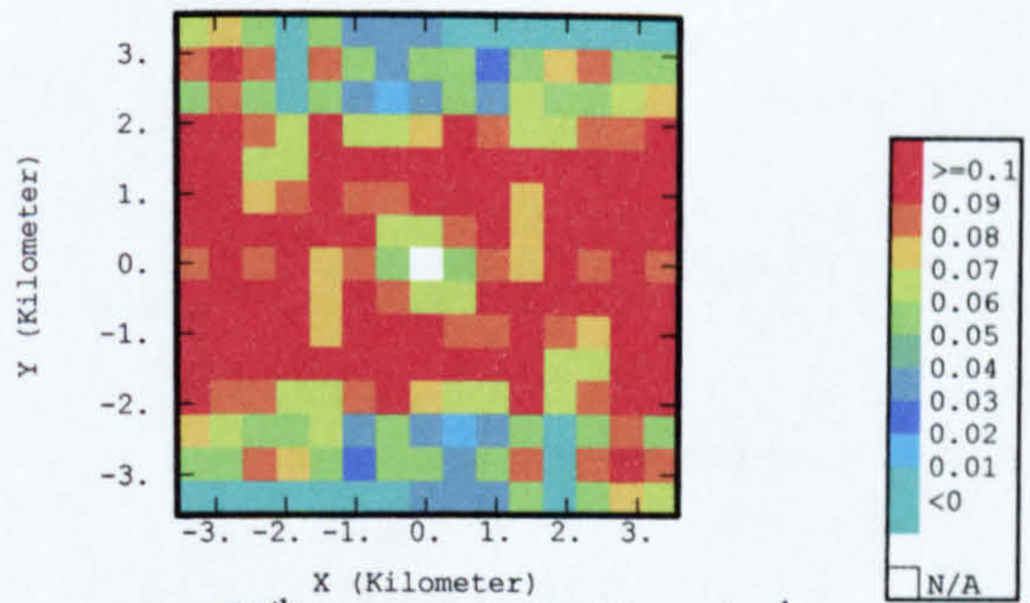
19th percentile (0.45mg.kg^{-1})



78th percentile (0.75mg.kg^{-1})

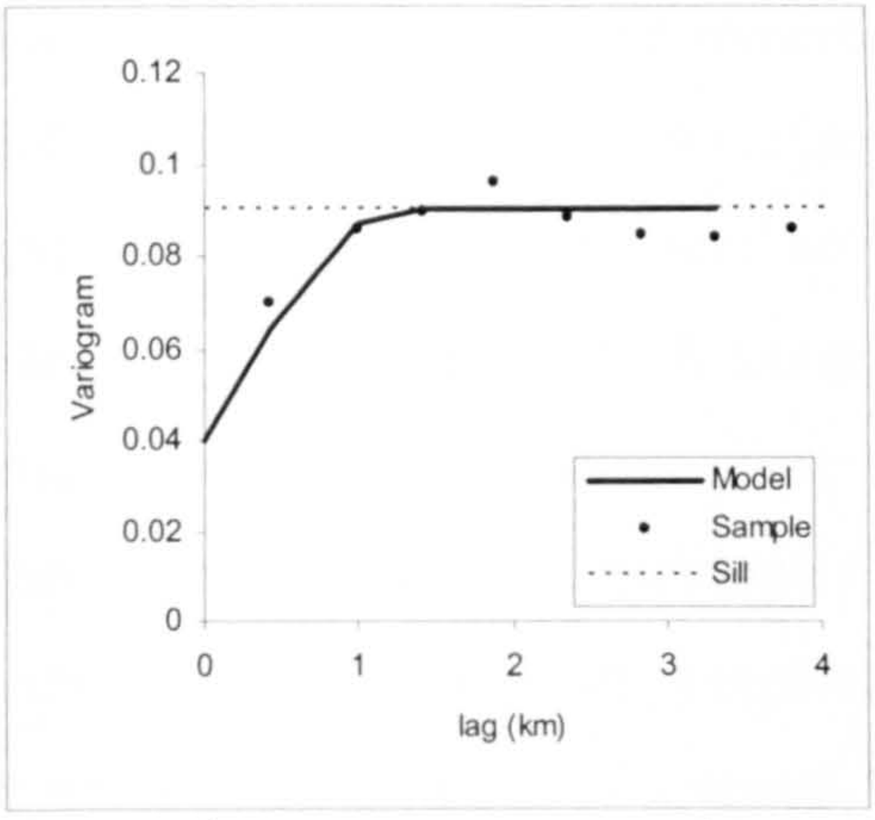


39th percentile (0.55mg.kg^{-1})

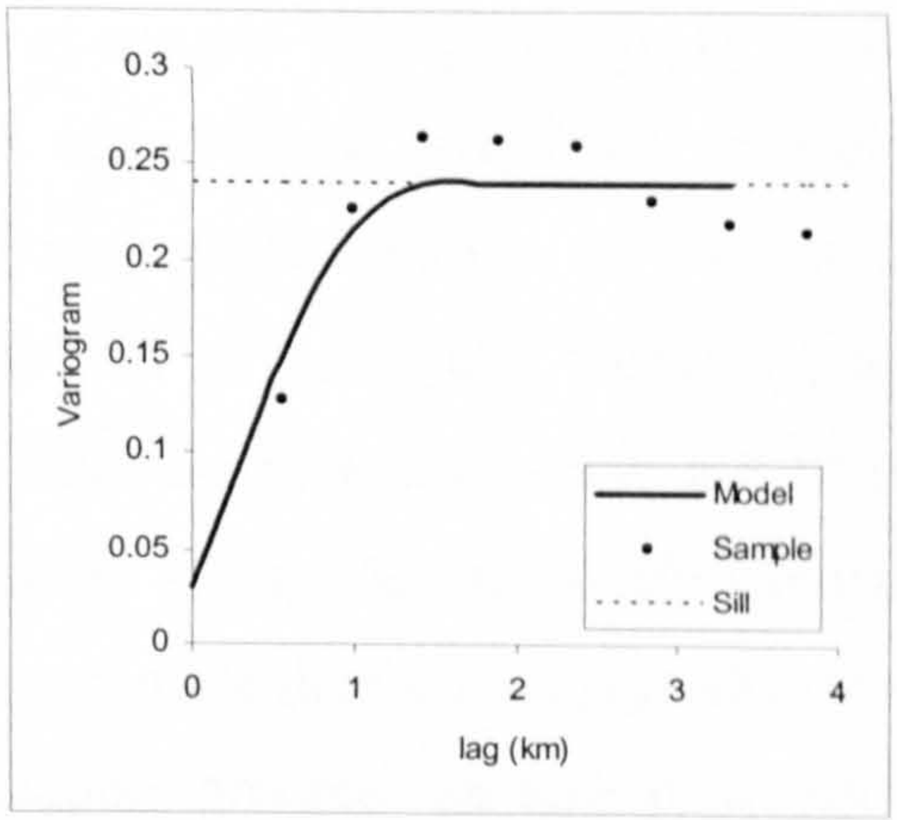


89th percentile (0.95mg.kg^{-1})

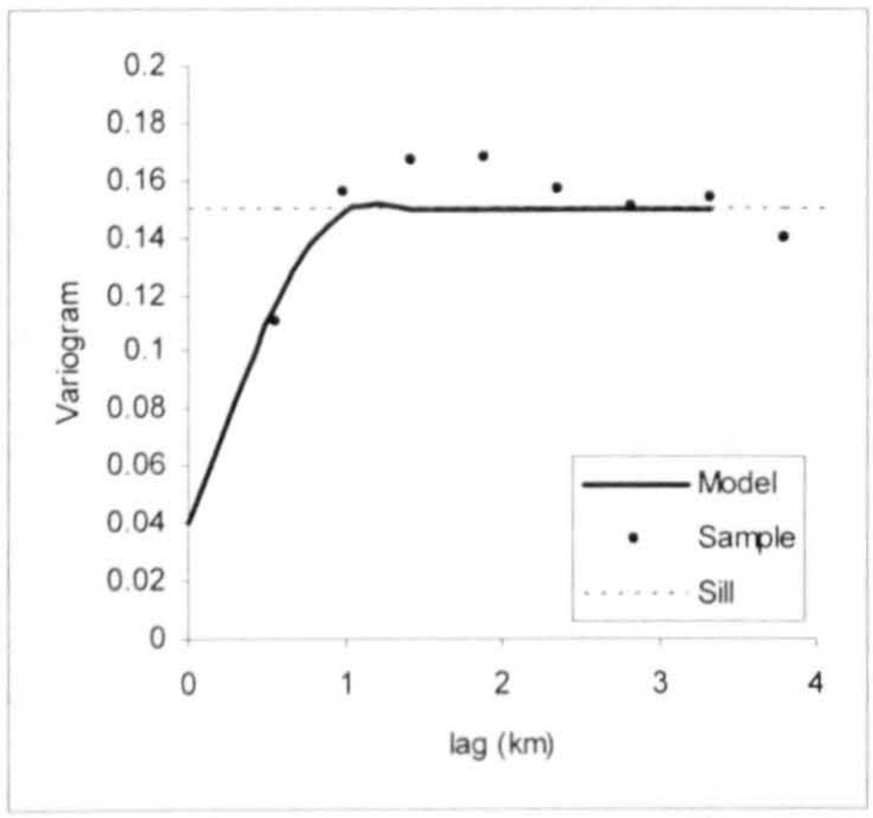
Figure C 2. Indicator variogram maps for Cd at six cut offs.



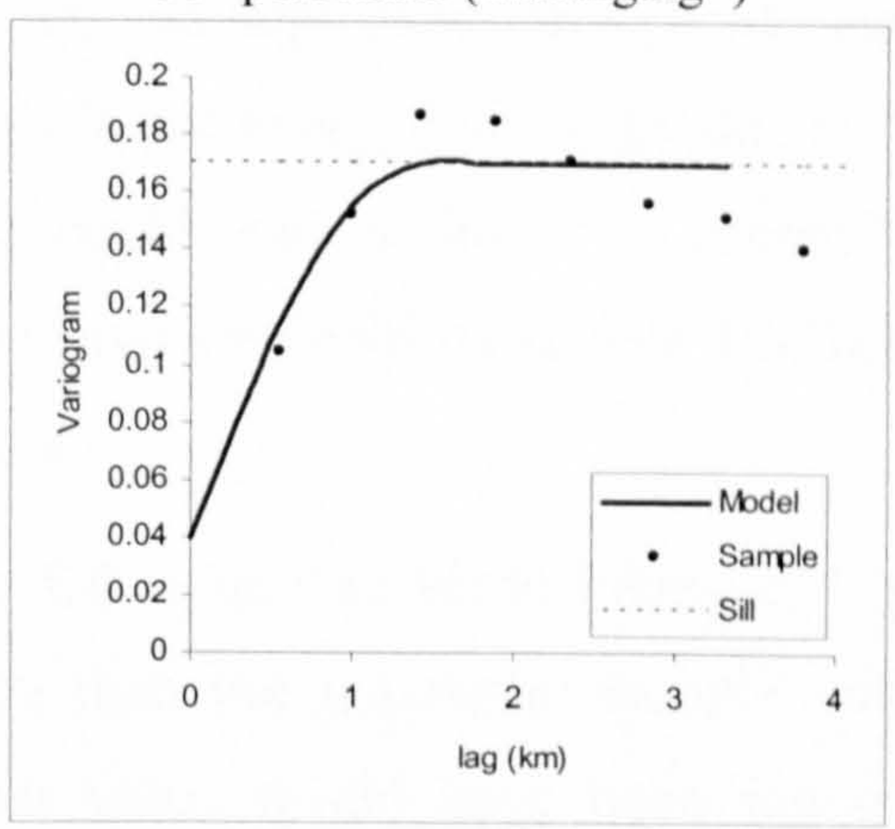
10th percentile (0.35mg.kg-1)



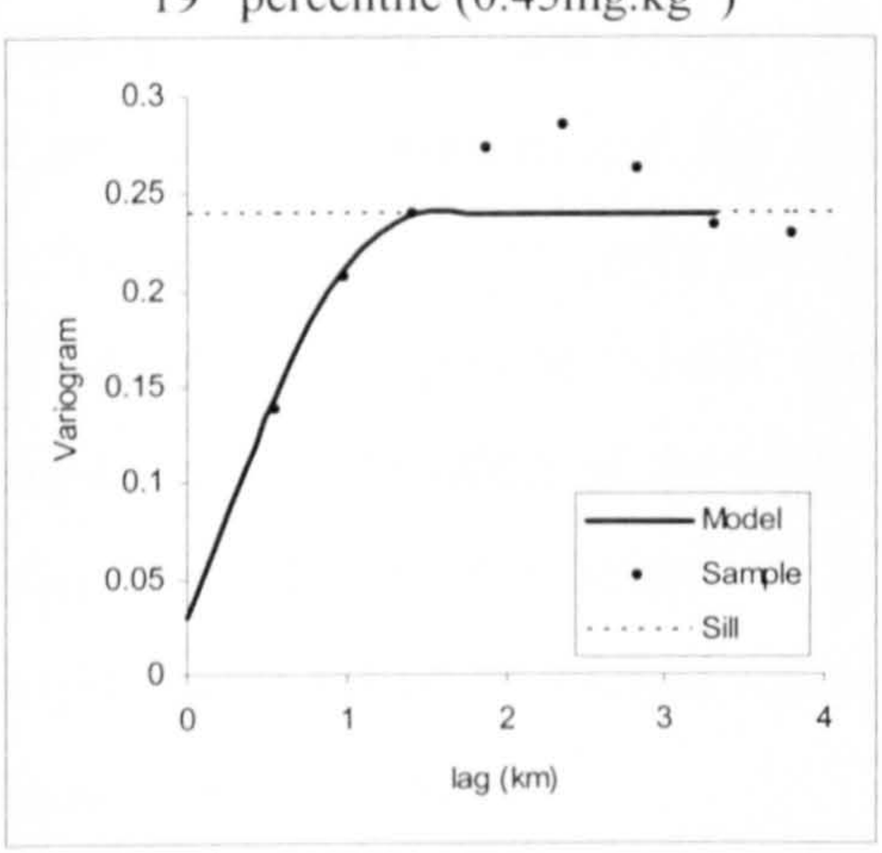
61st percentile (0.65mg.kg-1)



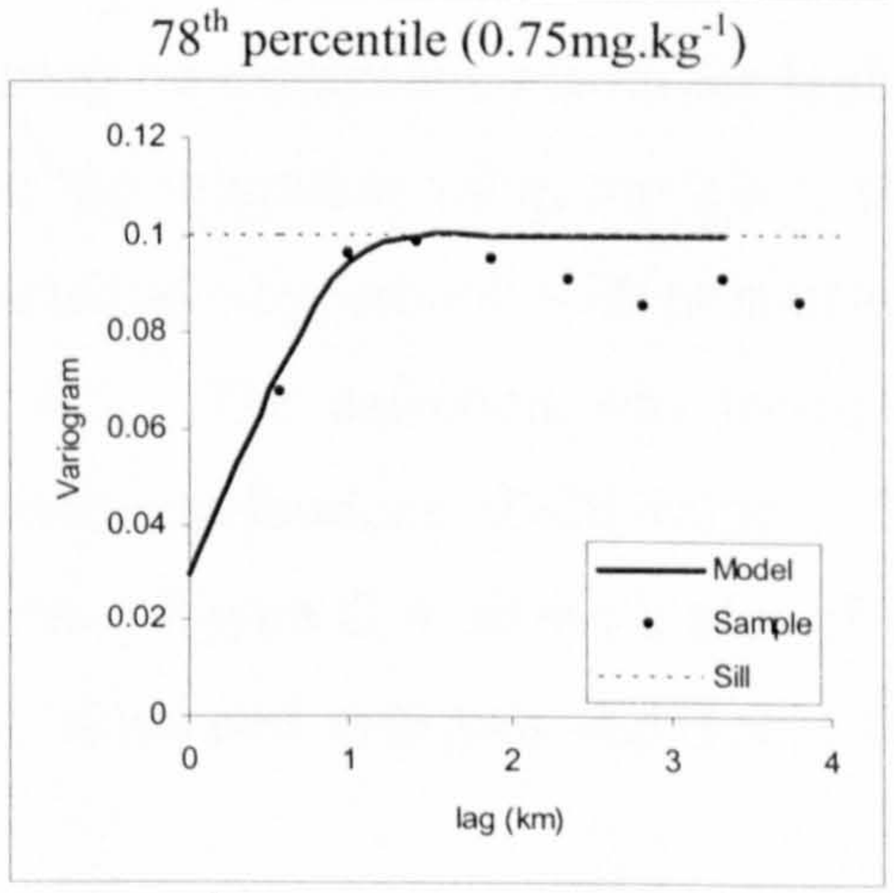
19th percentile (0.45mg.kg-1)



78th percentile (0.75mg.kg-1)



39th percentile (0.55mg.kg-1)



89th percentile (0.95mg.kg-1)

Figure C 3. Sample and model indicator variograms for Cd.

C.3 Running and validating simulations; Cd

Simulations were carried out on a 25mx25m square grid with 330 cells in the East-West direction and 160 North South using GSLIB SISIM routine (Deutsch & Journel, 1998). Each simulated value was taken to be representative of the entire 25x25m square. The ordinary kriging routine was used with a moving window neighbourhood of radius

3,000m. The minimum and maximum numbers of sample data points to be used in simulation of a grid point were set to 0 and 8 respectively, while the number of previously simulated nodes was set to 10. These parameters ensure that before any nodes are simulated, up to 8 sample data points will be used, and thereafter, once simulated nodes are available, they will be taken into account also, up to a 8:10 mix between sample data and previously simulated nodes. Using the previously simulated nodes ensures the joint ccdf is reproduced and using sample data assists reproduction of the sample histogram. To speed up the simulation process, sample data can be optionally allocated to grid nodes prior to simulation, and kept fixed throughout. As in previous case studies, this was deemed inappropriate since every simulated value at that point would be the same with no uncertainty and could lead to 'holes' in probability maps at a later stage i.e. that grid point will always have a probability of 0 or 1 of being over a threshold, no matter what the results around it.

On building the ccdfs for each cell, the minimum Cd value was set to 0.1mg.kg^{-1} (the detection limit) and the maximum to 10% more than the maximum sample value, assuming that it was unlikely that the maximum value would have been found in sampling. The upper and lower tails of the ccdf must be extrapolated from the highest cut off to the maximum and the lowest cut off to the minimum using models. After much trial and error, the upper tail model was selected as a hyperbolic with parameter 3 and the lower a power model with parameter 0.5. The selection was based on comparing the statistics of the sample and simulated distributions of Cd values. The closer the statistics, the better the model parameters. Figure C 4 shows a plot of the sample value statistics versus the average of the simulated statistics and Table C 3 summarises the statistics.

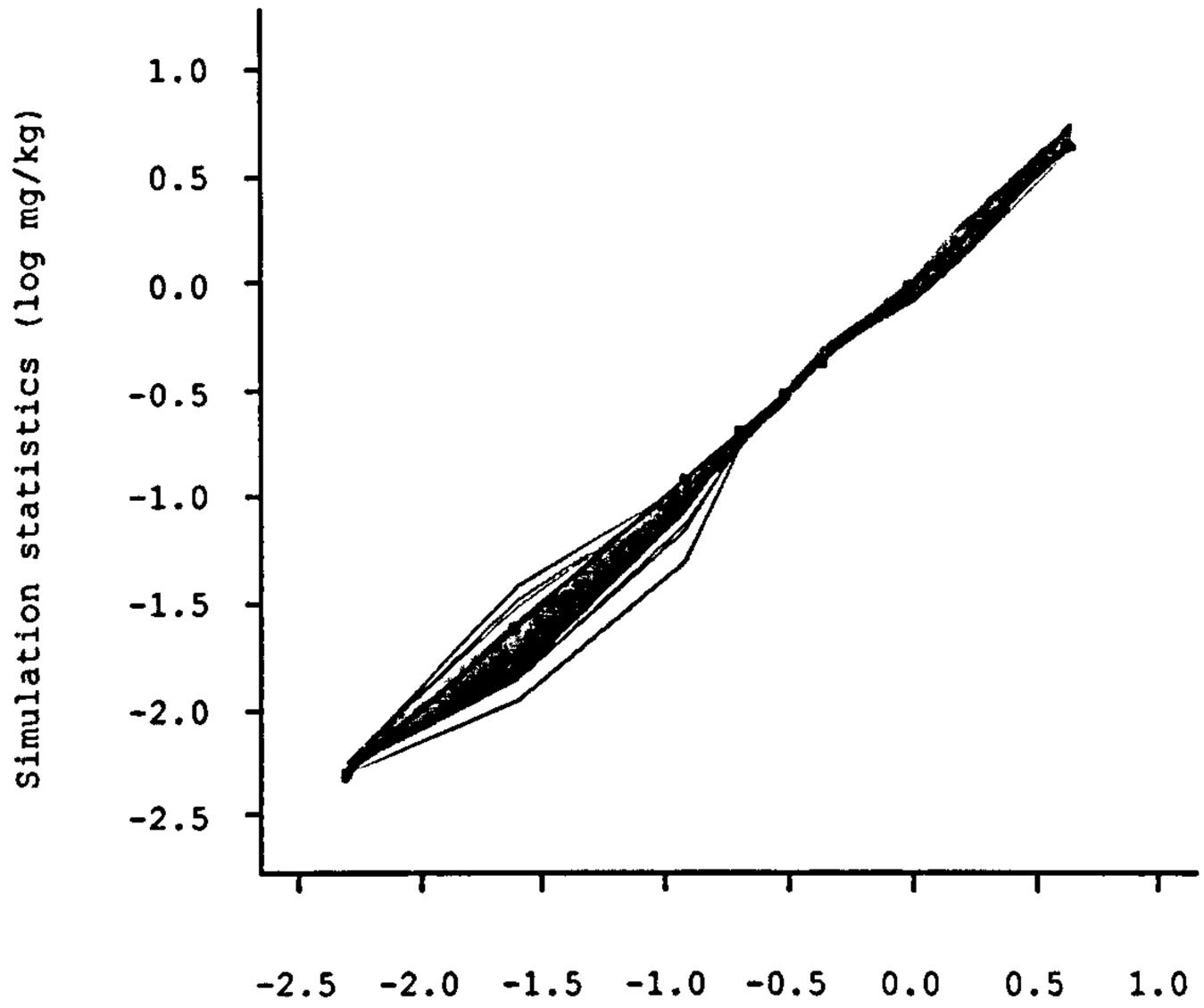


Figure C 4. Sample Cd statistics plotted against simulation statistics for first 50 simulations. Black diagonal represents perfect match. Statistics plotted are: min, P1, P5, P10, P25, P50, P75, P90, P95, P99, max.

Table C 3. Statistics of the Cd sample data and the average statistics from 100 simulations ($mg.kg^{-1}$).

	mean	stdev	min	p1	p5	p10	p25	p50	p75	p90	p95	p99	max
Samples	0.65	0.30	0.1	0.1	0.2	0.4	0.5	0.6	0.7	1.0	1.2	1.9	1.9
Simulations 25m grid	0.64	0.31	0.10	0.10	0.18	0.36	0.48	0.59	0.73	0.96	1.19	2.03	2.10
Simulations 250m grid	0.64	0.23	0.20	0.24	0.33	0.39	0.49	0.60	0.74	0.98	1.13	1.29	0.36

Indicator procedures at multiple cut offs with different indicator variogram models can result in order relation problems, whereby $(P > z_c) < (P > z_{c+1})$, which is clearly impossible. Order relation problems are corrected within the SISIM programme. In the order of 33% of ccdfs require corrections, with an average magnitude of correction in the order of 0.002.

In order to assess the goodness of the simulations the method devised by Deutsch was followed (Deutsch, 1997). This method produces Accuracy, Precision and Goodness statistics between 0 (worst case) and 1 (best case) by comparing the sample values with the simulated values at that point. An accurate probability distribution is one where the fraction of true values falling in the p-interval exceeds p for all p in [0,1] and a precise (and accurate) distribution is measured by the closeness of the fraction of true values to

p for all p in [0,1]. Goodness is a measure of departure from the 45degree line on the accuracy plot and weights inaccurate results twice, as they are more consequential than accurate ones. The results gave A=1, P=0.46 and G=0.73. An accuracy plot is shown in Figure C 5. In addition, a histogram is plotted of where in the distribution of simulations, the true value lies. First of all, the 135 grid nodes nearest to the sample points were selected as representing the simulated distribution of soil values for that sample point. The plot is shown in Figure C 6. A chi^2 test was significant, suggesting that the simulations differ significantly from the expected, and this looks to be the case, with the majority of sample values falling in the central percentile classes. This reflects the Accuracy and Precision results. Despite this problem, this simulation model was used, because an improvement could not be found. The simulations will suffice for demonstrating the spatially-evaluated quantitative intake assessment method.

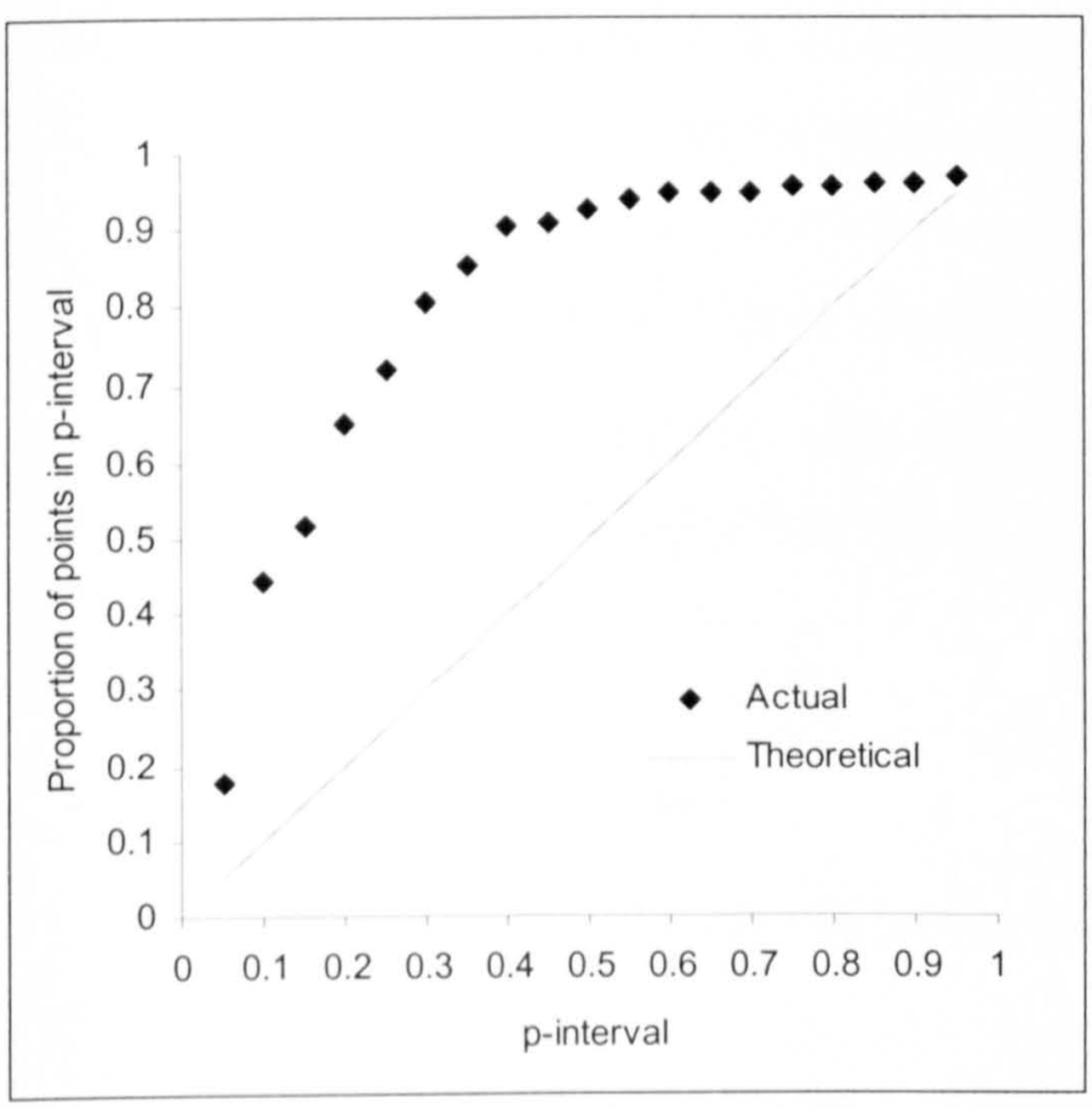


Figure C 5. Accuracy plot for Cd simulations.

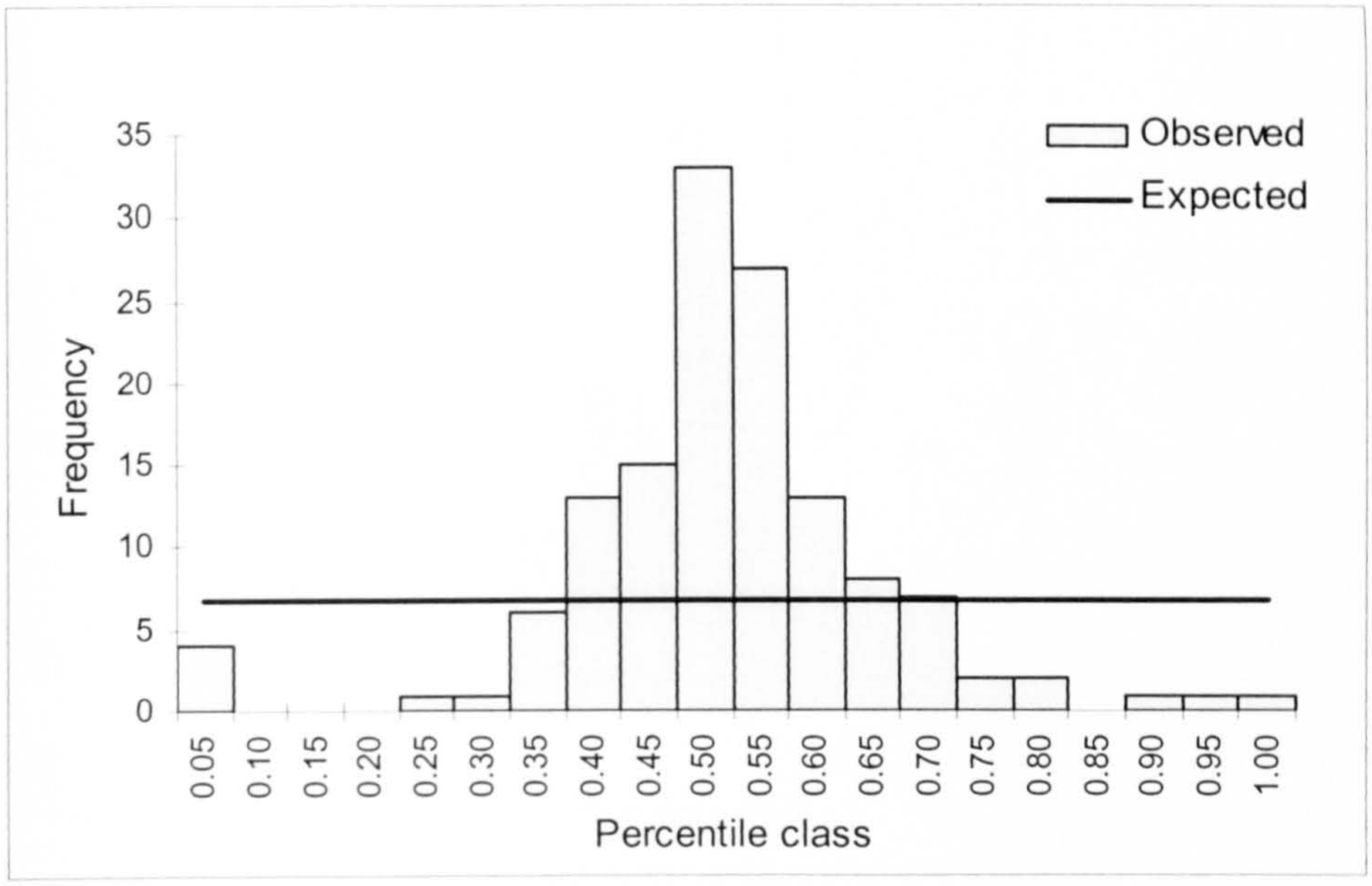


Figure C 6. Distribution of where in 100 simulated Cd values the true Cd value falls.

1,000 simulations were produced in all based on the fact that cumulative map statistics (mean, median and variance) had stabilised and at individual grid points the 'space of uncertainty' had reached a maximum width i.e. the interval between the minimum and maximum simulated value at that point increased no more.

To demonstrate the change of support effect, the 25x25m simulations were amalgamated into 250x250m blocks for use in the exposure analysis

C.4 Simple statistics; pH

Simple statistics of the 135 topsoil samples are given in Table C 4 and a histogram of the distribution in Figure C 7.

Table C 4. Simple statistics of the topsoil sample pH levels.

Mean	Stdev	Min	P1	P5	P10	P25	P50	P75	P90	P95	P99	Max
6.86	0.81	4.2	4.8	5.2	5.7	6.4	7.1	7.5	7.7	7.7	7.8	7.8

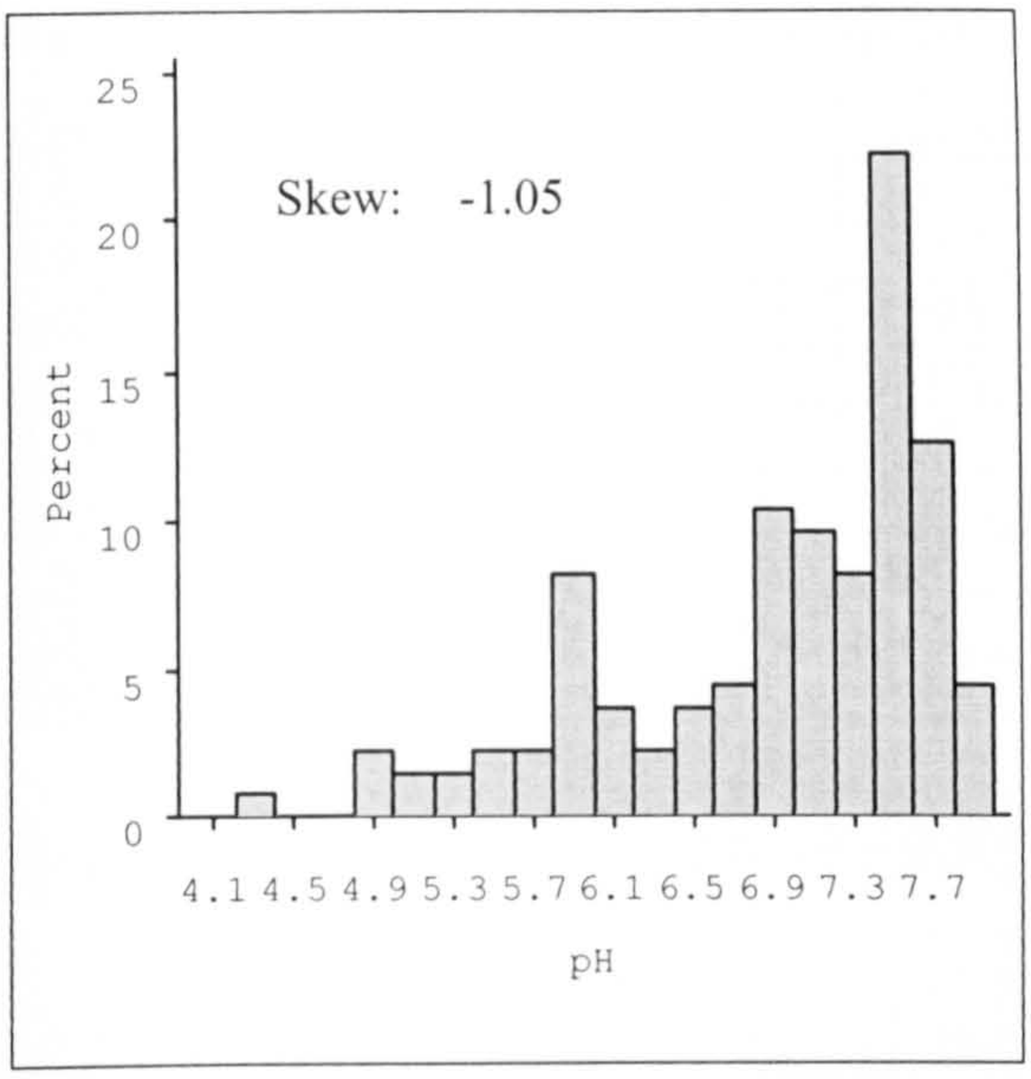


Figure C 7. Histogram of topsoil sample pH levels.

C.5 Variogram modelling; pH

Indicators were modelled using Isatis (Geovariances, 2002) at approximately the 9 deciles of the pH sample distribution. Actual percentiles modelled and the pH values at those percentiles are: 10th, 20th, 31st, 43rd, 53rd, 61st, 72nd, 83rd and 90th percentiles and pH 5.65, 6.05, 6.65, 6.95, 7.15, 7.35, 7.45, 7.55 and 7.65 respectively. The variogram maps are shown in Figure C 8. Most structure was discerned with a lag of 475m. Anisotropy is visible at some thresholds and an attempt was made to model it. However, the sample variograms were not particularly well structured and as the sills were forced to the theoretical variances to help reduce order relation problems, the variograms were hard to model in a multi-directional manner. The models fitted from one cut off to the next were not very consistent with anisotropic models and it was feared that this would lead to many order relation problems also. Given the findings in Chapter 5, i.e. that there appears little difference in results when modelled anisotropically and isotropically, omni-directional models were used, as it was felt that modelling the variograms without the anisotropy would lead to more consistent models from one cut off to the next. The models consist of spherical and nugget components for the lowest five cut offs and exponential and nugget components for the upper four cut offs. Sample and model variograms are shown in Figure C 9 and the model parameters are given in Table C 5.

Table C 5. Indicator model parameters for pH.

Cut off	model type	a	c	Nugget C_0
5.65	spherical[1]	1600	0.08	0.01
6.05	spherical[1]	2200	0.15	0.01
6.65	spherical[1]	2200	0.205	0.005
6.95	spherica[1]l	2500	0.24	0.005
7.15	spherical[1]	2900	0.23	0.02
7.35	exponential[2]	3100	0.21	0.03
7.45	exponential[2]	2900	0.15	0.05
7.55	exponential[2]	2900	0.08	0.06
7.65	exponential[2]	1300	0.05	0.04

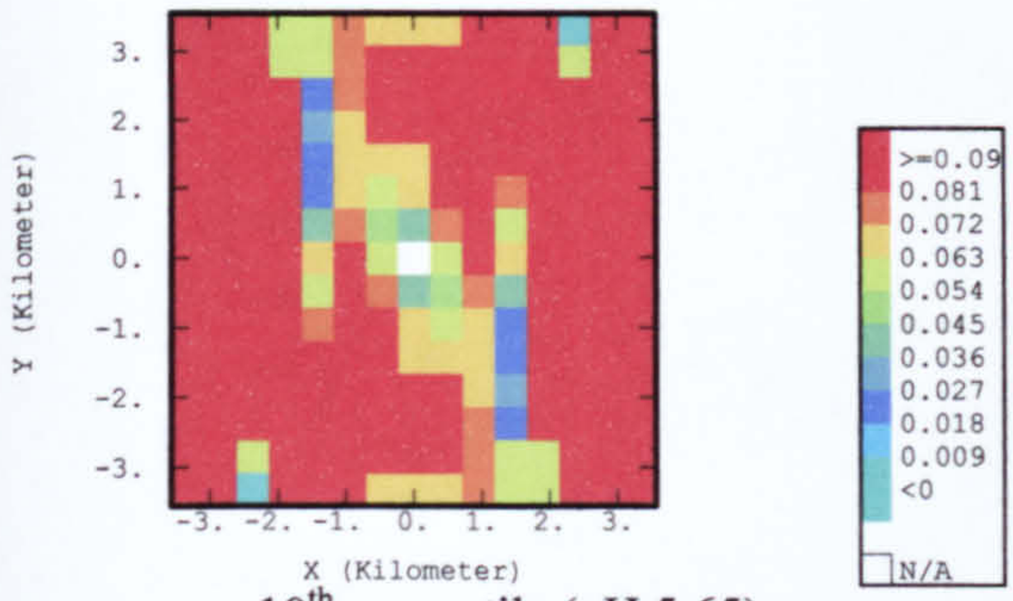
Spherical[1]: $\gamma(h) = C_0 + c.\left[1.5\frac{h}{a} - 0.5\left(\frac{h}{a}\right)^3\right]$ if $h \le a$

$\gamma(h) = C_0 + c$ if $h > a$

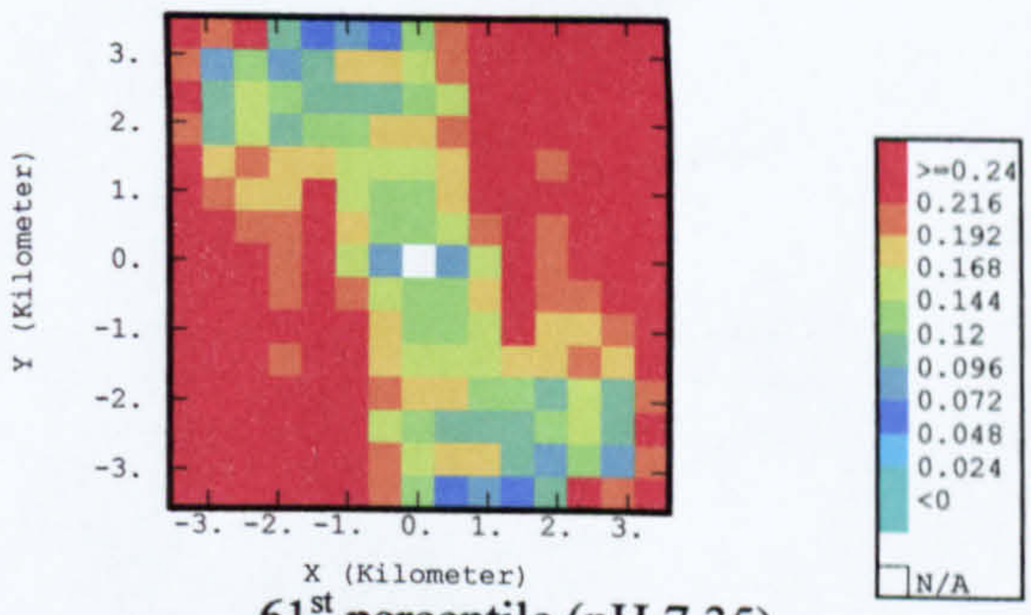
where h is the lag distance

Exponential[2]: $\gamma(h) = C_0 + c.\left[1 - \exp\left(-\frac{3h}{a}\right)\right]$

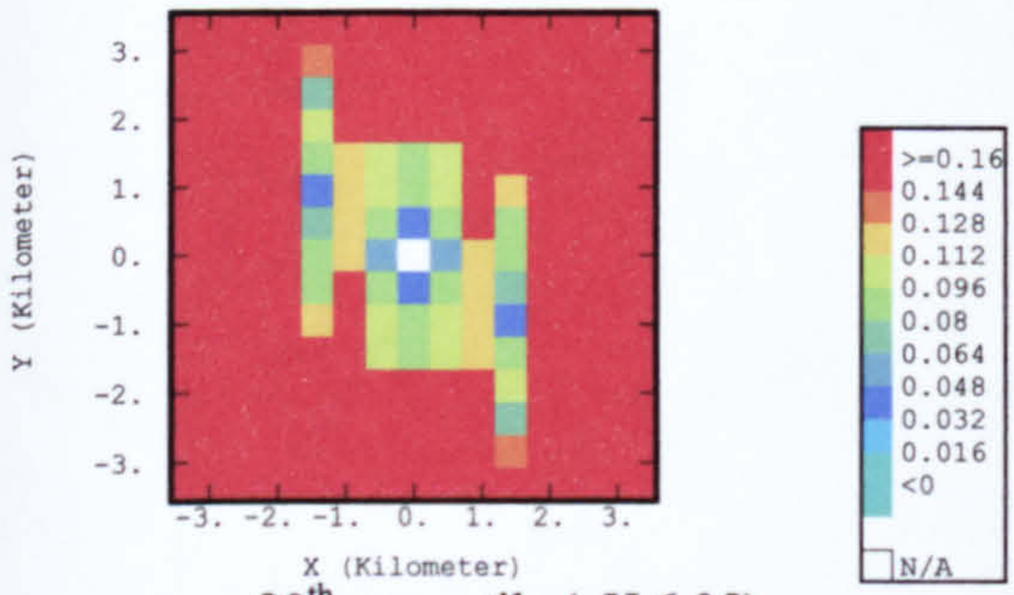
where h is the lag distance



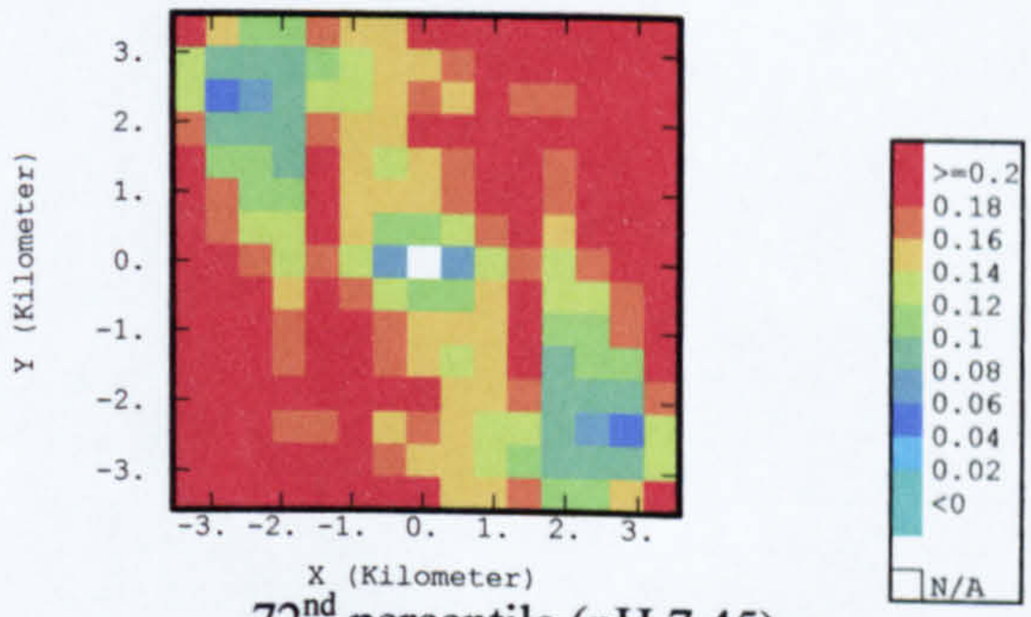
10th percentile (pH 5.65)



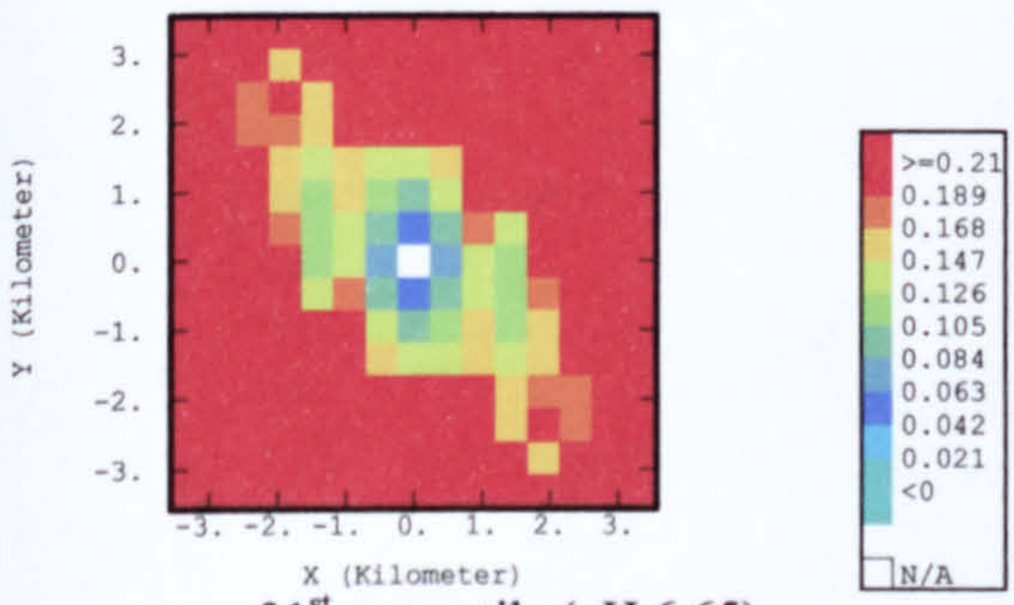
61st percentile (pH 7.35)



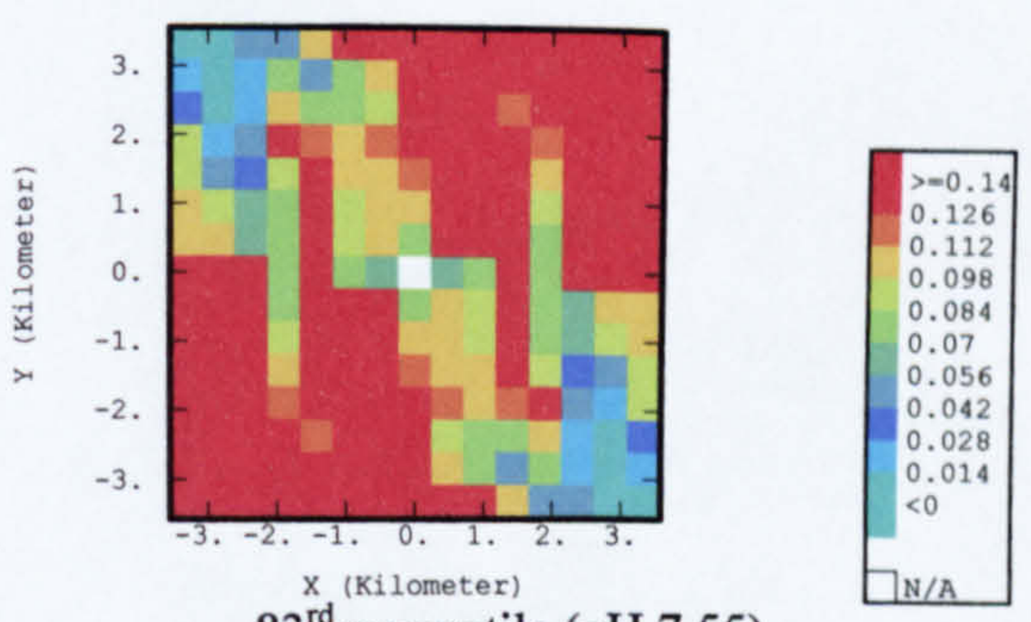
20th percentile (pH 6.05)



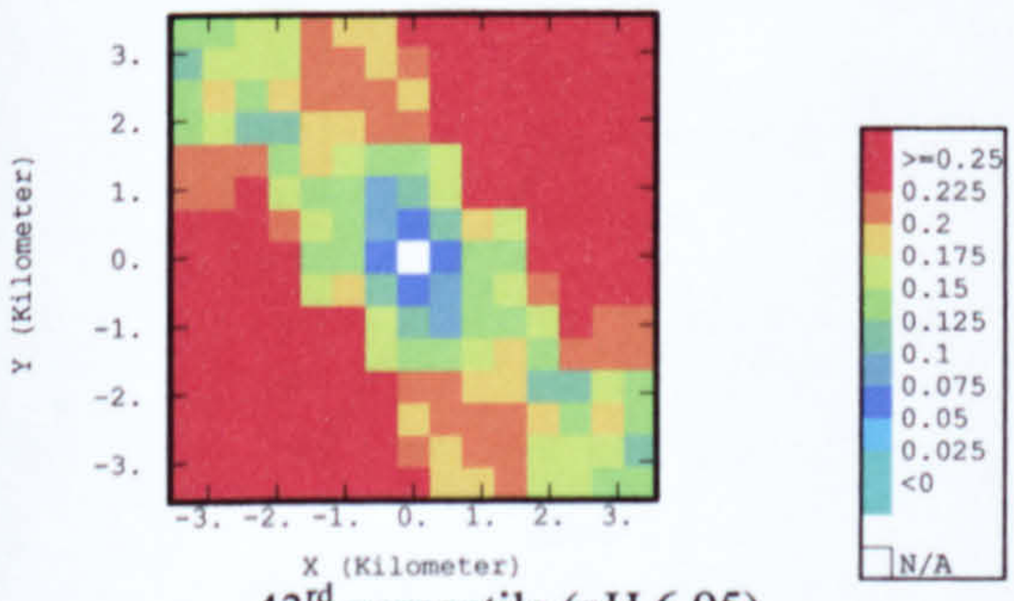
72nd percentile (pH 7.45)



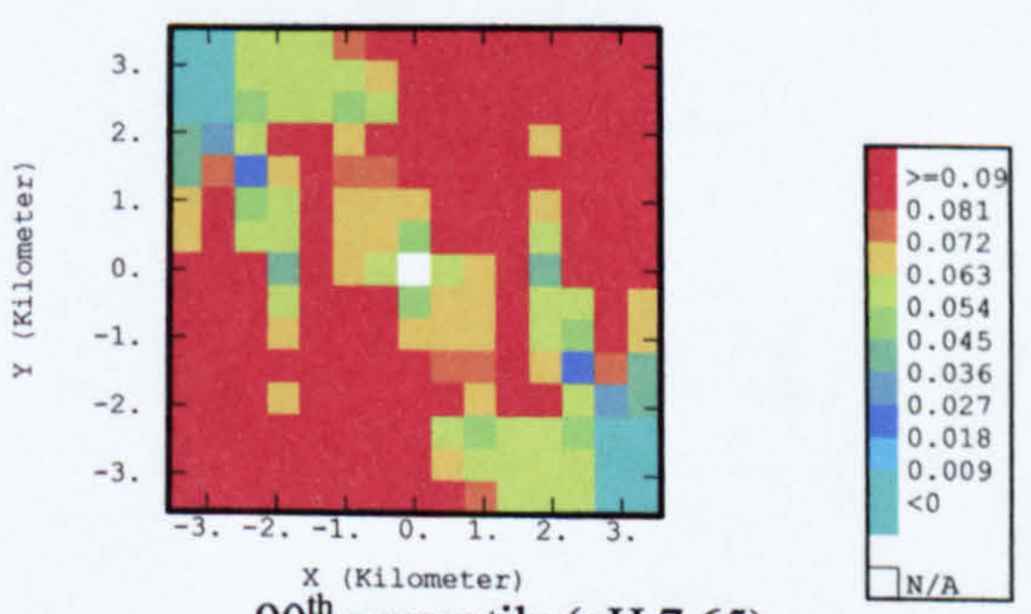
31st percentile (pH 6.65)



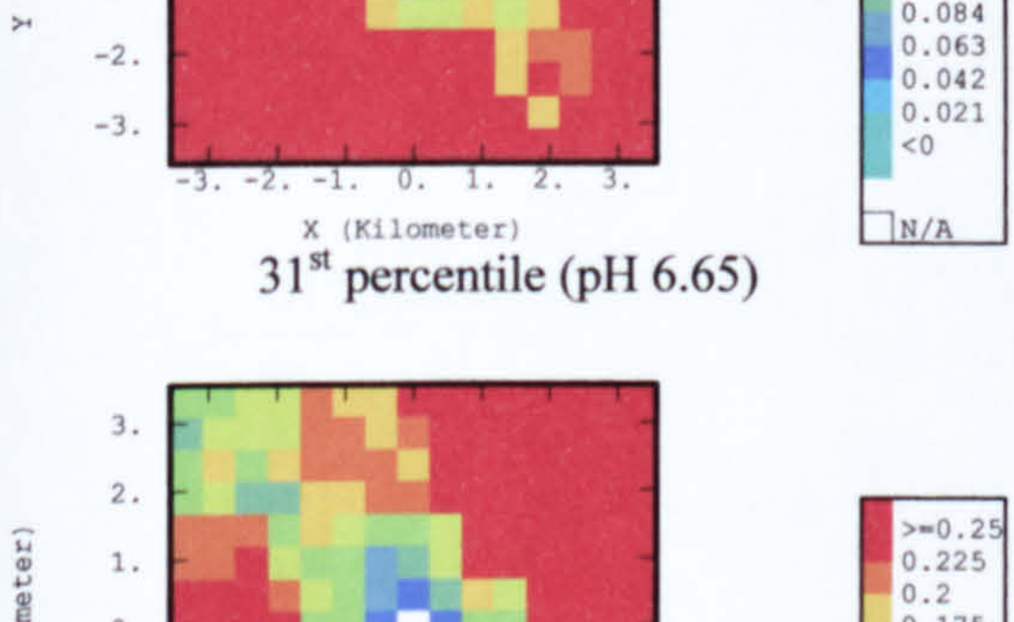
83rd percentile (pH 7.55)



43rd percentile (pH 6.95)

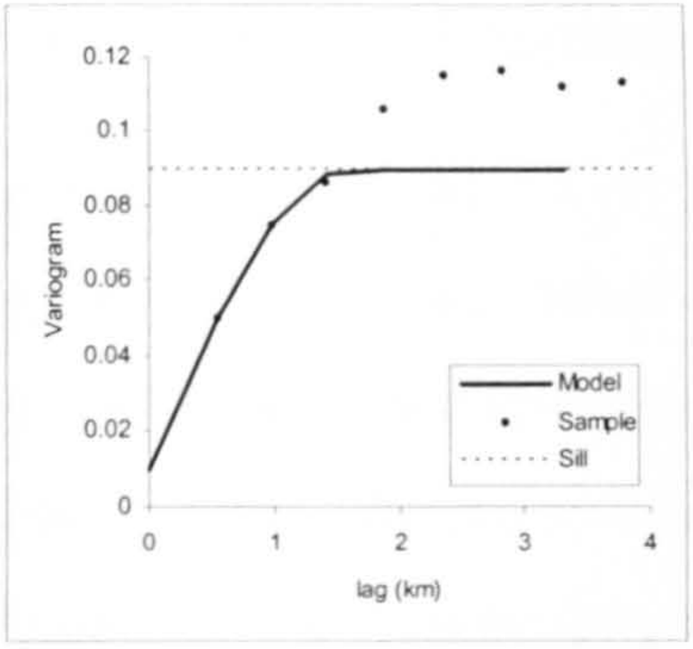


90th percentile (pH 7.65)

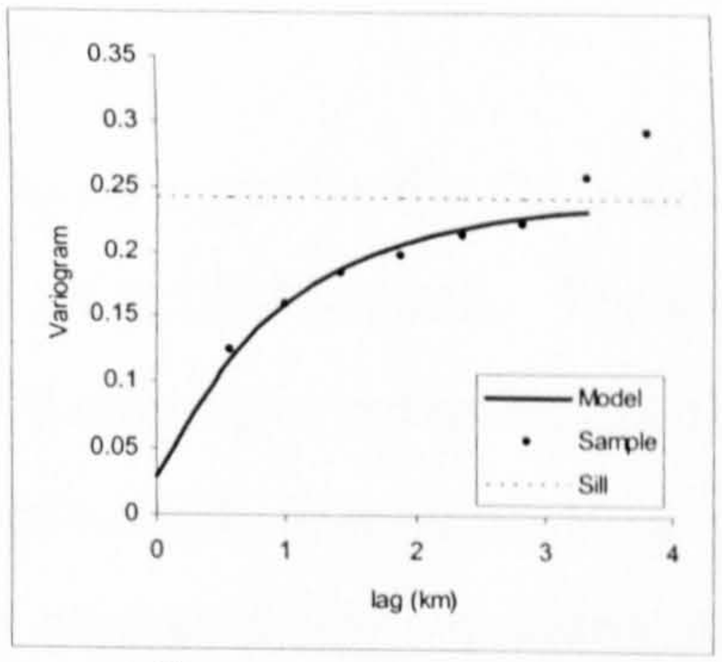


53rd percentile (pH 7.15)

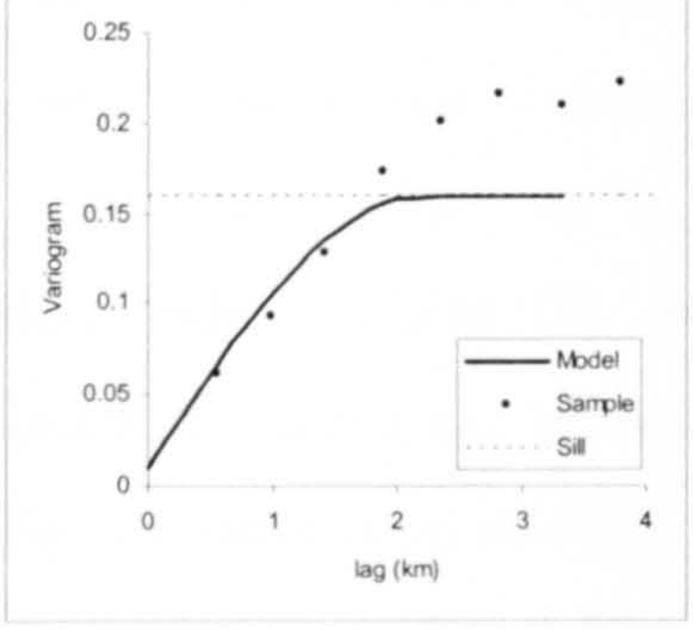
Figure C 8. Indicator variogram maps for pH at nine cut offs.



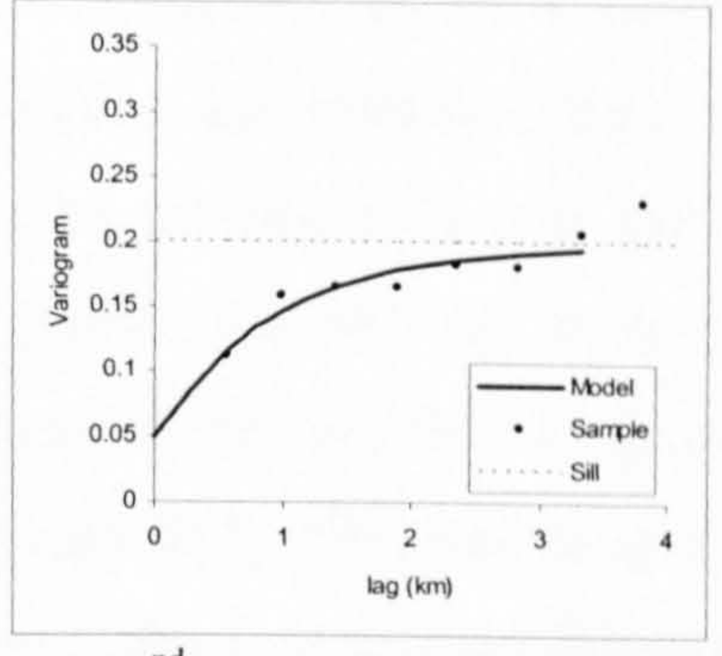
10th percentile (pH 5.65)



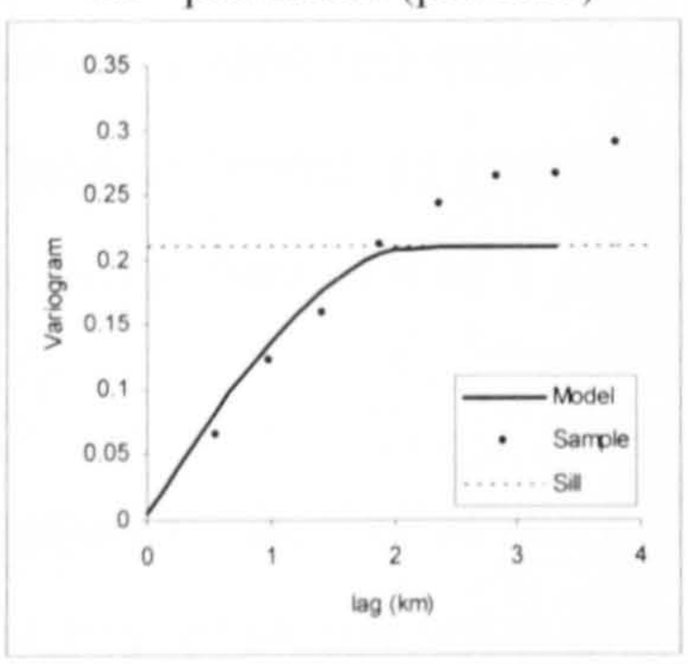
61st percentile (pH 7.35)



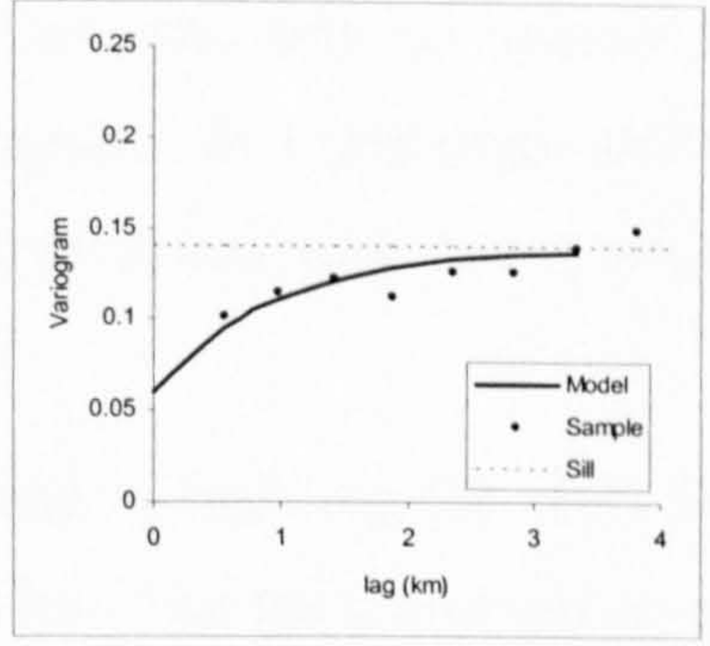
20th percentile (pH 6.05)



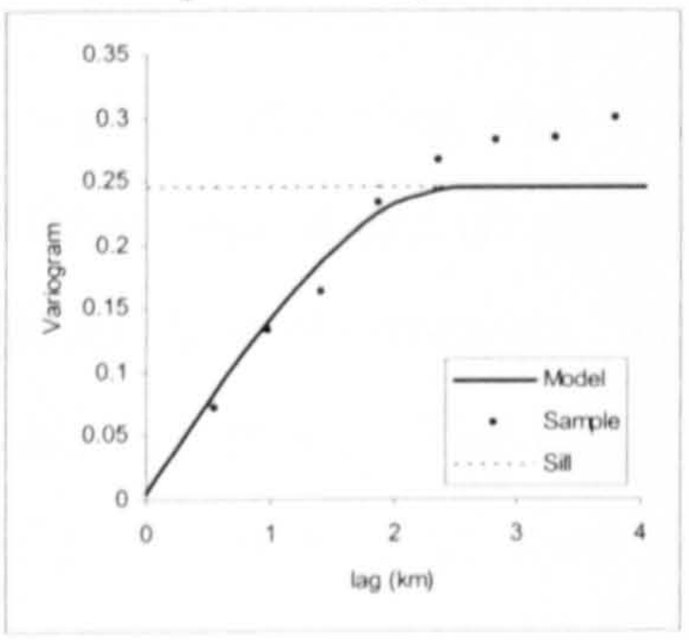
72nd percentile (pH 7.45)



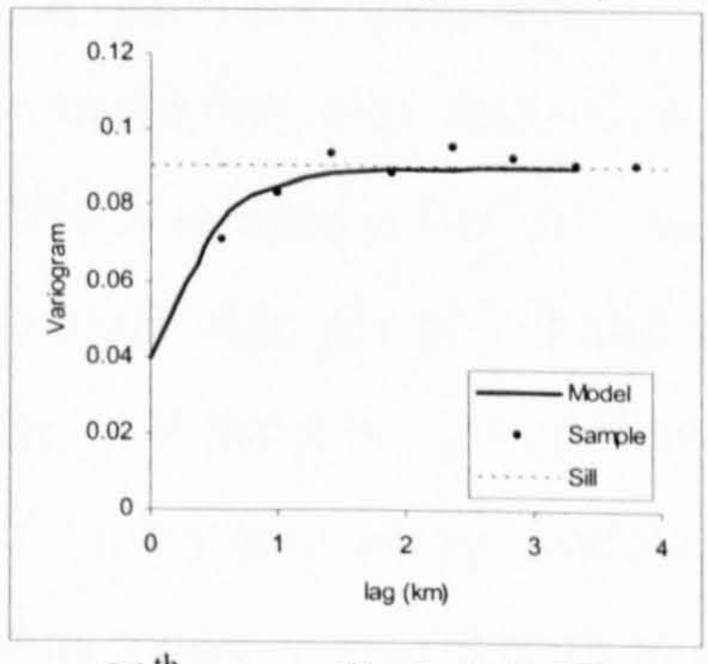
31st percentile (pH 6.65)



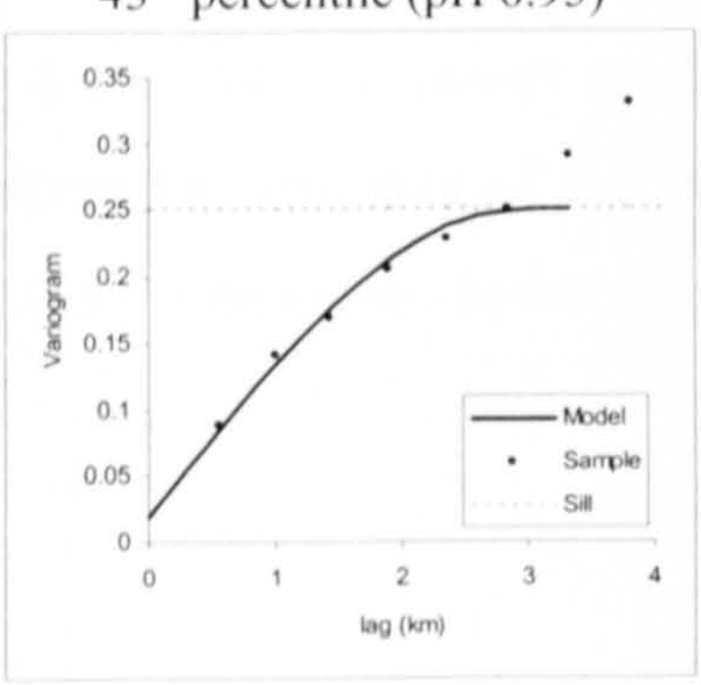
83rd percentile (pH 7.55)



43rd percentile (pH 6.95)



90th percentile (pH 7.65)



53rd percentile (pH 7.15)

Figure C 9. Sample and model indicator variograms for pH.

C.6 Running and validating simulations; pH

Simulations were carried out on the same 25mx25m square grid as the Cd simulations. Each simulated value was taken to be representative of a 25x25m cell. The ordinary kriging routine was used during the simulations, with a moving window neighbourhood of radius 3,000m. Minimum and maximum numbers of sample data points to be used in the simulations were set to 0 and 16 and the number of previously simulated nodes to 16. These parameters ensure that before any nodes are simulated, up to 16 sample data points will be used, and thereafter, once simulated nodes are available, they will be taken into account also, up to a 50:50 mix between sample data and previously simulated nodes. Using the previously simulated nodes ensures the joint ccdf is reproduced and using sample data assists reproduction of the sample histogram. To speed up the simulation process, sample data can be optionally allocated to grid nodes prior to simulation, and kept fixed throughout. However, this was deemed inappropriate since every simulated value at that point would be the same with no uncertainty and could lead to 'holes' in probability maps at a later stage i.e. that grid point will always have a probability of 0 or 1 of being over a threshold, no matter what the results around it.

pH is equal to the log of the H+ ion concentration, and on building the ccdfs for each cell, the minimum pH value was set to represent 10% less than the minimum sample H+ ion concentration and the maximum to 10% more than the maximum sample H+ ion concentration, assuming that it was unlikely that the minimum and maximum values would have been found in sampling. The figure of 10% is arbitrary, but may well lead to more realistic simulations. The result was an upper allowable pH of 7.9 and a lower allowable pH of 4.1. The upper and lower tails of the ccdf must be extrapolated from the 9th decile to the maximum and the 1st decile to the minimum using models. After much trial and error, the upper tail model was selected as a power model with parameter 0.5 and the lower a power model with parameter 2.5. The selection was based on comparing the statistics of the sample and simulated distributions of pH values. The closer the statistics, the better the model parameters. See Figure C 10 and Table C 6.

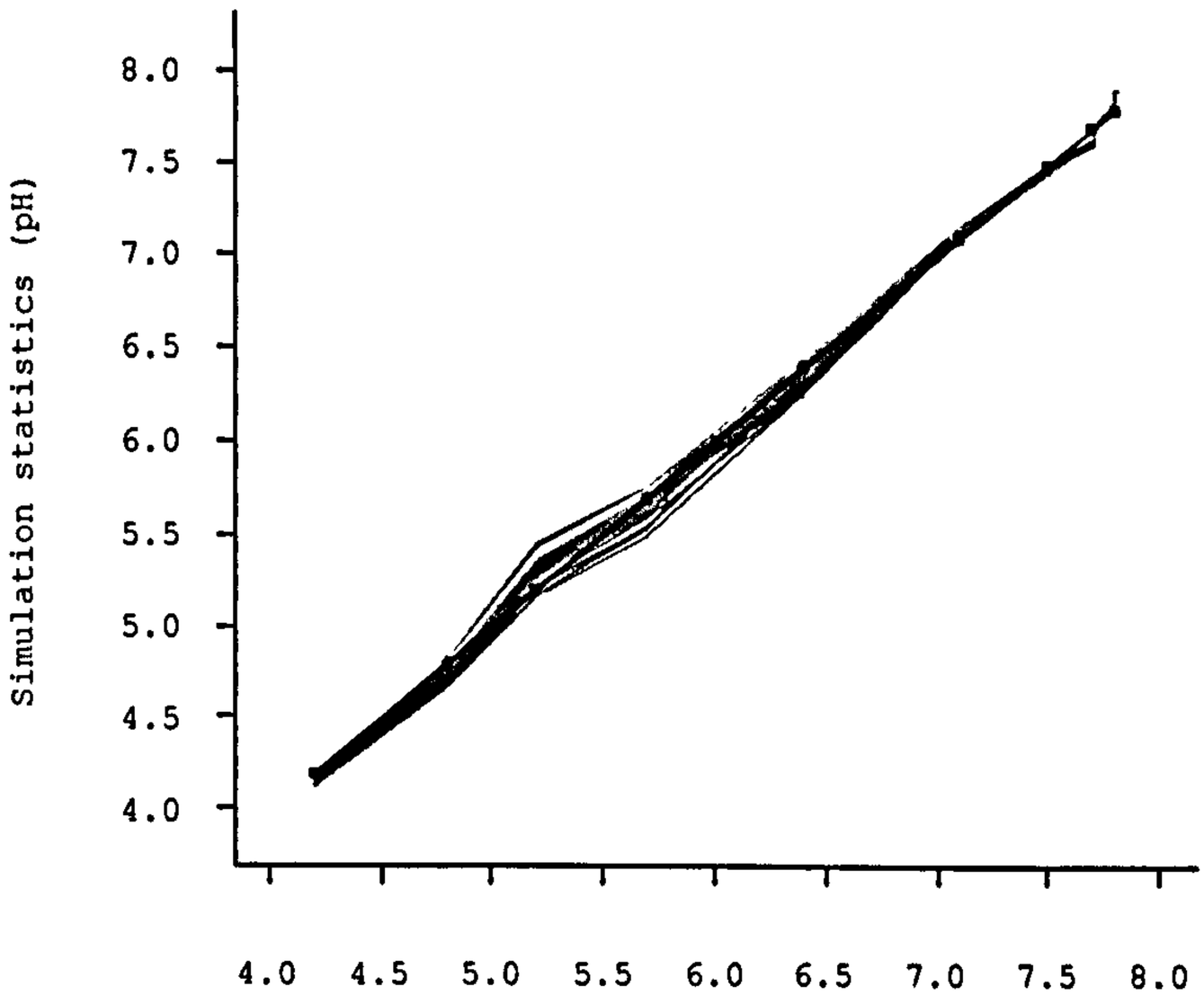


Figure C 10. Sample data statistics plotted against simulation statistics for first 50 simulations. Black diagonal represents perfect match. Statistics plotted are: min, P1, P5, P10, P25, P50, P75, P90, P95, P99, max.

Table C 6. Statistics of the sample data and the average statistics from the first 100 simulations for pH.

	Mean	Stdev	Min	P1	P5	P10	P25	P50	P75	P90	P95	P99	Max
Samples	6.86	0.81	4.2	4.8	5.2	5.7	6.4	7.1	7.5	7.7	7.7	7.8	7.8
Simulations 25m grid	6.86	0.79	4.1	4.7	5.3	5.7	6.4	7.1	7.5	7.6	7.7	7.8	7.9
Simulations 250m grid	6.86	0.71	5.2	5.2	5.4	5.7	6.4	7.1	7.5	7.5	7.6	7.7	7.7

Indicator procedures at multiple cut offs with different indicator variogram models can result in order relation problems, whereby $(P > z_c) < (P > z_{c+1})$, which is clearly impossible. Order relation problems are corrected within the SISIM programme. In the order of 55% of ccdfs require corrections, with an average magnitude of correction in the order of 0.001.

In order to assess the goodness of the simulations a method devised by Deutsch was followed (Deutsch, 1997). This method produces Accuracy, Precision and Goodness statistics between 0 (worst case) and 1 (best case) by comparing the sample values with the simulated values at that point. The results gave A=1, P=0.69 and G=0.85. An accuracy plot is shown in Figure C 11. In addition, a histogram is plotted of where in the distribution of simulations, the true value lies. First of all, the 135 grid nodes

nearest to the sample points were selected as representing the simulated distribution of soil values for that sample point. The plot is shown in Figure C 12. A chi^2 test was not significant, so it can be concluded that the results do not differ significantly from the expected. On the basis of these tests, the simulations were accepted as being satisfactory.

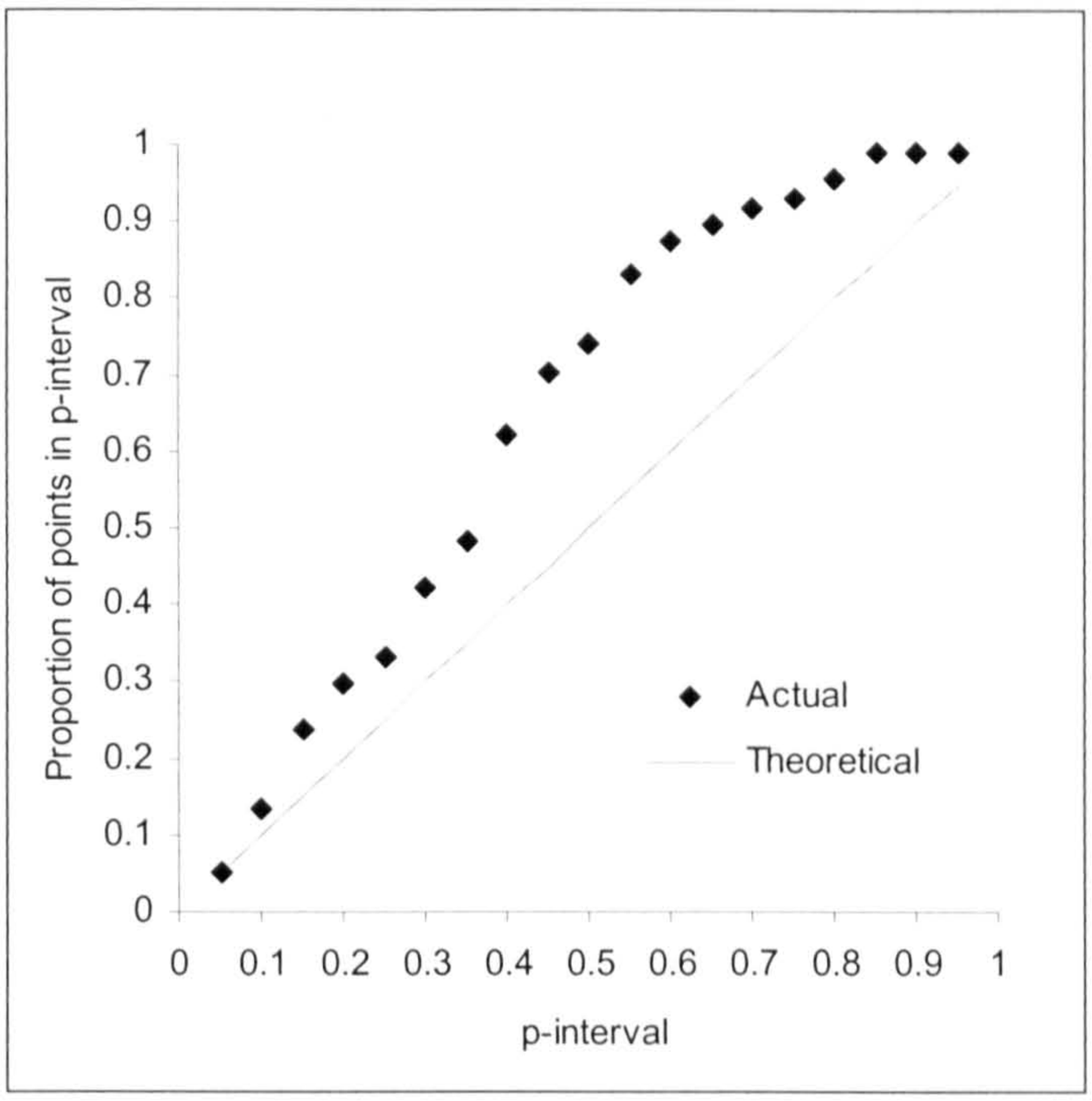


Figure C 11. Accuracy plot for pH simulations.

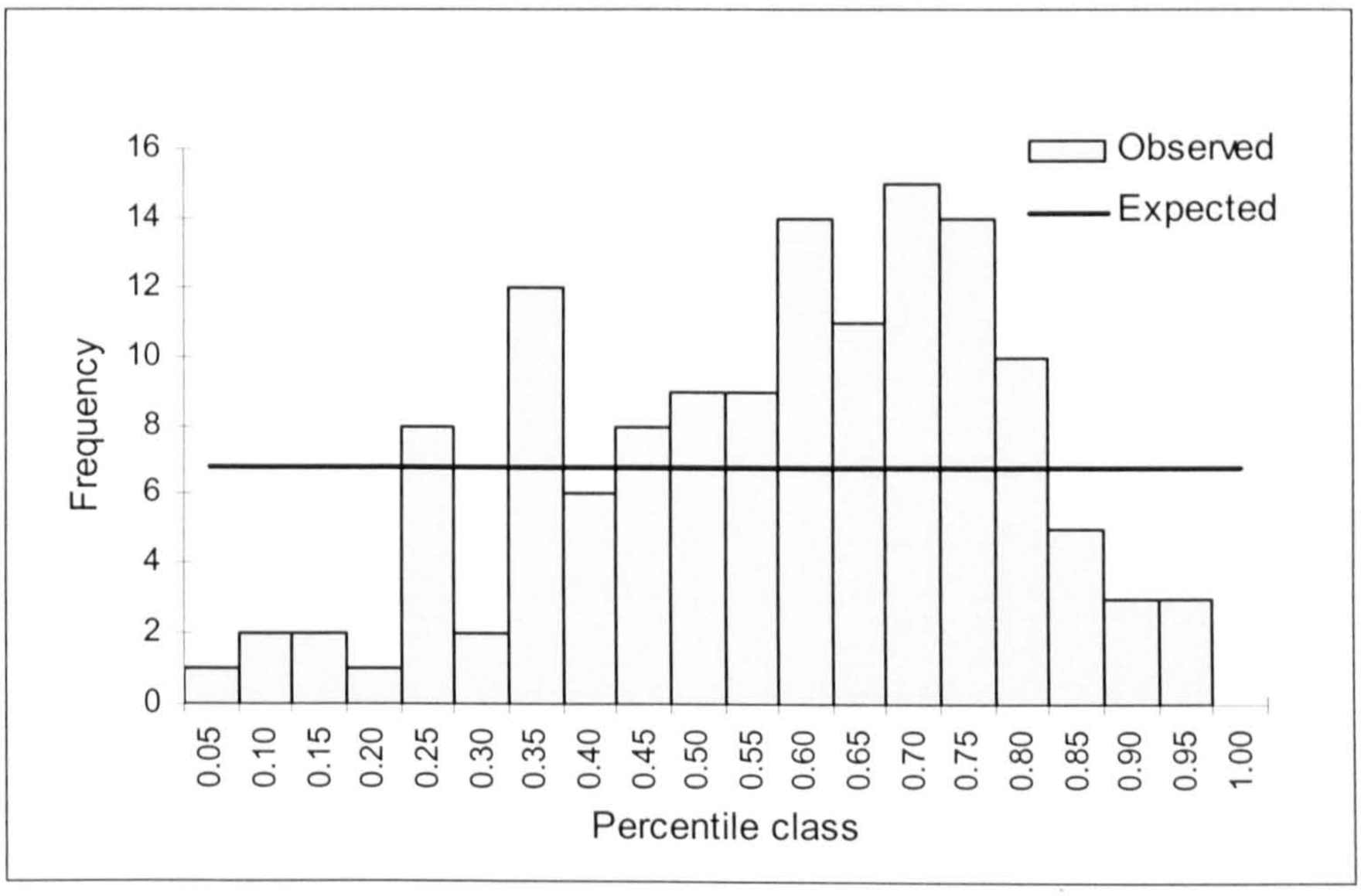


Figure C 12. Distribution of where in the 100 simulated pH values the true value falls.

1,000 simulations were produced in all based on the fact that cumulative map statistics (mean, median and variance) had stabilised and at individual grid points the 'space of uncertainty' had reached a maximum width i.e. the interval between the minimum and maximum simulated value at that point increased no more.

To demonstrate the change of support effect, the 25x25m simulations were amalgamated into 250x250m blocks for use in the exposure analysis. The statistics of the upscaled 250x250m simulations are also given in Table C 6. For the non-spatial analysis the site averages were used.

Appendix D Geostatistical procedure for Cadmium and pH in Stoke on Trent

D.1 Simple statistics; Cd

Simple statistics of the 744 topsoil samples are given in Table D 1 and a histogram of the distribution in Figure D 1.

Table D 1. Simple statistics of the Cd topsoil sample concentrations ($mg.kg^{-1}$).

Mean	Stdev	Min	P1	P5	P10	P25	P50	P75	P90	P95	P99	Max
1.70	2.01	0.4	0.4	0.4	1.0	1.0	2.0	2.0	2.0	2.0	5.5	43.0

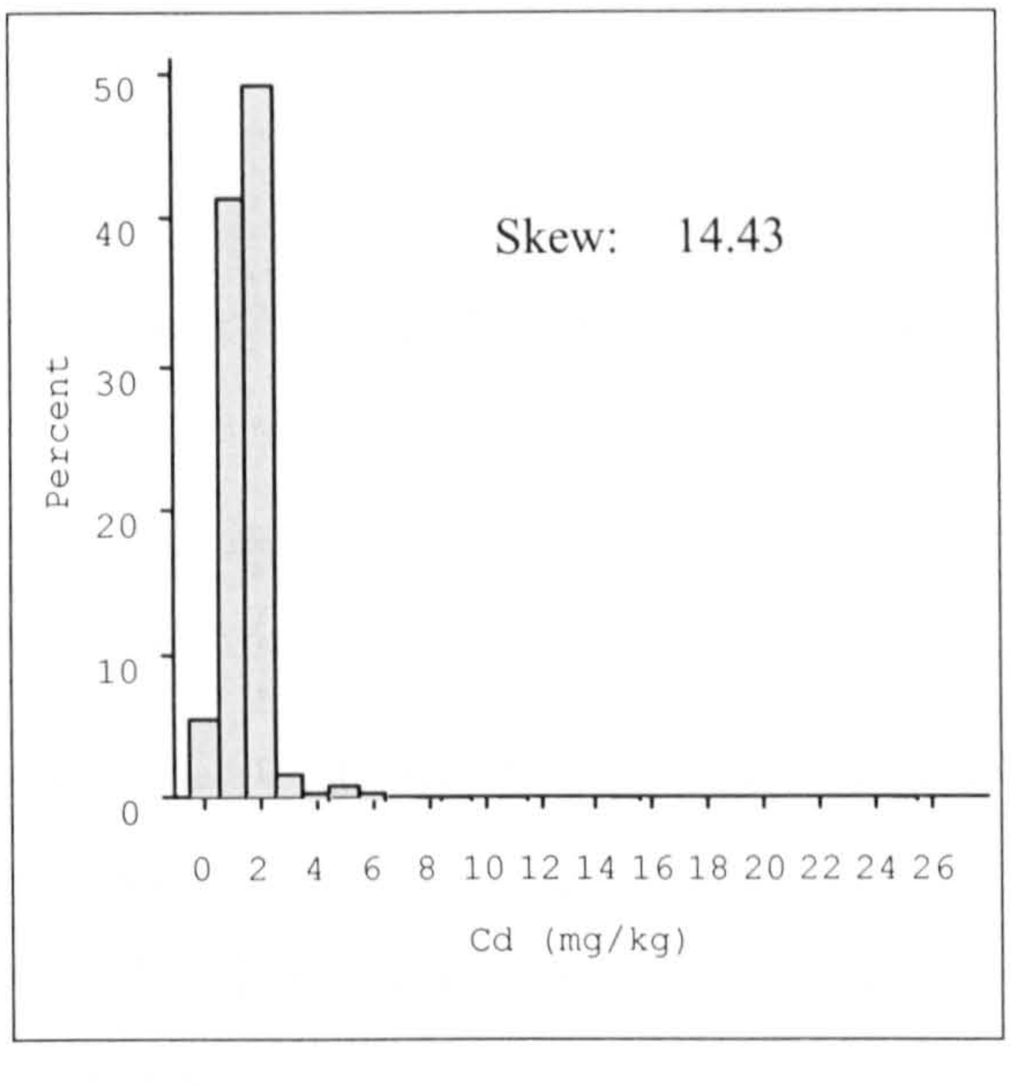


Figure D 1. Histogram of Cd topsoil sample concentrations ($mg.kg^{-1}$). One value of $43mg.kg^{-1}$ not shown.

D.2 Variogram modelling; Cd

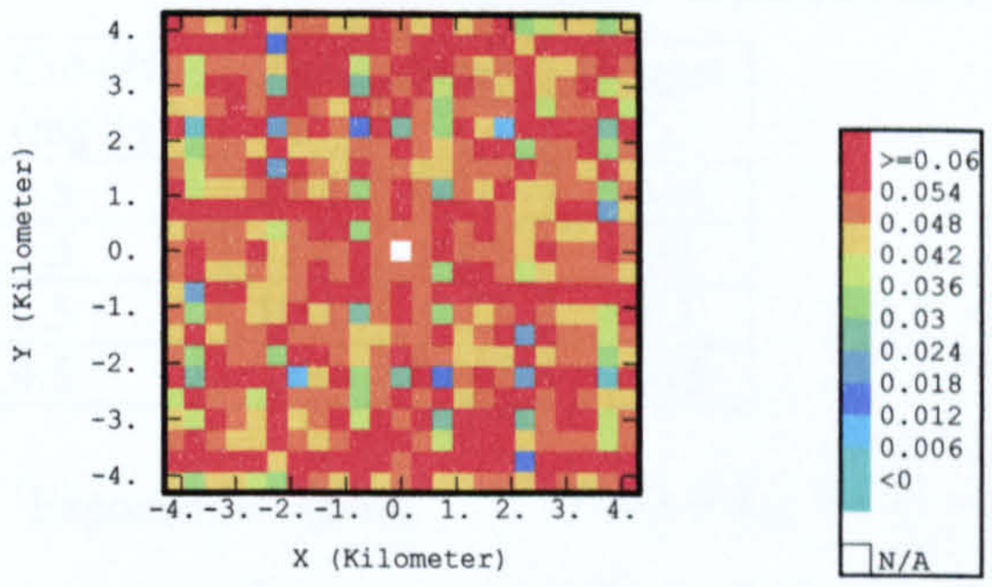
The minimum and maximum Cd sample values were recorded as 0.4 and $43mg.kg^{-1}$ respectively. In actual fact, all Cd measurements were made to the nearest $mg.kg^{-1}$, as

the analytical procedure was unable to make finer distinctions. 0.4 is approximately half the detection limit of the test and was used to represent 'non-detects' in the Stoke Cd data set. However, since all other readings were recorded to the nearest whole number, the 0.4 values were set to zero for the remainder of the analysis, as this was thought to be more consistent.

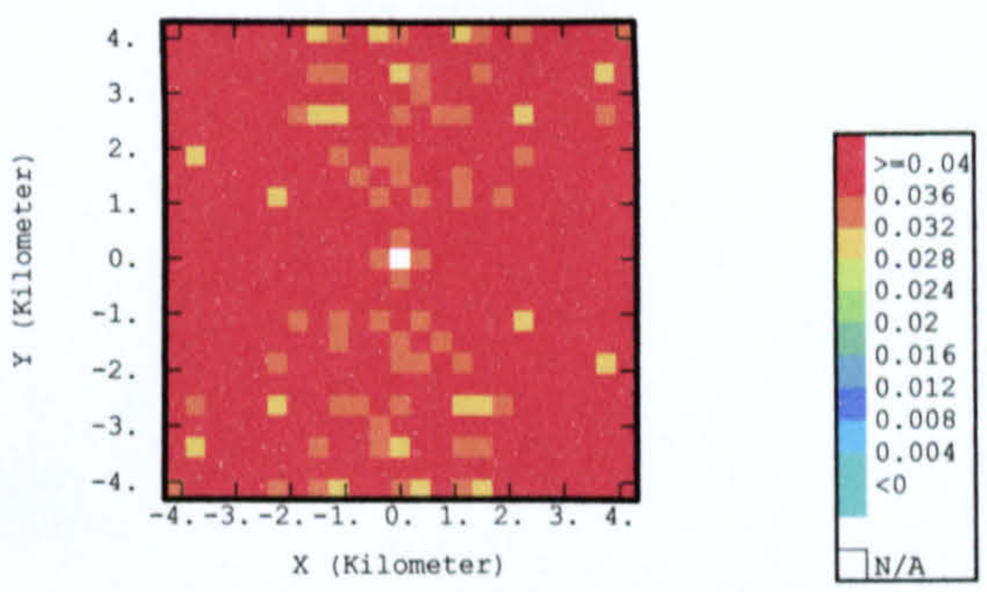
Since the analytical method is only accurate to the nearest $mg.kg^{-1}$, this led to many repetitive sample values and as such it is not practical to define the cdf through 10 deciles. Indicators were modelled at the 6^{th}, 47^{th}, 96^{th} and 98th deciles only, using cut off values of 0.5, 1.5, 2.5 and $4.5mg.kg^{-1}$ respectively. These cut off values were selected as being half way between the possible cut off values that would give the same proportion of 0s and 1s. For example, a cut off of 1.1 would give the same binary classification as a cut off of 1.9, since recorded values around the cut off are either 1 or $2mg.kg^{-1}$. $1.5mg.kg^{-1}$ was chosen as it is in the middle of the range of possible cut offs.

Another consideration is a practical issue of switching between the variogram modelling in Isatis (Geovariances, 2002) and the simulation procedure in GSLIB SISIM (Deutsch & Journel, 1998). Isatis uses the code of 1 for values greater than or equal to the cut off and GSLIB uses codes of 1 for values less than or equal to the cut off. Great care must be taken to ensure that the set of binary codes in Isatis is the exact mirror image of those used in GSLIB. Choosing a cut off which is equal to a recorded sample value will lead to inconsistencies between the two packages. Using a cut off of $1mg.kg^{-1}$, for example, in this case study would lead to very noticeable inconsistencies, since there are 307 recorded sample values of $1mg.kg^{-1}$ (41% of samples) which would mean 41% of codes would be inconsistent at this cut off.

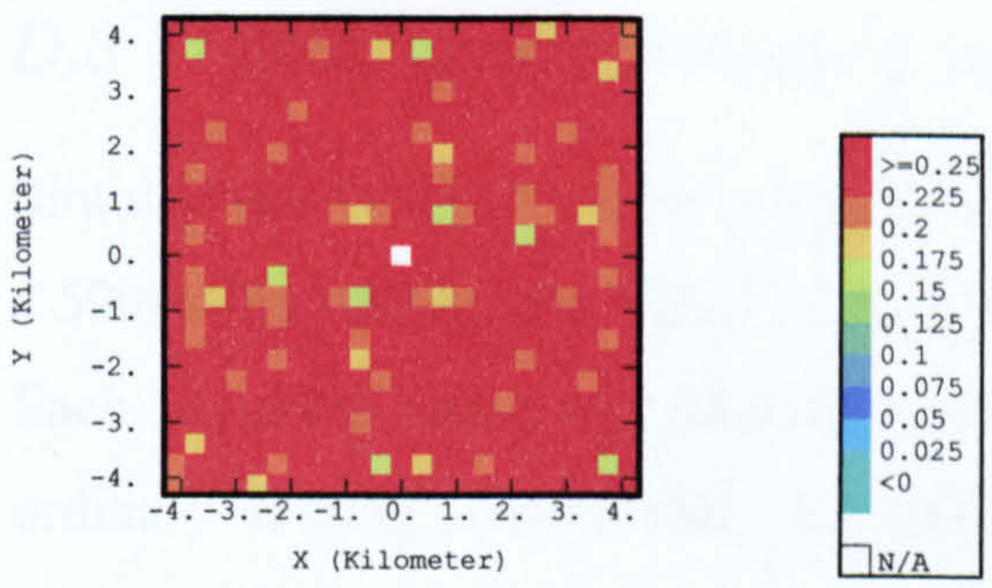
Sills of the indicator variogram models were set to the theoretical variances, to help reduce order relation problems. The sill values were 0.056, 0.25, 0.04 and 0.02 for the 0.5, 1.5, 2.5 and $4.5mg.kg^{-1}$ cut offs respectively. Little structure was found, most being discerned at a lag spacing of 375m and a lag tolerance of 50%. The models were made omni-directional. Variogram maps are shown in Figure D 2 and sample and model variograms in Figure D 3. Model parameters are given in Table D 2.



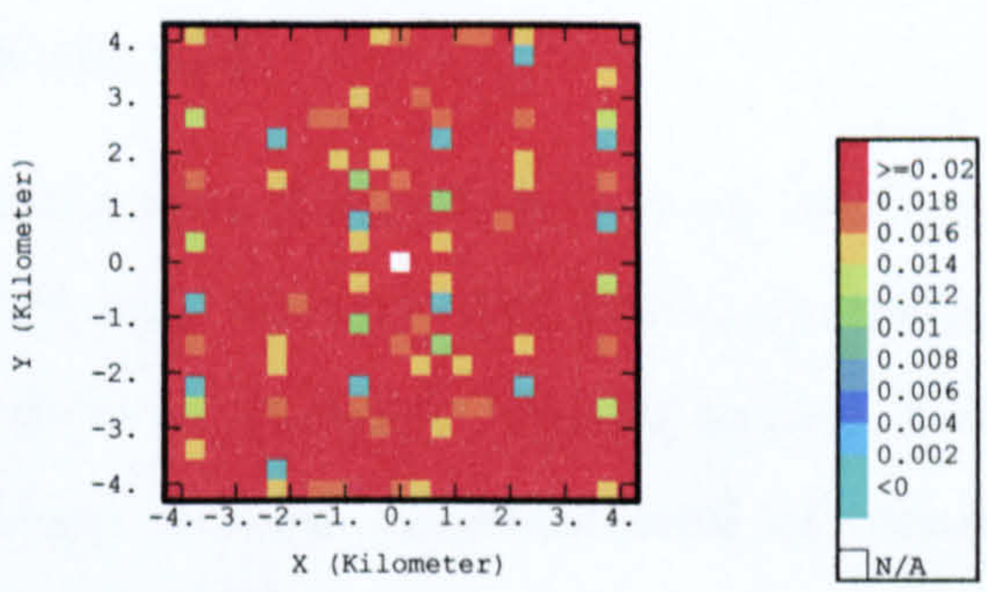
6th percentile (0.5mg.kg^{-1})



96th percentile (2.5mg.kg^{-1})

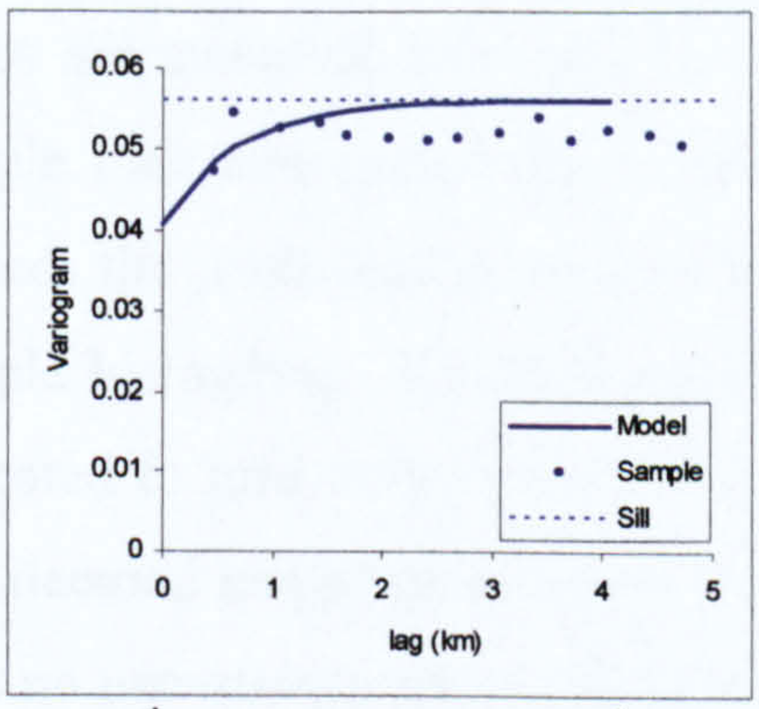


47th percentile (1.5mg.kg^{-1})

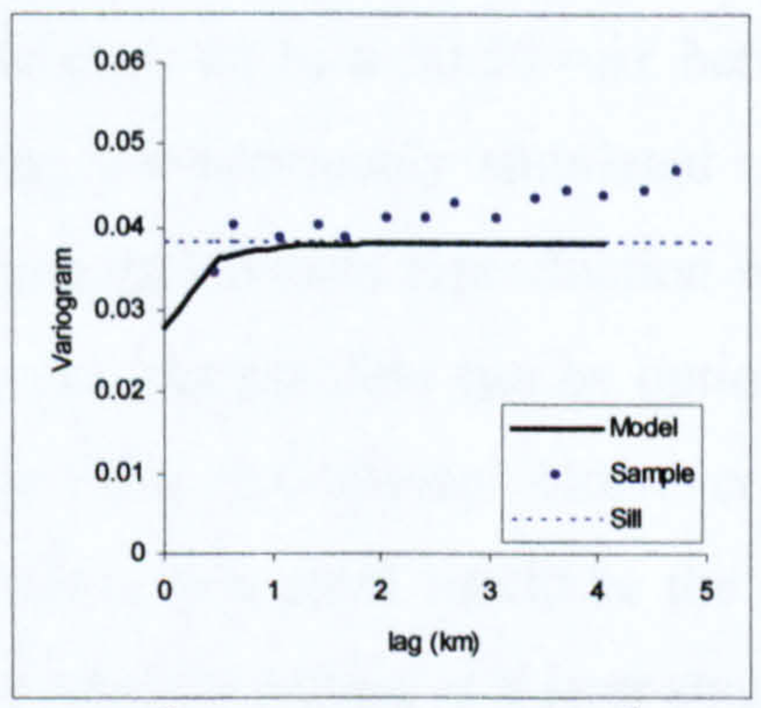


98th percentile (4.5mg.kg^{-1})

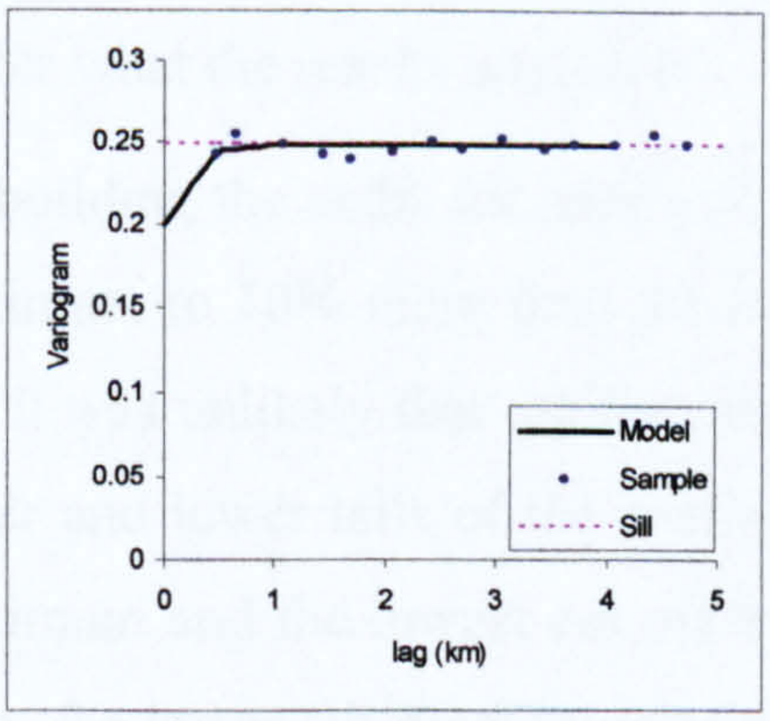
Figure D 2. Indicator variogram maps for Cd at four cut offs.



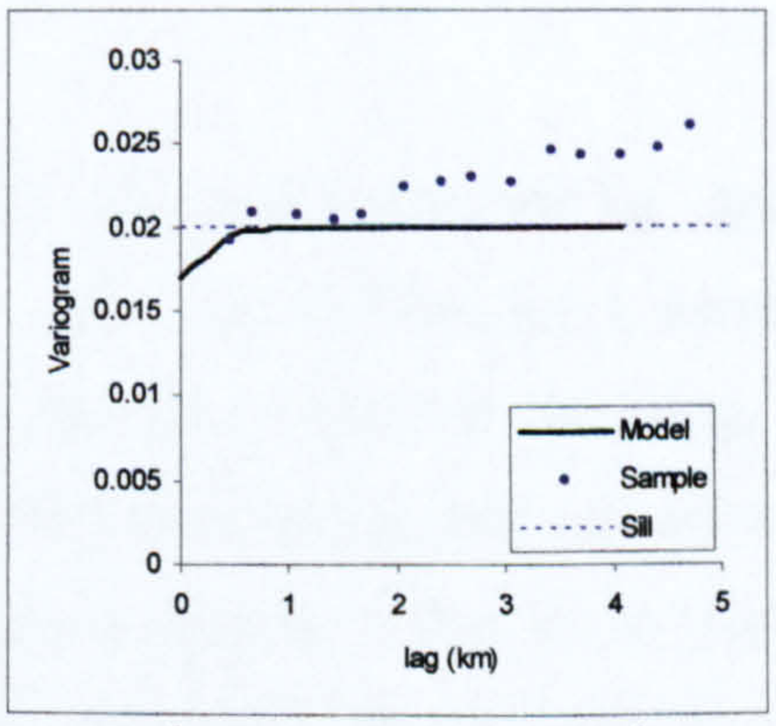
6th percentile (0.5mg.kg^{-1})



96th percentile (2.5mg.kg^{-1})



47th percentile (1.5mg.kg^{-1})



98th percentile (4.5mg.kg^{-1})

Figure D 3. Sample and model indicator variograms for Cd.

Table D 2. Indicator variogram model parameters for Cd

Cut off ($mg.kg^{-1}$)	Exponential[1]		Nugget C_0
	a	c	
0.5	2500	0.011	0.045
1.5	1100	0.03	0.22
2.5	1000	0.01	0.03
4.5	800	0.005	0.015

[1]Exponential model: $\gamma(h) = C_0 + c.\left[1 - \exp\left(-\frac{3h}{a}\right)\right]$

where h is the lag distance

D.3 Running and validating simulations; Cd

Simulations were carried out using GSLIB SISIM routine (Deutsch & Journel, 1998) on a 50mx50m square grid with 272 cells in the East-West direction and 312 North-South. Each simulated value was taken to be representative of the entire 50x50m square. The ordinary kriging routine was used with a moving window neighbourhood of radius 4,000m. The minimum and maximum number of sample data points to be used in simulation of a grid point were set to 0 and 8 respectively, while the number of previously simulated nodes was set to 8. These parameters ensure that before any nodes are simulated, up to 8 sample data points will be used, and thereafter, once simulated nodes are available, they will be taken into account also, up to a 50:50 mix between sample data and previously simulated nodes. Using the previously simulated nodes ensures the joint ccdf is reproduced and using sample data assists reproduction of the sample histogram. To speed up the simulation process, sample data can be optionally allocated to grid nodes prior to simulation, and kept fixed throughout. However, this was deemed inappropriate since every simulated value at that point would be the same with no uncertainty and could lead to 'holes' in the probability maps at a later stage i.e. that grid point will always have a probability of 0 or 1 of being over a threshold, no matter what the results around it.

On building the ccdfs for each cell, the minimum Cd value was set to $0mg.kg^{-1}$ and the maximum to 10% more than the maximum sample value (i.e. to $47mg.kg^{-1}$), assuming that it was unlikely that the maximum value would have been found in sampling. The upper and lower tails of the ccdf must be extrapolated from the highest cut off to the maximum and the lowest cut off to the minimum using models. After much trial and error, the upper tail model was selected as hyperbolic with parameter 2.5 and the lower a power model with parameter 2. The selection was based on comparing the statistics

of the sample and simulated distributions of Cd values. The closer the statistics, the better the model parameters. See Figure D 4 and Table D 3.

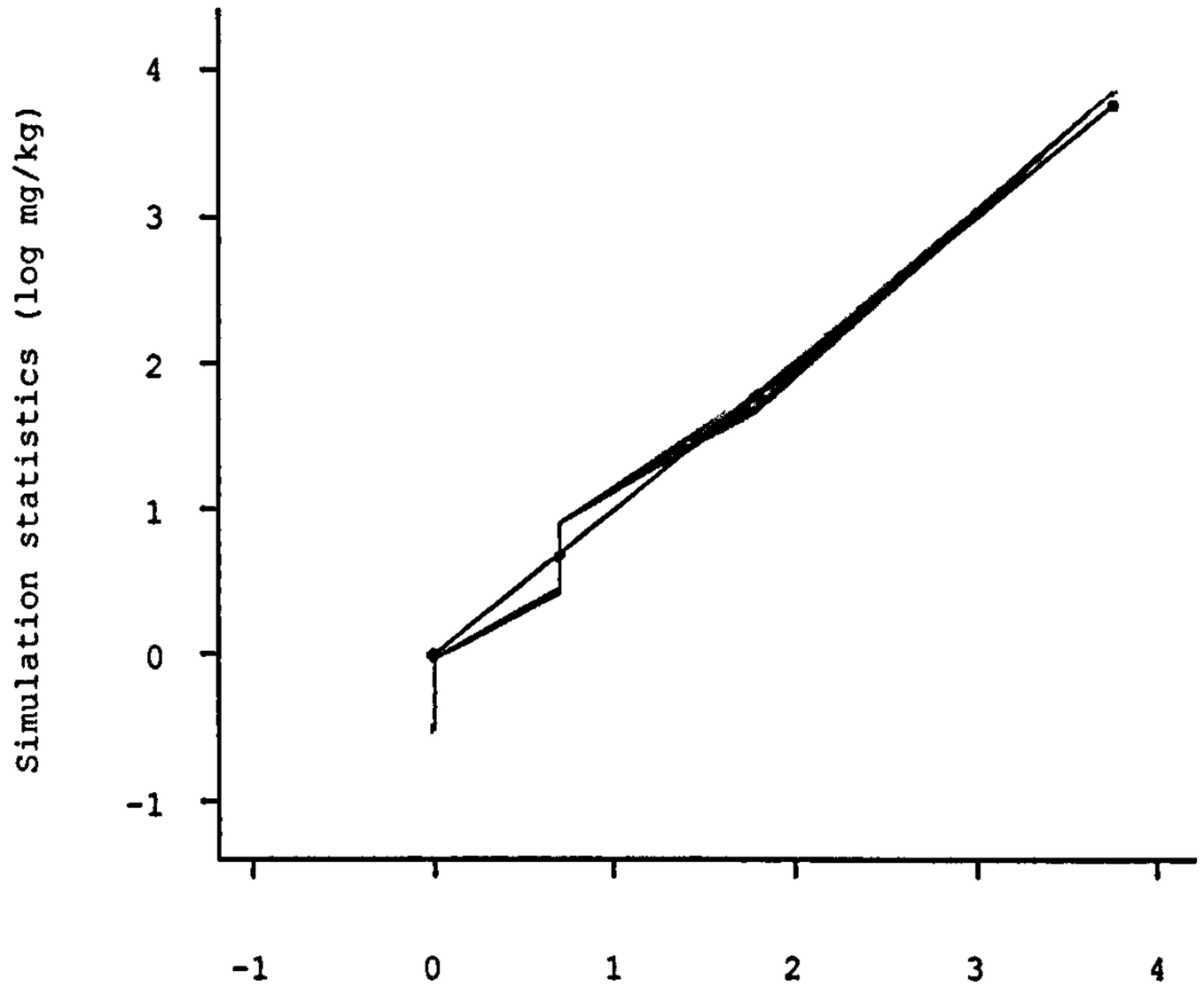


Figure D 4. Sample data statistics plotted against simulation statistics for first 50 simulations. Black diagonal represents perfect match, red the simulated statistics. Statistics plotted are: min, P1, P5, P10, P25, P50, P75, P90, P95, P99, max.

Table D 3. Statistics of Cd topsoil samples (with non-detects changed to zero) and average statistics from first 100 simulations ($mg.kg^{-1}$).

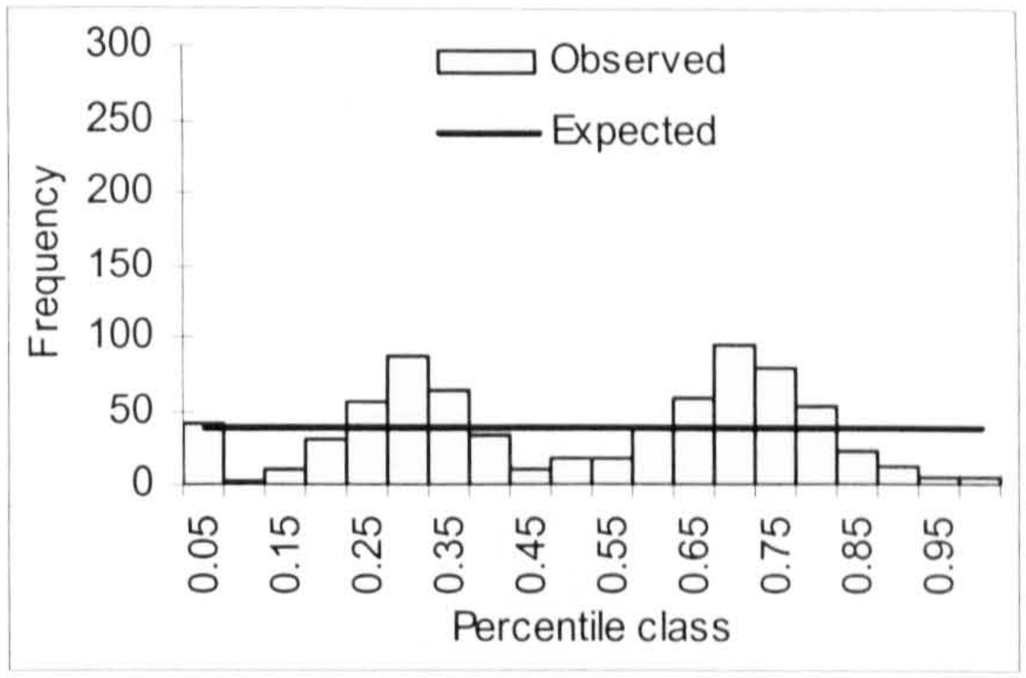
	Mean	Stdev	Min	P1	P5	P10	P25	P50	P75	P90	P95	P99	Max
Samples	1.67	2.03	0.0	0.0	0.0	1.0	1.0	2.0	2.0	2.0	2.0	6.0	43.0
Simulations 50m grid	1.63	1.25	0.0	0.2	0.5	0.6	1.0	1.6	2.1	2.4	2.5	5.9	46.9
Simulations 200m grid	1.63	0.53	0.6	0.9	1.1	1.2	1.3	1.6	1.8	2.1	2.5	3.9	7.5

Indicator procedures at multiple cut offs with different indicator variogram models can result in order relation problems, whereby $(P > z_c) < (P > z_{c+1})$, which is clearly impossible. Order relation problems are corrected within the SISIM programme. Order relation problems for these simulations are low, in the order of 10% of ccdfs requiring corrections, with an average magnitude of correction in the order of only 0.0005.

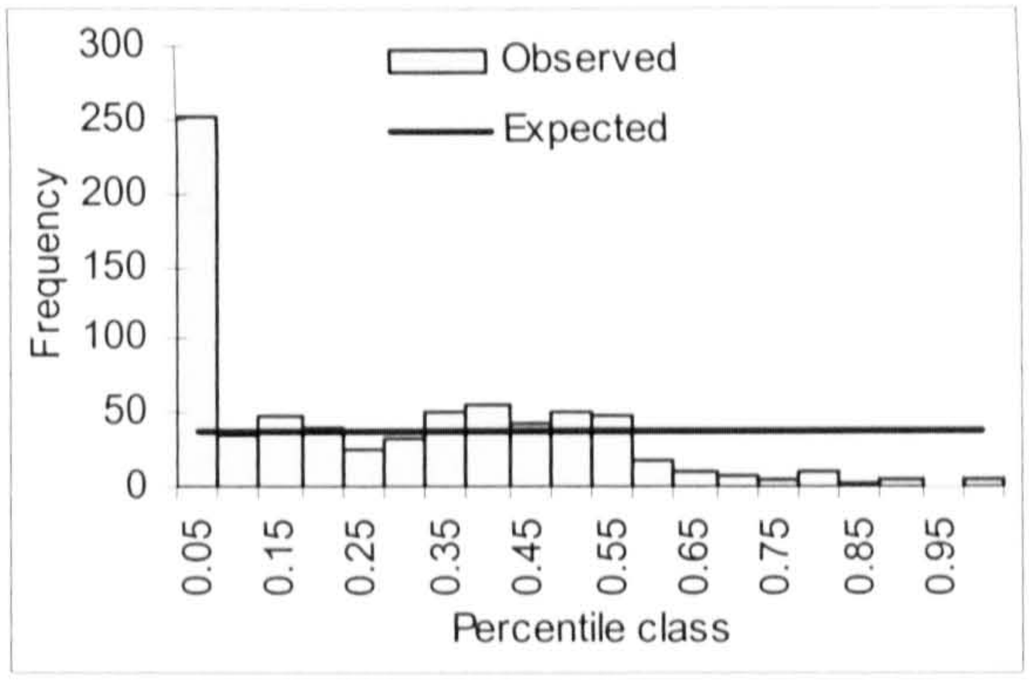
Assessing the goodness of the simulations in this case was less straightforward than in the previous case studies, due to the fact that the Cd samples are measured only to the

nearest $mg.kg^{-1}$ with 41% $1mg.kg^{-1}$ and 49% $2mg.kg^{-1}$ making up the bulk of the values. As usual, first of all, the 744 grid nodes nearest to the sample points were selected as representing the simulated distribution of soil values for that sample point. Then, the percentile at which each sample falls within the distribution of 100 simulated values is plotted in a histogram.

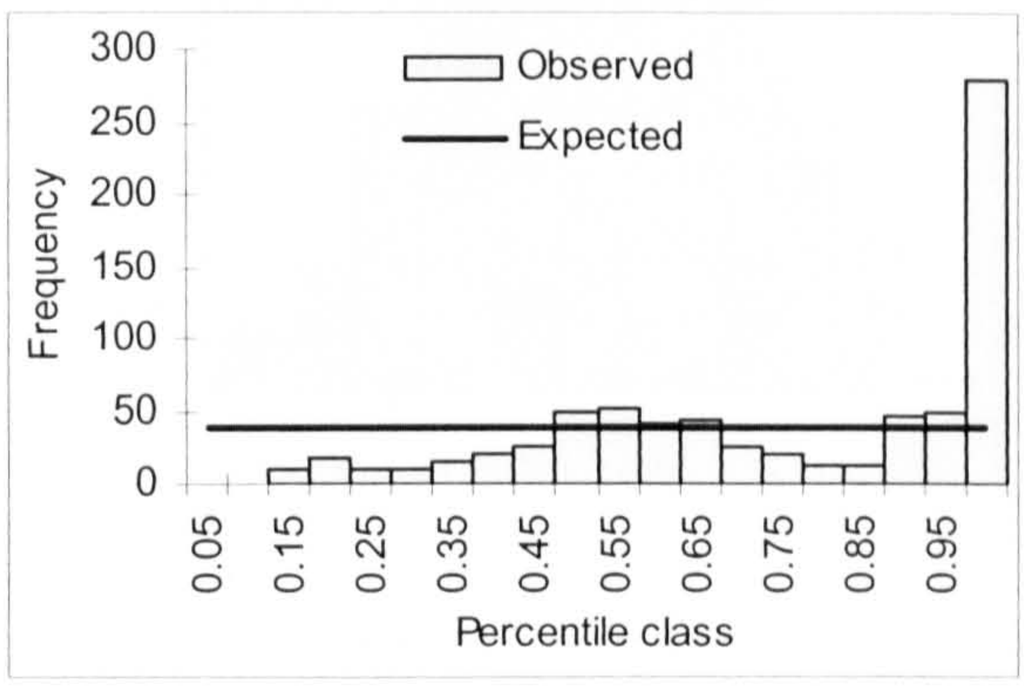
In this case, however, the plot is not straightforward. The problem with the Cd analysis is that a sample value recorded as $1mg.kg^{-1}$ could represent any Cd value between 0.5 and $1.5mg.kg^{-1}$, and a sample value recorded as $2mg.kg^{-1}$ any value between 1.5 and $2.5mg.kg^{-1}$ etc. A simple plot looks unsatisfactory with peaks at about the 26^{th} and 71^{st} percentiles (plot (a) in Figure D 5). Recall the first cut off, $0.5mg.kg^{-1}$, is at the 6^{th} percentile and the second cut off, $1.5mg.kg^{-1}$, is at the 47^{th} percentile. The 26^{th} percentile is midway between these two, so the peak at the 26^{th} percentile is due to repetitive $1mg.kg^{-1}$ values. Likewise, the peak at the 71^{st} percentile lies midway between the second and third cut offs (1.5 and $2.5mg.kg^{-1}$, the 47^{th} and 96^{th} percentiles) and is due to repetitive $2mg.kg^{-1}$ sample values. To get around this problem, the plot is produced twice more. First of all, it is assumed that all samples are recorded at the bottom of their range (e.g. all the $1mg.kg^{-1}$ values are set to $0.5mg.kg^{-1}$) and plot (b) in Figure D 5 is produced. Secondly all the sample values are set to the top of their range (e.g. all the $1mg.kg^{-1}$ values are set to $1.5mg.kg^{-1}$) and plot (c) in Figure D 5 is produced. Next the percentile is recorded for each '$-5mg.kg^{-1}$' result and the corresponding '$+0.5mg.kg^{-1}$' result and the number of percentile classes spanned is counted. The probability of a result is spread between the spanned classes. Finally the probabilities are summed and plot (d) in Figure D 5 is produced. This looks much more as expected, if the simulations are good. One would expect a uniform distribution of percentiles, as indicated by the 'expected' line in the Figure. In this case, there is a relatively good spread and a chi^2 test shows that the results do not deviate significantly from the expected. As a result of this, the simulations are taken to be acceptable.



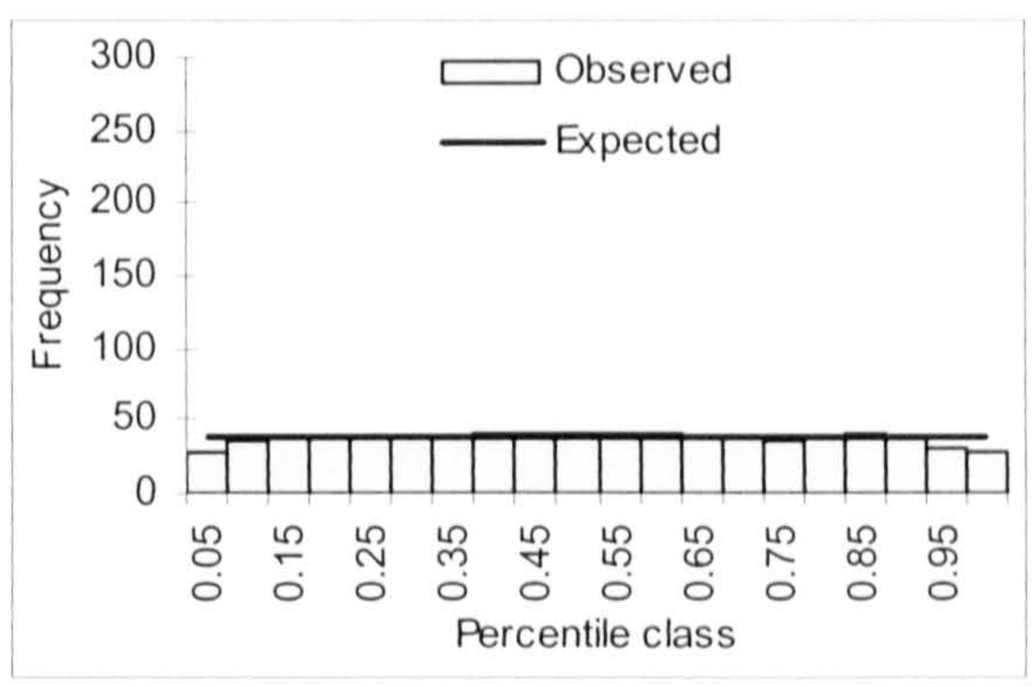
(a)



(b)



(c)



(d)

Figure D 5. Distribution of where in 100 simulated Cd values the true Cd value falls. (a) recorded sample value (b) recorded sample value minus 0.5mg.kg^{-1} (c) recorded sample value plus 0.5mg.kg^{-1} (d) weighted average of the three previous plots.

1,000 simulations were produced in all based on the fact that cumulative map statistics (mean, median and variance) had stabilised and at individual grid points the 'space of uncertainty' had reached a maximum width i.e. the interval between the minimum and maximum simulated value at that point increased no more.

To demonstrate the change of support effect, the 50x50m simulations were amalgamated into 200x200m blocks for use in the exposure analysis by taking the arithmetic mean of the 16 values within the larger block. The statistics of the upscaled 200x200m results are also given in Table D 3.

D.4 Simple statistics; pH

Simple statistics of the 742 topsoil samples are given in Table D 4 and a histogram of the distribution in Figure D 6. The data look normally distributed so it would be possible to simulate using a Gaussian based method. However, it was decided to use SIS again, since any continuity in high/low values would be preserved

Table D 4. Simple statistics of the topsoil sample pH levels.

Mean	Stdev	Min	P1	P5	P10	P25	P50	P75	P90	P95	P99	Max
5.46	1.05	2.9	3.2	3.7	4.1	4.7	5.4	6.3	6.9	7.2	7.57	8.01

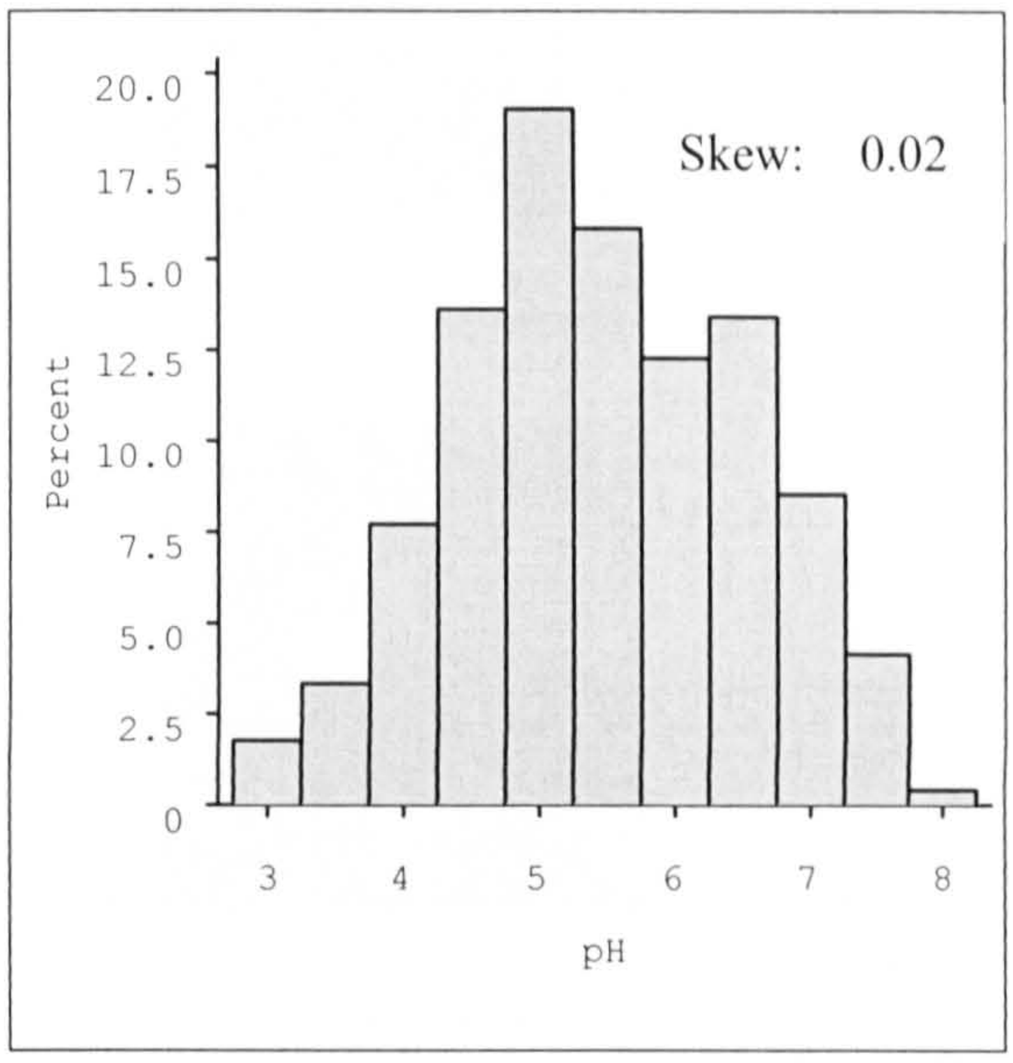
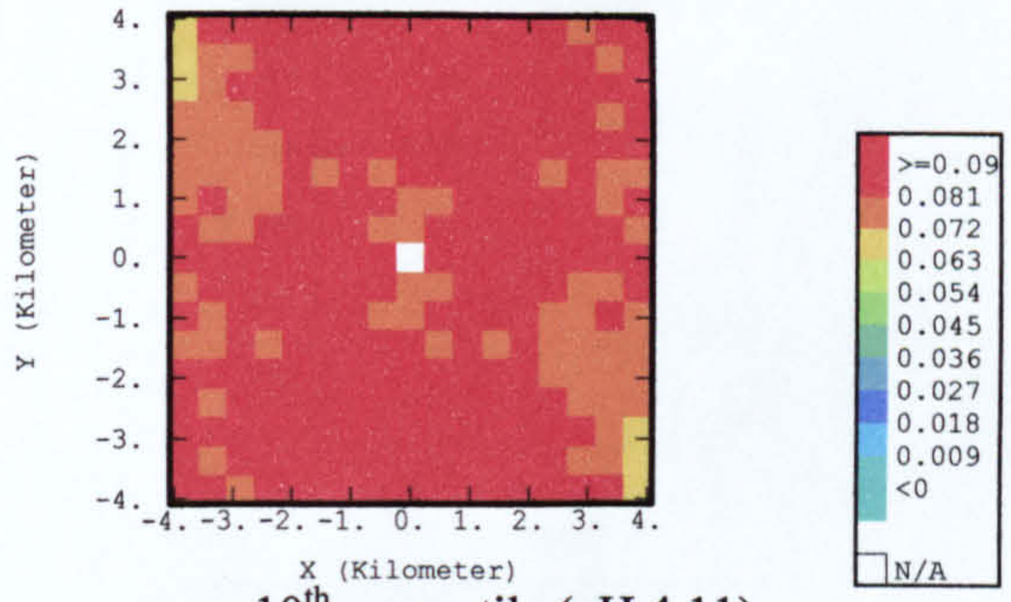


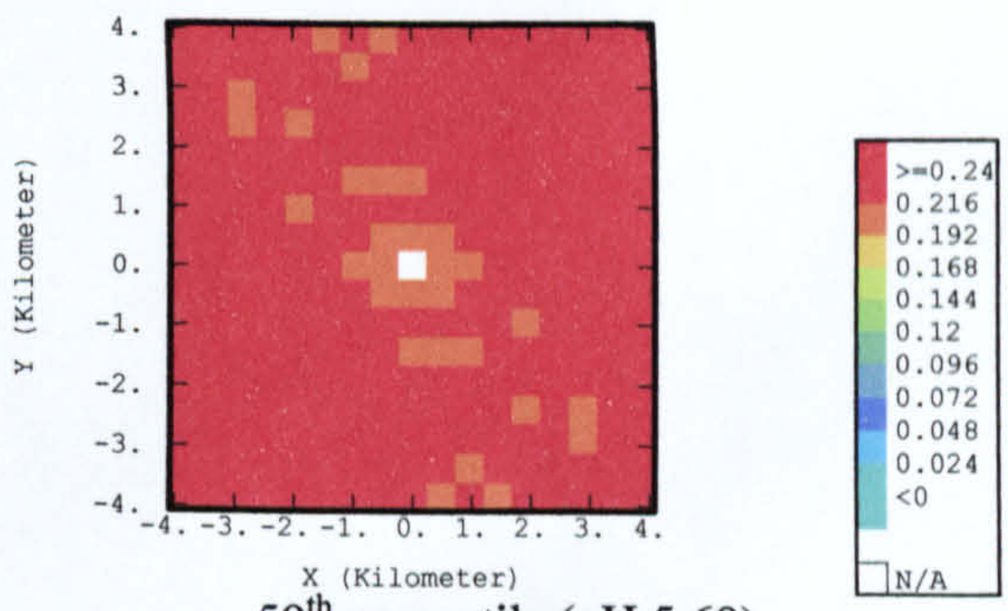
Figure D 6. Histogram of topsoil sample pH levels.

D.5 Variogram modelling; pH

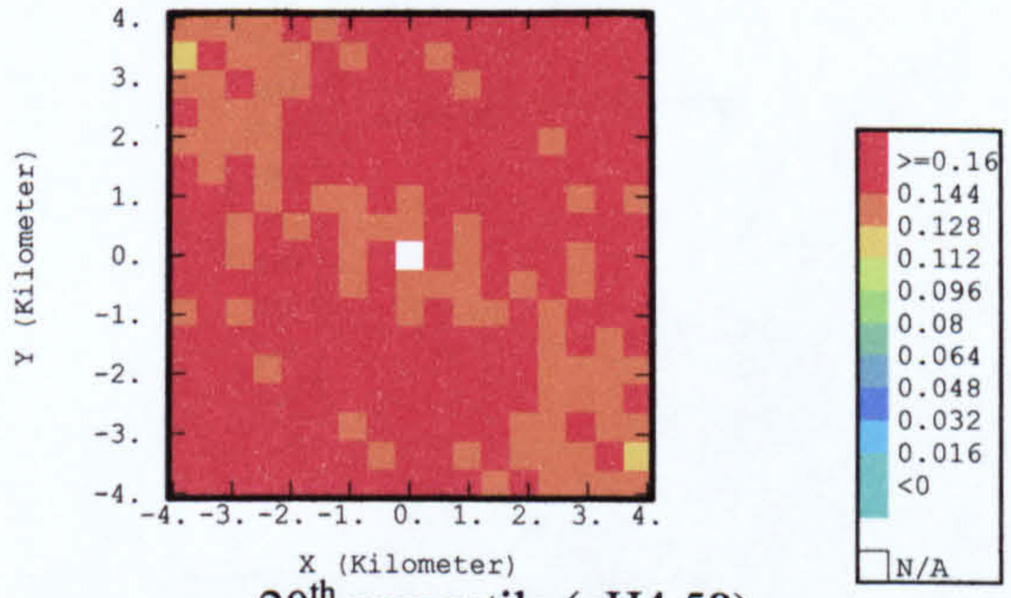
Indicators were modelled at approximately the 9 deciles of the pH sample distribution. Actual percentiles modelled and the pH values at those percentiles are: 10th, 20th, 33rd, 41st, 49th, 59th, 69th, 80th and 91st percentiles and pH 4.11, 4.58, 4.91, 5.11, 5.39, 5.69, 6.08, 6.46 and 6.92 respectively. The variogram maps are shown in Figure D 7. Little structure was found, but the most was seen using a lag of 480m and a lag tolerance of 50%. Because there was so little structure, the variograms were difficult to model and so were made omni-directional, despite there being a suggestion of anisotropy at some levels. Indicator sample and model variograms are shown in Figure D 8 and the model parameters given in Table D 5.



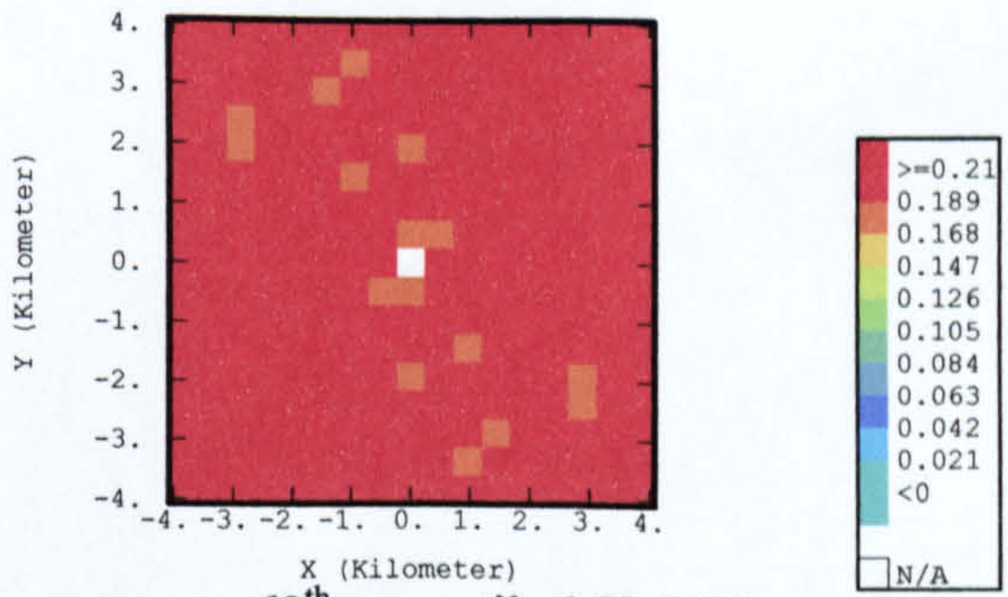
10th percentile (pH 4.11)



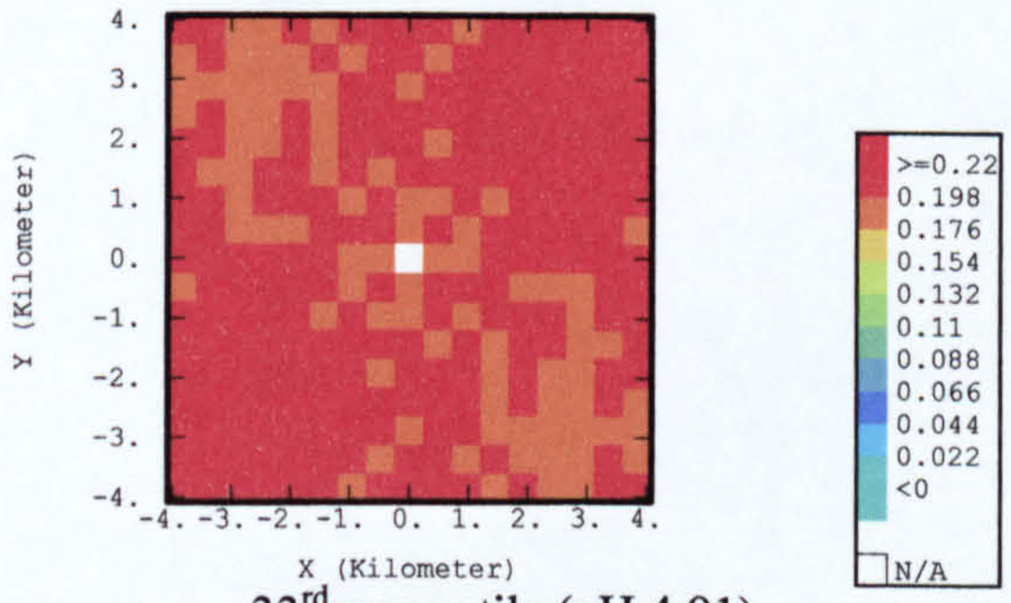
59th percentile (pH 5.69)



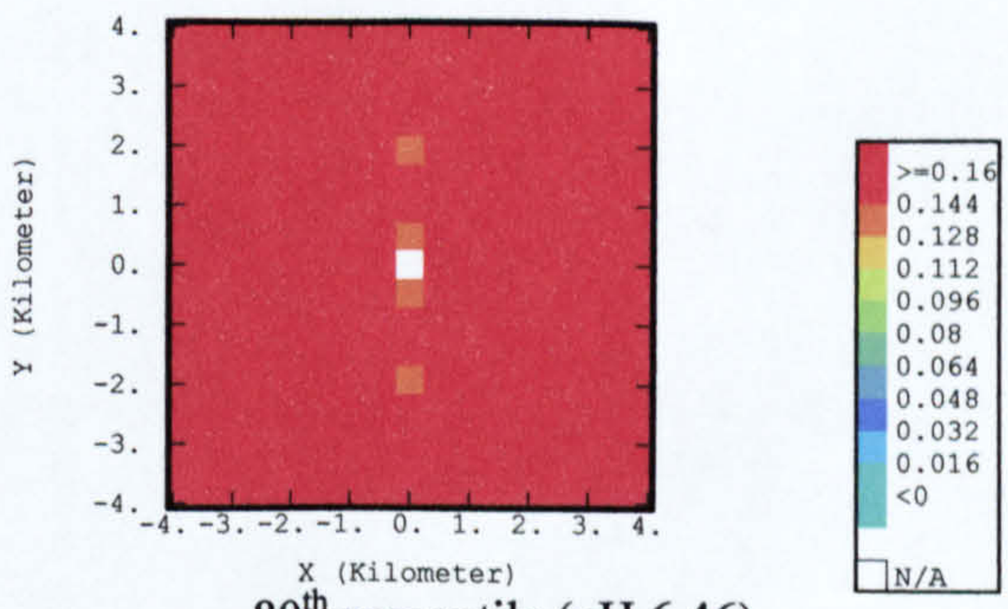
20th percentile (pH4.58)



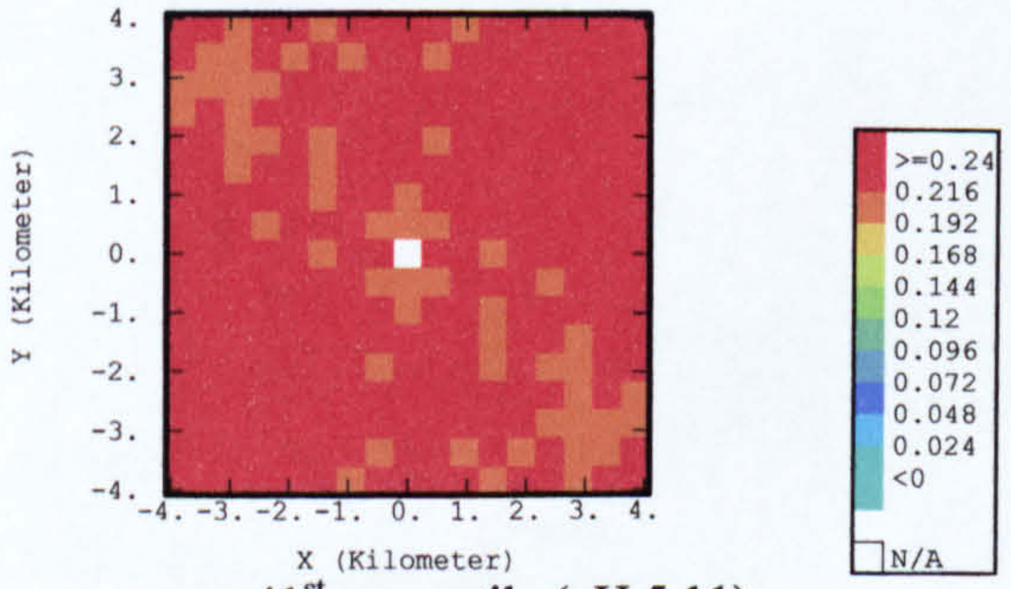
69th percentile (pH 6.08)



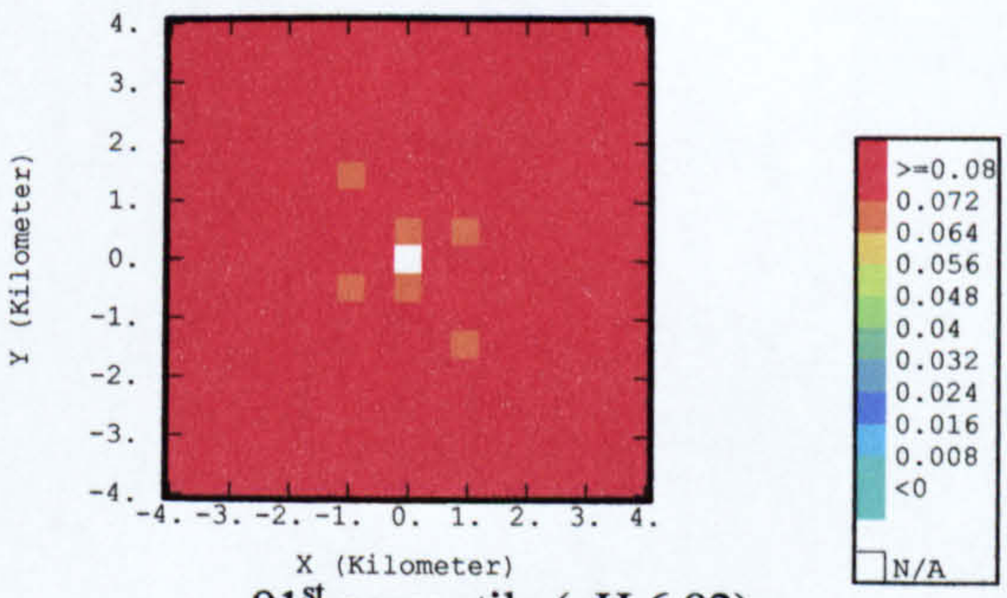
33rd percentile (pH 4.91)



80th percentile (pH 6.46)



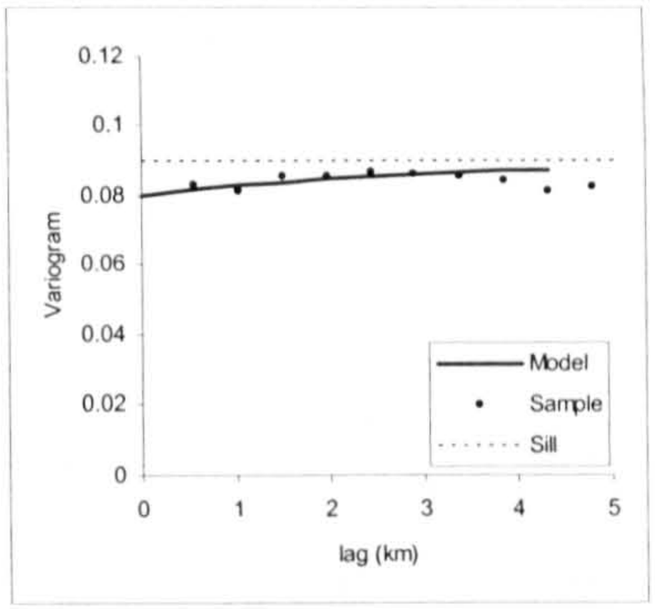
41st percentile (pH 5.11)



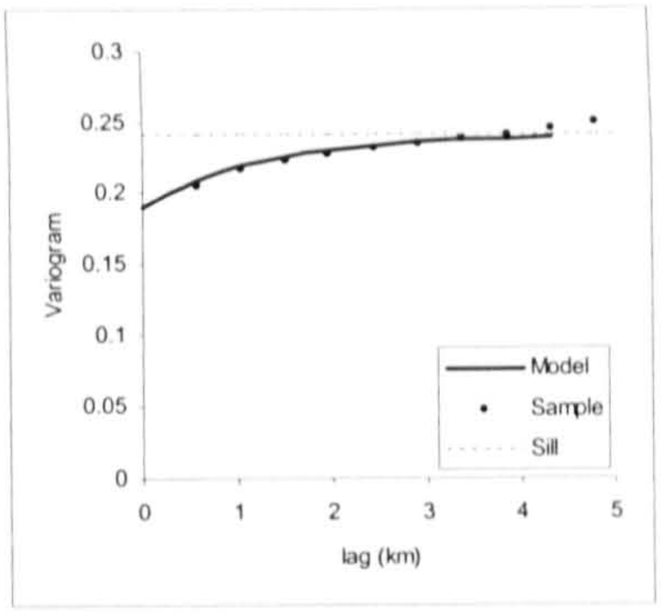
91st percentile (pH 6.92)

49th percentile (pH 5.39)

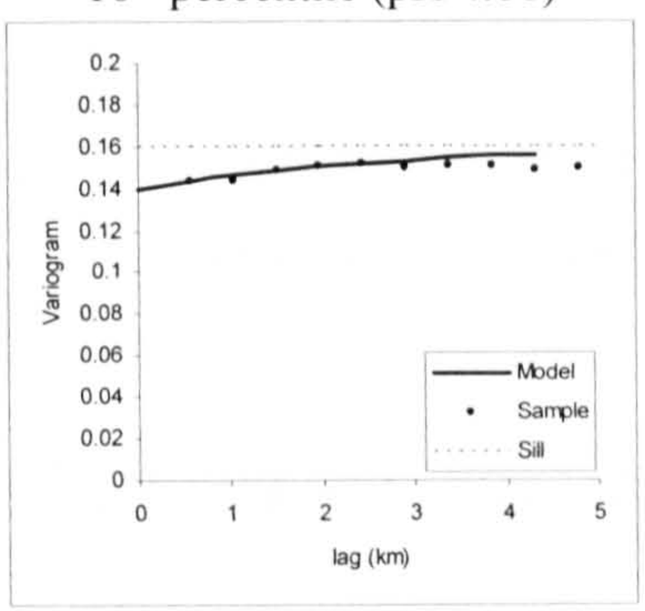
Figure D 7. Indicator variogram maps for pH at nine cut offs.



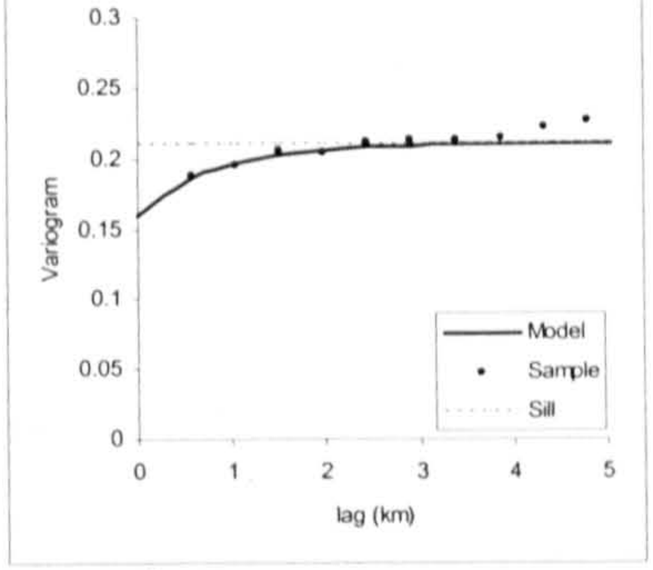
10th percentile (pH 4.11)



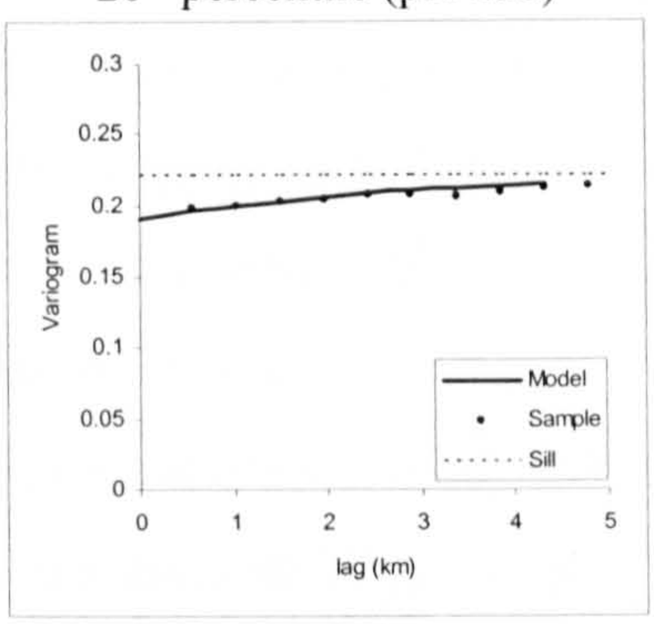
59th percentile (pH 5.69)



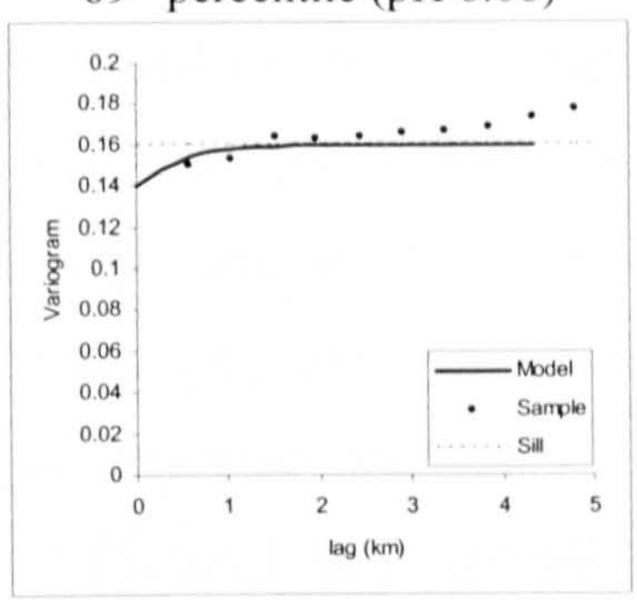
20th percentile (pH4.58)



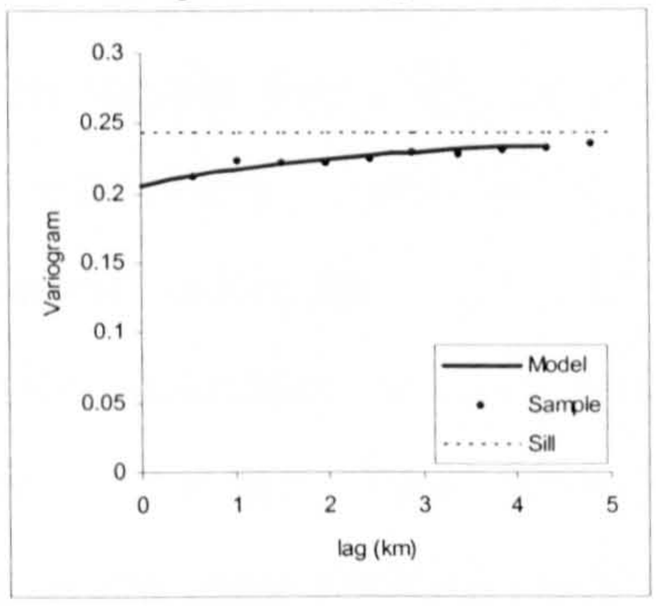
69th percentile (pH 6.08)



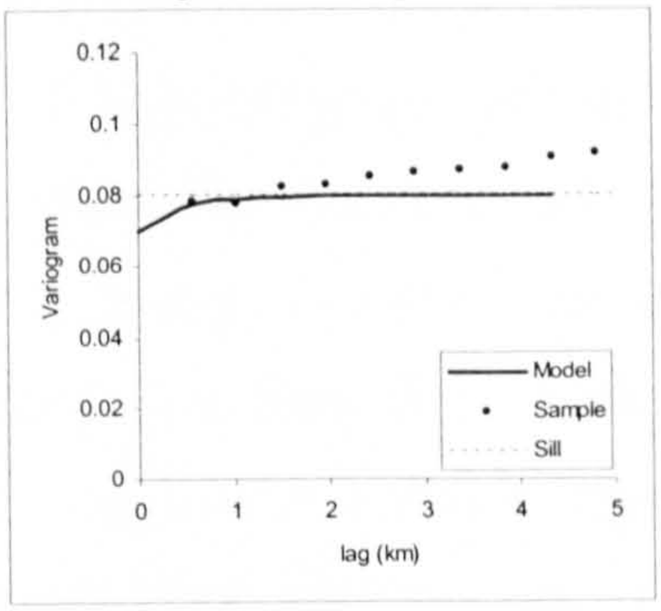
33rd percentile (pH 4.91)



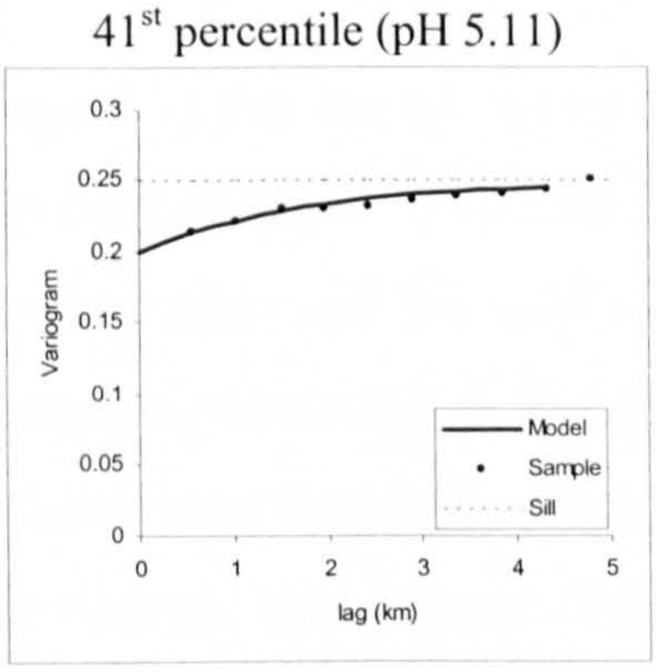
80th percentile (pH 6.46)



41st percentile (pH 5.11)



91st percentile (pH 6.92)



49th percentile (pH 5.39)

Figure D 8. Sample and model indicator variograms for pH.

Table D 5. Indicator model parameters for pH.

Cut off	Exponential		Nugget
	a	c	C_0
4.11	9000	0.01	0.08
4.58	8000	0.02	0.14
4.91	8000	0.03	0.19
5.11	7000	0.035	0.205
5.39	5500	0.05	0.20
5.69	3500	0.05	0.19
6.08	2300	0.05	0.16
6.46	1500	0.02	0.14
6.92	1200	0.01	0.07

Exponential[2]: $\gamma(h) = C_0 + c.\left[1 - \exp\left(-\frac{3h}{a}\right)\right]$

where h is the lag distance

D.6 Running and validating simulations; pH

Simulations were carried out on the same 50mx50m square grid as the Cd simulations. Each simulated value was taken to be representative of the entire 50x50m square. The ordinary kriging routine was used with a moving window neighbourhood of radius 4,000m. The minimum and maximum number of sample data points to be used in simulation of a grid point were set to 0 and 16 respectively, while the number of previously simulated nodes was set to 16. These parameters ensure that before any nodes are simulated, up to 16 sample data points will be used, and thereafter, once simulated nodes are available, they will be taken into account also, up to a 50:50 mix between sample data and previously simulated nodes. Using the previously simulated nodes ensures the joint ccdf is reproduced and using sample data assists reproduction of the sample histogram. To speed up the simulation process, sample data can be optionally allocated to grid nodes prior to simulation, and kept fixed throughout. However, this was deemed inappropriate since every simulated value at that point would be the same with no uncertainty, so data was not allocated to grid nodes.

pH is equal to the log of the H+ ion concentration, and on building the ccdfs for each cell, the minimum pH value was set to represent 10% less than the minimum sample H+ ion concentration and the maximum to 10% more than the maximum sample H+ ion concentration, assuming that it was unlikely that the minimum and maximum values would have been found in sampling. The figure of 10% is arbitrary, but may well lead to more realistic simulations. The result was a lower allowable pH of 2.79 and an upper allowable pH of 8.11. The upper and lower tails of the ccdf must be extrapolated from the 9th decile to the maximum and the 1st decile to the minimum using models. After

much trial and error, the upper tail model was selected as a power model with parameter 0.3 and the lower a power model with parameter 2. The selection was based on comparing the statistics of the sample and simulated distributions of pH values. The closer the statistics, the better the model parameters. See Figure D 9 and Table D 6.

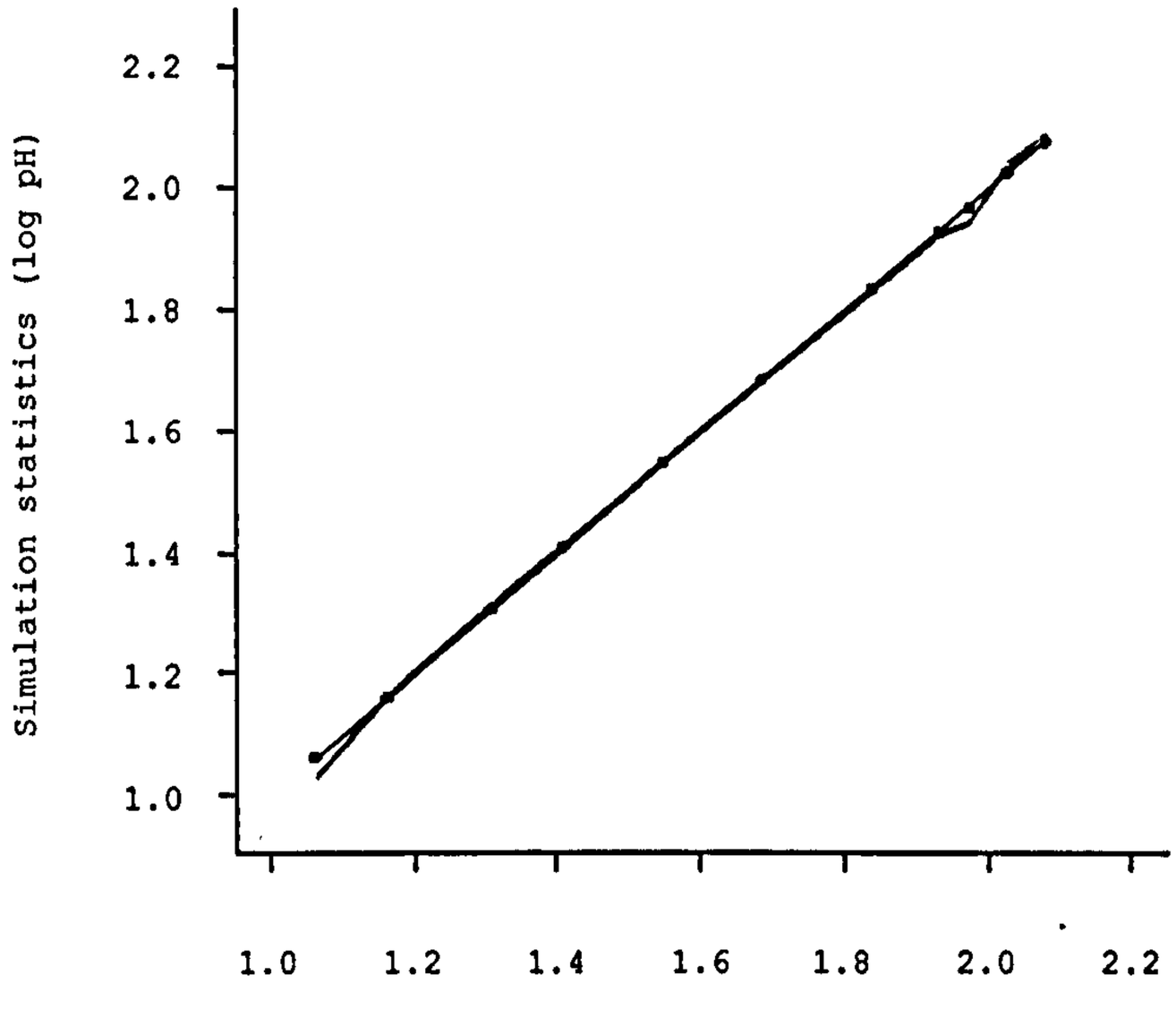


Figure D 9. Sample data statistics plotted against simulation statistics for first 50 simulations. Black diagonal represents perfect match, red the simulated statistics. Statistics plotted are: min, P1, P5, P10, P25, P50, P75, P90, P95, P99, max.

Table D 6. Statistics of pH topsoil samples and average statistics from first 100 simulations.

	Mean	Stdev	Min	P1	P5	P10	P25	P50	P75	P90	P95	P99	Max
Samples	5.46	1.06	2.9	3.2	3.7	4.1	4.7	5.4	6.3	6.9	7.2	7.6	8.0
Simulations 50m grid	5.44	1.04	2.8	3.2	3.7	4.1	4.7	5.4	6.3	6.9	7.0	7.7	8.1
Simulations 200m grid	5.44	0.45	4.0	4.5	4.7	4.9	5.1	5.4	5.8	6.0	6.2	6.5	6.9

Indicator procedures at multiple cut offs with different indicator variogram models can result in order relation problems, whereby $(P > z_c) < (P > z_{c+1})$, which is clearly impossible. Order relation problems are corrected within the SISIM programme. Order relation problems for these simulations are low, in the order of 3% of ccdfs requiring corrections, with an average magnitude of correction in the order of 0.007.

In order to assess the goodness of the simulations the method devised by Deutsch was followed (Deutsch, 1997). This method produces Accuracy, Precision and Goodness statistics between 0 (worst case) and 1 (best case) by comparing the sample values with the simulated values at that point. An accurate probability distribution is one where the fraction of true values falling in the p-interval exceeds p for all p in [0,1] and a precise (and accurate) distribution is measured by the closeness of the fraction of true values to p for all p in [0,1]. Goodness is a measure of departure from the 45degree line on the accuracy plot and weights inaccurate results twice, as they are more consequential than accurate ones. The results gave A=0.95, P=0.94 and G=0.97; an accuracy plot is shown in Figure D 10. In addition, a histogram is plotted of where in the distribution of simulations, the true value lies. First of all, the 742 grid nodes nearest to the sample points were selected as representing the simulated distribution of soil values for that sample point. The plot is shown in Figure D 11. A chi^2 test was not significant, so it can be concluded that the results do not differ significantly from the expected. On the basis of these tests, the pH simulations were accepted as being satisfactory.

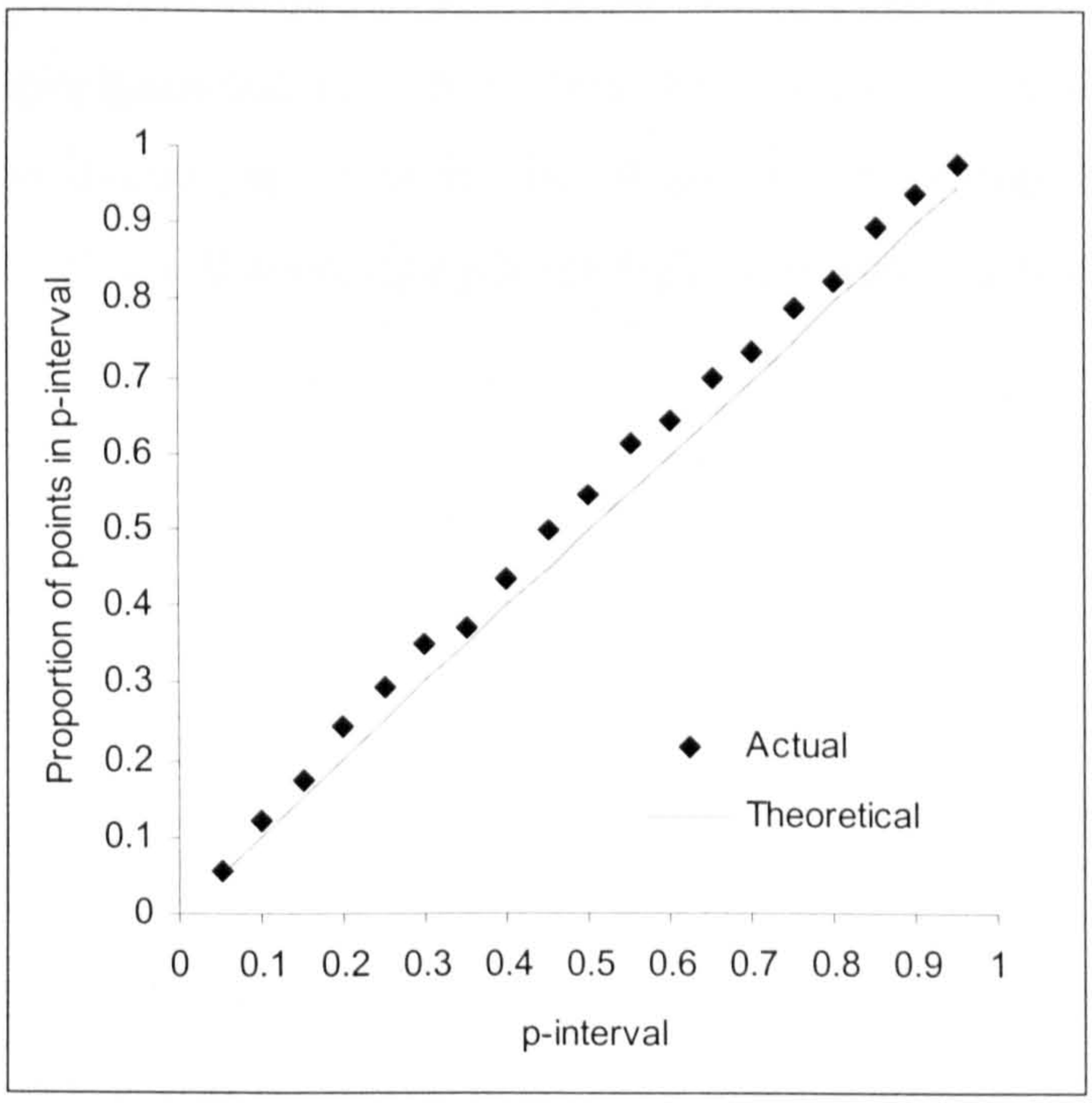


Figure D 10. Accuracy plot for pH simulations.

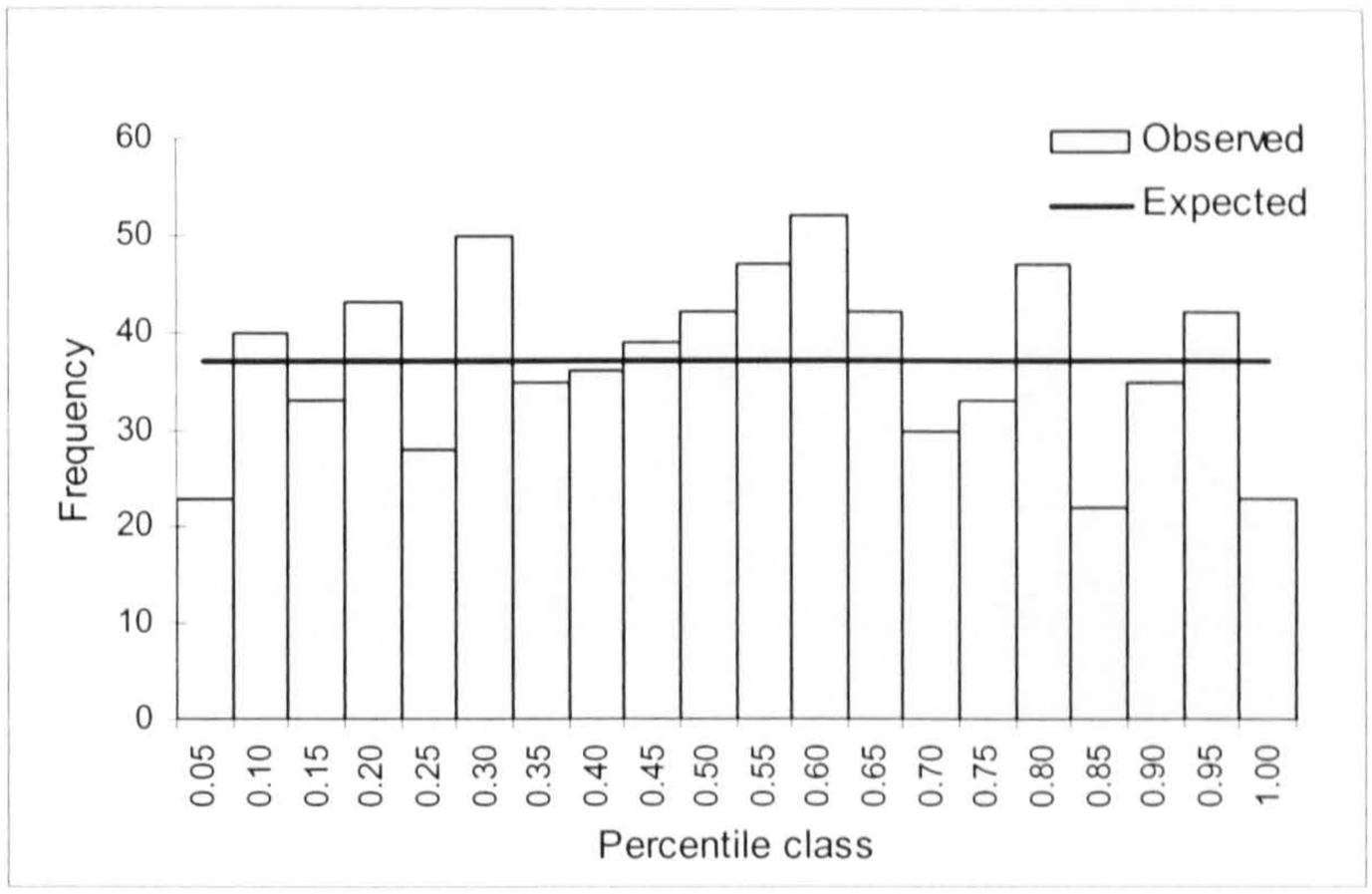


Figure D 11. Distribution of where in 100 simulated pH values the true pH value falls.

1,000 simulations were produced in all based on the fact that cumulative map statistics (mean, median and variance) had stabilised and at individual grid points the 'space of uncertainty' had reached a maximum width i.e. the interval between the minimum and maximum simulated value at that point increased no more.

To demonstrate the change of support effect, the 50x50m simulations were amalgamated into 200x200m blocks for use in the exposure analysis, by taking the arithmetic mean of the 16 smaller blocks within each large block. For the non-spatial analysis, the average pH for each map was also found.

BIBL.
LONDIN.
UNIV.