THE USE OF SOIL INFORMATION SYSTEMS IN LAND PLANNING:

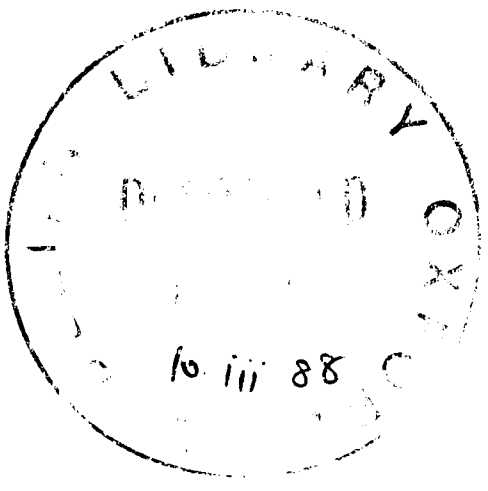
Procedures for generating information for properties, depths and sites not recorded.

A thesis submitted to the University of Oxford

for the degree of Doctor of Philosophy

by

R. Ponce-Hernandez



Trinity Term

1987

Balliol College

Oxford

THE USE OF SOIL INFORMATION SYSTEMS IN LAND PLANNING:
Procedures for generating information for properties, depths and sites not recorded

by

R. PONCE HERNANDEZ

A thesis submitted to the University of Oxford for the degree of Doctor of Philosophy

Trinity Term, 1987

BALLIOL COLLEGE

ABSTRACT

Computerized soil information systems (SIS) store and retrieve, much more efficiently than soil maps, the soil information collected from field survey which is essential for land-use planning. The soil is only observed and sampled at a limited number of locations, depths and for a limited number of properties. Information not recorded during survey is missing and if needed must be generated or predicted. New geostatistical techniques for spatial analysis and interpolation of soil data, i.e. the semi-variogram and Kriging, can now be included as on-line capabilities of SIS to equip users with a powerful tool for prediction of the missing information.

When there are no records of the wanted property, its values can be estimated by a function on other recorded properties acting as its surrogates. A strategy for model development is provided for the construction of surrogate functions based on multiple regression and curve fitting techniques, to generate the information missing.

When no records of the soil property at the depth wanted exist, the values at the required depths are interpolated by a function of the property on the soil depth. Equal-area spline curves reconstruct, piecewise, the depth function quite closely, enabling their use for interpolation of values and depths in a variety of formats. The equal-area spline algorithm is a capability of the Oxford SIS (OXSIS).

When the wanted site was unvisited during survey, the information missing is provided by spatial prediction. The predictions may come from means of soil classes or mapping units from conventional survey, or from Kriging interpolation based on spatial analysis by the semi-variogram. In order to select the best predictive tool, the success of these techniques in different situations of sampling effort and variablity were compared. Semi-variograms depicted the spatial structures of 8 selected soil properties. Anisotropic variation in 4 of them was induced by strong trends. Where the semi-variogram was isotropic Kriging was the best tool for prediction if spatial dependence is strong. Fitting elliptical functions to find a model for anisotropy did not give satisfactory results. Where anisotropies or trends precluded ordinary Kriging, map unit means and class means, in that order, gave the best predictions. Success in prediction was related to the structures in the semi-variogram, which when used for reconnaissance helps to infer which technique will give the best predictions so that survey is designed accordingly. Accounting for the trends removed anisotropies and Kriging of de-trended data was possible. Partitioning trends by stratification based in soil mapping units gave a greater improvement in predictions than modelling trends by bicubic spline surfaces and then Kriging the residuals from trend. After trend removal, Kriging did not always make the best predictions and means from classes seemed equally as good as Kriging or even better in some cases. These results indicate that a critical point is to ascertain how to best sample to estimate a reconnaissance semi-variogram for survey design to provide the information missing necessary for land-use planning.

Acknowledgements

I am grateful to the Consejo Nacional de Ciencia y Tecnologia (CONACYT) of Mexico and to Colegio de Postgraduados, Chapingo, Mexico, for their scholarship and the invaluable support granted for my studies at Oxford.

My most sincere gratitude goes to Dr. P.H.T. Beckett, my supervisor, for his correct and wise guidance, patience, constant encouragement and support without which the successful completion of this work would not have been possible.

I am grateful also to Dr R. Webster for kindly allowing me access to his computer software on semi-variograms and kriging interpolation.

I am deeply indebted to Michelle for her constant support, patience and unfailing encouragement and for having, welcomingly, introduced me to so many aspects of British life.

A thankful acknowledgement is extended to Dr. C.I. Goodwin-Bailey and Mrs. J Smith for their assistance in computational matters and for all the hardware and software facilities made available.

Mrs. Gillian Bendle, Ann Steadman, Ingrid Eades and Gladys Roper gave always invaluable assistance and support. To them and to so many others in the Plant Sciences Department, whose names for being many would be imposible to list, I am specially grateful.

To my family, in Mexico, I am grateful for their endless patience and constant encouragement.

****

To the loving memory of my mother.

To my father, with infinite gratitude.

CONTENTS

I.		INTRODUCTION	1
II.		SOIL INFORMATION AND LAND-USE PLANNING	7
	2.1	Agricultural planning	7
	2.2	Land-use planning	7
	2.3	Land evaluation and soil data	8
	2.4	Land evaluation and land productivity	10
	2.5	Partitioning variability	10
	2.6	The need for soil data in land-use planning	12
	2.7	Concluding remarks to the chapter	12
III.		SOIL VARIABILITY	13
	3.1	Sources of soil variability	14
		3.1.1 The natural landscape	15
		(a) Parent materials	15
		(b) Climate	16
		(c) Topography	16
		(d) Physical and chemical processes	16
		(e) Biological activities	17
		3.1.2 The cultivated landscape	17
	3.2	Variability in relation to the size of the area sampled	19
	3.3	Directional components of soil variability	29
		3.3.1 Vertical soil variability	29
		3.3.2 Spatial variability	31
	3.4	Methods for investigating the spatial variability of soils	34
		3.4.1 Soil maps	34
		3.4.2 Classical statistics	35
		3.4.3 Geostatistical techniques	36
	3.5	Concluding remarks to the chapter	37

IV. COLLECTION AND PROVISION OF SOIL INFORMATION:

Conventional Soil Survey 39

4.1 Aim and purpose of soil surveys 39

4.2 Kinds of soil survey 40

4.2.1 Very high intensity soil surveys 44

4.2.2 High intensity soil surveys 44

4.2.3 Medium intensity soil surveys 44

4.2.4 Exploratory and low intensity soil survey 44

4.3 Survey intensities 46

4.4 Map scale 47

4.5 Field procedures 48

4.6 Field records and sampling 51

4.7 Soil classification 54

4.8 Soil mapping 59

4.9 Use of soil classes, map units and memoir in prediction 61

4.9.1 Use of class means of the soil property as predictor 64

4.9.2 Use of the typical profile of the class as predictor 65

4.9.3 Use of the mapping unit mean as predictor 65

4.10 Concluding remarks to this chapter 66

V. SOIL INFORMATION SYSTEMS (SIS) 68

5.1 Information technology and SIS 68

5.2 Basic structure of SIS 71

5.2.1 Data Acquisition 72

5.2.2 Input 72

5.2.3 Query 74

5.2.4 Retrieval 74

5.2.5 Manipulation 76

5.2.6 Presentation 77

5.3 Review of Existing SIS 79

5.3.1 Pre-SIS volume of data 96

5.3.2 The DBMS (Data-Base) problem 96

5.3.3 Indexing and data modelling 96

5.3.4 The software compatibility problem 96

5.3.5 The problem of expertise 96

5.3.6 Free user/SIS communication language 97
5.3.7 The problem of data acquisition 97
5.3.8 The problem of up-grading SIS 97
5.3.9 The problem of time in SIS 97
5.3.10 The problem of SIS technology transfer and cost 98
5.3.11 The problem of data compatibility 98
5.3.12 Dimensionality and conceptual framework in SIS 98
5.4 Roles of SIS 99
5.4.1 Automation of the old tasks 99
5.4.2 The new capabilities in SIS 100
5.5 The Oxford Flexible Soil Information System (OXSIS) 101
5.5.1 The nature of the data in OXSIS 102
5.5.2 OXSIS design 103
5.5.2.1 Software 104
5.5.2.2 Hardware 104
5.5.3 Input Procedure 108
5.5.4 Query Procedure 112
5.5.5 Retrieval Procedure 117
5.5.6 Manipulation 118
5.6 Concluding remarks to this chapter 119

VI. PREDICTING MISSING INFORMATION: Special problems on interpolation 121
6.1 Missing soil information 121
6.2 Problems in predicting soil properties by means of a surrogate functions 124
6.3 Problems on soil depth interpolation 125
6.4 Problems in spatial interpolation of soil properties 129
6.4.1 Interpolation by boundaries 130
6.4.2 Interpolation by Classical Statistics 131
6.4.3 Interpolation by smooth surfaces 133
6.4.4 Optimal interpolation based on Spatial Autocovariance 138
6.5 Concluding remarks to this chapter 141

VII. MATHEMATICAL INTERPOLATION OF SOIL DATA 143

7.1 Surrogate Functions 143

7.1.1 Planning or pre-development stage 144

7.1.2 Model development stage 146

7.1.2.1 Surrogate Function when the model is known 146

7.1.2.2 Surrogate Function when the model is unknown 147

7.1.3 Model validation 149

7.2 Soil Depth Functions 152

7.2.1 Cubic Splines 154

7.2.2 Equal-Area Splines 157

7.3 Spatial Interpolation of Soil data 160

7.3.1 Thiessen Polygons 160

7.3.2 Interpolation by Smooth Surfaces 161

7.3.2.1 Global fitting methods 161

7.3.2.2 Local Fitting methods 162

7.3.3 Optimal Interpolation based on Spatial Autocovariance 169

7.3.3.1 Regionalized Variable Theory 169

7.3.3.2 Measuring Spatial Dependence: The Semi-variogram 173

7.3.3.3 Experimental Semi-variograms 174

7.3.3.4 Semi-variogram representation 176

7.3.3.5 Semi-variogram parameters 177

7.3.3.6 Semi-variogram models 179

7.3.3.7 Model selection and estimation of semi-variogram parameters 184

7.3.3.8 Anisotropy 187

7.3.3.9 Trends and changing drift 189

7.3.3.10 Semi-variogram interpretation 192

7.3.3.11 Spatial Interpolation by Kriging 195

7.3.3.12 The estimatoin variance and optimal sampling 200

7.3.3.13 Non-stationarity and Universal Kriging 202

7.4 Concluding remarks to this chapter 204

VIII. OXSIS DATA-MANIPULATION PACKAGE 206

8.1 Aim 207

8.2 General view of the package 207

8.3 Hardware 208

8.4 Software design 206

8.4.1 Files 210

8.4.2 Interactive dialogue 211

8.4.3 Facilities in the executive program 211

8.4.4 Data Flow 214

8.5 Using OXSIS-Manipulation facilities 217

8.5.1 Allocation 218

8.5.2 Sort Facility 218

8.5.3 Merge Facility 218

8.5.4 DEPTH Facility 219

8.5.4.1 The Spline Algorithms 220

8.5.4.2 The LAYER Algorithm 223

8.5.5 LISTING Facility 223

8.5.6 STATISTICS Facility 224

8.5.7 SEMI-VARIOGRAM Facility 225

8.5.8 KRIGING Facility 227

8.6 Concluding remarks to this chapter 228

IX. TRIALS IN SPATIAL PREDICTION OF SOIL PROPERTIES 229

9.1 Introduction 227

9.2 The problems 231

9.3 Aim 232

9.4 The area for the trials 233

9.5 Hierarchy of existent soil data in the area 236

9.6 The soil variables 250

9.7 Approach 254

9.8 Retrieval of data for the trials 255

9.9 Data manipulation to uniformize depth 255

9.10 Computation of experimental semi-variograms 256

9.11 Fitting models to experimental semi-variograms and estimation of parameters 256

9.12 The semi-variograms of the Yelford area 257

9.12.1 pH 257

9.12.2 Exchangeable Potassium 264

9.12.3 Exchangeable Magnesium 270
9.12.4 Total Carbonates 276
9.12.5 Clay 282
9.12.6 Silt 288
9.12.7 Coarse Sand 288
9.12.8 Organic Matter 293
9.12.9 concluding remarks about semi-variograms 299
9.13 Fitting models for anisotropy 305
9.14 Predicting values of soil properties from soil survey output 308
9.14.1 Class means as predictors 309
9.14.2 Typical profiles within mapping units as predictors 310
9.14.3 Mapping unit means as predictors 320
9.15 Predictions based on spatial interpolation by Kriging 320
9.16 Comparison of predictive procedures ignoring trends 332
9.17 The semi-variogram as a reconnaissance tool when trends are ignored 373
9.18 Removing trends 383
9.18.1 Modelling of trends by bicubic spline surfaces 383
9.18.2 Stratifying trends by soil mapping units 420
9.19 Semi-variograms of de-trended data 421
9.19.1 Semi-variograms of ressiduals from trend 422
9.19.2 Semi-variograms within mapping unit groupings 436
9.20 Spatial interpolation of de-trended data by Kriging 463
9.20.1 Kriging predictions using data de-trended by spline surfaces 463
9.20.1.1 Kriging of the residuals from trend 463
9.20.1.2 Evaluation of the trend surface at the random check points 464
9.20.2 Predictions by Kriging of de-trended data from stratification based on soil mapping units 482

9.21 **Comparison of predictive procedures when trends are accounted for** 486

9.21.1 Comparison of predictive procedures when trends are accounted for by modelling with bicubic spline surfaces 487

9.21.2 Comparison of predictive procedures when trends are partitioned by stratification based on soil mapping units 500

9.22 **The semi-variogram as a reconnaissance tool when trends are accounted for** 515

9.22.1 The semi-variogram of residuals from trend as a reconnaissance tool after trend modelling by spline surfaces 516

9.22.2 The within-stratum semi-variogram, after trend partitioning, as a reconnaissance tool. 517

9.23 **Concluding remarks to this chapter** 520

X. CONCLUSIONS 526

REFERENCES 536

Appendices in VOLUME 2.

List of APPENDICES

(in volume 2)

Appendix 3	3-D plots depicting spatial variability of some soil properties at Yelford.
Appendix 5.1	A comprehensive list of soil properties and other variables for which OXSIS offers data.
Appendix 5.2	MEXLARIS: Mexican Land Resources Information System. An outline of design.
Appendix 5.3	Interactive dialogue and information provided by OXSIS during the QUERY procedure.
Appendix 5.4	OXSIS Retrieval procedure. Interactive dialogue as displayed on VDU, and pro-formas to collate information to frame the retrieval.
Appendix 7.1	Constructing a surrogate function when the model is known: The non-linear Regression approach in mathematical notation, and flow-chart of strategy for model development.
Appendix 7.2	Results of an exercise in constructing a surrogate function for Cation Exchange Capacity in soils by means of regression.
Appendix 7.3	The mathematical solution for computing Equal-Area Spline curves.
Appendix 7.4	An improved method for reconstructing soil profiles from a small number of samples, (paper). It demostrates the advantages of Equal-Area Splines over other procedures for reconstructing soil depth functions and use them as depth interpolators.

Appendix 8 Description of the design and structure of the interactive and executive parts of the OXSIS Manipulation Package. Hierarchy of commands, macros, subroutines and files. Approached by individual Facilities. Examples, flow-charts and the whole set of programs are included.

Appendix 9.1 Listings of the data triads used for the trials in spatial prediction of soil properties. The listings include the random check points.

Appendix 9.2 Computer printouts with the fitting of elliptical functions to find a semi-variogram model for cases of anisotropy.

Appendix 9.3 List of soil classes to which individual points in the three sampling grids were allocated.

Appendix 9.4 List of sites grouped within each mapping unit in the trials.

Appendix 9.5 Estimation variances of Kriging estimates at the random check points for 8 soil properties, and 3-D plots of the Kriging estimates.

Appendix 9.6 Tables listing knots, observed values, fitted values and residuals from trend for Carbonate, Silt and Sand for the 3 sampling grids.

My propositions serve as elucidations in the following way:

anyone who understands me eventually recognizes them as nonsensical, when he has used them, as steps, to climb up beyond them.

He must trascend these propositions and then he will see the world aright.

.... **What we cannot speak about we must pass over in silence.**

L. Witgenstein

(Tractatus Logico-philosophicus)

I. Introduction.

In many areas of the world a rapidly increasing growth in population (about 90 million/year worldwide) is putting severe pressure on land resources. More demands are made on the productive capacity of the land to satisfy this ever increasing population. The alarming tendency of the population to increase at a geometric rate in contrast to the arithmetic rate of increase in food production brings Thomas Malthus' (1798) pessimistic views to the foreground again, particularly in recent years. Projections for the turning of the century indicate that an increase of 60% in agricultural production output in the next 20 years will be required to feed the world's population (FAO, 1981). Three lines of action can be visualized to meet the challenge:

i. Expansion of the land under cultivation.

ii. Increasing productivity of the land now under cultivation.

iii. Increasing cropping intensities on land now being farmed.

These activities require careful planning decisions concerning the selection of the best alternative land-use to maximize productivity. Thus, essential to achieving food production targets is the selection of optimal land-use patterns.

The planning of the land-use pattern in an area needs to be based on the assessment of the suitability of the land and its productive capacity for each of the alternatives. Land evaluation involves a comprehensive analysis of all relevant environmental and socio-economic factors influencing the production process and their interpretation in terms of the suitability of the land for a given purpose.

Any assessment of the land requires information on the land resources and on the potential uses. Current land evaluation methods are essentially based on physical attributes of the land.

Soil information constitutes the core of the land resource data conventionally used for land evaluation and land-use planning. Knowledge of soils is clearly integral to improving

management and increasing agricultural production in areas under cultivation. It is also integral to the planning of new arable areas.

This thesis is essentially on soil information and on procedures for providing soil information necessary for land-use planning.

Soil information is customarily obtained from and provided by soil surveys through collecting information from specific sites in the field. The role of soil survey as a form of land resources base is thus critical in any land-use planning program.

The soil varies in all directions at different rates and intensities. Soil variability strongly influences man's activities on the land and it is for that reason that many approaches for its study and representation have been suggested.

Since the soil mantle can only be observed and sampled at a rather small number of sites within a given area, only a minimal part of the total soil information of an area can be recorded during soil survey. Soil information from any unvisited location, unsampled depths, or unrecorded properties within such an area is missing information and if needed, must be predicted from the data recorded.

So, the soil is not the same everywhere and, in consequence, the variability of soils makes prediction of its properties necessary. Hence, the aim of soil survey is to provide means for predicting soil properties at any site and depth within the area surveyed. During soil survey the variability of soils is explored, the soil information collected from sites is then conceptually and spatially generalized into soil classes and mapping units. Thus, the products of conventional soil survey are soil classes, a soil map, its legend and a memoir. These are the means to convey the soil information necessary for planning to the user.

The success of a soil survey must be judged, therefore, by the extent to which predictions based on the products it generates are correct or acceptably close for a given purpose.

The utility of soil survey output relies on the quality and relevance of the soil map. Map quality depends indirectly on the purity of the mapping units in the map. On average, the purity of the mapping units produced from routine survey is not high. It is

doubtful whether the time and cost involved in producing map units of high purity are necessarily proportional to the use which can be made of them as systems for information retrieval, since the original information has been already generalized.

Soil classes, mapping units, legend and memoir are "static" or fixed products of soil survey in the sense that once decisions have been made about them they are "fossilised" in the result and any changes and updates, which require retrieval of the original raw data, are very difficult or impossible to make.

However, in spite of its shortcomings, soil survey is the standard and conventional method, to date, for providing the soil information needed for planning. It has served a purpose and has also numerous advantages.

It has been shown (Giltrap, 1981) that a soil map is a rather inefficient data retrieval system for the volume of point information collected during survey. A point in the map retains at most 6 to 7 bits of information from the not less than 100 bits in each of the ungeneralized point-data. So substantial information is lost due to generalization. This is a serious problem, since soil information is expensive.

In conventional soil survey information generalization in the form of classification and mapping is inevitable for it is a way of coping with pre-computer handling of enormous amounts of, essentially, spatially-distributed multivariate data.

The advent of computers into soil science and the subsequent development of soil information systems (SIS) offered a solution to this problem. A SIS provides powerful capabilities needed for storing and manipulating "hard" point-data and for performing the same routine tasks, as in conventional soil survey, but producing output faster and even better, yet retaining the valuable original point-data from survey.

Soil information systems thus represent an alternative to conventional soil survey procedures for the storage of large volumes of data, efficient retrieval, and powerful manipulations of data to provide the information needed for land-use planning.

The merits of relatively new techniques for spatial analysis and for spatial prediction of soil properties, and their potential in soil survey applications has been demonstrated (Burgess, 1979; Burgess and Webster, 1980a and b; Mc Bratney et al, 1982).

Spatial analysis of soil and its use in spatial prediction by interpolation are a fairly new field in soil science. Their great potential and enormous advantages have been realized and pointed out already by numerous workers.

Thus, provided the computing power and the software are available, prediction of values of soil properties with these techniques can take place facilitating and enhancing enormously the role of soil survey. However, these techniques are not yet part of the standard procedures of survey organizations. Therefore their virtues have to be tested in routine soil survey practice.

Since the fundamental aim of soil survey is to provide tools for prediction of missing information, this thesis makes a contribution into the line of enquiry of soil survey procedures, by exploring the predictive value of mathematical interpolation techniques and that of conventional survey output, and comparing them. It examines the conditions in which each provides the best tools for prediction, and explores also the possibilities of their complementary use to improve prediction.

So, this thesis deals with techniques for the solution to the problem of providing the missing soil information necessary for land-use planning. Thus, the focus of this thesis is prediction. Prediction of values of soil properties is taken here in the spatial and in the vertical sense. This work discusses special problems on interpolation of soil properties that arise during prediction.

Predictive models called here "surrogate functions", and a group of mathematical techniques for depth and spatial interpolation, namely equal-area spline functions, semi-variograms and kriging, are offered as tools for solving the problem of missing soil information in the multidimensional character space.

Thus, this work discusses the generation of missing values of soil properties from existing values of other properties acting as their "surrogates", by means of a predictive model, and offers a strategy for its generation illustrating with an example.

In the vertical sense the problem of depth interpolation is solved by furnishing an improved method based on the equal-area spline algorithm, for modelling soil depth functions and for interpolation.

Regionalized variable theory offers a sound body of theory for the quantitative description of soil spatial variability. Its main contribution is the semi-variogram. This is used here for the spatial analysis of some soil properties.

Amongst the mathematical techniques for spatial prediction, the Kriging technique is the most advanced and the best. Kriging is used in this thesis to predict values of soil properties at unvisited sites by means of the information on their spatial structures provided by the semi-variogram. The predicted values by Kriging are compared later with values predicted by conventional soil survey output under different situations of complexity of soil variability and sampling scales. Particular attention is placed on the effect that the presence of trends in the variation of some soil properties has on spatial interpolation by kriging, and also on the effect that de-trending by modelling, has on the accuracy of predictions.

The merits of both methods are determined and their possible complementary use is explored. The success of mathematical interpolation based on Kriging in predicting values of soil properties at unvisited sites is improved by the use of soil classes and mapping units as support for data stratification.

The matter of prediction is dealt with in this thesis in the context of soil information systems (SIS).

Related to SIS, this thesis examines the structural components, design and functionality of SIS. It provides a comparative review of SIS which are developed or in development in different countries. It discusses their main characteristics and some of the problems that the development of SIS has exposed. It also outlines the design of a SIS suitable to the conditions of a developing country like Mexico.

Semi-variograms and kriging are advanced techniques for spatial analysis and prediction. When incorporated as on-line capabilities in a soil information system they make a powerful tool for prediction.

This is demonstrated in this thesis by the inclusion of such capabilities in the software package for applications (the Manipulation package) of a SIS. Thus, this work contributes with the Manipulation package to the creation of a soil information system (OXSIS), at the Soil Science Laboratory, Oxford.

The importance of soil information in the context of land-use planning is discussed first in chapter 2. Understanding the nature of soil variability is essential in designing procedures for its study and representation. Soil variability is reviewed in chapter 3. Chapter 4 examines the conventional soil survey procedures for the collection of soil data. The use of soil survey output in the form of soil classes and map units in prediction is also discussed in this chapter. Soil information systems (SIS) are analysed and discussed in chapter 5, and the comprehensive review of documented SIS development for different countries is given and a comparative summary in tabular form is included. This chapter also introduces the design and capabilities of the soil information system (OXSIS), developed at the Soil SCience Laboratory, Oxford, which was used as a tool for providing the data and the data processing capabilities for the trials in prediction.

In chapter 6 special problems on interpolation are examined in some detail, and the problem of missing soil information discussed. The techniques for mathematical interpolation of soil data are formulated and discussed in detail in chapter 7. In this chapter the techniques for solving the problem of missing information are proposed.

Chapter 8 explains the capabilities of the Manipulation package developed for OXSIS. The algorithms proposed in chapter 7 are incorporated in the Manipulation package to equip users for predicting values of soil properties.

In chapter 9 the aim and the approach to the trials in spatial prediction are explained. All relevant information leading to the trials and the comparison of predictive procedures is included. The bulk of results and their discussion are included in this chapter and complemented with information held in the corresponding appendices. Finally, general conclusions are presented.

II. SOIL INFORMATION AND LAND-USE PLANNING

2.1. Agricultural planning

Planning is a scheme of arrangement of proposals of time and places of intended proceedings. In the agricultural context planning is the basis for the efficient allocation of resources in time and space to achieve production aims.

Agricultural planning may be considered at various scales, e.g. global, national, sub-national or regional and at local or community level. In order to meet the demands for food, fiber and fuel imposed by an increasing population, often, planning decisions regarding changes in the land-use pattern have to be taken. These decisions are generally formulated and implemented at the regional level.

2.2. Land-use planning

To be effective, regional land-use planning needs to be based on sound and reliable information about the existent natural resources, their present use and their productive potentials. The sophistication of regional planning procedures must be adapted to the quality of the natural resource data-base available and to the management capacities of the existent administrative system.

Land-use planning comprises three fundamental elements, namely, land, land-use and people. Their interrelationships can be appreciated in figure 2.1. Among them land suitability is of the particular concern of this thesis.

Regional land-use planning requires an assessment of the actual and potential land performance (i.e. land suitability) under the alternative land uses. Land evaluation concerns such assessments in terms of limitations and opportunities.

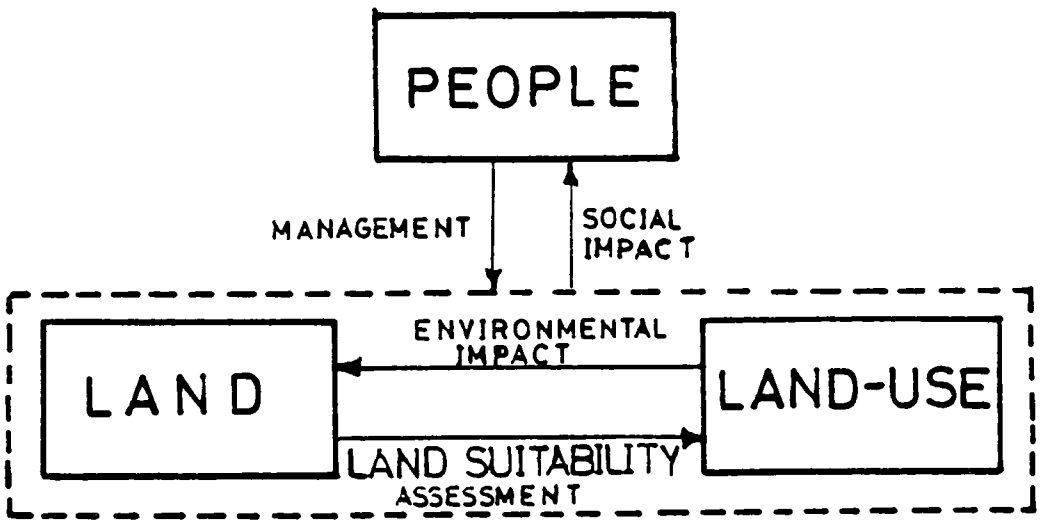


Fig. 2.1. Elements involved in land-use planning and their interactions.

2.3. Land evaluation and soil data

Land evaluation is the process of assessing land performance when used for specific purposes. It involves the execution and interpretation of surveys of land resources in order to identify and make a comparison of promising kinds of land-use in terms applicable to the objectives of the evaluation, (FAO, 1976).

Land is not the same as soil. The former is a broader concept and includes the latter plus other physical properties and attributes of the biosphere such as atmospheric conditions near the surface, plant and animal populations and human activities.

Land evaluation is based on physical land attributes, the so called land characteristics. In practice, data on land characteristics is collected in the course of soil survey, together with other meteorological, geomorphological and environmental information. These all form the resource data-base necessary for the assessment.

Soil survey data form the core of the information required for physical land evaluation. Figure 2.2 shows the place of soil data in land evaluation schemes.

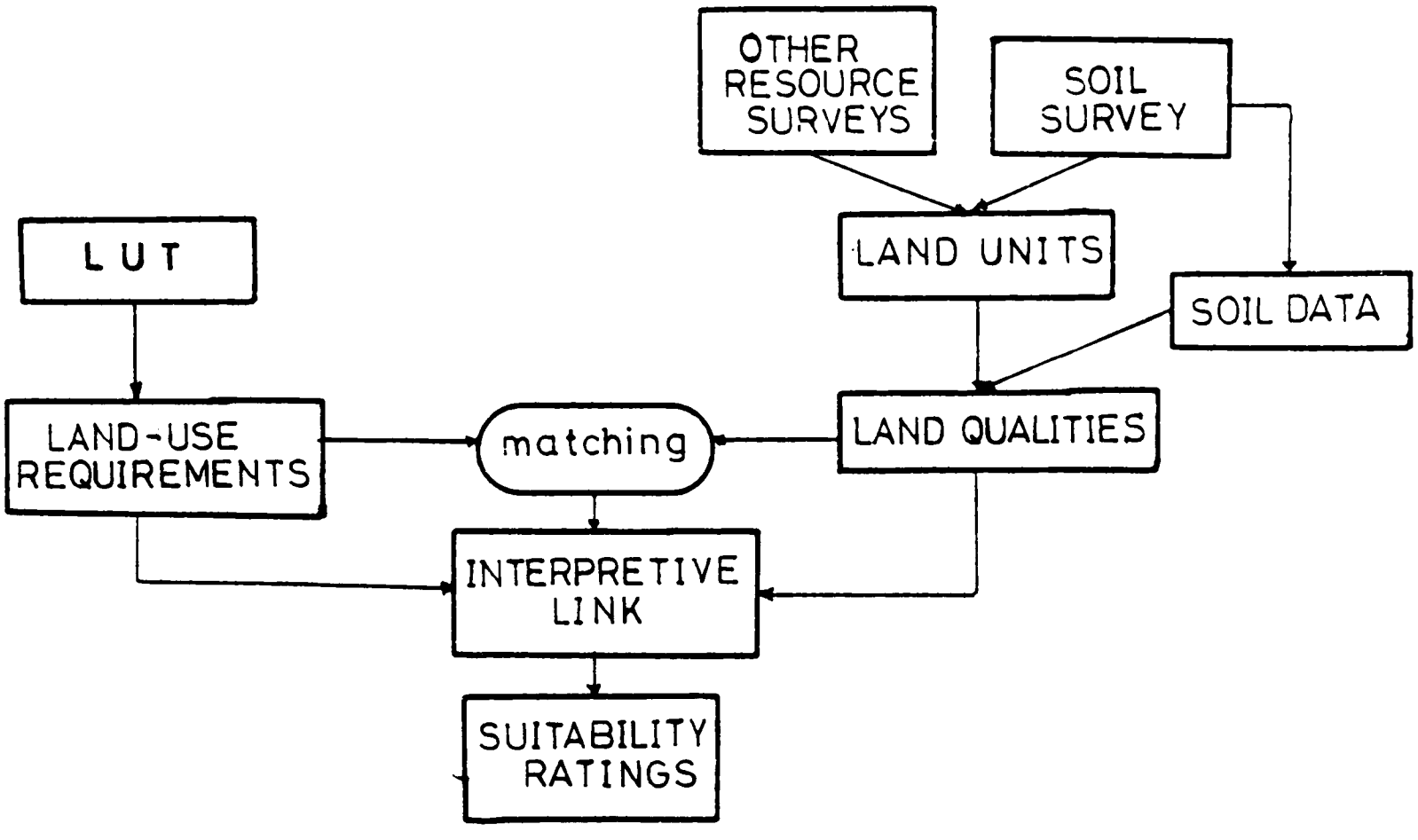


Fig. 2.2 The place of soil survey data in land evaluation schemes.

Land characteristics do not have a direct effect on land-use, but their interactions which define more complex attributes, i.e. the land qualities, affect directly the choice of alternative land-use.

The suitability of the land for a particular land-use is then assessed by matching the land qualities against the land-use requirements. During the matching procedure, criteria of "good land" are necessary and must be defined so that they can be used as a yardstick against which to compare land qualities.

For agricultural purposes a good land is a productive land. Land productivity in terms of crop yield per unit of area is a practically convenient criterion for the assessment of suitability.

Direct land evaluation methods use measurements of agricultural land productivity as indicators of land performance. Indirect land evaluation has to rely on estimations of productivity which are usually predicted from raw data, from soil

and land characteristics through an "interpretive link" between soil and land characteristics and the land-use intended.

2.4. Land evaluation and land productivity

Land productivity depends on many physical and socio-economic factors which cause variability in agricultural production conditions. In an agricultural production system, crop yield can be considered, in consequence, as a function of a universe of physical socio-economic and technological factors. In a simplistic form such a function can be written as follows;

$$Y = f\{(S_1,S_2,..S_n),(Cl_1,CL_2,..Cl_n),(T_1,T_2,..T_n),$$

$$(SA_1,SA_2,..SA_n),(EC_1,EC_2,..EC_n),(B_1,B_2,..B_n)\} \quad (2.1)$$

where Y represents physical yield per unit of area, S_n are the n soil properties influencing production. Cl_n are the n relevant climatic variables. T_n are the n topographic variables. SA_n the variables relevant to production in the social context. EC_n embrace all the economic factors in the productive process, and B_n the biological factors influencing crop production.

So the variability of this universe of factors and their multiple interactions originates the complexity of the agricultural production system.

Land suitability assessments use, implicitly or explicitly, predictions of the productivity of the land for a particular land-use as criteria of performance. The land evaluator, in effect, solves equation (2.1) intuitively or indirectly for each alternative land use and for defined tracts of land based on survey data.

The number of possible combinations defined by the status of the variables involved in the production process, (eq. 2.1), is immense and therefore their range needs to be partitioned so that productivity can be predicted.

2.5. Partitioning variability

In order to deal with the complexity in space of such a multifactor function, the land evaluator carries out in his mind a

sort of multivariate discriminant analysis. He is assisted in his task by procedures for partitioning variability in the factors of production included in equation (2.1).

The existence of several such procedures that have been developed in various fields can be demonstrated but their study in detail is not the concern of this thesis. Thus, geological and physiographic mapping provide the basis for stratification of rocks, topography and geoforms. Statistical procedures are standard practice for partitioning variability of meteorological data in time and space.

Particularly useful in land evaluation methods for partitioning the universe of socio-economic and technical conditions related to land use is the "land utilization type" (LUT) concept, (Beek, 1978). The land utilization type is an attempt to give structure to the complexity of land-use. It is a specific way of using the land, actual or alternative, and is described in terms of kind of crop, labour, capital, management, technology, and scale of operations. The LUT is the subject of land evaluation and land units are the objects.

In principle then the variability of each group of factors can be partitioned into strata within which they can be assumed to be constant. So, usually the evaluator deals with somewhat simpler equations than (2.1) representing strata in which some of the terms have been removed .

Information about soil and site properties in the form of land characteristics is the raw material for land evaluation, hence its importance. If all the other groups of variables are stratified and can be assumed to remain constant, crop yield and hence productivity would depend on the variability of soil properties only. Equation (2.1) could be transformed into:

$$Y = f\{(S_1, S_2, \ldots S_n)\};\ (Cl,T,B,SE,EC) = cnst \qquad (2.2)$$

where cnst=constant. More explicitly:

$$Y = f[(clay,pH,O.M.,P,K,..S_n)];\ (Cl,T,B,SE,EC)=cnst \qquad (2.3)$$

Hence, all predictions of productivity for land evaluation may depend on the variability of soil exclusively, being the variability of the other factors conveniently stratified.

2.6. The need for soil data in land-use planning

Since land suitability assessments are made essentially on its physical attributes, soil and site data are the core of the information needed in land evaluation for land-use planning.

All current land evaluation methods, invariably, require soil data provided from field surveys. This emphasizes their importance in the planning context, and at the same time provides the framework for this study.

This thesis, therefore, is essentially on soil data. While it is assumed that equations 2.2 and 2.3 remain valid in the course of this study and that the role of soil data in land-use planning is of primary importance, conventional and new methods for studying soil variability and for providing the soil information necessary for planning are examined in detail and form the central part of this thesis.

Having established that soil data are an essential element in land-use planning, the following chapter examines the nature of soil variability in the field in order to provide the necessary background for chapter 4 on the procedures for collecting soil information from field survey.

2.7. Concluding remarks to this chapter

i. Agricultural planning involves decisions about changes in the land-use pattern.
ii. Land-use planning requires an assessment of the suitability of the land for alternative land uses.
iii. Land evaluation is essentially in terms of physical atributes.
iv. Soil data are the core of the resource data-base for land evaluation.
v. land-use planning needs the provision of soil information.

III. SOIL VARIABILITY

The variability of soils is a fact that plays an important role on influencing man's activities on the land. Its importance for the land user is such that many economic activities rely upon a good knowledge of soil variability. Over time it has been the concern of many workers who have undertaken a variety of approaches for its study, quantification and portrayal.

This chapter deals with soil variability. It is the object of study by soil scientists and the main cause for developing procedures for predicting soil information at unvisited sites.

Soil variability is described in this chapter as it is, its sources, nature and scale in relation to the landscape, and the techniques, to date, for measuring it and displaying it. Such measurements and a good working knowledge about the structure of soil variability in an area are essential to the prediction of soil properties at a site.

In an ideal world, the soil mantle can be imagined as being totally uniform in origin, nature and in all its properties, within the bounds of an area of interest.

Prediction of soil conditions at unvisited sites in such a case would be unnecessary. A single soil sample would suffice to predict with certainty the nature of the soil everywhere else within that area. A landscape made up of a collection of such areas, showing well-defined, easily recognizable and mappable boundaries would make prediction of soil properties or behaviour a rather easy task. It would take as many soil samples as recognizable parcels like those described above, to provide complete information about the soil mantle in such a landscape. The soil mantle would be the same at any depth. The samples could be taken at any soil depth within such parcels or areas. Soil survey effort would be minimal and only used to locate and draw on a map those externally well-defined soil boundaries of the different parcels. The user would be fully equipped by such

maps to answer any questions on the condition of the soil, or to answer any queries based on soil information with absolute certainty.

Reality is not like that. The soil mantle is rather heterogeneous. It shows anisotropies in all directions. Some of the changes are reflected in external features on the ground but in most cases they are not so clear, or they do not appear at all.

The soil is the result of many interacting natural forces operating simultaneously at many rates and intensities, and over different spatial and temporal scales. These natural forces together with man's activities on the land are the sources of soil variability.

3.1. Sources of soil variability

The soil is an open, dynamic and exceedingly complex system. Its properties may change at widely different rates over space and over time. Some are more stable over time (e.g. physical properties) than others (e.g. chemical), but not necessarily over space.

Jenny, (1941), conveniently named "Soil Forming Factors" all those factors which give origin to the state of the soil system and its properties. The soil-forming factors are assumed to behave as independent variables in Jenny's equation of soil:

$$S = f\,(cl,\ o,\ r,\ p,\ t) \qquad (3.1)$$

In reality of course, climate (cl), organisms (o), relief (r), and parent material (p), interact in a variety of combinations over time (t) and space. These combinations originate soil variability which can be percieved at different spatial scales.

Hoping that such an equation could lead to predictions of the state of any given property of the soil, Jenny assembled most of the correlations and quantitative relationships known at the time. However, the postulates of this work remain largely theoretical and the equation of the soil forming factors remains unsolved. Nevertheless, Jenny's work directs attention upon the

main sources of soil variability, their scale of operation and their effect in producing changes in the soil.

3.1.1 The natural landscape

Similar sources of soil variability are applicable to both the natural and the cultivated landscape, where cultural activities contribute to soil variabilty also.

(a). Parent Materials

Soil parent materials may be of very diverse nature and origin and they may vary over short distances. Even within a primary rock the chemical and mineralogical composition may be variable. Variability is even greater within transported materials and sediments.

Some generalizations can be made. For instance soils formed in situ tend to be less variable than those formed on transported materials (Beckett and Webster, 1971). Volcanic materials tend to be chemically and mineralogically more uniform than transported sediments, e.g. a landscape of granitic plutons may produce less variable soils than a sequence of interbedded sedimentary rocks. Even if the origin of the soil is an apparently uniform sedimentay rock, this can produce differences due to the form and time of deposition, and the geochemical gradients associated with them.

Other geomorphic degradative processes like soil erosion may be added up contributing to short, medium and long range variation.

Jenny, (1941) showed the effect of parent material on soil variability. Within surprisingly short distances (30 ft intervals), the lime content of fine earth (< 2mm) varies enormously from 0 to 19% of $CaCO_3$. This is believed to be a consequence of the geologic origin of the soil matrix.

The nature of the elements released during in situ rock decay has a specific bearing on soil formation and variability, e.g. acid igneous rocks release considerable amounts of quartz and are rich in monovalent cations whereas basic igneous rocks are high in Ca and Mg. Different weathering rates of the component mineral also contribute to variability of soils, e.g. plagioclases weather more readily than the potasium feldspars;

biotites decay more rapidly than muscovites, and amphiboles are not as resistant as pyroxenes.

(b). Climate.

Climate produces gradual long-range changes in soil as a result of its continental gradations. The distribution of some soil properties such as CaCO3 and other soluble constituents, C/N ratio, etc., can be said to be strongly associated with climatic intracontinental changes. Soils of arid regions are richer in soluble constituents than those of humid zones. Jenny (1941) has shown interesting correlations between rainfall and clay content. He also noticed the distribution of nitrogen and organic matter as a function of temperature. Yost et al, (1982a) have also illustrated the effect of temperature and rainfall on long and short-range changes in soil properties. Continental gradients in climate also produce gradual changes in soil across the tension zone between two climatically determined plant communities, (Beckett and Webster, 1971).

(c). Topography

The configuration of the land surface produces regional differences in soil properties. Relief determines local differences in water penetration into soils, and hence runoff and erosion. This can produce close associations between recurrent patterns of different soils and recurrent patterns of land forms. Short-range variation in soil can be caused by erosion or drainage patterns associated with micro-relief. Differences in soil water regimes caused by topography have also an important effect on other soil properties, (Burrough, 1983a,b), e.g. organic matter accumulations and variation, plant nutrient concentrations in soils. Cuanalo and Webster, (1970) found regional differences in soil properties associated with regional contrasts in topography and geology. Modenesi, (1983) describes the effect of topography on soil morphogenesis and heterogeneity in a tropical environment.

(d). Physical and chemical processes

Many physical and chemical processes are taking place within the soil. These also have an effect on the short-range variability of some of its properties. Chemical processes determine the localized concentration of some elements in certain soils, for example, the tendency of iron-pans to form

pockets (Beckett and Webster, 1971), or the formation of CaC03 concretions in some arid soils. Processes such as diffusion, solifluxion, etc., and the chemical behaviour of some elements determine the fluctuations in mobility and concentration of inorganic and organic forms in the soil. Processes of this kind smooth out or induce lateral and vertical soil variability.

(e). Biological activities

Organisms, specially plants, play an important role in soil variability. Types of vegetation determine the amounts and nature of organic matter incorporated into soil, e.g. prairie soils, rain forest, etc. This also influences the kinds of bacteria, fungi and other micro-organinisms in soils and consequently their effect on soil reaction (pH), exchangeable bases and degree of saturation.

The variability of some plant-nutrient elements is particularly linked to biological activities, e.g. nitrogen fixation and availability. Macro-organisms produce soil heterogeneity too, e.g. concentrations of excreta, mechanical effect by wallowing, casts formation, mound building, all affect short-range variability, (Beckett and Webster, 1971).

Even in relatively undisturbed soils, e.g. forest soils, the variability of some properties as indicated by their CV is considerable. Total N, total P, extractable P, K, Ca and Mg showed ranges of CV's of 4-12, 4-46, 19-65, 11-41, 17-44 and 9-43 respectively, (Dahiya et al, 1984).

By contrast, in arid environments, the variability of chemical properties in desert soils is reported to be their most striking feature. Thus, variations in nitrate concentrations in a desert soil were found to be even greater than for a fertilized and cultivated soil (Hunter et al, 1982).

3.1.2 The cultivated landscape

Additional sources of soil variability are introduced by man's activites on the land. To the "intrinsic" soil variability that arises naturally, "extrinsic" soil variability is imposed as a result of distinct geometries used in soil-water-chemical management programmes. These introduce or smooth out natural variability in some soil properties.

Some generalizations can be made. Soil properties most affected by management (e.g. chemical) tend to be more variable than more stable properties (e.g. physical). Beckett and Webster,(1971), have summarized most pre-1970 investigations into soil variability. They used coefficients of variation (CV) as measures of total soil variability and found high CV's (60%) for properties such as available K, P, Mg and Ca, most affected by management, whereas for properties such as sand, clay or total P, the CV's were considerably lower (approximately 20%). Similar trends are reported by other workers, e.g. Courtin (1983), Dahiya et al (1984).

Wilding and Drees (1983), found that chemical parameters, especially exchangeable Ca, Mg and K are extemely variable, with mean CV's of 50 to 70% and ranges up to 160%. Most physical properties i.e. Atterberg limits, particle size distribution and bulk density were found in average less variable. Nielsen et al (1973), and Babalola (1978), report observed CV's for hydraulic conductivity ranging from 50% to 150%. Beckett and Webster (1971), have shown that soil moisture measured as pF varies considerably even over short distances, (i.e. CV = 3-10% within $1m^2$ and 11-26% within $10m^2$). Such variations also occur seasonally.

Dahiya et al (1984), summarize data on soil variability. From their work the effect of management on soil variability is clear. Relatively less affected or unaffected properties (i.e. soil texture, organic carbon, bulk density, total K and total N) showed CV's betwen 6-19% whereas more affected properties (i.e. nitrates, available K, electric conductivity) CV's varied between 23 to 81%.

Thus, additional heterogeneity on some less durable soil properties due to management is reflected in differences between fields that are superimposed on the variation already present in the native soil. Furthermore, much of the within-field soil variability, particularly of soil chemical propeties, is also a result of management factors. High within-field variability is introduced by recent fertilizing, soil amelioration and grazing, (Beckett and Webster, 1971). Thus, Biggar (1978), concluded that fertilizer applications tend to increase the variability of

soil nitrogen, and the CV value may be as much as 100-150% greater for fertilized plots as for unfertilized plots.

In general then, the variability of soils in cultivated landscapes is greater than in comparable areas of the natural landscape. A non-fertilized uncultivated soil should exhibit less variability than a non-fertilized cultivated one, and this in turn should be less variable than a fertilized cultivated soil. It might be supposed that in cultivated landscapes between-field soil variability is greater than within-field variability.

3.2. Variability in relation to the size of the area sampled

Numerous studies suggest that soil variability increases with the size of the area sampled, (Beckett and Webster, 1971; Dahiya et al, 1984), but the form of the relationship remains unknown. Theoretical and practical work to elucidate spatial variability of soil in relation to scale is fairly recent and still in progress, (Burgess and Webster 1980a, Burrough, 1983a,b).

Wilding and Drees (1983) observed that the magnitude of variability generally increases with increasing scale factor from pedons to polypedons to mapping units of a given series, to all soils within the survey area.

The scale factor (i.e. areal extent) associated with changes in spatial variability of soil depends on the property in question. Generalizations are difficult and some available data are still conjectural. However, some workers have provided some data. Thus, Nortcliff (1978), indicates that the increase in spatial variability with increases in the scale factor may be either linear or curvilinear.

At micromorphic scale, variability of important soil hydraulic properties like pore-size distribution may be considerable over distances of only a few millimeters, (Wilding and Drees, 1983). Even at the scale of a sampling site, variability of some soil propertiies within horizons or layers that appear morphologically uniform may still be significant.

Data from Oxfordshire have been used in this thesis to illustrate the variability of some soil properties with the size of the area sampled. The same data provide the material for spatial analysis and the trials in prediction dealt with in chapter 9. Table 3.1. shows changes in variability with the size of the area sampled in terms of values of CV. Table 3.2 shows changes of average within-class variability for the same area sizes.

From both tables it can be seen that the increase in variability with the area size is not as clear as could have been expected. This is due to the way the areas were sampled. The areas lay nested, so all the high values are concentrated in relatively small areas that occupy relatively high proportions of the smallest area (38 Ha) and therefore progressively smaller proportions of the larger areas. Nevertheless, the data are meaningful in terms of explaining soil-variability/landscape relationships (see chapter 9). Figure 3.1 shows graphically the data from table 3.1. The scale in the CV axis has to be compressed to accomodate the high CV values of Carbonate and Sand, hence the increase in CV with area size for other properties may not be quite noticeable.

Figures 3.2.(a) and 3.2.(b) show CV bars for 8 soil properties corresponding to total observed values (a), and average within-class CV's (b) for the same area. This allows one to examine the effect of classification on total variability as represented by CV values. It can be seen that for the most variable properties (e.g. Carbonate, Sand), the reduction in variability (i.e. CV) achieved by soil classes was considerable. In some other cases, particularly for less variable properties (e.g. clay, pH) the reductions achieved by classification were modest.

To illustrate variability, 3-D plots of property values on the plane for the two most variable (i.e. Carbonate and Sand), and the less variable (i.e. pH) properties in the areas sampled, are shown in figures 3.3.(a) to (c). Simmilar plots for the rest of the 8 soil properties studied (see chapter 9) are shown in appendix 3.

Table 3.1 Observed Variability (Coefficients of Variation) for some Properties from Oxfordshire Soils in Relation to the Size of the Area Sampled

Soil Property	AREA			
	50 m. Grid (38 Ha.)	100 m. Grid (111 Ha.)	200 m. Grid (372 Ha.)	1000 m. Grid (10,000 Ha.)
Clay	23.93	22.38	27.11	32.33
Silt	19.32	23.29	20.82	34.25
Coarse Sand	121.86	113.48	109.19	89.35 *
Organic Matter (Abundance)	30.75	34.51	35.49	42.81
Potassium (Exchangeable)	56.41	71.43	71.96	77.58
Magnesium (Exchangeable)	31.54	37.75	34.55	91.81
pH	6.30	6.25	6.81	11.05
Carbonate (Ca + Mg)	156.57	140.09	156.12	119.57 *

* In both these cases all the high values are concentrated in relatively small areas that occupy relatively high proportions of the smallest area C, and therefore progressively smaller proportions of the larger areas.

Table 3.2 Average Within-Class Variability (as measured by CV values) of some Soil Properties in Relation to Size of Area Sampled *

Soil Property	AREA		
	50 m. Grid (38 Ha.)	100 m. Grid (111 Ha.)	200 m. Grid (372 Ha.)
Clay	13.14	15.93	17.14
Silt	17.71	18.08	14.56
Coarse Sand	49.27	50.46	46.36
Organic Matter (Abundance)	24.35	27.46	26.33
Potassium (Exchangeable)	40.33	45.51	47.16
Magnesium (Exchangeable)	27.83	27.89	26.24
pH	3.54	3.96	3.82
Carbonate	91.96	67.13	55.22

* Values for the 10,000 Ha. area were not available.

Fig. 3.1

Variability of soil properties with the size of the area sampled in Yelford, Oxfordshire

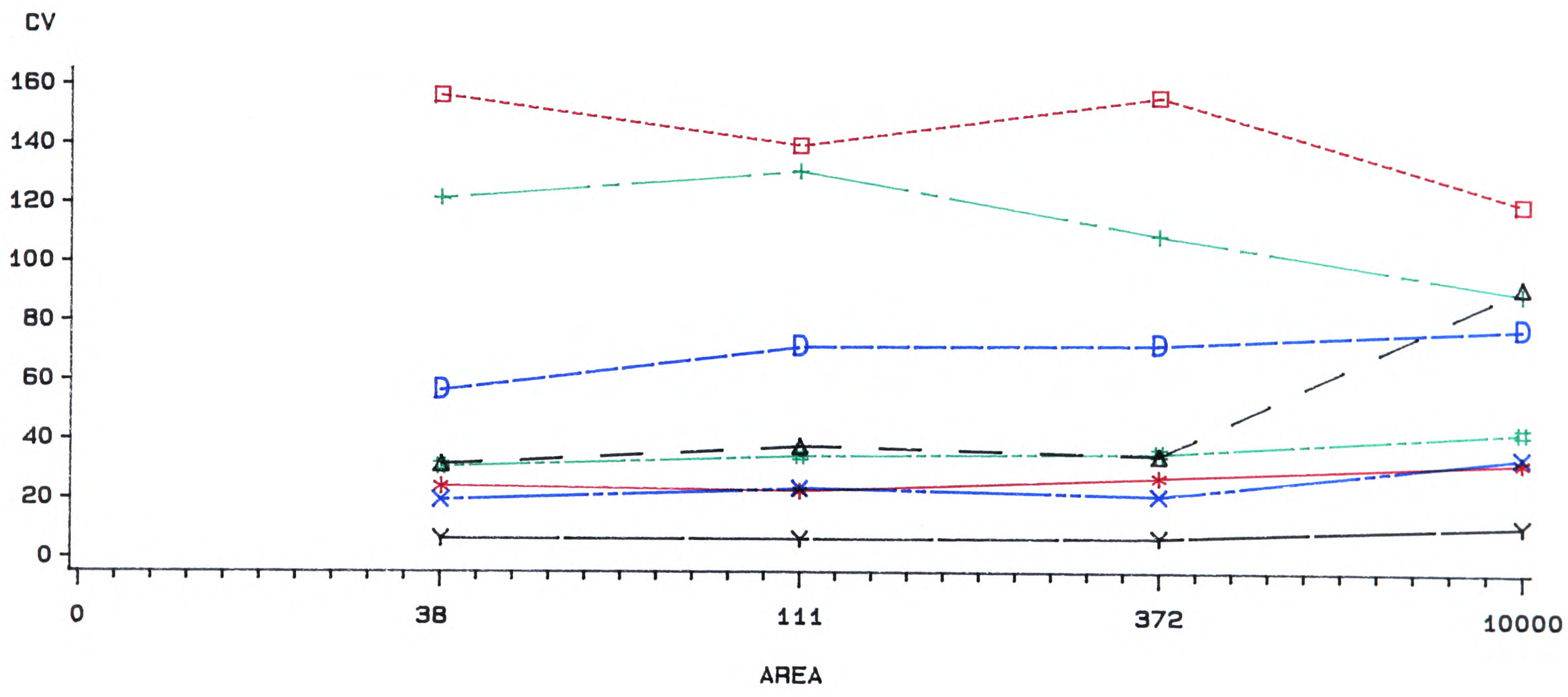
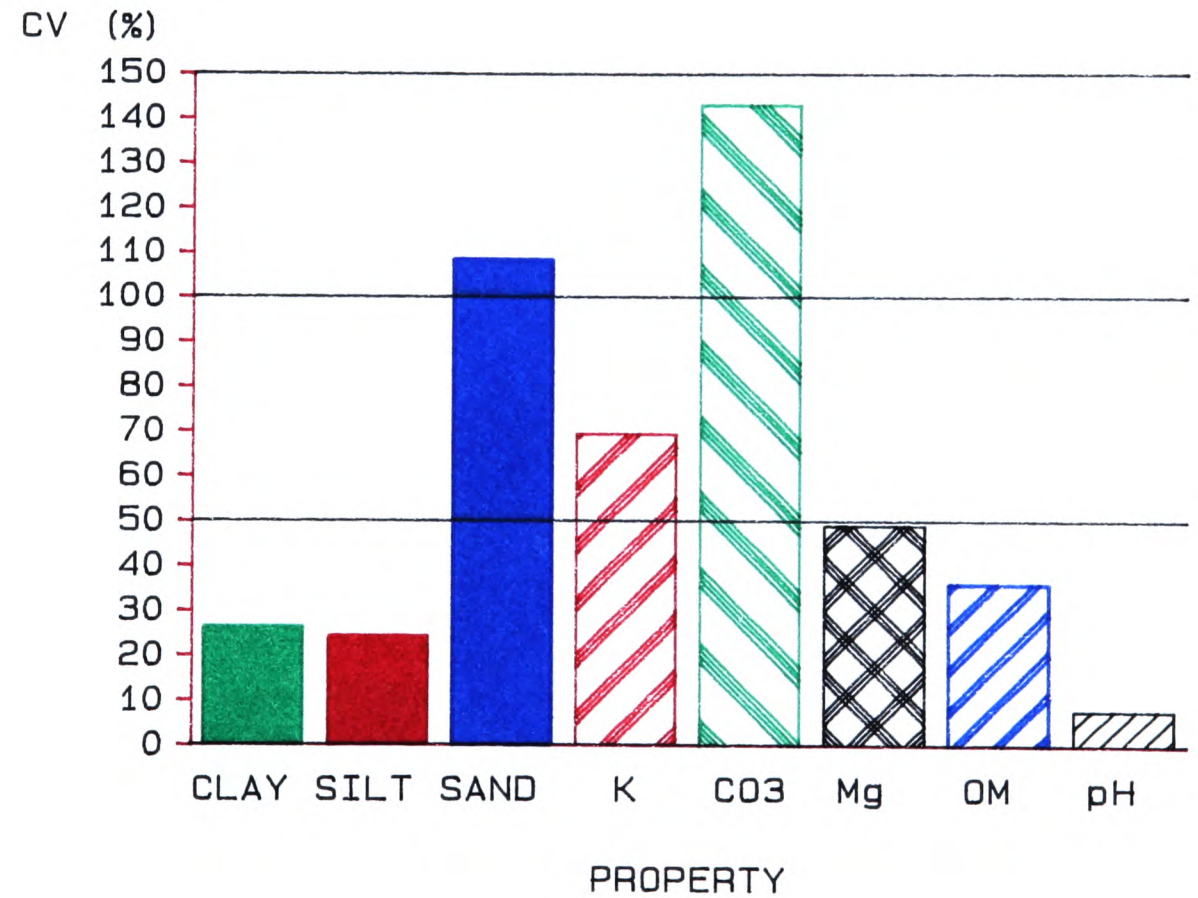
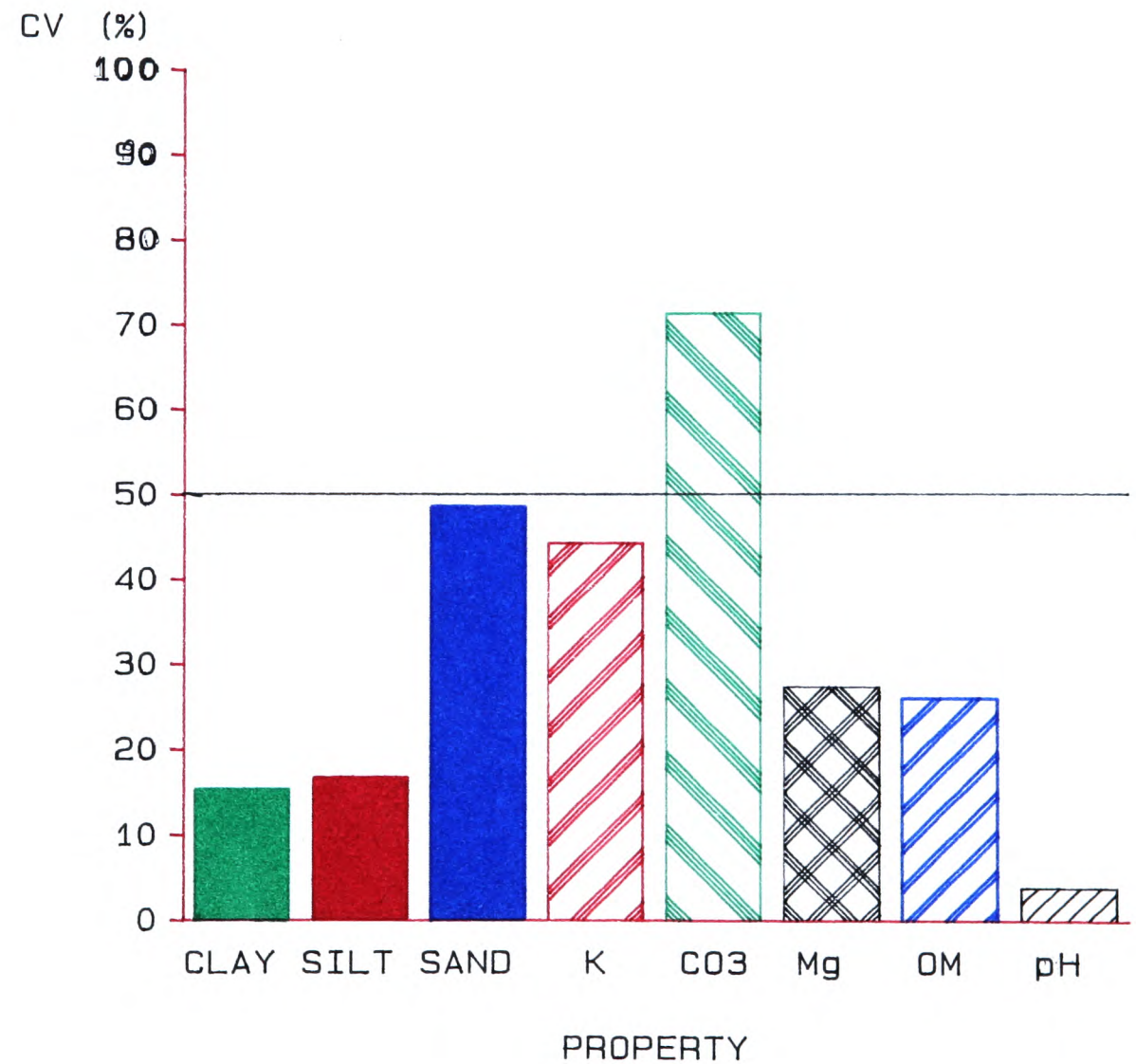


Fig. 3.2(a) *Observed variability (CV)* of some soil properties in the Yelford area, Oxfordshire*



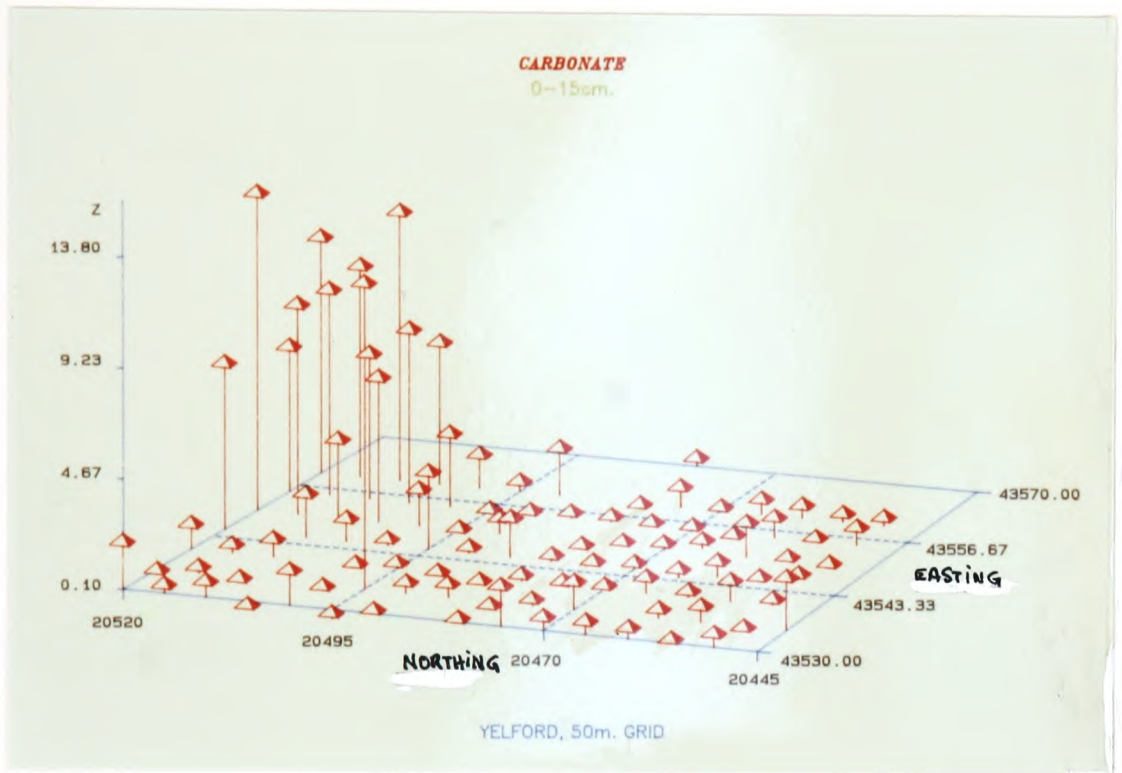
* Values pooled from areas of 38, 111 and 372 Ha.

Fig 3.2(b). *Within-class$^{\Psi}$ variability (CV)* of some soil properties in the Yelford area, Oxfordshire*



Ψ REFERS TO AVERAGE OVER ALL SOIL CLASSES (SEE CHAPTER 8)

$*$ AS IN FIG 3.2(a)


CARBONATE
0–15cm.
Z
13.80
9.23
4.67
0.10
20520
20495
NORTHING
20470
20445
43530.00
43543.33
43556.67
43570.00
EASTING
YELFORD, 50m. GRID

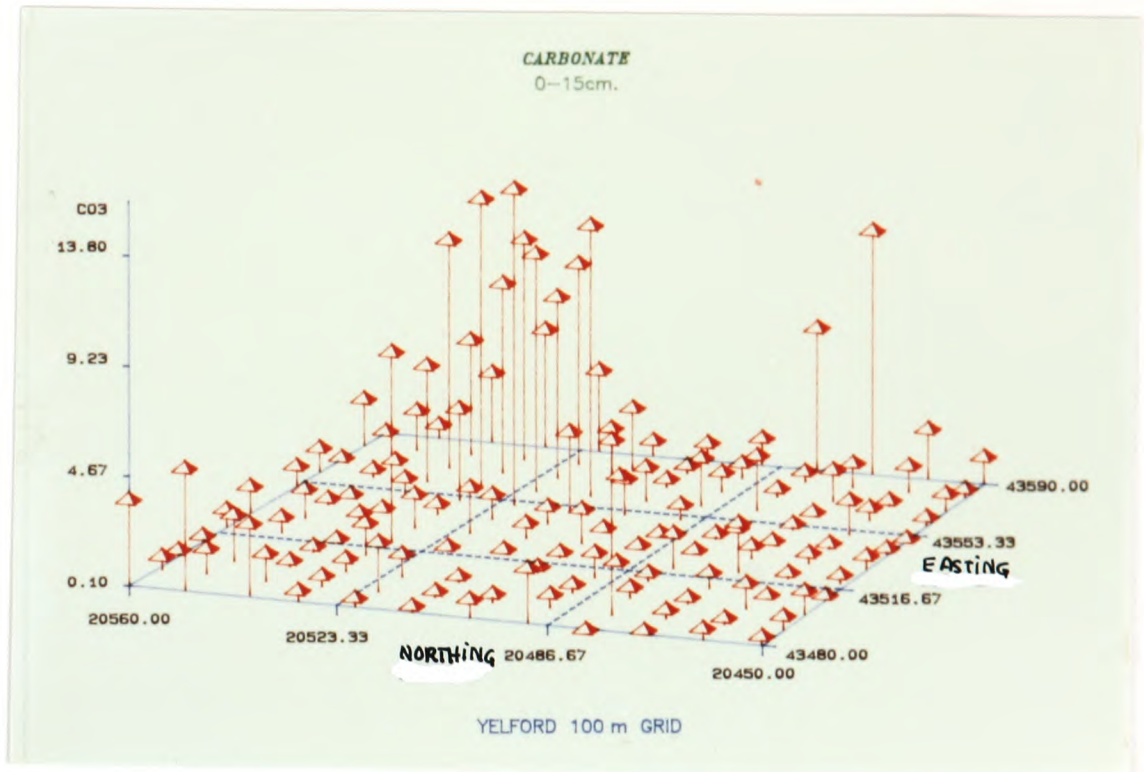
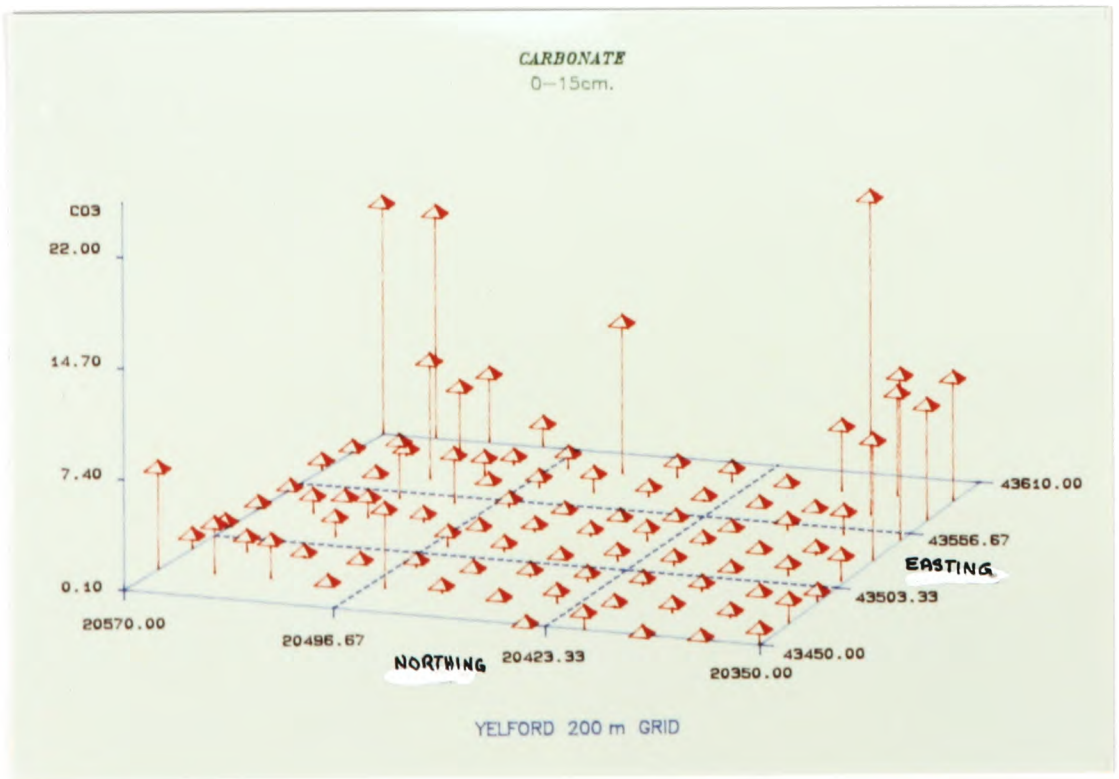
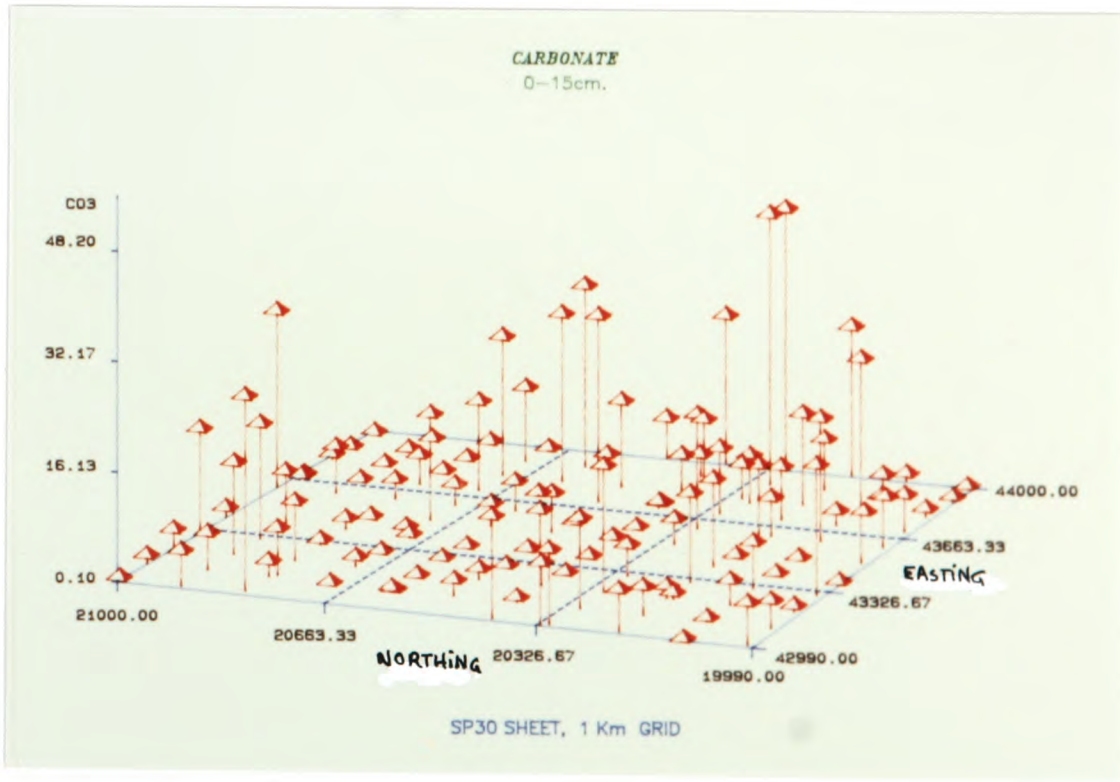
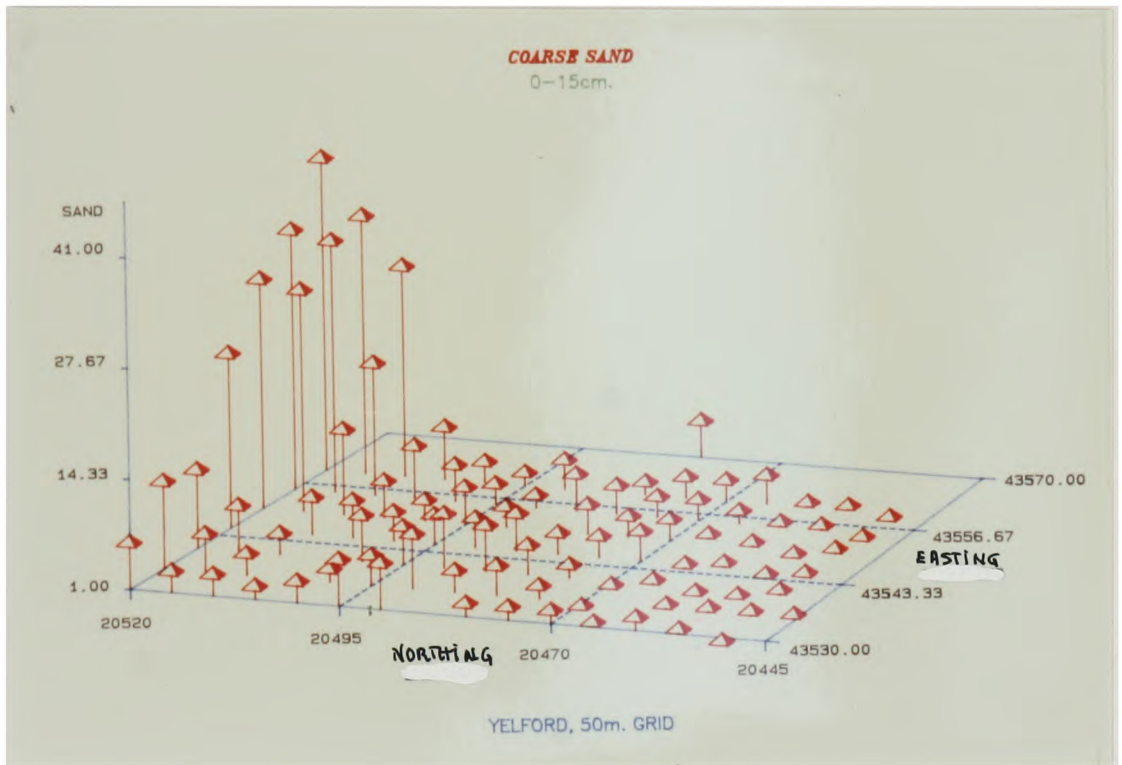
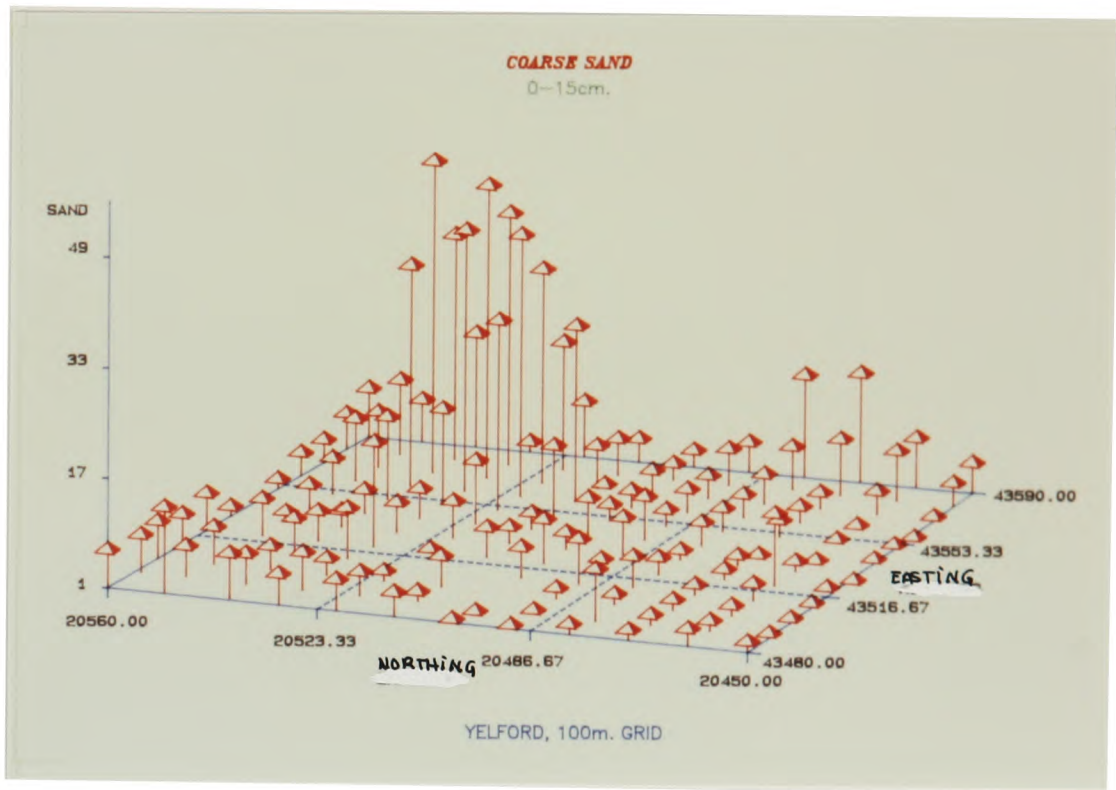

CARBONATE
0–15cm.
C03
13.80
9.23
4.67
0.10
20560.00
20523.33
NORTHING
20486.67
20450.00
43480.00
43516.67
43553.33
43590.00
EASTING
YELFORD 100 m GRID

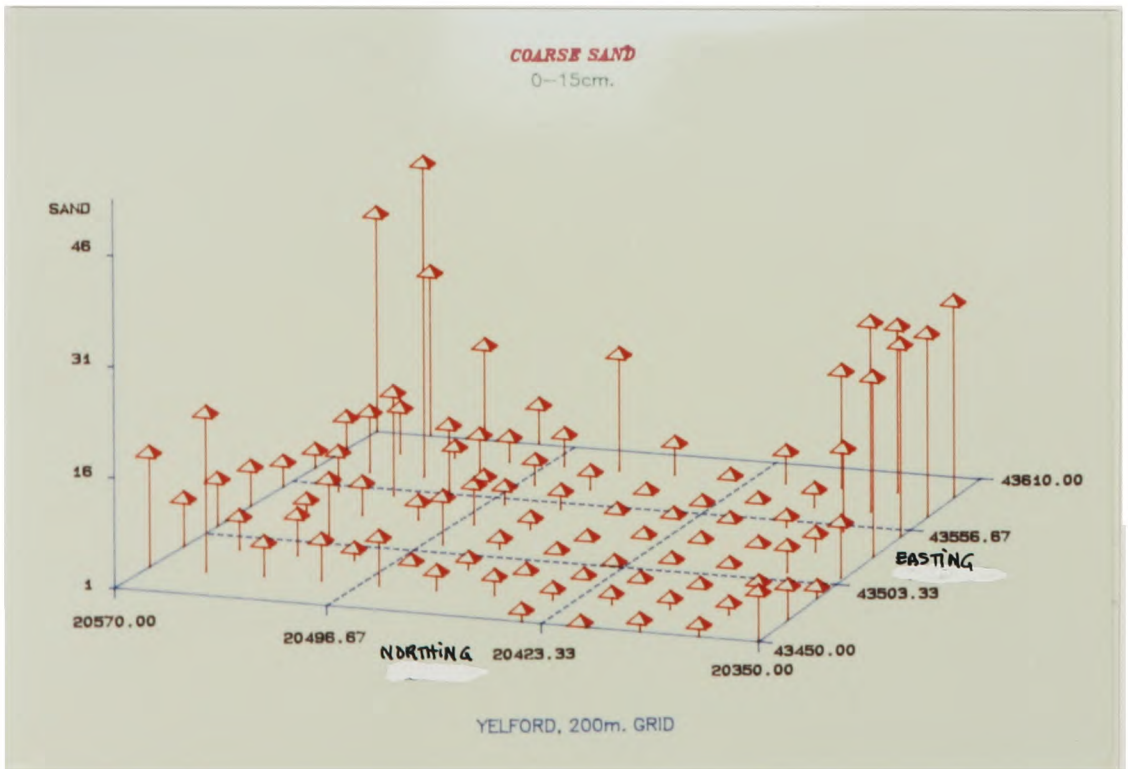
Fig 3.3(a)

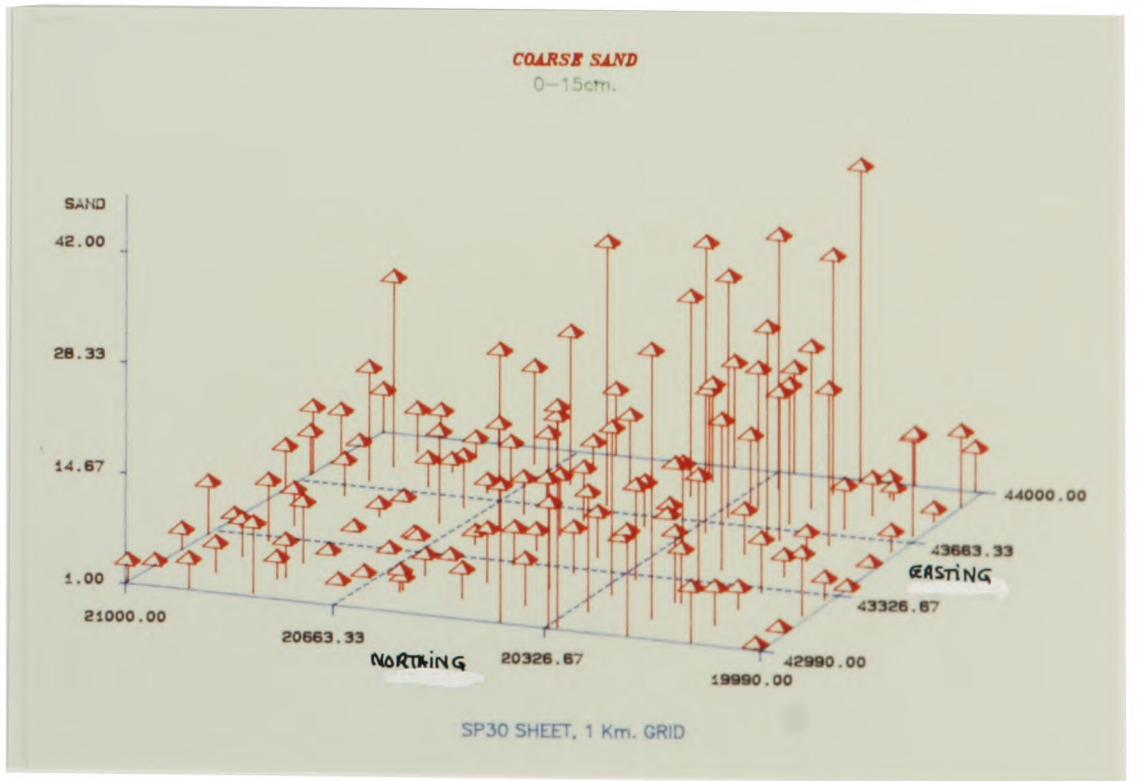

CARBONATE
0–15cm.
C03
22.00
14.70
7.40
0.10
20570.00
20496.67
NORTHING
20423.33
20350.00
43450.00
43503.33
43556.67
43610.00
EASTING
YELFORD 200 m GRID

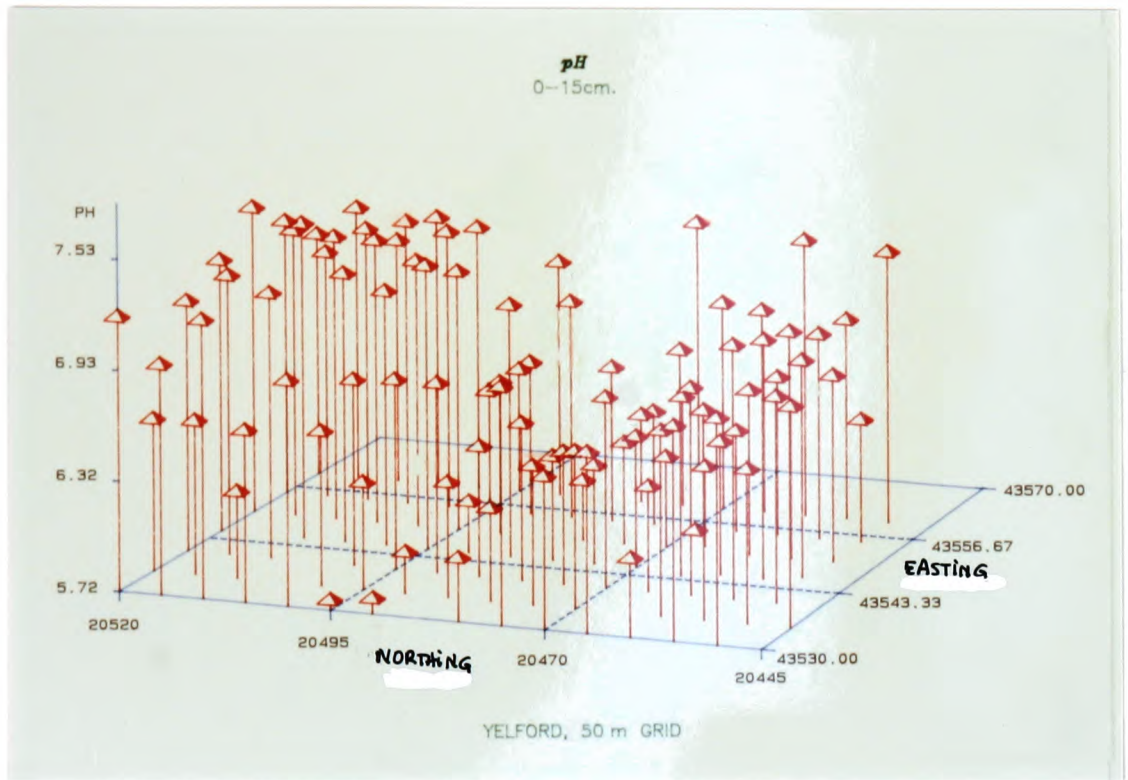

CARBONATE
0–15cm.
C03
48.20
32.17
16.13
0.10
21000.00
20663.33
NORTHING
20326.67
19990.00
42990.00
43326.67
43663.33
44000.00
EASTING
SP30 SHEET, 1 Km GRID


COARSE SAND
0–15cm.
SAND
41.00
27.67
14.33
1.00
20520
20495
20470
20445
NORTHING
43570.00
43556.67
43543.33
43530.00
EASTING
YELFORD, 50m. GRID


COARSE SAND
0–15cm.
SAND
49
33
17
1
20560.00
20523.33
20486.67
20450.00
NORTHING
43590.00
43553.33
43516.67
43480.00
EASTING
YELFORD, 100m. GRID


COARSE SAND
0–15cm.
SAND
46
31
16
1
20570.00
20496.67
20423.33
20350.00
NORTHING
43610.00
43556.67
43503.33
43450.00
EASTING
YELFORD, 200m. GRID


COARSE SAND
0–15cm.
SAND
42.00
28.33
14.67
1.00
21000.00
20663.33
20326.67
19990.00
NORTHING
44000.00
43663.33
43326.67
42990.00
EASTING
SP30 SHEET, 1 Km. GRID


pH
0–15cm.
PH
7.53
6.93
6.32
5.72
20520
20495
NORTHING
20470
20445
43570.00
43556.67
EASTING
43543.33
43530.00
YELFORD, 50 m GRID

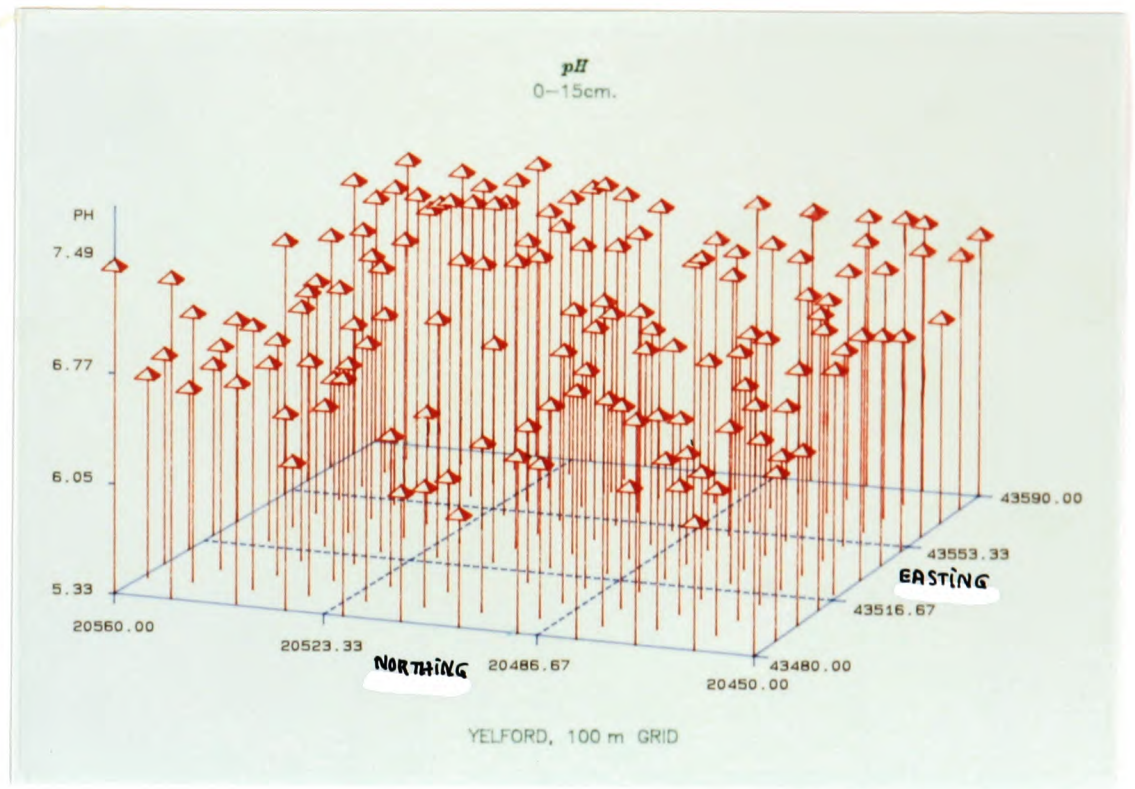
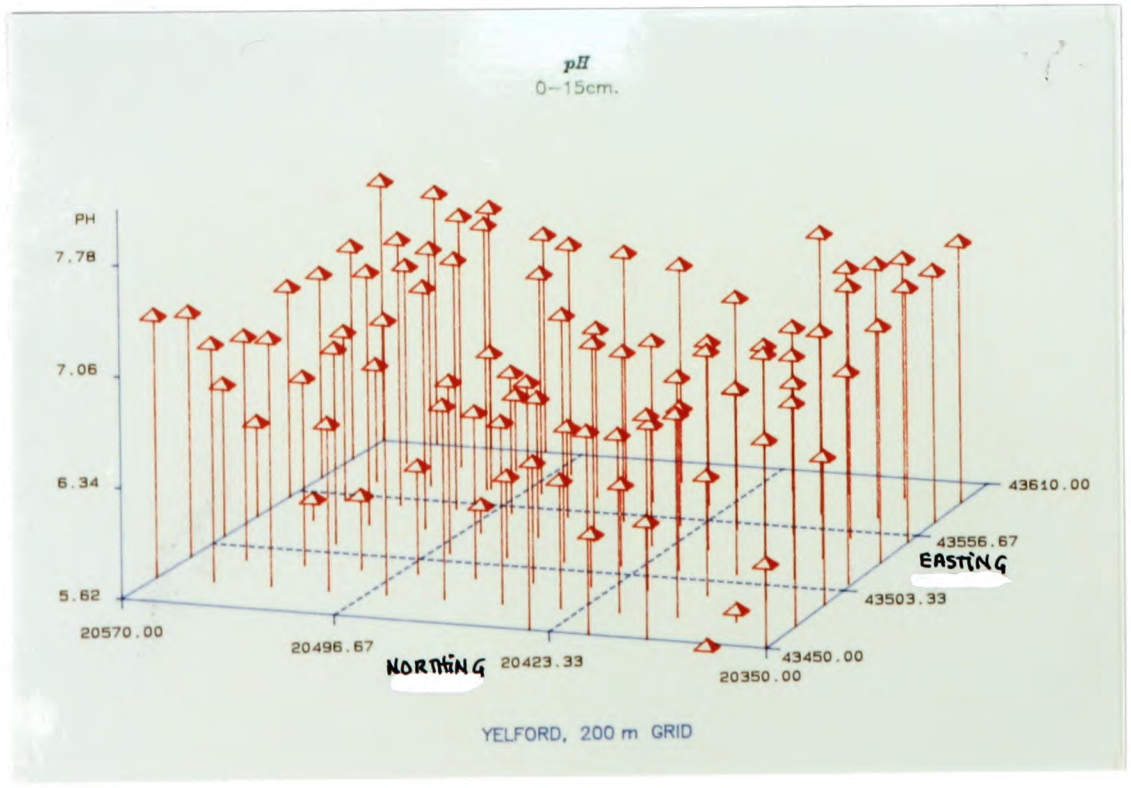
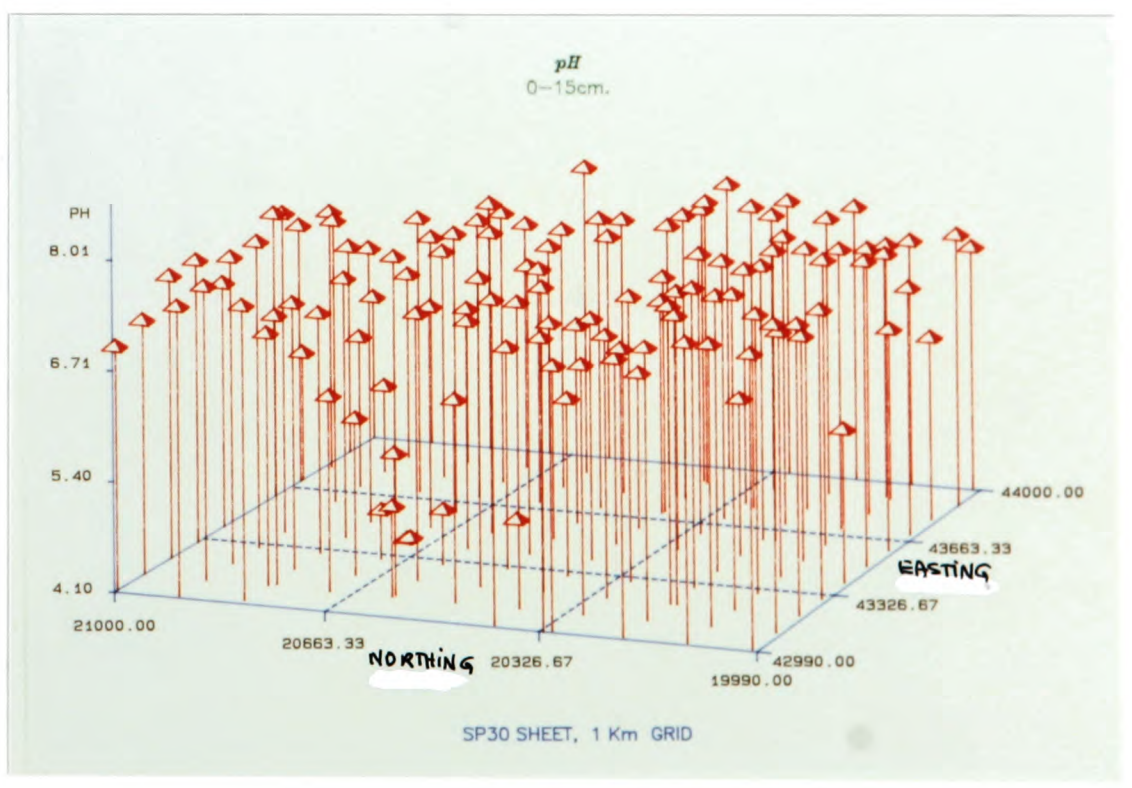

pH
0–15cm.
PH
7.49
6.77
6.05
5.33
20560.00
20523.33
NORTHING
20486.67
20450.00
43590.00
43553.33
EASTING
43516.67
43480.00
YELFORD, 100 m GRID

Fig. 3.3(c)


pH
0–15cm.
PH
7.78
7.06
6.34
5.62
20570.00
20496.67
NORTHING
20423.33
20350.00
43610.00
43556.67
EASTING
43503.33
43450.00
YELFORD, 200 m GRID


pH
0–15cm.
PH
8.01
6.71
5.40
4.10
21000.00
20663.33
NORTHING
20326.67
19990.00
44000.00
43663.33
EASTING
43326.67
42990.00
SP30 SHEET, 1 Km GRID

3.3. Directional components of soil variability

Soil shows anisotropies in all directions. For convenience vertical variability is distinguished from lateral or spatial soil variability . In nature they occur together and are inter-related.

3.3.1 Vertical soil variability

Soil properties vary also with the soil depth. Variability is not the same at all depths. Nor does it change with depth in the same way all seasons or for all properties.

A dug pit exposes a vertical section of the soil, the soil profile. Changes of soil properties are reflected in the soil profile face in the form of more or less defined strata or layers roughly parallel to the soil surface.

Conventionally, the soil profile is divided in layers, termed horizons using observable changes in morphological properties.

In dividing the soil profile into layers it is assumed that individual properties remain sufficiently uniform within the limits of each layer as defined in the field, i.e. that the soil is vertically isotropic within the thickness of each layer.

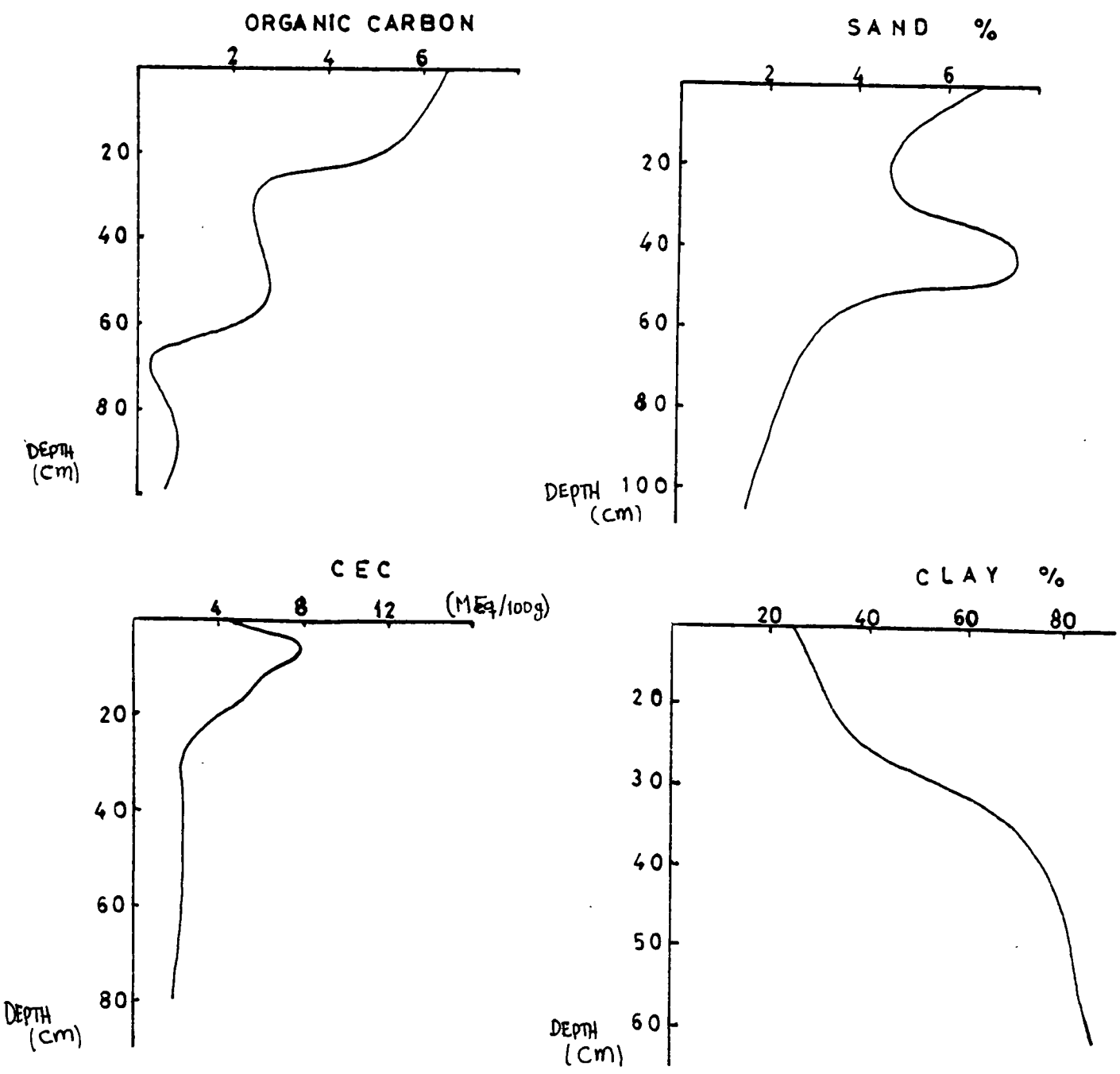
For the purpose of soil profile characterization, a sample is taken from each of the layers on the profile's face and analysed. In some other instances samples are taken ad-hoc at fixed depth intervals. A value for any given soil property is extracted and used to characterize a particular depth interval.

Many are the factors that affect vertical soil variability, making predictions difficult. However, some rough generalizations can be made.

Except in a few cases sub-soils are more variable than top soils. This is particularly true for static soil properties such as texture (Webster, 1981). Top soils are more variable than sub-soils only in some properties such as unsaturated hydraulic conductivity and diffusivity. The increase in variation with depth, for some properties such as soil texture, extractible and exchangeable cations, stoniness and solute

concentration is very systematic, (Dahiya, et al 1984). For some others variation with depth is more erratic, (e.g. hydraulic conductivity, silt, gravel).

Some typical examples are shown in this thesis in figures 3.4.(a) to (d), to illustrate in decreasing order a range of vertical variability that ocurrs and may be encountered in the field.



Figs. 3.4.(a)-(d). Observed variability of some soil properties with depth.

The vertical variation of some soil properties may be strongly correlated. For instance the decline in organic matter with depth is often accompanied by a similar change in cation

exchange capacity or in other properties that are related to biological activity.

For many practical purposes pedologists may want to be able to predict quantitively and accurately the values of any given soil property at any depth. The variation of a soil property down the soil profile can be described mathematically as a function of depth, (i.e. soil depth function), and the model used to interpolate the values wanted. Chapter 7 of this thesis discusses soil depth functions, and offers an algorithm to describe variation of a property with depth. The resulting curve can be used for interpolation and prediction.

3.3.2. Spatial Variability

The differences that occur in soil from place to place are of practical interest. Spatial variability of soils may be gradual or abrupt and its elucidation depends largely on the scale and intensity of observations. Soil spatial variability is, therefore, scale dependent.

The multiscale sources of variation are of various orders of magnitude. In general, differences in climate are associated with continental-scale soil differences. At increasing scales big tectonic formations and regional climatic changes within a major climatic zone are associated with soil changes in the order of tens of kilometres. Changes in soil within hundreds of metres or a few kilometres may occur mainly as a result of differences in rock type and local physiography. For instance, soils on river terraces contrasted against soils on hill slopes. At this scale, (i.e. changes taking place within distances ranging from a few tens of metres to a few kilometres),variability is of special practical importance for its impact on farming management and on engineering works and must receive special attention in land-use planning projects and advisory service. This is the scale and range of distances within which spatial soil variability has been most studied.

Within a given area, the variability of the soil is, therefore, the product of these effects, nested, and operating at such scales (Burrough, 1983a).

Two main components have been identified in the spatial variability of soil properties: a systematic component, and a random component, (Webster and Cuanalo, 1975, Webster and Burgess, 1983, Wilding and Drees, 1983, Burrough, 1983a).

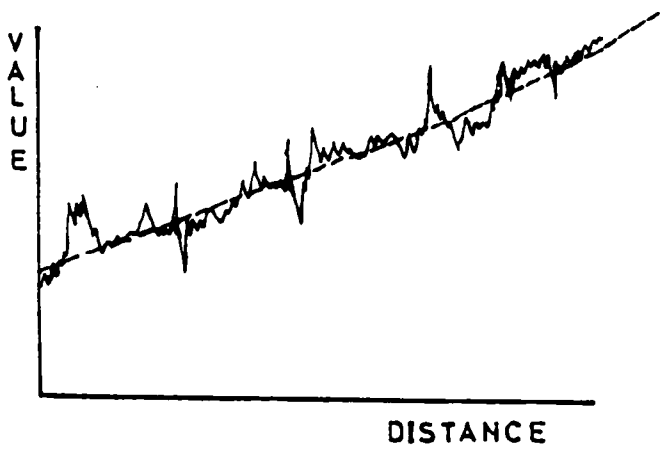
The systematic component presents itself generally as a long range gradual change in the form of trends at the scale of observation. When present at medium range it is also termed "drift" (Burgess and Webster, 1980).

The random component, often referred to as "noise", is the variability associated with the systematic variation which remains unexplained or unrelated to a known cause and cannot be discerned at the scale of observation. This noise can be substantial and reflect much point-to-point fluctuation of which only a small part may be attributable to measurement error, (Webster, 1983). The trend and the random component are illustrated in figure 3.5.(a).

Some local or medium range trends may recur with certain periodicity along distance, for example separate fields or clusters of fields which have been submitted to similar management and recur in tracts of usually a few hundred or metres. Alternated localised areas of contrasting parent materials may produce a periodical pattern too. These characteristic patterns are illustrated in figure 3.5(b).

Soil properties are characteristically "noisy" due to large random components, but on the other hand they may vary smoothly across the landscape. There is not sufficient conclusive data available yet to suggest general "laws" about the variability of different soil properties. The information available is still very much site-dependent. A block diagram of electrical conductivity of a Mexican soil is shown in figure 3.5.(c) to illustrate what could be said to be typical features of soil spatial variability.

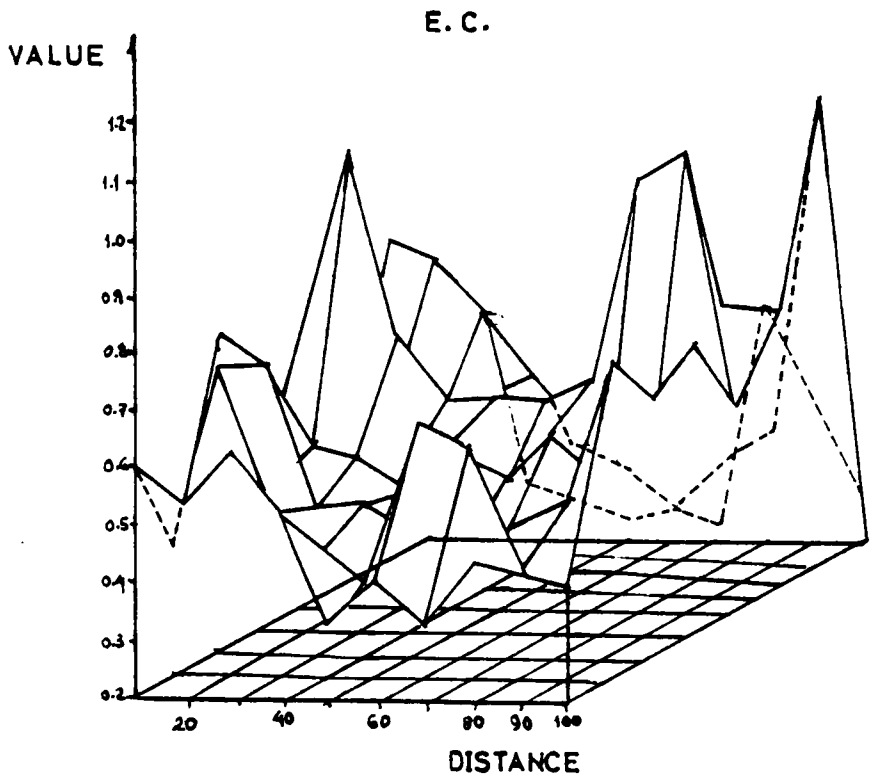
Trend, random noise, and possibly some degree of periodicity may be present at various scales of variation and may be encountered in the field depending on the scale at which the soil is observed.



(a)



(b)



(c)

Fig. 3.5 Characteristic components of the spatial variability of soils

(a) Random & Trend
(b) Periodicity
(c) An example from a Mexican soil.

In soil variability studies short, medium and long range variation are the terms of reference.

Short range variation usually appears in distances between a few metres to tens of metres, (or even few hundreds of metres). The medium range variation at which local trends (drifts) appear is present in distances between a few hundreds of metres, and the long range variation occurs at distances of a few hundreds metres up to several kilometres.

The soil variability measured in a field study then is a function of scale of observation and of the sampling intensity employed to obtain the data.

Clearly as the soil body is studied in greater detail, variation that may have been previously defined as random may be found to contain a systematic component. Thus medium or long-range variation that consists of a random component may find some systematic structure or trend if examined at a scale consistent with short-range variation.

The recent application of spatial analysis, based in regionalized variable theory (see chapter 7), to soil, has made it possible to elucidate more clearly the components of soil variability, (Webster, 1985). These are: (i) variation deriving from natural discontinuities. This is usually recognized and mapped during soil survey. (ii) a spatially dependent component that increases as separating distance increases, and (iii) variation present within the shortest sampling interval.

3.4 Methods for investigating the spatial variability of soils

Methods to investigate, understand and represent the variability of soils are useful in soil studies. The first method to be used was representational. Variability was depicted graphically in the form of crude maps. The soil scientist has refined this to the level of current, conventional soil survey. The statistical treatment of data extracted from soil samples has also been explored aided by the development of calculators and computers.

Recently, Geostatistics has helped to investigate and map soil variability and to use such knowledge in prediction.

3.4.1. Soil Maps

The production of soil maps from soil survey has, in effect, been a method to investigate soil variability by concentrating effort on systematic or abrupt changes and relating these soil distribution patterns to landforms and other landscape features. The conventional output from this procedure is a soil map its legend and a memoir. This has been the classical way of investigating and portraying soil

variability. It deals only with the type of variability which is apparent and can be recognised. Chapter 4 is devoted to discussing the various aspects of conventional soil survey and the use of its output in prediction.

3.4.2. Classical Statistics.

Quantitative statistical techniques have been applied to describe soil spatial variability. Firstly, to the analysis of field plot data in experimental designs. Later, in the evaluation of intra-class heterogeneity of soil classes and mapping units using some classical statistics parameters such as variances, coefficients of variance, intra-class correlation (Webster and Beckett, 1968), coefficients of relative variance, (Burrough, 1969), uniformity coefficients (Beckett and Burrough, 1971).

Most of the applications of classical statistics to soil variability in the univariate case rely upon analysis of variance. This implicity or explicitly assumes a certain probability distribution of the data (i.e. normal) and that the observations from the field are random and spatially independent (e.g. Beckett and Webster, 1979; Biggar, 1978; Cassel and Nelson, 1981; Wilding and Drees, 1983; and many others). The limitations of this classical approach are clear. More often than expected these assumptions could not be met.

It is quite possible that a normal distribution has been assumed in a number of situations, when more careful analysis of the data would have suggested otherwise.

Multivariate statistical methods have been applied to explore broad aspects of soil variability in space, considering more than one property at a time (Hole and Hironaka, 1960, Rayner, 1966, Webster and Cuanalo, 1975; Norris, 1970; Webster and Burrough, 1972a). The emphasis has been on reducing the dimensionality of the multivariate problem in soils and then exploring the pattern of joint variability of soil properties in few dimensions, and on exploiting the potentials of numerical classification and ordination on soil data.

However, multivariate statistics has the same shortcomings as univariate statistics concerning the assumptions of normality

and independence of the observations. Based on the evidence of recent work, the latter seems to be the most serious.

The limitations of the classical statistical techniques to describe soil variabilty are evident. Some alternative way of representing soil variation is desirable.

3.4.3. Geostatistical Techniques

In exploring spatial variability of soils, the separating distance between observations is of importance and needed to be included in the analysis. The first attempt consisted of fitting least-squares surfaces to soil data (trend surface analysis). This is a form of multiple regression of a soil property solely on the spatial co-ordinates as independent variables. However, as with any regression analysis technique a commitment to a particular model for the surface has to be made. Hence the fit at one end of the surface is affected by the other end with apparently very small or no relation at all. More importantly, there is no obvious functional relationship between a particular soil property and its geographic position or it is too complex to be represented by a workable mathematical function (Webster, 1985).

Early work by Webster (1973), and Webster and Cuanalo (1975) used auto-correlation functions (a sort of time-series analysis replacing time by distance), to examine the spatial dependence of soil properties. Using data collected along a transect the strength of association between pairs of observed values some lag distance apart was measured. The approach was found fairly useful in elucidating soil spatial variability.

Matheron (1965) introduced the theory of regionalised variables in the mining context to improve spatial analysis of ore bodies and local estimation of ore reserves. Campbell (1978); Burgess (1979); Burgess and Webster, (1980, a and b) introduced regionalised variable theory into soil science. Applied to soil this theory has provided a most fruitful line of inquiry and valuable tools for soil variability studies.

Regionalised variable theory builds on the semi-variogram which is a complementary function to the correlogram. The semi-variogram is a plot of semi-variance against lag distance.

The semi-variance is a measure of the similarity, on average, between points a given distance apart. The more alike the points the smaller the semi-variance.

The semi-variogram is an essential tool for spatial analysis and for optimal spatial interpolation. This is the most fruitful of its applications.

The semi-variogram is free of any assumptions related to probability distribution and independence for it is based purely on distance. The semi-variogram and the technique called Kriging as interpolator are the core of regionalised variable theory. This is considered as a breakthrough in soil variability studies and prediction (Webster, 1985). It is a fairly new field in soil science. There is a good amount of recent work on soil variability based on the semi-variogram and Kriging that indicates their potential for several applications in soils, (Webster, 1985)

Chapter 7 of this thesis deals with semi-variograms and kriging. Chapter 9 explores the efficiency of these techniques for predicting missing soil information at unvisited sites as compared with conventional soil classification and maps. It also attempts to examine the possible complementary use of the conventional and the new procedures for prediction.

3.5. Concluding remarks to this chapter

i. The soil is a rather heterogeneous mantle showing anisotropies in all directions.

ii. In general, soil variability as measured in terms of CV, is high for most soil properties.

iii. The sources of intrinsic soil variability are many, (i.e. rock, climate, topography, organisms), interacting apparently without a discernible pattern. In cultivated landscapes this variability is increased or smooth out by management.

iv. Soil properties show different rates of variability in space and in time.

v. In general, chemical properties tend to be more variable and less durable than physical and morphological properties.

vi. Variability increases with the size of the area sampled

but the variability in a field (e.g. 1 Ha), may be already present in a few m^2.

vii. Vertical variability is a function of depth. In general, subsoils are more variable than topsoils.

viii. The sources of variability operate at multiple spatial scales, so, the variability encountered in a particular area (i.e. short, medium or long range) depends on the scale and intensity of sampling.

ix. The spatial variability of soils occurs in terms of a systematic component (i.e. trend) and a random component (i.e. noise), but the pressence of either of them depends on the scale and intensity of sampling. Random variation may show structure by sampling more intensively.

x. From the methods for investigating spatial variability the geostatistical approach is, at present, the best. The semi-variogram is the best tool for this purpose. It is free of the assumptions of normality and independence of the observations that render classical statistics inadequate, and unlike soil maps, it accounts for both abrupt and transitional changes.

xi. It is against this background that predictions of values of soil properties at unvisited sites are required and need to be made.

IV. COLLECTION AND PROVISION OF SOIL INFORMATION: Conventional Soil Survey

Soil information is necessary for agricultural planning.

Conventional soil survey methods have been designed for the collection and provision of soil information.

This chapter reviews some relevant aspects of current soil survey methods and discusses the use of soil survey output to provide information about soils at sites of interest.

4.1. Aim and purpose of soil surveys.

Soil survey is the process whereby the variability of soils is explored, classes of soil are defined and described, data about its properties are recorded and their spatial distribution mapped.

Any soil survey attempts to provide information about the condition of the soil at any point within an area.

The aim of soil survey is to equip a user with the soil map and its legend and memoir, which will enable him to answer questions about the nature and distribution of soils with less trouble than he could have done without them, (Beckett and Burrough, 1971).

The questions that can be put are basically of the form:

i. What are the properties of the soil at any given site of interest in the area?

ii. Where soil with a given set of values of particular properties can be found in the area?

The aim of soil survey from this particular point of view is to provide means or tools (i.e. soil map, legend and memoir), for prediction.

Prediction is taken here in the sense that the survey based upon some observation sites, provides a means whereby the user or problem-solver can inform himself by the soil map, its legend and memoir, and predict the state of soil properties at any given site within the area without having to go there to observe.

This concept of prediction is important, and has been discussed broadly somewhere else (Webster and Beckett, 1968, Bie and Beckett, 1971a, Webster, 1977, Burgess, 1979, and Burgess and Webster, 1980a and b).

No soil map predicts soil conditions with absolute certainty, hence there is only one way in which the user can be totally confident about the values of soil properties at a site, and this is by going to the place to measure or observe them directly. Thus, the rate of success of a soil survey will be judged by the extent to which it accomplishes its aim, that is, the extent to which predictions about the soil properties at points of interest prove to be correct or fall within acceptable error bounds.

Every soil survey is carried out for a purpose, essentially of a practical nature.

The purposes of soil surveys can indeed be diverse. As diverse as are the possible land-related activites. The nature and the amount of soil information to be collected during the survey is related to the survey's purpose. This information is directed towards solving specific problems for which only a few soil properties may be relavant .

The surveyor should try to ascertain which soil properties are relevant so that they can be included in the list of those to be recorded during survey.

Any question then concerning the state of such properties in the area surveyed may be answered in terms of values of properties themselves or of soil classes defined from those values.

4.2. Kinds of soil survey

Soil surveys can be of specific or general purpose. Specific purpose soil surveys are undertaken to obtain only the necessary information for the solution to a particular practical problem (e.g. reclamation of soils affected by alkali or by excesive concentrations of soluble salts). The kind and number of soil properties recorded during a survey of this sort depends directly on the nature of the problem. The surveyor has to ascertain which properties are relevant to the solution of the problem, record them during field sampling, and form classes of soils and mapping units in terms only of the properties observed.

However, soil surveys are expensive. A substantial part of the total survey cost may be attributed to field work by professional and non-professional staff (i.e. recording observations and on transportation between observation sites; Bie and Beckett, 1970). Hence, once in the field, many soil properties relevant to a wide range of as yet unknown purposes, are recorded at each visited site in order to make maximum use of the survey effort available. Surveys of this kind are general purpose soil surveys. The resulting map and memoir serves to a broad range of users, essentially as basic soil inventory of the area concerned. General-purpose soil surveys are common practice in most survey organizations.

There are many different terms to refer to kinds of soil survey, which were created and are used by those who execute them. The terminologies used refer sometimes to a given range of scales. In other instances they use survey intensities to distinguish types of surveys, or survey purpose, or the nature of the mapping units and the information expected from the survey as criteria to categorize them.

Possible nomenclatures for soil survey have been suggested. Rourke, (1981) uses the concept of orders of soil survey in an attempt to bring consistency into the designation of surveys for different uses.

The nomenclature developed by FAO (1974), aims to standardize terms and is found the most adequate. This terminology is based on intensity of observations, and avoids the frequently misleading terminology of "detailed" , "semi-detailed", "reconnaissance".

The FAO terminology is shown in table 4.1.

It is possible to observe from table 4.1. that according to the range of scales, the "low intensity" and "exploratory" surverys both correspond to the "reconnaissance" type of survey. Similarly "semi-detailed", "detailed" and "intensive" correspond to "medium intensity", "high intensity" and "very high intensity" surveys of the FAO nomenclature.

Table 4.2. summarizes valuable additional information about other parameters like density of observations, scale, rate of progress, (FAO, 1974), related to the kinds of soil survey in the FAO system.

Table 4.1 FAO Terminology* of soil survey according to survey intensity in relation to survey scale and mapping units

Type of survey	Range of Scales	Kind of mapping unit
Very high intensity	Larger than 1:10,000	Phases of soil series; soil series occasionally soil complexes
High intensity	1:10,000 to 1:25,000	Phases of soil series: soil complexes
Medium intensity	1:25,000 to 1:100,000	Associations of soil series; physiographic units (enclosing identified soil series)
Low intensity	1:100,000 to 1:250,000	Associations of Great Soil Groups: occasionally individual Great Groups: phases of Great Groups. Alternatively, land units of various kinds enclosing identified Great Soil Groups
Exploratory	1:250,000 to 1:1,000,000	Land units of various types (preferably enclosing identified Great Soil Groups)
Syntheses	Smaller than 1:1,000,000	Great Soil Groups and phases of Great Groups (having essentially taxonomic significance)

* Source: FAO (1974)

Table 4.2 General Characteristics associated with different kinds of survey intensities*

Kinds of Survey	Scale	Area represented by 1 cm2 of map	Density of observation* (0.5 obs/cm2 of map)	Approx. average rate of progress (per 20 day month)	Accuracy of boundaries
Very high intensity	1:5,000	0.25 ha	1/0.5 ha	500 ha	Position of all boundaries checked throughout length on the ground
High intensity	1:20,000 1:25,000	4.0 ha 6.25 ha	1/8 ha 1/12.5 ha	1,250 ha 1,500 ha	Position of almost all boundaries checked throughout length on ground
Medium intensity	1:50,000	25.0 ha	1/50 ha	75 km2	Some boundary checking - most inferred
Low intensity	1:100,000	1 km2	1/2 km2	200 km2	Almost all boundaries inferred

* Source: FAO (1974)

The FAO terms can also be related to the purposes of the survey.

4.2.1. Very high intensity soil surveys

This type of survey is intended for the planning and implementaion of enterprises on small areas of perhaps a few hectares to a few hundred hectares, (e.g. highly intensive agricultural schemes, irrigation or engineering). Surveys of this type frequently are special-purpose surveys and consequently only a few properties are very intensively sampled.

4.2.2. High intensity soil surveys

Surveys of high intensity, (i.e "detailed"), are useful for determining the viability of specific projects in larger areas (e.g. irrigation, land reclamation, urban development), or for implementation of solutions to local problems, or for allocation of resources and agrotechnology extention.

They provide also information useful for decisions in the context of land evaluation and land-use planning. The mapping is carried out mainly by field sampling with the assistance of external ground features. This are general-purpose survey and the numerous soil properties recorded make it useful for a number of potential applications.

4.2.3. Medium intensity soil surveys

Medium intensity surveys (i.e. "semi-detailed"), are general purpose soil surveys most commonly executed to provide the necessary information for suitability assessments in land-use planning, for irrigation projects and for decision-making concerning broad aspects of land management within a region. They also provide the basis for extension and advisory work at this level. They are used as part of resource inventories in the assessment of potential for agricultural development. This kind of survey is the standard practice in the systematic surveys of developed and many developing countries. The mapping makes use of air-photointerpretation and external features to aid the location of boundaries which are based on a reasonable number of sampling points and field inspections.

4.2.4. Exploratory and low intensity soil surveys

This type of surveys corresponds to what has been known as "reconnaissance" survey. They are directed towards the identification of priority areas of high development potential.

They serve as resource inventories of large areas, and usually, but not always they preceed medium intensity or high intensity surveys in priority areas. Due to their scale, they make intensive use of air-photointerpretation supported by geomorphological surface features for the drawing of inferred boundaries.

Some maps from these surveys show individual classes of soil at a coarse level of an established national or international classification. Others show associations of individual classes of soil or in some instances just dominant soil types.

More detailed work should always be preceeded by exploratory and low intensity surveys to help in deciding on which most promising areas surveys of increasing intensity should be undertaken. Exploratory surveys also help in providing pre-survey information about the intricacy of the soil pattern and the structures of spatial variability.

The question about what is the most appropriate type of survey, for a particular purpose is the key to the solution of information-based problems related to soil. The choice of type of survey depends on the purpose which in turn determines the choice of sampling intensity and publication scale. The map scale must be adjusted so that it shows neatly the convolutions of boundary of the soil differences that are to be mapped.

In the planning context, survey intensity and mapping scale are subordinated to the purpose, that is, to the minimum area of management over which decisions will be made. Having this minimum area fixed, the map scale is a compromise between this and the minimum mapping area that can be shown, given the intricacy of the local soil pattern elucidated by the survey intensity that can be afforded. Table 4.2 provides information of this nature which helps to ascertain the kind of survey needed.

Clearly, the problem and the survey purpose must be carefully defined to avoid collecting an excess of expensive data without a clear idea as to their possible use. This is particularly true in the extremely detailed profile descriptions and chemical analysis that are product of routine surveys in many organizations.

4.3. Survey intensities

The usefulness of a soil map depends on the uniformity of significant soil properties described within a soil mapping unit (Webster and Beckett, 1968).

The correctness of soil boundaries and the precision of the delineation of the mapping units depend to a great extent on the density of point data, (i.e. the survey intensity).

The density of observations in a soil survey may be taken as a measure of its quality.

Survey intensity is defined in terms of the density of field inspections per unit area of published map. The ratio between field scale to publication scale relates this to the number of observations in the area surveyed (Bie and Beckett, 1970).

Survey intensity is determined by the survey purpose and by the publication scale suitable for such purpose.

The surveyor attempts to ascertain the necessary effort, i.e. density of observations to resolve the local soil pattern and to produce a satisfactory map for the survey objective.

Survey intensity must increase with the publication scale. Hence, detailed surveys should rely on a high density of observations.

There are considerable discrepancies between organizations about the specifications and rules of thumb on the optimal number of field observations per each cm^2 of published map, (Bie and Beckett, 1970). There are also few justifications for universally applicable specifications because of the need to adjust map scale and density of observations to the complexity of the local soil variability.

However, the figures offered in table 4.2, by FAO provide general guidance for surveys of various intensities.

In general, the closer the boundaries to be mapped, the greater the density of observations needed to locate them. The breadth of class and the intricacy of the soil pattern determine the density of soil boundaries. This in its turn determines the density of observations required in field survey to delineate them.

Information about the intricacy of the local soil variation and so an indication of the boundaries required for a given map scale can be obtained from pre-survey reconnaissance. Those two elements provide an estimate of the density of observations, the survey effort and consequently a gross estimate of the survey cost.

In practice, the "optimal" survey intensity often represents a compromise between the survey cost (i.e. the survey effort that can be afforded to achieve a certain boundary density), and the usefulness of the resulting map for predicting soil conditions.

In every landscape there is a certain intensity of survey effort beyond which the cost of further precision increases sharply and out of proportion to the level of precision already achieved, Beckett, (1981).

Thus there is a ceiling imposed by cost, and a bottom limit imposed by the minimum map precision acceptable. Between the two a compromise is usually reached.

Effort increases with the density of boundaries to be mapped, and may be reduced if air-photointerpretation (API) is used. The contribution of API decreases as map scale increases.

The intricacy of soil pattern is reported to account for 50-80 per cent of the variability of survey effort (Bie and Beckett, 1971a). This evidence emphasizes the importance of obtaining pre-survey estimates of the intricacy of soil pattern and of the structures of soil spatial variability. The semi-variogram (see chapter 7) is a useful reconnaissance tool for exploring the spatial structure of the soil pattern. It is also particularly suitable for the design of optimal sampling schemes.

4.4. Map Scale

The scale of a published soil map is often taken as an approximate indication of the amount of truth it shows.

The scale of a published map is often not the same as the scale of the base map used for field mapping.

Beckett (1981), indicates that the field scale should be about twice the publication scale.

The minimum scale for a soil map is the ratio between the smallest area distinguishable on the map to the smallest area of

soil that has been delineated during survey. The maximum scale is simply a practical limit to the area of a map sheet. There are obviously cartographic limits to the minimum size of a mapping unit that can be printed in a soil map. Bie and Beckett (1971,b), suggest that the minimum acceptable map scale is the ratio between the minimum feasible spacing between soil boundaries on the map and the minimum distance on the ground between the soil boundaries that are to be mapped. Clearly the scale of a soil map must increase as the minimum area to be distinguished decreases. The latter usually varies with the intensity of the land use.

Map scale should be subordinated to the area it portrays as well as to the density of the information on which the map is based. Often the map scale is determined by relating it to the density of observations. Vink (1963) suggests that scale should be adjusted to give four observations per cm^2.
The Soil Survey Staff (1951) prefers five observations per cm^2.
On the whole, independent organizations tend to have their own rules of thumb, Bie and Beckett (1970). Table 4.2 shows scales linked to particular observation densities, (FAO, 1974).

Map scale is an important factor in determining the required survey effort for a given purpose and also influences survey costs.

Survey effort (man-days/km^2) increases with the map scale, (Bie and Beckett, 1970).

Thus many factors affect the choice of the most suitable map scale. Among the most important in practical terms are the intricacy of the pattern of soils to be mapped, the effort (i.e. density of observations required to resolve and map such pattern, but, most importantly, the precision of the statements and predictions that can be made from the information provided by the survey that are related to the survey purpose. Some of these factors are in contradiction. The choice of scale represents therefore a compromise between them.

4.5. Field procedures

The way in which the surveyor proceeds to choose where to inspect in the field the soil he wants to map, is also a significant factor in the survey cost and efficiency.

Given a number of field observations that can be afforded the surveyor will attempt to locate them at places where he would expect each observation to provide as much information as possible about the structure of the soil pattern he is trying to map.

The surveyor will decide whether he uses intuition and knowledge of the local soil variability to place his observations, resulting in an irregularly distributed sampling pattern. This is free survey (Steur 1961, Buringh et al 1962). On the other hand, he may decide to place the sampling sites according to a regular pattern on a grid with uniform coverage over the area. This is grid survey.

The majority of observations in free survey are made to verify soil boundaries and the rest to check the soil within the boundaries that are being mapped. Human intuition and experience are aided by external features of the landscape to place soil boundaries. Free surveys then tend to be used when there is some external expression to soil boundaries (e.g. landform limits, changes in the land-use pattern). Air-photographs in this instance are an excellent tool to aid the placing of sampling sites and boundaries. The main stages in free survey (Burrough, et al 1971) are discussed in a following section.

Among the main advantages of free survey is that each observation may give maximum information about the soil around it if it is left in the hands of an experienced surveyor. It enables him to take maximum advantage of air-photointerpretation and ground-surface evidence to create in his mind a sort of soil-landscape model. Once such a model proves to be correct it can greatly reduce the number of sites to be inspected and consequently the survey effort.

Free survey method is less effective in featureless landscapes (e.g. extense alluvial flat areas) or in areas where access is difficult and navigation poor (e.g. tropical rain forest).

Grid Survey consists of regularly spaced observations along straight traverses which are laid parallel and regularly spaced to create a rectangular or triangular grid over the area. The observations are located at the traverses intercepts.

Grid survey provides a uniform coverage over the survey area. This is important for it makes grid surveys particularly suitable for a wide range of statistical analysis.

The decision about grid spacing is crucial as it must match both the expected precision of the soil map, and the complexity and intricacy of the local pattern of soils. Grid spacing is of importance too in elucidating the spatial variability of soil, and the design of optimal sampling schemes for mathematical interpolation by kriging (see chapter 7). The grid spacing is a way of choosing the scale of the structures of soil variability that the surveyor wants to elucidate by means of the semi-variogram (see chapter 9).

Grid surveys are particularly suitable for large-scale, high intensity surveys in which the ground evidence for soil boundaries is limited and the usefulness of air-photointerpretation reduced because of the scale, or due to other circumstances like the lack of adequate supporting material (i.e. base maps and air-photographs), or difficult access, or in situations where the forest canopy may obscure the ground evidence of soil boundaries. Under such circumstances grid survey represents a better alternative to free survey.

Controlled location survey (Western, 1978), is a compromise between grid and free surveys taking the form of a loose grid where the placing of observation sites is controlled but flexible. It is preferred in situations where a minimum density per unit area is wanted, and also where relatively uniform coverage of the area in a roughly rectangular distribution of sites is desirable to reduce the subjective element in mapping.

Burrough et al, (1971) found that, in general, free survey required 20-30% less effort (man-hrs/km^2) than grid survey. That is, for a given effort free survey could produce a map at a scale 20-30% larger.

Beckett and Burrough, (1971) and Burrough et al, (1971) compared the costs and the utility of grid and free surveys of the same areas. They found that at comparable observation density the mapping of series by free survey is cheaper on average than by grid survey, but an assistant can produce a soil series map by grid survey more cheaply than a scientific officer can by free survey. Thus their relative advantages in terms of costs depend on the kind of personnel to perform the survey, and on the ratio of salary between professional and non-professional staff.

In regards to the utility of the maps they concluded that the proportions of different series were better estimated from grid survey than from series maps by free survey. However, where soil boundaries had a clear external expression free survey was superior to grid survey. Moreover, on average the "purity" of maps by grid survey was higher than for maps by free survey at comparable survey effort.

At scales greater than 1:25,000 single-property maps achieve more uniformity of their property than general-purpose series maps. The latter seem to do better at medium scales.

The choice of survey procedure then depends on the purpose of the survey, budget, quality of the information required from the survey and the nature of the soil variability, its external expression and access. The final decision involves a compromise between those factors.

4.6. Field records and sampling

The location of the observation sites is determined by the field procedures. Measurements of soil properties of interest at the observations sites and samples for laboratory determinations are necessary to generate the soil data.

The term "sampling" as used here refers to the description of units of observation, to the recording of data from such descriptions and to the collection of soil samples for further analysis in the laboratory.

Conventionally, soil profiles are the units of observation on the three-dimensional soil landscape. However, particularly while locating soil boundaries, auger bores are the units of observation for checking the lateral extent of a given soil profile. In practice, more than 50% of the field observations during survey are auger borings made to locate soil boundaries.

The soil profile dimensions are to some extent arbitrary but usually it consists of a pit of about 1.5m wide and 2.5 to 3.0 m long and generally 2.0 m deep, unless rock or any other kind of impediment including water table is encountered while digging.

The soil profile exposes the vertical variation of the soil along depth. Such variation is stratified on the basis of a few

evident soil properties in the field. The layers or horizons defined are more or less horizontal and assumed to be relatively uniform.

When the survey is general-purpose the surveyor attempts to record as many soil properties as possible. The division into horizons is helped by the differences shown by soil properties of clear external expression easy to see and detect by hand on the profile's face, e.g. by differences in particle-size distribution, colour, soil structure, organic matter content, soil consistency, stoniness, drainage, moisture content, and some others.

The definition of horizons in the field inevitably carries some human intuition. Horizons are more or less homogeneous in the properties above and it is implicitly assumed that the rest of the soil properties will be equaly homogeneous.

In a specific purpose soil survey, the surveyor will place more emphasis on the relevant properties when dividing the profile into horizons. Particle size and mottling for instance would be important for a study on soil drainage.

The profile description is the very substance of the survey. Here, the needed information about soils is captured and recorded. Check-lists of selected soil features according to the survey purpose can be prepared beforehand. However, whenever possible the fullest descriptions are made, particularly in surveys of general purpose.

Guidelines and field manuals for site and profile descriptions are usually employed to achieve compatibility and uniformity in the collected data.

Descriptions must be fully informative, but for some purposes full descriptions may not be needed. The amount of information to be recorded is a matter of survey purpose and cost.

Horizon depth and thickness of each horizon are measured and recorded first. Morphological properties are the most frequently recorded in the field, notably colour (including mottling and its characteristics), organic matter, soil texture, structure, consistency, stoniness, soil moisture, roots, soil flora and fauna, and inclusions of pedogenic origin such as concretions, nodules, soft concentrations, cutans, pans, rock substrata and calcium carbonate and dolomite concentrations.

Having described the soil profile and recorded the field observations, laboratory analysis of chemical and physical properties complete the profile's data-set.

A bulked sample supposed to represent the soil horizon from which it was collected yields one value which is assumed to apply to the whole of the horizon. The original continuous depth function is transformed into the "stepped" depth function (see chapter 7) by this way of sampling. This is the simplest way of describing the changes in one soil property with depth. However, for some practical purposes interpolation of values at any soil depth may be required.

It can be shown that depth interpolation is feasible and convenient by means of procedures which reconstruct the original continuous depth function from stepped profile data from this way of sampling. These procedures are examined in chapter 7 of this thesis.

The methods for recording site and profile soil data have undergone some changes in the last few years to accomodate computer-oriented methods for data capture and input into information systems. Soil data may be recorded in the field to facilitate direct input into the computer. Ad-hoc computer-compatible field proformas are commonly used in soil information systems and other methods are being explored to reduce the steps between the field work and the computer (Bie and Schelling, 1975, Hodgson, 1978).

The advent of computers into soil science brought awareness of the enormous potential for automation that many of the soil survey tasks posses. The computer can perform faster and in some instances better many of the laborious tasks that are part of the conventional procedures in soil survey. These automated procedures when integrated become soil information systems. Chapter 5 of this thesis deals with computerized soil information systems. Most of the relevant aspects and stages in the information input into the computer, handling and output are discussed. Data capture and data input are important initial stages in a soil information system.

4.7. Soil classification

Soil classification is an essential operation in soil survey. Classification is the act of dividing a population, or the grouping of individuals or objects, i.e. soils, on the basis of their common properties.

Most of the information about soils is conveyed to the user via the soil class which in most survey reports are the only link between the information the user wants and the map. Thus, soil information is conveyed in terms of profile classes and mapping units.

Soils are classified for various reasons but essentially because classification facilitates the following operations.

a. **Communication.-** Information transmision about soils is faster in terms of class names than in terms of soil properties.

b. **Organization.-** Structured information enables speed in information retrieval and recall, improving efficiency in its use and better understanding of relationships between individuals.

c. **Generalization.-** Fairly large amounts of data from many soil descriptions can be reduced to manageable terms by generalizing the information enabling efficient communication.

d. **Prediction.-** The role of the surveyor is to provide a means (i.e soil map, legend and memoir), with which a user can inform himself and predict the soil conditions without having to go to the field to observe. In short, a user is expected to be able to predict soil conditions at a given place by identifying the soil class and using information from that class.

This thesis concentrates on the use of soil classes as a means for predicting soil properties at unvisited sites.

Any system of classification is useful if it defines groups, i.e. classes, about which a good number of important statements can be made concerning the properties of the individuals grouped, to serve best the purpose of the classification. Hence, classes must be relevant for the particular problems of management or planning that the survey is to help solve. In doing this the first necessary step after problem definition is to ascertain what are the soil properties most closely linked to the solution or understanding of the problem or the purpose intended.

A single system of classification will rarely serve two purposes equally well. Ad-hoc classifications are single-purpose and are usually locally defined, in contrast to the so called "natural" or taxonomic classifications which are general-purpose, and usually of national scope.

In soil survey practice the surveyor may use either of two alternative systems of classification depending on the purpose. The decision is a matter of a suitable level of generalization. So, in attempting to classify the soils of an area there are two alternative ways of action, either to rely on empirical evidence about the local population of soils and design a classification to accomodate the local soil variation and the purposes of the survey, or else to adopt an already designed classification system and proceed to identify and allocate the soils inspected into the classes conceived by such a system. The former is an "ascending" method where reasoning is from facts to concepts, and the latter is a "descending" method, where the higher levels in the hierarchy are conceived in terms of hypothesis and theoretical principles of pedogenic nature. The system is further elaborated as observation proceeds, to accomodate lower categories into an already existing hierarchical structure.

There is a bridging problem between the rather conceptual higher categories and the lower ones, based on what is found directly in the field. As a result this may cause misallocations of soil individuals into given categories.

Butler, (1980) proposes the recognition of "taxonomic hiatus" between any general classification system and the local classes of any local soil survey classification. This hiatus will also apply between one soil survey project and a neighbouring one. This hiatus presents itself in terms of lack of matching between the differentiating properties or differentiating of local classes and those of the general classification.

A natural classification is one with classes phylogenetically determined, Cuanalo, (1972). Natural systems are structured hierarchically as galleries of conceptual classes at various levels of generalization. Some properties are chosen as the basis for grouping at each level. These are the differentiating characteristics or identification keys.

Natural or taxonomic systems are constructed to serve the largest number of purposes and therefore they are expected to produce broad classes so the properties used as identification keys may often not be associated to properties relevant to the problem that the soil survey attempts to solve.

If a locally constructed classification is to be used the differentiating characteristics must be chosen to produce classes with the greatest predictive power for other and preferably useful properties. The properties chosen for the key must produce groups which reveal the degree of orderliness or structure present in the population. Orderliness generally is reflected in well defined clusters of points in a scatter diagram, where the attributes used for the key are axes. So it appears to be more convenient to select differentiating characteristics locally as opposed to acquire them from nation-wide or international classifications. The locally developed key is more likely to result in the classes having greater predictive power.

In contrast, natural classifications use in their elaborate hierarchies, differentiating characteristics that are morphological and resulting from soil genesis. Genetic criteria are often considered too propositional to be used as a basis for prediction. Moreover, elaborate hierarchies are inappropriate in soil classification as they rely on too many asumptions on the orderliness of the soil population without rigourous analysis of the data before subdivision in classes, (Webster, 1968). Hence the resulting sub-classes are unlikely to improve predictions.

Ideally, a good local classification must be taken from the landscape i.e. there must be some degree of agreement betwen the limits of the classes and the landscape units that the surveyor sees. In addition, emphasis must be placed on the selection of the differentiating characteristics which is crucial in obtaining a classification with good predictive value to serve the survey purpose.

The main drawback of a local classification is the possibility of a taxonomic hiatus when attempting to fit the local classes into the upper levels of generalization of a wider national or international classification in order to improve communication and correlation with other systems. The discordance or break between the general classification and the local classes

is essentially due to mismatch between the differentiating characteristics of the two systems.

The higher in the hierarchy of conceptual classes a soil entity is identified, the greater is the degree of generalization and so is the breadth of the class, which becomes more inclusive. Prediction of soil data using very inclusive classes is obviously less accurate than from narrow and exclusive ones.

In this thesis locally constructed classifications are seen as being more useful for prediction purposes than general or nation-wide systems. In the former, their classes are less inclusive, generalization occurs to a lower degree, and the properties for the key are good discriminants fitted to the conditions and particularities of the local area.

The surveyor constructs a local classification in the field by inspecting the soil at contrasting places to discover how presumed differences correspond to differences in the profile's face. A broad perspective of the soils is acquired by augering or digging pits. Mental correlations are established between classes of profile and units of landscape. Once a comprehensive overall picture of the soil variation and its relation to external features on the landscape is complete, sites of the proper survey sampling are located. At this point the surveyor is able to propose some classes, his gallery of conceptua1 soil classes forms a provisional classification. This is subject to modifications if classes not recognised before are found during the field sampling. The surveyor's classes usually correspond to soil series. A soil series is a collection of profiles with a similar or comparable sequence of horizons. The soil series is the basic unit of general purpose classification. This collection of profiles is made around a modal profile with individuals (i.e. profiles), tied to it by bonds of varying strength. The class as a whole can be defined in terms of values of differentiating and accessories characteristics, and deviations from those values within the range of the class, Cline (1963). The differentiating characteristics are necessary and sufficient conditions for memnbership to a class. The range or breadth of the class is stablished a priori by convenience. It is an indication of the level of precision with which predictions of soil data are required and this is determined pre-eminently by the survey purpose and the nature of

the problem it attempts to solve. The class breadth is to some extent, subject to compromise between the local soil variability, the external features of the landscape, the survey purpose and the budget.

Classification involves information generalization. Loss of information from the orignal data collected for each sampling site occurs due to generalizations. The amount of information lost when "hard" field data collected at a site gets transformed into general statemets concerning only ranges of values of the class can be considerable and is quantifiable.

The point of information loss due to generalization is a very important one and deserves attention, for data from soil survey are expensive. It may be pointless to devote effort to obtain detailed observations and measurements if part of this information is lost by generalization. Generalization was inevitable for it is a pre-computer means to cope with multiple observations in large data sets. The computer presents alternative ways for storing and handling efficiently non-generalized or "hard" soil data produced by survey. Computerized soil information systems offer capabilities for storing and manipulating "hard" soil data and producing the same conventional soil survey output but without losing the original "hard" data. Moreover, the computer also enables the on-line use of the non-generalized soil data for prediction by means of mathematical interpolation. Chapter 5 of this thesis examines the main aspects of soil information systems and Chapter 7 reviews the techniques for mathematical interpolation.

Class means of several soil properties are used as predictors of such properties at unvisited sites in the trials described in chapter 9. A soil classification locally designed in terms of ad-hoc differentiating criteria, was used to define the soil classes in the area of Yelford (see chapter 9).

So, the surveyor classifies the soil in the field in three main ways, Webster (1977). First he divides the profile into horizons, then into classes of soil profiles, and finally into tracks of country or areas in which one or more profile classes occur, these are the mapping units.

4.8. Soil mapping

Soil classification and soil mapping are the central and most important part of soil survey. Any soil map has a purpose. In principle, the purpose has been identified in the survey planning stage. According to that purpose and the problem it is to help solve, the surveyor aims to provide a tool for prediction. The final result of survey is a map of simple or compound mapping units accompanied by a legend which indicates the soil classes to be expected in each mapping unit and explains any symbols or inscriptions.

The purpose of the survey determines the breadth of the soil classes into which the soil population is divided. This is an indication of the precision wanted in the prediction of soil information by means of the soil classes.

Simple mapping units are defined by drawing soil boundaries so that the soil enclosed by the boundary corresponds to only one of the soil classes.

In many instances this is not possible due to the intricacy of the soil pattern which may be irresolvable for two or more soil classes so that they will probably be mapped together as a complex or compound mapping unit. This may be so, particularly when the differences between them are of little significance for the purpose of the survey.

The boundaries drawn on the map are a simplification of the tortuosity of those real boundaries in nature. So, the area covered by a mapping unit will not be completely pure but will contain inliers of other classes. This is true even if the surveyor locates the boundaries very carefully and uses repeated auger borings.

The purity of a mapping unit is the fraction of a number of inspections in the area covered by a mapping unit, where the soil class as stated by the legend of the map corresponds to that mapping unit.

In general, there are poor relationships between soil series profile classes and the mapping units which are supposedly representing them. The levels of purity are in some instances as low as 40%.

Hence, it is important to distinguish soil series profile classes from soil series mapping units as two different concepts, the former are soil-profile entities and the later are soil-area entities. So, the grouping of similar soil entities (wherever they occur), into a soil class is a different operation from the grouping of adjacent soil entities into a mapping unit.

The scale of the soil map is an important aspect. It is determined by various factors. Firstly, the class width is fixed in accordance to the survey purpose. The class width determines the frequency of the boundaries that are to be located. It may be necessary to seek a compromise between the class width and the intricacy of the soil pattern. From this compromise the surveyor is able to work out a likely minimum boundary spacing in the field. From this he is able to determine the appropriate map scale that would show neatly the minor convolutions of soil pattern, and calculate the density of the observations required and the survey procedure to follow.

The density of observations must increase with the publication scale. There is little point in drawing broad maps at large scale if they are based on only a few observations. If the soil map is to be detailed, a high density of observations and a large map scale are required to be able to show the convolutions of the soil boundaries clearly.

The mapping is carried out essentially in two alternative ways.

i. Finding areas within which one or other of a combination of profile classes from a provisional legend predominates. These become provisonal mapping units subject to further inspection by auger and modified if necessary.

ii. Resolving first the soil-landscape model directly into provisonal mapping units.

External features of the landscape are used for the drawing of boundaries to separate areas of one kind from areas of a different kind regardless of the procedure chosen. So, landscape boundaries can be used if it can be proved that they also divide classes of soils.

Clearly the association between soil-landscape boundaries and soil-class limits is too important to go untested. The advantage of this way of mapping is that once the soil-landscape/soil-class

model in the mind of the surveyor proves to be correct it greatly reduces survey effort and facilitates the mapping. Conversely, an imperfect or partial association between landscape and soil boundaries will greatly reduce the accuracy of his soil map.

Air-photointerpretation at this stage may be of considerable help to see patterns of soil-landscape. The map legend is prepared to link the classification to the landscape. Mapping units are usually described in terms of their constituent profile classes except for miscellaneous landscape units such as "badlands" and "shallow/rocky lands".

The field mapping is carried out by identifying profile classes by auger or spade, usually working along loose traverses and looking for changes in external features that can be shown to be associated with soil boundaries. Possible soil boundaries have to be verified by making some soil checks at both sides of it. If confirmed, the soil boundary is extended by either repeated observations or by looking for external features known to be associated with it.

Confronted with a boundary which lacks clear external expression, the surveyor may decide either to spend extra time and effort on more observations to locate all the boundary or to embrace the two adjacent soils into a compound mapping unit. Additional observations may be made to confirm the uniformity of the soil within a simple mapping unit.

Sometimes, constraints on map scale or resources force the surveyor to generalize two or more distinguishable mapping units into larger compound mapping units made of associations of series or associations of series-phases.

The final product is the soil map, accompanied by the legend. The legend records the soil classes to be expected in each mapping unit and gives information about symbols and other inscriptions. It is the key to the map and helps a user to identify the soils in the field.

4.9. Use of soil classes, map units and memoir in prediction

Soil surveys provide the soil information necessary to assess the suitability of the land for alternative uses. The soil is observed and sampled in a relatively small number of places in

the area of interest. Hence, in order to provide information for unvisited sites, values of soil properties at those locations need to be predicted.

For a given site X, or a particular land-use Y, the questions regarding the assessment of its suitability can be of any of the two following forms:

a. What can be done/built/grown at site X ?

b. Where can "Y" be done/built/grown ?

The answer to questions of this sort requires raw soil data and also an interpretive link between the data and the land-use.

In terms of soil data these questions may take the following elementary form.

a. What is the value of soil property Z at site X ?

b. Where is the soil with a value V in property Z ?

In effect, the answer to the latter question (b) demands that question (a) must be answered first for every site in the survey area. So both (a) and (b) may be simply subsumed to the standard question

i. What is the value/state of property Z at site X ?

The answer to this question requires an exercise in prediction. Conventional soil survey offers some tools for this. The soil map, its legend and the corresponding memoir are the means by which the user can predict the information wanted.

The first step is to locate as accurately as possible the position of site X on the map. Its location will rarely correspond exactly to that of a site with recorded information, which has been visited in the field. Next is the identification of the symbol of the delineation within which site X is. The map legend is examined to discover the map unit enclosed by such delineation. There may be more than one class in a map unit, or some inclusions of other classes that are taken as impurities. In any event, the soil class at site X has to be predicted.

The successful prediction of the soil class at a site of interest determines to a large extent the accuracy with which soil survey output can be used for predicting soil data.

Mapping units vary in kind and therefore in content. Hence, the prediction of the soil class depends on the type of mapping unit where the site X is located. The range of possibilities are:

(a). A simple mapping unit containing one dominant soil class and a few minor soil classes as impurities that may be stated in the legend.

(b). A compound mapping unit which includes two or more significant soil classes mixed into an unresolvably intricate pattern. This is a soil "complex". The unit is described in the legend only in terms of the proportions of the soil classes it contains.

(c). A compound mapping unit, but unlike (b), the pattern of occurrence of the soil classes is resolved and is described in their association with external landscape features so that the spatial occurrence of individual classes can be traced. Units of this sort are soil associations.

Simple mapping units, as in (a), provide grounds for a higher rate of success in the prediction of the soil class. In a soil association, as in (c), the rate of success is lower and depends on the ability of the map user to trace the occurrence of the corresponding class at X on the basis of class/landscape associations. The soil complex, as in (b), is the more problematic. The soil class at X can only be predicted in probabilistic terms drawing from the proportions of each class within the map unit as stated in the legend and memoir.

The breadth of class influences also the success in class prediction. In general, narrow classes are more difficult to predict than broad classes, but as the latter become more inclusive their predictive power of soil properties decreases.

The predicted class leads to the source of the data stored in the memoir. The utility of such information in prediction of property values is only limited by the terms and the amount of information i.e. field observations, used in its definition.

Clearly then, the success of soil survey output in predicting soil information at a given site depends on two essential and equally important factors: One is the capability of the map to predict the soil class at the wanted site within a map unit. The other, once the soil class is known, is the extent to which raw data have been generalized into soil classes. This determines the range of variation (i.e. the precision) of the information held in that class.

The generalization of information that occurs during classification and mapping represents the main disadvantage in the use of soil survey output for accurate prediction of properties. Raw point-data from the field gets generalized in two stages. The first is through classification of soil profiles. The values of soil properties get transformed into ranges for the class. The second stage of generalization occurs while mapping. The pressence of inclusions of other classes in the form of impurities in a mapping unit amounts to a form of further information generalization. This decreases the probability of success in prediction.

There are several items of information from soil survey output that can be used in a combination of forms to predict the value of property Z at site X. The mean value of a class, values from the typical profile of the class, or the mean value of a map unit all can be used as predictors. The approaches followed in this thesis to predict soil properties make use of those items. They are as follows:

4.9.1. Use of class mean of the soil property as predictor

The procedure consists of the following sequence

i. Locate site X on the soil map.
ii. Predict the soil class at X.
iii. Extract from the memoir the profile data available pertaining to the same class as site X, (i.e. means or raw data).
iv. Calculate the mean of the class of property Z from such profile data.
v. Apply class mean of Z at site X.

The predictive success of this approach is a function of the accuracy in the prediction of the soil class at X, the breadth of the class, and the density of point-data (i.e. the number of observed profiles within the class to calculate the mean), used in survey to define the class. Narrow classes based on sufficient observed profiles would, on the whole, yield better estimates of Z at X than broad classes defined on few profiles.

4.9.2. Use of typical profile of the class to provide the value as predictor.

This approach may be considered as a variant of the previous procedure in that it uses a central parameter too, i.e. the mode instead of the mean, but the value used as predictor is commonly from only one profile considered the typical of the class, i.e. modal profile. So, points i and ii of the procedure are as in 4.9.1, and the rest of the steps are as follows:

iii. Select the soil profile most typical of the class.
iv. Extract the value of Z from the typical profile as stored in the memoir.
v. Apply the value of Z to the soil at site X.

The predictive success of this approach depends strongly on the accuracy with which the class at X was predicted, on the breadth of the class, and on the choice of the profile thought as the typical or modal of the class.

4.9.3. Use of the mapping unit mean as predictor.

This approach can be said to be similar to 4.9.1 above, but the number of profiles used in calculating the mean is expanded to include profiles within the map unit which inevitably carries some impurities from other classes. So, the approach could be expected to yield poorer estimates than 4.9.1, unless the mapping unit possesses a greater degree of uniformity than the class. On the other hand, prediction of the class is not necessary for it suffices that a profile is located within the mapping unit for it to be included in the calculation. The steps are as follows:

i. Locate X on the soil map.
ii. Record the mapping unit symbol containing X.
iii. Extract profile data within the same mapping unit as stored in the memoir, i.e. means or raw data.
iv. Calculate the mean of property Z from such profile data.
v. Apply the mean value to the soil in X.

The main limitation of these three approaches appears to be the availability of raw point-data as reported in the memoir of

the survey. For these ideal procedures to be applicable there must be some point-data in the memoir. Further, to make good predictions, there must be a sufficient number of within-class or within-map-unit point-data to obtain good estimates of the class or map unit mean. Often the memoir presents means, not raw data.

Ideally the three approaches calculate their predictors by going back to ungeneralized information, i.e. the raw point-data. In effect, they perform a sort of spatial interpolation at site X on the raw point-data. These approaches exploit some measure of similarity existent between the soil at X and the point-data available to draw from the latter an estimate of the value of Z at X. Thus, the similarity link between site X and the raw point-data is provided by their pertaining to the same class or map unit. So, provided the soil at X and some point-data available belong to the same class or mapping unit, prediction can take place. Often, however, the memoir only gives the mean values.

In effect, the predicted value can be visualized as a form of local average from values within a finite neighbourhood (i.e. the soil class or the mapping unit). In this sense prediction by conventional soil survey output is comparable to spatial interpolation by kriging (see chapter 7), which is a sort of weighted moving average on nearest neighbours, where the weights and the neighbourhood are defined by the degree of spatial dependence between site X and its surrounding points. In the case of class or mapping unit means the weights are equal for all points within the delineation and the neighbourhoods are defined by the breadth of the class or by the mapping unit boundaries on the map.

From the three approaches, 4.9.1 appears to be the best, provided the soil class at the site has been predicted correctly and there is a good estimate of the mean. Since the mean is the natural expectation of any value, therefore means provide the best estimates of the value of Z at site X.

4.10. Concluding remarks to this chapter

i. The aim of soil survey is to provide means for prediction of the state of soil properties at any site within an

area.

ii. The success of a soil survey must be judged by the extent to which these predictions prove to be correct or fall within acceptable error bounds.

iii. A clear definition of the problem linked to the survey purpose, and knowledge about the intricacy of the local soil pattern are key to the adequate choice of survey type. Hence, to the definition of breadth of soil classes, scale and sampling intensity. These, in turn, are reflected in the survey cost.

iv. The density of observations in a soil survey may be taken as an indirect measure of its quality.

v. Every landscape has a limiting survey intensity beyond which the cost of further gains in precision over the one aleady achieved, increases sharply and out of proportion.

vi. Reconnaissance survey is fundamental in determining the intricacy of the local soil pattern. Hence, in designing the best survey strategy for a given cost.

vii. Soil information is conveyed to the user in terms of soil profile classes and mapping units.

viii. Classification and soil maps are inefficient systems for storage of information. They involve information generalization. A proportion of the information collected during field survey, (which is expensive), is lost due to generalization. So, there may be little point in devoting effort to obtain detailed observations and measurements in the field if this information will be generalized.

ix. The products of soil survey, i.e. soil classification, legend and map, are fixed. Once decisions have been made about them, and data generalized, they are "fossilised" in the result.

x. The computer offers an alternative to conventional soil survey for storing "hard" point-data and producing the same conventional soil survey output more efficiently.

xi. The success of soil classes and mapping units in predicting the state of soil properties at unvisited sites may be improved by mathematical interpolation of soil data. The predictive success of the conventional and the new methods needs to be compared and their complementary use explored.

V. SOIL INFORMATION SYSTEMS

Computerized soil information systems (SIS) provide an alternative to the procedures of conventional soil survey for storing manipulating and providing the soil information needed for land-use planning.

The development and implementation of SIS following the expansion of computer capabilities and information technology has posed various fundamental problems to soil science challenging the conventional and somewhat rigid concepts and procedures. So, developing soil information systems presupposes radical changes in the methodology of research and in the entire approach to the soil science problem (Ciric, 1976).

5.1. Information Technology and Soil Information Systems

Until fairly recently the handling of quantitative and descriptive soil information lagged considerably behind the technological advances in microelectronics and communications.

Information transmission enabling communication has only experienced three major changes. These have been of radical nature and profound consequences; i.e. the development of written language, the discovering of arithmetic operations enabling performance of repeatable operations with quantative data, and the invention of the press, some 500 years ago. With the press, information took mainly the form of characters and letters on paper, and a single message could then reach a wider audience. The very fast developments in technology, particularly in the fields of microelectronics in the last 40 years brought the advent of computers. The computer is basically a machine for the automatic storage and processing of information in electromagnetic form. This is perhaps one of its most important features, for information stored in such form can be recalled as many times as necessary, bearing on it the burden of repetitiveness, a limitation of printed information.

Another important feature of the computer is its ability to choose one of a number of different sequences of instructions, depending on the results it has calculated previously, so that it can perform a list of operations; i.e. a computer programme. A collection of computer programmes is called in computer jargon "software". To function in this way a computer requires some elements, namely, a store to hold the instructions and any data or operands, a processing unit which is a control unit to ensure instructions are obtained from store correctly and in proper sequence order, and a set of input-output peripherals to communicate with the machine. These are the basic components of the "hardware".

Information in a digital computer must be represented in a digital form. Analogue communication systems have been superce ded by digital communication systems. These handle information in digital form as does the computer. Thus, the convergence of these two technologies, i.e. what is known as Information Technology, has made available the provision of a low-cost computing element, that is the microprocessor. It has also enabled computers to make use of advanced electronic communications. Thus, due to its electromagnetic nature digital information can be stored, retrieved and communicated in enormous quantities and at phenomenal speeds.

Digital information in the computer can also be selected, edited, rearranged and transformed, activities which were, until fairly recently, the sole province of the human brain.

Information is the meaning assigned to data by the known conventions used in data representation, (Dumanski and Kloosterman, 1978). The process of extracting some meaning from data can be said to provide information.

A soil information system is a collection of activities and procedures involving systematic, geographically referenced soil data collection, storage, retrieval and manipulation (Dumanski and Kloosterman, 1978). SIS is an orderly method of procedure imparting to others instructive knowledge of particular facts or occurences pertaining to soil (Moore, 1978). Information systems in soil science are not new. What is relatively new is the use of computers as support for providing soil information.

In essence, SIS is derived from the application of computers to the storage, processing, display and presentation of soil data.

A combination of factors has promoted the development of SIS:

i). The realization that soil maps are a rather inefficient system for storage and retrieval of soil information. Giltrap (1981), has estimated that a point in a map retains at most 6 to 7 bits of information from the not less than 100 bits of information in each of the original ungeneralized point-data as observed in the field. This represents only around 10% of the total information collected during survey.

ii). The enormous amount of data that soil surveys generate and that needs to be handled, which invites automation so that computer based systems are becoming increasingly essential i.e. the computer as a notebook or data bank.

iii). The availability in recent years of the necessary and cost-accessible computer hardware and software to execute the conventional tasks of routine production of soil survey output, i.e. the computer as a clerk.

iv). More importantly perhaps is the great flexibility in the use of data which is possible in SIS. This is demonstrated by the literally unlimited re-usage that can be made of primary data, for example in producing interpretive ad-hoc classifications and maps from reappraisals of such data. The on-line access to built-in applications programs for data processing adds even greater flexibility to SIS.

The development of SIS appears to have occurred in three stages:

1st Stage.- The computer was first used as calculator to execute lengthy operations as part of statistical and numerical techniques applied to soil characterization and soil variability studies, e.g. Bidwell and Hole, (1964), Russell and Moore, (1968), Norris, (1972), Webster, (1971). The use of a computer is particularly associated with its capability to store large volumes of data and to perform long and even complex operations at high speeds.

2nd Stage.- Computers are used to support the laborious tasks in the production of conventional soil survey output,

e.g. report writing, automated classification and mapping of soils from the application of cartographic procedures and devices to soil. The solution to the problems of data base, data base management systems and data acquisition represent the major thrust in SIS development at this stage. The computerized handling of soil data broadened from simple manipulation of small and specific purpose data-sets and files to the large and multiple file integrated systems.

3rd Stage.- The solution to the spatial problem in soil data represents the main thrust on the development of SIS at this stage. The capability offered by SIS to reapprise "hard" point data to obtain ad-hoc interpretive classifications is also exploited. New techniques for spatial analysis and interpolation of data, (see chapter 7), which represent a breakthrough in soil science for dealing with the spatial problem, become standard algorithms incorporated into SIS bringing about a different approach to the mapping of soils and to the apprissals of soil data for interpretive classifications, (Giltrap, 1984).

5.2. Basic structure of SIS.

Physically, a soil information system has two basic components; namely, software and hardware. However, functionally SIS is made up of a number of distinctive components or modules which define the necessary operations that have to be sequentially performed on raw soil data. These are, essentially, the different stages or modules of the internal data-flow in the system.

All SIS have the modules input, retrieval and manipulation in common. According to Beckett et al (1972), data banks for soils and/or related data present data collection and input, data storage and retrieval, and applications, including display as standard procedures.

In general, then, six components of SIS can be recognized. These components are:

a). **Data acquisition**

b). **Input**

c). **Query**

d). **Retrieval**

e). **Manipulation**

f). **Presentation and output**

The six components listed above are briefly discussed next, individually.

5.2.1. Data Acquisition

This operation consists of the transformation of the observed features and attributes of the soil mantle into abstract alphanumeric and character codes guided by established conventions and units. The source and nature of the data jointly with the degree of sophistication of the SIS determine the kind of mechanism that is used for the acquisition. For field point data, acquisition by fill-in or tick-off proformas have began replacing the conventional long textual descriptions in the field notebook. Some devices such as specially fitted tape recorders, calculators or even portable remote terminals are in use for data acquisition in some organizations. For map data, digitizers and raster methods are also used, particularly in SIS with capabilities to handle map data.

Data acquisition is an important operation for it reflects the way in which the real world is perceived and abstracted as data. It has a significant effect on subsequent operations but mainly on the modelling of data in the computer in terms of data base configuration. Some data acquisition procedures of existing SIS are reviewed and briefly discussed later in this chapter.

5.2.2. Input

The input procedure consists of the necessary operations to get the raw data into the computer checking that the data format matches the internal data base configuration designed for the system.

So, on input the data are validated to achieve a high degree of uniformity in features, codes, definitions, formats, and on indexing of records and data file organization ensuring that the data are ready for storage.

One fundamental problem in the development of SIS is encountered during data input, that is, the problem of data compatibility. Lack of compatibility in data betwen contributors and SIS restricts the scope and range of applicability of SIS.

A SIS is said to be flexible when it can accept soil data coming from different individuals and organizations in a variety of formats, codes and definitions and with other idiosyncracies.

Data compatibility has been the concern of many workers, e.g. Moore, (1978), Dumanski and Kloosterman (1978), Riquier and Chidley (1978), Cormack and Moore (1981), Lynch (1981), Craze (1981), McDonald (1981), and Van Berkel and Eyles, (1981), and is a problem of primary importance for it is related to other subsidiary problems posed by the development of SIS. Data compatibility is discussed a little further at the end of this chapter.

The methods for data input vary according to the data sources, Lynch (1977). For field recorded data input is by punch cards or by keyboard to tape or disk. Areal or map data are input by digitizer or by grid cell methods.

Digital satellite data are input via tape-to-tape or tape-to-disc transfer operations with previous susbtantial validation.

The utility of a number of input devices has also been explored. Optical character readers, tape recorders, portable terminals by telephone line transfer to an off-line receiver or directly into the computer.

Input in SIS is the more time-consuming operation and is probably more expensive than any manual system, (Lynch, 1977).

Validation and editing are considered to be an important part of the input procedure although sometimes they are considered as separate modules.

Once in the computer the storage of data and the data-file organization is the province of data base and data base management systems.

5.2.3. Query

Once data are in SIS, the system has to be queried to retrieve information. The function of the Query procedure is to allow the user to ascertain what sorts of data and information the SIS contains. During query a range of questions posed to SIS by the user in relation to the various aspects of the data which are of relevance to him are answered. A "user-friendly" approach to query procedure is achieved through a man-machine interactive dialogue.

A suitable query language to query and update the data base is the most important feature of SIS concerning the scientist, Mackenzie and Smith (1977).

There are three major query language levels. At a high level a user could frame his query in natural language. Queries can also be expressed concisely and mathematically using set theory and logic, e.g. Poseidon, the system at Orstom (Van Der Driessche et al, 1978). At the third level language the model of data structures in the data base is used to guide the query.

Most SIS operate the Query procedure at the third level language. Nevertheless, the user will have to find out anyway what information SIS has to offer before data can be retrieved. This is independent of whether it is done fully through the vehicle of SIS documentation or in an interactive man-machine session, through a hierarchy of menus and questionaires. However, at the end of his query the user should know whether SIS contains information on his interest and if so, which data sets are relevant to him.

5.2.4. Retrieval

The retrieval procedure consists of the necessary operations to extract from the data base the information relevant to the users query.

The complexity of the retrieval procedure varies with the data base (DB) and with the nature of the data base management system (DBMS). These in turn are consequence of SIS design. Two alternative approaches to DB and DBMS in SIS can be used for retrieval operations.

i. Use of commercially avaible DB and DBMS.

ii. In-site development of software based on ad-hoc retrieval algorithms.

There are some trade-offs in either choice. The former relieves the designer of the burden of much programming but commercially available DB and DBMS are expensive and may be too loose and general for ad-hoc retrieval purposes. The latter is easier to control and cheaper but it is rigid and the range of its applicability in information retrieval is narrow.

The efficiency of retrieval is improved by a proper and accurate framing of the query and by an efficient retrieval algorithm. The correctness with which data structures are modelled in the data base and the volume of data in it determine also retrieval efficiency. Proper indexing and selection of identifiers for the data are the key to avoid unnecessary data redundancies and to keep searches through the data base files to a minimum improving efficiency. Thus, most factors that affect retrieval efficiency are somewhat subordinated to the adequate modelling of the data in the data base.

Each data base model carries intrinsically certain advantages, shortcomings and costs inherent to information retrieval.

The modelling of land resources data for computer storage and retrieval as related to DB and DBMS structures, and its importance in SIS has been discussed at length, e.g. Cormack and Moore (1981), Mackenzie and Smith (1977), Smith (1981), Alagic et al (1984), Bluhm et al (1984). This is the sole province of the fast advancing field of DB and DBMS and it is not of the particular concern of this thesis.

On the other hand, in-site developed retrieval rules may also be ad-hoc and purpose orientated, e.g. retrieval algorithms predicting the suitability of the land for various uses (Cook, 1977), or algorithms that search for areas of land satisfying certain criteria.

Having retrieved successfully the relevant data, the user may want to produce output of the raw data.

Files containing the retrieved data are sent directly to the presentation procedure without further processing. On the

other hand, retrieved data may be further processed in the SIS manipulation procedure.

5.2.5. Manipulation

The manipulation module comprises any operation performed on retrieved data before their presentation and output. Data may be manipulated a number of times as necessary to extract meaning and the wanted information.

In this thesis, Manipulation is the package of applications programs for data-handling to operate on retrieved data.

The variety of built-in applications programs in the manipulation module provides SIS with capabilities for on-line data processing, accomplishing tasks that were carried out separately before SIS, demanding much time and effort. This makes SIS a flexible and powerful tool.

The number of possible applications programs of various degrees of sophistication that can be fitted into SIS, is only limited by SIS design, management decisions and the nature of the data itself.

The capabilities define the alternative uses that can be made of SIS for a number of purposes.

In principle, the manipulation module should be designed to be as flexible as possible to allow for expansion and growth in any direction. New facilities and interfaces could be added as necessary to increase the capabilities of the system and to meet the demands on it.

The manipulation package in most SIS is developed with emphasis on a particular orientation. Some SIS are more statistically orientated than others. In other SIS the emphasis is on cartographic output. However, on the whole, the possible facilities that a SIS manipulation package can offer, to date, can be grouped into five categories. Existing systems offer one or more of these kinds of facilities, namely:

i. **Report writing on retrieved data.**

Descriptive statistics, tabulations and graphical output (i.e. the computer as a "notebook").

ii. Data Processing.

Statistical analysis and all other mathematical treatment of data (i.e. the computer as a powerful calculator).

iii. Production of conventional soil survey output.

Output on the memoir of soil survey, soil classification and soil and interpretive maps, cartographic facilities (i.e. the computer as a "clerk").

iv. Interfacing other types of land resources data.

Bridging packages to interface climate data-files, land-use data-files, remote sensor digital data-files or other external files (i.e. the computer as a communications link).

v. New procedures for spatial analysis and interpolation of soil data for the production of soil and interpretive maps.

Spatial analysis, kriging and other spatial interpolation routines and their application in soil and interpretive mapping.

A multipurpose system would include most, if not all of the facilities above indicated, covering a wide spectrum of applications. Figure 5.1 illustrates a very comprehensive manipulation module. Most existing SIS offer only part of these facilities. The complete multipurpose SIS has not been built up yet. The manipulation capabilities of existing SIS respond to more somewhat narrowly defined purposes.

The development of the manipulation package for a soil information system at the Soil Science Laboratory, Oxford (OXSIS), is an important part of this thesis. The OXSIS manipulation package consists of a set of built-in applications programs and interfaces to external statistical packages. Its philosophy, development and capabilities are fully described in the next chapter of this thesis.

5.2.6. Presentation

This module comprises all the necesary operations to transform output from retrieval or manipulation procedures into an easy-to-read format.

Retrieved data from a SIS can be either reported directly to the user on request or else used as input for the manipulation module. In either case, the data may not be found in a format from which users can derive interpretations.

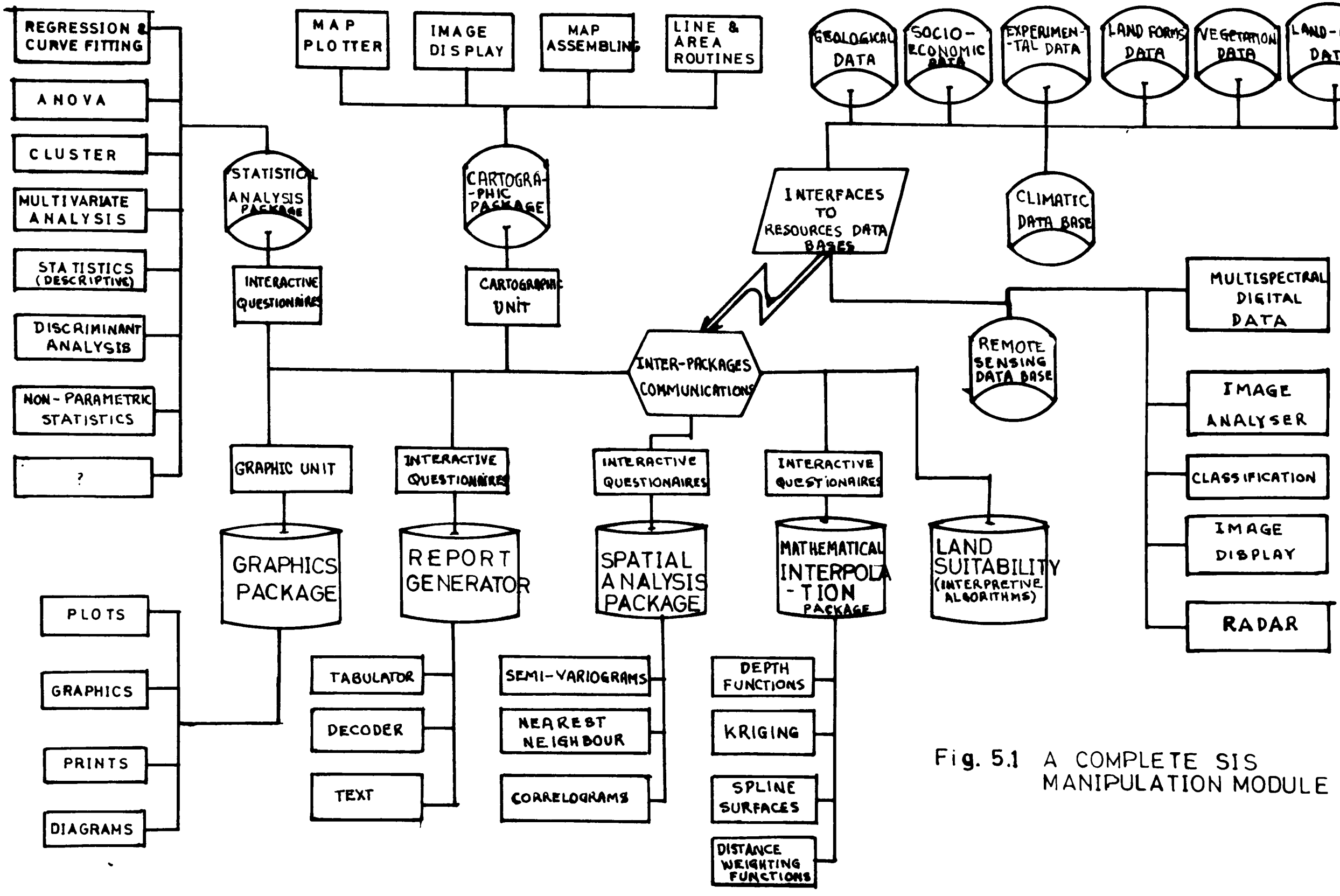


Fig. 5.1 A COMPLETE SIS MANIPULATION MODULE

The presentation package is then needed for transforming the data formats output from the different operations and packages to a readable and easy to understand output format.

The complexity of the presentation module depends on the variety of output formats from the manipulation and retrieval modules and on the variety of the available output devices. Line printers, visual display units (VDU), graph plotters and mappers are the usual output devices.

The modules briefly reviewed so far constitute the distinguishable components of a SIS. They are illustrated graphically in figure 5.2.

5.3 Review of Existing SIS

Work on developing SIS has taken national and international interests. A review of the components and capabilities of systems already in operation or in development is presented here.

SIS have been set up to solve a variety of different problems in different environments from various countries. Consequently, the approaches to input, retrieval, manipulation and presentation differ substantially too.

The creation of SIS took place first in some industrialized countries. France and The Netherlands pioneered their development in Europe and Canada and the U.S.in America. Developments occurred almost in parallel in Australia. In Oxford, early work by Beckett et al (1972) indicated some relevant aspects of the creation of data banks for the storage of terrain information.

Soil information systems have evolved into a very complex field. This complexity is the response to the varied nature of the problems which SIS have attempted to solve. Differences between existing SIS are found in aim and purposes, scopes, scales of operations and in the nature of the data they contain. Consequently, their sizes, components, design and capabilities are quite different too. SIS have grown unorderly in attention to local needs, resources and local

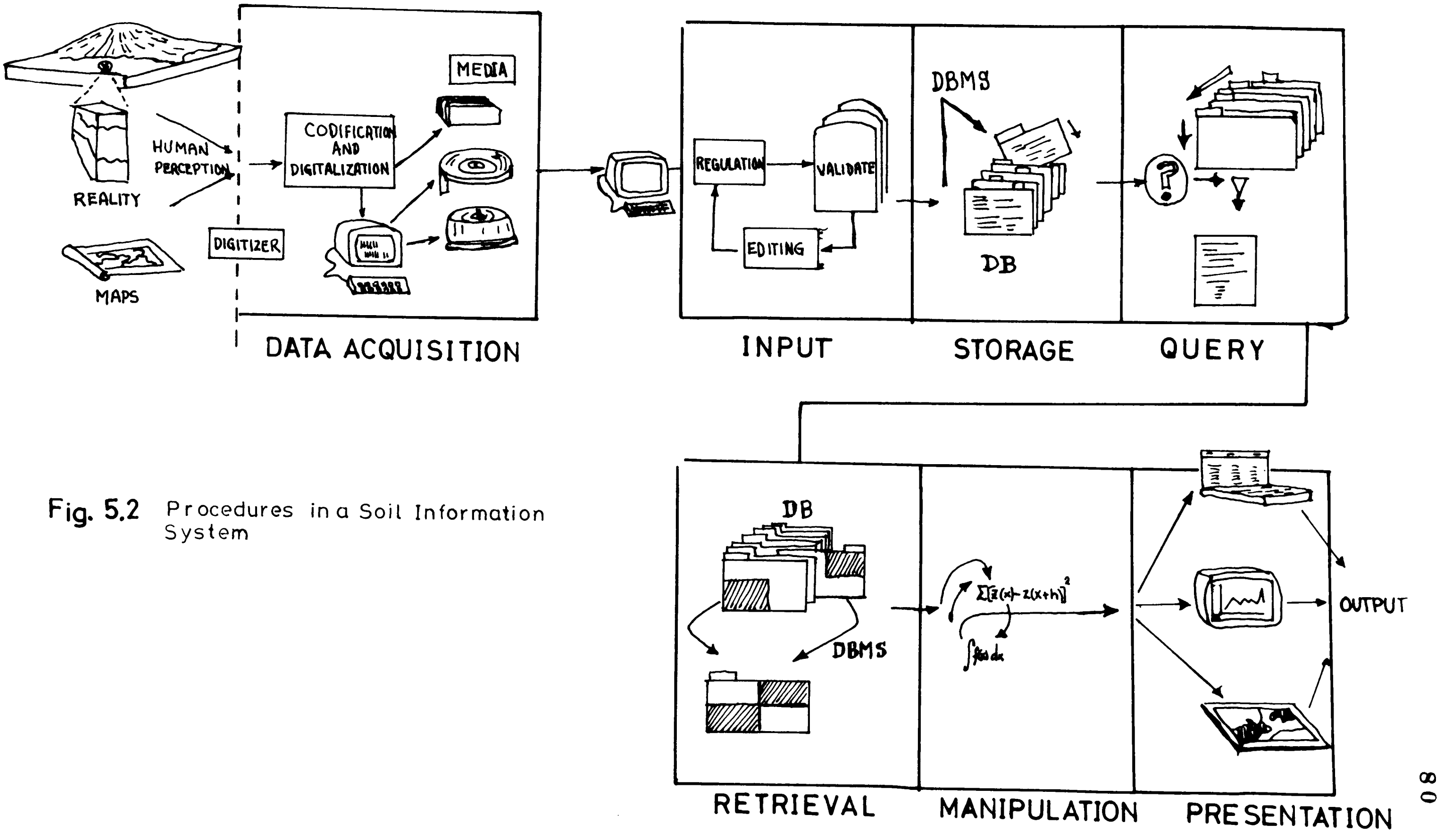


Fig. 5.2 Procedures in a Soil Information System

ideosyncracies. There is no reason why this should not have happened.

In designing SIS, identifying the actual and potential users has remained one of the key problems. The design and capabilities of SIS are determined by the questions it is to answer. Where the questions that can be anticipated have been guessed wrongly, the result has been "an information system which is looking for users", Moore, (1978).

Ideally, the potential users and their questions should have been clearly identified first. This would have indicated the capabilities that the SIS needed in order to answer such questions. Such capabilities would in turn point to the nature of the data to be held, to the dimensions of the system, and to the scale at which it is to operate. This information decides the modular components of SIS and the specific hardware requirements within the budget constraints. Figure 5.3 illustrates the ideal sequence of activities in designing a SIS.

Often SIS are built up around data which have been already collected, or around a particular set of hardware capabilities already available, or else with a given DB and DBMS already in mind. The first case is the most common of the three. However, soil data are expensive and so the designer of SIS must exploit to the full existing sets of data and use the hardware and software available to him.

In the final section of this chapter an outline of the structure and capabilities of the soil information system at Oxford (OXSIS) is given. OXSIS is the case of a system built up, having data already collected but with a flexibility such that it accepts data from any source and in any format. Its purpose is to exploit to the full existing sets of data.

Existing SIS around the world range from the simple-task orientated systems which hold soil data only from a particular locality and serve only narrowly defined purposes, to very complex and inclusive systems intended to provide information at a national or international scale and with sophisticated statistical and cartographic capabilities. Such systems may include not only soil data but other land resources data. Some of these even attempt to integrate digital satellite information of the earth's crust with soil data. Some others interface

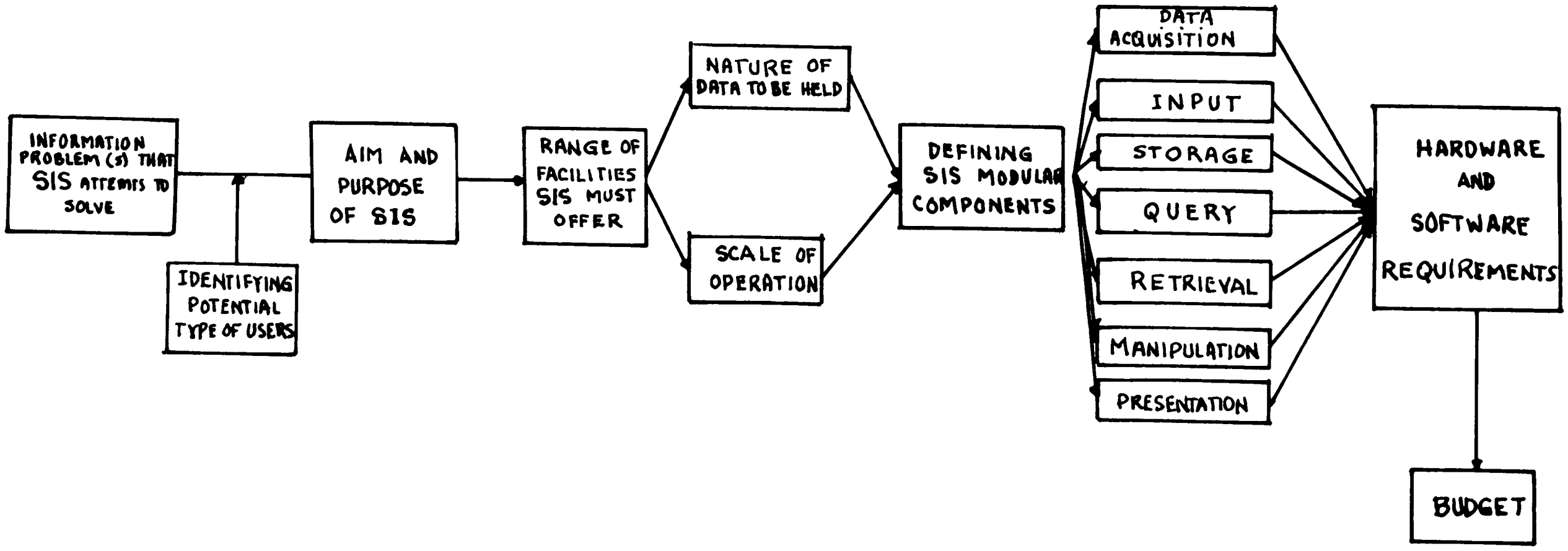


Fig. 5.3. Ideal sequence for designing SIS

external data bases holding other environmental information such as climate, vegetation, land-use, bringing them together into one general system for crop monitoring or for land evaluation and land-use planning.

Among the simplest of the SIS are those developed to concentrate only on a specific and narrow range of soil properties. Sleeman, (1977), presents a system for soil fabric analysis. A Rumanian data-bank (Canarache et al, 1978), is a system for soil moisture properties. A data bank specialized in clay minerals in soils is in operation in Bulgaria, (Behar and Evtimov, 1978). In Australia a system specializes in commercial soil analysis, (Price, 1977).

These and other similar systems serve to illustrate one, and the simplest use of the computer, as a "notebook". The machine is only used as a data bank, primarily for simple information retrieval. Data banks are basically set up for storing acquired data for its future retrieval and use.

In other places the capabilities of SIS are stretched a bit further beyond those of a simple data bank, to provide descriptive statistics of the data they hold or even some sort of statistical analysis of a certain degree of sophistication. The system developed at Orstom, France, known by the acronym of Poseidon, (Van Der Driessche, 1978), and the Yugoslav system, (Ciric and Palc, 1978), are examples of this type. From these relatively simple systems to the more complex ones, SIS become more and more inclusive in terms of types of data and hardware and software capabilities reflected in a wider range of facilities available.

The Netherlands Integrated Information System For Soil Survey (Bie and Schelling, 1978), was one of the first systems designed to put together data bank, data processing, and cartographic facilities, by allowing storage and retrieval of point, line, area and time information in on-line and off-line modes. The system consists of a set of already existing modules obtained from different sources and put together in a building-block design. It is claimed that its modular design makes it largely portable.

The system can be used in soil suitability classifications, (Bie et al, 1976). This system is an example of an attempt to

build a national information system for environmental sciences by assembling modules which represent the up-to-date knowledge in Europe and North America, in a building-block approach to the design. These modules were purchased and put together by in-house developed interfaces. This approach to building a SIS is the most expensive. In contrast, CanSIS, the Canadian Soil Information System (Dumanski, Kloosterman and Brandon, 1975) is a collection of co-operative soil data systems, national and provincial, linked together through a hierarchical orgnization of files and codes co-ordinated at the national level. Regional or provincial data systems are made up of pertinent proportions of the national system, and by agreement, all are compatible. This hierarchial structure on the basis of files records and modules appears to be the most efficient approach to designing a national system. The underlying objective of CanSIS is soil resource interpretation, correlation and parametric land evaluation. CanSIS seems to be one of the most successful existing soil information systems.

Its hierarchical design, flexibility and expandability appears to have been carefully planned. This system has served as an example for development work elsewhere, (e.g. Craze, 1981).

The latter two systems use the computer in three ways; as data storage and retrieval, as a powerful calculator in data processing, and as a "clerk" in the production of soil and interpretive maps. There is no use in these systems of the new procedures for spatial interpolation of data and production of maps based on them.

The Danish integrated data-processing system of the Bureau of Land Data, (Platou, 1984) is used for handling soil and other natural resource data in parametric or line-area form.

Geographic information and digitized maps are stored in the line-area unit. The system is intended for use in regional agricultural, water resource and land-use planning. It does not produce final interpretive maps. Its aim is to produce low cost maps sp ecially for planning work.

The Bulgarian system (Garbouchev et al, 1978) is among the most complete. Its capabilities have been extended to interface data from meteorological observations and other environmental

data. Differentiated soil and climatic regions are the data sources. The system aims at crop monitoring based on soil information. The data bases include soil, climatic, biological, agro-technical, socio-economic and experimental data. The continuous growth of such a system has caused it to evolve into a so called integrated computerized agriculture management system. The aim of such a system is to predict yields of the main crops grown in the country and provide information for making the technical decisions to obtain such yields, (Sadovski, 1983). Being such a system in operation and standing for the claims that are made about it, then, systems like this represent the most advanced stage in the development of SIS towards integration for decisions and policy making in agriculture. Systems of this sort can be developed for decision support in the form of an "expert system" for land evaluation and land-use planning.

The map image display generator (MIDGE) of the Soil Bureau of New Zealand, (Giltrap, 1984), demonstrates that self-contained low-cost systems which incorporate spatial interpolation techniques for the production of single-factor, or interpretive soil maps are feasable and obtainable. MIDGE is an exceptionally interesting system in many respects. It is implemented in a low-cost microcomputer which accesses maps stored in floppy discs in a run-length coded grid format. The maps may be displayed in a colour VDU at any scale. The map production is achieved by mathematically interpolating the observed soil properties from random sites to "target" points on an arbitrarily fine grid. Giltrap (1981), has shown that the technique called kriging can be applied routinely as interpolator at a relatively low cost for a comprehensive range of soil properties from survey. From reappraisals of point data, users can create their own interpretive classifications, and generate the resulting soil map by spatial interpolation. Maps may be overlayed to produce composites for interpretations. Data from digitized soil maps may also be used.

The MIDGE system is typical of the third stage in the SIS development. It offers a solution to three problems:

a. The hardware problem.- By demonstrating that low-cost and yet powerful microcomputer-based soil information systems are possible.

b. The spatial problem.- By incorporating "built-in", the new techniques for spatial interpolation of soil properties.

c. The problem of reappraising or reclassifying sites ad-hoc for a specific purpose.

When the solutions to problems b and c are combined in a system the result is a very powerful tool to assist decision-making by land planners and farmers.

Differences between the SIS around the world are also due to differences on scale of operation. Existing SIS are designed to operate at three levels; regional, national, and international.

Some regional systems are demonstrated by Vevers (1978), Pachepski (1978), Scott (1980), and Emery (1980). Others are planned to operate nationally, e.g. CanSIS, (Dumanski et al 1975), the Dutch system, (Bie and Schelling 1978), the Bulgarian system, (Garbouchev et al 1978), and the Danish system, (Platou 1984). Systems like the FAO international soil data bank, (Riquier and Chidley, 1978) represent an effort to create a truly international data bank. However, the problem of data compatibility at this scale may be great. Fundamentally, the problem consists of having very few national systems considering the possible exchange of data between different banks.

The problem of data compatibility is a general and important problem in SIS at all scales of operation, and is preventing the exchange of valuable soil data and information between nations, institutions and even between individuals.

Differences in entities recorded, codes, definitions and units reflect the institutional approaches and ideosyncracies, Moore (1978).

FAO, (Riquier and Chidley, 1978) has taken some steps to tackle the data compatibility problem at international level. Bayane and Legros, (1978) in France, have also aimed to solve the problems related to data exchange. In their soil data processing international network project (SDPIN) they propose the definition of an international soil data transfer format

(FITSOL). Tests on FITSOL have been carried out between France, Belgium and Canada. FAO has already proposed guidelines for coding soil data in an attempt to achieve standardization. They purport to create a very flexible system with two formats (free and fixed) for data input, glossary of terms and to standardize soil data coding.

The position towards data compatibility problems adopted by many SIS has been in many instances, to impose rigorous data formats for entry. To some extent such is the case of the FAO system. So far, the bewildering array of field proformas and similar forms of data capture, frighten many would-be contributors. This also constrains the SIS to be rigid and of limited scope. Much valuable and expensive data are left out due to incompatibility. Also some information is excluded when records have to be "squeezed" into rigid formats with which they are not wholly compatible.

The "flexible" soil information system at Oxford (OXSIS), operates locally and has been designed to tackle data compatibility problems. It can accept soil data about any features and in any format, as any data is better than none. OXSIS design and capabilities are outlined at the end of this chapter.

Another important source of differences between existing systems is the nature of the data they hold. Many of the SIS that are already in operation are not strictly soil systems. They are part of the land resources information systems family. Although some of them are based on soil information, they are designed to store other land resources data, i.e. geology, climate, vegetation, hydrology, landforms, e.g. the South Australian environmental data banks, (Scott, 1981). In such systems, soil information is integrated to other data within the same structural model in the data base.

Some other systems are typically SIS but have the capability of interfacing other kinds of data, e.g. Garbouchev (1978).

There are also systems designed to store the large volumes of remotely sensed digital information of the earth's surface. These are not properly SIS. However, attempts are being made to interface soil and digital information from satellite (McCloy, 1981).

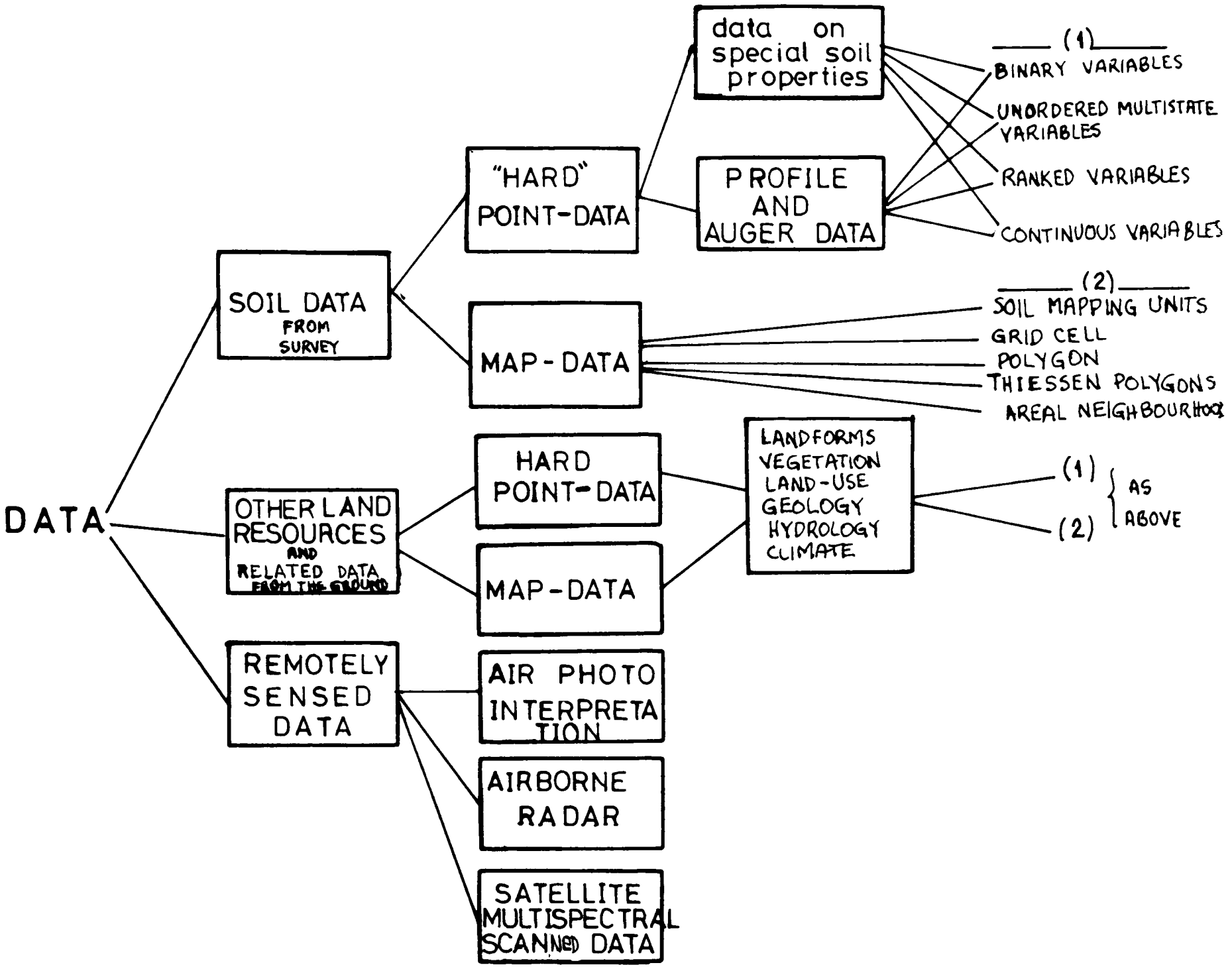


Fig. 5.4 Sources and forms of data for SIS

Remotely sensed multiband digital data can be also useful in the SIS for monitoring the changes in soil surface, crop growth, soil salinity and soil moisture conditions. Venkataratnam, (1984), suggests a SIS model for India incorporating remotely sensed data for such purposes.

On the whole the sources and forms of data that existing SIS and related systems contain can be summarized graphically in figure 5.4.

A summary of the features of existing SIS around the world, and their capabilities is provided in table 5.1. This summary is not totally inclusive. There must be various SIS in operation or under development for which reports have not been published yet or which have escaped this review. However, table 5.1 attempts to put those reported all together under the same headings so that they can be compared, and trends in SIS development analyzed.

The data from table 5.1 reflects the complexity of variations in approaches to SIS development. It also shows a wide range of combinations of software capabilities and hardware devices. It is difficult to pick up a trend from such complexity. However, some general remarks can be made.

The importance of automation in the handling and provision of land resources information seems to have percolated research and academic circles into bureaucracies and offices of goverments of all political ideologies. SIS are in operation or in development in both, western and east-european countries. However, there are not many such cases in the non-industrialized world. In some developing countries SIS are only proposals and projects on paper. The prohibitive costs involved in obtaining the appropriate hardware and software capabilities, the lack of expertise to set up, manage and maintain SIS, or simply unawareness or indifference to automation of the tasks which are conventionally manual, can be counted as the main causes of the absence of SIS in developing countries.

SIS development strategies and methods that proved successful in Europe and North America have yet to be put to the test under the limiting conditions and constraints of all kinds characteristic of developing countries. The challenges of building up SIS in a non-industrialized nation are still great

Table 5.1

SYSTEM	ORIG COUNTRY	DAT
Regional SIS Armstrong and Wetherby (1976)	Australia	Va ti & li ca
National Fertility Soil Data Bank, Colwell (1976)	Australia	La ar
Computerized soil testing service	Australia	
Storage and retrieval of micromorfological descriptions Sleeman (1976)	Australia	SO sy ph
SIS for crop monitoring, & computerised agricultural management system	Bulgaria	Se da cc hc pi
Integrated information system for soil survey (cost = more than fl.1m)	The Netherlands	2 sj re sj tl n f
FAO (in development)	United Nations	
Can SIS (Federal Provincial Co-operative system)	Canada	7 1 o f s e a g g
Information system for Pedo-climatic pattern & land evaluation	Roumania	2 c c (2 c i t 1 (f

Table 5.1

SYSTEM	ORIGIN		OF
	COUNTRY	REGION	or I
Poseidon	France	General	Orst
Yugoslav Soil Information System	Yugoslavia	Sarajevo	Elek Facu Sara
New Zealand Resource Inventory	New Zealand	National coverage	Nat & S vat zat
INRA-6 computer-assisted cartographic system	France	Montpellier	Ins de Agr
South Australian Environmental data banks	Australia	South Australia	Div Lan Res
MIDGE A microcomputer soil information system (map	New Zealand	Nationwide	Soi Dep Sci Ind Res
ADK Integrated data processing system of Denmark	Denmark	Nationwide	Bur Dat Min Agr

Table 5

SYSTEM	ORIGIN		
	COUNTRY	REGION	ORGANIZATION OR INSTI
IRIS Integrated Resource Information System	U S A	Nation-wide	Soil Conse Service USD
KIBIS Kiel Soil Information System	Federal Republic of Germany	Kiel	Institute o Nutrition & Science, CA Al Brecht U sity, Kiel

in spite of the availability of the necessary hardware at falling prices, and the skills to create the software.

The systems reviewed contain information which ranges from vary particular aspects of soil such as micromorphology, to systems storing or interfacing information from several natural resources, technical and socio-economic aspects, and even digital satellite information. However, the core of the information in all systems reviewed consisted of site and soil profile descriptions or auger samples from survey.

For data acquisition several methods are in use. Field sheets, fill-in, tick-off and mark-sense proformas, optical character readers and field cards. The field sheets and fill-in proformas with fixed format appear to be the most commonly used for point data.

For map data the decision between raster or vector methods has lost importance since conversion programs from one to the other are now available. Nevertheless the use of digitizer appears to be the most frequently used method for map data acquisition, although grid-cell methods are used by many, particularly where the hardware for digitizing is unavailable.

Punched cards and card readers are still one of the most common input methods. But these are being replaced increasingly faster by remote terminals on-line to main frame or minicomputers. Microcomputers are also taking off and are being used as remote stand-alone terminals for off-line input into magnetic floppy discs. The information from these is later transferred onto the main frame computer if required.

The software used on input consists of programs for reading data from external media, and programs for checking, validating, re-formatig and up-dating data. Almost invariably each organization possesses its own set of in-house developed input routines. Few standard and commercially available packages for such purpose are reported in use.

The hardware for storage in most SIS are magnetic tapes and discs fitted into a variety of main frame and minicomputers. The problem of preservation of information stored in media for long periods of time seems to have found an adequate solution.

The choice of DB and DBMS model to suit the type of data in a SIS seems to have lost importance. Emphasis in SIS

development has moved away from the controversial topic of DB and DBMS most adequate for soil and land resources data, to other problems such as the spatial interpolation problem and the problem of reapprising point data.

The DB and DBMS problems have been solved, in most instances, by organizations developing their own software taylored to their particular needs, since early commercially available DB and DBMS were found not particularly suitable for modelling soil and other land resources data. However, software packages ad-hoc designed for soil and land resources data modelling are in development and becomming commercially available and some are currently in use in some SIS.

Few SIS were reported to have the "query" facility as a separate module. For most, it is imbedded in the retrieval procedure. The hardware for the query procedure is VDU terminal and keyboard which allow for interactive dialogue. However, in most cases querying a SIS was still carried out by punched cards and card readers, but desk top microcomputers and floppy disks were also found as part of the hardware for this module. Some systems intended the development of a complete query language. Queries in plain language then, could make a SIS totally "user-friendly", but it does not exist yet. Both built-in and commercial software for query are in operation in SIS. Most of them, such as "SOLFAB", "GRASP", "FINDER", "INFOL" and the "PL1-ENQUIRY" programs (see table 5.1), work interactively.

For the retrieval procedure the devices used are punched cards, card readers and VDU terminals and keyboards. Some desk-top microcomputers also do the job by submitting an already elaborated retrie val function as a batch job in a main frame. The media on which retrieval operates are magnetic tapes and discs for both, point and map data. Selection by Boolean algebra using logical operators is present in almost all the retrieval algorithms reviewed.

The variety of approaches to retrieval algorithms is as great as the variety of designs of DB and DBMS that are in use. These range from the single search-file programs, to laborious and sophisticated DBMS. Single-letter commands, elaborated Boolean equations or "macros" made up of several words operate for submission to the retrieval procedure and for the searches

through the data base. For some systems the retrieval of point data is carried out separately from the retrieval of areal or map data.

Manipulation procedures on retrieved data are carried out by a wide variety of applications programs. From simple report writing and tabulations to the most complicated statistical analysis, mathematical interpolations ad-hoc land classifications for land evaluation, cartographic analysis and display.

In terms of software these applications employ in-house developed routines and packages, when the application in turn is quite particular to the system. But, most of the statistical analysis capabilities in the SIS reviewed are provided by interfacing commercially available statistical packages, although a good number of SIS created their own statistical routines.

For the production of cartographic output, many existing SIS have resorted to commercial standard cartographic packages consisting of map overlay, display and map plotting routines.

Outstanding among the in-house developed cartographic packages is the map overlay and image manipulation and display system (MIDGE) in a micro-computer, (Giltrap, 1984). Soil or interpretive maps can be overlayed and composites displayed on the micro's colour VDU screen.

The varieties of output from SIS in operation, ranges from print-outs, listings, tabulations and reports from retrieved data to reports from complex statistical analysis and other mathematical treatment of retrieved data, cartographic output in the form of soil or interpretive maps, graphical output and plots. Most systems tended to use the line-printer as a main hardware device for output. The more sophisicated SIS use their own map and graphical plotters in their cartographic units and the VDU for visual displays.

The development of computerized soil information systems has exposed many conceptual and practical problems related to the handling of soil data in the computer. Solutions to some of the problems exposed challenge conventional views and perceptions of soil and even entire approaches to soil

information. A selected list of the most relevant ones is briefly discussed below.

5.3.1. The problem of pre-SIS volume of soil data.

The volume of pre-SIS collected soil data from survey operating for many years accumulates. In some cases it is immense and has required pre-classification or re-grouping at a very early stage resulting in the loss of much primary data.

5.3.2. The DBMS problem.

The decision between whether to buy commercial DBMS or to develop own DBMS software for a SIS is of importance. Commercial DBMS (developed for business), are not well suited for the more complex structure of soil data. As a solution, some workers have started developing DBMS software, but particularly tailored to their needs. Hence, difficult to make commercially available. A more general approach is required and there are not reports as yet of DBMS commercially available, developed specially for soil and earth sciences that have been found completely satisfactory by many.

5.3.3. Indexing and data modelling problems.

This subordinated problem is part of the more general DBMS problem. The choice of adequate indexing and identifiers that would improve data modelling in the computer and consequently SIS efficiency is somewhat arbitrary. There are no standard procedures for soil data indexing or guidelines for data modelling in SIS.

5.3.4. The Software compatibility problem.

Software, hence knowledge, transfer between SIS is a problem. It is desirable but difficult to achieve successfully due to the diverse nature of computer hardware capabilities and the consequent specificity of software packages developed for - them by soil organizations. There is a considerable duplication of software development effort around existing SIS. The method-base system proposed by Hatje and Lamp (1984), is an indication of the increasing awareness to this problem.

5.3.5. The problem of expertise.

The knowledge and expertise of a soil surveyor and land evaluator is progra able in the computer for the creation of a knowledge-based systems for problem-solving and decision making

in land-use planning. It is possible to envisage the creation of expert systems, i.e. knowledge-based, self-learning systems, in which the contribution of intuitive judgement and expertise in soil data interpretations can be incorporated.

5.3.6. The problem of free user/SIS communication language

Substantial improvements are yet to be made to achieve a complete free-language communication with SIS. On average the commands language to use SIS is highly restrictive and inflexible. The painstaking effort required to learn how to communicate fluently with a system, frightens many would-be users and contributors. More improvements on high level languages for more powerful statements and functions to allow easy communication with a SIS can be achieved.

5.3.7. The problem of fast and efficient data acquisition.

This problem consists of minimizing time, effort and the number of intermediate devices between data as collected in the field and its input into the computer. There is still a lot of room for improvement in data transmission from the field to the computer. In spite of the progress made, the task of data capture still remains laborious and error prone.

5.3.8. The problem of up-grading SIS.

Information technology changes at an exceedingly fast pace. After a relatively short period of time some hardware components become obsolete or are superce ded by a new generation of devices which do things faster and better. Some soil survey organizations which committed themselves heavily to the purchasing of costly equipment and software from different sources, have had to regret the apparent overlooking of this problem, i.e. The Dutch system, (Bie and Schelling, 1978). How "piece-meal" and how flexible, the design of a SIS ought to be to allow for changes and upgrading remains a key question.

5.3.9. The problem of time in SIS.

Existing SIS are time-invariant. If SIS development aims at becomming total information systems for crop monitoring and decision-making in planning, the dimension time must be introduced. There is plenty of time-related phenomena in soil science and crop production. Dynamic simulation models will soon find a place in the data processing tasks of SIS for land evaluation. Such models are essentially time-variant.

5.3.10. The problem of SIS technology transfer and cost.

SIS have their roots in industrialized nations where they are well established, but their design is based on assumptions about technology and capital that most developing nations can not meet. Since SIS have not yet penetrated the developing world, they have not been tested in such restrictive environments. It is not clear how much of the SIS technology is realistically transferable to developing countries, nor what is the contribution of creativity and imagination as substitutes for capital in the implementation of SIS in such nations. As a contribution to this thesis regarding this particular topic, a proposal of design of a soil and land resources information system for a developing nation, i.e. Mexico, is outlined in appendix 5.2. This model proposal is about the system's design in terms of hardware components, data structures and operational strategy. Figures and tables are used for the outlining.

5.3.11. The problem of data compatibility.

Data incompatibility can be tackled by either; imposing rigorous conditions upon the data on entry to match the system's format, or by designing a bank in SIS to have such flexibility as to be able to accept data of any kind and from any source. So far, the conventional solution has been the first as it is the easiest. Proposals for an international standardization of soil data coding has been the major thrust for tackling this problem so far. However, it is possible to design flexible SIS. OXSIS is an example of such a kind of systems.

5.3.12. The problem of dimensionality and conceptual framework in SIS.

The space framework in soils perception has conventionally been three-dimensional. Generalization has helped to deal with the multivariate nature of soils. This process has its roots in the need for reducing to something managable the bulk of information the human brain and its conventional devices had to process and could cope with.

However, the computer in SIS allows for retention and manipulation of large volumes of data. Soils no longer have to be visualized as discrete entities in a three-dimensional space. The old should give place to a new framework in which soils in digital form can be visualized in multi-dimensional

character space. The use of the old framework's concepts and conventions in computerized SIS no longer finds justification. Soil data perceived in multi-dimensional space and in digital form offer many advantages as an alternative to the old soils perception in SIS. One and each of the soil properties can represent a dimension in addition to the two spatial dimensions, altitude and soil depth.

Almost any other problem described above is related to this fundamental problem of how soils are perceived from nature and on what theoretical constructs they base their representation in the form of data. Bie (1984), has discussed some of the aspects of this problem. Formalisation of a new theoretical framework to soil science in multidimensional digital space is needed.

5.4 Roles of SIS

The main role of a soil information system is to provide information about soils. To do this, the computer in SIS can be employed in a number of ways most of which have already been described in this chapter and they will only be listed here.

Some SIS operate by getting the computer to do the old conventional things that in pre-computer information systems were carried out manually or by calculator. Some other SIS in addition to this, get the computer to perform new tasks which were not conventionally carried out in pre-computer information systems.

5.4.1. The Automation of the old tasks.

A number of old tasks customarily executed manually are now carried out automatically by SIS. These roles have been discussed above in this chapter and need no further elaboration.

i. The computer as a notebook
ii. The computer as a super-clerk
iii. The computer as a high powered calculator

5.4.2. The new capabilities in SIS.

The tasks listed above existed already in the absence of computers. SIS came in conveniently, only to perform them faster and perhaps better. However, SIS have also a role to play in offering powerful new procedures that have been incorporated lately.

i. The computer for producing ad-hoc interpretive classifications and maps.

Soil information is of great value only if interpreted or translated in terms relevant to the intended purpose. For a given practical use then soil information requires a standard interpretive link which is essentially an algorithm.

Ungeneralized point data in SIS can be used to create ad-hoc interpretive classifications using a standard interpretive algorithm or link. With such a coded algorithm apprisals and reapprisals of all available point-data can be made then for any conceivable land-use automatically in the computer. To produce the interpretive map then, either, the values of the relevant soil properties are interpolated on to a very fine grid and then every point is ad-hoc reapprised or else the original sites are ad-hoc reaprised and then a way to interpolate the assessment on to a fine grid is found so that ad-hoc boundaries can be drawn.

ii. The new geostatistical techniques for spatial interpolation of soil data and prediction in SIS

For many practical purposes, scientists increasingly require quantitative estimates of soil properties for areas of different sizes.

The theory of regionalized variables (Matheron, 1965), provides a sound theoretical background and the necessary tools, i.e. the semi-variogram, for the analysis of the structures in the spatial distribution of soils. Knowledge about these structures is used in the form of a semi-variogram model for spatial interpolation and local estimation by the technique called Kriging. Estimates of soil properties at any site within a sampled area of any size can then be obtained quantitatively and mapped confidently by kriging provided there are sufficient number of sampled points as support.

Studies on soil variability and on interpolation of soil properties are not new in soil science. What is relatively new, is the use of geostatistical tools based upon a sound body of theory to exploit spatial structures in the data and to interpolate and predict soil properties. Such techniques are being incorporated into SIS, (e.g. Giltrap, 1984), as standard algorithms for more accurate on-line local estimation and prediction which could not have been possible by conventional soil survey technology, (see chapter 8).

The system at the Soil Science Laboratory, Oxford (OXSIS), posseses facilities for the on-line computation of semi-variograms and kriging. These are a part of the manipulation procedure. The development of software for standard routines and interfaces is part of the work on the manipulation package for this thesis. OXSIS and its manipulation facilities are described in the next section.

5.5. The Oxford Flexible Soil Information System (OXSIS).

An information system for soils (OXSIS), has been in development at the Soil Science Laboratory, Oxford. OXSIS is a system for maximizing the use of existing data. The primary consideration for its development is that data are expensive and that much information is lost when records with different formats, properties, codes or conventions to that of a system have to be rejected or changed to fit into standard formats or conventions with which they are not totally compatible.

OXSIS represents an approach to tackle data compatibility problems. Since any data are better than none, OXSIS is a flexible system that has been designed to accept data about any properties in any units and in any format.

So far OXSIS is a local system. It holds soil data that covers the administrative area of Oxfordshire with some minor overlaps with the county limits.

Work on developing software for OXSIS was carried out as part of this thesis, and in agreement with the system's general design. A good number of applications algorithms and interfaces to other packages were written. These are now integreted into a

user-friendly package which became the OXSIS manipulation module. So, the contribution is reflected in the facilities for data manipulation that OXSIS offers.

To demonstrate the value of a soil information system in providing the on-line capabilities for predicting missing soil information for land evaluation and land-use planning, OXSIS was the tool used in this study. OXSIS provided both, the data through its data retrieval capabilities and the data processing required for prediction.

The system was designed to be user-friendly, that is, the user makes decisions through interactive dialogue with the machine in quasi-English command language. So no programming skills are required.

5.5.1. The nature of the data in OXSIS.

The data in OXSIS so far, is soil data exclusively. Nevertheless, other environmental features related to the site description may be included. A point on the land's surface represents a site. This is the unit of data. In the system's internal data-files, all data corresponding to a site are put together in a **block** of information (see figure 5.9). To each site/block is assigned a code number. Such a number is unique to the site within the information system.

Data-sets or data-files comprise a collection of blocks, all of which share the same block structure. OXSIS will be elaborated later to hold map data. So far it only contains point-data.

Thus, OXSIS has been designed to hold two types of data:

1). The values or states of soil and site features:
   - i. At a site
   - ii. Averaged over a small area around a site
   - iii. Averaged over all sites in a soil class or mapping unit.

2). The legend and location of maps, (e.g. soil, geology, land capability).

A fundamental characteristic of the data in OXSIS is that it is mixed data. That is, data which to other systems might

have been considered incompatible have been accepted in OXSIS. Mixed data means that the system accepts data which may show inconsistencies in the naming and definitions of soil features. The same name may be used for more than one feature, or for different definitions of the same feature. There may be more than one name (synonym), for the same feature, and the same feature may be measured in different ways and recorded in terms of different units or different states (or manifolds). Also, data may be collected in different ways or in different kinds of sampling framework, that is, it may have different status. Later, it is explained how OXSIS deals with such mixed data.

There are a number of variables that are mandatory and ought to be included in each site/block. These are essential for identification and indexing of the data; site number, site location (i.e. easting and northing from the national grid-reference), status, and source of the data. In addition to these, the data block may include; site description, e.g. topography, geology, climate, landform, relief, land cover and land-use, soil class or map unit. It may also include soil profile data, e.g. depth of soil horizons, their pedological classifications, soil horizon descriptions, chemical analysis and/or depth of each of those horizons is specified. Appendix 5.1 shows a fully comprehensive list of the variables for which OXSIS offers data.

5.5.2. OXSIS Design.

The design and structure of OXSIS may be explained in terms of the recognizable stages or operations from the acceptance of data to the selective retrieval and transformation through manipulation and output to provide the information wanted. The components of OXSIS are therefore, the following main groups of procedures:

i. **Input**
ii. **Query**
iii. **Retrieval**
iv. **Manipulation**
v. **Presentation**

Most of the data for banking in OXSIS had already been collected and stored in media at the time of initiating development work for the system. Hence, the data acquisition module for which field fill-in proformas were used, (Hazelden et al 1976), will not be treated here. Fig. 5.5 illustrates graphically the main five groups of procedures that make up OXSIS design. These procedures or modules comprise specific software and hardware capabilities.

5.5.2.1. Software.

Strictly from the software point of view, OXSIS is made up of:

i. Data
ii. Packages of programs
iii. Supporting Files

Programs packages and suporting files are related to the five OXSIS procedures. They control and manipulate the system. Fig. 5.6 shows the software components as related to the four of the main procedures.

5.5.2.2. Hardware.

In terms of hardware the system consists of:

i. Set of pro-formas to collate information
ii. VAX minicomputer and peripherals
iii. ICL-2988 main-frame computer and peripherals
iv. Storage media (i.e. magnetic discs and tapes)
v. Graph plotter and Tektronic graphic display terminals.

OXSIS computer power is provided by Oxford University Computing Service in a time-sharing basis with many other users from other faculties and departments. Hence, the computers and other peripherals are not exclusively for OXSIS service. This brings some constraints in the operation of the system, particularly in storage and speed of processing.

Whilst the philosophy behind OXSIS design has aimed at achieving maximum flexibility on the input, query, retrieval and manipulation procedures, on the other hand, operationally the

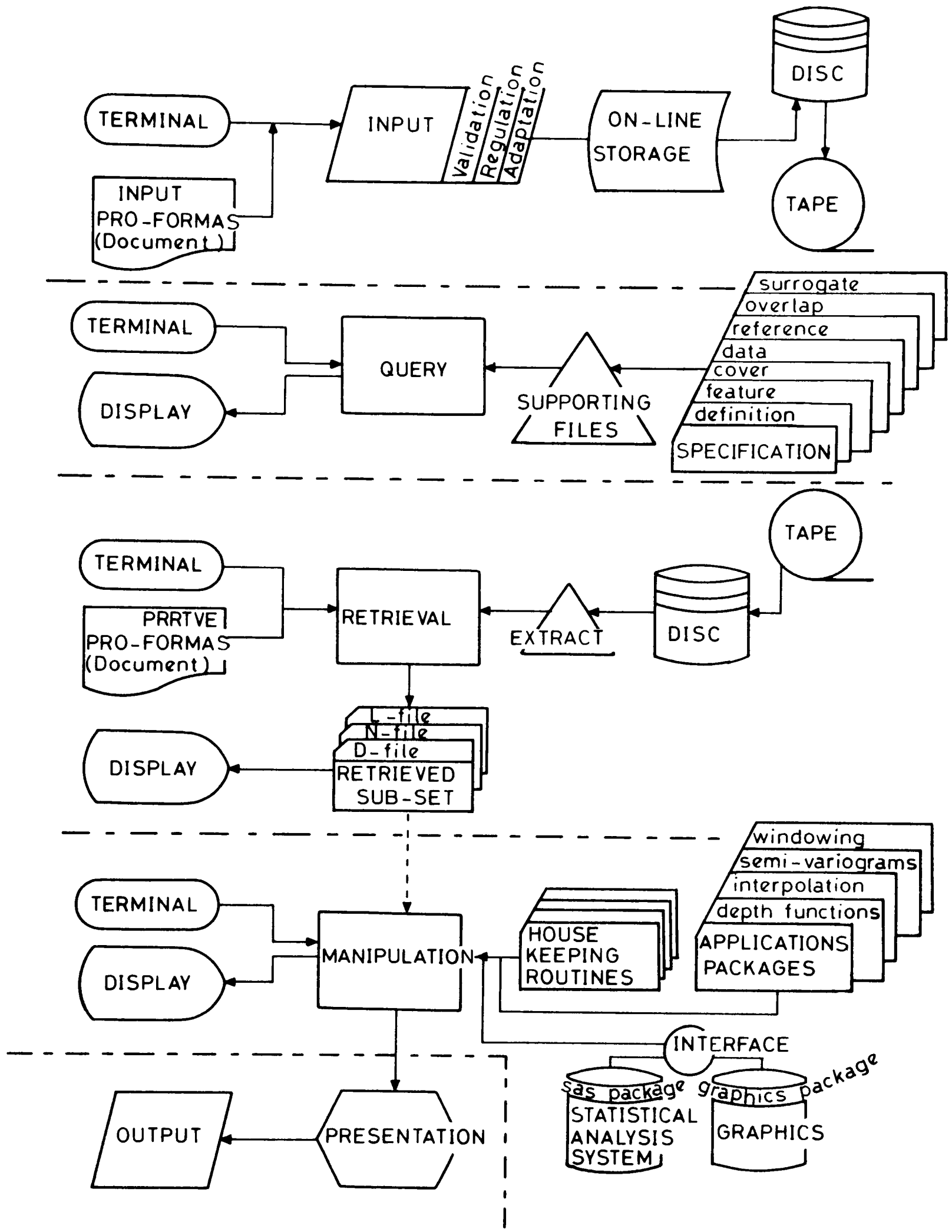


Fig. 5.5 The main 5 procedures in OXSIS

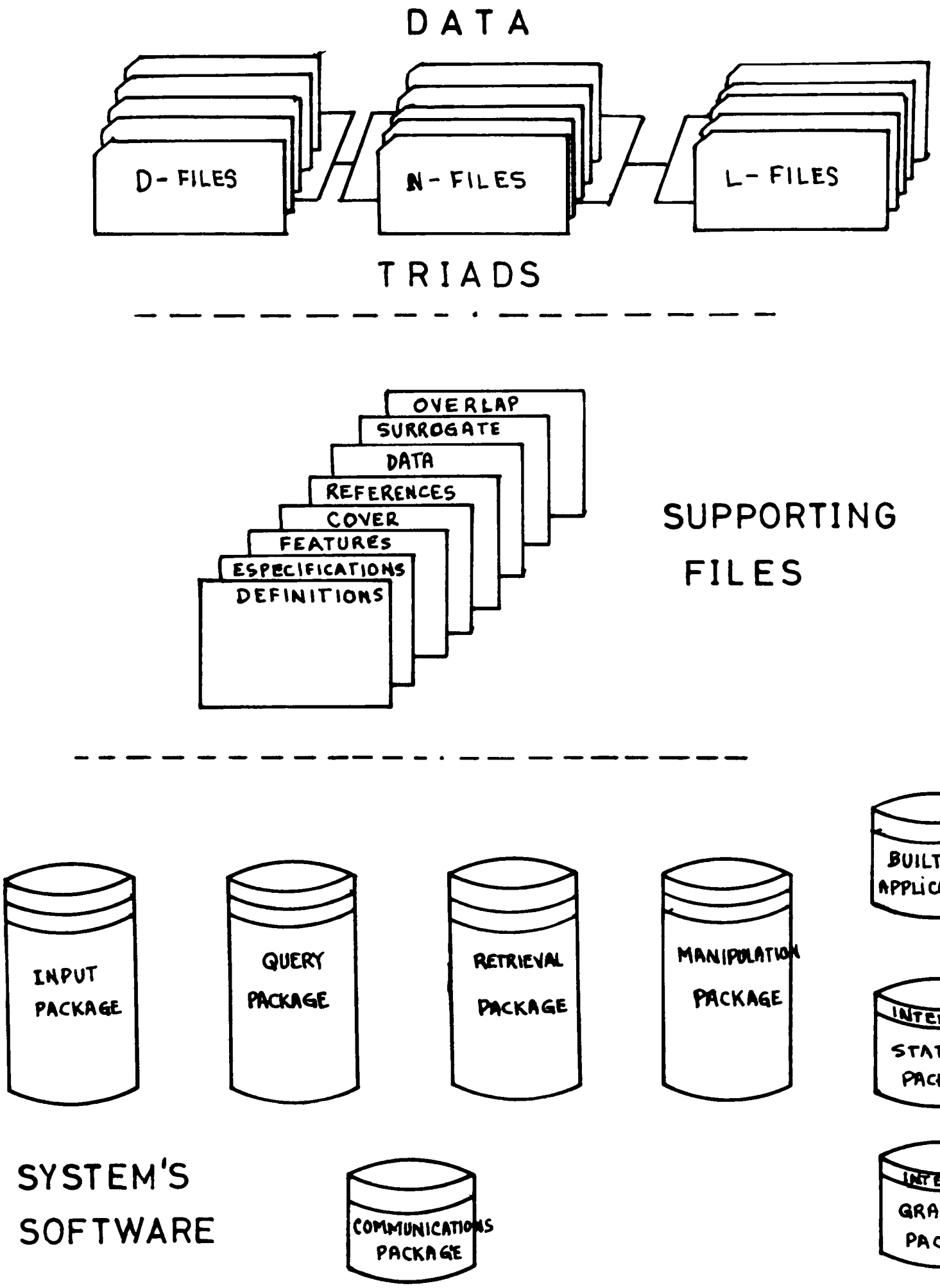


Fig. 5.6 OXSIS software components

system attempts to simulate or mimic the way in which a system of national scale would operate.

To achieve flexibility, the operation of OXSIS is effectively carried out through various modes. The flexible decision making stages that require the user's interaction, are separated from the long searches, number-crunching and other time-consuming manipulations through large data files.

To the user's view, OXSIS appears to present four modes of operation:

i. **Manual Mode**. Users are encouraged to facilitate the decision making stage in the interactive dialogue with the system by using pro-formas to collate information and decisions as much as possible before entering the interactive mode.

ii. **Interactive Mode**. Decision making stages and flexible choice are carried out interactively by dialogue with the VAX minicomputer. Input, query, retrieval and manipulation procedures are in effect, carried out by presenting to the user a hierarchy of menus and interactive questionaires, of multiple choice, in a "user-friendly" fashion. The VAX minicomputer holds in its filestore, i.e. supporting files, all the information about the data in OXSIS required to allow flexible interactive dialogue. The dialogue takes place by invoking a series of commands and prompts in quasi-English language. Usually, but not always, the operation on this mode ends with the creation of an instructions file, which is submitted to a batch or interactive job according to its contents.

iii. **Batch Mode**. Some stages requiring long manipulations and number-crunching of large data files are carried out on the ICL-2988 main-frame computer. The instructions for such jobs are submitted in the form of instruction files from the interactive mode in the VAX computer via the VAX/ICL-2988 communications link.

iv. **Archive Mode**. Back-up copies of the data-files, the whole of OXSIS software and of other relevant archival material are kept on magnetic tape. They can be reached through the main-frame (2988-T) computer if required.

The operation on these four modes on the two computers somewhat simulates the operation of a national SIS (see appendix 5.2. for a model of a national SIS for Mexico). The interactive mode on the VAX minicomputer via remote terminals represents the regional centres within the country. Meanwhile, the batch mode on the main-frame represents the head centre. The communications link between computers simulate the communications between regional centres and the head office, where the bulk of the national data is held. Regional centres may keep a copy only of their own data. Major batch jobs are submitted to the head centre (i.e. 2988 main-frame), from regional centres (i.e. remote VAX terminals), and results are sent back via the intercomputer communications link.

The arrangement of hardware devices in OXSIS is shown in fig. 5.7. The five OXSIS procedures are discussed briefly next.

5.5.3. Input Procedure.

Since OXSIS is designed to accept any kinds of soil data as a way of tackling problems of data compatibility, the input is the most elaborate and time-consuming of the five procedures. It is also organized around a flexible interactive framing of instructions. Decisions are based on the information held in the supporting files which describe the nature of the data already in OXSIS.

So, on inputing data into OXSIS all the relevant associated information about the data (i.e. how they were collected, analysed, units, definitions, e.t.c), must be entered and recorded into supporting files too.

The main aim of the input procedures is to check data for errors, to clean and validate them (validation), to regulate them (re-code into decimal numbers where necessary), and to standardize them as far as possible (adaptation). At the end of this procedure, the data-files contain only data on defined features, in real numbers and in units and manifolds standard to each different feature entered.

The input procedure in OXSIS consists of three stages, namely: validation, regulation and adaptation. Each of them in

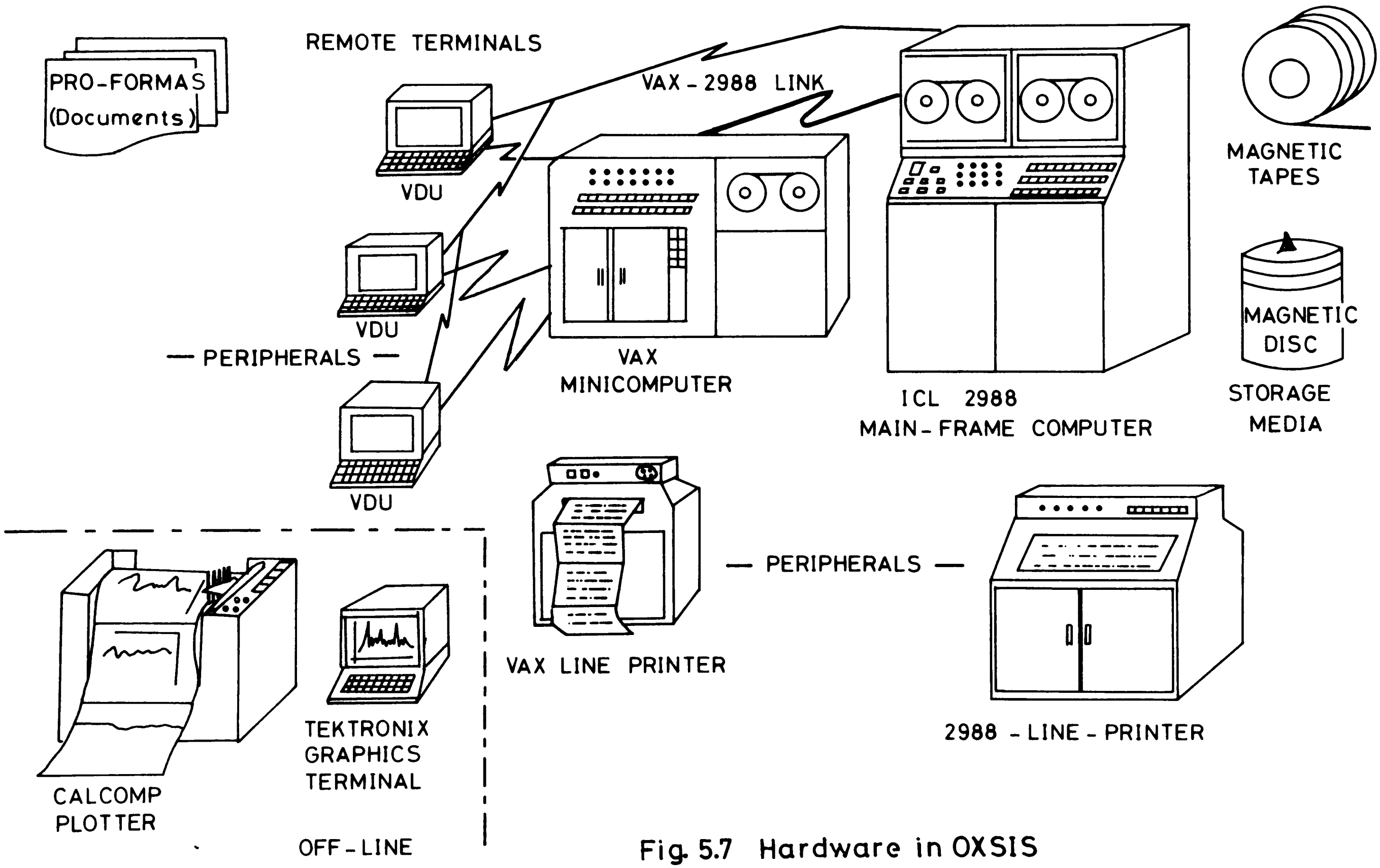


Fig. 5.7 Hardware in OXSIS

terms of software is a sub-package of algorithms that execute the corresponding operations. They are activated from the interactive stage in the VAX once the correct instructions have been framed.

In terms of hardware the input procedure is accessed from the VAX minicomputer by login into SIS and the job is submitted to and executed in the ICL-2988 main-frame batch mode.

In order to explain how OXSIS deals with mixed data, some aspects of the data storage must be discussed.

It has been stated that a point on the landscape represents a site and that a site is the unit of data. A site forms a block of information.

Storage of soil data is carried out in hierarchical structures as follows: The value or state of a soil or site feature is represented in the block by a string of characters. This is a record and it is the smallest unit of information, (fig 5.8). Usually, a record contains features similar in kind i.e. only morphological descriptions, or only laboratory analysis. One or more records in turn then form a block. The field occupied by a feature in the record is fixed and is the same for all similar records. A block (i.e. set of records) will always have all the information from one site (fig 5.9).

The records in one block may all be of the same format, but usually a block contains data records of more than one type (e.g. one record for site features and others for horizon descriptions and possibly a third kind of record for laboratory analysis) together in one block for they are data from the same site.

All data from a block is stored sequentially in a data file. So a data file comprises a number of blocks (at least one), all of which share the same block structure and cover the same range of features. This way OXSIS brings some sort of uniformity into the data files.

It was arranged by design that the structure of a block will always be laid in OXSIS format, which consists of a set of numerical headings preceeding the records. The headings explain how many records of each type of data are to be expected in the block. The D-file headings preceed the block headings and apply to all blocks in the file as they are uniform in their block

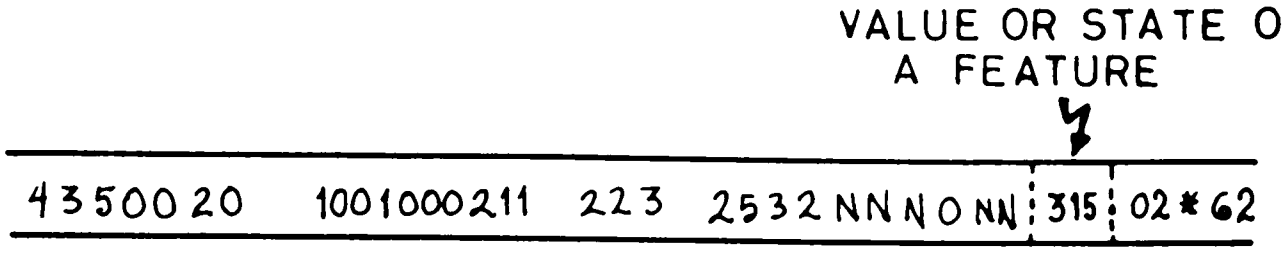


Fig 5.8 Representation of a record in a block of information.

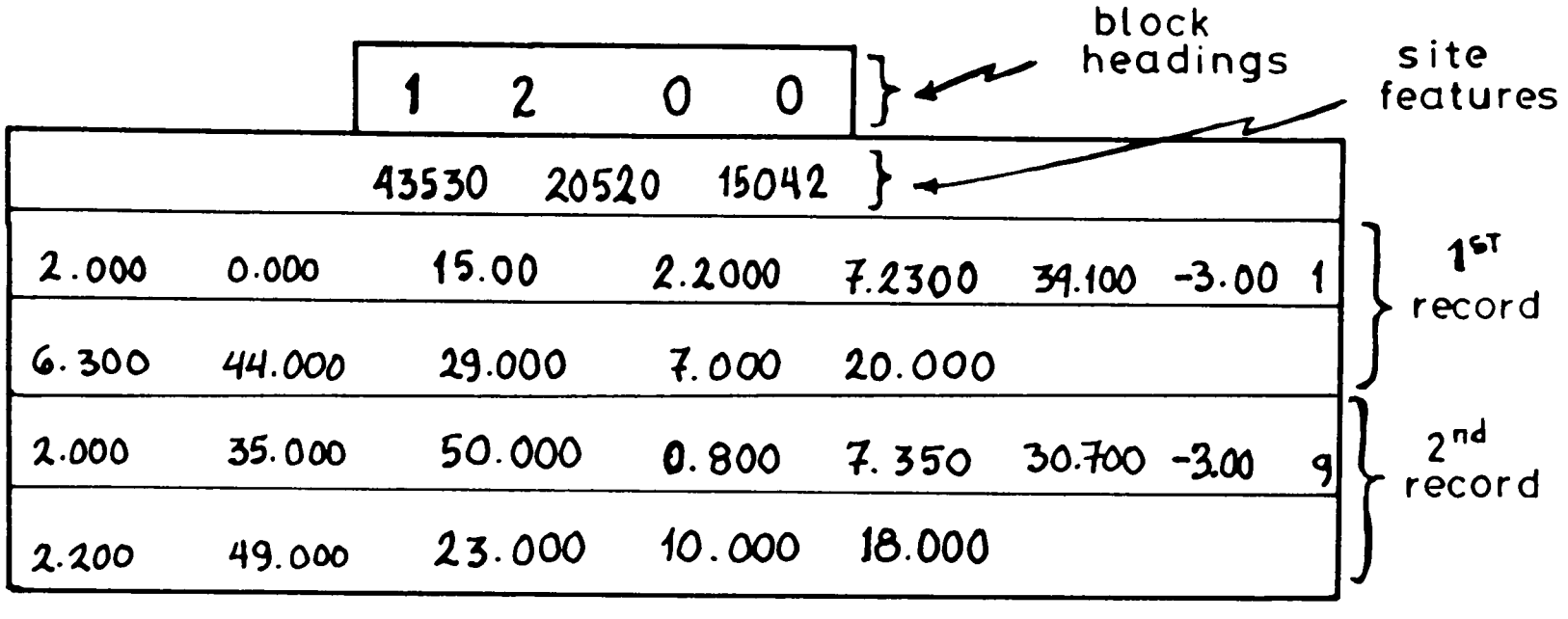


Fig 5.9 A data block in Oxsis.

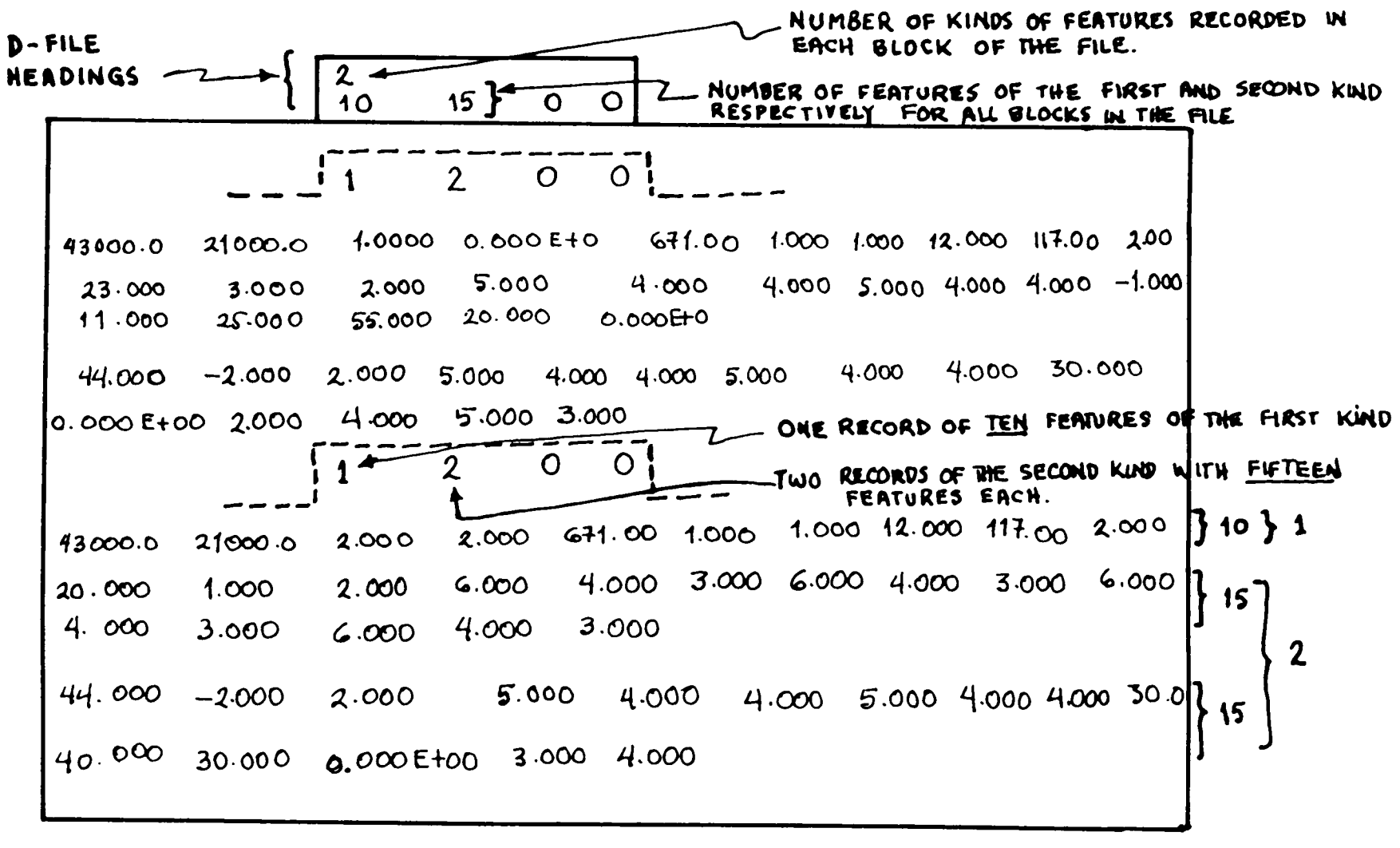


Fig 5.10 A data file (D-file) showing the Oxsis-format and the block structure used for data storage. (File of two blocks).

structure. The D-file headings explain how many types of information (i.e. formats), are held in the file's blocks and how many features (i.e. fields), for each type. Such file headings apply to all the blocks in the file (fig 5.10)

A data file (D-file), must be accompanied by its two other sister files. A names-file (N-file), which lists the names of the features held in its corresponding D-file, and a locations-file (L-file), which lists the easting, northing and site number of each of the blocks in the D-file. These form a "triad" in OXSIS. The number of items in the N and L-files, lists and their order is totally in agreement with the OXSIS format of the D-file. So each file in the triad functions as a table in a relational data base. This concept of TRIAD is central to the flexibility of OXSIS, hence it is at the core of the system's justification. It permits to hold data on any feature provided its name has been recorded and listed in the N-file in accordance with the block format of the D-file. In addition to this, the geographical position of each block is also recorded and listed in the L-file, according to the sequence in the list of blocks in the D-file, very much as in a ralational data base. Fig. 5.11 shows an OXSIS triad.

All the internal handling of data in OXSIS is done in terms of TRIADS. The OXSIS response to the possible confusion that handling mixed data might create, is to acquire, compile and organize during the input procedure all the necessary information about the data-triads and about the data held in them. For reference such information is held in a package of supporting files.

5.5.4. Query Procedure.

To retrieve data from OXSIS the system is usually questioned firstly, to ascertain on what data it can provide information.

The query procedure is a package of routines which allow the user by interaction with the machine to find out whether the data he requires exists in OXSIS, and if so, to ascertain in which triads the data are held. The OXSIS supporting files have

D-FILE

```
          2
          6          10          0          0
          1           4          0          0
   -3554.00    20563.00    95001.00    3.000000    1173.000    2.000000
    15.00000    7.000000    4.000000    3.000000   -2.000000   -2.000000   -2.000000    13.00000    0.0000000E
+00   1.000000
    35.00000    7.000000    5.000000    4.000000   -2.000000   -2.000000   -2.000000    17.00000    0.0000000E
+00   1.000000
    70.00000    7.000000    5.000000    4.000000   -2.000000   -2.000000   -2.000000    17.00000   -2.000000
       2.000000
    100.0000    9.000000    5.000000    3.000000   -2.000000   -2.000000   -2.000000    17.00000   -2.000000
       2.000000
          1           3          0          0
   -3503.00    20540.00    95002.00    2.000000    1173.000    2.000000
    15.00000    7.000000    4.000000    2.000000   -2.000000   -2.000000   -2.000000    11.00000    1.000000
       1.000000
   -5.00000    7.000000    5.000000    4.000000   -2.000000   -2.000000   -2.000000    17.00000    0.0000000E
+00   1.000000
    70.00000    8.000000    5.000000    3.000000   -2.000000   -2.000000   -2.000000    17.00000    0.0000000E
+00   2.000000
          1           3           3          3
   -3571.00    20533.00    95003.00    1.000000    1173.000    4.000000
    20.00000    6.000000    4.000000    2.000000    4.000000    4.000000    2.000000    12.00000    0.0000000E
+00   2.000000
    50.00000    6.000000    4.000000    3.000000    3.000000   -4.000000    3.000000    16.00000    0.0000000E
+00   3.000000
    100.0000    5.000000    4.000000    3.000000   -2.000000   -2.000000   -2.000000    12.00000    0.0000000E
          1           4           6          3
   -3547.00    20511.00    95004.00    3.000000    1173.000    3.000000
    20.00000    7.000000    3.000000    2.000000    7.000000    3.000000    2.000000    13.00000    0.0000000E
+00   2.000000
    45.00000    7.000000    5.000000    4.000000    7.000000    5.000000    4.000000    17.00000    0.0000000E
+00   2.000000
    75.00000    7.000000    5.000000    4.000000   -2.000000   -2.000000   -2.000000    17.00000    0.0000000E
+00   3.000000
    100.0000    5.000000    5.000000    3.000000   -2.000000   -2.000000   -2.000000    17.00000    0.0000000E
+00   3.000000
          1           3          0          3
   -3512.00    20502.00    95005.00    0.0000000E+00    1173.000    2.000000
    15.00000    6.000000    3.000000    2.000000    5.000000    3.000000    2.000000    10.00000    1.000000
       1.000000
    50.00000    7.000000    5.000000    3.000000   -2.000000   -2.000000   -2.000000    13.00000   -2.000000
       1.000000
    100.0000    5.000000    5.000000    4.000000   -2.000000   -2.000000   -2.000000    13.00000   -2.000000
       2.000000
          1           3          0          0
   -3572.00    20556.00    95006.00    0.0000000E+00    1173.000    2.000000
    15.00000   -2.000000   -2.000000   -2.000000   -2.000000   -2.000000   -2.000000   -2.000000    1.000000
       1.000000
    60.00000   -2.000000   -2.000000   -2.000000   -2.000000   -2.000000   -2.000000   -2.000000    0.0000000E
+00   2.000000
    100.0000   -2.000000   -2.000000   -2.000000   -2.000000   -2.000000   -2.000000   -2.000000    0.0000000E
```

N-FILE

```
EASTING
NORTHING
SITE_NO
SLOPE_GRADIENT A01
DATE
DRAINAGE,_PROFILE A01-M01
HORIZON_DEPTH,_LOWER *01
COLOUR,_MATRIX,_HUE A01-M01
COLOUR,_MATRIX,_VALUE A01
COLOUR,_MATRIX,_CHROMA A01
MOTTLES,_COLOUR,_HUE A01-M01
MOTTLES,_COLOUR,_VALUE A01
MOTTLES,_COLOUR,_CHROMA A01
TEXTURE_CLASS B01-M01
ORGANIC_MATTER,_ABUNDANCE A01-M01
CARBONATE A01-M01
```

L-FILE

```
2056343584 95001
2054043503 95002
2053343571 95003
2051143547 95004
2050243512 95005
2046643572 95006
```

Fig 5.11
An OXSIS Triad

all the information about what data is in OXSIS. So they have to be queried through an interactive dialogue.

The most common questions that can be anticipated are in terms of area, features, or status. These questions are of the kind

- Are there any data within area X ?
- Are there any data with records on property Y ?
- Are there any data collected by sampling method Z ?

The query procedure is performed interactively via a VDU terminal to the VAX. The user's questions are framed by using a set of commands or "macros", which basically, are keywords in English. Through each of the commands the user gains access to a particular interactive questionaire relative to his query. Table 5.2 shows important commands used during the query procedure, and what sort of information the user gets back through them. Fig. 5.12 illustrates graphically a perception of the query procedure. Appendix 5.3 shows some of the information displayed by OXSIS on the VDU screen after using some of the query commands.

The ability to query a system about the information it holds is of primary importance and it is central to its use. It says much about the usability, utility and justification of a system. The query commands in OXSIS can be invoked in any order, interactively, and as many times as necessary to let the user have sufficient knowledge about the data so to decide on retrievals from the bank. However, users should work out in advance, their best sequence of questions for their particular query.

The end of a successful query session should provide the user with a list of relevant data triads to him. Once the user knows exactly where in OXSIS is what he wants, he can then move on to the retrieval procedure.

Table 5.2 - List of OXSIS commands used in the query procedure to interrogate the supporting files on data available in OXSIS

Command	OXSIS Response
"COVER"	Produces a list of any triads which are fully within or overlap the area of interest. The area at this stage is only crudely defined.
"DATA"	Displays a list of the features recorded in a particular triad of interest. The triad's name has to be known.
"DEFINITION"	Informs the way a given feature of interest is defined in OXSIS together with its units or codes. It could also produce a list of all the features defined in OXSIS.
"FEATURE"	Displays a list of the data triads which hold records on the feature of interest. The feature's name must be known to OXSIS.
"OVERLAP"	Produces same effect as with "COVER" but in this case the area of interest has been more clearly defined.
"REFERENCE"	Produces a list of references for any definition in OXSIS. This helps to explain terms used in OXSIS by further consultation.
"SPECIFICATION"	Informs about the general specifications of any data triad of interest and offers comments on it.
"SURROGATE"	Displays a catholog of existing algorithms which can be used to calculate the value(s) of a missing feature from the values of features existing in OXSIS.

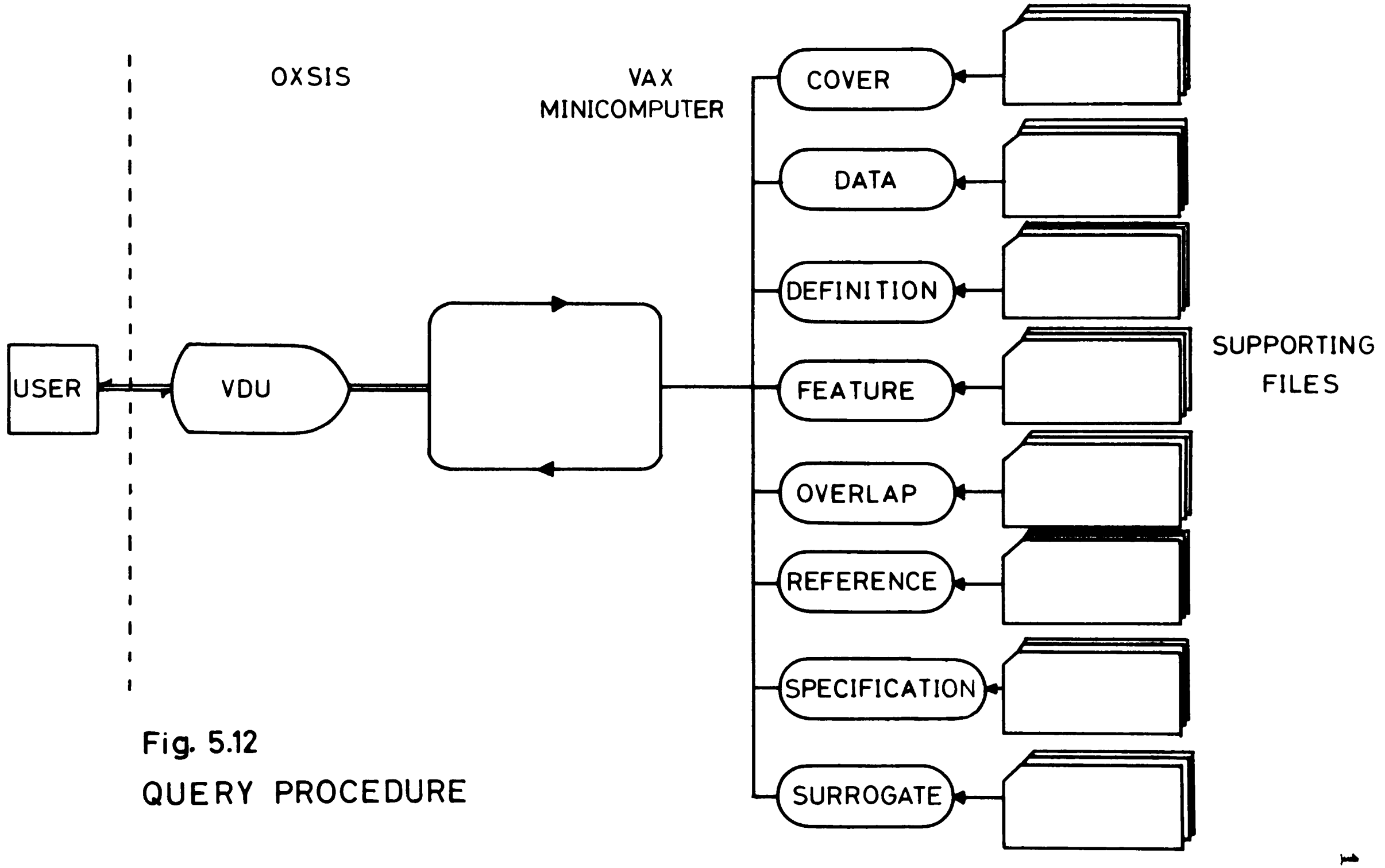


Fig. 5.12
QUERY PROCEDURE

5.5.5. Retrieval Procedure.

After the query, the relevant data triads have been identified. They cover the area of interest or have records on features wanted or their data was collected in the desired way.

Because a triad usually contains data from more than one block and blocks often hold site descriptions, horizon descriptions and analytical data, it is not unusual that the triads have more information than what is wanted. Hence, only the wanted data will have to be extracted from the relevant triads.

The retrieval procedure extracts the required data from the selected triads, one triad at a time, according to a set of retrieval rules.

Not all the blocks in a triad may be wanted and even more precisely, not all the features contained in the correct blocks may be wanted. So, through the retrieval rules the user specifies first, what are the wanted blocks. Having selected the correct blocks, then the wanted features which are contained in the selected blocks are specified. The block selection is done one block at a time by imposing certain conditions as discriminants. The selected blocks must meet them all. Thus, block selection can be done by:

a. Specifying the limits of an area of interest to which the blocks must pertain.
b. Specifying a list of the site numbers of the wanted blocks.
c. Specifying several conditions related to characteristics of a site or a soil profile as a set of Boolean statements, combined in a logical relation and expressed in the form of a Boolean equation. Such statements must be met by the wanted blocks.

The retrieval procedure permits up to eight Boolean statements as selection criteria linked by logical operators, e.g. "or", "and", in the equation, e.g. the statements:

```
A: = ALTITUDE A.GT.20
B: = DRAINAGE,_ SITE A01-M01.LT.3.0
C: = SLOPE_GRADIENT A01.GT.1.0
```

The relation: (A and C) or B.

Ad-hoc proformas help to collate information off-line for decision-making during retrieval. The information on the proformas is transcribed into OXSIS by answering the interactive questionaires. These questionaires first tackle the block selection and then prompt the user to specify what to do with the blocks that meet the condition. Some features within the block are always retained automatically. These fix the location (i.e. national grid reference), identification (site number)

The retrieval procedure ends by generating a retrieved sub-set (RS), in the form of a triad, i.e. a D-file, an N-file and an L-file, just as in the original triad but containing only the blocks and features wanted. Appendix 5.4 shows the interactive questionaires displayed on a VDU terminal during the retrieval procedure and the pro-formas to collate information before invoking the retrieve command.

The retrieved sub-set (RS), is ready then to be manipulated. This is achieved in the manipulation stage.

5.5.6 Manipulation.

The result of the retrieval procedure is a retrieved sub-set (RS), to which a name has been assigned by the user. The RS is in OXSIS format like its parent triad, namely, a triad made up of D, N and L-files. These files remain in the main-frame 2988-D computer.

The manipulation procedure comprises all operations that are performed on the retrieved sub-set. The result from these operations is allocated to a file, the manipulated sub-set (MS), to which the user has also assigned a name. Depending on the operation the MS can be further manipulated or else is left ready for presentation and output.

The manipulation package (i.e. set of applications programs) constitutes the data processing capabilities that OXSIS has to offer. It has been designed so that it can be expanded by adding new facilities. Knowledge of Fortran programming is necessary for this purpose. So far, the package has been

elaborated to comprise four kinds of facilities: i. "housekeeping" routines; ii. data-processing routines; iii. interfaces to external statistical routines, and iv. interfaces to graphic packages.

The elaboration of the manipulation package is an important part of this thesis.

For the execution of the trials in predicting soil information (see chapter 9), effort in developing software was necessary. These routines which had to be developed were designed so that they could be accomodated into a package for OXSIS. The software development effort was aimed at two ends. Firstly, to provide the necessary statistical and numerical tools for data analysis so that an exercise on the depth and spatial prediction of missing data could take place. Secondly, to elaborate the routines to a point from where they could be put together and become a user-friendly OXSIS facility. The manipulation package purports to equip the soil surveyor with the suitable software to operate various useful applications on data from survey, notably depth and spatial interpolation, spatial analysis and parametric statistics. The structure and design of the OXSIS manipulation package are described in detail in the next chapter of this thesis. Data from Oxfordshire currently stored in OXSIS are used to illustrate each of the manipulation facilities with examples.

5.6. Concluding remarks to this chapter.

i. Soil data from survey are expensive and considerable information may be lost due to the generalization implicit in the production of conventional soil survey output.

ii. A soil map is an inadequate data retrieval system for the volume of point information collected from soil survey.

iii. Soil Information Systems represent an alternative to conventional soil survey procedures for the storage of vast volumes of data, efficient retrieval, and powerful manipulations on data and for providing the

information needed for land-use planning.

iv. There are many advantages to the automation of the routine tasks in producing conventional soil survey output. SIS can carry out these tasks faster and in many instances better.

v. In designing SIS, identifying clearly its potential users is a key factor to its capabilities, efficient architecture and costs.

vi. Since data, software and hardware are expensive, the design of SIS should be as flexible as to reduce to a minimum problems of compatibility in data, and software and to allow for easy up-grading of the hardware components of the system.

vii. A new, imaginative and inexpensive approach is needed to face the challenges of building-up SIS in developing countries.

viii. The development of SIS has posed many theoretical and practical problems. The most fundamental of all lies in the need for a change in our pre-computer perception of soils, its conceptual framework, and of soil information. This may even lead to a change in our entire approach to soil science.

ix. SIS can be elaborated to interface other land resources data bases, offering an enormous potential as a tool for decision-making in land-use planning and crop monitoring. Its capabilities can even be expanded to become knowlwdge-based or expert systems for problem-solving and advisory work in those fields.

x. So far, the new on-line capabilities for spatial analysis and interpolation of soils in SIS make them a powerful tool for predicting the condition of the soil at unvisited sites.

VI. PREDICTING MISSING INFORMATION: Special problems on interpolation.

Soil information is necessary for land-use planning and it is obtained from field surveys. Chapter 4 has reviewed the conventional procedures for soil survey and the nature of their output. Soil classes and mapping units are defined from a limited number of field observations on which inferences have to be made in the form of generalizations. In effect, the process produces, implicitly, a form of interpolation or prediction of values of soil properties at all those sites unvisited and unsampled during the field survey on which the generalizations were made. This chapter deals with some special problems that emerge during prediction and that demand special procedures for interpolation of soil data.

6.1. Missing soil information.

From chapter 4 it was observed that most soil information derives from observations or measurements at points (sites) on the landscape, and that the unit of observation of soil is its profile, i.e. a vertical section usually 1-2m deep from which properties are measured at several depths within it. Although the soil profile has an areal extent this is such that it could be conveniently assumed to represent only a point on the Landscape. Properties may be measured over layers thin enough to be considered as point depths or as averages when taken from layers 2-30 cm thick.

The total soil information of an area would comprise the values of every soil property that can be measured at every site and at every depth. If it could be possible to plot such values in a multidimensional space in which every property was one dimension, and depth and the three geographical co-ordinates were one dimension each, the total soil information would be defined by the hypersolid produced. The total soil variability of the area

could be said to be completely described. If available and represented in this way it would form a complete data-set. There would be no need to predict soil conditions at unvisited sites or to make generalisations to assist predictions.

In real life it is practically and economically imposible to obtain the total soil information from an area. The soil surveyor and the user of the data he collects must always work with incomplete data-sets from a limited number of samples. They represent only a minute portion of the total soil information. In effect, the hypothetical hypersolid in the multidimensional space as described above contains immense empty spaces between the plotted data. The information not available and not plotted is missing. This constitutes the problem of missing information.

In practice the soil surveyor working with incomplete data-sets will seek to provide the information missing by generating or predicting it from the point-data available. Hence the problem is typically one of interpolation. Depending on the case, the surveyor's task is to develop a suitable equation or to choose the most appropriate interpolation procedure which will give the best predictions.

In a practical context, on the other hand, the missing information needs to be predicted for it is usually related to the answer to practical questions in which either the purpose is known and the suitable soil is to be found or else the soil conditions are known and its best alternative use is to be assessed. The format of the questions may read:

a). "can Y be done/grown/built at site X ?"

b). "where can Y be done/grown/built ?"

Questions of this sort may be sufficiently answered by a combination of soil information from site X plus information on the soil requirements for purpose Y. A suitable interpretive algorithm between soil properties and a particular use (conveniently called the "link") is recognized, but this could also be part of the missing information.

Four categories of missing information related to soils and their use can be clearly distinguished:

i. Soil property missing.

There are data for the right site and at the right depth but no records available for the property wanted. The soil surveyor in this case will attempt to estimate the missing values of the property of interest from available records of other properties. Quantitative relationships between these properties may be based on common knowledge derived from experience or from their expected assosiation based on observations elsewhere. Multiple linear and nonlinear regression techniques are used in this context to investigate the degree and nature of the quantitative relationships bet ween the variables and to generate predictive models. These may be called surrogate functions.

ii. Data available but from the wrong depth.

Information is available for the right property at the right site but there are no records from the depth wanted. The soil surveyor will have to interpolate the missing values at the wanted depth from the records available for the same property at other depths. A suitable function of the wanted property on soil depth is used as interpolator. Any quantitative expression that relates the vertical variation of a soil property to its depth will be called a soil depth function. This mathematical algorithm models the distribution of the property through depth and interpolates either values of the property at specified depth(s) or the depth(s) at which the property has specified values.

iii. No data from wanted site, or sites.

The wanted site was not visited during survey and consequently no records exist for any property at any depth. This is by far the most frequent problem in practice. The soil surveyor will have to use any method for spatial interpolation available to him. If records for the property wanted at the depth required are available for sufficient neighbouring sites the matter is a problem of spatial interpolation. If such is not the case then information will have to be either generated or interpolated depending on the situation by using "surrogate" functions or depth interpolation, or both, first, and then to interpolate the values of the property at the wanted site(s) by using spatial interpolation techniques. From the range of interpolation

techniques available, the problem then consists in choosing the best technique for the data and situation.

iv. Interpretive link missing.

Data exist for the property wanted at the right site and at the right depth but no "link" is available between soil data and land-use requirements to produce meaningful interpretations. The soil surveyor will seek to find a suitable algorithm to assess the quality of the land for the intended use by translating the values of the soil property(ies) into categories or suitability classes. If a number of sites are being assessed, in effect he will attempt to find the best discriminants related to the use intended which will partition the sites into classes or categories. Interpretations in terms of suitability class can then be made for the site(s) wanted.

Point (iv) above is not the concern of this thesis. The present chapter is concerned with special problems that arise on the mathematical interpolation of soil data required to provide the missing information to solve problems (ii) and (iii), and with the generation of predictive models to solve problem (i).

6.2 Problems in predicting missing soil properties by means of surrogate functions.

The generation of values of unrecorded soil properties necessitates the calculation of a predictive model (the surrogate function) based on recorded properties which are presumed to bear some degree of association with the wanted property. Usually a function of this sort is constructed by multiple linear or non-linear regression or curve fitting techniques. There are a number of technical problems in estimating such a function. These are similar to the problems in any model-building exercise by regression and curve fitting. Most of them, however, are part of any of the two general problems:

i. The selection of the independent variables to appear in the model.

ii. The mathematical nature of the model itself.

If the model is known in advance these two problems vanish and the only problem then consists on defining the criteria for the

best equation which will become the final model. The strategy to develop the model will be targeted to find the parameters of the best equation. Thus the best equation can be defined by either, the set of parameters that yield the closest fit of the known equation to the data, or the set of parameters that yield the best predictions, or both.

If, on the other hand, the form of the model is unknown in advance, then problems (i) and (ii) above have to be solved. These problems of selection of variables and model type are well known and have been sufficiently studied in statistical analysis practice, where numerous procedures involving a variety of selection criteria are available in the form of computer algorithms and packages. However, since the model is unknown, most of these algorithms have a common tendency to assume and fit linear additive polynomials and find the combination of variables and their parameters that produce the best fit. Nevertheless, in nature, many relationships between variables of this sort tend to be rather complex and most of them of the exponential type.

In this thesis, the general problem in constructing the surrogate function is to devise a strategy for model-building using the best of the standard procedures available and combinations of them, to tackle all possible situations arising during the model development process and particularly related to the solution of (i) and (ii) discussed in this section. Chapter 7 discusses and offers such strategy for both situations, when the model is known and when it is unknown. In appendix 7 the results of an exercise in constructing a surrogate function can be found in the form of computer print outs.

6.3 Problems in soil depth interpolation.

Procedures for solving problem (ii) require depth interpolation. A range of quantitative techniques are available for this purpose, but even the best of them present some problems of relevance in practice.

It is a known fact that any soil property varies with depth. This variation is acknowledged to be smooth and not discontinuous. Continuity and smoothness are the two properties of soil vertical variation that facilitate modelling by any approximative

function. The degree of complexity of the approximative function required may vary considerably depending on the property in turn and the nature of its vertical variation.

Ideally, if the soil property was measured without error, the continuity and smoothness of the "true" curve could be known from an infinite number of consecutive point-samples down the soil profile. In reality the "true" function is never known. The soil profile is sampled by taking non-contiguous segments of the same or different thicknesses from pits or bores. The data obtained this way always contain some error even if introduced only by laboratory procedure. So, depth functions will always be approximations.

However, in reality profile data is obtained essentially from three ways of sampling:

(a). By horizons defined by limits according to observed similarity or uniformity on the profile's face.

(b). By dividing the profile in strata at fixed depths.

(c). By small samples representing a point (point-depth) on the profile's face.

These sampling procedures all yield discrete data. Samples from (a) and (b) procedures are bulked and from analysis a single value for each property is obtained to represent the thickness of the horizon or stratum. In effect, from the "true" continuous curve (see fig 6.1a), values are averaged, and since only one value is available they produce the "stepped" curve of figure 6.1b. Values from point-depth samples represent only a set of non-contiguous points which in the best of the cases lay somewhere along the true curve.

The interpolation problem consists of reconstructing the "true" smooth function as close as possible, starting from the stepped profile and its limited number of observations and using such a reconstructed curve as interpolator by evaluating its value at the required points along the depth axis.

The reconstruction must be as close and accurate as possible to serve a number of practical purposes where accuracy is important. Yet any reconstruction of the "true" profile will start off from values which have already been "damped" by averaging (see figure 6.1c).

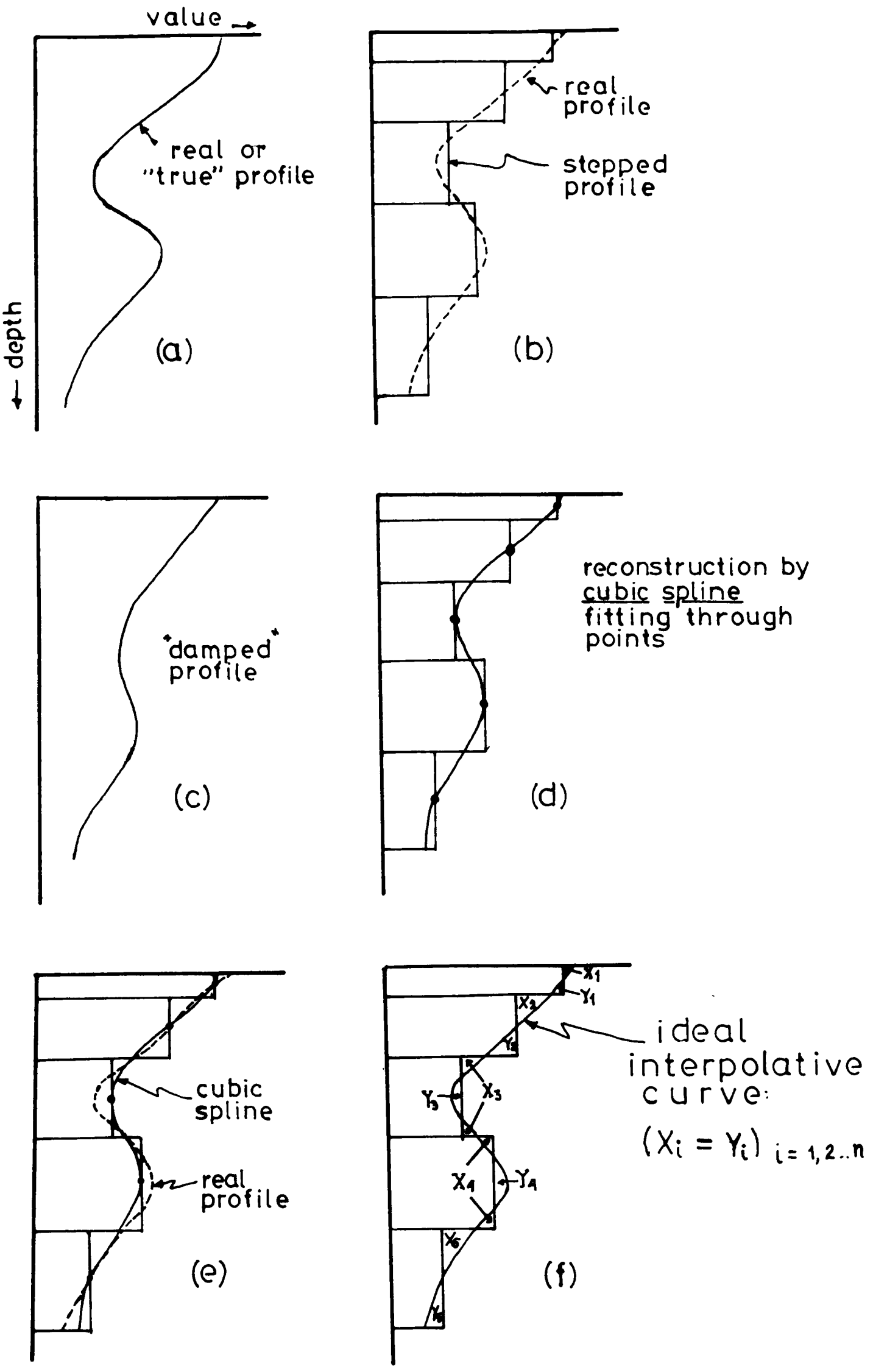


Fig. 6.1 The depth interpolation problem.

Atempts have been made to fit polynomials to model and reconstruct the smooth depth functions. Moore et al (1972) fitted orthogonal polynomials of fifth degree to profile data by least squares. The goodness of fit was inconsistent and overall poor. Russell and Moore (1968), also fitted exponential decay models to properties like organic matter which tend to decrease from the topsoil down. Their fit was reasonably good but based on many points which is uncommon in this type of data. Further, there is no theory from which to assume that any soil property should vary exponentially with depth. Functions which express physical relationships in soils are frequently of a disjointed or disassociated nature. Their behaviour at one depth may be totally unrelated to their behaviour in another layer, and polynomials have just the opposite property. Their behaviour everywhere maybe determined by their behaviour in any small region.

Spline curves (see chapter 7) overcome the limitations of global polynomials by reconstructing the curve piecewise. Webster (1981) has shown that cubic splines posses many advantages over other procedures for approximating depth functions. Cubic spline curves can be made to pass exactly through the data points, i.e. the layer midpoints, (see figure 6.1d), but Webster (1981) indicates that a least squares cubic splines fit to profile data represent the general behaviour of the depth function better than an exact spline.

However, cubic splines have the disadvantage that they fit curves through horizon averages. That is, they start off from data whose maxima and minima have already been damped by bulking during sampling. So they reconstruct profiles which will inevitably show smoother convolutions than the original, (see figure 6.1e). Yet for some applications extrema are what matter, e.g. maxima and minima of a soil property at a given depth may be criteria for soil suitability classification.

Hence, better reconstructions of smooth profiles (see figure 6.1f) are required to be used as depth interpolators. Two alternative approaches can be suggested:

i. Modify the sampling procedure of soil profiles in the field so that it exploits best the virtues of the cubic spline fitting, or;

ii. Modify the curve-fitting procedure so that it accounts

for the oversmoothing produced by the bulking or averaging that occurs during the sampling of horizons the conventional way.

Following approach (i), and with the purpose of helping the ordinary cubic spline fitting in reconstructing the true profile, an alternative sampling procedure was explored, (Ponce-Hernandez et al, 1986). This consisted of taking samples at point-depths down the soil profile which were placed where it was believed that they could pick-up extrema in the values of the property. The efficacy of these two procedures (i and ii) in reconstructing the original true curve is compared and the results included in appendix 7.

Following approach (ii) a modified spline fitting procedure specially designed to overcome this limitations is proposed in chapter 7. Details about its mathematical solution are also included. An algorithm in FORTRAN 77 to solve the modified spline and compute the set of coefficients of the curve is described in chapter 8. This algorithm also evaluates the curve at any point permiting its use for a range of interpolations related to practical purposes. Hence, the algorithm has been included as a depth interpolation facility in the manipulation package of OXSIS described in chapter 8.

6.4 Problems in spatial interpolation of soil properties.

Missing soil information is more evident in the spatial sense where the distances between observations are considerable larger than in the vertical direction.

The spatial problem is universal and characteristic of any natural body distributed with and over the earth's surface. It can be stated as follows:

Given a number of points representing observations or measurements, in this case of soil, which have a particular distribution over space, use this information to predict the value of a soil property at unvisited sites within the domain of the sample. This problem is typically of interpolation over space.

Spatial interpolation then refers to the problem of estimating the variable Z at some point P(X,Y) from the known values of Z at a number of data points usually arranged randomly.

Thus, there will always be some sampled data points. Their number and location are relevant to the solution . With such a data-set the problem, in effect, is to select the best interpolation technique for the given data and purposes.

The problem of spatial interpolation is not new. It has been addressed in various disciplines and there is a large number of approaches to its solution but relatively few, however, have been tried on soil. They are grouped according to their main features, and the problems that they pose to interpolation of soil properties are briefly discussed below, and dealt with in more detail in chapter 7.

6.4.1 Interpolation by boundaries.

The simplest approach to interpolation is a graphical one. The area of interest can be divided into smaller parcels of land of regular or irregular shape and these parcels used as domains to interpolate information within them. Inevitably, generalization of information occurs. The following are variants to this approach.

i. Thiessen polygons.

The delineation of boundaries is based on the nearest neighbour rule. This is a simple but safe mapping rule. The value of any given property or the class at any unobserved point is drawn from the values or the class of the nearest observed point. The procedure to find the nearest neighbour to an unobserved point amounts to drawing boundaries midway between neighbouring observations with different values of the property or contrasting class. For each observed point then a tile or polygon is defined so that it contains those unobserved points that are nearer to it that any other unobserved point, (see chapter 7).

Applied to soil data, Thiessen polygons help in defining interpolative neighbourhoods from observed points. Their goodness for this purpose however, has not been evaluated.

Burgess (1979) used Thiessen polygons in the context of soil sampling. The polygons serve in devising a mechanism for assessing the efficacy of sampling grids.

Thiessen polygons can be relatively accurate for interpolating soil data only when the pattern of soils is quite simple and the support observations are closely spaced, but in such circumstances

conventional soil mapping procedures or other mathematical interpolation techniques offer a better alternative.

ii. Parcellation.

Dividing the land into irregularly shaped and to some extent arbitrary boundaries can also yield interpolative neighbourhoods. This is the realm of soil mapping and the procedures are part of the conventional soil survey method reviewed in chapter 4.

The resulting parcels are soil mapping units that represent distinctive soil classes or combinations of them. Interpolation takes place by ascertaining the mapping unit corresponding to the point of interest and predicting its class. Values of the property are estimated from either the class mean or from the typical profile representative of the class. So, at the most, averages from the available observations within the corresponding mapping unit are the predictors, even though the units may carry many impurities.

There is much to be said about this way of interpolating, and the matter has been discussed also in chapter 4. On the whole, the low purity of the parcels or mapping units as defined in soil survey practice combined with the generalizations made to define the units may present serious problems in the obtaining of accurate estimates.

The need to improve the products of conventional soil survey as predictors has motivated the search for mathematical representations of soil variability to be used as predictors of soil properties.

6.4.2 Interpolation by Classical Statistics.

Classical statistics have been applied to soil data in many ways and for many purposes. This has been summarized by Beckett and Webster (1971), and later by Wilding and Drees (1983). However, their application has been without due attention to location and spatial arrangement of the observations. This approach assumes that the expected value of property Z at location X can be given by $Z(X) = m + e(x)$. Where $m = E\{ Z(x) \}$, the mean or expected value of $Z(x)$ from the sample, and $e(x)$ are random errors, independent and spatially uncorrelated. So, two points $P_1(X_1,Y_1)$ and $P_2(X_2,Y_2)$ a few meters apart are

expected to have a common mean (m) but fluctuations around it which are expected to be totally independent from one point to the other. In most instances soil variation seems to show just the opposite property.

Classical statistics also assumes that the errors are normally distributed. Within a given sampling area interpolation takes place by extracting the mean of the observations irrespective of their position on the field and using it to estimate properties at unsampled locations or for representing the whole area. Measures of dispersion around the mean, i.e. variance, standard deviation, standard error, confidence limits, are used to indicate the precision of the mean as estimator.

Estimations by this sort of interpolation totally ignore the spatial arrangement of, and dependence between, the values of the sampled points. The estimates may be rough.

The success of classical statistics in prediction depends on the complexity of the pattern of soils itself, the optimal number of samples required to adequately represent it, and the extent to which the assumptions made about the data are met.

Classic statistical parameters can only provide reasonable estimates if the pattern of soils is fairly simple or largely uniform and yet there are abundant observed points to represent those minor variations so that the fluctuations around the mean are small. Such conditions and particularly the assumptions about independence of the observations are their weakest point and difficult to meet in practice. Abundant evidence arising from recent work (Webster and Cuanalo, 1975, Campbell, 1978, Burgess, 1979, Burgess and Webster, 1980a, McBratney and Webster, 1981a, Yost et al, 1982a, Oliver, 1984), contradicts such assumptions. Most properties exhibit spatial dependence according to direction and distance within some localized region.

Estimation by classic statistical parameters can only be improved on if stratification or subdivision of the sampling area accounts for all spatially dependent variance (McBratney et al 1982).

6.4.3. Interpolation by Smooth Surfaces.

This group of techniques have in common that they attempt to incorporate the location of the samples in the interpolation procedure. They interpolate by fitting, globally or piecewise, a smooth surface to data on the plane. The smoothness is inherent to the mathematical procedure that fits the surface. So in general they are implicitly based on the idea of gradual change of soil properties with distance. They describe information which is mixed with noise without knowing which is which. Hence these techniques may smooth out or filter much of the short range variation with the noise.

Some of these techniques are best thought as filters, that is, intended for smoothing purposes but capable of use for interpolation. Their approach to the fitting splits them in two groups: global and local methods.

i. Global fitting methods.

The procedures in this group fit a common surface to the whole space covered by the observed points. Linear least squares are used as the fitting method, (a form of multiple regression). Trend surface analysis is the best known in this group, (see chapter 7).

The trend can then be represented by a global algebraic polynomial on the geographical co-ordinates of the points.

There have been only few attempts to apply trend surface to soil data (Walker et al 1961). The results have been inconsistent. In some segments of the landscape the fit may be reasonable but not so in others.

Probably the main drawbacks of using a global fitting method of this sort as interpolator are the poor fits than can be obtained in some regions of the sample area where local variation may be of importance, and the fit in this region may be affected by the fit in other totally unconected regions. Yet, soil data almost invariably adds short range or local variation together with the sampling error to the total or regional pattern. The former may be of considerable importance in prediction, yet trend surface analysis removes these two components for the sake of exposing the general pattern.

In general, though, there is no theory to support the existence of functional relationships between geographic coordinates and soil, (Webster, 1985).

Global fitting methods on their own are found not adequate for spatial interpolation of soil properties unless the variability they attempt to portray is sufficiently smooth. However, they are better used as spatial filters for smoothing or for modelling systematic variations or trends. These can be used for interpolation if accompanied by a method based on the autocovariance functions of the residuals, (see chapter 7).

Thus, from the shortcommings of all the techniques discussed so far and given the nature of soil variability (see chapter 3), it can be concluded that **the fundamental problems in spatial interpolation are two:**

(1). Any successful method of interpolation must be local, so that it accounts for the short range and local variations in soil. This makes necessary the definition of local neighbourhoods, by subdivision of the area, within which an interpolative procedure can operate.

(2). The procedure for interpolation within the local neighbourhood and the criteria for the definition of the neighbourhoods themselves must be based on a sound theory of what is known about soil variability and its structures.

ii. Local Fitting and Interpolation Methods.

To overcome the shortcomings of global fitting methods, surfaces are best fitted locally. The local variaton then can be better described. These methods consist of dividing the sampling area into a number of neighbourhoods represented by contiguous panels or subareas within which either a local polynomial is fitted to the data, or the point at the centroid is estimated by any form of weighting of the neighbouring observations.

In general, in any local estimation two factors connected with (1) and (2) above determine the success of an interpolative procedure, namely:

- The criteria for neighbourhood definition, which decide the size and shape of the "window", and;
- The procedure for interpolation that operates over the

neighbourhood.

Thus, implicitly or explicitly, **the fundamental problem in spatial interpolation has been reduced to finding out the best predictive combination of neighbourhood definition and the interpolative procedure to operate on it.**

A good number of procedures arisen recently have endeavoured to find out the most successful combination of those two factors. Attached to any fitting or interpolative technique there is a particular way of defining the neighbourhood and vice versa, (see chapter 7 for tabular summaries of tried methods for defining neighbourhood and interpolative procedure).

Most local fitting methods, including low degree polynomials, produce some kind of moving average or linear sum. Local polynomials and distance-weighting functions in the form of moving averages have been the most explored techniques for interpolation. Among the most successful local polynomial methods for interpolation are the patchwise bicubic splines which fit surfaces to neighbourhoods or panels.

ii.a Patchwise Bicubic Spline Interpolation.

Bicubic splines are cubic splines in two variables, in this case the coordinates (x, y) on the plane, i.e. constrained cubic polynomials in two dimensions. They fit a smooth surface over each panel or neighbourhood by a low degree polynomial. Each patch or subsurface on a neighbourhood or panel is joined to the neighbouring surfaces, giving continuity and smoothness to the total surface, (see chapter 7 for the mathematical background).

As with the one-dimensional splines, (Ponce-Hernandez et al 1986), the best surface fit is but obtained iteratively. Trials to discover the positioning of the knots (see chapter 7) that gives the best fit have to be carried out. In that sense it is not a direct method as some intermediate fits have to be obtained before the last best one.

The value at a given point is interpolated by substitution, by evaluating the bicubic spline equation which describes the corresponding "patch" or local surface where the point lies, at its coordinates x,y.

Local fitting methods, particularly bicubic splines, are more accurate as interpolators than global fitting methods because they

display the local variation which is generally corrupted or smoothed out by global polynomials.

However, the main drawback of this method is that the definition of the size and form of the panels or neighbourhoods does not obey any law or rule related to the structures of spatial variability of soils. It is to some extent arbitrary and as a general rule these are achieved by positioning the knots close to points of inflection in the data.

It can be shown (see chapter 9), that bicubic spline surfaces are most useful particularly for modelling systematic trends in soil data, and so allowing the spatial analysis of the residuals from the surface and their interpolation by Kriging.

ii.b Moving Averages and Distance-Weighting Functions.

These are procedures for local interpolation. In contrast to polynomial surfaces, moving averages and distance-weighting functions do not fit a local surface but directly estimate values at target points within a shifting window or neighbourhood by weighting or averaging or both. The target points are usually but not always at the centroid of the neighbourhood and this moves position to a new centroid after each calculation, until the whole sampling area is covered.

Moving averages are based on the assumption that the value at a point is somehow related to the values of its surrounding nearest neighbours and that this dependence relation is such that the value at the point can be estimated from them. The form of spatial dependence on each of the surrounding points is implicitly assumed and represented by the "weights" or coefficients affecting each point in the calculation of the average that estimates the centroid. In effect the weights indicate the importance of each neighbour in determining the value of the centroid. A summary of weighting procedures and interpolating functions based on distance is discussed in chapter 7, and presented in tabular form.

Criteria for the assignment of weights are based purely on distance. Tobler (1975), in agreement with what has been stated above, indicates that there are basicaly two problems in local estimation: First, is to decide the neighbourhood. Second, is the choosing of optimal weights for particular purposes.

In relation to the first problem, Burgess (1979) has shown that the most convenient way of defining neighbourhoods for local

estimation is one based on the spatial autocovariance of the property. The analysis of such is called structural analysis and its tool is the semi-variogram (see chapter 7).

Related to the second problem, there is no theory behind distance-weighting functions, as most are developed empirically from arbitrary expressions. They deal with a ratio between distance to a neighbouring point and maximum neighbourhood distance. Therefore, in the choosing of the weighting function the definition of maximum distance, i.e. size, of the neighbourhood is critical. The weight assigned to a point is a monotonic decreasing function of the distance to that point from the point to be interpolated. The weighting function must approach zero asymptotically or nearly so, when the distance to a reference point approaches the specified maximum neighbourhood distance. In these methods the form of the function describing this behaviour is unknown. Its representation is not unique and needs a theory to back it.

There is little experience with the use of weighting functions in soil data even as filters (Ponce-Hernandez, 1975), and experiments to find out the best combinations of neighbourhood size and weights are needed (Webster, 1977).

None of these methods provides satisfactory estimates of the error in the interpolation and therefore Webster (1977) argues that they may not be the best from a purely mathematical point of view.

Clearly, no general rule can be derived from experiments on particular data and point configurations. Consequently the choice of an empirical distance weighting function is more a matter of personal belief, tradition or confidence in the advice of "influential authorities" (Delfiner and Delhome, 1975).

So, it can be concluded so far, that although local fitting and interpolation methods represent an improvement on survey output and the other procedures examined previously, a generally sound theory of soil variability to support techniques for local interpolation is needed.

6.4.4. Optimal Interpolation based on Spatial Autocovariance.

Burgess (1979), Burgess and Webster (1980a and b), and Webster and Burgess (1980) have introduced a range of geostatistical techniques for the spatial analysis of soil variability and interpolation.

These techniques offer an optimal solution to the two fundamental problems in spatial interpolation for local estimation, namely:

i). The sound definition of a neighbourhood based on the structures of spatial autocovariance functions.

ii). Interpolation derived from optimal weighting of the neighbouring points according to their spatial dependence as indicated by the autocovariance function. This type of interpolation is optimal in the sense that it is unbiased and gives values of the minimum estimation variances.

These techniques so far provide the best tools for interpolating soil data and are backed-up by a sound theory of soil spatial variability (i.e. the theory of regionalized variables, see chapter 7). They are just the theory and interpolation techniques that soil scientists were looking for.

Burgess (1979), postulates the semi-variogram as the most useful of the autocovariance functions that describe soil variability in terms of the distance between sampling points. The semi-variogram is unambiguous and is based on the most general of the hypotheses concerning soil variability.

The semi-variogram is a technique for describing spatial dependence and spatial structure in the data. It can be thought of as a measure of the degree of dissimilarity between two places or as half the error of estimating the value at one place from an observation at the other. The parameters of the semi-variogram help in defining convenient neighbourhoods of spatially-dependent points for interpolation, i.e. the range of influence, and in the weighting procedure for interpolation by the Kriging technique. The assignment of weights is based on the spatial relationships, i.e. the form of the increase of the spatial-dependence variance, between the neighbouring points and the point to be interpolated. So, the semi-variogram and kriging together solve problems of neighbourhood defin ition and assignment of weights. The theory

and mathematical formulation of these techniques are fully discussed and elaborated in chapter 7.

Kriging is clearly an advanced technique. It has been shown (Burgess, 1979, Burgess and Webster, 1981a and b, Webster, 1985), that it is superior to any other mathematical technique for spatial interpolation. It is a powerful method for optimal local estimation thus, avoiding the weaknesses of other methods. It is reliable for it is based on a sound theory about the nature of regionalized phenomena which accomodates what is known about soil spatial variability. It also provides measurements of the reliability of its results in the form of estimation variances.

Algorithms for these two techniques were made available to OXSIS by Webster (1985), in the form of FORTRAN 77 computer programs. These have been incorporated as facilities in the OXSIS manipulation package to equip users for predicting soil properties anywhere within a sampled area. These are discussed in chapter 8.

Further, it can be shown (Burgess et al, 1981, and McBratney et al, 1981) that an optimal sampling spacing, i.e. one for which the variance of the estimations is minimum, can be found by semi-variograms and Kriging. Thus, optimal sampling schemes can be designed if the semi-variogram is known in advance from reconnaissance sampling. Hence, the semi-variogram and kriging may also be used as basic tools for planning surveys, by estimating the precision in interpolation that can be expected from a given configuration and sampling effort, and consequently a given survey cost.

Semi-variograms and kriging offer an optimal solution to the problem of missing information in the spatial sense.

However, the predictive success of kriging depends on the semi-variogram and its correct interpretation. Kriging can be used to interpolate values at unvisited sites from values at neighbouring points provided the sampling is sufficiently dense, the property to be interpolated shows spatial dependence, and the data satisfies the intrinsic hypothesis (see chapter 7).

The feasibility of calculating a semi-variogram, and of interpolating by kriging depend strongly on the amount of point-data available or, more precisely, on the sampling density at the scale of sampling. This is their most limiting factor, for in ordinary survey practice sampling densities contribute

substantially to the survey cost, hence they tend to be low and point-data sparse.

Further, it is not clear down to what survey scale kriging can still be applied with confidence. The contribution of intuition in the form of mapping units and classes to kriging interpolation, particularly in large areas surveyed at small scales, and the benefits of their combined use have not been fully explored.

There are a number of situations in which kriging cannot be directly applied. For instance, where the semi-variogram shows no spatial dependence between data-point values and all variation can be said to be random noise. Also, when the property is non-stationary, or when point-data are too sparse to allow the calculation of reliable semi-variograms. In circumstances like these, since soil maps and classes are the common products of any survey and are abundant, they may help mathematical interpolation or even provide better interpolated values.

Hence, the predictive power of kriging techniques is not yet fully incorporated into conventional soil survey practice. It needs first, to be assessed against the merits of the conventional soil survey output, i.e. soil classification, map and legend, for providing the missing soil information essential for land evaluation and land-use planning.

Thus, many aspects of the joint use of conventional soil survey output, semi-variograms and kriging need to be explored to improve interpolation and prediction of soil properties.

An immediate use of soil maps and classes in an interpolation exercise is as a means to stratify long term variability in the form of trends. Trends need to be removed by partitioning, so that the stationarity assumptions necessary for kriging (see chapter 7) can be met and interpolation by kriging carried out with confidence.

Moreover, the effect of using divisions of the landscape in the form of soil or land mapping units on kriging estimations can be compared with mathematical procedures such as bicubic spline surfaces for modelling the trend (see chapter 7). The effect of removing these trends, either by mapping units or by spline surfaces on the structures of the semi-variogram as a reconnaissance tool and on prediction, needs to be assessed and compared too.

With these problems in mind, trials in the prediction of soil properties were carried out in an area near Oxford using the wealth of data already collected from several grid surveys. The object of the trials was to compare, first, the predictive value of kriging against that of soil classes and mapping units. Second to explore their combined use and that of other analytical procedures, i.e. bicubic spline surfaces, for modelling trends and their effect on prediction. This trials are the object of chapter 9.

6.5 Concluding remarks to this chapter.

i. Only a minimal part of the total soil information of an area can be provided by soil survey.
ii. Information at unvisited sites, unsampled depths, and for unrecorded properties is missing information.
iii. Prediction of soil information requires procedures for estimation of properties, and depth and spatial interpolations.
iv. The values of missing soil properties can be estimated by a function on recorded properties (a "surrogate function").
v. There is a sufficient number of statistical procedures and criteria to construct a surrogate function. What is needed is a strategy for their best use to tackle a variety of problems during the model development stage.
vi. The problem of depth interpolation is solved by furnishing a suitable depth function to model the vertical distribution of soil properties and to be used as interpolator.
vii. Local methods are to be preferred to global methods for spatial interpolation.
viii. The two fundamental problems in spatial interpolation are the definition of a neighbourhood that makes sense to the spatial structures of soil variability, and the finding of an optimal weighting procedure to operate over the neighbourhood based on such structures.
ix. The semi-variogram and kriging are the best techniques for spatial analysis and interpolation of soil properties.

x. The predictive value of kriging needs to be compared against that of conventional soil survey output.

xi. Many aspects of the complementary use of kriging and soil classes and maps, and other analytical procedures for modelling soil variability demand investigation to improve procedures for predicting missing soil information.

VII. MATHEMATICAL INTERPOLATION OF SOIL DATA.

Land-use planning needs soil information but most of it is missing and has to be generated or predicted. Chapter 6 has explained the problem of missing information and reviewed special problems in providing such information by estimation and interpolation. This chapter examines methods for estimating the surrogate function, but particularly mathematical techniques for spatial and depth interpolation emphasizing relatively new geostatistical techniques, e.g. semi-variograms and kriging.

The techniques discussed in this chapter provide the theoretical background to the algorithms that are part of the OXSIS Manipulation package presented in chapter 8. These algorithms furnish the necessary tools for the trials in prediction of missing data presented in chapter 9.

7.1 Surrogate Functions.

The aim in developing predictive models is to construct quantitative functions as predictors to tackle problem (i) in chapter 6, i.e. soil property missing. This is not a problem of mathematical interpolation but rather one of generating missing data from existing information on different properties. However, due to its quantitative nature it is included in this chapter.

The starting point is that records for other properties exist within the area of interest. These are ungeneralized or "hard" point-data and may or may not be accompanied by another types of generalized information such as the soil mapping units and soil classes at the sites.

It is assumed that information is available for the sites and depths of interest, but not for the property of interest. Hence, no spatial or depth interpolation is needed and the problem reduces itself to finding a suitable function of the missing property on non-missing properties by multiple linear or nonlinear regressions based on least squares.

The strategy for developing a predictive model can be divided into three stages:

Stage 1. Predevelopment or planning.

" 2. Development.

" 3. Validation and Maintenance.

The flow chart of fig. 7.1 shows graphically these three stages of a general plan for development of predictive models.

7.1.1 Planning or pre-development stage.

The first stage includes the phases of problem definition and goals, data completion and organization, selection of potential predictors, data quality checks, transformations and scaling of data if applicable, the checking of computer and software facilities and obtaining the matrix of correlation coefficients, scatter diagrams and descriptive statistics of the response and potential predictors. First runs of preliminary regressions take place also at this stage, to narrow down the potential variables and models.

In order to develop a usable predictive model the analyst should aim at meeting the following goals:

i. The final equation should explain at least 80% of the total variation of the response (R-square > 0.80).

ii. The standard error of the estimates should be less than 5% of the average of the X's used in the model.

iii. It would be desirable that all the estimated coefficients in the final model were statistically significant at the level of alpha = 0.05, but this is not essential provided the model is shown to be a good predictor.

iv. There should be no discernible patterns in the residuals.

v. The number of predictor variables should be in agreement with the number of observations used to calculate the regression. As a general rule there should be about ten complete sets for each of the potential predictor variables to be included in the model. Thus, for a model with four predictors plus a constant, the number of observations should be at

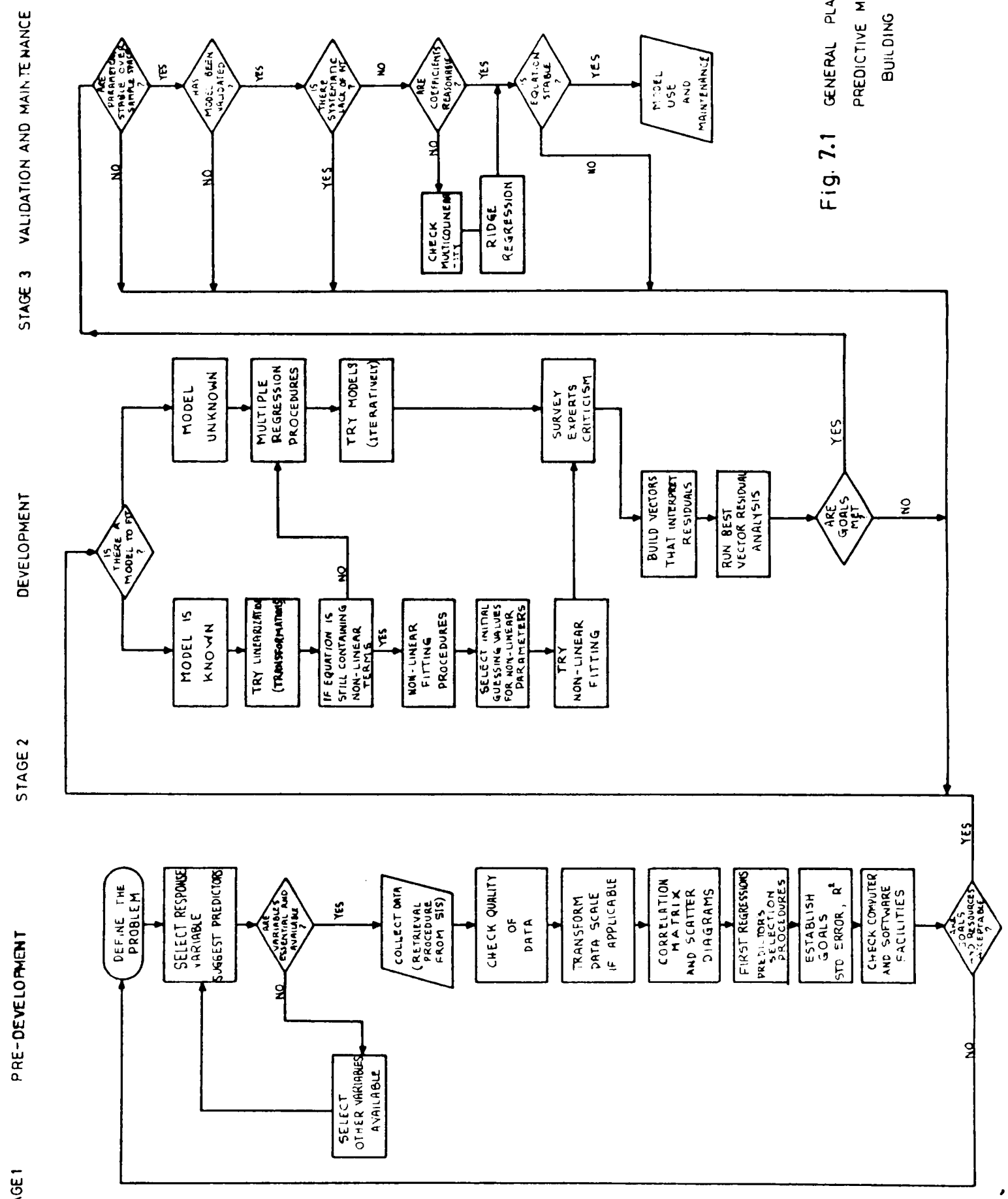


Fig. 7.1 GENERAL PLAN FOR PREDICTIVE MODEL BUILDING

least forty (n=40).

7.1.2 Model development stage.

The strategy in the development stage depends directly on what is known about the model. There are two possible situations:

a). The form of the model is known.

b). The form of the model is unknown.

Situation (b) is the most common. Occasionally for instance, it may be known which are the predictor variables that should be in the model but the nature of their association and exponents could be unknown. For convenience this case should be treated as an advanced stage of situation (b) above.

7.1.2.1 Surrogate function when the model is known.

When theory or prior knowledge suggests that a particular function links the wanted variable to the known predictor variables then it may be best to fit the function to the data to determine the appropriate constant terms. Nonlinear estimation is used to determine the parameters required by the known equation. In effect the analyst "fits the observations" to a known curve.

Most known models tend to be of exponential form, and some of them highly complex. This type of model can only be fitted to data by nonlinear least squares. Given complete knowledge of the model it is better to fit the nonlinear model rather than a less realistic linear polynomial. A model of this sort may be known through a full description of the terms of the "true" equation, or through a set of differential equations which the model must satisfy.

However, there are models susceptible to linearization after transformation. These are intrinsically linear. The intrinsically nonlinear models are impossible to convert into a linear form even by transformation. In a practical situation and particularly with very complex models the analyst will attempt to simplify the equation and to linearize it as much as possible by transforming the variables. If the model becomes linear after transformation, multiple regression procedures may be used to estimate the parameters. If the model still contains some nonlinear terms, nonlinear fitting procedures are best applied.

It is worth noticing that there is an important drawback in the use of transformations in that they may lead to biased

estimates. Furthermore, the R^2 found after the fitting will account only for the parameters in the transformed model. Consequently, when desirable and available the analyst should prefer maximum likelihood methods instead.

The problem of building a surrogate function from nonlinear least squares procedures is presented in mathematical notation in appendix 7.1.

7.1.2.2 The surrogate function when the model is unknown.

Constructing a surrogate function when there is no known model is the most common situation in practice.

In contrast with situation 7.1.2.1, in this case the user has to select the variables, their right powers, interactions between such variables and the right powers of those interactions, as well as the general form of the equation in which all those factors are related. Thus, the range of possibilities is enormous.

Any model is only an approximation to an ideal function. Hence, the approximate model and its predicted values will be found in error. Two factors contribute to that error:

- The model has too few independent variables so that the effect of the other variables not included appears as residual term or error.
- The fitted function is of the wrong form with terms of the wrong power.

The first factor is well known in model fitting by multiple regression. The second tends to be forgotten since analysts usually tend to fit additive polynomials, i.e. a simple addition of simple terms.

These two factors and particularly the latter must be considered during model development. Thus, given the large number of possible combinations of independent variables, their interactions, their powers and the powers of the interactions a model building strategy should be adopted based on some premises:

i. A model for the surrogate function can be said to be acceptable if it consistently predicts values of the independent variable within conveniently pre-established limits.

ii. The emphasis in building the model is on predictive power. Relationships to explain causality take lower

priority. In this sense the model can be empirical rather than explanatory, but must meet at least the targeted R^2 in the goals as above.

iii. The final model must be as simple as possible without compromising its quality as a good predictor.

iv. As far as possible the model must be based on common knowledge about the behaviour of the property of interest with respect to the independent variables.

A great part of the effort on model development is spent on selecting the best predictor variables, powers and interactions. To narrow down the number of potential predictors a first selection should be made by matching variables suggested in published work and experience elsewhere to the data available. Further selection from this group relies on extensive computing effort using statistically-based procedures for variable selection available in most software packages for regression analysis.

An extensive examination of scatter diagrams and correlation coefficients should be made to guide the first choice of variables and exponents to enter any standard procedure for variable selection. Ideally, the potential predictors should be highly correlated with the response but not correlated or poorly correlated with each other.

Depending on the selection procedure the improvement on the fit at each step is evaluated by R^2 so that at the end only the most significant variables are considered for the model. Trials for the powers of the terms in the model can be run concurrently. The starting point is of course the simplest case, the additive model with multiple terms.

When sufficient data exist some data may be excluded from these runs to be used later for checks on the predictive capacity of the final model.

There are a number of standard procedures for variables selection, namely:

i. **All possible regressions approach.**

ii. **Backward elimination procedure.**

iii. **Stepwise elimination procedure.**

iv. **"Best k" subset of regressions.**

v. **PRESS,(PREdictive Sum of Squares) procedure.**

vi. Latent Root regression.

Ridge regression (Hoerl and Kennard, 1970) is offered as a technique for overcoming the problem of multicolinearity, i.e. the presence of highly intercorrelated predictor variables. **Principal Component regression** (Hotelling, 1933) offers an alternative to Ridge regression in that it is a procedure that analyzes the correlation structures in some detail. An extension of principal component regression is the Latent Root regression that can be used as a selection procedure in combination with backward elimination.

The stepwise elimination procedure is a suitable technique for the first rounds of predictor selection. When the number of variables has been reduced to a consistent set then the PRESS technique should be used to ensure that only the variables which improve the predictive capacity of the model as determined by the predictive sum of squares, will remain in the final equation.

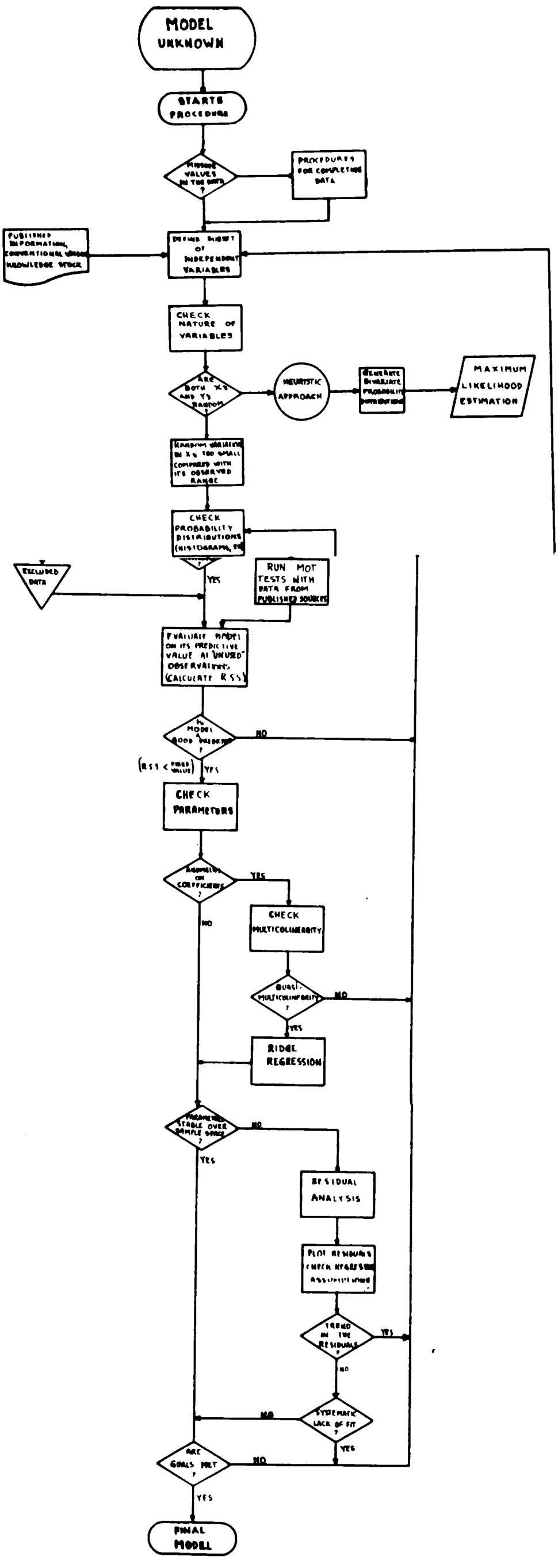
Thus in practice a **combination of stepwise and PRESS regression techniques is recommended for constructing the surrogate function when the model is unknown and selection of predictor variables is necessary.**

At the end of this stage only a few model proposals enter the stage of testing. The sequence for model development here suggested, which includes the procedures for constructing a surrogate function when the model is unknown is shown in the form of a flow chart in figure 7.2.

7.1.3. Model validation.

The testing of the proposed models involves checks on various aspects. The most important of them is the consistency of their predictive success. The model must be checked for other anomalies too, for instance the sign on the regression coefficients which may be an indication of multicolinearity. The residuals for each candidate model must also be checked. Regression analysis is based on initial assumptions concerning these errors,i.e. that they are independent, normally distributed and that they have zero mean and a constant variance. This last assumption is required to make valid F-tests. Thus, if the candidate model in turn is

Fig 7.2 **STAGE 2** MODEL DEVELOPMENT



correct its residuals should tend towards confirming such assumptions.

The purpose of the analysis of residuals is threefold; they allow checks on the assumtions of the model, and to diagnose the existence of "lack of fit" due to systematic trends in residuals and they illustrate the behaviour of the residuals of predictors not yet included in the model and which could be included.

The analysis of residuals can be applied to any situation where the fitting of a model has left some degree of unexplained variation. The analysis can be graphical and/or analytical. The overall plot of residuals should resemble a normal distribution of observations with zero mean. The plot may exhibit slight irregularities but should not appear abnormal for a sample of a size of the observations. To judge, a plot with values from a table of random normal deviates can be compared against the residuals plot.

Residuals can also be examined by plotting them against the predicted values of the response (Y). A satisfactory residuals plot should give the general impression of a horizontal band asymptotic to the axis of the predicted response, indicating a fair degree of independence in the observations. This plot also helps to check if the variance is constant, as assumed, and in general on the adequacy of the model. Unsatisfactory residual behaviour is characterized by trends in the scatter of residuals or by any form other than the horizontal band.

Plots of the residuals can also be drawn against the independent variables Xji, i=1,2,....n. This plot would point out also if the variance is not constant and if there is a need for extra terms in the model, for example a quadratic term or an interaction.

If there are repeat runs in the data, (multiple observations on Y for the same value of X), or a prior estimate of the variation is available. The candidate model can be tested for lack of fit by splitting the residual sum of squares (SS) into SS due to "lack of fit", and SS due to "pure error", and conducting F tests. If lack of fit is significant the model will have to be reconsidered for it is inadequate in its present form. Tests of this sort are part of the validation of the model and are included in the proposed flow chart of figure 7.1

The checks suggested correspond to stage 3, model validation, in the flow chart of the general plan for predictive model building in figure 7.1.

To illustrate the points discussed above, an exercise on constructing a surrogate function to estimate Cation Exchange Capacity (CEC) from various other "surrogate" soil properties was carried out. A limited number of observations (n=20) from Oxfordshire soils was used. The statistical package GENSTAT, (Lawes Agricultural Trust, Rothamsted,1983), was used for this purpose. Appendix 7.2 shows the printouts with the results.

7.2 Soil Depth Functions.

The problems of missing information and interpolation down the soil depth have been discussed in section 6.1.ii of the previous chapter. The general situation was that information was available on the wanted property at the right site, but at the wrong depths.

The interpolation problem consists of reconstructing the "true" smooth function as closely as possible and using such a function to provide data at the required depths.

Thus, given a set of values of Xi (depth), Yi (property values), Y = f(X) is the true depth function. An interpolator is any function Y = Pn(X) that maps a set of real numbers which closely approximates to Y = f(X), and provides estimates of Y within certain bounds at any point within the X space.

An interpolator can be a graphical or a mathematical representation of the approximation. Graphical interpolators can be of three forms.

(i). Free-hand drawing of smooth curves through the data points.

This assumes the form and smoothness of the curves drawn through data points and is determined to a great extent by intuition. The interpolated values can be no more than gross approximations.

(ii). Joining data points with straight lines.

Interpolation takes place as above. The interpolated values inevitably contain errors for there is nothing to back the assumption that changes between points are linear.

(iii). Drawing a "stepped" curve.

This assumes that each data point is the average value over the whole depth interval that it represents. Interpolation occurs by assigning the same average value to any depth lieing in the depth interval.

Graphical interpolators have all the shortcomings characteristic of any arbitrary procedure so their use should be restricted to cases where precision is not too important.

The procedures for mathematical interpolation can be grouped in the following categories:

i. Global polynomials.

ii. Local or piecewise polynomials.

iii. Exponentials and other non-linear functions.

Polynomials, particularly local polynomials are the most suitable as interpolators for they uniformly approximate continuous functions. They are defined as a set of functions of the form:

$$P_n(X) = a_0 + a_1 X_1 + \ldots\ldots a_n X^n \qquad (7.1)$$

where n is a non-negative integer and $a_i \ldots . a_n$ are real constants.

If f(X) is defined as being continuous on a depth interval [a,b] and error bounds are fixed such that e>0, then there exists a polynomial P defined on [a,b] with the property that

$$[f(X) - P(X)] < e \text{ , for all } X \text{ in } [a,b] \qquad (7.2)$$

That is, given any function that is defined continuous on a closed interval there is a polynomial that approximates as close as possible to the function desired as permited by the error bounds.

Some of the limitations of global polynomials have been discussed in the previous chapter in section 6.2.

An alternative approach to global polynomials consists of dividing the depth interval [a,b] into a collection of subintervals and constructing a different polynomial for each subinterval. Functions of this type are piecewise polynomials.

It has been shown (Webster, 1981), that amongst piecewise polynomials, spline functions possess many advantages over any other procedures for approximating depth functions and can be used as interpolators.

7.2.1 Cubic Splines.

Spline functions are piecewise polynomials of order n. The pieces join in the so called knots and fulfil continuity conditions for the function itself and for the first n-1 derivatives. The term "spline function " is intended to suggest that the plot of such a function is similar to the curve drawn by a mechanical spline, i.e. a strip or rod of some flexible material constrained by weights to pass through points on a graph. Indeed, it has been shown that to a first order of approximation the curve produced by a mechanical spline is a cubic, i.e. third degree spline function.

So, given the function f(X) defined on the profile´s depth with limits [a,b], and a set of numbers, called knots such that $a=K_0<K_1<K_2<....<k_n=b$, a cubic spline interpolant, S, for f(X) is a cubic polynomial

$$Sj(X) = a_j + b_j(X_{j+1} - X_j) + c_j(X_{j+1} - X_j)^2 + d_j(X_{j+1} - X_j)^3 \qquad (7.3)$$

wich satisfies the following conditions:

(7.3.1) S is a cubic polynomial, denoted Sj , on the sub-interval $[X_j, X_{j+1}]$, for each j = 0,1,...n

(7.3.2) $S(X_j) = f(X_j)$ for each j= 0,1, ...n-2;

(7.3.3) $S_{j+1}(X_{j+1}) = S_j(X_{j+1})$ for each j = 0,1,..n-2

(7.3.4) $S'_{j+1}(X_{j+1}) = S'_j(X_{j+1})$ for each j=0,1,..n-2

(7.3.5) $S''_{j+1}(X_{j+1}) = S''_j(X_{j+1})$ for each j=0,1,..n-2

(7.3.6) Boundary conditions: (one of the following set must be satisfied).

$S''(X_0) = S''(Xn) = 0$ (free boundary).

$S'(X_0) = f'(X_0)$ and $S'(Xn) = f'(Xn)$ (clamped)

It can be seen from conditions (7.3.4) and (7.3.5), that S(X) and its first and second derivatives are continuous so the set of cubic polynomial arcs are joined smoothly end to end. The smoothness is provided by continuity up to the second derivative, but the third derivative may have a discontinuity at each of the points X = Ki, i = 1,2,...m.

Cubic splines are piecewise locally constructed functions with continuity and smoothness. These two properties permit a fairly accurate reconstruction of the distribution of soil properties with depth.

Usually, when curve-fitting, the number of knots, m, and their positions Kj ; j = 1,2, ...m, are the user's choice. These define the number and size of the set of sub-intervals and the number of splines required to fit them. Inevitably, this has a direct effect on the overall fit of the curve. The number and positon of knots is critical. Wold (1974) compares it to the selection of functional type in an ordinary curve-fitting problem and suggests that knots should be chosen to correspond to the general behaviour of the data, i.e. number of points, inflexions, maxima and minima. He suggests rules of thumb applicable where there are fewer than 30 observations, which is the case for most soil profile data. The rules are:

i. Have as few knots as possible and at least 4 or 5 points per interval, since the great flexibility of splines can make over-fitting a problem.

ii. Have not more than one extremum point and one inflexion point per interval (a cubic could not approximate more variations).

iii. Place knots close to inflexion points.

iv. Have extrema centred on the intervals.

Knot positioning for a good fit is then a matter of trial and error. This may raise serious difficulties for a non-specialist user who may want to process a large number of profiles by an automated spline-fitting routine as a part of a computerised SIS.

The cubic splines can be made to pass exactly through the data points by imposing further restrictions but exact splines, due to their flexibility, often tend to overfit the data. Webster (1978) indicates that in these circumstances a model that approximates to the data points may be better than one that fits exactly. He

fitted least squares cubic splines to profile data resulting in a better representation of the general behaviour of the depth function than with an exact spline fit which may be describing information and "noise" together. Wold (1974) would limit exact splines to sets of not less than 10-15 points with relatively uniform distribution. Soil profile data seldom meet such conditions.

Cubic splines are one of the best current procedures for producing a smooth curve over the whole profile which can be used as an interpolator. However, there is still room for ad-hoc improvements in the conventional cubic spline method to allow for even better reconstructions of soil depth functions. These were suggested in the previous chapter. It was concluded from there that two possible ways of improving the reconstruction of smooth profiles can be explored:

i). Modify the procedure for sampling soil profiles so that it suits best the cubic spline fitting, or:

ii). Modify the curve fitting procedure to take account of the oversmoothing produced by the bulking or averaging that occurs when horizons or layers are sampled the conventional way.

Ponce-Hernandez et al (1986) have attempted both. Appendix 7.4 includes the details of their study. They compared the efficacy of procedures for reconstructing smooth original profiles of varied complexity. It was found that there is little advantage in collecting point-depth samples rather than whole-horizon samples in order to reconstruct the profile more precisely.

On the other hand, and as a contribution to the solution of the depth interpolation problem, this thesis proposes a modified spline as the best procedure for reconstructing true profiles from the horizon mean values produced by conventional sampling. The modified or "equal-area" spline function solves the problem of averaging created by sampling the conventional way, i.e. whole-horizon (see chapter 6). It produces the curve of fig. 6.1 f, and has been found to give better reconstructions of true profiles from horizon mean values than the ordinary cubic spline (Ponce-Hernandez et al, 1986). Further, the procedure is direct in that it accomplishes the final fitting in one stage without the successive approximations required by ordinary cubic splines,

since it does not require the user to suggest knots or any other parameters as do the ordinary splines.

7.2.2 Equal-Area Splines.

The modified equal-area spline procedure fits a spline curve to the mean values of each of the sequence of horizons that comprise the whole soil profile. Each horizon defines an interval and consequently the knots are placed exactly on the inter-horizon boundaries; X_i, X_{i+1}, ..to..X_{n-1}.

The spline on each interval must meet the equal-area condition. That is, that the area under the fitted curve in each interval equals to the area of the rectangle defined by the interval, i.e. the horizon mean represented by the values Vi in figure 7.3.

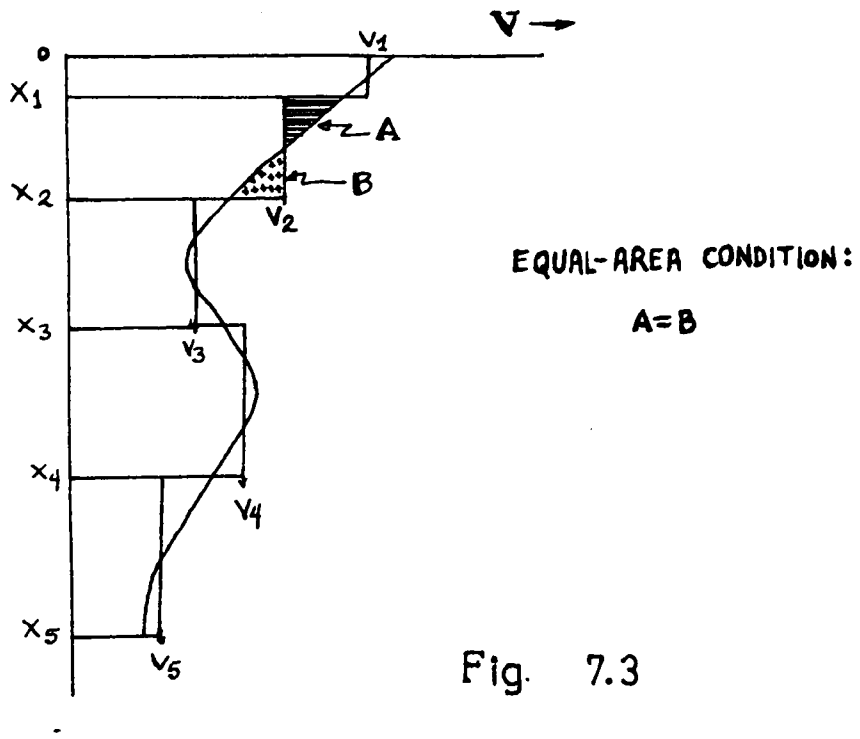


Fig. 7.3

So we can write

$$V_1(X_1 - X_0) = \int_{x_0}^{x_1} f(X)d(X)$$

$$V_2(X_2 - X_1) = \int_{x_1}^{x_2} f(X)d(X)$$

in general:

$$V_n(X_n - X_{n-1}) = \int_{x_{n-1}}^{x_n} f(X)d(X) \qquad (7.4)$$

Thus the change in the value of the variable within interval i described by the curve is represented by an horizon mean Vi.

A cubic polynomial can be used to describe the area within each interval. The equal-area condition is then expressed as

$$V_i(X_i-X_{i-1}) = a_i + b_iX + c_iX^2 + d_iX^3 \tag{7.5}$$

then we can write

$$\int_{x_{i-1}}^{x_i} f(X)d(X) = a_i + b_iX + c_iX^2 + d_iX^3 \tag{7.6}$$

Hence, the spline function within each interval is described by a quadratic polynomial

$$f(X) = b_i + 2c_iX + 3d_iX^2 \tag{7.7}$$

Thus, if we chose an interval defined by the horizon limits X_i and X_{i+1}, the equal-area condition is

$$b_i(X_{i+1}-X_i) + c_i(X_{i+1}^2-X_i^2) + d_i(X_{i+1}^3-X_i^3) = V_i(X_{i+1}-X_i) \tag{7.8}$$

The continuity condition is

$$b_i + 2c_iX_{i+1} + 3d_iX_{i+1}^2 = b_{i+1} + 2c_{i+1}X_{i+1} + 3d_{i+1}X_{i+1}^2 \tag{7.9}$$

hence the curve has no breaks and has a continuous derivative

$$c_i + 3d_iX_{i+1} = c_{i+1} + 3d_{i+1}X_{i+1} \tag{7.10}$$

Now, if there are n horizons (i.e. intervals) in the profile with limits $0,X_1,X_2,\ldots X_n$; there are n equations of the first type, and n-1 of the second and third type, i.e. the continuity conditions. There are thus 3n-2 equations in 3n unknowns. By fitting two further end-conditions these equations may be solved. Notice that if cubic polynomials were fitted to each interval instead of quadratics this would have yielded 4n unknowns and only 4n-3 equations which would require 3 extra values. The two extra values will have to be guessed. For convenience in computation these values are calculated at the top and the bottom of the profile and thus inserted. Ponce-Hernandez

et al (1986) found good overall fits by using two suggested procedures to calculate top and bottom values. These are explained in detail in appendix 7.4

The algorithm for fitting equal-area splines to data collected from profile horizons consists of solving a system of equations to calculate the coefficients b, c, and d for each piecewise polynomial. The algebraic solution is shown in appendix 7.3. A set of computer routines has been written in Fortran 77 to calculate equal-area splines and to fit them to profile data. These algorithms are part of the facilities of the Manipulation procedure of OXSIS, (see chapter 8).

It has been shown that the algorithm fits a continuous smooth spline-like curve piecewise through the whole profile depth, but it consists of a set of separate spline quadratic polynomials for each interval or profile horizon.

To be used as interpolator, this depth function has to be evaluated. Thus, the depth interval or horizon within which a value is to be interpolated is first ascertained. The coefficients of the spline for that interval are then used to interpolate the value of the property at the specified depth.

The curve can also be evaluated in the opposite sense. Given a soil property value it is easy to find the depth at which it occurs, by a similar procedure. Some other ways of evaluating the depth function to answer practical questions are feasible and have been explored.

The equal-area spline algorithm has been implemented in OXSIS as a part of a software package for depth interpolation, (see chapter 8). Such a package offers five ways of interpolating data based on depth functions from equal-area spline reconstructions.

The full description of the comparisons and their results can be seen in appendix 7.4. On the whole, it was found that the equal-area spline produced a better reconstruction of true profiles from horizon mean values than the conventional cubic spline. In regards to the sampling, it was found that collecting point-depth samples instead of whole horizon samples represented little advantage in reconstructing the true profile.

7.3 Spatial Interpolation of soil data.

It has been seen from chapter 6 that the problem of missing information is present in practically any soil data-set. The soil surveyor attempts to generate or predict the missing information from available observations.

Spatial interpolation is the process of estimating the variable Z at some point P(X,Y) from the known values of Z at a number of data points, which may be arranged systematically or irregularly around it.

The methods described next have been briefly outlined in chapter 6. Their mathematical formulation concerns this chapter. However, since it was concluded that semi-variograms and kriging are the best techniques for spatial analysis and interpolation of soil properties, the chapter emphasizes their formulation and discusses their various aspects concerning prediction of missing information.

7.3.1. Thiessen Polygons.

Thiessen polygons are a subdivision of the plane determined by a finite set of distinctive points. Each point has associated with it the region of the plane nearer to that point than to any other. Such region is convex and polygonal and represents in practice the "area of influence" of the given point. So the value of any given property or the class at any unobserved point can be drawn from the value or the class of the nearest observed point within the same polygonal region.

Mathematically the polygons can be defined in terms of sets of points (Green and Sibson, 1977). So, let P_1, P_2,.... P_N be a finite number of points in the plane, no two of which coincide. A polygon (i.e. tile) of P_n is the set T_n defined by

$$Tn = \{x : d(x,P_n) < d(x,P_m) \text{ for all } m = n\} \qquad (7.11)$$

where d is the Euclidean distance of point x to its nearest neighbours P_n and to the other points P_m. Each polygon T_n is the intersection of the open half planes bounded by the

perpendicular bisectors of the lines joining P_n with each of the other P_m.

Every point of the plane with a unique nearest neighbour among $P_1,...P_n$ will lie in the tile of that nearest neighbour.

This technique has found application in the construction of probabilistic models of spatial phenomena in ecological studies of plant and animal communities and in interpolation of climatic data. Its application to soil data has been discussed in chapter 6, where its main role is seen as a method for defining interpolative neighbourhoods from observed points.

7.3.2 Interpolation by Smooth Surfaces.

The common underlying principle of smooth surfaces is to let the values of soil property Z at point P(x,y) on the plane be a function exclusively of the location given by its geographical coordinates (x,y). The interpolation problem is reduced to finding

$$Z\ (X_i,Y_j) = f\{\ X_i,Y_j\} \qquad (7.12)$$

were f{ Xi,Yj } is a function on the plane such that

$$[Z(X_i,Y_j) - \hat{Z}(X_i,Y_j)]^2 = \text{min ; for all i,j.} \qquad (7.13)$$

So

$$E\{[Z(X_i,Y_j) - \hat{Z}(X_i,Y_j)]^2\} = 0 \qquad (7.14)$$

where $\hat{Z}$ are points on the fitted surface at locations (Xi,Yj) and Z are the observed values at the same locations.

According to how the fitting is done these techniques can be divided into (a). global fitting methods, and
(b). local fitting methods.

7.3.2.1 Global fitting methods.

These procedures fit a common surface to the whole space covered by the observed points.

Trend surface analysis is the best known in this group.

i. Trend Surface Analysis.

This technique assumes that the systematic component or trend m(x,y), is sufficient to describe the variations of soil across the landscape, and that the random variation or error e (x,y) is unimportant and can be ignored. Thus, a value of Z (X,Y) is defined by

$$Z\ (X,Y) = m\ (x,y) + e\ (x,y) \qquad (7.15)$$

Trend surface analysis uses a series of filtering processes to identify the function Z = f (x,y) that represents the essential features of the regional variation pattern of Z, i.e. the trend, and as much as possible of the random noise included in the observations.

The trend can then be represented by a global algebraic polynomial giving the surface:

$$Z = a_0+a_1x+a_2y+a_3x^2+a_4xy+a_5y^2+..+a_nx^n+..+ e \qquad (7.16)$$

where the ai are the coefficients estimated by least squares, e is the random variation and x and y are the geographical coordinates of the point with the value of the property Z.

Surfaces of this sort filter or smooth out much noise mixed with information without knowing which is which.

The drawbacks of trend surface analysis as interpolator of soil properties have been discussed in chapter 6. It is clear that this is not the best alternative method for the task of predicting the soil information missing.

7.3.2.2 Local Fitting Methods.

To account for the local variations in soil, methods that fit the data locally are necessary. For this purpose the area has to be divided into local neighbourhoods or panels, and a mathematical procedure is made to operate on them to fit either a local surface or to estimate directly a value at the centroid by any weighting procedure.

Implicitly or explicitly most recent efforts in spatial interpolation have endeavoured to find the best predictive

combination of neighbourhood definition and interpolative method to operate on it. A summary of these procedures is given in tables 7.1 and 7.2.

Table 7.2 shows that most local fitting methods are some kind of moving average or linear sum, including low degree polynomials. Local polynomials have been fitted by least squares to sliding windows using constrained and unconstrained methods. These have found use in digital terrain models, (Schut, 1976).

It was established in chapter 6 that constrained patchwise polynomials in the form of bicubic splines are among the most successful local polynomial methods. Smooth surfaces are produced by fitting bicubic spline polynomials to the neighbourhoods or panels.

i. Patchwise Bicubic Spline Interpolation.

Bicubic splines are cubic splines on two variables on the plane.

The limits of the total area defined on x and y co-ordinates are a, b, c, d, such that $a < x < b$; and $c < y < d$ in the plane. The ranges [a,b] and [c,d] are subdivided by sets of knots ki and li respectively.

The points (ki,lj) partition the whole area R into rectangular panels or neighbourhoods, Ri,j, i=1,2,..h+1 and j=1,2..q+1, such that

$$a = k_0 < k_1 < \ldots < K_{h+1} = b \; ; \qquad (7.17)$$

and

$$c = l_0 < l_1 < \ldots < l_{q+1} = d \; , \qquad (7.18)$$

and Rij is the panel $ki-1 < x < ki$, $lj-1 < y < lj$. It can be shown that a bicubic spline on R with respect to the given mesh is a polynomial function of the form (Ahlberg, 1967):

$$S(x,y) = \sum_{s=0}^{3} \sum_{t=0}^{3} Bs,t \; X^s Y^t \qquad (7.19)$$

in each panel Rij, where B are the set of spline coefficients
This function has continuity up to the second partial derivatives everywhere in R.

Table 7.1 CRITERIA AND METHODS FOR NEIGHBOURHOOD DEFINITION

Determining Factor	Definition	Form or Method	Neighbourhood Limits
Neighbourhood Shape & Size	Arbitrary	Centroids at Target Points	(a) To include a Fixed Number of Neighbouring Points
			(b) By Overlay of a Regular Grid
	By Goodness of Fit	Panels of Various Sizes	(c) Guided by Inflation Points and Extrema in the Global Surface
	Based on Spatial Auto-correlation	Quasi-Circular Isotropic Regions Given by Range of Influence	(d) Given by Distance Equal to Range of Spatial Influence

Table 7.2. DISTANCE-WEIGHTING FUNCTIONS AND DISTANCE-BASED EQUATIONS FOR INTERPOLATION.

Determining Factor	Procedure	Method	Weighting	Variants
Procedures for local Interpolation	Triangulation	Linear Interpolation between vertices of triangle		
	Moving Surface Methods	Moving Averages	Unweighted	(I)
			Distance weighting functions	(II) Distance-weighted averages to eliminate systematic trend
		Local Polynomials fitted to a moving window by least squares	Unweighted	(III) (- 12 and 16 term bicubic polynomials; (- 8 term biquadratic polynomial; (- 4 term bilinear polynomial
			Distance weighting functions	(IV)
	Simultaneous patchwise polynomials	Bicubic splines	Unweighted	(V) Least-squares bicubic splines
			Weighted on the observations	(VI)
	Local interpolation based on spatial auto-coveriance	Kriging: point, block, co-kriging	By optimization based on auto-covariance function	(VII)

The bicubic spline can be obtained from the tensor product of two one-dimensional cubic spline sets (de Boor,1962). The cross-product of one function from each set provides the basis for the bicubic spline. These can be represented as sets of B-splines (Hayes and Halliday, 1974). Let Mi(x) be the B-spline based on the knots ku and the corresponding Nj(y) based on the knots lv. The bicubic spline can be represented (Hayes and Halliday, 1974) by the form

$$S(x,y) = \sum_{i=1}^{h+4} \sum_{j=1}^{k+4} \gamma_{ij}\, Mi(x)\, Nj(y). \qquad (7.20)$$

where $\gamma_{i,j}$ are the B-spline coefficients. The set of knots has been extended outside the ends of each range so that the B-splines can be fully defined within them. Thus

$$u < i < u+3 ; \qquad v < j < v+3$$

are the set of external knots that need to be inserted.

A stable algorithm for a least squares fitting and evaluation of the B-splines that make the bicubic surface is given by Hayes and Halliday (1974), and has been used in this thesis (see chapter 9). The surface is not constrained to pass through the data values.

Each patchwise surface on a panel is joined to the neighbouring surfaces. Continuity in the first and second derivatives gives continuity and smoothness to the total surface.

As with the one-dimensional splines, the best surface fit is best obtained iteratively. Trials to discover the positioning of the knots that gives the best fit have to be carried out. In that sense it is not a direct method as some intermediate fits have to be obtained before the last best one.

The value at a given point may then be interpolated by evaluating, at the cordinates x,y, the equation of the bicubic spline subsurface that contains the point of interest located at those co-ordinates.

The merits of bicubic splines as a method for interpolation have been discussed in chapter 6. This technique is used in this thesis for the modelling of trends (see chapter 9), and combined

with structural analysis and Kriging for the interpolation of soil properties.

ii. Moving Averages and Distance-weighting functions.

These are procedures for local interpolation. In contrast to polynomial surfaces, moving averages and distance-weighting functions do not fit a local surface but estimate directly at a target point, by weighting or by averaging or both. In that sense they could be considered as point-fitting methods.

The estimation is based on a quantitative expression of the spatial relationship between the target point and its neighbours. The true form of such relationship is generally unknown, but it is assummed and left implicit in the expression. In this sense these procedures are arbitrary. Thus, they lead to expressions in which the contribution to the value at the target point from its neighbours depends solely on their separating distance. This contribution is measured in two ways:

(a). **By the weights assigned to the neighbouring points based on their separating distances to the target.**

The weighted points are then included in a linear sum or average which estimates the value at the target point. A distance weighting function is used to assign the weights to the neighbours. Such a function makes implicit assumptions about the spatial relationships between points. It also determines indirectly, the accuracy of the estimates at the target point.

(b). **By quantitative functions on distance which estimate directly the value at the target point and which do not require the assignment of weights.**

The spatial relationship assumed is implicit in the form that such an expression takes.

In both cases (a and b), once the target point has been estimated the equation or average is shifted to a new neighbourhood and applied. The procedure is repeated for all neighbourhoods until the whole area is covered.

Numerous distance-weighting functions and distance-based equations have been suggested for local interpolation in other disciplines, but only a limited few have been applied to soil data. Schut (1974) has reviewed those concerning applications to

digital terrain models. Table 7.2 presents a summary of both distance-weighting functions and distance-based equations. Their virtues and shortcomings as spatial interpolators have already been discussed in chapter 6.

The linear sum or average referred to in (a) above, can be expressed in its most general form as:

$$\hat{Z} = 1/n \sum_{i=1}^{n} w_i z_i \qquad (7.21)$$

where w_i are the weights assigned to the n neighbours, z_i are the values of the property at those locations and Z is the estimated value at the target point in each of the neighbourhoods. In the simplest case $w_i = 1$ for all i, hence the target point is estimated by a simple average. In other situations the factor 1/n may be already implicit in the weights w_i.

Most arbitrary distance-weighting functions are based on a ratio r between the distances from the target point to any neighbour and from the target point to the boundary of the neighbourhood. The most common expressions of these functions are, in structure, similar to any of the following:

$$w = 1/ r^q \qquad (7.22)$$

$$w = 1 - \S\, r^q \qquad (7.23)$$

$$w = \exp(- \S\, r^q) \qquad (7.24)$$

where § are arbitrary coefficients (real numbers) from trial-error, and q are the powers found also in the same way.

The distance-based equations in (b) above, have a general form which can be said to be represented by the expression:

$$\hat{Z} = \frac{\sum_{i=1}^{n} (z_i/d^q)}{\sum_{i=1}^{n} (1/d^q)} \qquad (7.25)$$

where the d_i are the corresponding distances between the target point and each of the i neighbouring points with values z_i. It is easy to see that these expressions are structurally a sort of moving average. Some other variants to this type of expressions are found in table 7.2.

On the whole, it can be said that both these procedures (a and b above), use criteria of distance to interpolate at the target point. The equations based on such criteria reflect the notion of a monotonic decrease in the strength of association with the separating distance between the target point and its neighbours. However, such criteria are arbitrary.

Clearly, no general rule can be derived from the range of experiments in published work involving weighting functions, distance-base equations and various data configurations. Their common features are that they are arbitrary, and that they lack a theory to explain the nature of the changes in spatial dependence between neighbouring points with the separating distance.

7.3.3 Optimal Interpolation based on Spatial Autocovariance.

Spatial autocovariance functions represent the approach needed to solve satisfactorily and theoretically soundly the two fundamental problems in spatial interpolation, i.e. the definition of neighbourhoods that make sense to the spatial structures in soil, and an optimal weighting procedure based on such structures.

Burgess (1979) presents the semi-variogram as the most useful of the functions that describe soil variability in terms of the distance between sampling points. The semi-variogram is a kind of spatial autocovariance function based on regionalized variable theory. This function is unambiguous and is based on the most general of the hypotheses concerning soil variability.

7.3.3.1. Regionalized Variables Theory.

Consider the observations or measurements of a soil property Z taken from samples that can be considered points on the landscape. Each point is defined in terms of two spatial co-ordinates (x_i, x_j) denoted both by the vector (x). Thus, the points are

denoted by x_1, $x_2 \ldots x_n$, and the observed values are $Z(x_1)$, $Z(x_2)$, $\ldots Z(x_n)$.

If the observed values come from sites sufficiently close together, they will exhibit a certain degree of spatial continuity and cannot be regarded as totally independent of one another.

Regardless of the method of measuring spatial variability the following statements can always be made.

i. The observed values $Z(x_i)$ of any point x_i near the point of interest x are related to its value Z(x).

ii. The observed values of the sites $Z(x_i)$ located closest to the point x are more closely related to the value Z(x) of x than those for points further away.

iii. There exists a relationship between observed values from the sites that is a function of the distance between the sites.

The statements above presuppose that there is spatial structure in the sample distribution. The samples are said to have been obtained from a regionalized phenomenon and consequently the observed values $Z(x_i)$ are a regionalized variable.

In addition to the observed values Z(x) which are the sole interest of classical statistics, position and configuration are also of much importance in a regionalized variable.

Matheron (1965) has brought together a body of theory to deal with regionalized variables. This is the theory that supports autocovariance functions.

Apart from spatial dependence, a regionalized variable may exhibit directional dependence. This is to mean that the nature of the spatial dependence can change with direction .

A regionalized variable is said to be isotropic if the nature of the spatial dependence is the same or similar in all directions.

A regionalized variable possesses both random and structured aspects. Both of these are expressed if $Z(x_i)$ is defined as a random function and the values $z(x_1)$, $z(x_2)$, ... $z(x_n)$ are the realizations of such a random function, so that for any specific value of x, say x_i, $Z(x_i)$ is a random variable.

The spatial structure of the phenomenon can be represented by a model and such a model will be used in statistical inference to estimate values at unvisited sites from the spatial relationships of the known values at visited sites. Such inference is valid

only if the model represents the spatial properties of both the sampled field and the point to be estimated. Some stationarity conditions must also be satisfied. These involve assumptions about the first order moment of these random variables, that is, about the mean value of $Z(x_i)$,i=1,2...n, and about the second order moments which correspond to the joint distributions of pairs of random variables $Z(x_i),Z(x_j)$,i,j = 1,2,...n. Such assumptions are the **hypotheses of stationarity**.

A random function Z(x) is said to be first order stationary in the area of a neighbourhood V, if the expected value of Z at any place within V is independent of location, i.e. is the mean. So

$$E\ [\ Z(x)] = m\ , \quad \text{for all } x \text{ in } V \tag{7.26}$$

where m is a constant, and is usually known as the drift.

A random function is said to be second order stationary if the covariance of values from realizations at two points at distance h apart, say x and x+h in V, depends only on the magnitude and direction of vector h and not on x.

$$Cov\ \{\ Z(x),Z(x+h)\ \} = c(h) \text{ for all } (x,x+h) \text{ in } V \tag{7.27}$$

If there is sufficient evidence to show that a regionalized variable is isotropic then the covariance depends only on distance [h] and not on the direction of the vector h. If so

$$c(h) = c([h]) \tag{7.28}$$

The second order stationarity assumption is difficult to meet when dealing with soil data. It has been shown (see chapter 3), that soil variability increases with the size of the area sampled. Furthermore, some properties may even show a tendency to increase their capacity for dispersion. Second order stationarity assumes the existence of a covariance c(h) and thus of a finite *a priori* variance: Var{ Z(x)} = C(0).

For soil data *a priori* variances or covariances cannot be assumed. Hence, a weaker hypothesis is needed.

The intrinsic hypothesis describes the first and second moments of the differences [Z(x + h) - Z(x)] instead. It has been found suitable for describing spatial variations of soil

properties. The first and second order moments of the differences depend only upon h and not on x.

The intrinsic hypothesis states that a random function is said to be stationary when:

i. The mathematical expectation exists and does not depend on x.

$$E\{Z(x)\} = m, \quad \text{for all } x; \tag{7.29}$$

ii. For all vectors h the increment [Z(x+h)- Z(x)], has a finite variance for all x;

$$\text{Var}\{Z(x+h) - Z(x)\} = E\{[Z(x+h) - Z(x)]^2\} = 2\gamma(h) \tag{7.30}$$

which does not depend on x. The function $\gamma(h)$ is the semi-variogram.

The intrinsic hypothesis then implies that for any random function for which a semi-variogram can be defined, the existence and stationarity of the semi-variogram can also be assumed. Thus, when the intrinsic hypothesis holds, the same degree of difference can be expected between two points h distance apart, regardless of what the values of the property are. Further, under the intrinsic hypothesis the values Z(x) are described by the drift and by a residual term:

$$Z(x) = m + e(x) \tag{7.31}$$

where m is the mean in V, and e(x) is the residual term from the drift. This residual term, under the intrinsic hypothesis, is a spatially dependent random component with zero mean and variance:

$$\text{var}[e(x)-e(x+h)] = E\{[e(x)-e(x+h)]^2\} = 2\gamma(h) \tag{7.32}$$

so the function (h) depends only upon the variance of the differences of residual terms. This function is also the semi-variogram and has enormous practical implications.

However, in practice the structural function $\gamma(x,x+h)$ cannot be safely assumed to be stationary except locally, for limited distances [h] < b . The limit b may represent the radius of a neighbourhood for local estimation or may be the extent of a

homogeneous zone. The regionalized variable then is said to be quasi-stationary, and the intrinsic hypothesis (if only the semi-variogram is assummed) for distances [h] < b corresponds to a quasi-stationarity hypothesis.

In practice moving neighbourhoods of size b can be defined for local estimation within which the expectation and the covariance can be assumed stationary and this is sufficient for inference, provided the semi-variogram exist and is known.

The hypothesis of quasi-stationarity works well for soil data for it is really a compromise between the scale of homogeneity of soil and the amount of available data.

7.3.3.2 Measuring Spatial Dependence: The Semi-Variogram.

The structure and components of the spatial variation of soil properties must be measured and modelled. Such a model has two important practical applications. It indicates the extent ([h]<b) in distance within which a quasi-stationary neighbourhood can be defined, and it contributes information about the nature of the spatial variation of the property to the weighting of points in the neighbourhood so that the property at a point can be estimated.

Instead of using the correlogram (Webster and Cuanalo, 1975), to present the relationship between sample values at a distance h, Matheron (1971) has proposed, in the mining context, a more powerful tool, the semi-variogram. Burgess (1979) and Burgess and Webster (1980a) have used it for soil.

The semi-variogram exists and can be estimated wherever the intrinsic hypothesis holds. It has the added advantage of defining parameters needed for local estimation by Kriging.

By definition the value of the variogram $2\gamma(h)$ for a given distance h is the expected squared difference between the values of the samples separated by distance h.

$$2\gamma(h) = E\{[z(x) - z(x+h)]^2\} \qquad (7.33)$$

The semi-variogram is half of this quantity

$$\gamma(h) = 1/2\ E\ \{[\ z(x) - z(x + h)]^2\} \qquad (7.34)$$

The semi-variogram can be thought of as a measure of the degree of dissimilarity between two places or as half the error of estimating the value at one place from an observation at the other (Webster and Burgess, 1983).

Given the intrinsic hypothesis, an estimator of the semi-variogram (h) from field data is:

$$(h) = 1/2M\sum_{i=1}^{n} [Z(X_i) - Z(X_i - h)]^2 \qquad (7.35)$$

where h is the vector of distances between pairs of observations X_i and M is the number of such pairs at lag h.

It can be shown that, under the full second order stationarity assumption, the semi-variance is related to the autocovariance and autocorrelation by

$$(h) = C(0) - C(h)$$
$$= \sigma^2[1 - p(h)] \qquad (7.36)$$

where C(h) is the autocovariance of a property at points separated by lag h. When h = 0 , C(0) defines the variance σ^2. The ratio C(h)/C(0) is the autocorrelation denoted by p(h). They are complementary. The greater the differences in the values at lag h the larger the semi-variance and the smaller the autocovariance and the autocorrelation. However, the semi-variogram can exist without the corresponding autocovariance and that is the reason why the semi-variogram is preferred for spatial analysis.

7.3.3.3 Experimental Semi-Variograms.

A semi-variogram can be calculated from data collected along soil transects or on the plane in two or more directions.

A plot of the semi-variances in one direction against the lag distances is known as the sample or experimental semi-variogram. it is readily drawn and if desired the plotted points may be joined by straight lines.

Soil survey concentrates on spatial variation, therefore the semi-variogram on the plane defined by the two lateral dimensions is the most useful. If variation is isotropic then the same semi-variogram can be assumed for any direction on the plane. The experimental semi-variogram can then be drawn as a simple graph. In cases where the experimental semi-variograms from the same data change with direction on the plane, the soil property in that field is said to be anisotropic. Anisotropies are examined in more detail later in this chapter.

Experimental semi-variograms can be calculated as two-dimensional functions on the plane from samples taken at regular intervals as in a rectangular grid or in a less neat situation when the sample points are irregularly scattered.

For samples on a regular grid of, say, m rows and n columns, Webster (1985) suggests the estimator:

$$\hat{\gamma}(p,q) = 1/[2(m-p)(n-q)]\sum_{i=1}^{m-p}\sum_{j=1}^{n-q}[z(i,j) - z(i+p,j+q)]^2 \qquad (7.37)$$

$$\hat{\gamma}(p,-q) = 1/[2(m-p)(n-q)]\sum_{i=1}^{m-p}\sum_{j=q+1}^{n}[z(i,j) - z(i+p,j+q)]^2$$

where p and q are the lags in the two dimensions.

Usually only half of the total set of semi-variances is needed as a semi-variogram tends to be symmetrical about its center. Usually the number of pairs used to compute the semi-variances decreases as the lag distance increases. Thus, the semi-variances for larger lag distances are somewhat less reliable. The closer to the origin of axes on the graph the more reliable are the semi-variance estimates, hence the tendency to symmetry, and the neglect of the extreme right part of the plot as insignificant. The length of the maximum alignment of the field defines the significant section of the semi-variogram; $[h] < L/2$. This is part of the so called "practical rule" (Journel and Huijbregts 1978). The equations above permit the computation of half of the semi-variances for lags -q to q and from 0 to p.

In soil survey practice there are often missing data even in regular grids. In such cases the denominators above will have to be replaced by the number of paired comparisons in each sum.

The most common situation is to have data from irregularly distributed samples on the plane. Webster (1985) describes an algorithm to compute the experimental semi-variogram by grouping the point data into angle classes each representing one direction, followed by distance classes. A set of angular regularizations is chosen in specific directions. Within the space limited by the arc subtended by such an angle all points are considered as being in the same direction. The lag distances are defined by classifying them in ranges. The mid-point of each range is a nominal lag distance. Each squared difference contributes to the semi-variance for the lag class within which it falls by virtue of its actual separation. The same occurs for each angular regularization representing a direction.

This algorithm was made available to OXSIS to compute the experimental semi-variograms of chapter 9.

In practice, it is usually enough to consider four directions in the two-dimensional space (Journel and Huijbregts, 1978).

In any semi-variogram calculation the size of the support, i.e. the volume of each individual sample, always has to be stated as the amount of variation present in a sample or core depends on the size of the core and this affects the semi-variogram. The bigger the the area of the sampled core the more variation it will embrace. Thus, the semi-variogram depends on the size, shape, and orientation of the individual samples on which measurements are made. Semi-variograms from auger pits refer specifically to that type of support and may not be comparable to those obtained from bulked samples from open pits.

Another important aspect of the semi-variogram calculation is the total number of samples included in the computation. Reliable semi-variograms cannot be expected when calculated from less than about 40 pairs, (N> 30,50 pairs, is part of the "practical rule"). So in between 100 to 200 observations are necessary to obtain a reliable estimate of a semi-variogram (Webster,1986).

7.3.3.4 Semi-Variogram representation.

The computed semi-varinces can be represented graphically in various forms. In the one dimensional case, i.e. transects, a straightforward plot of semi-variances against lag distance

suffices. The plotted points may be joined by straight lines since they represent continuous functions. For the two-dimensional case other forms of representation exist, e.g. isarithmic maps, perspective block diagrams, cylindrical projections (Webster, 1985), and others. Alternatively, the semi-variogram for the four or more directions may be represented in a joint plot on a single pair of perpendicular axes, and if desired, the points corresponding to each direction could be joined by distinctive straight lines for easy identification.

However, if the soil property varies isotropically there is no need for separating the points of each direction. A single semi-variogram will be sufficient to describe variation in all directions and can be drawn as a simple graph. This is the experimental semi-variogram. This form of representation has been chosen for the semi-variograms shown in chapter 9.

When the semi-variogram represented in this way shows distinctive departures of the points in one direction from the rest, so that it cannot be described by the same model as the others then the variation is anisotropic.

7.3.3.5 Semi-Variogram parameters.

The shape of an experimental semi-variogram can take many forms. Ideally the semi-variance increases with the distance between observed sites rising up to a point where it levels out to a constant value known as the **Sill**. Such semi-variograms are said to be transitive. A transition phenomenon is a regionalized variable with a semi-variogram that increases smoothly as the lag distance h increases until h reaches a value **a**. The value **a** is the **Range** of spatial dependence and is also called the **Range of influence**. The range is the lag distance at which the sill is reached. All values from pairs of points which are separated by a distance shorter than **a**, are said to be spatially dependent, and those otherwise are independent. The range of influence defines the neighbourhood size within which the structure of variation can be modelled and thus used in weighting for interpolation.

The semi-variance at zero lag is theoretically zero itself, so ideally the experimental semi-variogram should pass through the origin at zero lag. Nevertheless many soil properties appear to

have non-zero semi-variances as lag distance approaches zero, and smooth curves fitted to experimental semi-variograms often show a non-zero intercept Co. This variance is known as the Nugget Variance, and the feature as the Nugget effect. It represents all variabilities occurring at distances shorter than the minimum sampling interval, plus the variation due to measurement error. So this unexplained variation is considered as random variance.

The nugget variance is a true form of variation and can only be elucidated by sampling more closely. Usually but not always, it is a matter of the scale of sampling. Some structure may be found from seemingly random variance by examining the variation at a bigger scale.

The nugget effect Co, added to the spatial covariance C, approximately equals the sill variance which estimates the sample variance S^2. So it can be written:

$$S^2 = Co + C \qquad (7.38)$$

A semi-variogram may appear completely flat suggesting total absence of spatial dependence. Variation of this sort is considered as completely random at the scale of sampling and corresponds to a pure nugget effect. Pure nugget effect arises from very large point-to-point variation at short distances of separation. In this case the best estimator at any point within the sample is the mean.

The nugget variance can also be expressed as a proportion of the sill variance to allow comparisons between different semi-variograms.

The nugget variance sets lower limits to both the precision of the sampling or measurement technique if it can not be reduced by further sampling (Burrough, 1983a), and to the size of the estimation variance by Kriging, thus defining the precision of interpolations.

Since the nugget effect is considered as a scale effect, the elucidation of successive nugget effects by further sampling at closer intervals makes it possible to study the nested structures of variation caused by multiscale sources of spatial variability in soils (Burrough, 1983a).

In summary, the significant parameters of semi-variograms of isotropic transition phenomena are the range of influence, the nugget variance and the sill variance when it is present.

Nested structures of variation will yield nested sets of semi-variograms if various sampling scales have been used. These parameters are estimated for each structure present.

7.3.3.6 Semi-variogram models.

To characterize the spatial structure of variation in a soil property a model for the experimental semi-variogram is needed. In principle, a constellation of functions could be fitted to experimental semi-variograms by a variety of procedures. However, if they have been calculated from enough data, most experimental semi-variograms present a relatively simple curve. To be useful an equation or model for this curve is needed. In practice, only a small number of such equations are currently used. Though other equations could be fitted, they tend to be very complex, and to restrict mathematical manipulation. More importantly, any equation fitted must meet certain conditions. This reduces the range of choices considerably to only a few theoretical models of stationary regionalization, most of which have become standard. The analyst is restricted to choice from these models.

The conditions for a semi-variogram model are:

i. **The model must be a conditional negative semi-definite function, (CNSD).**

A regionalized variable Z must have a positive or zero variance but not negative. Its covariance function C(**h**) if it exists, must be positive semi-definite (Webster, 1986).

Recalling eq. (7.36) the semi-variance relates to the covariance by:

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

Hence since C(h) is conditional positive and on removing the sign the semi-variogram must be conditional negative semi-definite (CSND):

$$C(h) = C(0) - \gamma(h)$$

The property of a function C(h) to be positive semi-definite changes with the dimensions of the space in which the vector [h] is referenced. So, concerning the semi-variogram a CNSD in one dimension may not be so in two or more dimensions, but a CNSD in n dimensions is always so in every p < n dimensions.

ii. The model must describe the two main features of a stationary variogram.

These are related to the behaviour at the origin, i.e. linear, parabolic, or nugget effect, and the presence or absence of a sill in the increase $\gamma(h)$ and $\gamma(h)$=constant when [h] > a, (the range).

A vast family of CNSD can be found to meet condition (i), and it is not easy to test if any particular model is CNSD. There exist some procedures but they are either elaborate or too restrictive, so functions not CNSD are best avoided by using only recommended or "authorized" models.

Webster (1986) divides them according to the presence or not of a finite a priori variance into:

i). Transitive models

ii). Unbounded models

Transitive models are models with a sill. They possess a finite a priori variance. Spherical and exponential models are of this sort. Unbounded models do not posses an apparent finite a priori variance within the area covered by the samples. Power functions and the linear-without-sill model are models of this type.

Models can also be categorized (Webster, 1985) into:

i). Safe models

ii). Risky models

i). Safe models.

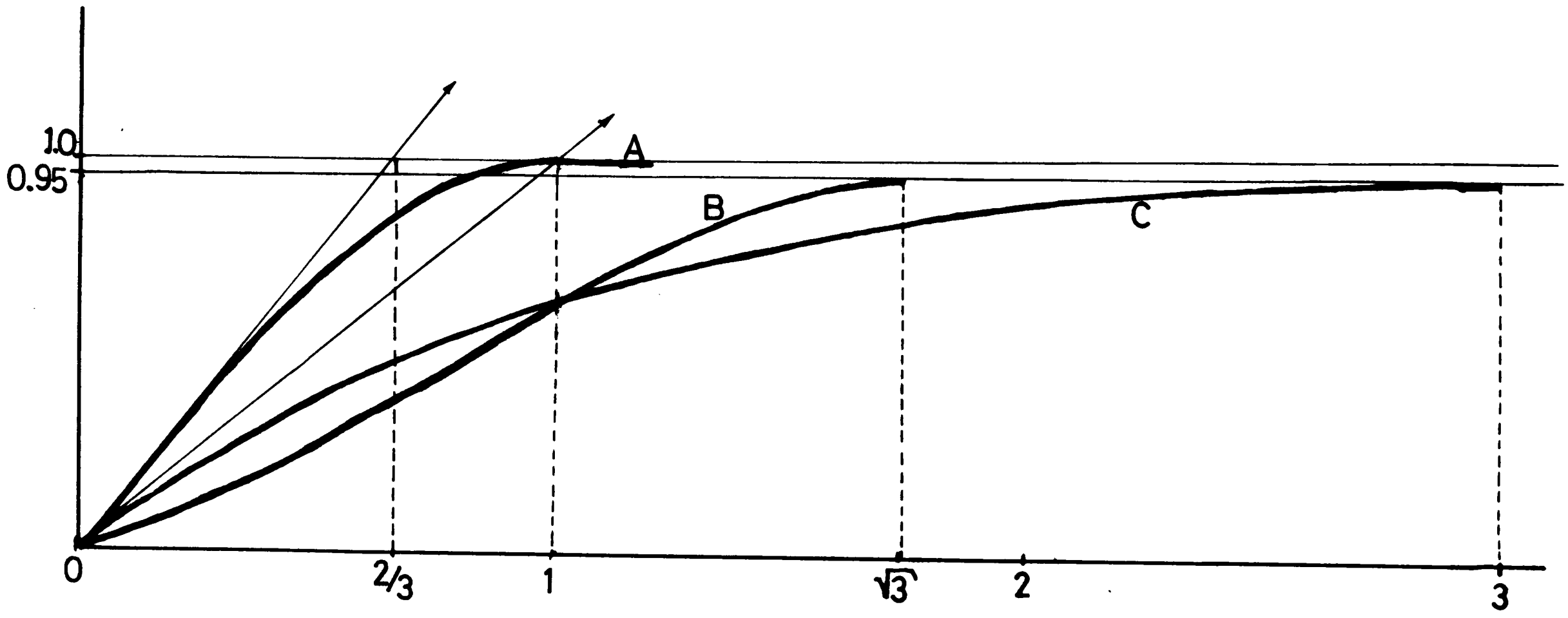
Safe models are models that experience has shown can be recommended for soil properties, they are CNSD in two and three dimensions. The curves generated by the safe models are shown together in the plot of figure 7.4. The models are:

1. Linear models.

$$\gamma(h) = Co + gh \qquad \text{for all } h > 0; \text{ and } \gamma(0) = 0. \qquad (7.39)$$

where Co is the intercept or nugget variance and g the slope.

Fig. 7.4 Semi-variogram models



A = Spherical
B = Gaussian
C = Exponential

2. **Spherical models**.

$$\gamma(h) = Co + C\,[3/2\,(h/a) - 1/2\,(h/a)^3] \quad \text{for } 0<h<a \qquad (7.40)$$

$$\gamma(h) = Co + C \quad \text{for } h > a$$

$$\gamma(0) = 0$$

where b is the range of influence C is the Sill and Co is the nugget variance.

Spherical models describe well transition phenomena that occur in soil. This model fits the experimental semi-variograms of many soil properties and it is much used. The spherical model was assumed for the experimental semi-variograms of some soil properties from the Yelford area, in chapter 9.

3. **Exponential models**.

Exponential models are represented by the equations:

$$\gamma(h) = Co + C\,[\,1 - \exp(-h/r)] \quad \text{for all } h>0 \qquad (7.41)$$

$$\gamma(0) = 0$$

where r is a distance parameter controlling the spatial extent of the function, Co, C and h are as above. The exponential model approaches the sill asymptotically and in consequence there is not a finite definition of the range of influence. A practical estimate to determine the range is to consider it as a function of r. Thus, $a = 3r$ defines a practical range for this model. This rule of thumb is very important as it helps to relate this model to the spherical model. Such relationships, in practice, provide graphic criteria for a first visual selection of the model to fit the experimental data.

The difference between the spherical and the exponential models is the distance (lag) at which their tangents at the origin intersect the sill. Thus the spherical model reaches its sill faster than the exponential model. These sort of relationships can be seen in figure 7.3, and were used as key to the first

choice of a model for some experimental semi-variograms in the Yelford area (see chapter 9).

3. Gaussian model.

The form of its mathematical representation resembles that of the exponential model. The sill is reached asymptotically too and there is a practical rule to define the range of influence. The model is:

$$\gamma(h) = 1 - \exp(- h^2/r^2) \qquad (7.42)$$

where the terms in the equation are the same as for the exponential model. A practical range is reached when a = r/3.

ii). Risky Models.

Some models can be considered risky for they are not CNSD in the dimensions of importance for soil survey where they can find application. In particular the bounded linear model or linear with sill model.

1. Linear-with-sill model.

The model writes:

$$\gamma(h) = Co + C\,(h/a) \;; \qquad \text{for } 0 < h < a. \qquad (7.43)$$

$$\gamma(h) = Co + C \;; \qquad \text{for } h > a$$

$$\gamma(0) = 0$$

This model represents the extreme form of a transitive model. It produces good fits particularly in situations when the semi-variogram appears to curve to a sill more sharply than a spherical model does. However, this model is only CNSD in one dimension. It can be used with confidence in transects, but in two dimensions it is not CNSD and thence should not be used in soil survey situations.

2. Circular model.

This model is CNSD only up to two dimensions but for soil survey that is fairly convenient. It produces sharper approximations to data than the spherical model, particularly in the region of the curve where the monotonic increase bends on reaching the sill. The mathematical expression of this model is somewhat elaborate and it is not presented here. There are not

many accounts of its use on soil properties, for in most practical instances a spherical model achieves similarly good descriptions. Webster (1985) presents a reasonable fit of a circular semi-variogram to values of topsoil pH.

Some other models have been fitted to experimental semi-variograms of soil properties. Webster (1985) presents some examples. Since these models have not been sufficiently investigated to allow their safe application to experimental data, they are not the concern of this thesis.

7.3.3.7 Model selection and estimation of semi-variogram parameters.

The model of the semi-variogram contains the parameters that need to be estimated. Thus there are two aspects of this stage. On one hand is the choice of the semi-variogram model, and on the other is the estimation of the parameters by fitting the model to the sample data.

Model selection is still one of the most delicate aspects in the application of spatial statistics to soil studies. Initially it rests on a combination of intuition, experience and knowledge of the regionalized phenomenon under study. One of the safe or authorized models has to be selected, and it has to be one that describes the sample semi-variogram best. Inevitably the procedure is iterative, and there is no unique technique or method for choosing and fitting a theoretical model. Nor is there a way of completely knowing a semi-variogram.

Least squares techniques weighted for the number of sample pairs appear to be preferred by some workers for fitting semi-variogram models and for parameter estimation (Webster, 1985, Trangmar et al, 1985). Between two models that look equally satisfactory Webster (1986) chooses the one that left the smallest residual mean square, and for models involving different numbers of parameters he uses the so called Akaike information criterion to chose the best model.

Journel and Huijbregts (1978) warn strongly against the use of automatic fitting by least squares methods on various grounds. The estimator (h) of each experimental semi-variogram at the

point separated by h distance, is subject to errors of estimation and fluctuation which are particular to that point and which vary with the lag distance h, locality, and other factors. It is considered that the best method for assessing the reliability of each estimate (h) comes from practical experience, common sense, knowledge of the physical phenomenon and critical review of the data, and that these are as essential to structural analysis as the calculated experimental semi-variogram itself.

Thus the fitting of a model cannot be blind and an initial choice has to be fed into these techniques. Visual inspection of the experimental semi-variogram combined with a priori knowledge of the regionalized phenomenon guide this initial choice.

Procedures for finding the nugget variance and the sill of transitive models have been suggested by Journel and Huijbregts (1978). These recommendations were applied in this thesis. The geometric relationships between parameters of safe models, (i.e. spherical, exponential, gaussian) as displayed on the joint plot of figure 7.4, can be conveniently exploited to guide the initial choice of model. If any of these models is pertinent then the initial choice must be supported by experience and local background knowledge of the factors affecting soil variability.

For the estimation of soil properties the region of transitional monotonic increase in the experimental semi-variogram is to be considered the most important. This must be the area of concern when fitting a particular model and estimating parameters. If two types of semi-variogram model approximate reasonably close to the data in this region it is unlikely that their discrepancies may produce considerable differences in the geostatistical results.

Many experimental semi-variograms of soil are approximately linear over the first few lags so the intercept to the ordinate corresponding to the nugget variance can be found by fitting straight lines and extrapolating. For the estimation of the sill variance, the dispersion variance as calculated from the same sample is used together with the sample semi-variogram. Bearing in mind that the dispersion variance underestimates the sill variance for transitive semi-variograms, adjustments are guided by the sample semi-variogram.

In the Yelford area the initial choices of semi-variogram models and parameters were made for various soil properties, guided by the type of relationships discussed above (see chapter 9). For the estimation of parameters of the semi-variograms in this thesis, a computer program that plots the curve of the model onto the scatter of points of the experimental semi-variogram was used interactively to display the fit on the screen of the VDU terminal. The procedure was iterative. Several combinations of values of parameters for the model in turn were input into the program and tried interactively until the best fit of the model on the points, as determined by visual inspection, was achieved. The set of parameters of the final and best fit was used as the parameters of the model for subsequent calculations.

Journel and Huijbregts (1978) have demonstrated the robustness of geostatistical results with respect to the chosen type of semi-variogram model by comparing two types of model (e.g. spherical vs. exponential). They found very small differences between the results derived from the two models.

As far as this thesis is concerned the main purpose of spatial analysis is to provide the semi-variogram parameters for local estimation by kriging (see later in this chapter). In this context minor discrepancies in estimation of semi-variogram parameters make little difference to the reliability of the interpolation because of the robustness of kriging (David, 1977).

The estimation variances in kriging are the most sensitive parameter to changes in semi-variogram model type. The effect of a particular model type on subsequent kriging estimation variances can be used as a criterion for chosing between competing models. Cross-validation of the precision of kriging estimates as compared with the original values at check points is the most practical criterion to assess the success of any choice of semi-variogram model.

The variances of the kriging estimates were used to cross-validate the semi-variogram models of the Yelford area in this thesis, which were originally chosen by the procedures described above (see chapter 9).

7.3.3.8 Anisotropy.

A regionalized phenomenon is said to be anisotropic when its variability is not the same in every direction. In this case the function (h) depends not only on the distance modulus [h], but also on directional parameters. The semi-variogram is not the same in all directions, so the structural function

$$\gamma(h, \Theta, \phi) = ([h]) \quad \text{for all } \Theta, \phi. \qquad (7.44)$$

that characterises isotropic phenomena can no longer hold.

In practice, anisotropy is to be interpreted as the presence of preferential directions of regionalization during the genesis of the soil or the effect of external factors on soil variation. Some of these preferential directions can be known or inferred a priori from analysis of the landscape.

Ideally, an isotropic soil property will define a circular neighbourhood of radius equal to its range, which is the same in all directions. The same anisotropic soil property, with the same sill and varying range will define an ellipsoid or elliptical shaped neighbourhood on the plane due to the preferential directions of regionalization. This is said to be geometric anisotropy. When nested structures of variation are characterized by corresponding semi-variograms and each component structure may have its own anisotropy, zonal anisotropy is present.

Geometrical anisotropy can be recognised by examining the semi-variograms of individual directions separately and comparing them. Usually, the semi-variograms in different directions can be described by the same model, but with different parameters. In some cases anisotropy is so strong that it may be apparent even on the joint plot. Some experimental semi-variograms obtained from the Yelford area evidently show this type of anisotropy, (chapter 9).

Procedures for finding a model for anisotropy and so to allow for kriging, are suggested by Journel and Huijbregts (1978), Burgess (1979), and David (1977). These procedures are all based on the transformation of the elliptical or ellipsoid shape described by the contours of the semi-variograms in different directions due to their different ranges of influence, into a

circular one. Thus, a single semi-variogram model can be used to represent the semi-variograms in different directions, if each directional semi-variogram can be approximated by a straight line with slope m_i such as:

$$\gamma(h,\theta) = Co + w(\theta)[h] \qquad (7.45)$$

where θ is the direction of the vector [h], relative to the X-axis, and $w(\theta_i) = m_i$, being θ_i the angle of the ith direction.

The quantity $w(\theta)$ can be estimated by an elliptical function:

$$w(\theta) = A\cos^2(\theta-\phi) + B\sin^2(\theta-\phi) \qquad (7.46)$$

where θ is the direction of maximum variation, A is the gradient of the semi-variogram in that direction, and B is the gradient in the direction $\phi+\pi/2$.

The latter equation can be fitted by non-linear least squares procedures using the slopes of the individual directional semi-variograms (m_i) and the direction (θ_i) in radians. Burgess (1979) found a good fit with four points (directions), and this enabled him to interpolate by kriging using the transformed semi-variogram model of eq. (7.45).

Strict geometric anisotropy is defined by David (1977). It is represented by the following equation:

$$w(\theta) = [A^2\cos^2(\theta-\phi) + B^2\sin^2(\theta-\phi)]^{1/2} \qquad (7.47)$$

where the terms are the same as in the previous equation. This permits the estimation of the parameter $w(\theta)$ necessary for eq. (7.45), since the representation of $w(\theta)$ need not be unique.

For the anisotropic semi-variograms encountered in Yelford attempts were made to fit both of the geometric functions above using four points, i.e. directions, (see chapter 9). No good fit could be obtained with any of them. The fitting by non-linear least squares using a standard algorithm from a statistical package was lengthy and expensive.

Background information on the area suggests that the anisotropies follow a more general source of variation such as a drift or general trend which causes the presence of a preferential

direction of regionalization. To deal with this source of anisotropy such trend must be accounted for by modelling and so removed. This is shown in chapter 9.

7.3.3.9 Trends and changing drift.

It has been seen in section 7.3.3.1 that the intrinsic hypothesis does not require strict stationarity of the variable Z. It holds provided quasi-stationarity can be assumed. This means that very gradual and smooth changes can be tolerated. The semi-variogram exists and can be calculated as if the local trend or "drift", given by the expected value of Z(x); E [Z(x)] = m, were constant (m). The model of soil variation is then (recalling equation 7.31):

$$Z(x) = m + e(x)$$

being e(x) a random residual term with zero mean.

However, even over the neighbourhood V, for any given soil property the drift (m) may change gradually in the form of a local trend. When these changes in the form of trends are sufficiently strong quasi-stationarity cannot be safely assumed for the drift (m) within V is no longer constant but is a function of position. So, not being strictly a constant it is best represented by m(x). Thus, the model of soil variation is now;

$$Z(x) = m(x) + e(x). \qquad (7.48)$$

These trends may be part of a global recognisable trend over the entire area, or a local strongly changing drift, i.e. m(x). In any case, the presence of such trends indicates a departure from first order stationarity in the intrinsic hypothesis. To best describe the variable Z(x) then a procedure is needed that will reveal both the drift m(x) and the deviations from it e(x), allowing for interpolation of non-stationary data.

If the slope of a semi-variogram appears to be parabolic or higher power near the origin and the extrapolated curves near the

origin show insignificant nugget variance, these are the indications of the presence of changing drifts, Burgess (1979).

In the presence of a local trend or drift the residuals depend not only on the observed values but on the drift values too and such values are unknown. Thus the e(x) are now a deviation from the expectation E[Z(x)]:

$$e(x) = Z(x) - m(x) \qquad (7.49)$$

and the expectation of the squared differences of Z(x) and Z(x+h) is no longer a valid estimator of the semi-variogram for it contains values incorporated by the drift. So equation (7.30) does not hold under these circumstances.

If the semi-variogram is calculated as if the drift were constant, but in fact a local drift (say linear) is present within V the computed semi-variogram will be approximately parabolic near the origin and this is the diagnostic criterion of the presence of a linear drift. Higher order drifts are detected when the semi-variogram shows higher order polynomials near the origin.

A detailed structural analysis is needed to reveal both the drift and the residuals which are the deviations from it.

It is worth mentioning that semi-variograms with parabolic behaviour near the origin and insignificant nugget variance are exceptional and only a few cases are reported for soil properties in published work.

The procedures for detailed structural analysis are fully documented in Olea (1975), Burgess (1979), and Webster and Burgess (1980).

Structural analysis aims at both characterizing the changing drift by a model to account for it, and to analyze any spatial structure remaining in the residuals from the modelled drift. The procedures involve:

(i). Calculate a function to estimate the drift m(x)

(ii). Obtain the residuals from the calculated drift.

(iii). Compute the semi-variogram of the residuals.

(iv). Compare the semi-variogram of the residuals with an assumed theoretical semi-variogram. A satisfactory match should be found.

The procedure is somewhat trial-error based on combinations of drift power and neighbourhood size in each trial. If the semi-variogram of the residuals matches the theoretical one the combination is judged appropriate. If the match is poor the size of the neighbourhood is changed or the drift power is changed or both. The best fit of the trial drift is chosen by graph inspection.

The solution to a detailed structural analysis may not be unique. It depends upon the neighbourhood size and scale. Samples at different intensities can provide different solutions to structural analysis and different expressions to the drift.

This situation has important practical consequences. It leaves room for alternative procedures to enable interpolation in the presence of trends or changing drifts. Burgess (1979) makes reference to a particularly important case in practice. Semi-variogram models for anisotropies that attempt to describe stationary phenomena, in some instances it may be better to replace them by a single isotropic semi-variogram plus a linear or quadratic model for the drift, which in turn carries the variation that was previouly accounted for by the anisotropy. This is precisely the approach followed in this thesis to deal with the anisotropic variation of some soil properties in the presence of trends to allow interpolation by kriging (see chapter 9).

In cases where it is not possible to find a good fit of the geometrical functions of equations 7.46 or 7.47 to find a model for anisotropy, as in the case of the Yelford data, a simple model for the drift can be fitted and a semi-variogram of the residuals from the drift calculated to allow spatial interpolation.

In chapter 9 of this thesis a bicubic spline surface (eqs. 7.19 and 7.20) was fitted piecewise to model the drift and so to account for the evident anisotropy as revealed by the semi-variogram. Semi-variograms of the residuals were calculated and from them kriging estimates of the anisotropic property were obtained by adding them to the values of the surface.

In practice, when significant nugget variance is present even if there is an approximately parabolic behaviour near the origin a detailed structural analysis is not needed for it is difficult to separate random noise from a presumed changing drift or local

trend. The local variations can be explained by the semi-variogram calculated as if the drift were constant.

7.3.3.10 Semi-variogram interpretation.

There is not a standard method or procedure to interpret semi-variograms. The existence of established soil property/semi-variogram or landscape/semi-variogram relationships would be of much use in soil survey practice, but there is not enough knowledge derived from field experience yet.

A semi-variogram of soil must make pedological sense. The evidence available tends to indicate that the regional and local controls that affect the shape of the semi-variogram are rock outcrop, landform, rainfall, hydrology, land use and soil management.

Interpretations of semi-variograms are a function of, primarily the scale of observations, then of the soil property, the sampling density, the spatial arrangement of the samples, the field configuration, the size of support and of knowledge about the spatial regional and local controls that influence their shape.

The semi-variogram splits the total variance into two parts. One represents the spatial dependence between values from points at increasingly larger distances. The other represents local random variances, the nugget variance. Hence, a semi-variogram is interpreted in terms of how the different combinations of nugget variance, sill and range of influence affect its shape for a particular property with reference to the ground structures of a particular landscape.

i. Nugget Variance.

Rather small and negligible nugget variances are not common in soil data for they have much noise. However, if present, they can be associated with properties showing a great degree of continuity along transitions, but the most common case in soil properties will be to find a reasonable amount of nugget variance.

Pure nugget effect indicates total absence of spatial correlation, but checks to ensure that the oscillations are not due to experimental fluctuations (too few pairs used in the

calculation) have to be made. The oscillations round the sill in a pure nugget effect are usually small as compared to those due to experimental fluctuations.

ii. Region of monotonic increase.

A monotonic increase in the semi-variogram rising from the nugget variance to the sill is characteristic of transition phenomena and indicates the proportion of the total variance which is explained by spatial dependence between points. It describes the strength of spatial association between neighbours at increasing separating distance by means of its slope.

Most soil properties exhibit linear monotonic increases in this region. However, for a few others the increase may not be monotonic but oscillatory. These semi-variograms are said to display a "hole effect" (Journel and Huijbregts, 1978). Webster (1985) suggests that these semi-variograms should be called periodic and suggests that some periodic phenomena of this sort can be represented by the transformation of the semi-variogram into the frequency domain by Fourier analysis giving a function that expresses the distribution of variances with frequency and which is known as the power spectrum.

This type of phenomena occurs in certain landscapes where it characterises the presence of rich and poor zones in the variable studied, or where some other phenomenon induces periodicity, e.g. across gilgai terrain, or as in sand depositions in a direction perpendicular to two rivers, or in semi-variograms of properties like electric conductivity computed in directions perpendicular to irrigation lines.

iii. Variation round the sill.

The position of the sill is not necessarily given by the value of the dispersion variance. Nevertheless, it can be guided by the value of the dispersion variance bearing in mind that, as Webster (1985) has shown, this underestimates the sill variance.

The fluctuations round the sill are merely experimental fluctuations. A resumed increase after a first sill signifies the possible presence of a second structure which may become fully evident at a different scale. Depending on the purpose of the study this second structure can be ignored or modelled. Checks

can be carried out to ascertain whether this structure is not due to an insufficient number of pairs used in the calculation or due to ignoring the "practical rule".

If a series of nested samplings have been undertaken a second structure may be revealed as a part of a nested model. Nested semi-variograms for some soil properties have been obtained by McBratney et al (1982) and Webster and Nortcliff (1984). Semi-variograms of this sort represent variation at more than one spatial scale.

In the Yelford area, models of nested structures for some soil properties were fitted to experimental semi-variograms calculated from a set of overlapping grids of varying sampling spacings. Different spatial scales of variation are revealed from these semi-variograms. They are shown in chapter 9.

The part of the semi-variogram where the monotonic increase bends towards the sill has a practical interpretation too. If the turn is sharp towards the sill this indicates an extreme form of a transitive model.

The lag at which the semi-variogram turns towards a sill, i.e. the range, may have a clear physical interpretation. This lag may coincide for example with the width of outcrop of main geological formations or with an abrupt change in soil from one unit to another.

iv. Range of influence.

The lag distance where the monotonic increase reaches the sill and spatial dependence ceases defines the range of influence. So it is a function of slope and curvature of the monotonic increase and of the smoothness with which the latter reaches the sill. Short ranges are usually associated with local variation at the scale of sampling or with rapid changes in the values of the property over distance particularly if the nugget variance is small and the slope of the increase is steep. Short ranges can be expected in properties that tend to be dynamic and mobile such as some chemical elements in their soluble forms or in the form of minerals of low concentration.

Long ranges of influence indicate smooth long transitions and very gradual changes over distance. They represent a good degree of uniformity in the regionalized variable. This is usually associated with properties influenced by alluvial deposition of

sediments and interbedding in plains and basins and by the influence of river drifts and fluvial depositions too. The exponential and the gaussian models are characterized by long ranges of influence and are also frequently associated with this type of spatial phenomena.

The existence of significantly different ranges of influence along different directions even if the sill is virtually the same is the most clear indication of geometric anisotropy. Corrections for this type of anisotropy have been discussed above.

The interpretation of semi-variograms, fitting of models, and estimation of their parameters are the information required to elaborate a plan for spatial interpolation by the technique called kriging.

7.3.3.11 Spatial interpolation by Kriging.

The sound definition of a local neighbourhood based on spatial structures of variation, i.e. from the range of the semi-variogram, and the selection of an optimal procedure to interpolate using the information of such spatial structures within such neighbourhood are the two fundamental problems in local estimation.

The technique called kriging combined with information from the semi-variogram provide an optimal solution to both.

Kriging can be used to interpolate values at unvisited sites from values at neighbouring visited sites provided the the data satisfies the intrinsic hypothesis and the property shows spatial dependence.

Kriging is essentially a weighted local moving average of property Z which has been measured at sites (points) X_1, X_2,... X_n within the neighbourhood V, giving the values $Z(X_1)$, $Z(X_2)$, ... $Z(X_n)$. The value of Z at a location Xo , Z(Xo), can then be estimated by a linear sum:

$$\hat{Z}(Xo) = \lambda_1 Z(X_1) + \lambda_2 Z(X_2) + ... + \lambda_n Z(X_n) \qquad (7.50)$$

or

$$\hat{Z}(Xo) = \sum_{i=1}^{n} \lambda_i Z(X_i) \qquad (7.51)$$

where the λi are the coefficients or weights associated with the sampled points x_i within V.

The dimensions of the location Xo are such that it can be considered as a point of dimensions comparable to those of the samples or support (point kriging). Alternatively the location can be an area or block (block kriging). Point kriging is regarded as a special case of block kriging.

i. **Point Kriging**.

The neighbourhood V in kriging is centered at the point Xo and its radius is less than the range of the semi-variogram calculated from the same samples and containing a number of sampling points which, it has been shown, do not need to be more than 25.

Kriging is an optimal interpolation procedure in the sense that the weights λ_i of the neighbours of Xo are chosen so that the estimates are unbiased, and the estimation variances are minimized. In this sense kriging is optimal for interpolation.

To obtain an unbiased estimator $\hat{Z}(Xo)$ of Z(Xo) the expectation must be:

$$E\,[Z(Xo) - \hat{Z}(Xo)] = 0 \qquad (7.52)$$

This condition is satisfied if the sum of the weights λ_i is 1:

$$\sum_{i=1}^{n} \lambda_i = 1. \qquad (7.53)$$

The estimation variance is then the expected squared difference of the estimate and the true value and it is to be minimized:

$$E\,\{[Z(Xo) - \hat{Z}(Xo)]^2\} = \text{minimum}. \qquad (7.54)$$

It can be shown that the expectation is:

$$E\{[Z(Xo) - \hat{Z}(Xo)^2]\} = 2 \sum_{j=1}^{n} \lambda_j \gamma(Xo, X_j) - \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j \gamma(X_i, X_j) \qquad (7.55)$$

where (X_i, X_j) is the value of the semi-variogram along the line joining X_i and X_j which are separated by a distance [Xi-Xj]. Direction is immaterial if the property varies isotropically and the semi-variogram depends only on the intervening distance. Thus;

$$\gamma(X_i, X_j) = \gamma([X_i - X_j]) \qquad (7.56)$$

The optimization procedure consists then of minimizing the estimation variance with respect to each λ_i subject to the constraint that $\sum_{i=1}^{n} \lambda_i = 1$. The problem is one of constrained minimization which is solved by finding the partial derivatives of equation 7.55 with respect to each coefficient λ_i and by introducing a Lagrange multiplier L.

The unique combination of weights λ_i for which the estimation variance is minimized can be obtained when:

$$\sum_{i=1}^{n} \lambda_i \gamma(X_i, X_j) + L = \gamma(X_j, Xo) \quad \text{for } j=1,2,..n \qquad (7.57)$$

The solution of the n+1 simultaneous equations in the kriging system, (eq. 7.52) for each λ_i and L enables the estimation of Z(Xo) by equation 7.51. The estimation variance can then be determined by solving for:

$$\sigma^2 = \sum_{i=1}^{n} \lambda_i \gamma(X_i, Xo) + L \qquad (7.58)$$

The kriging system is best solved by a matrix approach. The coefficients or weights can be estimated by:

$$\begin{bmatrix} \lambda \\ L \end{bmatrix} = A^{-1} b \qquad (7.59)$$

where

$$A = \begin{bmatrix} \gamma(x_1,x_1) & \gamma(x_2,x_1) & \cdots & \gamma(x_n,x_1) & 1 \\ \gamma(x_1,x_2) & \gamma(x_2,x_2) & \cdots & \gamma(x_n,x_2) & 1 \\ \cdot & \cdot & & \cdot & \\ \cdot & \cdot & & \cdot & \\ \cdot & \cdot & & \cdot & \\ \gamma(x_1,x_n) & \gamma(x_2,x_n) & \cdots & \gamma(x_n,x_n) & 1 \\ 1 & 1 & & 1 & 0 \end{bmatrix},$$

and

$$b = \begin{bmatrix} \gamma(x_1,x_0) \\ \gamma(x_2,x_0) \\ \cdot \\ \cdot \\ \cdot \\ \gamma(x_n,x_0) \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \lambda \\ L \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \cdot \\ \cdot \\ \cdot \\ \lambda_n \\ L \end{bmatrix} \tag{7.60}$$

the vector of coefficients λ can then be found from inversion of matrix A. Similarly, the estimation variance can be obtained from a single inversion of such matrix by:

$$\sigma_E^2 = b^T \begin{bmatrix} \lambda \\ L \end{bmatrix} . \tag{7.61}$$

Thus, there is always a unique solution if matrix A has an inverse. This is conditional on the validity of the semi-variogram. If the semi-variogram for the same data is CNSD then matrix A has an inverse. Further, if the soil property varies isotropically and there are no changing local drifts punctual kriging is an exact interpolator, that is $\hat{Z}(Xo) = Z(Xo)$. On this being true then the estimation variance σ_E^2 is zero if the nugget is zero and they are equivalent otherwise.

The interpolated value at any unsampled location Xo is the most precise possible from the available data. Measures of the confidence of this value can also be known through the estimation variance. The estimation variance depends only on the semi-variogram and on the configuration of sampled points to kriged points and not on the values of the observations.

The size of matrix A is of much importance. The inversion of too large a matrix for every point may make computing time prohibitive. For interpolation onto a regular grid kriging becomes a formidable task. However, matrix A is inverted not for every point X_i, X_j. In practice the support for the interpolation is confined to a certain number of nearest neighbours only. The coefficients λ_i associated with each neighbouring point become more dependent on the configuration of the sampling points as the nugget variance decreases. So the nearest few samples to the interpolated point are the most heavily weighted. When there is no significant nugget variance the weights of the points out in the periphery of the first ring are very small (Burgess, 1979), and for those far apart the weight is nothing. This has very important practical implications. It means that by ommitting points of no consequence, distant to Xo, whose weight is rather small the size of matrix A can be maintained within manageable dimensions, say no larger than 20x20. An improved algorithm is still needed to minimize the number of matrix inversions.

Burgess and Webster (1980a) found that for data on a square grid the nearest 16 to 25 points are usually sufficient to give an accurate estimate. For irregularly distributed points the 3 nearest points within any octant, for a number of octants surrounding Xo may suffice. In any situation, less than 10 neighbouring points should never be used for kriging.

Since the few nearest points to be interpolated are the most heavily weighted, consequently, the semi-variogram needs to be fairly accurate only over the first few lags.

It can be shown that the kriging technique is robust in that over the first few lags practically the same estimates are obtained regardless of the procedure used for fitting the semi-variogram model. Least squares techniques and the procedures suggested by Journel and Huijbregts (1978), for semi-variogram

model fitting would yield somewhat similar kriging estimates provided the parameters of the model are correctly estimated.

Punctual kriging does not interpolate strictly at points but at small blocks. These blocks have exactly the same size and shape as the size of support,i.e. few inches diameter if sampling was made by auger.

Although block kriging is not the concern of this thesis, it can be shown that, in mapping, block kriging will produce smoother maps or 3-D block diagrams than those by punctual kriging which displays local irregularities around sampling points. When the whole pattern of variation within an area is of more interest than the detail, block kriging is to be preferred. An in-depth presentation of block kriging can be found in Journel and Huijbregts (1978), and Webster (1985).

7.3.3.12 The estimation variance and optimal sampling.

A key feature of kriging for practical applications in soil survey is that revealed by equations 7.58 and 7.61: The estimation variance depends only on the semi-variogram and on the configuration of the sampling points and the distances between these and the point to be estimated and not on the values Z of the observations.

Hence, if the representation of the true semi-variogram by a model is accurate the estimation variance is a reliable estimate of the interpolation error. Thus, provided the semi-variogram is known already the estimation variances from kriging for a given sampling configuration can be determined in advance even before sampling takes place.

The estimation variance tends to increase with the distance (d) between the interpolated point or block and its observed neighbours. Hence given a configuration of samples, a sampling spacing for which the estimation variance is minimum can be found by shortening the distance (d), which amounts to sampling more intensively. Sampling densities and configurations can then be combined in designing optimal sampling shemes for local estimation. A good sampling scheme is one which minimizes the maximum kriging variance. This can be used as criterion to judge

sampling schemes for soil surveys for given budgets, i.e. sampling densities.

Burgess et al (1981) and McBratney et al (1981) have developed methods for optimal sampling for local estimation based on these principles. Their strategies are optimal in the sense that the sampling effort they yield is the least possible for achieving a desired precision.

The semi-variogram is a fundamental tool for planning surveys. So, given the semi-variogram the value of n, i.e. the sample size, associated with a particular sampling spacing can be read from a graph of estimation variances against the sample size n, for a predetermined acceptable error.

McBratney et al. (1981), provide a computer program and sampling theory for the design of optimal sampling schemes. McBratney and Webster (1983), report gains of upto 3 to 9 fold in sampling efficiency over that estimated by classical statistical theory.

In survey practice the semi-variogram will be unknown in most situations. Pre-survey reconnaissance sampling will have then to be undertaken as a portion of the total sampling to estimate the semi-variogram. Various alternative sampling procedures for tackling this problem have been suggested. When the semi-variogram is unknown McBratney and Webster (1983), recommend as the best strategy to sample on a regular grid with grid spacing determined by the number of observations that can be afforded. As an alternative, sampling along transects is recommended to obtain the reconnaissance semi-variogram and from its interpretation the optimal sampling for kriging can be determined.

Based on the principles outlined above, the semi-variogram as a reconnaissance tool is used with the Yelford data in chapter 9 to find out sampling spacings that would have been required to achieve a given level of precision in the estimation of properties at unvisited sites by kriging.

The semi-variogram and kriging provide the best tools so far for soil survey design and optimal sampling schemes. However, the cost-effectiveness of sampling schemes in practical survey operations has yet to be evaluated in the field against conventional soil survey sampling techniques for different types of terrain and under different budget constraints.

7.3.3.13 Non-stationarity and Universal Kriging.

Ordinary kriging is a suitable interpolation method when data are stationary or quasi-stationary following the intrinsic hypothesis.

It is not clear, however, when a trend present in the data can be considered sufficiently strong to require a method other than ordinary kriging for interpolation.

Universal kriging or unbiased kriging of the kth order (Journel and Huijbregts, 1978) is a method designed to permit kriging of non-stationary phenomena due to the presence of trends in the data.

Universal kriging assumes that the local trend or drift within each neighbourhood V is linear or parabolic at most. So the property Z is now a non-stationary random function (equation 7.48) which has an unknown drift. This drift is of the form

$$E\,[\,Z(X)] = m(x) = \sum_{i=1}^{n} a_i f_i(x) \qquad (7.62)$$

where the functions fi(x) can be known.

The estimation of the drift has been approached in several ways. Global polynomials fitted to the entire region have been used in the past to remove the trends and estimate the semi-variogram from the residuals (David, 1977). The disadvantages of these surfaces have been discussed in chapter 6. Burgess (1979), and Webster and Burgess (1980c) use detailed structural analysis as proposed by Olea (1975) to estimate changing drifts.

The drift is removed from neighbourhoods rather than for the whole region by finding suitable expressions for it within each neighbourhood, so estimating it locally with independent coefficients for each neighbourhood.

Once the expressions for the drifts are found universal kriging can be performed by incorporating the expressions fi(Xi) of the trends into the expanded matrix A (eq. 7.59). Details of the matricial representation can be found in Olea (1975) and Webster and Burgess (1980c).

Universal kriging is designed to overcome the dependence of simple kriging on stationarity of the data. So it can be thought universally applicable. However, for soil data it has not been found of particular benefit for interpolation. Webster and Burgess (1980c) argue that since soil data are noisy they usually show large nugget variances. Point-to-point variation is so large that it obscures the distinction between changing drifts and constant drifts. The coefficients for the drifts can not be estimated confidently by a complicated structural analysis. Only in circumstances where nugget variances are negligible such as in data taken from contiguous parcels or blocks of soil, is there scope for universal krigi ng. In any other circumstances, where there are considerable nugget effects, point or block kriging are best used.

Yost et al (1982b) found that universal kriging resulted in very little improvement over ordinary kriging when comparing the observed and the estimated points by "jack-knifing", i.e. removing one point and estimating it from its neighbours and then comparing.

The effects of ignoring trends in data on ordinary kriging have been determined by Faith and Sheshinski (1979), Starks and Fang (1982), and Yost et. al. (1982b). Their results indicate invariably that ordinary kriging is fairly robust to the presence of trends and non-stationarity.

Based on this evidence universal kriging is not used in this thesis to interpolate with non-stationary data. Cases of non-stationarity in the area of Yelford are first identified by analysis of background soil, geology and landscape information. The form of the semi-variogram was then used to determine non-stationarity due to the presence of trends. Various other procedures have been suggested to check for non-stationarity, e.g. Starks and Fang (1982) suggest to check for non-stationarity by inspection of the data for trends and a procedure called "jack-knife", i.e. comparing kriging estimates with the actual data and the variance of actual kriged values with the estimation variances.

The trends identified in the Yelford area (see chapter 9) were removed by fitting local bicubic splines on intervals approximately of the same size as the neighbourhoods and so

avoiding the complications of the structural analysis. The residuals from the spline surfaces were then used to estimate the semi-variogram and for kriging. The kriged estimates were then added to the values interpolated by the surfaces at the wanted points to provide the final estimates.

7.4 Concluding remarks to this chapter.

i. Mathematical interpolation techniques and surrogate functions offer a solution to the problem of missing soil information in the multidimensional character space.

ii. Missing values of soil properties can be generated from the existing values of other properties acting as their surrogates, by means of a predictive model on the surrogates.

iii. The equal-area spline algorithm provides the best method so far, for reconstructing soil depth functions and for interpolation of properties through soil depth.

iv. Regionalized variable theory offers a sound body of theory for the quantitative description of soil spatial variability.

v. The semi-variogram is the main contribution of regionalized variable theory to spatial analysis of soil variation, local estimation and optimal sampling. It is an excellent tool for discovering spatial structures in the data.

vi. The range of spatial dependence in the semi-variogram serves two main purposes: it defines the neighbourhoods for local estimation and it conveys essential information for the kriging estimation.

vii. The interpretation of the semi-variogram against background landscape information is the key to successful spatial prediction by kriging.

viii. Anisotropies as detected by the semi-variogram can be modelled by a single linear semi-variogram with slope estimated from geometrical functions enabling interpolation by kriging.

ix. Amongst the mathematical techniques for spatial prediction

of missing information kriging is the most advanced and the best. Its success depends on the semi-variogram and its correct interpretation.

x. There is evidence to suggest that kriging is robust to the presence of local trends.

xi. Kriging is sufficiently robust to the choice of semi-variogram model from two or more equally close alternatives, and to the procedure for fitting the model, provided the region within the range can be approximated reasonably well and the parameters estimated with accuracy.

xii. The modelling of drifts and local trends allows for local estimation by kriging from the semi-variogram of the residuals from the trends.

xiii. Semi-variograms and kriging are advanced techniques which equip the soil surveyor with powerful tools for predicting soil data at unvisited sites, and to provide the missing soil information essential for land evaluation and land-use planning.

VIII. OXSIS DATA-MANIPULATION PACKAGE.

The capabilities of the Soil Information System at the Soil Science Laboratory, Oxford (OXSIS), have been outlined in chapter 5. Input of data, query about the information in the system, and retrieval of data are operations that can be performed in OXSIS.

The virtues of the semi-variogram and kriging for analysing the spatial structures in data and for interpolation have been discussed in previous chapters. It is clear that in order to equip users/soil-surveyors with capabilities to predict values of soil properties at any site and at any depth within a surveyed area, software for depth-functions, semi-variograms, and kriging must be part of the information system. The advantages of incorporating these capabilities on-line have been discussed in chapters 5, 6, and 7. Such capabilities make SIS a powerful tool for prediction.

In this chapter the data manipulation capabilities of OXSIS are described. Apart from the statistical analysis and graphical output packages or interfaces to packages which are almost standard to any SIS, the OXSIS Manipulation package contains the new capabilities for spatial analysis and spatial and depth interpolation of soil properties.

The theoretical background to these algorithms has been discussed at some length in various sections of chapter 7. These algorithms provide solutions for the special problems on interpolation discussed in chapter 6.

Thus, at the end of the retrieval stage a user of the system has the data-set of interest in the form of a retrieved subset (RS) in a file named by himself. The RS is in OXSIS format like its parent triad, that is it is made up of data, i.e. D-file, list of features recorded, i.e. N-file, and a list of locations, i.e. L-file. The user now wants to manipulate it.

Manipulation comprises the whole set of operations on a retrieved subset which enables the re-arranging and transformation of retrieved data and the generation of new data from these to

extract more information and meaning about particular aspects of the original data.

8.1 Aim.

OXSIS manipulation package aims at providing the necessary capabilities to equip users and particularly the soil surveyor, starting from retrieved data-sets, to make statements about or to predict the conditions of any property of the soil at any depth, and at any point within the area covered by the data in the system. Predictions of this sort at unvisited sites involve spatial and depth interpolation capabilities to cover soil variation in the three dimensional space.

In addition to the facilities above, file-handling and organizing routines (i.e."house-keeping") complete the capabilities of OXSIS for data manipulation.

8.2 General view of the package.

In physical terms the manipulation package in OXSIS is a set of coded master files organized hierarchically. The software in such files is a set of nested and modular structures consisting of main programs and subroutines written in FORTRAN-77 computer language.

As well as the input and retrieval packages, manipulation in OXSIS is "user-friendly". It enables the user to specify the operations he wants to perform on a RS through a man-machine interactive dialogue. To the user's view the manipulation stage consists merely of a set of questionnaires. Questions regarding the nature of and the parameters necessary for a particular analysis are posed on the VDU terminal screen to the user, to be answered in their turn. The questionnaire's configuration is simple and the user only needs to know exactly what operations he wants to perform.

Except for the high resolution graphics, which are device-dependent most manipulation facilities are on-line. To the users perception the manipulation package is black-boxed within OXSIS and appears as an integrated unit.

8.3 Hardware.

Manipulation is implemented and performed wholly in the Oxford University Computing Service's (OUCS) VAX mini-computer on a time-sharing basis. Therefore the high hierarchy commands it uses have to be compatible with the VMS-4.3 operating system commands implemented in such a machine. OXSIS retrieval procedure is executed in the OUCS ICL-2988D mainframe computer to produce the RS. Therefore this has to be transferred to the VAX mini-computer via the communications link between the two machines before entering the manipulation procedure. The three files (D, N, L) which make-up the RS can be "fetched" from the main-frame computer while in a VAX interactive session by using the OXSIS command "FETCH". This activates and runs a command file with the instructions for the transference of each file of the RS at a time. Names of the RS files are supplied interactively. Once successfully run this leaves the RS in the VAX machine ready to be manipulated.

The manipulation procedure uses the same peripherals as the rest of OXSIS procedures, namely, remote terminals (VT100 and Dacol) and line-printers (i.e., OUCS user-area line-printer and the Plant Sciences Department's line-printer). In addition to these, Tektronix-4010 high resolution terminals and Calcomp-1012 and Hewlett Packard-7475A plotters in OUCS user-area are used off-line, for the production of graphical output. Figure 8.1 displays the array of hardware components and the data flow between them during the manipulation stage.

8.4 Software design.

One of the main considerations in designing the OXSIS manipulation package has been that it should allow for growth by the inclusion of new capabilities at any time. This is achieved through giving a modular design to the software within the framework of a hierarchical structure of command -files. Such modular design also permits easy debugging, editing and maintenance.

"User-friendliness" has been the second main consideration in the design of the package, so that the average user may have easy

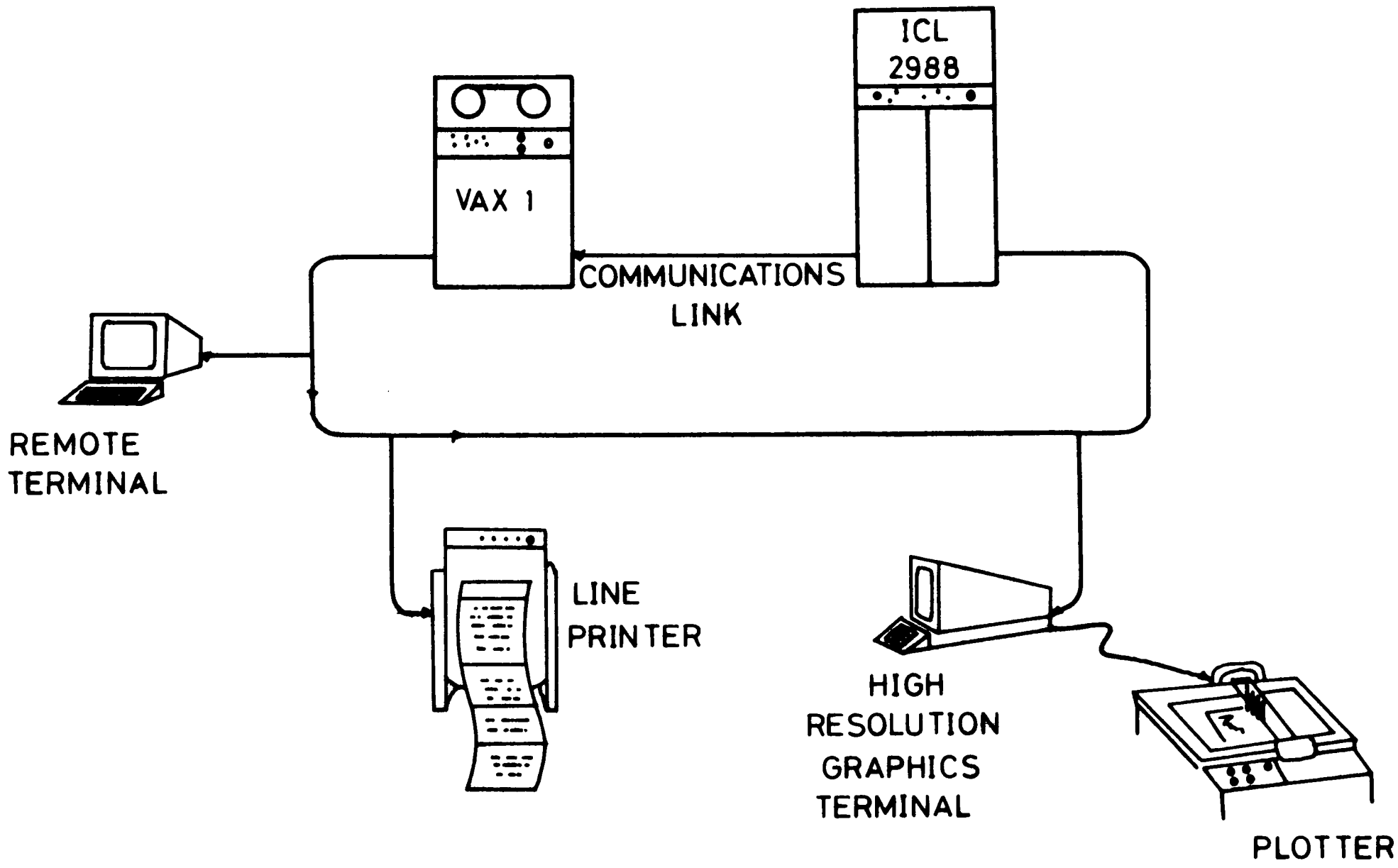


Fig 8.1 Hardware in OXSIS-Manipulation Procedure.

access even to the most elaborated applications programs. This is achieved through "menus" requiring a simple man-machine interactive dialogue.

In summary then, the potential for expantion and user-friendliness have been the chief features that have guided the software development and are now reflected on its design.

The manipulation procedure has two main programs, one is designed to work interactively, to collect instructions and parameters from the questionaires answered by the user. The other is an executive program in the sense that it consists of the algorithms themselves which execute or perform the calculations and "number crunching" as read from a instructions file created at the end of running the interactive program. The structure of these two main programs in terms of files, commands, macros and subroutines is described in appendix 8.

8.4.1 Files.

In essence, the manipulation package is a set of files of different types. Some of them are part of the system itself and some others are created and even deleted during its running.

Five file types can be conveniently distinguished within this procedure:

i. Instruction files (IF).
ii. Data files (DF) including D, N and L files.
iii. Command files (CF).
iv. Results files (RF).
v. Temporary files (TF), from intermediate steps during processing.
vi. Program master files (PMF).

If the RS or TF have been generated during the OXSIS manipulation procedure they are considered internally as manipulated files (MF) and to distinguish them from the rest the first letter of their name is automatically substituted by an "M" letter by the system. Any changes of this sort are notified to the user on the terminal screen. The PMF are the two main programs (interactive and executive) and contain only FORTRAN-77 executable statements and algorithms. The algorithms are structured in terms of main programs and numerous sub-routines.

Each of these sub-routines is associated to a particular facility offered by the package. In their turn these sub-routines have their own numerous sub-routines resulting from the division of tasks performed by the facility. They each form a sub-package in their own right.

8.4.2 Interactive dialogue.

Two main components can be distinguished in the manipulation package, i.e. an interactive PMF and an executive PMF. These two programs are activated by the command "MANIPULATION". The interactive part contains a "MENU" and a set of questionnaires. There is an ad-hoc questionnaire or set of questionnaires for each facility offered by the package.

The design of both the interactive and executive PMF's is such that they can act as a framework to which new facilities can be added.

8.4.3 Facilities and the executive program.

The facilities available in OXSIS for data manipulation can be considered in 3 categories: "built-in" or internal data processing routines, house keeping routines and "bridging" routines or interfaces to external software packages.

The built-in routines operate on data in OXSIS-format and when required for further processing some of them produce output in OXSIS-format too. These internal routines provide capabilities for spatial statistics, depth and spatial interpolation and other data analysis.

"House keeping" routines are also "built-in". They assist in file-handling, sorting and in general re-structuring and organizing data-sets. House-keping algorithms also operate on input and output of OXSIS-formatted data.

There is a wide range of commercially available software packages for statistical analysis and graphical output. Depending on their flexiblty some are more difficult to interface than others. The Statistical Analysis Sytem (SAS, SAS Institute Inc. 1981), available in the OUCS-VAX machines was chosen to provide extensive statistical analysis capabilities to OXSIS via a

suitable interface. Bridging routines had to be developed to transform data from OXSIS-format into SAS-readable format. In parallel, routines were also developed for getting the system to "translate" the user's instructions in the IF and to "write" its own SAS-commands program onto a SAS file according to such instructions. SAS statistical procedures have their own output format and there is no access to them. So, SAS output has to be an end product and somewhere at the end of the data processing line, for no reverting of the output back into OXSIS for further processing is possible.

In this category also are the two graphical packages that can be used "off-line" on OXSIS data. GHOST, a graphical output system (Calderbank and Prior, 1980), and SAS/Graph, (SAS Institute Inc. 1981). GHOST programs can be run "on-line" to create grid files but the display and plotting of the actual graphs is device-dependent and consequently "off-line". By necessity SAS/Graph is run interactively, hence developing plots and graphs from OXSIS output has to be done interactively too via the high resolution graphic terminals. The manipulation interface facility for SAS/Graph only has to transform the OXSIS format or the data to a SAS readable format.

OXSIS-Manipulation has been developed so far to include the following list of facilities:

1. **"ALLOCATION"**
2. **"AVERAGE"**
3. **"KRIGING"**
4. **"LISTING"**
5. **"MERGE"**
6. **"SORT"**
8. **"DEPTH FUNCTIONS"**
9. **"STATISTICS"**
10. **"SEMI-VARIOGRAM"**

"MERGE", "SORT", "SELECTION" and "ALLOCATION" are "house-keeping" facilities. Their result is always other data subsets in OXSIS-format.

"AVERAGE", "DEPTH FUNCTIONS", "SEMI-VARIOGRAMS" and "KRIGING" are built-in internal manipulation facilities. They are the main core of OXSIS capabilities for spatial statistics and depth and

spatial interpolations. "LISTING" transforms data from one or more soil properties in OXSIS-format into a simple format of columns suitable for listing such information or for input to SAS statistical package. Hence this facility functions as a bridging routine for interfacing SAS procedures or SAS/Graph graphical output routines.

"STATISTICS" is a facility which encapsulates the whole range of classical statistical analysis capabilities. It permits access to both univariate and multivariate statistics from within the frame of SAS procedures. An elaborate set of nested questionnaires is used in the interactive stage to guide the user to specify all the characteristics and parameters of the statistical analysis wanted from the large number of statistical procedures offered by SAS. "STATISTICS" facility does not perform the data-processing itself. This is done by the SAS routines. "STATISTICS" collects the user's instructions and transforms them into SAS commands by "writing" its own SAS-commands program. Simultaneously, it changes OXSIS-formatted data into SAS-formatted data.

The manipulation package is not complete yet. Capabilities for handling cartographic and map information are needed. Also algorithms for ad-hoc interpretive classifications of soils. However, the framework is present and could be elaborated further to accomodate vector or raster information and the necessary cartographic display capabilities for automatic map production, and the ad-hoc interpretive routines.

At present the new OXSIS manipulation capabilities are algorithms for spatial analysis and the spatial interpolation of soil properties to equip the soil surveyor with a tool for predicting soil conditions at unvisited sites, and for making statements about the soils in an area.

The complete descriptions of the algorithms of the individual facilities and their use in package are illustrated with examples in appendix 8. The following sections describe the facilities available in the package only in general terms.

8.4.4 Data Flow.

The flow of data through the manipulation procedure may be seen as divided into three major stages. The first stage is mainly concerned with data organisation and preparation. The second stage is fact-finding, i.e. more facts about the original data are found out and knowledge is gained to support decisions for further data processing. The third stage is predictive, i.e. it consists of the use of facilities for predicting soil properties.

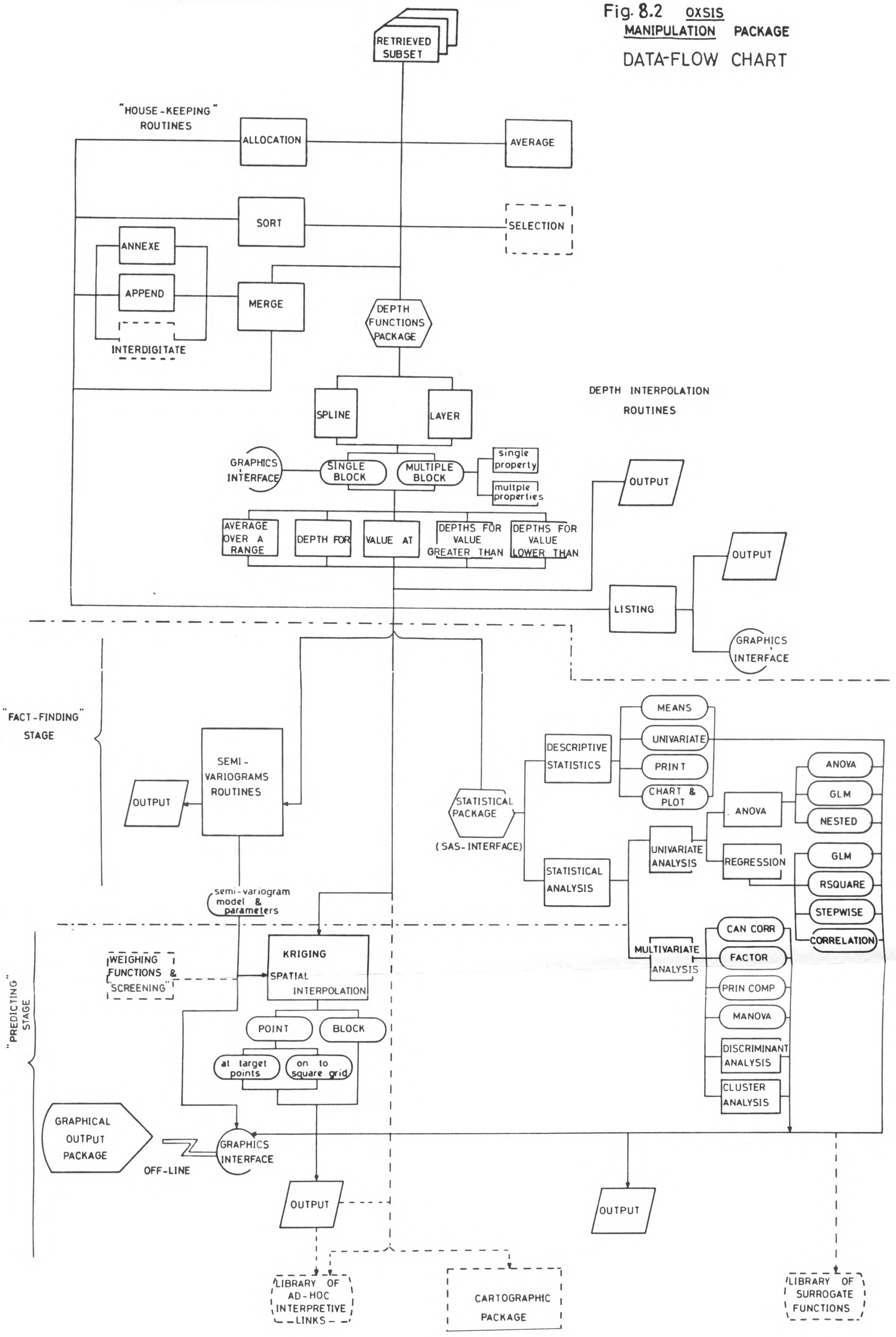
Although in its present state the package has not been elaborated to include it, a fourth stage in the data flow could be envisaged, in which interpretations of soil data could be made and a variety of specific-purpose maps produced. The general data-flow chart of this procedure is shown in figure 8.2. It includes in components not yet in the system, in dashed lines.

From the flow-chart of figure 8.2 it can be seen that the "house-keeping" routines are characteristically of the first stage. At the core of the data-flow is the depth-function package. Since most statistical and spatial analysis and interpolation computations are performed on data from uniform depths, before any of these operations can be performed, the depth-function package needs to be used to interpolate and produce data files with uniform depth. So, data must pass through the depth-functions package before any other data analysis facility. Hence, it is situated at the center of the operations. Output from the "DEPTH" facility can be reverted back to any of the "house-keeping" routines for data restructuring as it is in OXSIS-format, or else reformatted by the "LISTING" facility for output or for graphical display.

On entry to the "fact-finding" stage the manipulated subset (MS) from the "DEPTH" facility contains data on one or more soil features of interest and at a convenient uniform depth or range of depths. It is at this point that the user decides whether he needs to use classical statistics or geostatics to achieve his ends. The data-flow may then turn to the "SEMI-VARIOGRAM" facility to explore spatial dependence and structure and to model it, and to use the model parameters as input for spatial interpolation at the following stage. On the other hand the data-flow may continue at this stage to the "STATISTICS"

Fig. 8.2 OXSIS MANIPULATION PACKAGE

DATA-FLOW CHART



facility. On invoking such facility the OXSIS-SAS interface uses the SAS package, from which either descriptive statistics and summaries of statistical parameters, or a statistical analysis can be performed to provide useful information and knowledge about particular aspects of the data to help in supporting decisions on further data analysis, if desired.

The "SEMI-VARIOGRAM" routines and the "STATISTICS" package constitute the two tools for the fact-finding stage within "MANIPULATION".

The predictive stage in the data-flow line is reached when all the information needed for interpolation and prediction has been obtained. Such information may be in the form of code or a data file, but also in the form of knowledge about the nature of the data. Naturally there are operations that take place outside the computer. Evaluation of results, decision-making in data analysis and many other steps which need human intervention, are reflected implicity in the data-flow chart of figure 8.2.

The predictive stage is constituted typically by the "KRIGING" facility, but depth interpolation and regression models could also be considered part of this stage. To interpolate and predict by "KRIGING" data need to be contained in a MS which has already been depth-interpolated. The MS comes directly from the "DEPTH" facility. Other not explicit concomitant information is also needed for an interpolation plan. This information is drawn from spatial analysis by means of the semi-variogram computations.

The "KRIGING" facility interpolates values of a given soil property at points, (i.e. Point Kriging) or over small parcels or blocks (i.e. Block Kriging). When interpolating at points, the interpolated values can be located on a square grid of spacings decided by the user or at "target" points specified interactively by a pair of co-ordinates for each target.

The output produced by "KRIGING" includes the grid point co-ordinates, the estimates of the soil property and the variances of the estimates. An output file with a listing of co-ordinates and estimated values for each point serves as input for the graphics interface. The graphical output is produced "off-line" and may consist of a map of contours over the estimated values, or a three-dimensional block diagram of such values. This so far, is the final product of the predictive stage. Although it is not

indicated in figure 8.2, an algorithm for the computation of bicubic spline surfaces has also been included as a late addition into the manipulation package. The algorithm interfaces two external algorithms in the library of routines of NAG (Numerical Algorithms Group), for the computation of the surfaces and for their evaluation at any point for interpolation. This routine can also serve as interpolator in its own right but at the predictive stage it is most useful as a tool for modelling trends in the data during structural analysis of non-stationary soil properties, (see chapter 7).

As stated above, the design of the manipulation procedure allows for expansion. We may envisage the growth of the package to include a library of routines for determining ad-hoc interpretive algorithms of soil data for various practical purposes. Interface to other data-bases (e.g. crop yield records, experimental data, climate, engineering data) may also be found necessary in the future.

The need to output information in map form could lead to the inclusion of a whole cartographic package of routines for production of soil and ad-hoc interpretive maps.

A library of "surrogate functions" for estimating values of unrecorded properties from their surrogates will also be an addition to the manipulation package.

These as yet unbuilt facilities are shown in the data-flow chart of figure 8.2 with dashed lines and they represent the end of the flow of the data through the system.

8.5 Using the OXSIS-Manipulation Facilities.

This section describes briefly the individual facilities offered by the Manipulation procedure and their use. The algorithms for each of the facilities in turn are outlined and examples of the kinds of output produced by the various options within each facility are given in appendix 8. There, the use of Manipulation is illustrated by "screens" or frames representing the output viewed by the user on the VDU of the computer terminal.

8.5.1 Allocation.

This facility allocates blocks of information from a RS into subgroups or classes specified by the user. For each class the corresponding blocks are placed into a new D-file and its N and L-files that form the triad are also created. There may then be as many new triads from the parent RS as classes indicated by the user. The algorithm and flow-chart for this facility together with some examples are given in appendix 8.

8.5.2 Sort Facility.

The "SORT" facility comprises a set of routines for re-arranging the internal order of blocks within a retrieved subset. Data blocks may be sorted according to the increasing value of a soil or site property directly specified by the user during an interactive session. The L-file order in the RS is also altered in accordance with the sorting in the D-file.

The algorithm, its flow-chart and displays of the interactive and executive parts of this facility are given in appendix 8.

8.5.3 Merge Facility.

The contents of two triads from different RS can be merged into one by using the facility "MERGE".

Three possible ways in which the two triads could be merged have been considered as sub-facilities:

a). **Annexe.** The contents of one of the triads comes from the same sites as that of a second triad but their records are of different variables or soil properties. They can be said to be complementary for they have data from the same site numbers but on different properties and formats. Hence they can be annexed to each existing data block identified by its site number. In this case the L-file remains unchanged but new additions to the N-file have to be made to include the annexed variable names.

b). **Append.** The two RS have records on exactly the same properties and in the same formats but their data come from different sample sites. They do not overlap. In this case

additions have to be made to the L-file but not to the N-file. Because of the lack of areal overlap the data in one RS is enlarged by appending the data from the second RS at the bottom of the file. The merged data is stored in a new named file.

c). **Interdigitate.** The two RS are in a similar situation as in (b) above but in this case the areas covered by their data overlap. That is, the two files have data in the same format but they do not contain the same blocks (site numbers). "Interdigitate" merges the two RS in site number order and places the result in a new named file.

The N-file remains unaltered and only the L-file has to be added-up to.

The algorithm for the "ANNEXE" sub-facility serves to illustrate this facility in terms of the flow-chart and examples in appendix 8.

8.5.4 Depth Facility.

The "DEPTH" facility is a set of algorithms for solving problems of depth interpolation. The package is divided into two main sets of algorithms to answer questions of practical significance which need a variety of forms of depth interpolations, namely, the SPLINE and the LAYER subpackages.

The SPLINE set of algorithms is of much wider and more general application than the LAYER algorithms. The former uses a formal numerical interpolation procedure. The latter only extracts the mean values of the horizons or layers, (assigned to their mid-points) within which the depth submitted by the user is found.

The SPLINE and LAYER subfacilities are sufficiently flexible in regard to the data in that they can be applied to either, a single block (soil profile) or to a part of all of the blocks in a RS. These "SINGLE" or "MULTIPLE" options are integrated as part of both subfacilities. Further, in the "MULTIPLE" option there is an extra choice that one single or many soil properties may be interpolated in the same run. This array of nested facilities and options within the DEPTH package is better illustrated in the general data-flow chart of figure 8.2.

The philosophy behind the "DEPTH" package design is that most interpolation cases can be solved by the "SPLINE" subfacility. Within this, the package can provide two contrasting levels of detail in the output of information about the interpolation. For the single-block case the output contains detailed information. In contrast there is the "industrial" or "MULTIPLE" processing of interpolations on, say, all blocks in a RS. Here there is a trade-off between detail and number of blocks that can be processed and not much detail can be afforded. In the "industrial" mode what is wanted are only the interpolated values of each of all or part of all the blocks in a RS for further manipulations.

The DEPTH package can provide different sorts of information derived from interpolation to answer the range of most common queries involving depth-distributed soil data. Such queries are grouped here into five categories. Users may want to find:

a). The value of a property at a given depth.

b). The average value of a property over a range of depths

c). The depth(s) at which the soil property shows a given value.

d). The depth(s) at which the value of a property is greater than a known value.

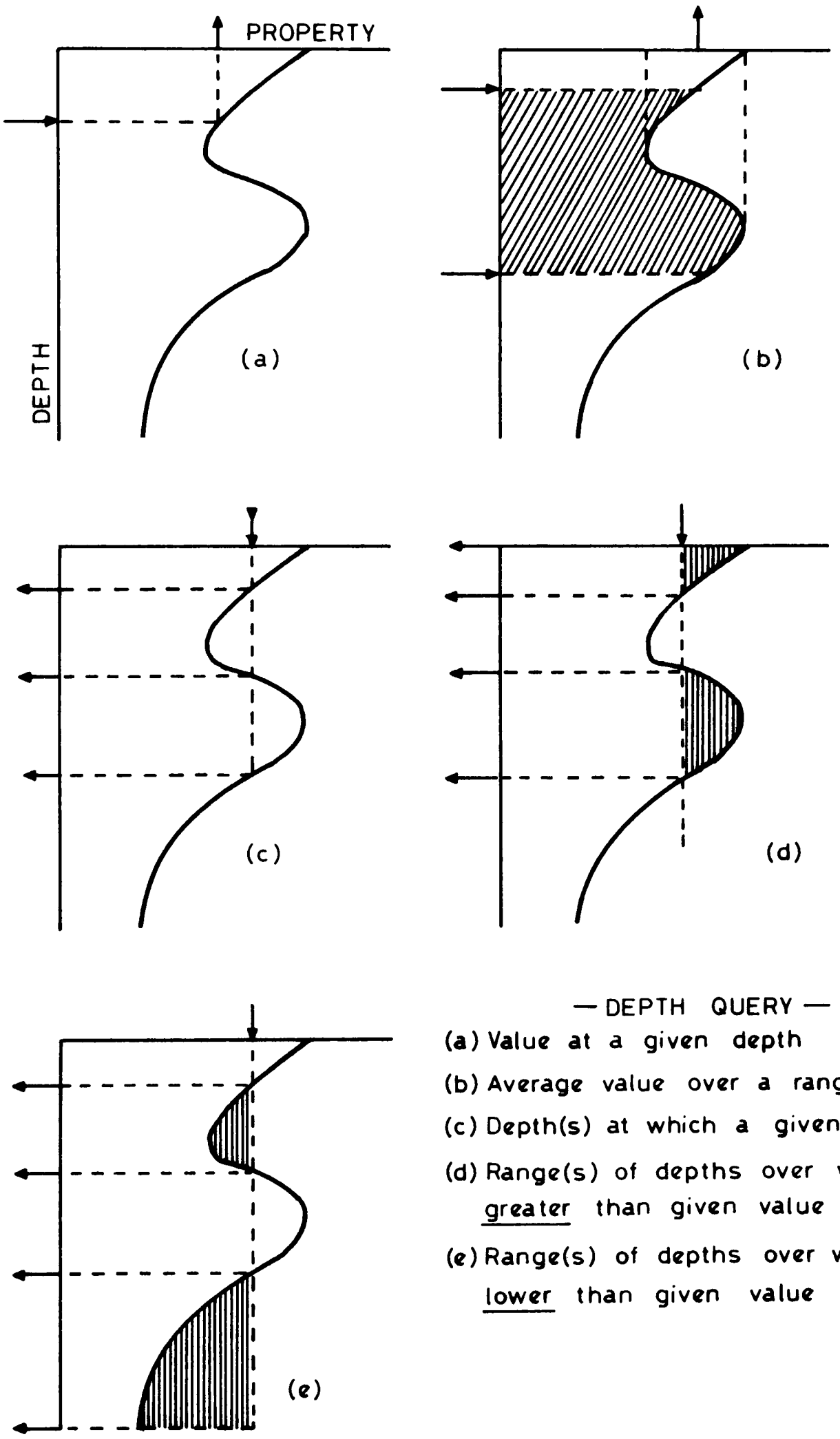
e). The depth(s) at which the value of a property is lower than a known value.

In the DEPTH facility five algorithms have been created to provide the sort of information required in each of the five categories above. These algorithms shall be called here the Depth Queries. They are applied to both the "SPLINE" and the "LAYER" interpolation subfacilities. So, information to answer any of the five queries above may come from a SPLINE interpolation or from a LAYER extraction. They are illustrated graphically in figures 8.3, and 8.4. The algorithms are also shown in the general data-flow chart of figure 8.2.

8.5.4.1 The SPLINE Algorithms.

The SPLINE algorithm used as interpolator in the DEPTH package operates on a similar principle as the curve-fitting by cubic splines (see chapter 7).

Fig. 8.3 — SPLINE INTERPOLATION —



— DEPTH QUERY —

(a) Value at a given depth

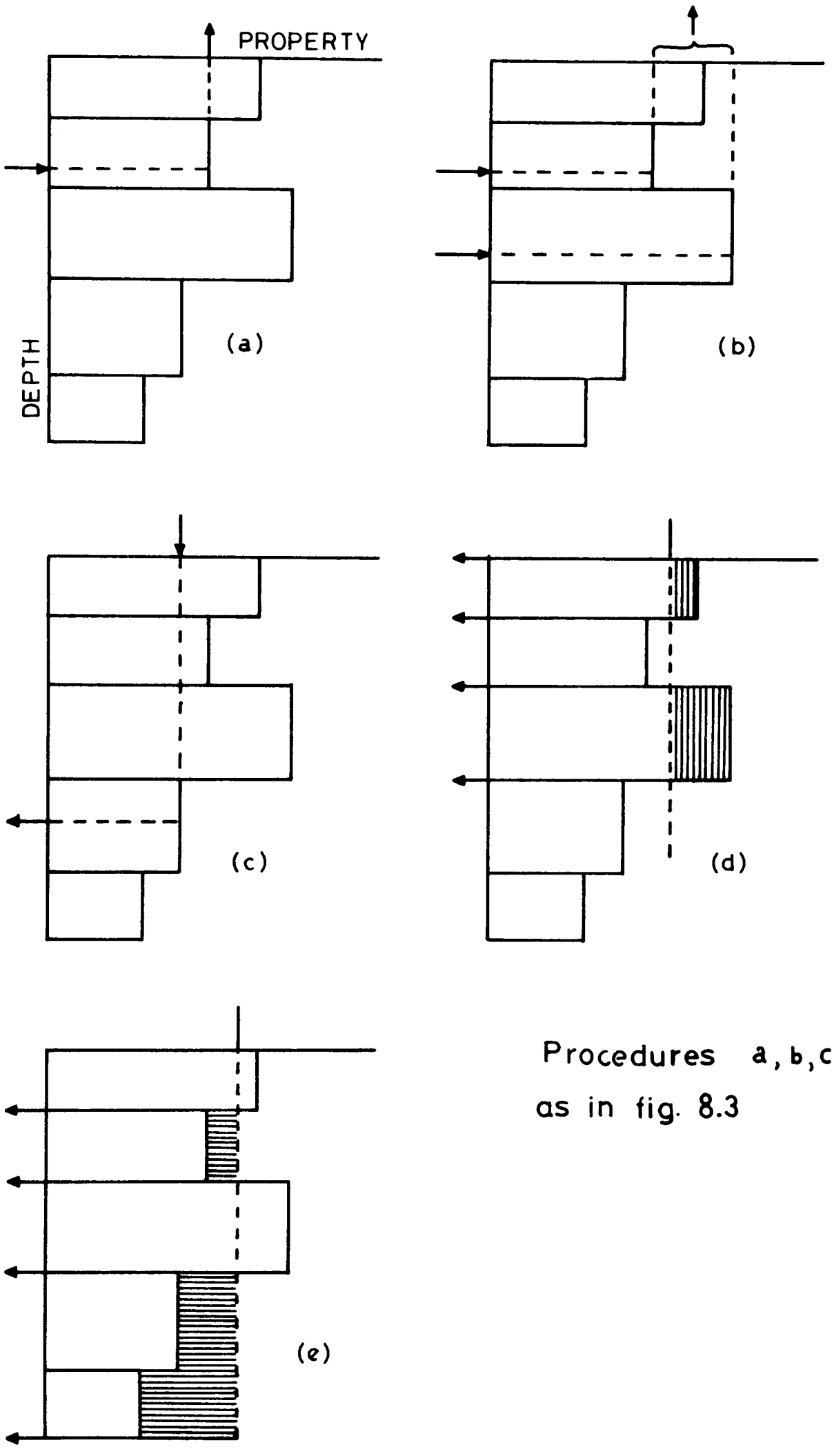
(b) Average value over a range of depths

(c) Depth(s) at which a given value

(d) Range(s) of depths over which feature greater than given value

(e) Range(s) of depths over which feature lower than given value

Fig 8.4 — LAYER EXTRACTION —



Procedures a, b, c, d, e
as in fig. 8.3

The SPLINE algorithm uses a modified spline or "equal-area" spline function (Ponce-Hernandez et al, 1986, see appendix 7.3) to fit smooth curves describing the behaviour of a property through soil depth.

The background theory and mathematical treatment of equal-area splines which serve as the basis for the SPLINE algorithm is fully described in chapter 7 in the section of depth functions.

8.5.4.2 The LAYER algorithm.

As an alternative to SPLINE, the LAYER algorithm is used to extract the values of a soil property or the soil depth(s) for a given value from analytical data. It uses the layers or horizons recorded in the data as uniform brackets within which the analytical value for such layer is assumed to be the same for the whole of its thickness and hence assigned to any depth within it.

The LAYER algorithm is based upon an instruction to find out whether the depth at which a value is wanted lies within the limits of any of the defined layers or horizons in the blocks and if so the value of the feature attributed to such layer is extracted and output as the answer to the depth query. Alternatively if a depth is requested for an input value, the values of all the layers are compared with this and where applicable the horizon's mid-point is extracted as the depth wanted and then output.

All the five depth queries that are used in SPLINE are also used in the LAYER algorithm though the values are not interpolated ones but merely extracted when applicable according to the depth query. This is illustrated in figure 8.4. above.

8.5.5 LISTING facility.

This facility in manipulation serves two main purposes. On the one hand it transforms the data format from OXSIS-format to an easy to read output listing of columns. Co-ordinates (i.e. eastings and northings) and site numbers are output together with the corresponding values of a property of interest for each site. On the other hand this facility is used as a bridging routine to interface the Statistical Analysis System package

(SAS), and its graphics package SAS/GRAPH by transforming to a SAS-compatible format.

The use of this facility presupposes that the values of the property come from uniform depths and therefore that the data in the RS have been depth-interpolated. This can be appreciated in the data-flow chart of figure 8.2.

8.5.6 STATISTICS facility.

As stated above the STATISTICS facility in manipulation offers capabilities for a wide range of statistical analysis of data. OXSIS capabilities build on an interface to the SAS package (Statistical Analysis System, SAS Institute Ltd, 1981).

Descriptive statistics as well as univariate and multivariate analysis can be performed by this facility in two modes of operation; interactively and automatic (i.e. OXSIS-controlled).

The range of capabilities of the STATISTICS facility is displayed in figure 8.2.

The STATISTICS interface was designed with two types of user in mind. The first is a user not acquainted with the statistical package and with very limited knowledge of computing and programming. For this type of user the statistical package (SAS) is black-boxed and to perform his analysis his efforts are limited to answer a set of OXSIS interactive questionnaires. To this type of user it would appear as if OXSIS is totally in charge of the "number-crunching" operations which are carried out by SAS. This is the automatic or OXSIS-controlled mode of operation.

The second type of user has some knowledge and skills in using the computer and is acquainted with SAS. He is prepared to write a set of SAS commands during an interactive session to instruct the package in some detail on the sort of options and analysis he wants. This is the interactive mode of operation.

For a successful run of STATISTICS interfacing SAS two things are required. First, the existence of a SAS-formatted data file and second the necessary SAS commands to operate on such data.

On entry to this facility the user selects between interactive and automatic operation modes. The interactive mode would usually be chosen by the second type of user. The OXSIS interface then selects and extracts the properties and sites required from the

depth-interpolated RS (now a MS) and re-formats them onto a SAS-readable format in a SAS-data file. The SAS commands are supplied by the user interactively.

If the automatic mode was selected OXSIS creates the SAS data file as above and additionally a SAS-commands file is created to run and operate on the data. The SAS commands file is created internally by "translating" the instructions from the IF to SAS-commands and by writing these commands in the appropriate order in a new file created for that purpose. In that sense the interface is a subroutine which" WRITES" its own SAS-program.

The facility is divided first into two major subroutines one for Descriptive Statistics and one for Statistical Analysis. The latter makes several calls to other subroutines depending on the SAS procedure (e.g. regression, correlation, analysis of variance) requested. There is one subroutine for each major SAS procedure and these in their turn may proceed further to call other subroutines needed to complete and round out the analysis as specified in the IF. This can be seen in figure 8.2. It is worth mentioning that the hierarchy of fig 8.2 is identical for both the interactive part of the facility which captures and writes instructions in the IF, and for the executive part which reads instructions, creates the SAS data-file and writes SAS commands on to a SAS commands file in agreement to the procedure requested. The structure of the subroutines, that make up this facility, their flow-charts and the programs are placed in appendix 8, together with an example.

8.5.7 SEMI-VARIOGRAM Facility.

The SEMI-VARIOGRAM facility is a group of subroutines which compute the experimental semi-variogram for data from irregularly distributed samples in two-dimensional space in the plane.

The Semi-variogram is a technique for measuring the spatial dependence of soil properties. It is a function for describing the degree of similarity or dissimilarity between neighbouring sampling points, which is essential for the structural analysis of spatial phenomena and its use in spatial interpolation. The relevant background theory to the semi-variogram, structural analysis and interpolation are reviewed in chapter 7. This

section concentrates only on the semi-variogram as a facility of the Manipulation procedure.

Two kinds of information are output by this facility: the semi-variances calculated from the points in the data-set and their conditions of estimation, and a plot of the semi-variances against lag distance.

The SEMI-VARIOGRAM facility in OXSIS-Manipulation builds on an algorithm by Webster (1985), to compute the experimental semi-variogram from irregularly distributed samples in the plane. Dr. R. Webster (Rothamsted Experimental Station) kindly provided software to compute the semi-variances and to draw semi-variogram models. His software suffered some modifications so it could be adapted to OXSIS. More "connective" software had to be developed for such adaptation. Other programs were developed too to make this algorithm work interactively.

The semi-variogram computation presupposes uniformity in depth and uses the data achieved by depth interpolation previous to using this facility (see data-flow in fig. 8.2).

The lag distances and directions are chosen in regular progressions. For irregularly distributed data a range for the classification of distances equal to the class interval between successive lags is chosen. The nominal lag distance lies at the centre of this range. The squared differences of any points within such a range then contribute to the semi-variance for that particular lag class by virtue of their actual distance. The directions are chosen in degrees counter-clockwise from east. A range in direction is also chosen so that the nominal direction passes at the centre of the arc subtended by such range. This is the angular regularization and in this routine sets it to 22.5 degrees by default.

The algorithm has a capacity for computing the semi-variograms for up to 3500 samples in 10 directions and up to 40 "steps" or distance classes in each direction. As stated above, this algorithm has been modified into a set of subroutines and adapted to the design of the Manipulation procedure to operate on OXSIS data. This facility is interactive like the rest of OXSIS. The interactive and executive parts of this facility are described in appendix 8. The flow-charts and examples are also included in appendix 8.

8.5.8 KRIGING Facility.

The "KRIGING" facility is a group of subroutines for the spatial interpolation of soil properties at any point within a set of recorded sites in an area. It uses the technique called Kriging which is fully described in chapter 7.

The place at which a property is to be estimated can be a "point" which corresponds to an area of the same size as the area of soil from which a sample was taken, i.e. the size of support. It can also be an area or block. Point Kriging and Block Kriging are the procedures for estimation in those two instances.

The Kriging alogrithm thus consists of solving the "Kriging System" (eqns. 7.51, 7.59, 7.60 and 7.61), i.e. a set of n equations in n unknowns which are represented in matrix form, and finding the weights of the contribution from neighbouring points to the site being estimated. The weights depend on the structures of the semi-variogram and on the geometry of the sampled points in the field. So, to run the KRIGING facility in OXSIS, it is essential to have computed the semi-variogram for the data with the use of the SEMI-VARIOGRAM facility.

The kriging facility interpolates at a point or block centred on given co-ordinates. The points or blocks may lie on a rectangular grid of spacing chosen by the user.

This facility builds too in an algorithm by Webster (1985) which was made available to OXSIS in the form of software. The original algorithm was also modified and adapted to operate on OXSIS data. In addition, the original kriging program has also been expanded to allow for interpolation at specific "target" points by pairs of co-ordinates specified by the user. The subroutine returns the interpolated value and its estimation variance.

Each of the points or rectangular blocks on the grid or at target points are thus interpolated and the estimation variances provided by the algorithm. This output may serve as input for isarithmic mapping, i.e. contouring, or 3-D block plotting using the graphics interface off-line.

The flow-chart showing the struture of the software for the "KRIGING" facility is provided in appendix 8. Several screens

illustrate the questionnaires and the interactive dialogue for this facility with an example. These are included also in appendix 8.

Contour maps and 3-D plots can be prepared with the output from "KRIGING". A secondary output file contains a listing of co-ordinates of the points, the Kriged estimates of the property of interest and their estimation variances. This information is in a format suitable for reading in the SAS/Graph package. The contour map and the 3-D plots have to be developed interactively using SAS commands directly on the keyboard of a high-resolution graphics terminal.

The facilities so far described and illustrated in this chapter form the present state of the OXSIS-manipulation package. Once complete, they provided the data processing facilities used for all the calculations in this thesis. The manipulation package in its present state is fully operational and provides the necessary tools to equip the soil surveyor to make predictions of the conditions of soil properties at unrecorded depths at visited or unvisited sites, and even for unrecorded properties which can be estimated from their surrogates.

8.6 Concluding remarks to this chapter.

i. The design and development of applications programs in a manipulation procedure of a SIS must allow for expansion by the inclusion of new capabilities. This has been demonstrated with the development of the Manipulation package in OXSIS.

ii. The operation of this "user friendly" Manipulation package has been demonstrated with examples, and its structure explained.

iii. OXSIS is now a tool which equips the user/soil-surveyor with the capabilities for predicting values of soil properties at any site and at any depth within a surveyed area.

iv. With the help of this tool, the problem of missing soil information necessary for land-use planning can be solved.

IX. TRIALS IN SPATIAL PREDICTION OF SOIL PROPERTIES.

9.1 Introduction.

Land-use planning requires information about soils, (chapter 2). Soil variability, (chapter 3), needs to be studied and portrayed to convey information. The role of the soil surveyor is to provide such information reliably and efficiently. Conventional soil survey is the standard procedure for the collection, analysis and representation of soil information, (chapter 4). Soil information is transmited in the form of soil classes, mapping units, a legend and an accompanying memoir.

Computerized soil information systems, (chapter 5), are an improved alternative to conventional soil survey to assist the soil surveyor in accomplishing his tasks. Mathematical interpolation techniques, (chapter 7), specially semi-variograms and optimal interpolation by Kriging can now be incorporated into information systems, (chapter 8), to equip the soil surveyor with a powerful tool for analysis of the spatial structures of soil variability, and for prediction.

However, various aspects concerning the use of these techniques in ordinary survey practice demand investigation.

The solution to the problem of missing soil information, (chapter 6), requires the study of the structures of spatial variability in the area of interest at different scales.

It has been seen in chapter 3 that, in general, three scales of structure in soil variability can be recognised.

The short range structures contain narrow ranges of soil properties and are usually indicated by significant levels of auto-correlation as described by the semi-variogram. The medium range structures are the structures that remain after the "noise" and perturbations of soil properties have been "filtered" out and they loosely correspond to soil classes. Ideally, The limits to such classes are established by procedures, (see chapter 4), that maximize the ratio of

between-class/within-class variances. Long range structures can be intuitively perceived as assemblages of soil classes characteristically associated with patterns of topography or any other land attribute forming landscape patterns.

On the other hand, it has been stated that kriging is the best technique for spatial interpolation but it is not a completely automatic technique. It depends on the existence of spatial auto-correlation and requires that stationarity assumptions be sufficiently obeyed. Its range of applicability is therefore limited mainly by the density of point-data (scale), the presence of trends, and by accidents in the landscape that disrupt variability.

Factors that determine the applicability of kriging are then the density of point-data, the size of the area sampled and the structures in the variability of the soil property that are embraced by the size of the area. Scale is reflected in the way sampling density and area size combine.

The soil surveyor is interested in the precision of the estimations that can be made. Kriging requires relatively high densities of sampling points to maintain the precision above acceptable limits. These sampling densities may not be affordable in ordinary survey practice.

For any situation of spatial prediction there must necessarily be some point-samples. If the points are close enough so that the density is high, and provided the property can be shown to be spatially dependent, they can be used for local estimation by kriging, i.e. exploiting the short range structures in soil data. The nugget variance in this case will indicate the extent to which noise should or can be reduced by substituting area samples for point samples, and whether or not mathematical interpolation of values may provide a better predictor than conventional soil survey output.

If the points are sparse or the property is not spatially dependent the estimation will have to be based on the mean values of soil classes or mapping units, i.e. exploiting the medium range structures in soil data. The level of precision in this case depends on the narrowness of the classes and on the purity of the mapping units.

For moderate densities of point-data, prediction will rest on a combination of interpolation and intuition, with a significant contribution from the latter.

Moreover, it has been seen that kriging is a suitable interpolator when data are stationary, but soil data shows a tendency to be naturally non-stationary (Burrough, 1983a). The presence of trends in soil variability is commonplace in the field, (chapter 3). They assocciate with medium-range structures of variability. The necessary theory and tools to permit interpolation in the presence of trends are available, but it is not clear when trends in the data must be considered sufficiently strong to merit forms of interpolation other than ordinary kriging.

9.2 The problems

Given an area of interest in which the three structures of variability overlap, and a sampling effort limited by survey cost:

(1). The main problem consists of finding out whether mathematical interpolation of values by kriging provides a better predictor of soil properties than means from classes or mapping units, and in what situations of variability, landscape and sampling density this is most likely to occur.

In a practical context, the problem consists of finding out how to decide in advance and as quickly and economically as possible, for a given area and its variability and a density of point-samples, which method or combination of these two methods a soil surveyor can use to make predictions of values of soil properties at unvisited sites as a solution to the problem of missing information. The choice depends on the predictive success of each technique under different situations of landscape and density of point-data, and is related to the spatial structures of variability of the soil property as depicted by a reconnaissance semi-variogram.

So, the predictive methods need to be compared under different situations of variability and sampling effort and ascertain their predictive value in each set of circumstances.

The resulting knowledge can be particularly useful in soil survey design.

The structures of soil variability that may be revealed by a semi-variogram, while tackling problem (1), lead to the study of other three equally important problems.

Problem (2). There is evidence to suggest that ordinary kriging is fairly robust to the effects of ignoring the presence of trends (see chapter 7). Nevertheless, given their importance as an inherent component of soil variability, the effects that ignoring these trends or accounting for them may have on the predictive success of the procedures to be compared needs to be explored too.

Problem (3). Alternative procedures for removing or accounting for trends, i.e. stratification by means of map units, modelling by local polynomial surfaces, or any other quick and cheap means, and their effect on the predictive success of kriging also merits investigation.

Problem (4). The utility of the semi-variogram as a reconnaissance tool to help deciding whether to produce a map by conventional procedures and use it for predictions or whether to base the predictions on kriging in the presence of accounted or unaccounted-for trends, needs also to be evaluated.

In order to study the main problem and its subsidiaries, trials on the spatial prediction of soil properties at unvisited sites were carried out with data from an array of nested sampling grids at different spacings in the area of Yelford near Oxford in west Oxfordshire. The aim of these trials, the area studied and the approach followed are discussed next.

9.3. Aim.

This study aims to compare the relative merits of the conventional and the new procedures for predicting missing soil information.

The predictive value of soil classes and soil mapping units is compared against that of spatial interpolation by kriging under different conditions of sampling densities, scales, and structures in the variability of soil properties.

The use of a Soil Information System, (OXSIS), and its capabilities for predicting missing soil information is also demonstrated by its use as the tool for the comparison.

As a by-product of the comparison, this study attempts to elucidate the role of the semi-variogram as a reconnaissance tool to help make decisions regarding predictions, and to determine the maximum sampling spacing at which, in this area, interpolation by the kriging technique is conveniently possible. It also attempts to assess the contribution of suitable combinations of intuition modelling and interpolation to tackle a range of problems that arise in spatial prediction and that make kriging either not applicable or not possible, e.g. notably, non-stationarity caused by trends. Such problems are intrinsic to the nature of the structures of soil variability in the field. So, complementary ways of combining mathematical techniques with soil classes and maps are explored to define practical strategies for generating missing soil information.

9.4 The area for the trials.

The trial area, (about 400 ha), is located approximately 2.5 miles, (4 km) south of Witney, and 12km west of Oxford, (map sheet SP30, Jarvis and Hazelden, 1982; see figure 9.1). It lies across the contact between the clay hills of the upper Jurassic, (Oxford Clay), and terrace remnants of the Upper Thames Valley, the Thames alluvium and low terraces, (Pleistocene and Recent), of the limits of the Thames floodplain. The area is dominated by two hills, Claywell and Rickless hills, and crowned by the calcareous sands and limestone gravels of the 100 ft. terrace (300-310 ft) above the river Thames (200-210 ft).

The floodplain is punctuated by well-drained areas of low (10 ft) gravel terraces. It contains deposits of two ages. An older alluvium which consists of clayey or loamy material over coarse loams, sands and gravels, and lies 3-5 ft above a younger alluvium. The younger alluvium lies outside the trial area in this study and consists of clayey alluvium, (see figure 9.1).

The whole area was extensively affected by peri-glacial processes. Windblown loess and fine cover-sand from

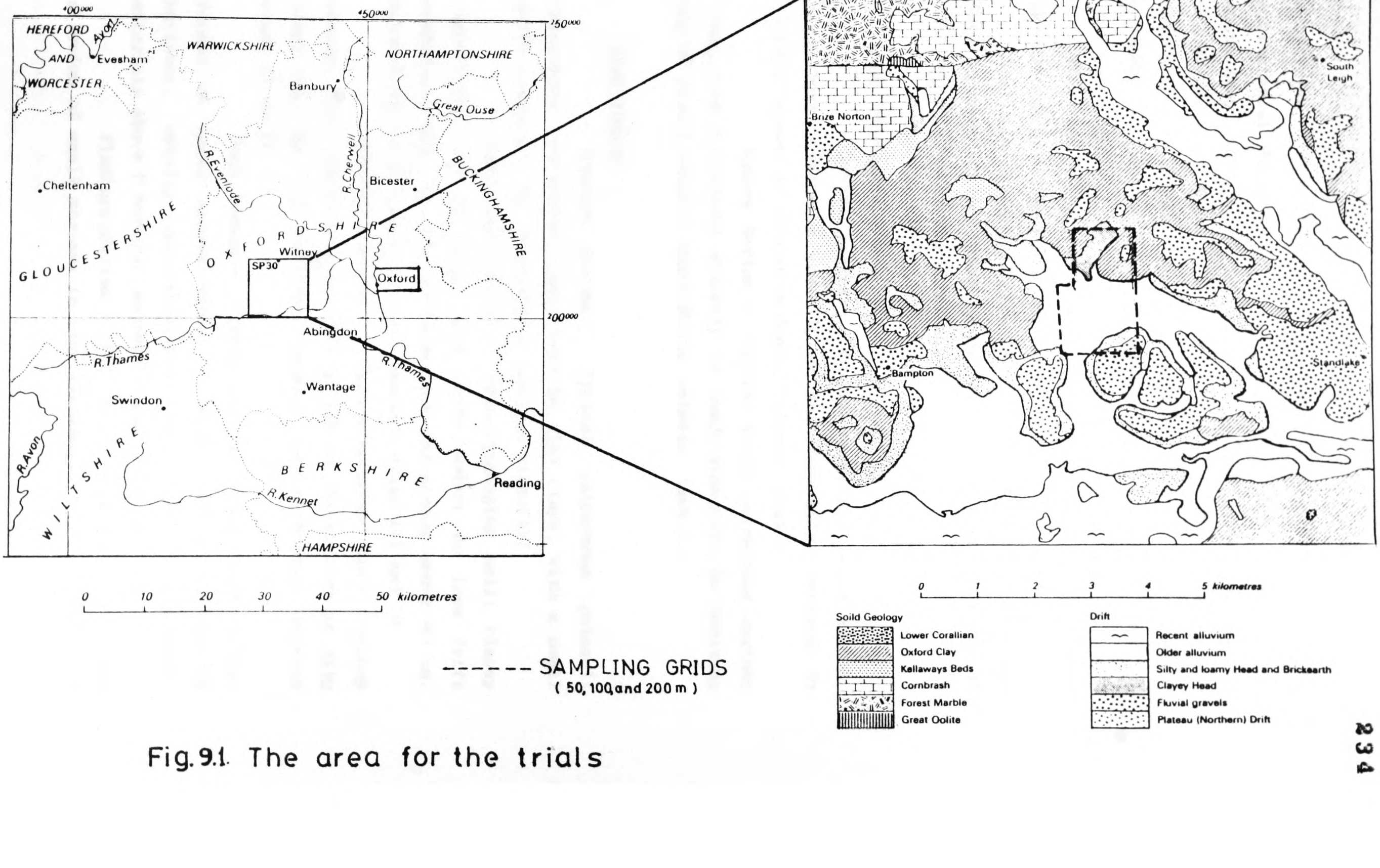


Fig. 9.1. The area for the trials

peri-glacial floodplains locally form deposits as much as 1 metre deep, and has increased the silty and loamy component in the topsoil of most parts of the landscape of both hills and floodplains. Freeze-thaw solifluction has led to downslope movement and to the extensive mixing of materials particularly on lower slopes.

The land is predominantly used for arable crops with some pastures in small portions, particularly on the hills.

The main characteristics of the soils from the hills and high terrace to the flood plain can be described in terms of soil series (Jarvis and Hazelden, 1982). Thus:

High Terrace:

Sutton Series.- Typical argillic brown earths, loamy over calcareous gravelly or sandy-gravelly; textural Bt horizon; gravel at 50-100 cm depth; wetness class I.

Badsey Series.- Typical brown calcareous earths; loamy over calcareous gravelly or sandy-gravelly; Bw horizon; top of gravel usually above 60 cm; wetness class I.

Clay slopes:

Evesham Series.- Typical calcareous pelosol; calcareous; calcareous clayey over *in situ* clays, with a small drift component; Bg, Bc horizons; wetness class III.

Denchworth Series.- Pelo-stagnogley soil; clayey (smectitic) over *in situ* clays, with modest or less drift component; Ahg, Bg, BCg horizons; non-calcareous above 40 cm, subsoil may be calcareous or not; wetness class III and IV.

Lawford Series.- Pelo-stagnogley soil; stoney clayey, over clayey stoney drift 50-150 cm thick, over *in situ* clay; Ahg, Bg, BCg horizons; usually non-calcareous; wetness class II or IV.

Hook Series.- Gleyic argillic brown earth; silty, formed on local accumulations of loess; Bt or Btg and Cg horizons; usually non-calcareous over loamy or gravelly materials above C horizon; wetness class II or III.

Fladbury Series.- Pelo-alluvial gley soil; clayey alluvium of small streams in clay landscape, over gravels or *in*

situ clay; Ahg, Bg, BCg horizons; noncalcareous except at depth; wetness class IV.

Low Terraces:

Badsey Series.- (as above).

Bampton Series.- Gleyic brown calcareous earth; silty or loamy formed from loess or fine alluvium over terrace gravels; Bg and/or Cg horizons; slightly calcareous, becoming more calcareous at depth; wetness class II.

Flood plain (older alluvium):

Kelmscott Series.- Calcaro-cambic gley soil; loamy or silty over calcareous gravelly or sandy-gravelly deposits; Apg, Bw(g) or Bg, Cg horizons; more or less calcareous throughout; wetness class III or IV.

Caswell Series.- Pelo-cambic gley soil; clayey river deposits over coarse loamy or coarser (terrace deposits) from 60 cm or deeper;Bg, BCg horizons; non-calcareous above the coarse subsoil material; wetness class III.

The spatial distribution of these soil series and their relation to the landscpe is shown in terms of mapping units in figure 9.2, (from Jarvis and Hazelden, 1982).

9.5 Hierarchy of existent soil data in the area.

The trial draws upon data from several stages of an earlier study that covered the trial area (the "Yelford area").

In the first stage Ir. B. Marsman from STIBOKA (Wageningen, The Netherlands), mapped the whole of the SP30 sheet at 1:1/4 M. He then remapped the Yelford area (520 ha) in two stages:

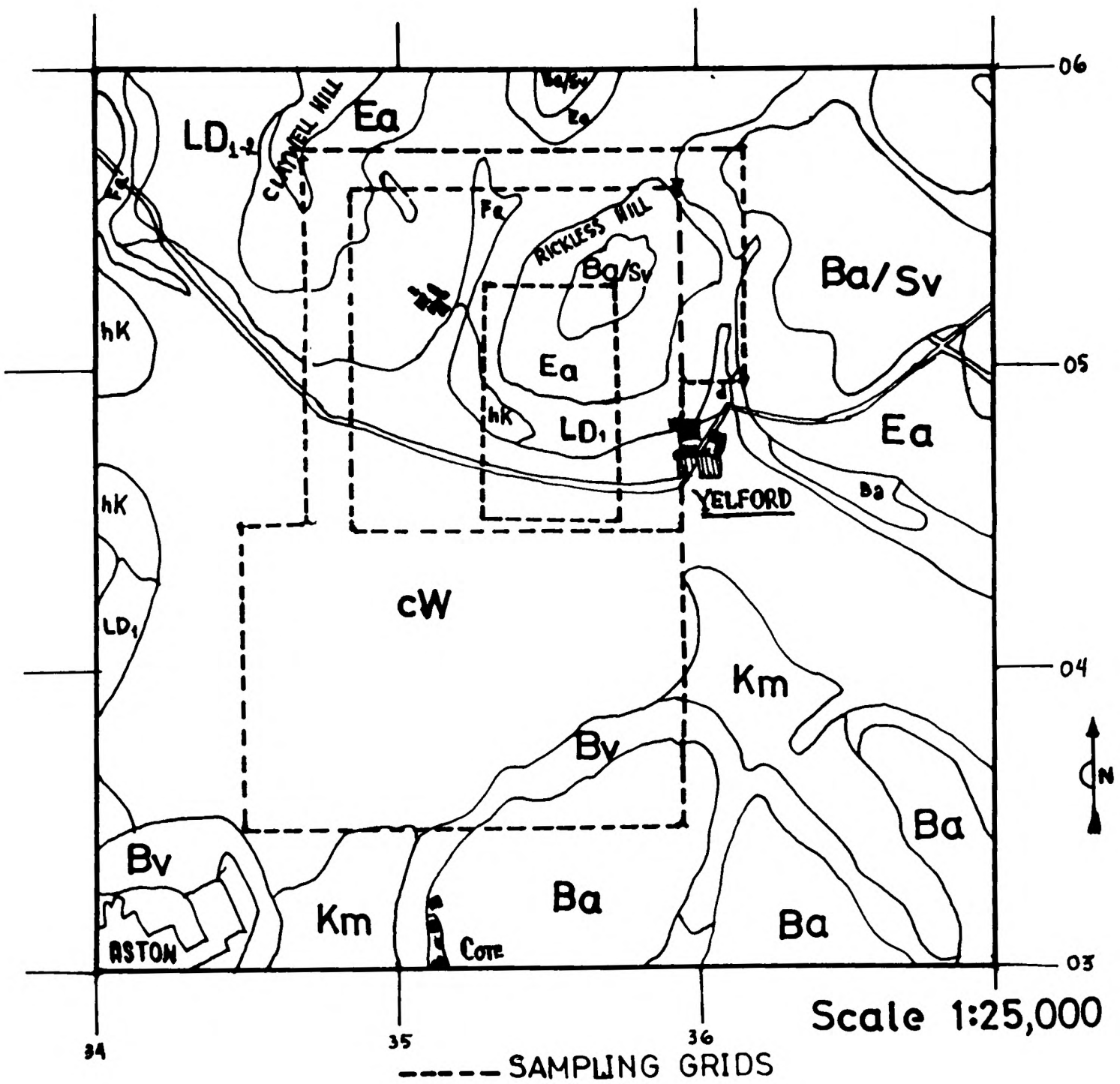
(i). Its central core (97 ha) at 1:10,000.

(ii). The remainder (423 ha) at 1:35,000, to which was added the core map (i), now generalized to the scale of map (ii).

The profile classes shown in table 9.1 were distinguished during these surveys.

Some of them were divided on their texture, e.g. NGG1 is a clay loam, or a clay loam over clay, and NGG2 is a clay, or clay

Fig. 9.2 Soil-series mapping units in the area of study as defined by Jarvis and Hazelden (1982)



cW (CARSWELL) Pelo-cambic gley soils

Ba (BADSEY) Typical brown calcareous earths

Bv (BAMPTON) Gleyic brown calcareous earths

Ea (EVESHAM) Typical calcareous pelosols

LD (LAWFORD) Pelo-stagnogley soils

Km (KEMSCOT) Calcaro-cambic gley soils

Sv (SUTTON) Typical argillic brown earths

Fa (FLADBURY) Pelo-alluvial gley soils

hK (HOOK) Gleyic-argillic brown earths

Table 9.1 Profile classes distinguished in the Yelford area from Marsman's surveys

Soil	Label	Soil Class
Freely drained soils	1	
1 + with a weathered B horizon	3	
Poorly drained gley soils	2	
2 + without a weathered B horizon	4	
3 + with gleying in the profile	5	
3 + without gleying in the profile	6	
4 + with characteristics of surface water gleys	7	
4 + with characteristics of ground water gley soils	8	
5 + calcareous		BCg
5 + non-calcareous		BEg
6 + calcareous		BC
6 + non-calcareous		BE
7 + calcareous		CSG
7 + non-calcareous		NSG
8 + calcareous		CGG
8 + non-calcareous		NGG

over clay loam, and on their drainage class (1-5, in which 5 is the best drained). A further drainage symbol (g) was inserted between g and freely drained.

Then, Marsman defined mapping units at two scales. These are shown in table 9.2. The fractions, e.g. 4/5, indicate that the units contain profiles of two drainage classes.

K. Rowe (Soil Conservation Authority, Victoria), took the project a further stage. Firstly, he produced a key by which to identify Marsman's profile classes in the field. Then he established three sampling grids with grid spacings of 50m, 100m, and 200m (from which other two grids of spacing 283m could be derived, but they are not included in these trials). The grids cover progressively larger areas of 38 ha (100 sites), 111 ha (140 sites), and 372 ha (105 sites). Figure 9.3 shows the way these grids were laid on the landscape of the trial area and their relation to soil mapping units as defined by Marsman.

Data from a fourth grid of 1 km spacings covering the whole area corresponding to the SP30 sheet (Jarvis and Hazelden, 1982), hence embracing the previous smaller grids, were also available from OXSIS data-bank. Data from this grid, however, were not considered directly for these trials in prediction and were used only to elucidate nested structures of variation from semi-variograms.

The location of each sampling site in the four grids, in terms of grid reference eastings and northings, is shown in the computer plots of figures 9.4a to d.

Two important features of the location of these grids on the landscape must be noticed. The first is that, in the east-west direction, they lay so that smaller grids are increasingly concentrated on one of the hills, (Rickless hill), and cover a part of the east slope of the other (Claywell hill). Thus they go from the calcareous sands and limestone gravels of the high terraces on the hill top, down to the clayey or loamy material of the slopes and the older alluvium and up again to the clay slopes. The north-south and northeast-southwest directions of the grids follow the direction of the drift. This is a very important feature and must be remembered while interpreting semi-variograms calculated from these data, as the presence of trends in the same direction can be expected and anticipated.

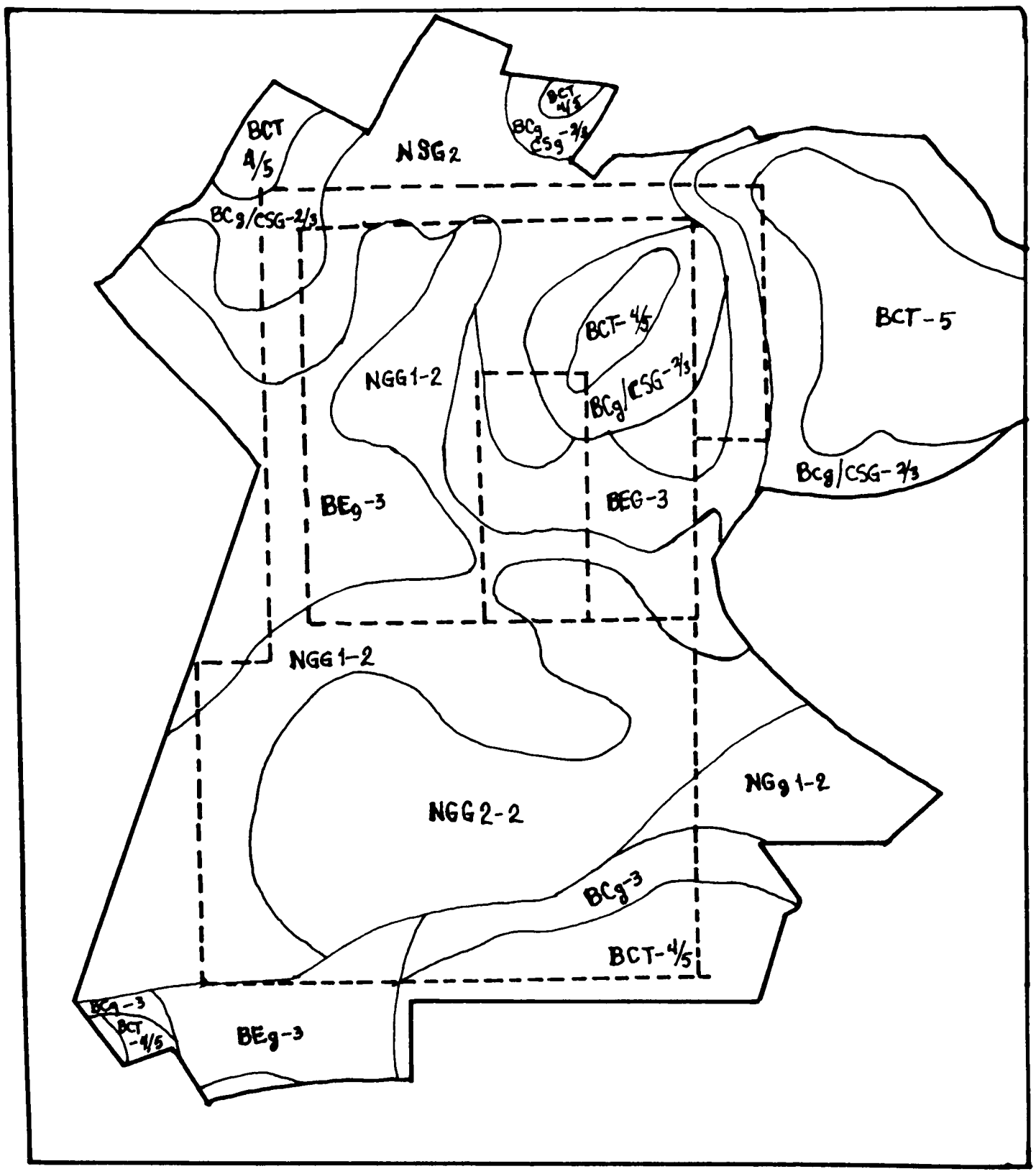
Table 9.2 Mapping units defined from Marsman's surveys at two scales

Mapping Unit	Scale 1 : 35,000	Scale 1 : 10,000
Brown calcareous soils (BC)	BC - 5	BC - 5
Brown calcareous soils with gleys	BCg - 3	BC(g) + - 3/4 BCg - 2/3
Brown earths with gleying	BEg - 3	BEg - 3
Calcareous surface water gleys		CSG - 2
Non-calcareous surface water gleys	NSG - 2	NSG - 2
Non-calcareous ground water gleys	// NGG1 - 2 NGG2 - 2	NGG1 - 2 NGG2 - 2
Soil Complex	BCg/CSG - 2/3 (40 : 60)	

// - Division on their texture (e.g. NGG1 = clay loam or clay loam over clay; NGG2 - clay or clay over clay loan)

* - X/Y Indicates that units contain profiles of two drainage classes.

\+ - Drainage symbol inserted between g and freely drained.



----- SAMPLING GRIDS

——— MAPPING UNITS BOUNDARIES

APROX. SCALE 1:20 681

BCg Brown calcareous soils with gleying

BCT Brown calcareous soils on terrace gravel

BEg Brown earths with gleying

BEG Brown earths gley soils

CSG Calcareous surface water gley soils

CSg Calcareous surface water with gleying

NGG Non-calcareous ground water gley soils

NGg Non-calcareous ground water with gleying

NSG Non-calcareous surface water gley soils

Sub-division according to one of the following aspects: texture, soil depth, drainage class (1-5), ripening, parent material or consolidated rock at 1m. Depth

Fig. 9.3 Position of the 3 sampling grids and their relation to soil mapping units as defined by Marsman at Yelford.

Fig. 9.4a

YELFORD 50M GRID

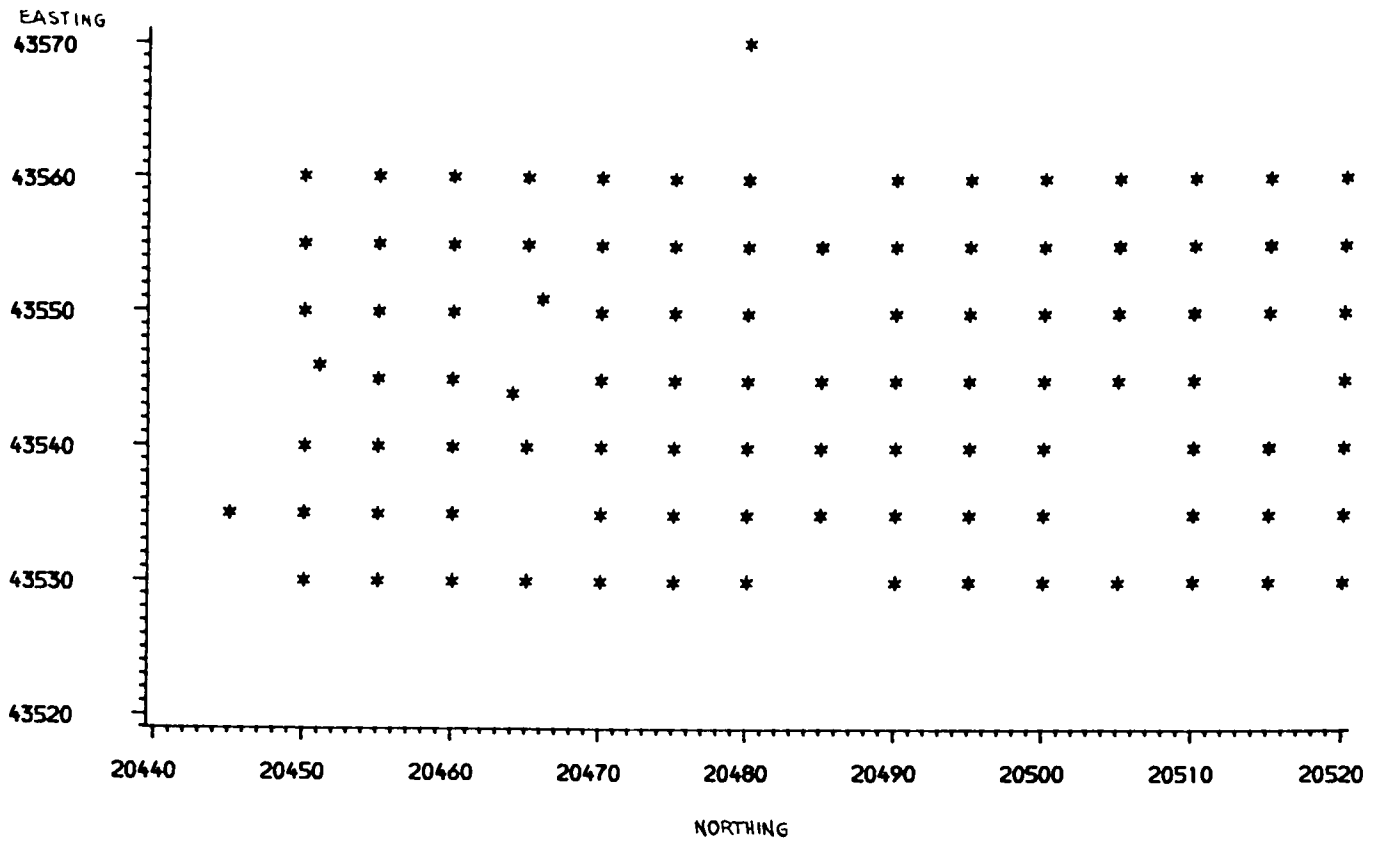
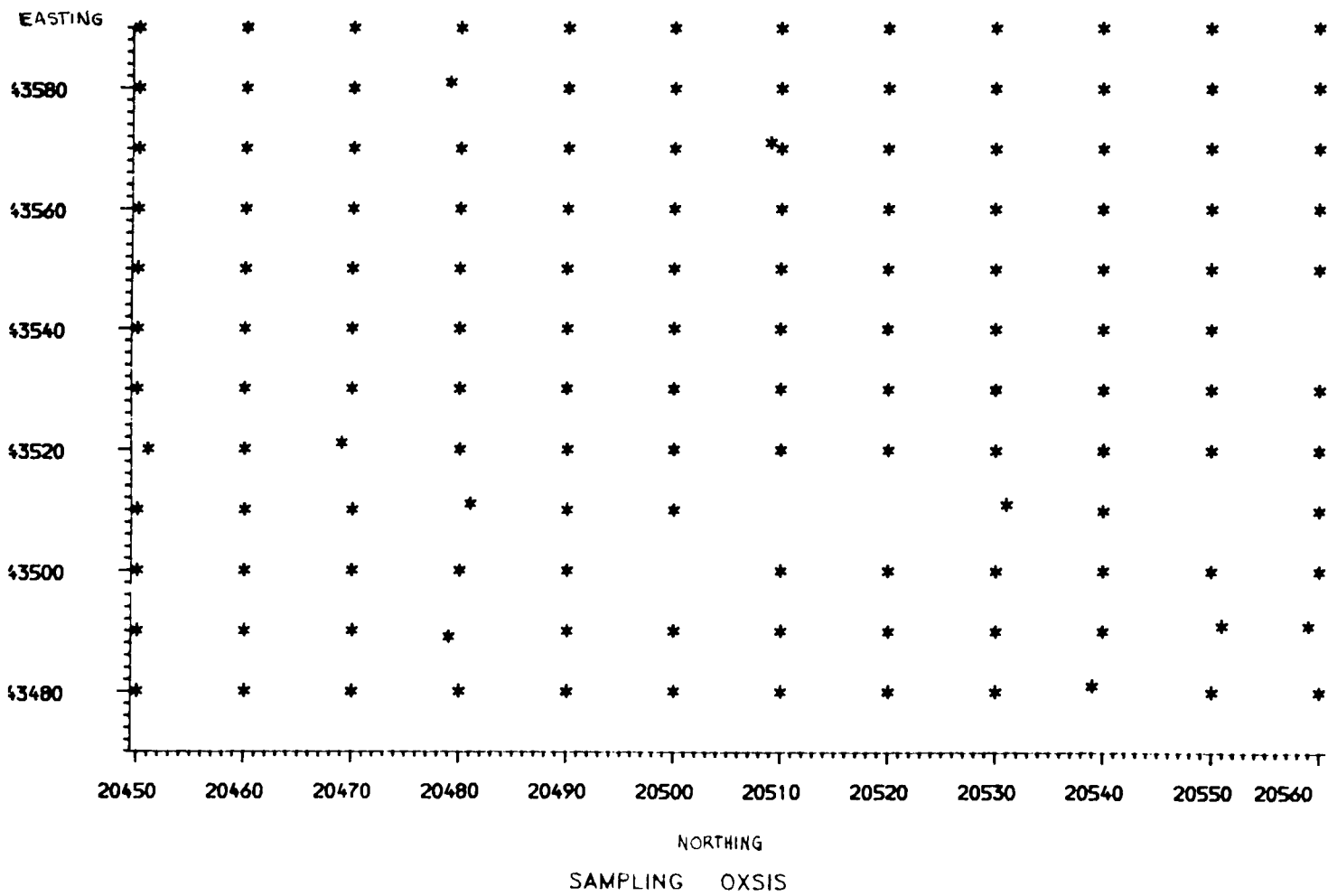


Fig 9.4 b

YELFORD 100M GRID



SAMPLING OXSIS

Fig 9.4c *YELFORD 200M. GRID*

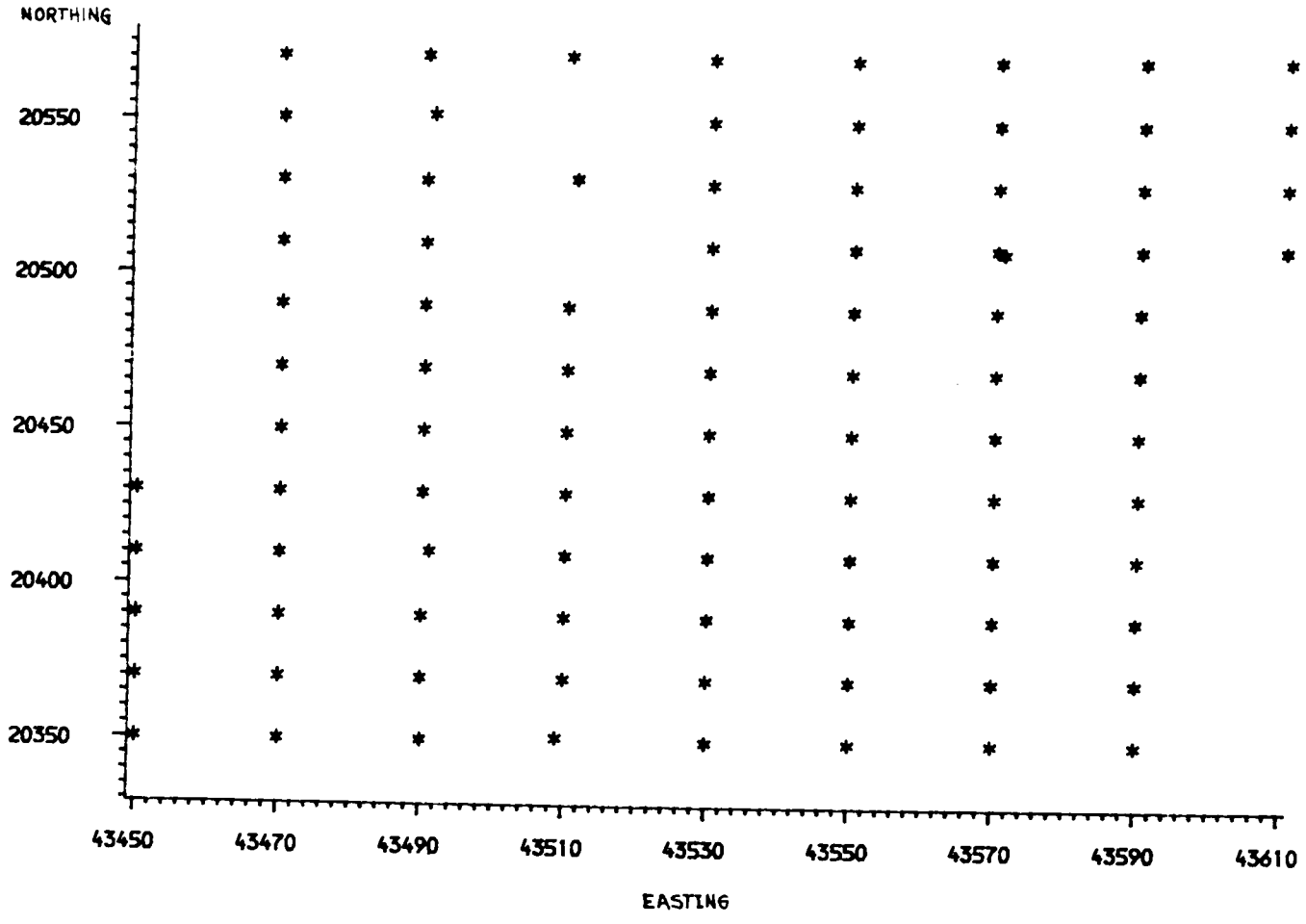
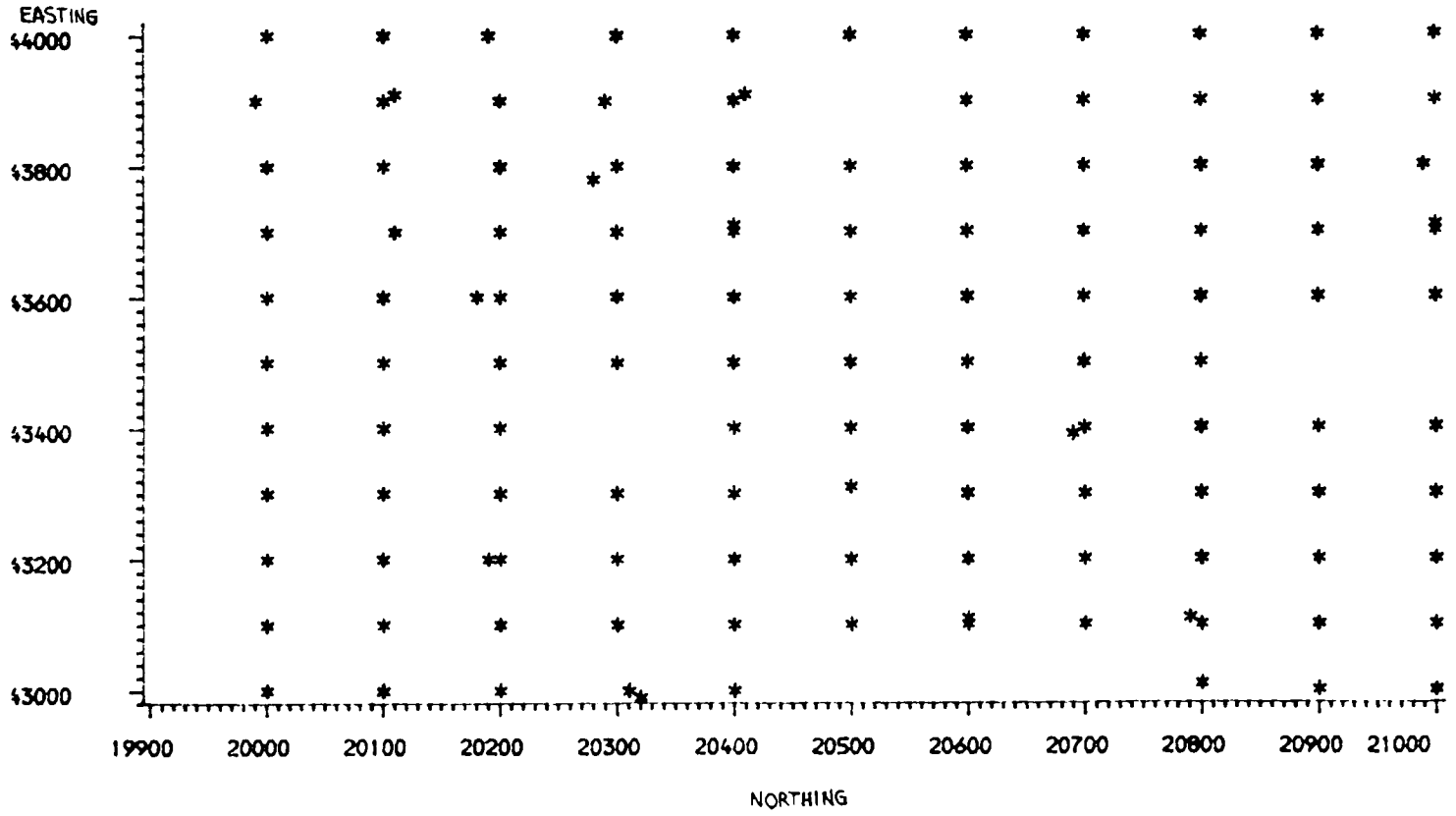


Fig. 9.4d

SP30 SHEET 1KM. GRID



SAMPLING SCHEME OXSIS

So the drift runs from the hill-top on the high terrace to the clay slopes and down to the older alluvium.

The second feature worth noticing is that as the grid spacing becomes shorter, the representativeness of the variability of the soil-landscape provided by these grids decreases too. Thus the smallest grid only samples one part of the area sampled by the largest grid. For instance, the 50m grid lies mostly on the clay slopes (Oxford Clay), and part of the top of Rickless hill (gravels), on the high terrace, while only a small portion of it covers the older alluvium. However, due to the variable landscape it covers, rapid changes in the values of soil properties may be expected from the data-set provided by this grid. In contrast, the 200 m grid evenly covers the spectrum of variation given by the contact between the clay hills and the older alluvium.

The profile class was recorded by Rowe at every grid site, an abreviated profile description was made, and samples were taken for analysis. He also established a set of "random" check sites within each sample grid. Their soils were identified and sampled. Advantage was taken of the availability of these check sites in this thesis. They were used as "target" points to test the validity of predictions of values of soil properties by various procedures, and to compare them. The list of the "random" check points corresponding to each sampling grid, together with their co-ordinates and profile classes are given in tables 9.3a to c. Their approximate location within the three grids is shown in figure 9.4e.

Thus, from his observations Rowe produced families of soil maps based on grid or free survey for the three sample areas using legends broadly similar to those of Marsman. These are listed in table 9.4, where 4/5, etc. indicates a state intermediate between drainage classes 4 and 5, while 4(5) indicates a mixture of 4 and 5, with 4 the more frequent. With these maps and records from random sites, Rowe was able to calculate the purities of the units mapped by various procedures and with various degrees of effort.

In due course the soil samples were analyzed for a range of chemical and pedological properties.

Table 9.3a - Random Check Sites within the 50m Sampling Grid

Site	Northing	Easting	Site Number	Soil Class
A4	20511	43547	95004	BCg-3
B13	20513	43542	95024	BEg-3/2
C1	20519	43560	95027	BCT-4/5
C1	20502	43538	95028	NSG-2
C3	20500	43542	95029	BEg-3/2
C5	20494	43548	95030	NSG2/BEg-2
C6	20491	43538	95031	BEg-3
C9	20487	43558	95032	BEg-3
C11	20468	43557	95033	CGG-2/Bcg-2
C12	20468	43561	95034	NGG1-2
C13	20462	43544	95035	NGG1-2
C14	20461	43547	95036	NGG1-2
C16	20458	43560	95037	NGG1-2
E10	20454	43544	95047	NGG1-2
E14	20511	43560	95051	BCT-4/5
E15	20503	43545	95052	BCg-2

Table 9.3b - Random Check Sites within the 100m Sampling Grid

Site	Northing	Easting	Site Number	Soil Class
A1	20563	43584	95001	BEg-2
A2	20540	43503	95002	BEg-2
A3	20533	43571	95003	BCT-4/5
A4	20511	43547	95004	BCg-3
A5	20502	43512	95005	NGG1-2
A6	20466	43572	95006	NGG1-2
B1	20544	43572	95012	5Eg-3
B2	20531	43504	95013	BEg-3
B3	20521	43495	95014	BEg-(4)
B4	20532	43544	95015	BEg-3
B5	20532	43590	95016	BC(g)-3/4
B6	20524	43539	95017	BCg-3
B7	20508	43491	95018	BEg-3/4
B8	20500	43579	95019	NSG-2
B9	20491	43582	95020	NSG-2
B10	20461	43509	95021	BEg-3/2
B12	20455	43576	95023	CGG-2
B13	20513	43542	95024	BEg-3/4
C1	20519	43560	95027	BCT-4/5
C2	20502	43538	95028	NSG-2
C3	20500	43542	95029	BEg-3/2
C5	20494	43548	95030	NSG-2
C6	20491	43538	95031	BEg-3
C9	20487	43558	95032	BEg-3
C11	20468	43557	95033	CGG-2/BCg-2
C12	20468	43561	95034	NGG-2
C13	20462	43544	95035	NGG1-2
C14	20461	43547	95036	NGG1-2
C16	20458	43560	95037	NGG1-2
E4	20513	43575	95041	BCg-2
E5	20453	43572	95042	CGG-2
E6	20560	43552	95043	NSG-2
E7	20450	43588	95044	BCg-3/2
E8	20528	43524	95045	NGG1-2
E9	20553	43541	95046	NSG-2
E10	20454	43544	95047	NGG1-2
E11	20515	43567	95048	BCg-3
E12	20500	43563	95049	BC(g)-3/4
E14	20511	43560	95051	BCT-4/5
E16	20503	43545	95052	BCg-2

Table 9.3c - Random Check Sites within the 200m Sampling Grid

Site	Northing	Easting	Site Number	Soil Class
A1	20563	43584	95001	BEg-2
A2	20540	43503	95002	BEg-2
A3	20533	43571	95003	BCT-4/5
A4	20511	43547	95004	BCg-3
A5	20502	43512	95005	NGG1-2
A6	20466	43572	95006	CGG-2
A7	20446	43558	95007	NGG2-2
A8	20415	43481	95008	NGG1-2
A9	20397	43501	95009	NGG1-2
A10	20382	43572	95010	BCg-3/2
A12	20372	43563	95011	BCg-3
B1	20544	43572	95012	BEg-3
B2	20531	43504	95013	BEg-3
B3	20521	43495	95014	BEg-(4)
B4	20532	43544	95015	BEg-3
B5	20532	43590	95016	BC(g)-3/4
B6	20524	43539	95017	BCg-3
B7	20508	43491	95018	BEg-3/4
B8	20500	43579	95019	NSG-2
B9	20491	43582	95020	NSG-2
B10	20461	43509	95021	BEg-3/2
B11	20475	43572	95022	BEg-3/2
B12	20455	43576	95023	CGG-2
B13	20513	43542	95024	BEg-3/4
B14	20501	43559	95025	BCg-3
B15	20481	43538	95026	BEg-2/3
C1	20519	43560	95027	BCT-4/5
C2	20502	43538	95028	NSG-2
C3	20500	43542	95029	BEg-3/2
C5	20494	43548	95030	NSG-2
C6	20491	43538	95031	BEg-3
C9	20487	43558	95032	BEg-3
C11	20468	43557	95033	CGG-2/BCg-2
C12	20468	43561	95034	NGG-2
C13	20462	43544	95035	NGG1-2
C14	20461	43547	95036	NGG1-2
C16	20458	43560	95037	NGG1-2
E1	20418	43468	95038	NGG1-2
E2	20567	43605	95039	CSG-2
E3	20567	43578	95040	NSG-2
E4	20513	43575	95041	BCg-2
E5	20453	43572	95042	CGG-2
E6	20560	43552	95043	NSG-2
E7	20450	43588	95044	BCg-3/2
E8	20528	43524	95045	NGG1-2
E9	20553	43541	95046	NSG-2
E10	20454	43544	95047	NGG1-2
E11	20515	43567	95048	BCg-3
E12	20500	43563	95049	BC(g)-3/4
E13	20515	43561	95050	BCT-4/5
E14	20511	43560	95051	BCT-4/5
E15	20503	43545	95052	BCg-2

Fig. 9.4e Location of the random check points in the three sampling grids at Velford.

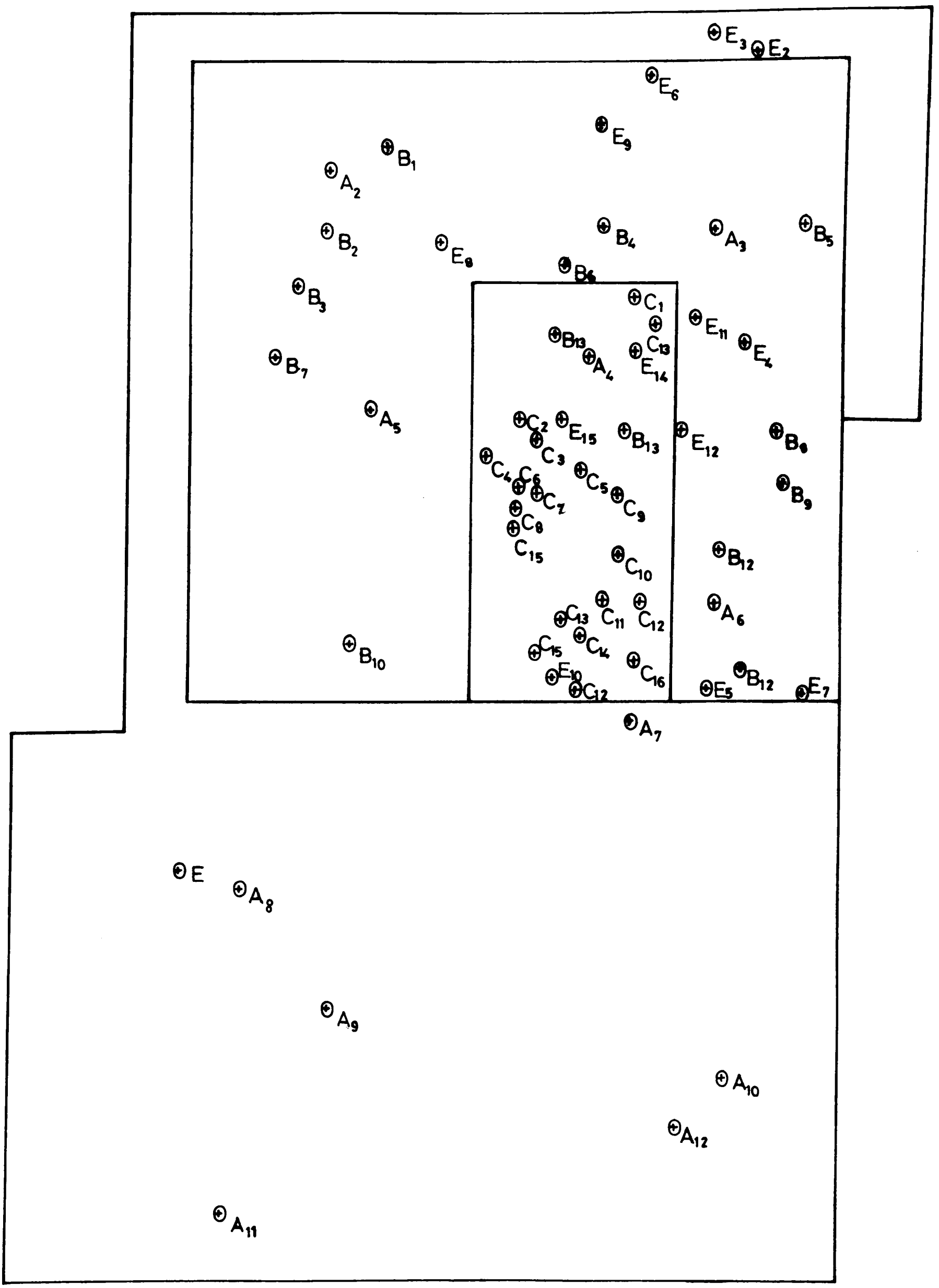


Table 9.4 Soil mapping units defined from Rowe's surveys and found within the three sampling grids in the trial area, Yelford

GRID		
50m. Grid	100m. Grid	200m Grid
BCT - 4/5*	BCT - 4/5	BCT - 4/5
BC(g) - 3/4	BC(g) - 3/4	
BCg - 3(2)@	BCg - 3(2)	BCg - 3(2) BCg - 3
BEg - 3(2)	BEg - 3(2)	BEg - 2(3) BEg - 3(2)
NSG - 2	CGG - 2 NSG - 2	NSG - 2
NGG1 - 1	NGG1 - 2	NGG1 - 2
NGG2 - 2	NGG2 - 2	NGG2 - 2

* X/Y - Indicates an intermediate state between drainage classes X and Y

@ X(Y) - Indicates a mixture of X and Y with X the more frequent.

So, the prediction study started with 3 sets of data, one for each of the sampling grids, and another set corresponding to the check points. Each of these data-sets is stored in OXSIS in three files in the form of a data-triad (chapter 5). Listings of these triads including the triad of the check points are included in appendix 9.1.

The concepts of Marsman's profile classes and Rowe's map units were in a few cases widened to give the modified classification in table 9.5. The main characteristics of the soils in these classes are given in table 9.6.

9.6 The soil variables.

A total of 8 analysed properties were selected from the sets to serve as the variables to be predicted. The first four properties in the list below were chosen because they seem to be much affected by land use and management. The bottom four were chosen because they are expected to show some degree of spatial order not affected by management. Thus the soil variables selected were:

1. pH
2. Exchangeable Potassium (k)
3. Exchangeable Magnesium (Mg)
4. Carbonate, (total calcium and magnesium carbonates)
5. Clay content in percentage
6. Silt content in percentage
7. Sand (coarse) content in percentage
8. Organic Matter (quantity of humified matter in soil)

The three grids were sampled by shallow pit and auger. Only the analytical data from these samples was used in this study. Table 9.7 shows the meaning of each property and the method of determination.

Records for these properties are stored in OXSIS data-bank in separate data triads (see appendix 9.1). It is pertinent to state here that small pits and auger bores are the size of support from which the samples were taken.

Table 9.5 New soil classes from widening Marsman's profile classes and Rowe's map units concepts. The new classes are used for the calculation of class means of soil properties as predictors in the trial area, Yelford

Soil Class	Brief Description	Grouped Sub-Division	Class Identifier
BCT	Brown calcareous on terrace gravel	-4, -5, -4/5	1
BCg	Brown calcareous with gleying	-3, - 3/4	2
BCg	Brown calcareous with gleying	-3/2, -2, -2/3	3
BEg	Brown earths with gleying	-3, -4, -(3/4)	4
BEg	Brown earths with gleying	3/2, 2, 2/3	5
CSG	Calcareous surface water gley soils	2	6
NSG	Non-calcareous surface water gley soils	2	8
NGG1	Non-calcareous ground water gley soils	2, 2/3	8
NGG2	Non-calcareous ground water gley soils	2	9
CGG	Calcareous ground water gley soils	2	10

Table: 9.6. Main characteristics of Soil Classes found in the Yelford Area*

Map Unit	Description
BCT - 4 or 5	Weathered B horizon. Calcareous, shallow to moderately deep (100 cm), loam or loam over clay loam. Lighter texture mainly where gravel is shallow. Terrace gravel or drift/terrace.
BCg - 2 or 3 or 4	Weathered B horizon and gleying. Calcareous; moderately deep to deep. Topsoil of loam over clay loam to clay over clay at 50 cm and locally over loam or clay loam. Lighter textures and better drainage associated with greater depths of drift; roughly 4 on loam, 3 on clay loam, 2 on clay. Drift, drift/Oxford clay, thin drift on Oxford clay (possibly drift/alluvium). Drainage roughly follows topsoil texture, i.e. 3/4 on clay loam or loam/clay loam; 2/3 on clay to clay loam.
BEg - 2 or 3	Weathered B horizon and gleying. Non-calcareous, deep. Loam or clay loam or clay loam/clay over clay loam or clay loam/clay or fine loam, sandy loam, silt loam below 50 cm associated with alluvium. Drift over Oxford clay, alluvium (small occurrences of colluvium). Assume drainage follows topsoil texture.
CSG - 2, NSG2 - 2	Weathered B horizon and gleying. Surface water gley, calcareous (CSG) and non-calcareous (NSG). Clay or clay loam over clay over clay from 50 cm (except locally topsoil loam or clay loam and subsoil loam or sandy loam in SCG - 2). Oxford clay or drift over Oxford clay.
NSG 1 - 2	Non-calcareous surface water gley. Deep. Clay loam or clay loam/clay, locally with loam over washed, over loam or sandy loam at depth greater than 70 cm. Alluvium.
NGG 2 - 2	Non-calcareous ground water gley. Deep. Clay or clay/clay loam from 60 - 80 cm over loam or fine sandy loam at depth. Alluvium.

* From widening Marsmans concepts and Rowe's mapping units

Table 9.7 - Soil Properties involved in the trials in prediction in the Yelford Area

Soil Property	Description and Measurement
pH C01	Measured in M/100 calcium chlorian (soil/solution - 1/2.5) stirred and left to equilibrate overnight (Tomlinson et al 1977)
Potassium,-Exchange.A01	Potassium ions exchangeable to M.Amm. Acetate at pH 8.1 (Tomlinson et al 1977). in me/100 g. air-dry soil.
Magnesium,-Exchange.A01	Magnesium ions exchangeable to M.Amm. Acetate at pH 8.1 (Tomlinson et al 1977). in me/100 g. air-dry soil.
Carbonate A02	Total calcium and magnesium carbonates as proportion of soil, estimated as percent of $CaCO_3$ equivalent. Estimated in the laboratory on loss in weight on treatment with HCl (Bauer, Beckett, Bie (1972)). In percentage of air-dry fine earth.
Clay *02	Particles with equivalent spherical diameter less than 2 microns. Estimated in suspension of fine earth dispersed in Calgon by plummet method. (Bannister, Tomlinson,Beckett 1974). In percentage of air-dry fine earth.
Silt A02	Soil particles of equivalent spherical diameter 2 - 20 microns: International silt. Estimated in suspension of fine earth dispersed in Calgon, by plummet method (Bannister, Tomlinson, Beckett 1974). In percentage of air-dry earth.
Sand,-Coarse.A02	Soil particles of equivalent spherical diameter 200 - 2000 microns: International COARSE SAND. Estimated by a sieving suspension of fine earth in Calgon. In percentage air-dry fine earth.
Organic-Matter,-Abundance.A02	Organic matter determined in the laboratory by wet oxidation with dichromate. Multiplied by 1.33 for incompleted oxidation and multiplied by 1.73 to convert organic C to organic matter. (Walkley, 1947). in percentage of oven-dry fine earth.

9.7. Approach.

The general approach to this study uses semi-variograms and kriging capabilities built into OXSIS to exploit stored point-data. It demonstrates the use of an information system and the benefits of its on-line capabilities for prediction, and it also compares the predictive value of kriging with that of soil classes and mapping units.

The sets of point-data were provided by the three sampling grids. The generalized data was calculated as the mean values for the mapping units and classes described in the previous section. Extra sets of point-data, the "random check sites" that had been collected by Rowe, served as "targets" for checking the values of predictions. Both the point-data from the grids and from the check points were already stored in the OXSIS data-bank.

In brief, <u>the approach to these trials</u> consisted of the following sequence:

i. Computation of experimental semi-variograms.

ii. Fitting models to experimental semi-variograms.

iii. Semi-variogram interpretation.

iv. Fitting models for anisotropy.

vi. Predicting values of soil properties from soil survey output.

vii. Predicting by Kriging interpolation.

viii. Comparison of predictive procedures ignoring trends.

ix. The semi-variogram as a reconnaissance tool ignoring trends.

x. Removing trends.

a. Modelling the trends by fitting bicubic spline surfaces.
b. Stratifying trends by mapping units.

xi. Computing semi-variograms of de-trended data.

xii. Interpolation by Kriging of de-trended data.

xiii. Comparison of predictive procedures when trends are accounted for.

xiv. The semi-variogram as a reconnaissance tool when trends are accounted for.

The calculations and analysis in the approach listed above, were executed as described next. However, before any operations on data occured these had to be retrieved from the data-bank and prepared by extracting values at uniform depths.

9.8 Retrieval of data for the trials.

Initially the relevant data were identified through interrogating the "supporting files" of the system by means of the "Query" procedure and then retrieved using the OXSIS "Retrieve" procedure, (chapter 5). Records on the selected soil properties were present in the data sets of grids and the check points. They were extracted during retrieval and became the data in the "retrieved sub-set".

The retrieved sub-sets then entered the "Manipulation" stage. The "Depth", "Statistics", "Semi-variogram" and "Kriging" facilities of the manipulation procedure, (chapter 8), were used to uniformize depths, compute statistics, semi-variograms and the kriging estimates necessary for this study. Graphic interfaces were used for graph plotting.

9.9 Data manipulation to uniformize depth.

Before any calculation on data took place the values of the properties had to be extracted at uniform depths through the

"DEPTH" facility and in particular by the use of the "LAYER" algorithm. No spline depth-interpolation was required since the data of interest were recorded at two fixed depths (0-15 and 35-50 cm). The "LAYER" algorithm extracted values for the top soil (0-15 cm) for all the data sub-sets of interest and stored them in a manipulated sub-set ready for the computation of statistics and semi-variances.

9.10 Computation of experimental Semi-variograms.

Experimental semi-variograms of the 8 properties chosen were computed for each of the three sampling grids by the procedures discussed in 7.3.3.2, (eqs. 7.35 and 7.37). The algorithm of the "SEMI-VARIOGRAM" facility, (chapter 8), in OXSIS was used for this purpose.

In addition to the 24 experimental semi-variograms calculated, (8 properties for each of three sampling grids), and when it was found pertinent, 2 more semi-variograms for each property were necessary to elucidate the nested structures inherent in the overlapping sampling grids. One nested semi-variogram was calculated with data from the three grids together (50, 100, and 200m), and the other semi-variogram was calculated with data from the 1 km grid, which overlaps the smaller grids and contains some of the sites from these grids.

Each semi-variogram was calculated along 4 directions: east-west (0°), north-south (90°), north-east - south-west (45°), and north-west - south-east (135°). Rough line-printer plots were made for each direction for individual analysis, but a global plot of the semi-variances for all directions was produced on the graph plotter. When the global plot shown that a single semi-variogram model could well represent the shape of the semi-variogram in all directions, then a suitable model to fit the points was sought.

9.11 Fitting models to experimental semi-variograms and estimation of parameters.

The selection of a suitable model for each experimental semi-variogram was made by procedures described in 7.3.3.7, and

involving several iterations. First choices of models were guided by observation of the individual directional semi-variograms, the global plot, and the geometric relationships of safe models displayed in figure 7.3. These were helped by background information about the area. An algorithm that fits the model chosen onto the scatter of the experimental semi-variogram was used interactively, displaying each plot on the screen of the VDU graphics terminal, until a satisfactory fit was found. Sets of tentative model parameters were input in each iteration. So, the parameters were estimated simultaneously, when the best fit was found.

9.12 The Semi-variograms of the Yelford area.

Sets of semi-variograms of the 8 soil properties in the Yelford area resulted from the application of procedures 9.10 and 9.11 above. The models fitted and their parameters were ploted onto the points of the experimental semi-variograms. Since the variability of soil properties is not the same from one to the other, these are shown and interpreted by individual properties.

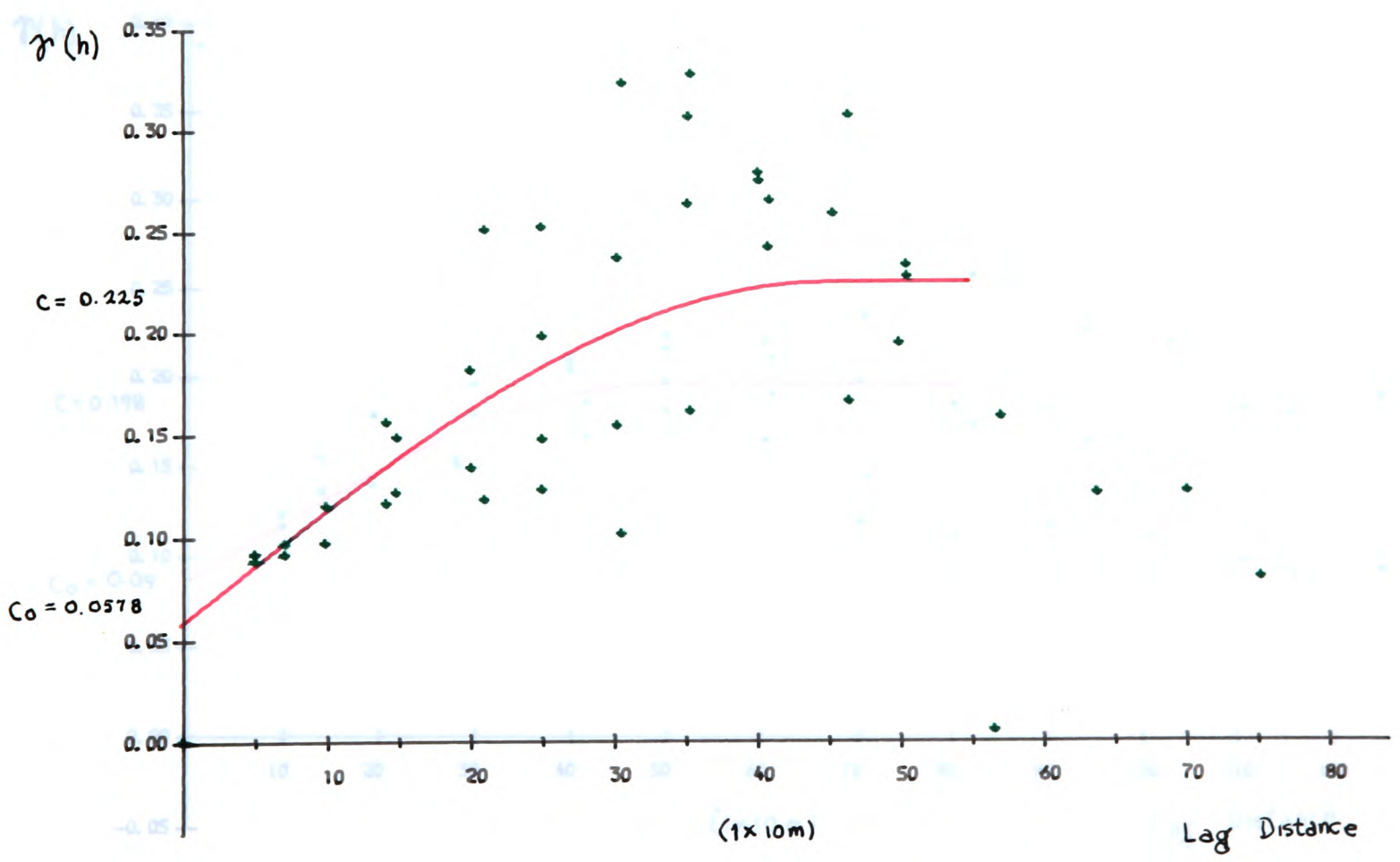
9.12.1 pH.

The semi-variograms of pH are shown in figures 9.5a to e. The semi-variogram for the 50m grid (fig. 9.5a), shows transitional variation for points within distances up to 200m. For distances in between 200m and 400m the points appear more scattered. This reflects the changes in pH that occur with changes in soils, geology and, more likely, the effect of the drift. The dispersion of points is caused mainly by low values of the semi-variances along the directions N-S and NE-SW suggesting that the transition from brown calcareous soils in the terrace gravels to the brown earths and non-calcareous soils of the alluvium, more evident for the other directions, might have been smoothed out only lightly by the drift. However, a spherical model could be fitted with a range of 450m.

The tightly defined shape of the semi-variogram for the 100m grid indicates a well defined spatial structure of transitional change at these lag distances. However, the nugget variance is

$$\gamma(h) = 0.0578 + 0.167 \left[\frac{3}{2}\frac{h}{450} - \frac{1}{2}\left(\frac{h}{450}\right)^3\right]$$

Fig. 9.5a



$$\gamma(h) = 0.09 + 0.108 \left[\frac{3}{2} \frac{h}{500} - \frac{1}{2} \left(\frac{h}{500} \right)^3 \right]$$

Fig. 9.5b

SEMI-VARIOGRAM of pH, Yelford, 100 m GRID

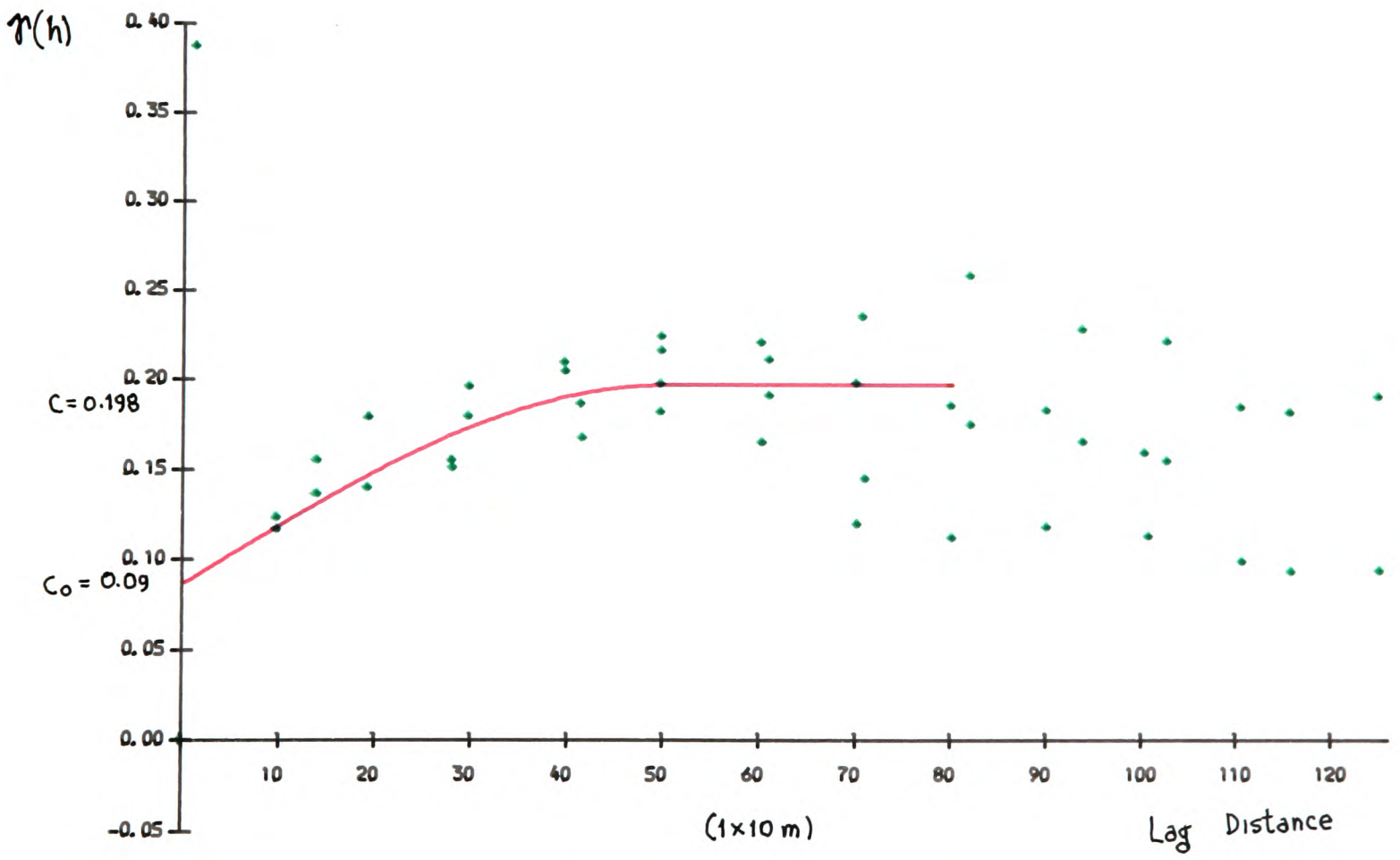
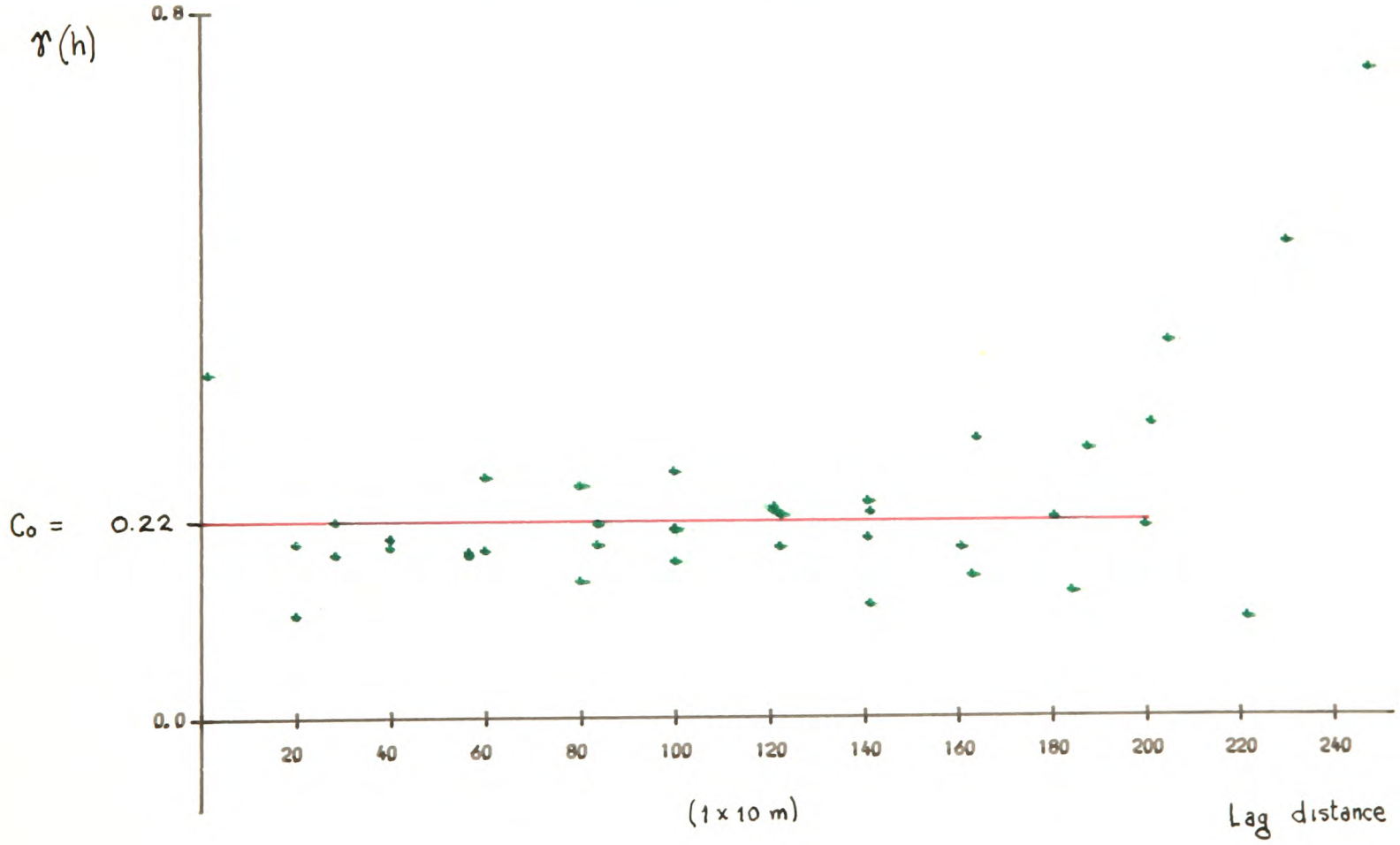


Fig 9.5c. SEMI-VARIOGRAM of pH, Yelford, 200 m GRID



$$\gamma(h) = 0.07 + 0.15 \left[\frac{3}{2} \frac{h}{55} - \frac{1}{2} \left(\frac{h}{55} \right)^3 \right]$$

NESTED STRUCTURES of pH, Yelford

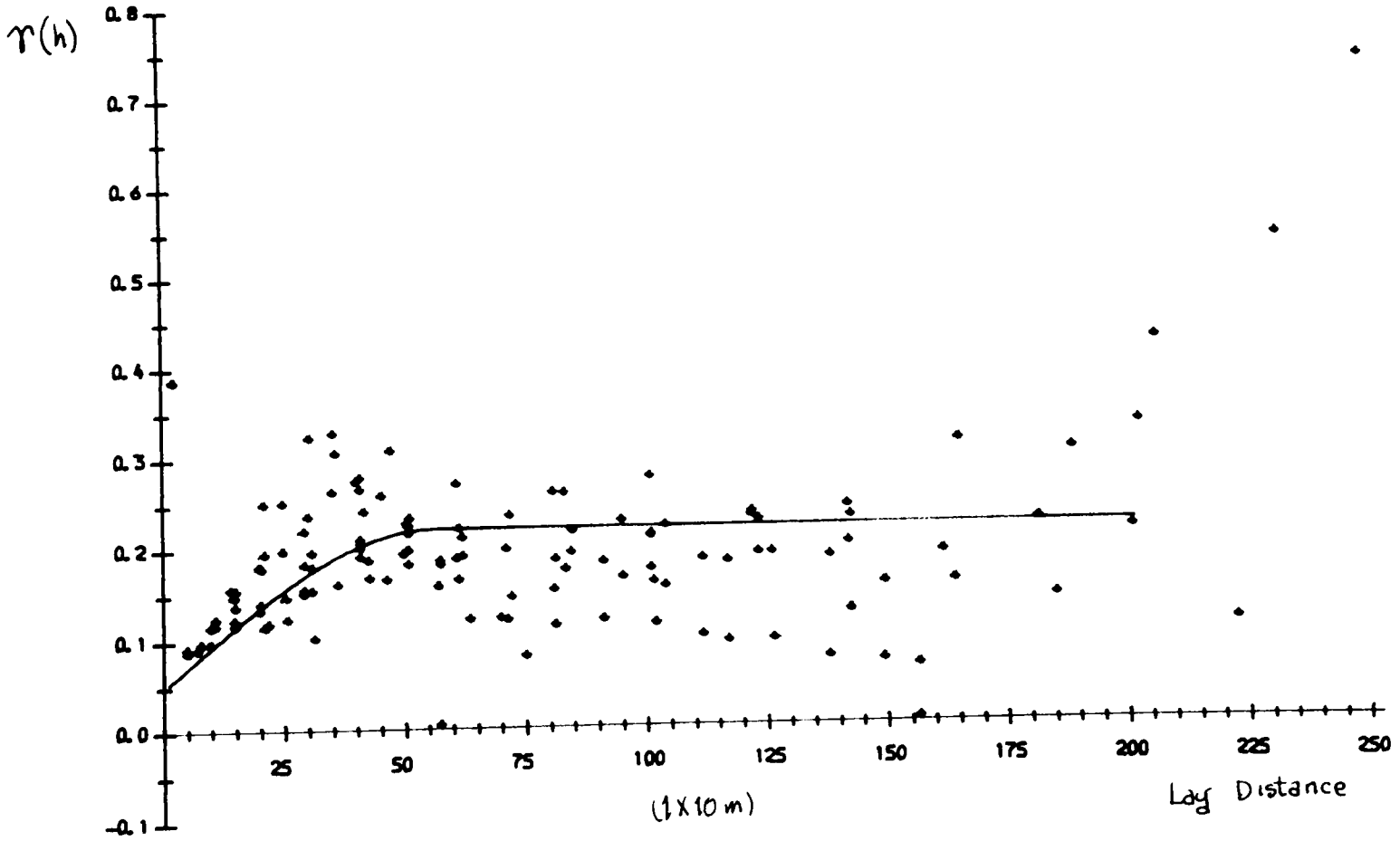


Fig. 9.5 d

$$\gamma(h) = 0.07 + 0.15\left[\frac{3}{2}\frac{h}{55} - \frac{1}{2}\left(\frac{h}{55}\right)^3\right]$$

pH , NESTED STRUCTURES (50, 100, 200m GRIDS), Yelford

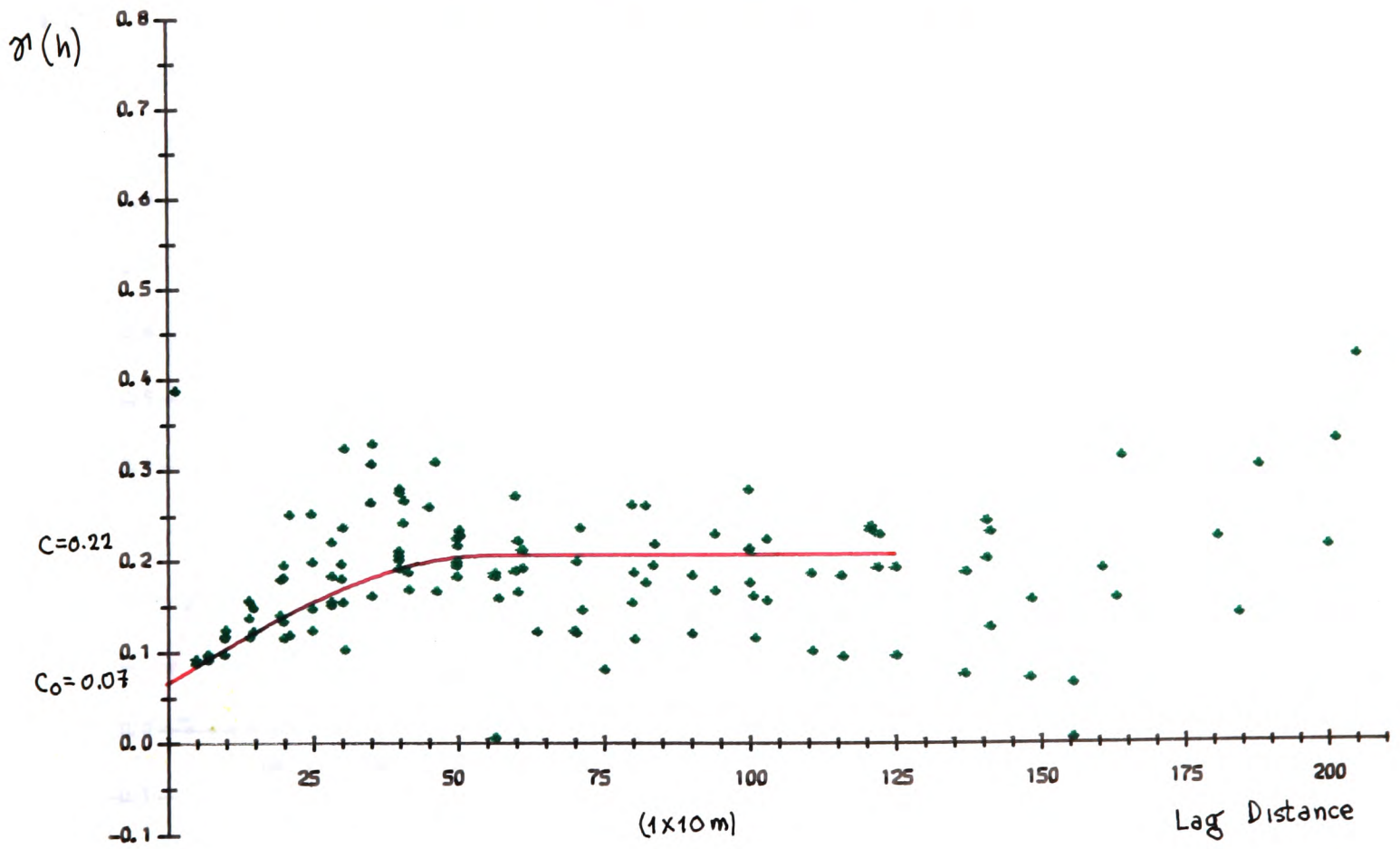


Fig. 9.5d

$C_0 = 0.047$

$C_1 = 0.373$ $\quad C_2 = 0.32$

$a_1 = 110.0$ $\quad a_2 = 0.470$

SEMI-VARIOGRAM of pH 1 Km GRID (NESTED STRUCTURES)

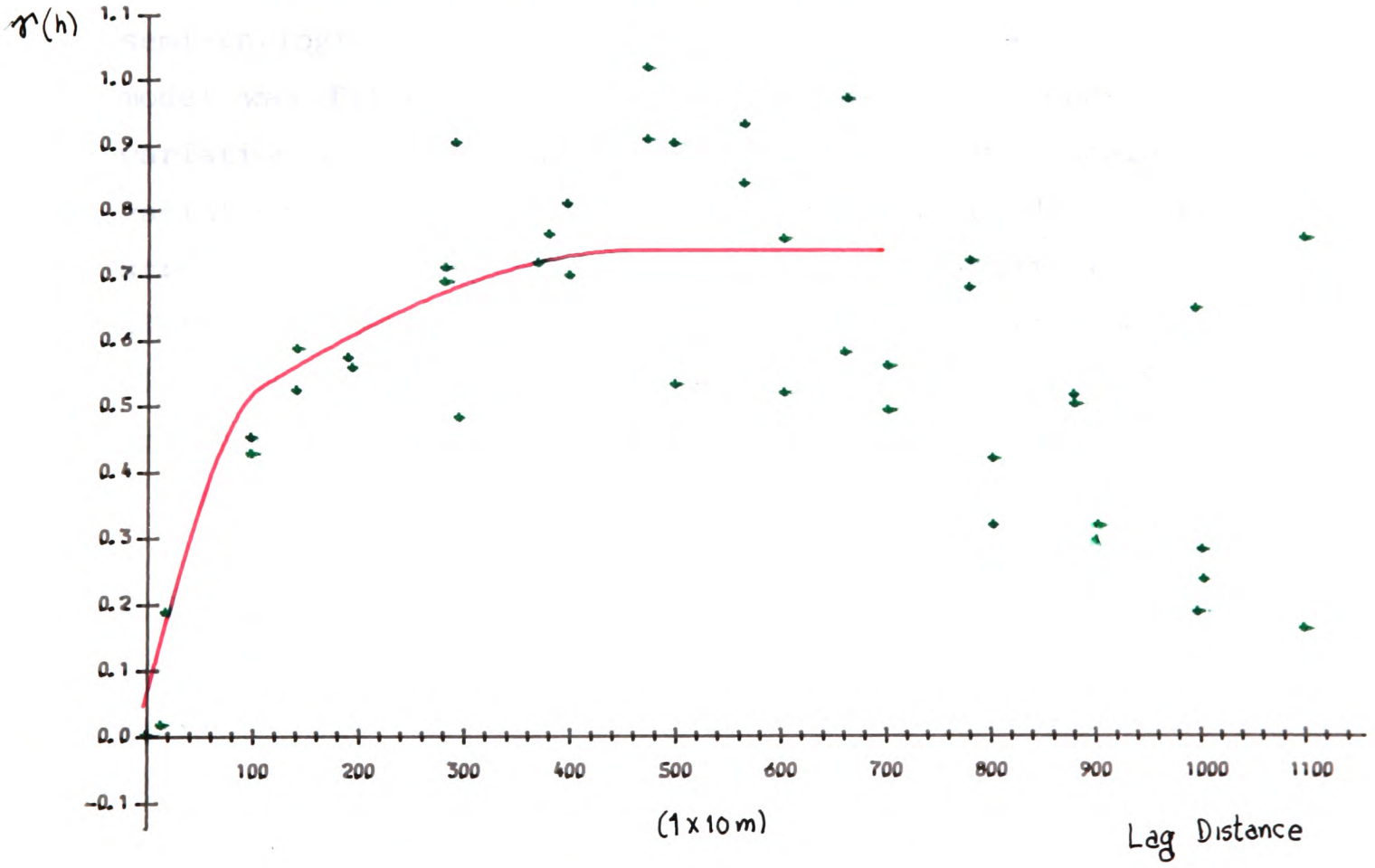


Fig. 9.5e

proportionally high and the spatial dependence comparatively small. The soil pH at this sampling scale is either "noisy" or there are structures inbeded in the nugget variance which require elucidation by sampling closer. A spherical model was found to fit this semi-variogram defining a range of 500m.

For the 200m grid the semi-variogram (fig 9.5c) of pH shows a shape that was interpreted as "pure nugget effect". This indicates total absence of spatial structure at this scale of sampling. Neighbourhoods as wide as 2 km can be defined within which the best estimator of pH is the mean of the samples. The apparent upsurge after lag 10 (2000m) is not a clear parabolic effect which would indicate the presence of trends. There are not sufficient number of pairs at the lags corresponding to the upsurge as to make these semi-variances reliable.

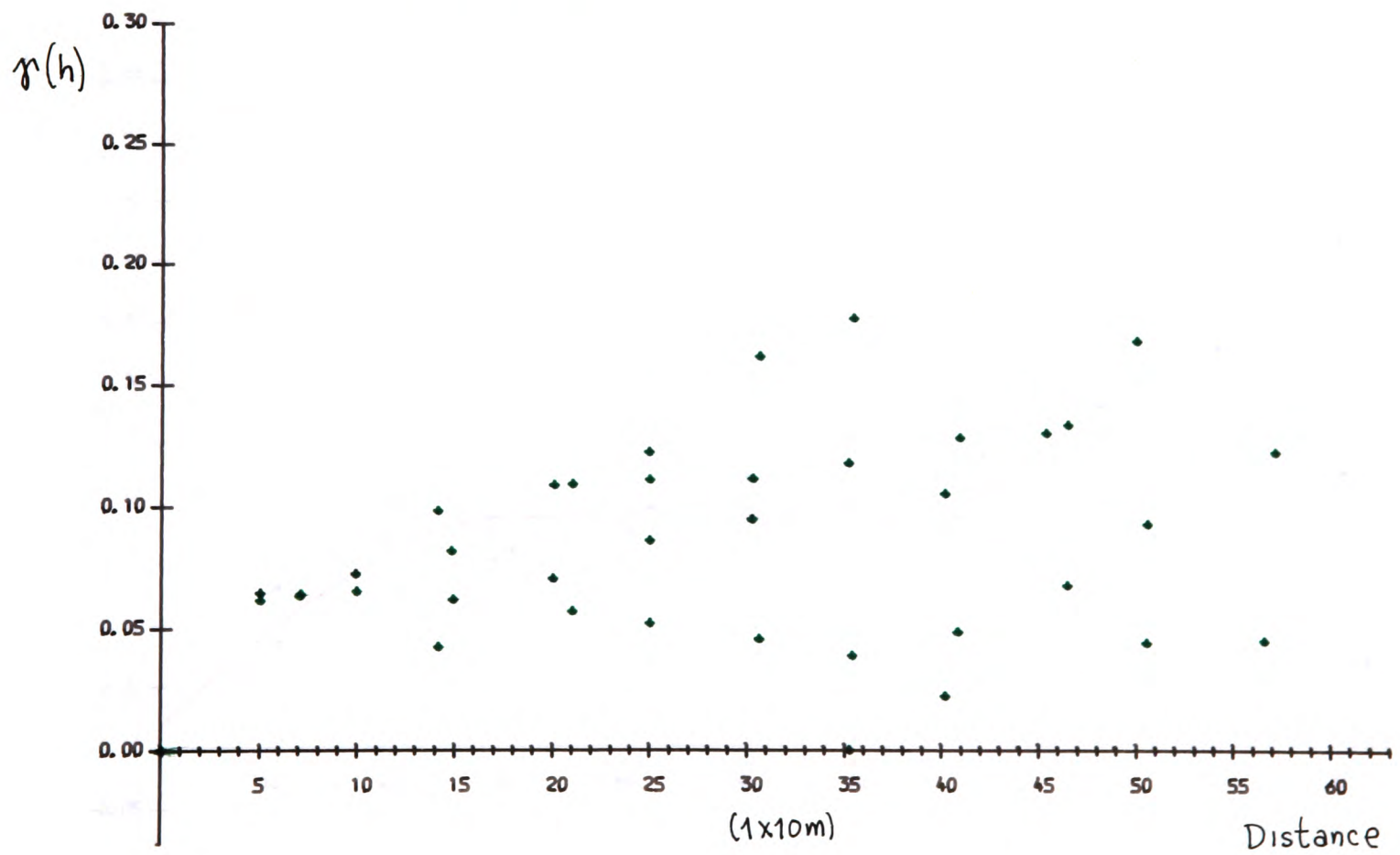
The semi-variogram of nested structures (fig 9.5d) is such that can be represented by a single structure and a spherical model. It indicates that there is a clear spatial structure in the variation of soil pH within these sampling scales.

Two nested structures of pH were found from the semi-variogram for the 1 km grid (fig 9.5e). A double spherical model was fitted. The first structure corresponds to spatial variation within 1 km. This is precisely the scale at which variation is elucidated by the smaller grids. The second structure corresponds to spatial variation occurring at distances between 1 km and 4 km. The latter may be thought as a limit to spatial autocovariance of pH in this area giving this scale of sampling.

9.12.2 Exchangeable Potassium (K).

The semi-variogram for the 50m. grid is shown in figure 9.6a. No isotropic model could be found to fit adequately this sort of variation. Although exchangeable potassium is a property most affected by management, and this may be reflected in the between-field variability at this scale of sampling, variation in this case is clearly anisotropic and due to the influence of the drift. The semi-variogram in the direction northeast-southwest showed a different pattern than those in the other directions. The semi-variances in this direction tended to be on the whole lower but without discernible pattern. Hence the broad scatter of points in the plot.

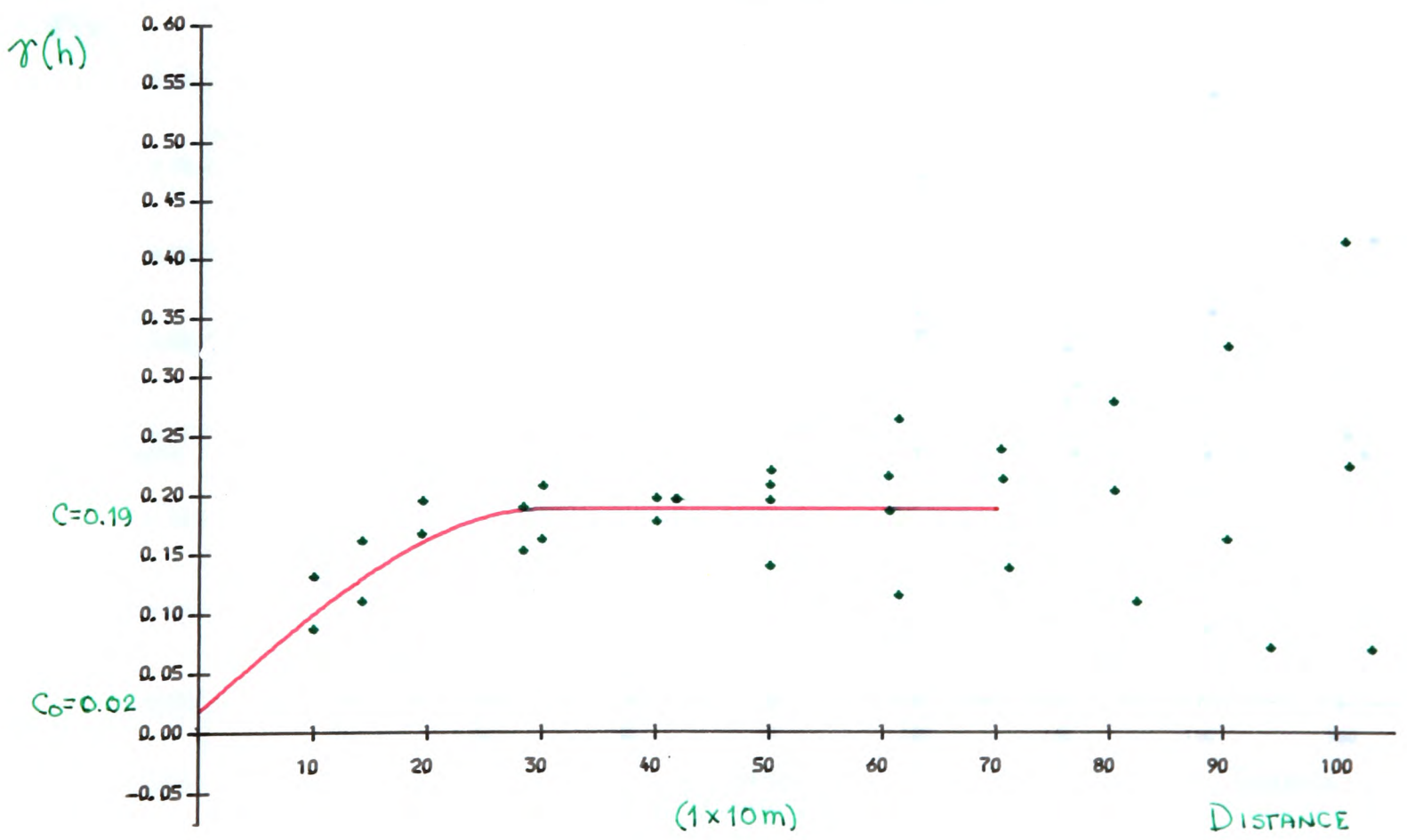
Fig. 9.6 a


POTASSIUM _ EXCHANGE, 50 m GRID, YELFORD
γ(h)
0.30
0.25
0.20
0.15
0.10
0.05
0.00
5
10
15
20
25
30
35
40
45
50
55
60
(1x10m)
Distance

$$\gamma(h) = 0.02 + 0.19\left[\frac{3}{2}\frac{h}{30} - \frac{1}{2}\left(\frac{h}{30}\right)^3\right]$$

Fig. 9.6.b

SEMI-VARIOGRAM of POTASSIUM, 100 m GRID, YELFORD



$$\gamma(h) = 0.145 + 0.21 \left[\frac{3}{2} \frac{h}{80} - \frac{1}{2} \left(\frac{h}{80} \right)^3 \right]$$

Fig. 9.6c

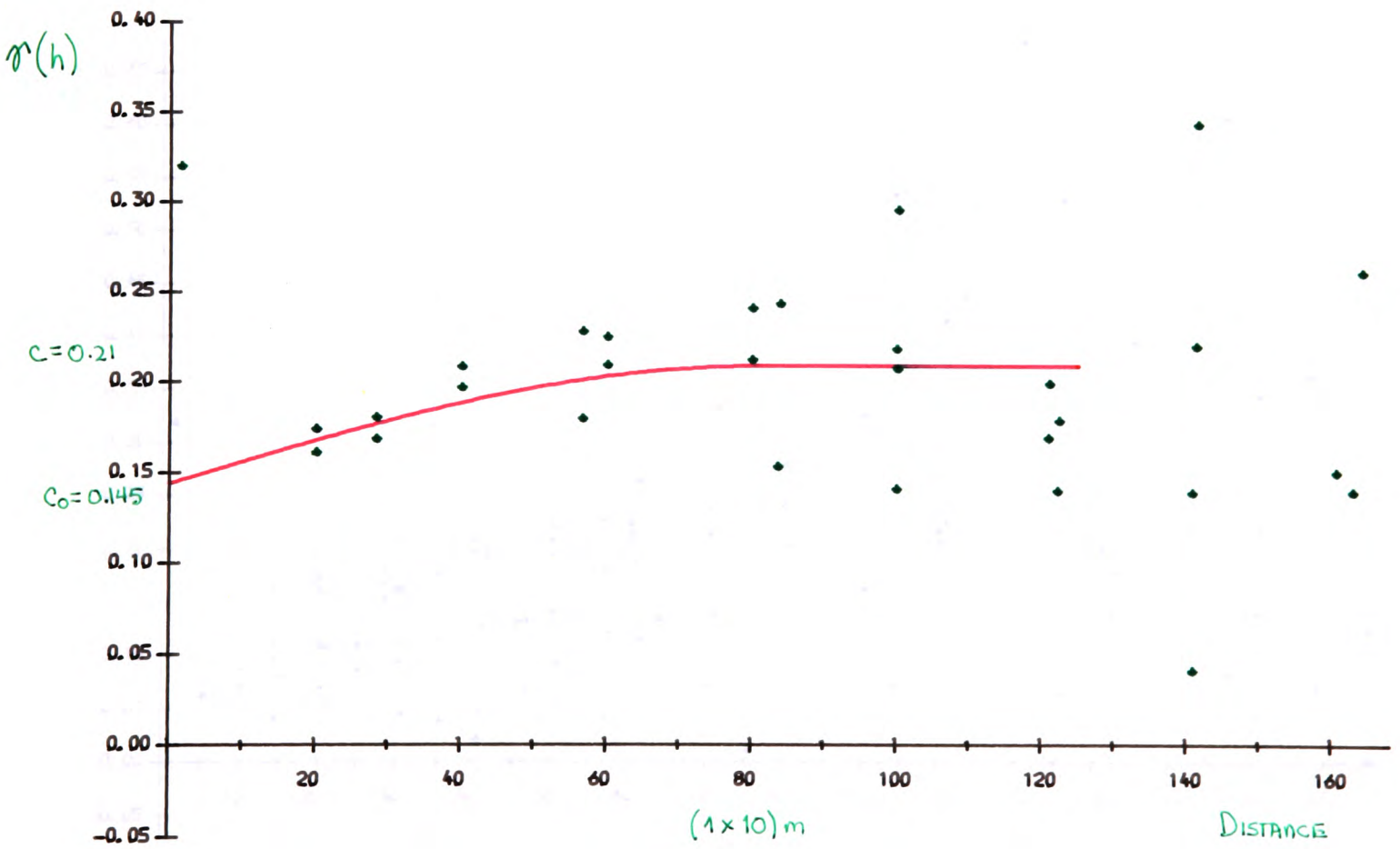


Fig. 9.6.d

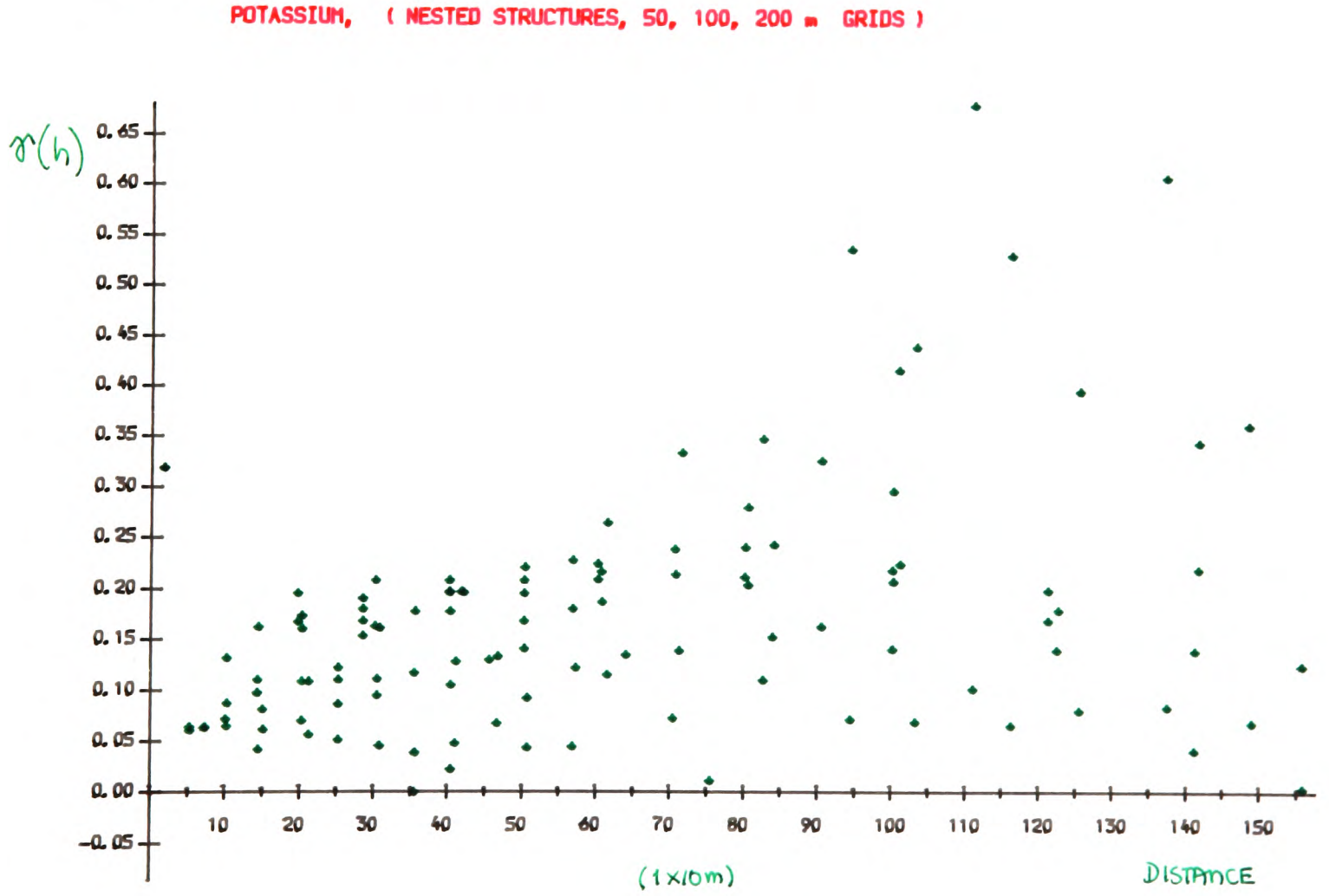
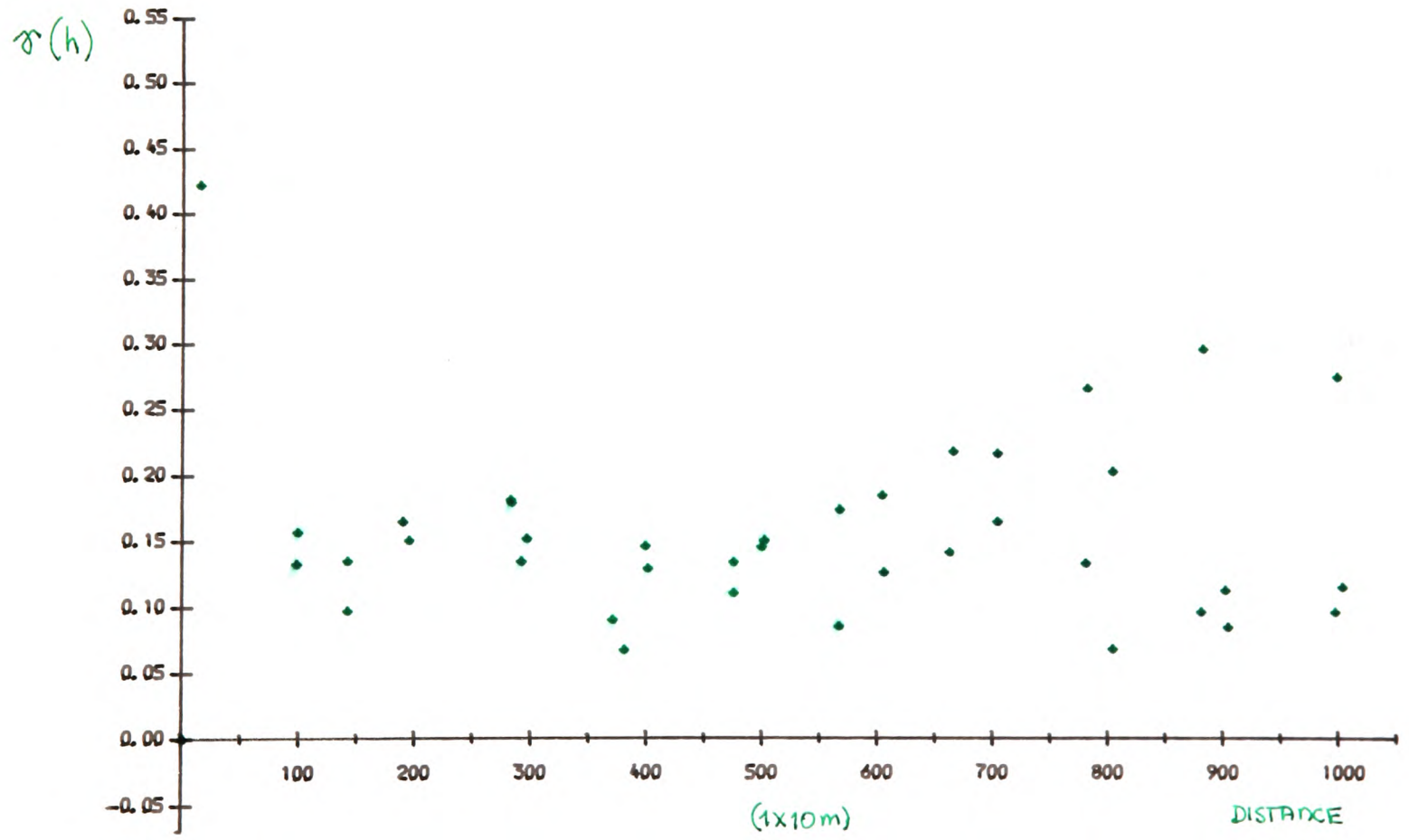

POTASSIUM, (NESTED STRUCTURES, 50, 100, 200 m GRIDS)
γ(h)
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
-0.05
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
(1x10m)
DISTANCE

Fig. 9.6 e POTASSIUM, 1 Km GRID, SP30 SHEET, (NESTED STRUCTURES



In contrast, variability of Potassium at another scale (100m grid spacing), showed some spatial structure for lag distances up to 300m (fig. 9.6b), and relatively small nugget variance. However, spatial dependence is still somewhat weak. This weakness is even clearer in the semi-variogram for the 200m grid, where spatial dependence is even weaker and the nugget variance is considerable. A range could be found at about 800m for this semi-variogram.

The semi-variograms of figures 9.6d and 9.6e calculated so that they show the nested structures of variation, serve to illustrate this point even further. No apparent spatial structure can be found in the semi-variogram of the three smaller grids, and the semi-variogram for the grid of 1km spacings appears to be characterized by pure nugget variance.

The semi-variograms of Potassium suggest, in general, the presence of anisotropic variation at the shortest sampling spacing, and a weak spatial structure in the variability of this property which may be interpreted as the effect of management, notably additions of chemical fertilizers, superimposed to variability caused by other sources.

9.12.3 Exchangeable Magnesium (Mg).

The variability of Mg for the 50m grid showed a good degree of spatial dependence within distances of 250m (fig. 9.7a). The nugget variance was relatively low in this semi-variogram.

Unlike K, Magnesium is not as much affected by management but by variation in parent material and the local geology. The range found for this semi-variogram corresponds roughly to changes in soil type, i.e. between the calcareous soils on the terrace gravel and the brown earths gley soils downslope.

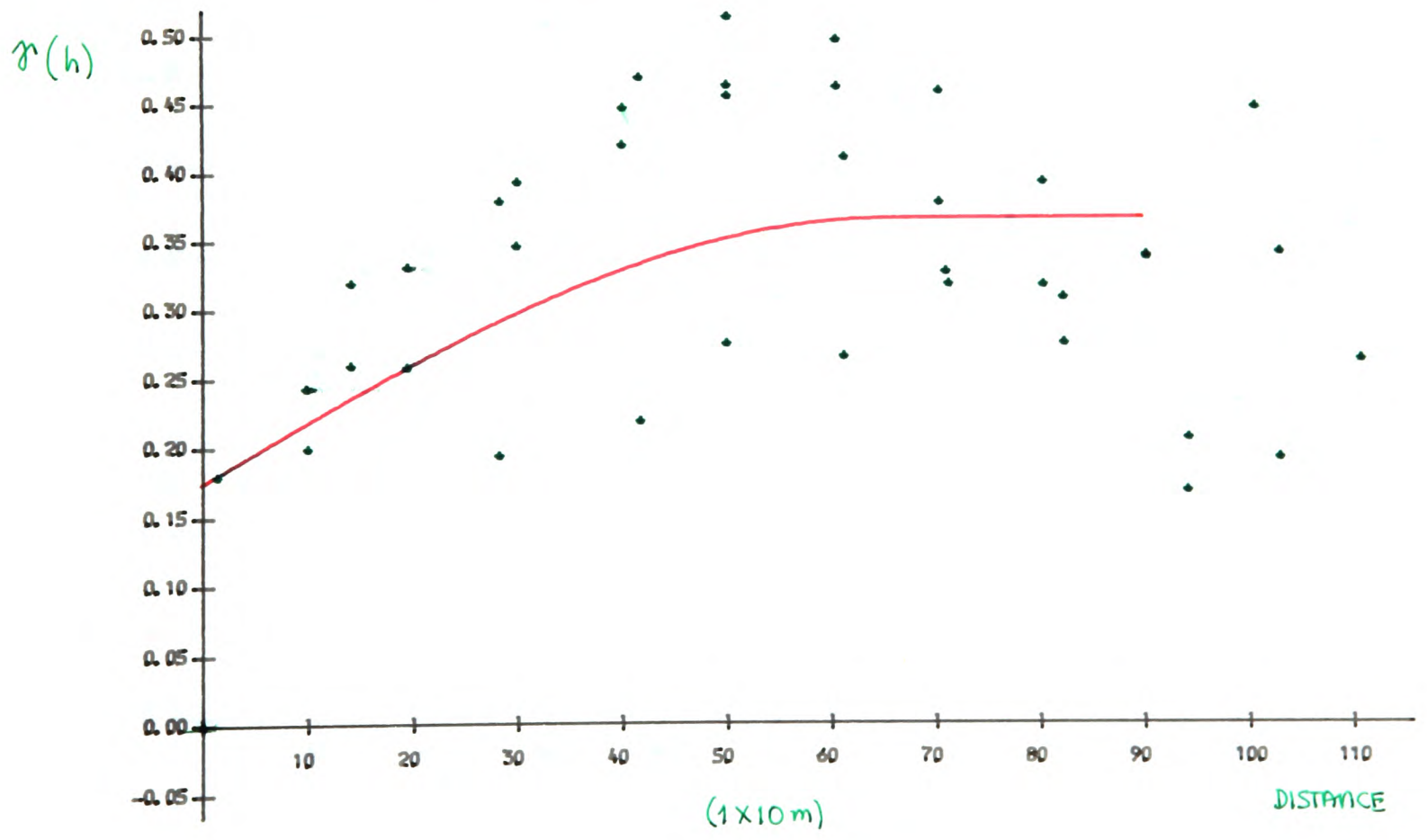
The semi-variogram for the 100m grid, (fig. 9.7b), showed spatial dependence, but in the direction NE-SW, (the direction of the drift) a tendency towards anisotropic variation represented by most of the low semi-variances beneath the curve of the model in figure 9.7b.

Since this departure was not judged strong enough to be considered anisotropy, a spherical model could still be fitted to the points. A range was found at 650m and a relatively high nugget varince.

$$\gamma(h) = 0.173 + 0.192\left[\frac{3}{2}\frac{h}{650} - \frac{1}{2}\left(\frac{h}{650}\right)^3\right]$$

Fig 9.7 b

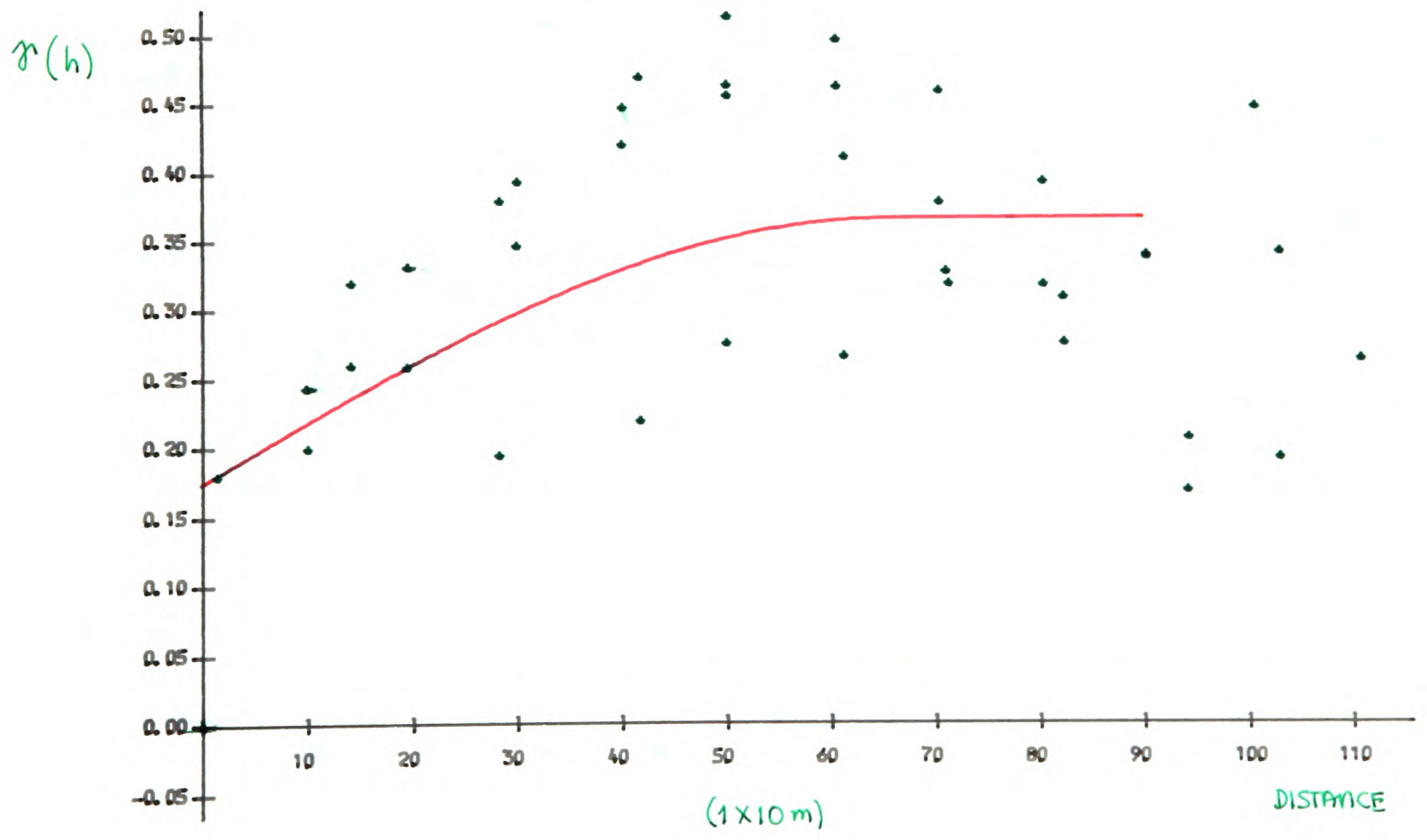
SEMI-VARIOGRAM of Mg exch., 100 m grid, Yelford



$$\gamma(h) = 0.173 + 0.192\left[\frac{3}{2}\frac{h}{650} - \frac{1}{2}\left(\frac{h}{650}\right)^3\right]$$

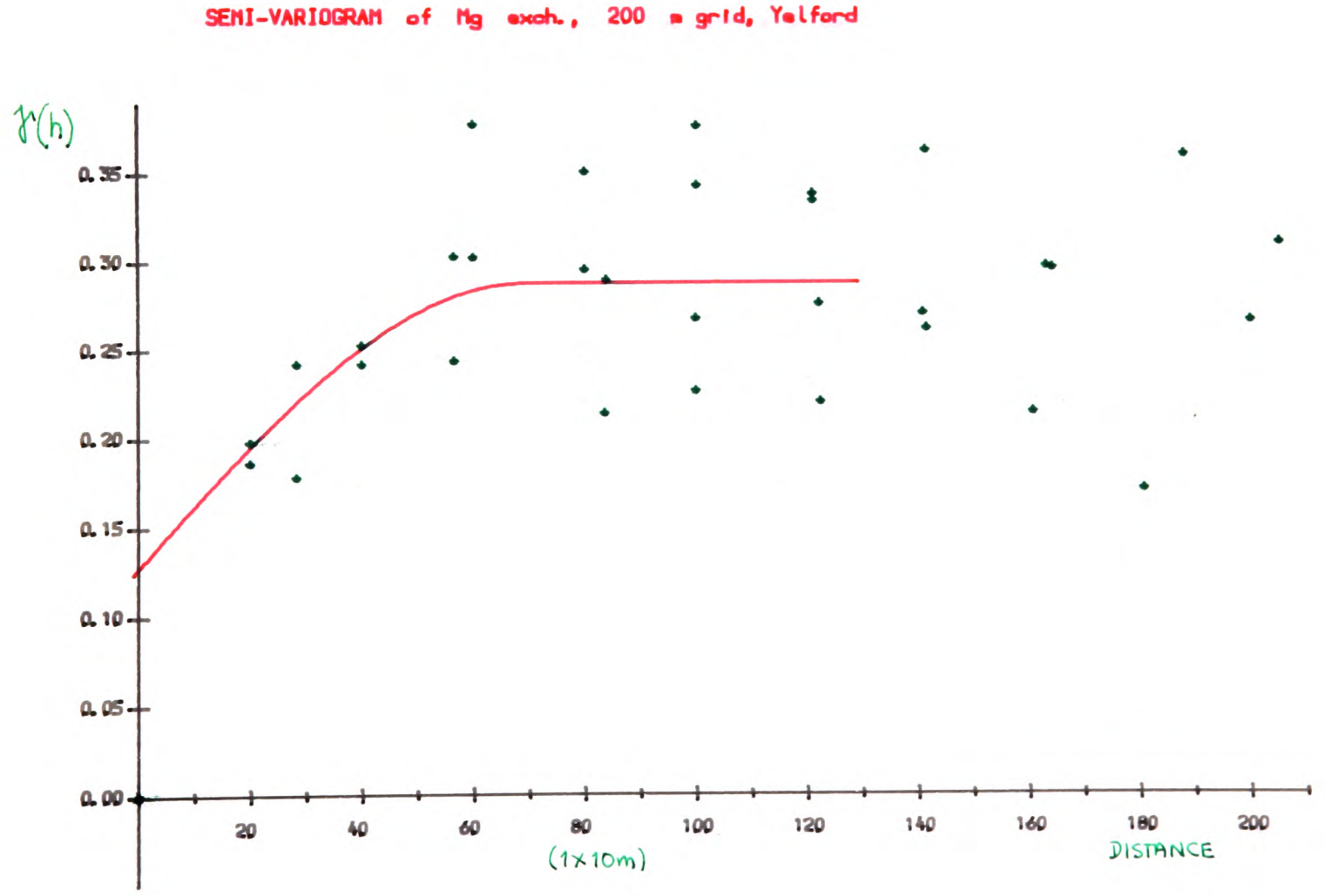
Fig 9.7 b

SEMI-VARIOGRAM of Mg exch., 100 m grid, Yelford



$$\gamma(h) = 0.125 + 0.160 \left[\frac{3}{2} \frac{h}{700} - \frac{1}{2} \left(\frac{h}{700} \right)^3 \right]$$

Fig. 9.7c



$C_o = 0.0685$

$C_1 = 0.1515$ $C_2 = 0.145$

$a_1 = 33$ $a_2 = 70$

Fig. 9.7d

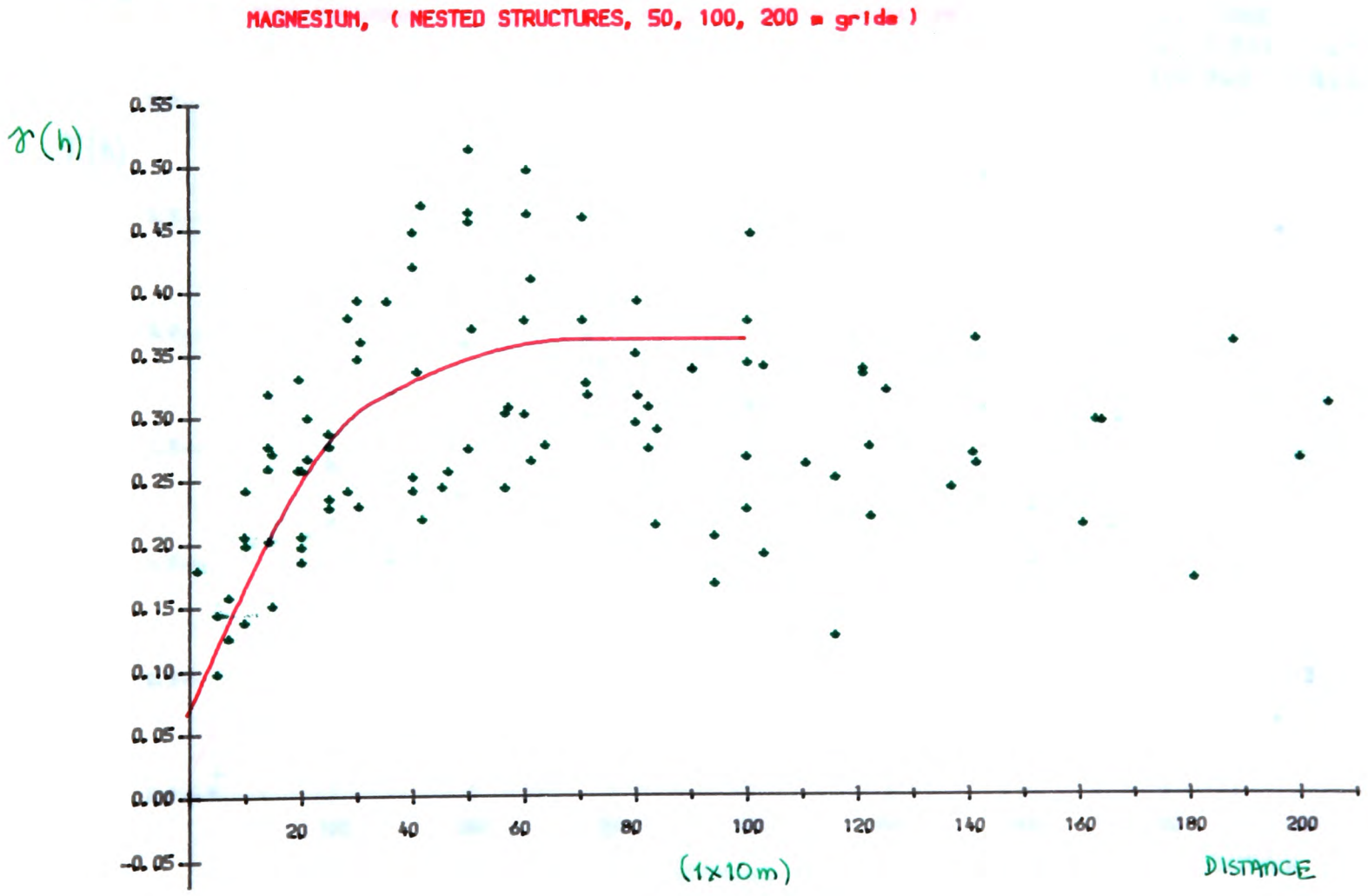
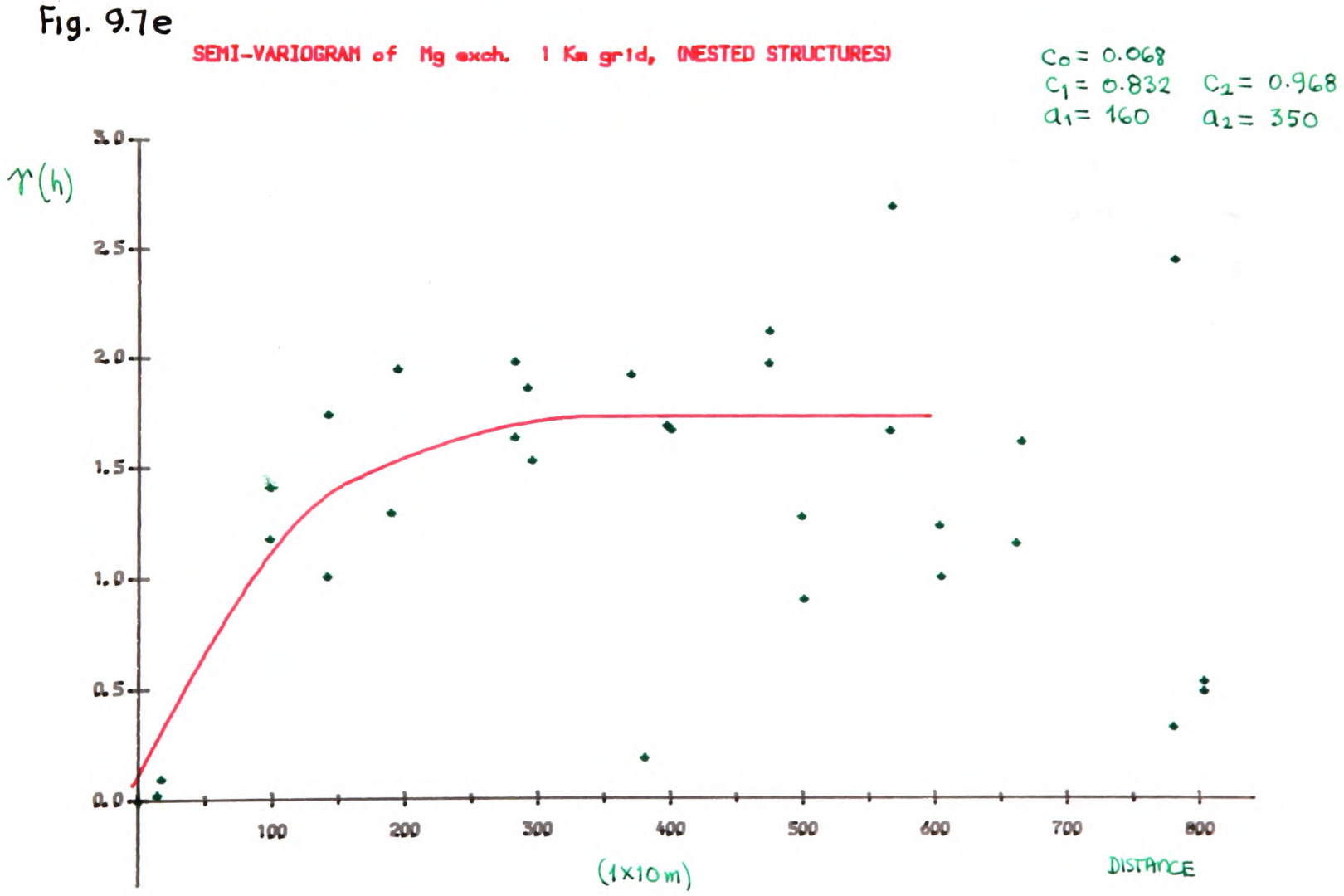


Fig. 9.7e


SEMI-VARIOGRAM of Mg exch. 1 Km grid, (NESTED STRUCTURES)
$C_0 = 0.068$
$C_1 = 0.832$ $C_2 = 0.968$
$a_1 = 160$ $a_2 = 350$
γ(h)
3.0
2.5
2.0
1.5
1.0
0.5
0.0
100
200
300
400
500
600
700
800
(1x10 m)
DISTANCE

Spatial dependence at a smaller scale, (200m grid), was less marked than in the previous grid. The property at this sampling spacing showed more variation round the sill, and the range found at 700m corresponds approximately to the range of distance within which changes in geology occur, i.e. from gravel terrace and Oxford clay to the older alluvium on the lower terrace. That is from calcareous landscape to non-calcareous and gley soils. The variations round the sill indicate random effects due to any other sources.

The semi-variogram of the nested structures (50m, 100m, and 200m) showed clearly (fig. 9.7d) the two structures discussed above. The range of a first structure at about 300m corresponding to the spatial dependence found in the smallest grid, and a second structure with a range of 700m which matches major changes in geology and soils.

9.12.4 Total Carbonates.

Anisotropic variation on Carbonate is clearly evident in the semi-variogram for the 50m grid in figure 9.8a. Spatial dependence is evident for directions NW-SE and even N-S in the semi-variogram, running from the calcareous terrace gravel down to the brown earths with gleying soils. The set of lower points in the semi-variogram correspond to direction NE-SW which runs with the drift, so the Carbonate values are less different for they have been smoothed out and "averaged" by the drift.

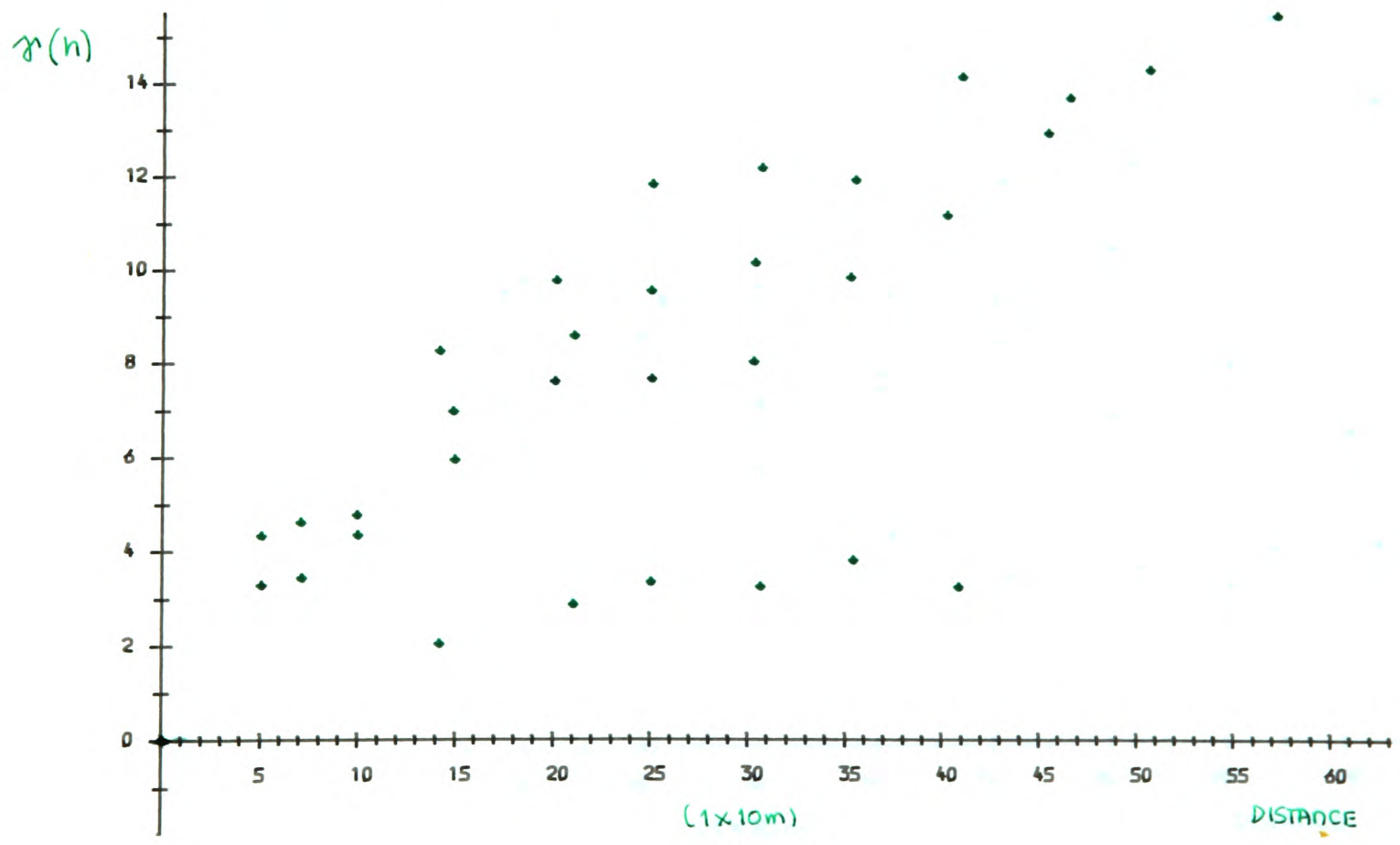
No isotropic model can be fitted to this type of variation. Cases of anisotropy were approached as described later in this chapter.

The semi-variogram for the 100m grid showed spatial dependence for distances up to 350m. This range may be associated with the transition from the calcareous gravel terrace to the alluvium. The nugget variance was relatively small indicating the transitional nature of the variation of this property at this sampling scale.

The exponential model fitted to the semivariogram for the 200m grid is an indication of smooth gradual changes and long transitions on Carbonate at this scale of sampling. Spatial dependence goes as far as 1600m. This can be explain by the fact that the 200m grid covers the whole transition from

Fig 9.8a

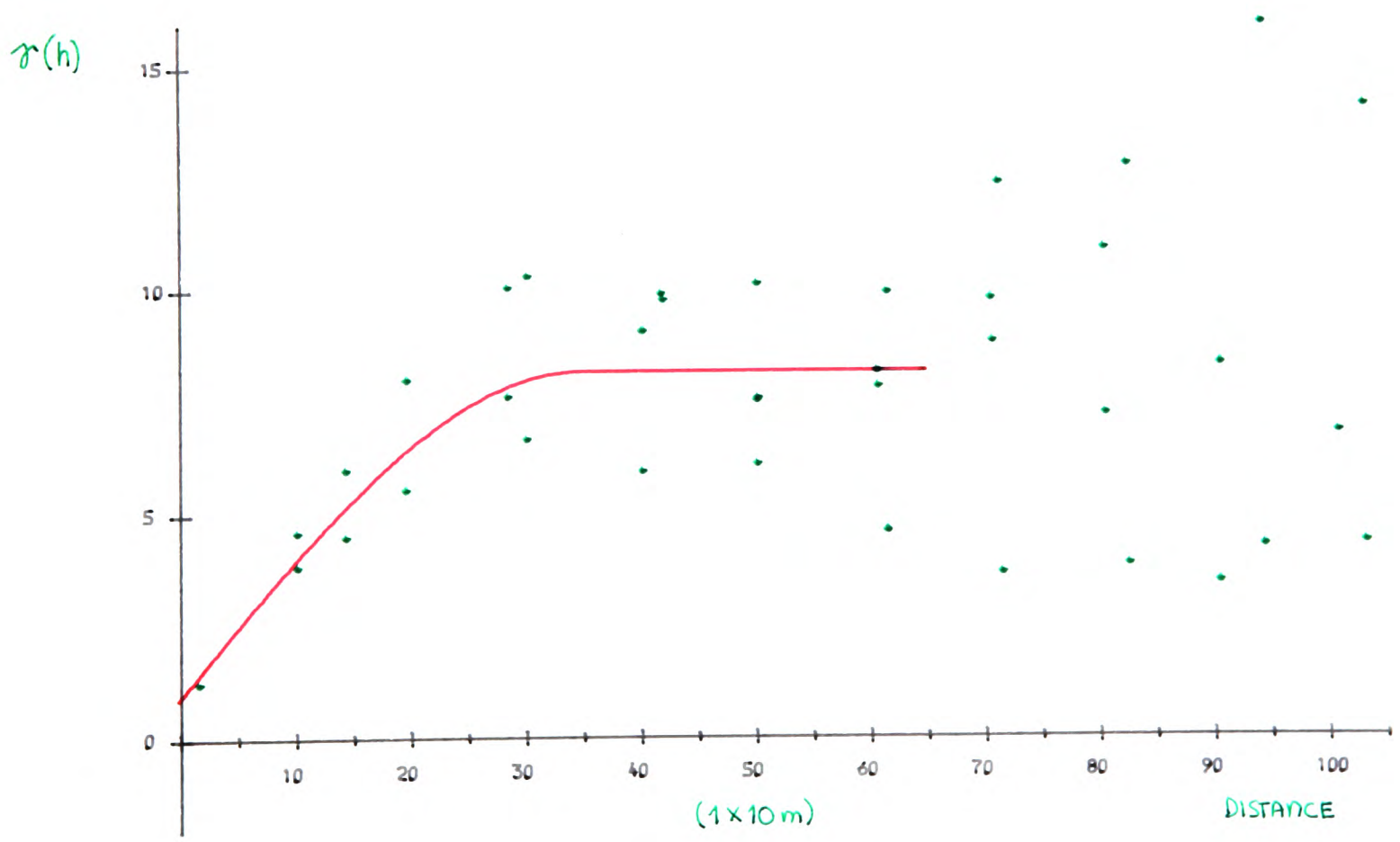
CARBONATE, 50 m GRID, YELFORD



$$\gamma(h) = 1.0 + 7.25\left[\frac{3}{2}\,\frac{h}{350} - \frac{1}{2}\left(\frac{h}{350}\right)^3\right]$$

Fig 9.8b

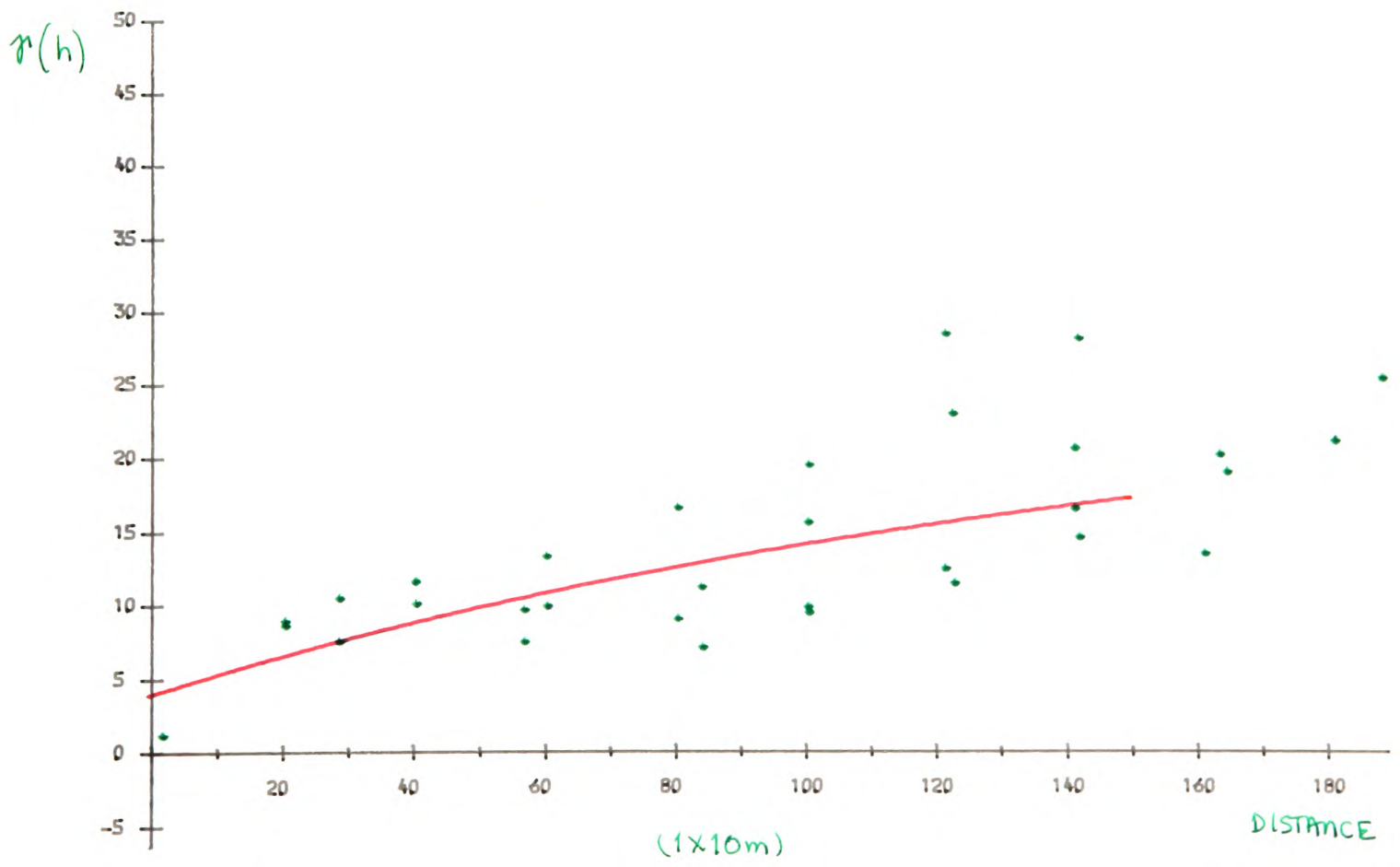
SEMI-VARIOGRAM of CARBONATE, 100 m grid, Yelford



$$\gamma(h) = 4.0 + 22\left[1 - \exp\left(\frac{-h}{1600}\right)\right]$$

Fig 9.8c

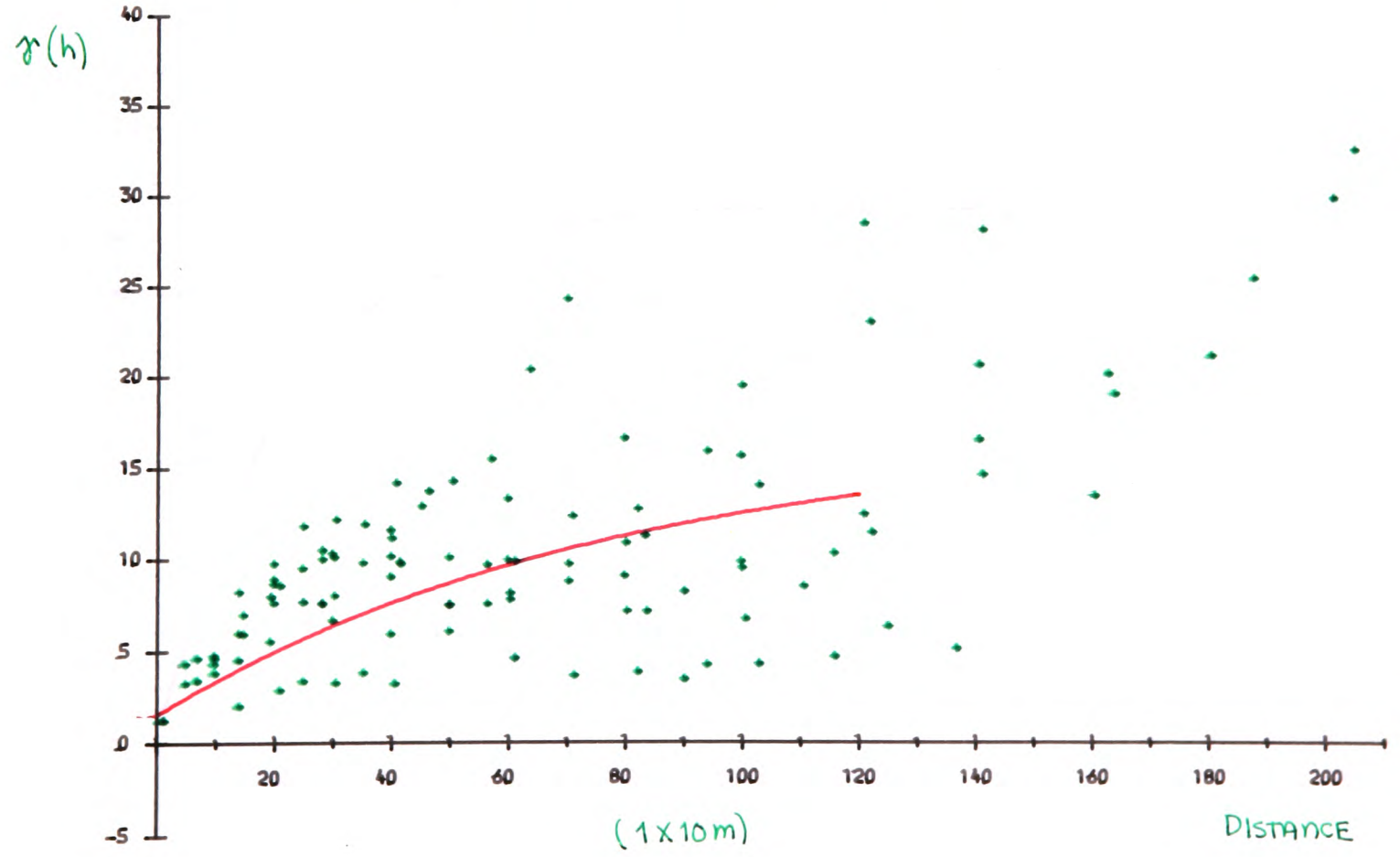
SEMI-VARIOGRAM of CARBONATE, 200 m grid, Yelford



$$\gamma(h) = 1.5 + 15.5\left[1 - \exp\left(\frac{-h}{1.60}\right)\right]$$

Fig. 9.8 d

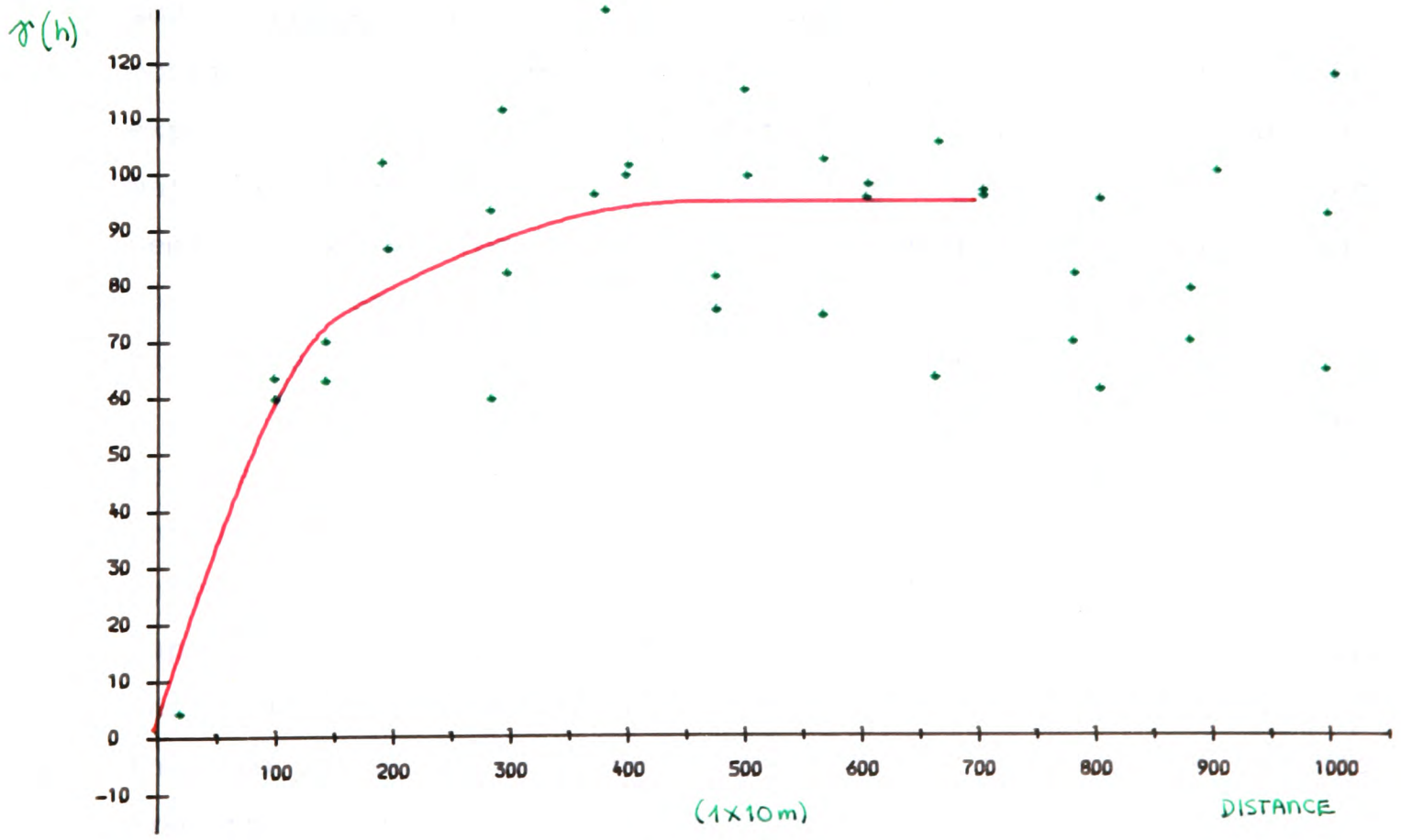
CARBONATE, (NESTED STRUCTURES, 50, 100, 200 m grids)



$C_0 = 1.5$
$q_1 = 160$ $C_2 = 40$
$C_1 = 53.5$ $q_2 = 460$

Fig. 9.8e

SEMI-VARIOGRAM of CARBONATE, 1 Km grid NESTED STRUCTURES



calcareous gravels on the terrace down to the non-calcareous alluvial soils.

The exponential model fitted to the semi-variances from the nested sampling grids (fig 9.8d) provide more evidence of the transitional variation of this property.

The first structure of the nested model of fig 9.8e, remains conjectural since there are few points and variation that would correspond to this structure (50m grid) is anisotropic. However, the second structure of this semi-variogram levels up smoothly to the sill defining a range of 4.6 km within which spatial autocorrelation is still present and Kriging of this property is possible.

9.12.5 Clay.

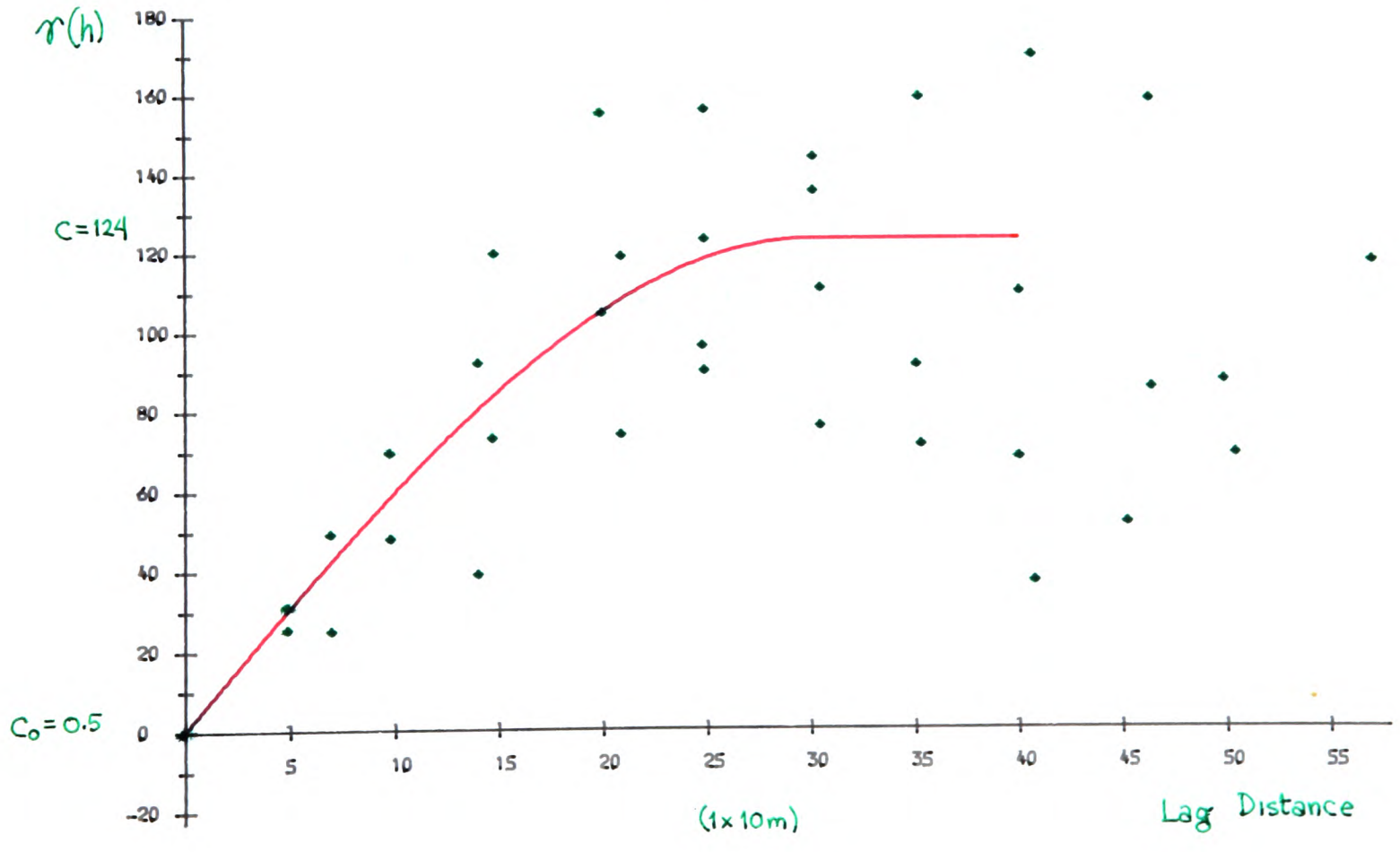
The semi-variograms of clay in the Yelford area all are bounded and show spatial structure characteristic of transition phenomena. A set of spherical models with distinct parameters was fitted to the semi-variograms of the three grids. On the whole, this property exhibits a greater degree of spatial structure than properties more affected by management. At the 50m grid, (fig 9.9a), the semi-variogram shows a considerably small nugget variance which indicates a good degree of continuity in the variation of this property. This was to be expected given the clayey nature of the geology, parent materials and soils in this area. The nugget variances of the semi-variograms for the 100m and 200m grids are small and comparable and reflect, predominantly, structure impossible to elucidate at the corresponding sampling spacings.

It is worth noticing that the ranges of influence found for both semi-variograms: for the 50m and 100m grids, (figs. 9.9a and 9.9b), are almost the same (300m and 330m). These distances correspond closely to the average inter-boundary spacings of the mapping units in Rowe's map (fig 9.3), and soil series mapping units of Jarvis and Hazelden (fig 9.2), over the area covered by the two grids. Thus, the neighbourhoods that can be defined from those ranges of influence make some sense to the pattern of variability as resolved by the boundaries of the mapping units.

The well-defined shape of the semi-variogram for the nested grids (fig 9.9d), and its neglegible nugget variance illustrate the continuity and strong spatial structure of variation of this

$$\gamma(h) = 0.5 + 123.5 \left[\frac{3}{2} \frac{h}{300} - \frac{1}{2} \left(\frac{h}{300} \right)^3 \right]$$

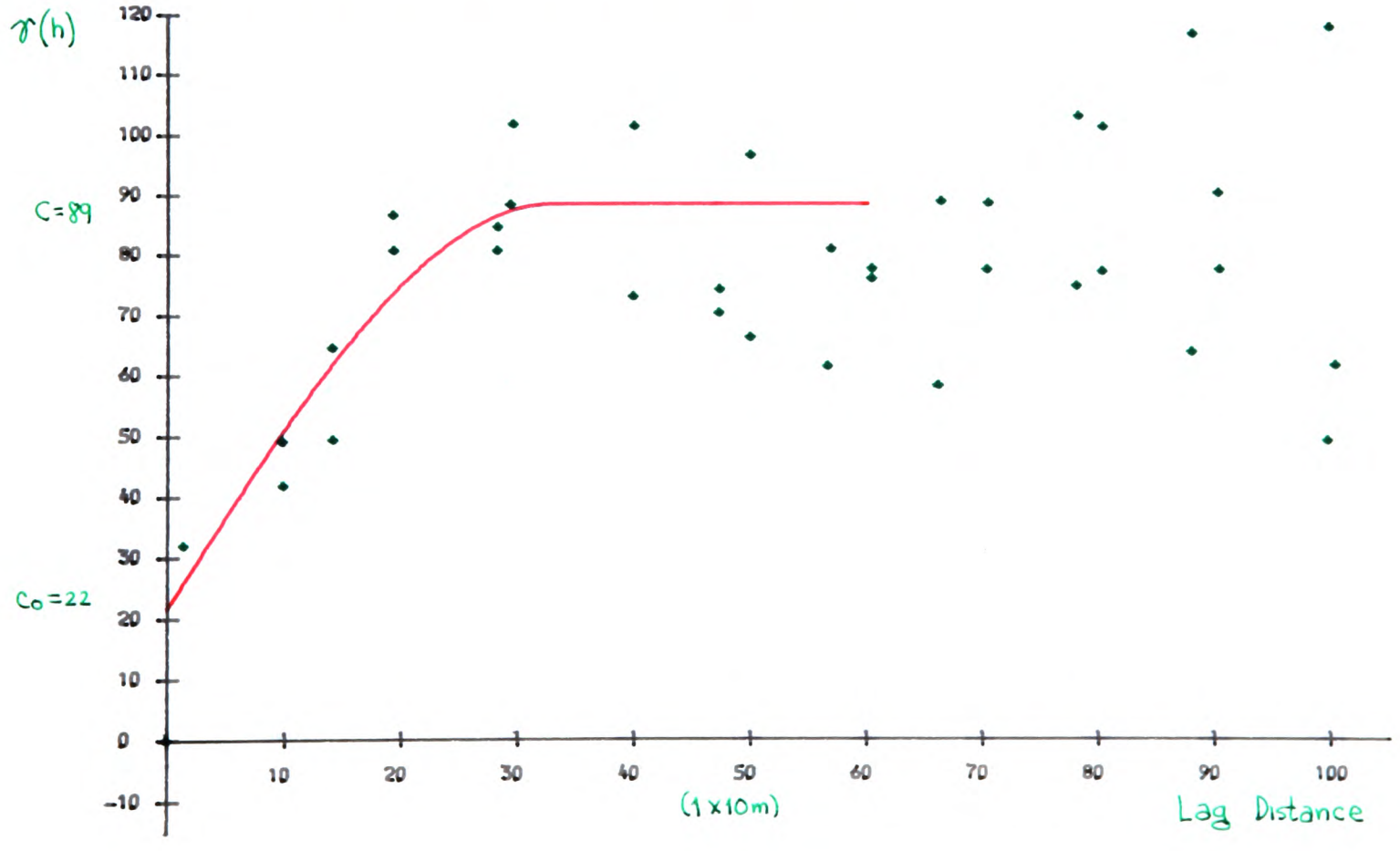
Fig 9.9 a SEMI-VARIOGRAM of CLAY, 50 m grid, Yelford



$$\gamma(h) = 22 + 67\left[\frac{3}{2}\,\frac{h}{330} - \frac{1}{2}\left(\frac{h}{330}\right)^3\right]$$

Fig 9.9b

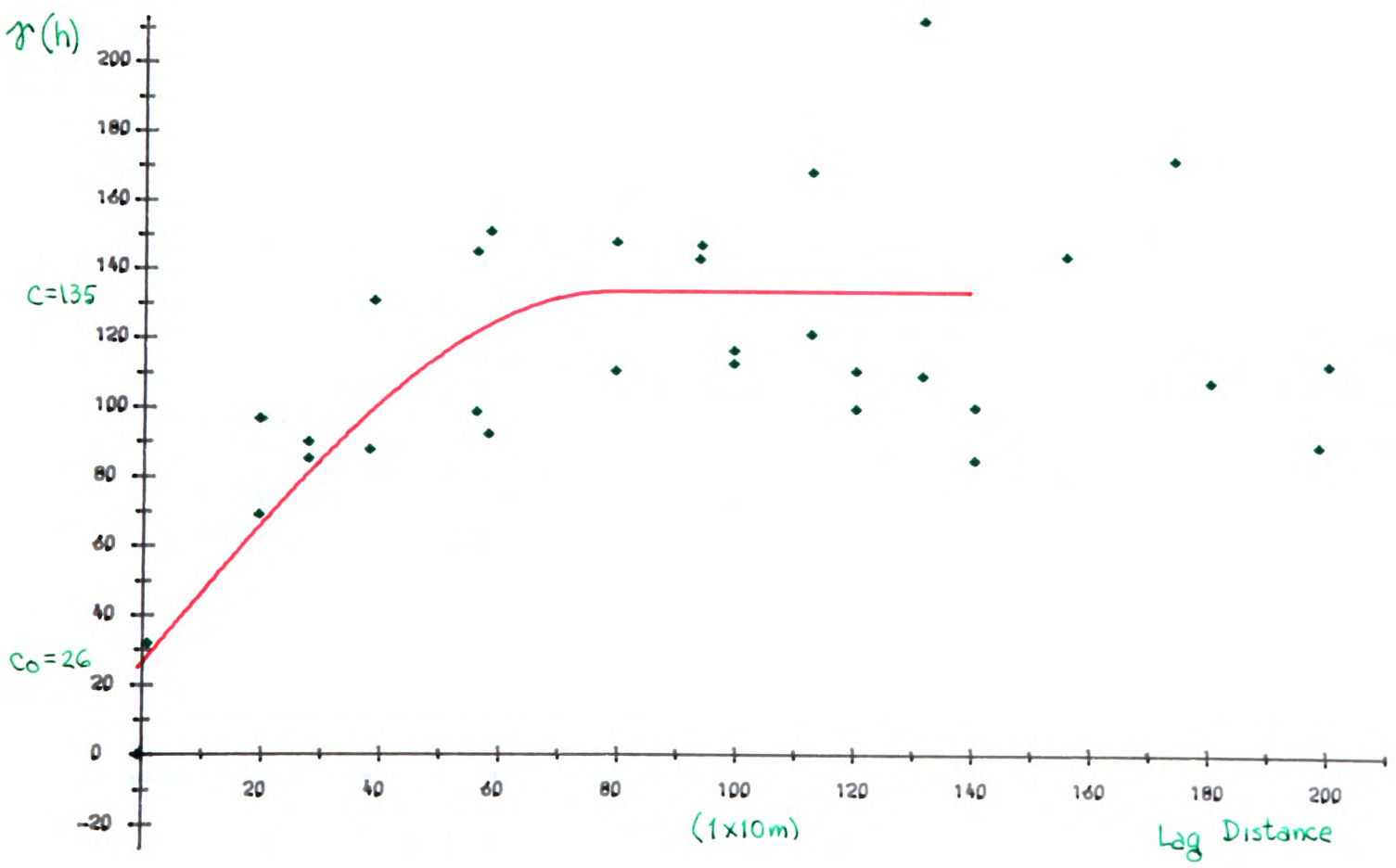
SEMI-VARIOGRAM of CLAY, 100 m grid, Yelford



$$\gamma(h) = 26 + 109\left[\frac{3}{2}\frac{h}{790} - \frac{1}{2}\left(\frac{h}{790}\right)^3\right]$$

Fig 9.9 c

SEMI-VARIOGRAM of CLAY, 200 m grid, Yelford



$$\gamma(h) = 6 + 94.\left[\frac{3}{2}\frac{h}{300} - \frac{1}{2}\left(\frac{h}{300}\right)^3\right]$$

Fig 9.9d

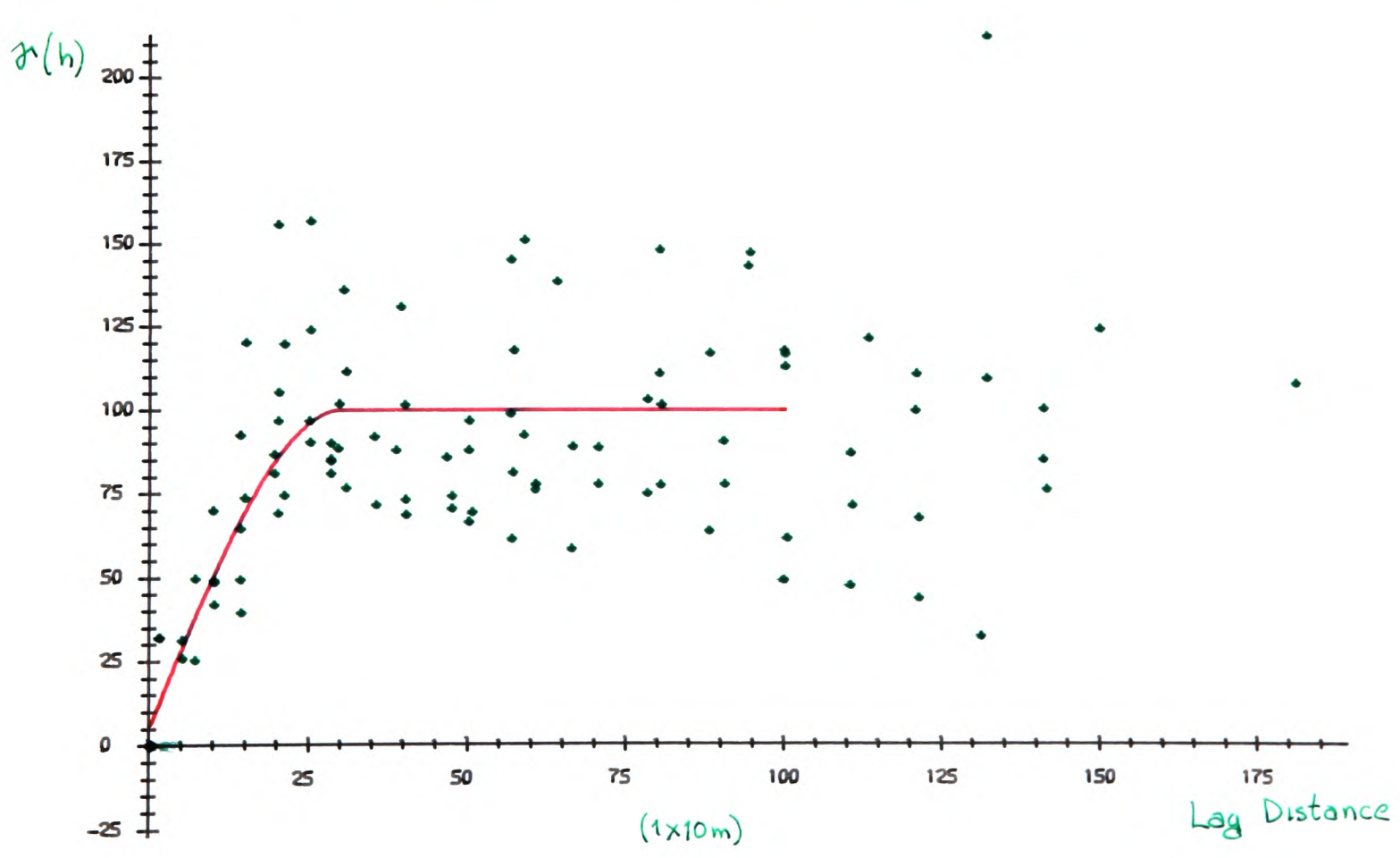
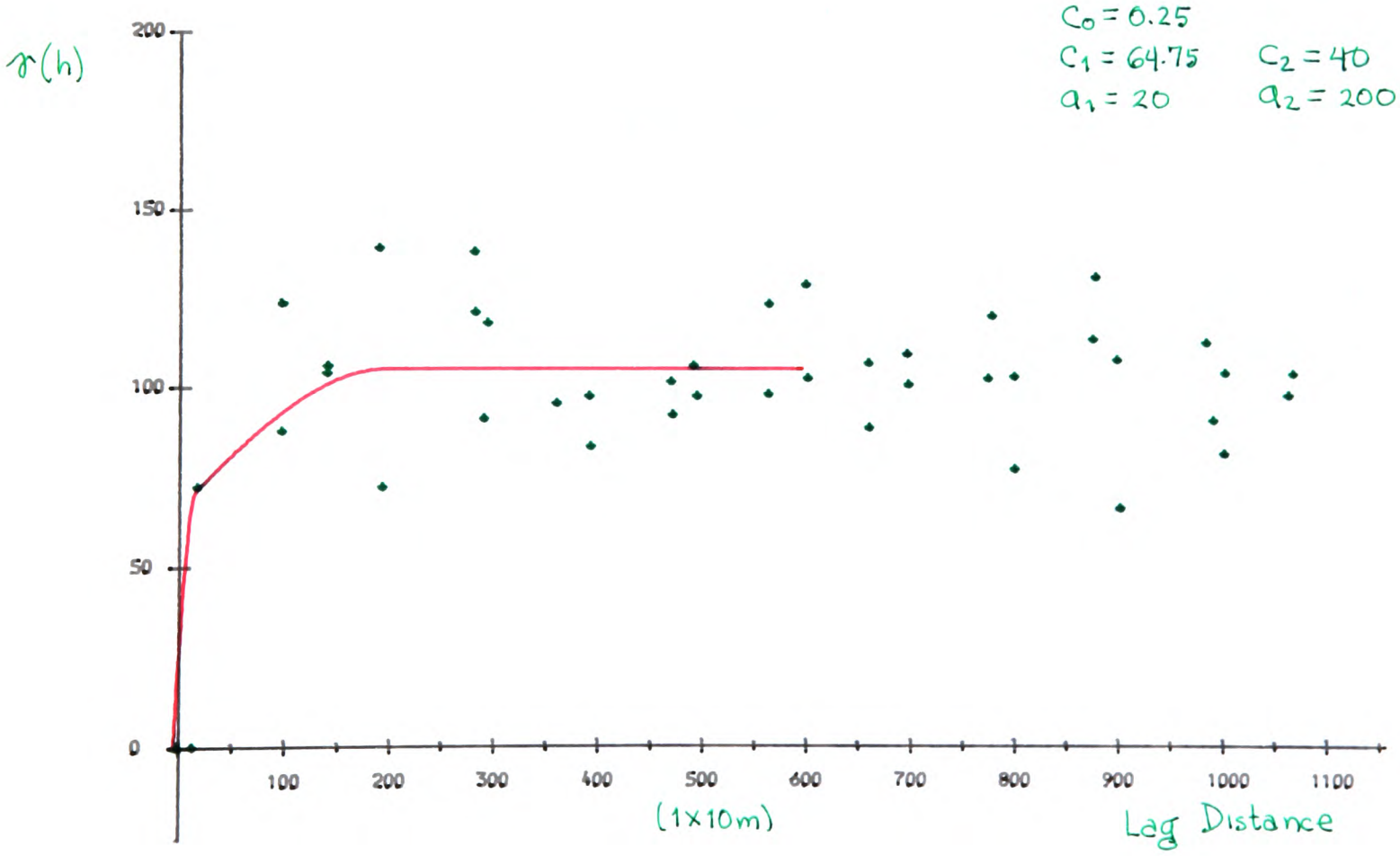


Fig 9.9 e

SEMI-VARIOGRAM of CLAY, 1 Km grid, (NESTED STRUCTURES)

$C_0 = 0.25$
$C_1 = 64.75$ $C_2 = 40$
$a_1 = 20$ $a_2 = 200$



property in the area. A simple spherical model was used to describe the semi-variances from the three grids. The range of this model is again the same as for the smallest grid (300m) and comparable to the inter-boundary spacings of mapping units as described above.

9.12.6 Silt.

The variation of Silt is anisotropic at a minimum sampling spacing of 50m, as shown in the semi-variogram of fig 9.10a. This appears to be characteristic of the soils in the area covered by the 50m sampling grid. Most soil properties in this study show anisotropic variation at this sampling scale.

The direction of anisotropy is again, as with the other soil properties, the direction NE-SW which is the same as the direction of the drift. This is shown by the set of points in the plot that depart from the rest and that lie almost horizontally.

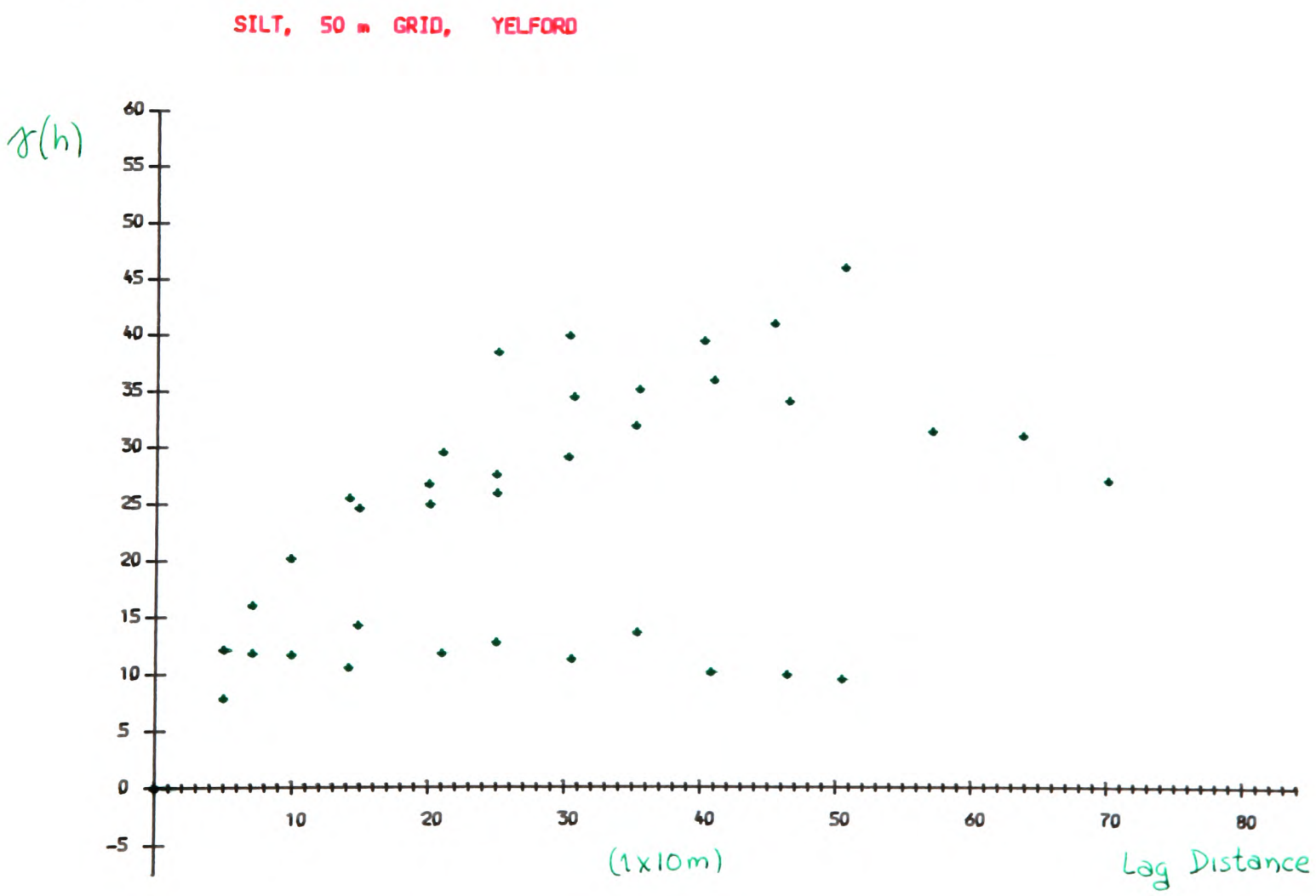
In contrast, the semivariogram for the 100m grid (fig 9.10b) showed a neat shape and clear spatial dependence along the transition within distances up to 450m defined by its range of influence. This distance can be said to reflect differences in soil type as represented by the frequency of soil boundaries within the area covered by this grid, and their average inter-boundary spacings.

The exponential model, found to fit the semi-variogram for the 200m grid (fig 9.10c), indicates a much more gradual change. This can be explained as the effect of the long and smooth transition characteristic of alluvial depositions. Most of the area covered by this grid is on the older alluvium of the lower terrace, and Silt is quite susceptible to alluvial transport. The long range of influence, (1250 m), corresponds also to much broader and spacious mapping units and lower frequency of soil boundaries. The semi-variogram of the nested grids (fig 9.10d) showed a very scattered set of points. The nested model fitted (double spherical) does not offer a satisfactory representation, but no other model can do better.

9.12.7 Coarse Sand.

The semi-variogram of coarse sand for the 50m grid (figure 9.11a) showed two interesting features. It is clearly anisotropic. The anisotropy is represented by the total absence

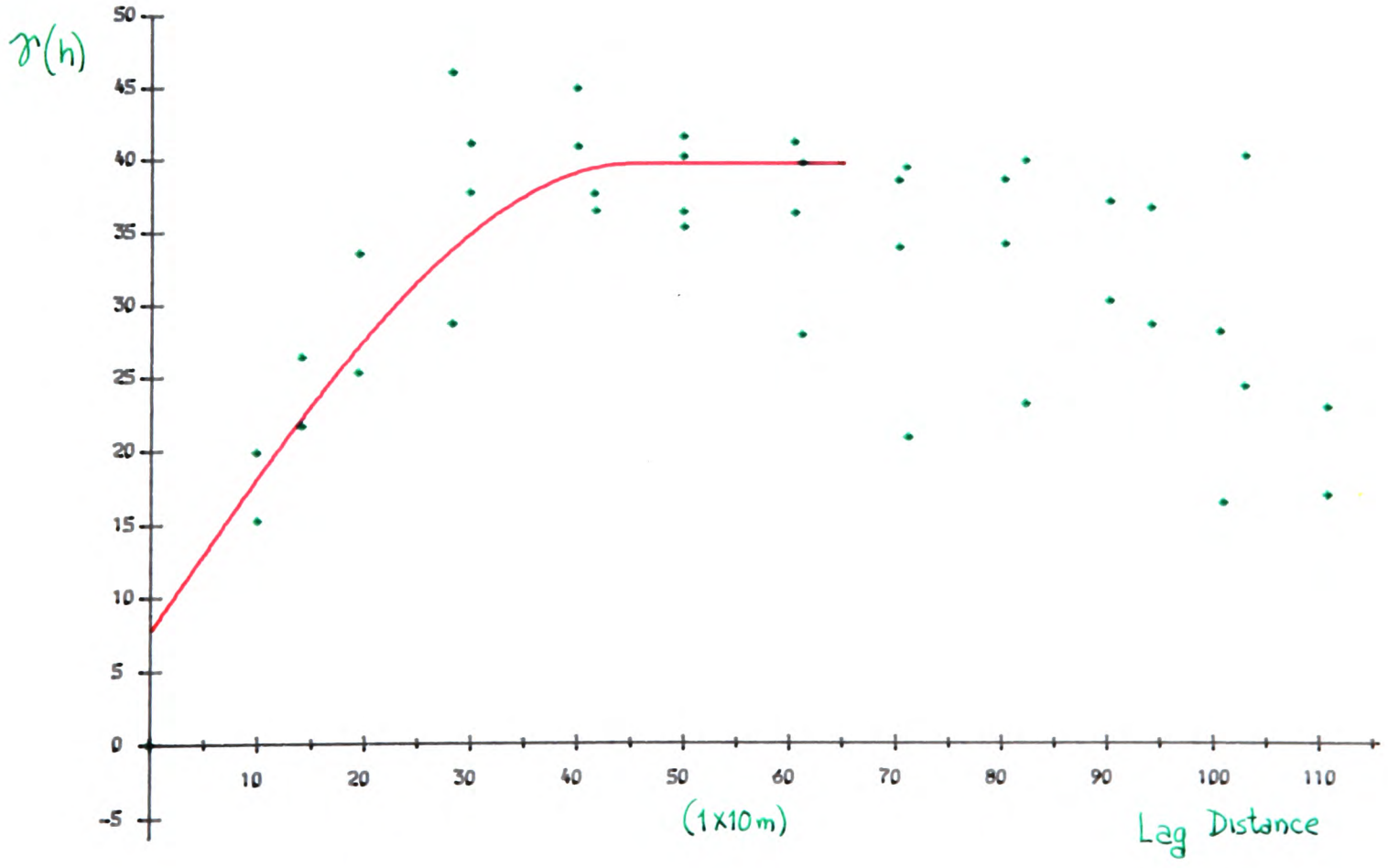
Fig 9.10 a


SILT, 50 m GRID, YELFORD
γ(h)
60
55
50
45
40
35
30
25
20
15
10
5
0
-5
10
20
30
40
50
60
70
80
(1x10m)
Lag Distance

$$\gamma(h) = 8 + 32\left[\frac{3}{2}\frac{h}{450} - \frac{1}{2}\left(\frac{h}{450}\right)^3\right]$$

Fig 9.10 b

SEMI-VARIOGRAM of SILT, 100 m grid, Yelford



$$\gamma(h) = 13 + 19\left[1 - \exp\left(\frac{-h}{1250}\right)\right]$$

Fig 9.10c

SEMI-VARIOGRAM of SILT, 200 m grid, Yelford

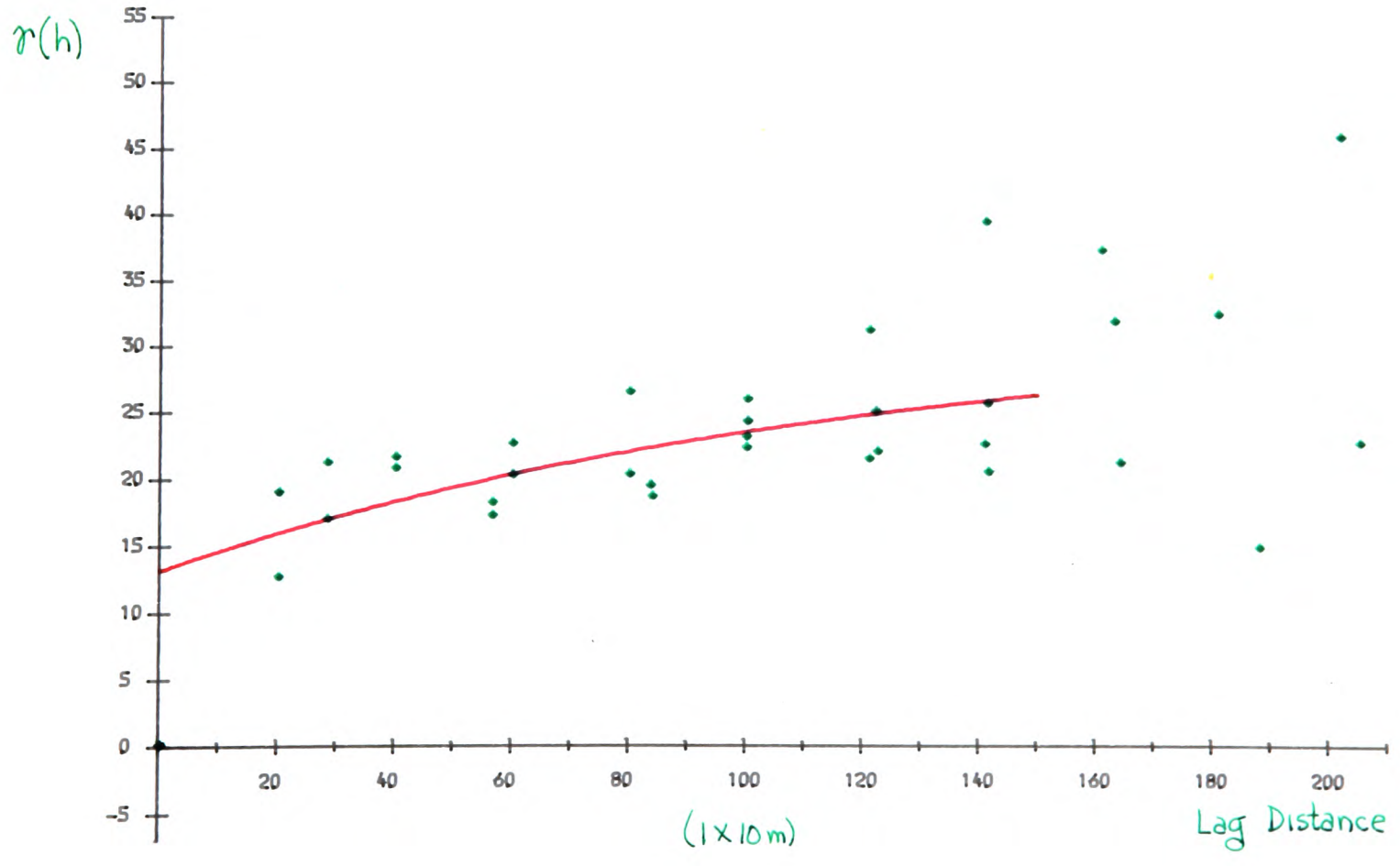
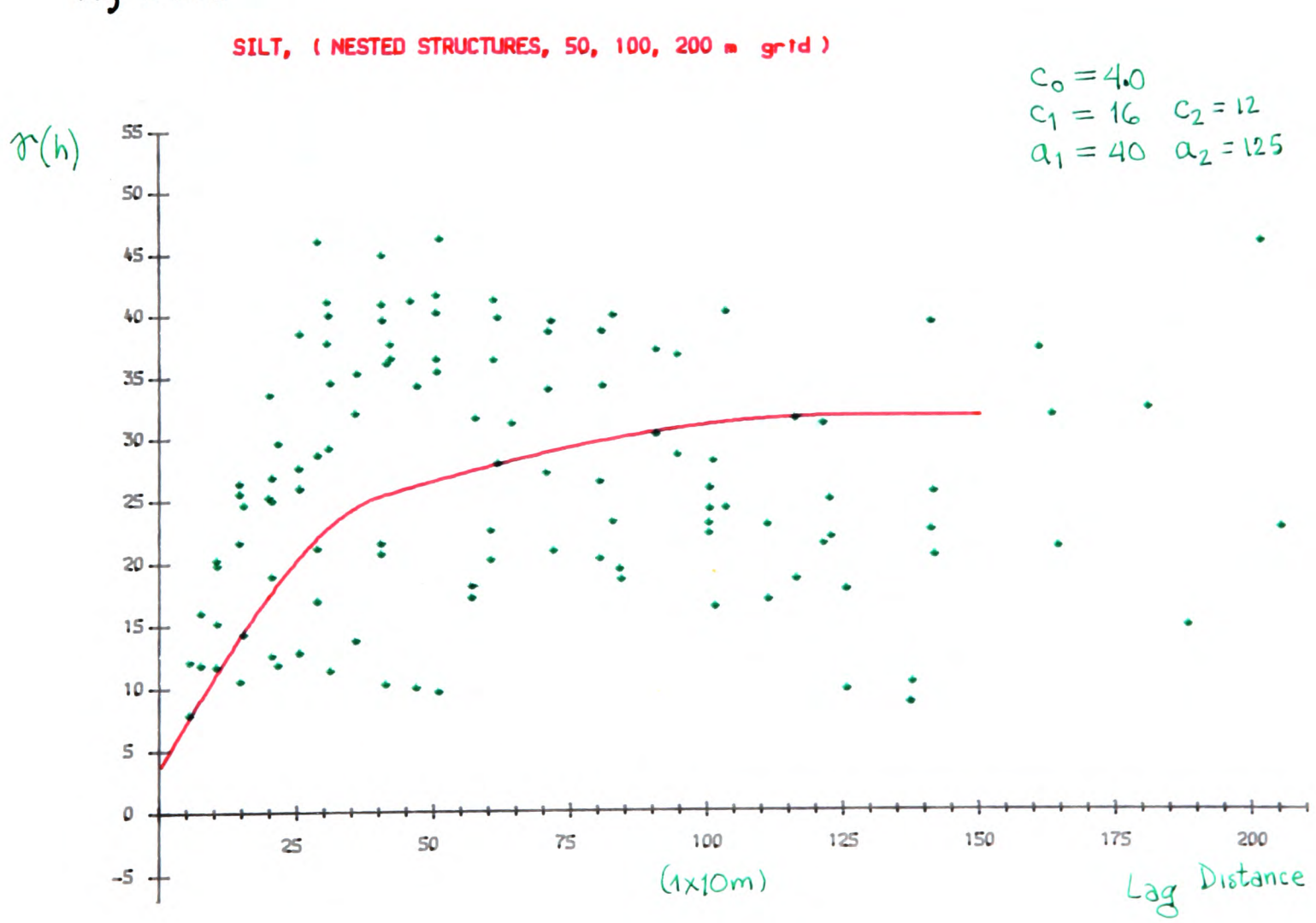


Fig 9.10 d


SILT, (NESTED STRUCTURES, 50, 100, 200 m grid)
$c_0 = 4.0$
$c_1 = 16$ $c_2 = 12$
$a_1 = 40$ $a_2 = 125$
$\gamma(h)$
55
50
45
40
35
30
25
20
15
10
5
0
-5
25
50
75
100
125
150
175
200
(1x10m)
Lag Distance

of spatial dependence in the NE-SW direction, the direction of the drift. The points representing the semi-variances in this direction lay flat and parallel to the axes of distance. In contrast, the semi-variograms for the other directions showed parabolic behaviour of the monotonic increase, indicating the presence of trends or changing drifts. So, evidence of anisotropic variation is combined with the presence of trends.

Most affected by alluvial deposition, the values of coarse sand were smoothed out in the direction of the drift, but in the other directions the presence of the trends is clear.

As with most of the other properties at this scale of sampling, the semi-variogram of coarse sand for the 200m grid (fig 9.11b) showed a good degree of spatial dependence in a transition along distances up to 325m (the range of influence). A good fit was obtained with a spherical model.

Coarse sand showed, on the whole, a pattern of variation at all scales which was quite similar to that of silt. Thus, an exponential model was also found to fit well the semi-variogram for the 200m grid (fig 9.11c) of this property. This type of semi-variogram indicates the long gradual transition characteristic of alluvial depositions of these materials, evident at this scale of sampling, and which can also be explained by the large area on the old alluvium covered by the 200m grid.

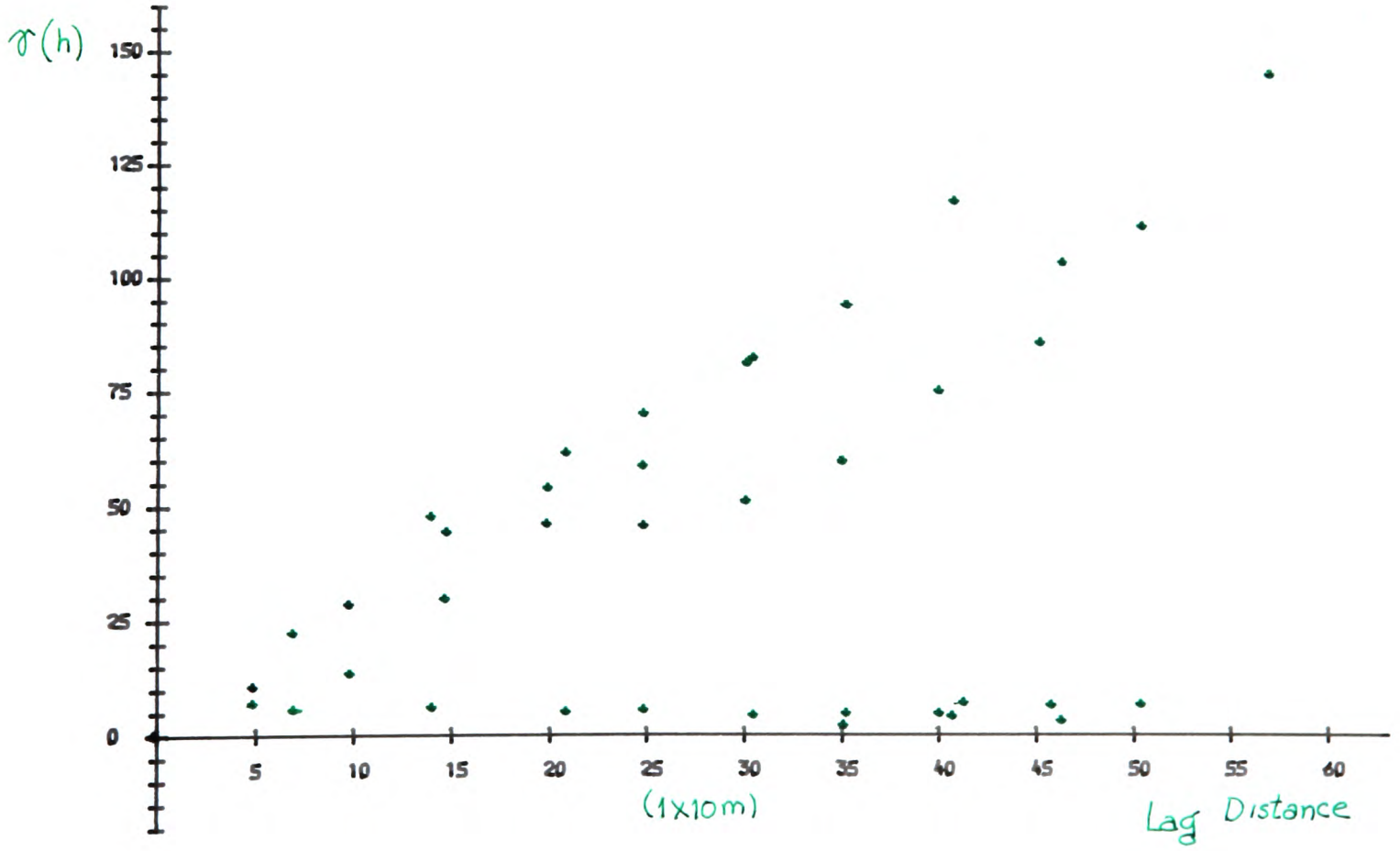
The anisotropic variation of this property and the nested structures, can be appreciated in the semi-variogram for the three grids in fig 9.11d. The different scales of variation are also evident in this plot. No model can be fitted to this scatter of points.

The first structure of the nested semi-variogram for the 1 km grid (fig 9.11e), corresponds to variation depicted by the semi-variograms of the three smaller sampling grids. The first range is comparable to the range of the 200m grid. The range of the second structure (5.5 km) gives the limit to spatial autocorrelation of this property in the area of study.

9.12.8 Organic Matter.

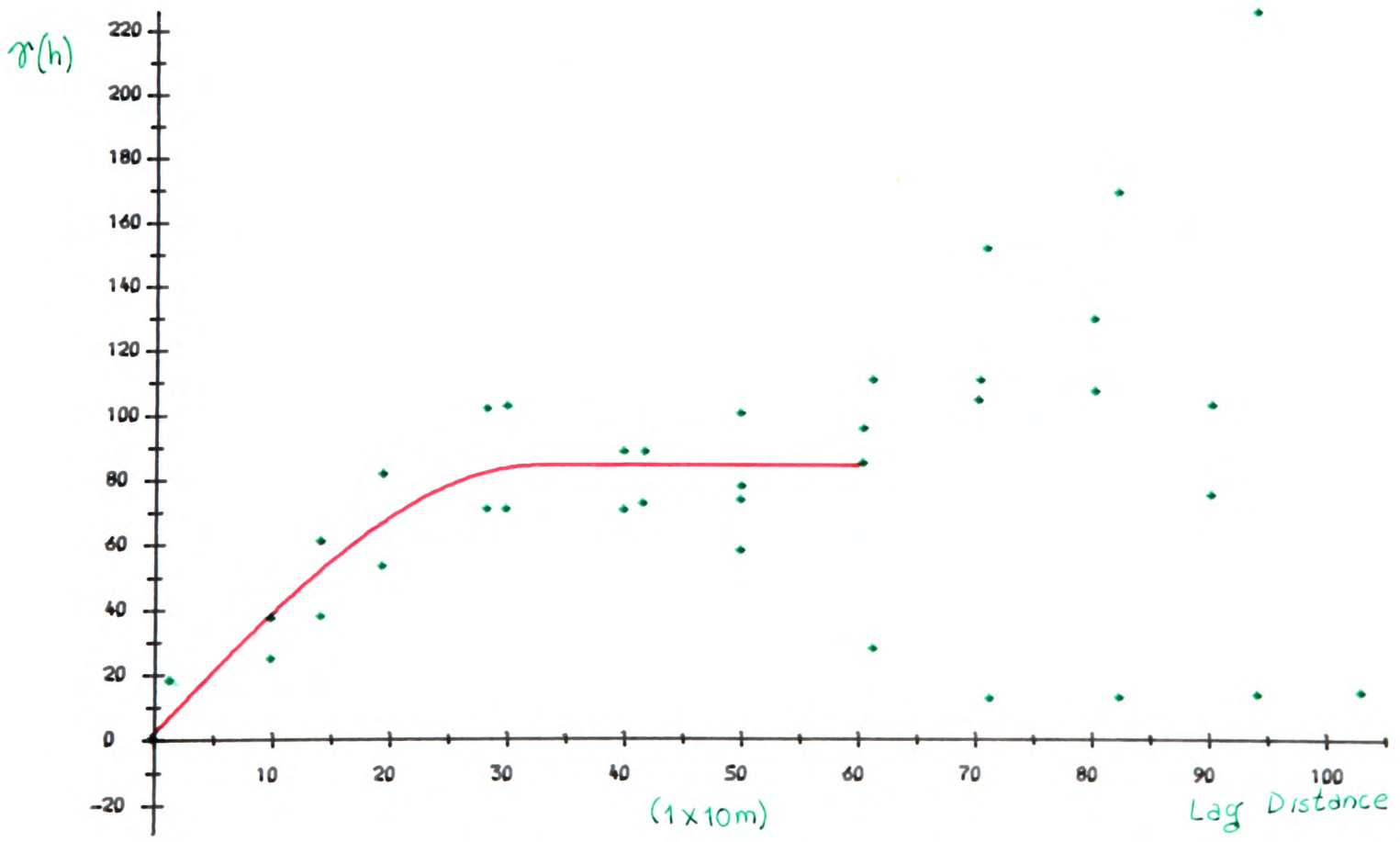
This property possesses a relatively high degree of spatial structure in the Yelford area at all scales of sampling. The semi-variograms calculated for all sampling grids were isotropic

Fig 9.11 a COARSE SAND, 50 m grid, Yelford



$$\gamma(h) = 2.5 + 82.5\left[\frac{3}{2}\frac{h}{325} - \frac{1}{2}\left(\frac{h}{325}\right)^3\right]$$

Fig 9.11b SEMI-VARIOGRAM of COARSE SAND, 100 m grid, Yelford



$$\gamma(h) = 30 + 120\left[1 - \exp\left(\frac{-h}{2000}\right)\right]$$

Fig 9.11c SEMI-VARIOGRAM of COARSE SAND, 200 m grid, Yelford

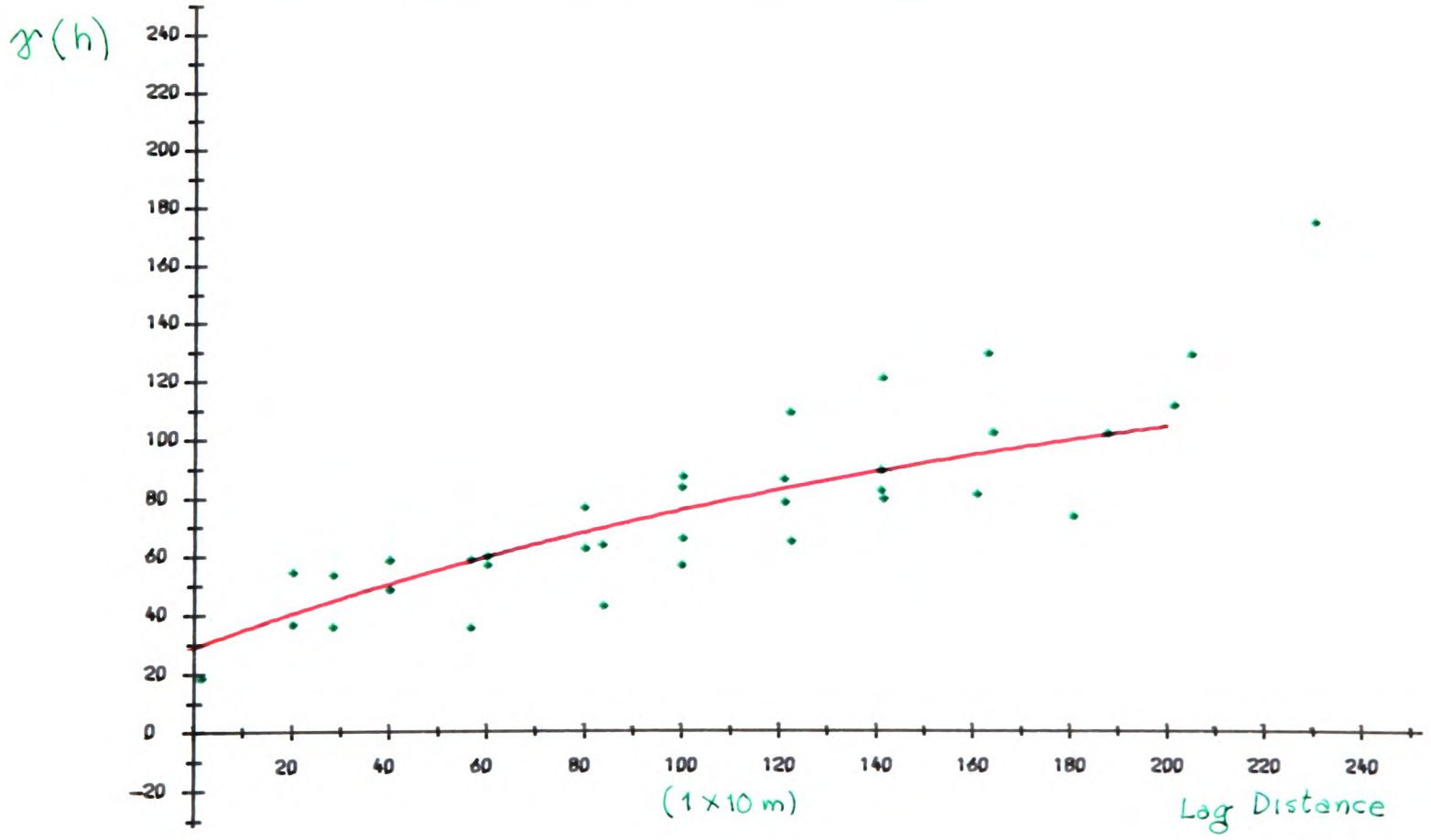
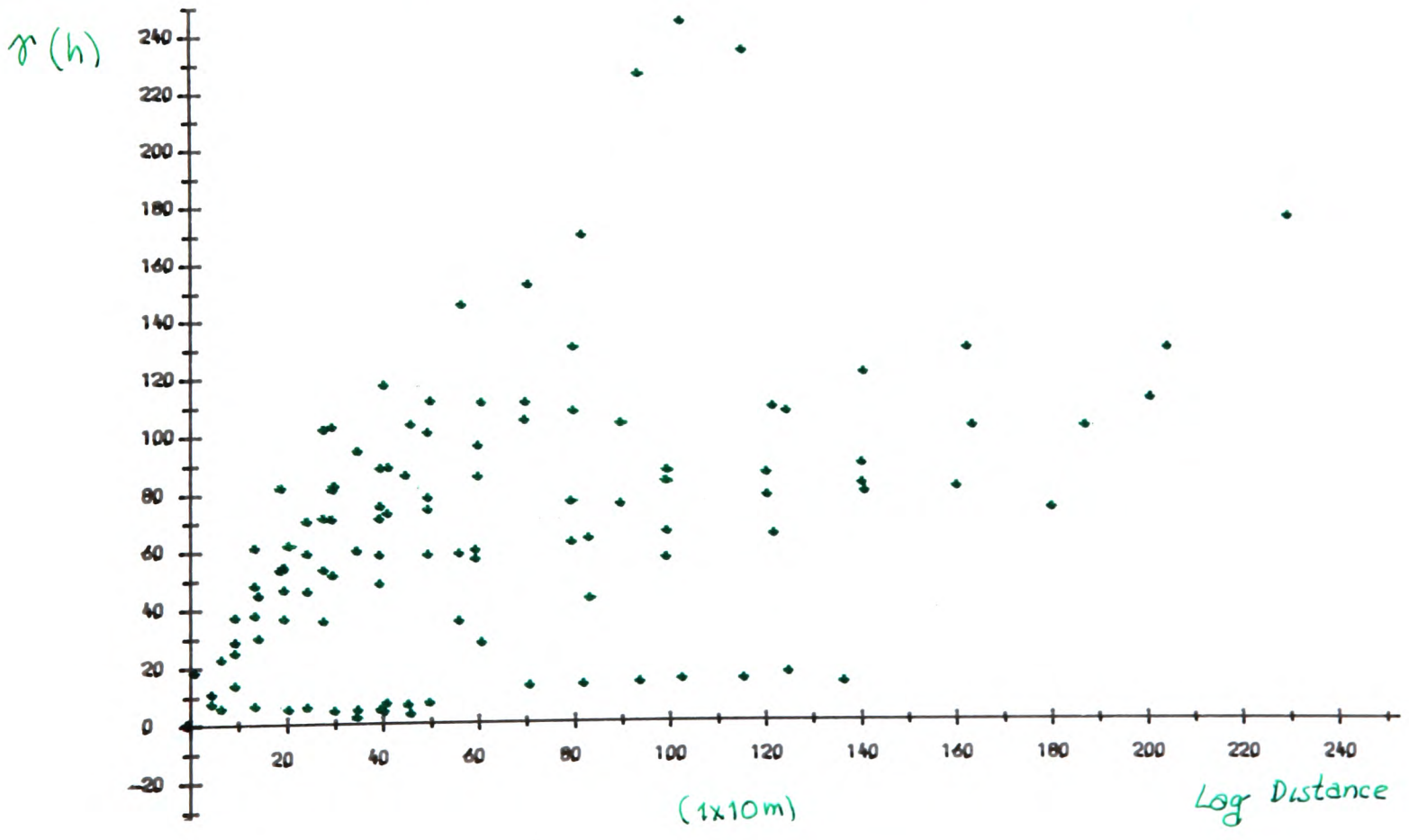
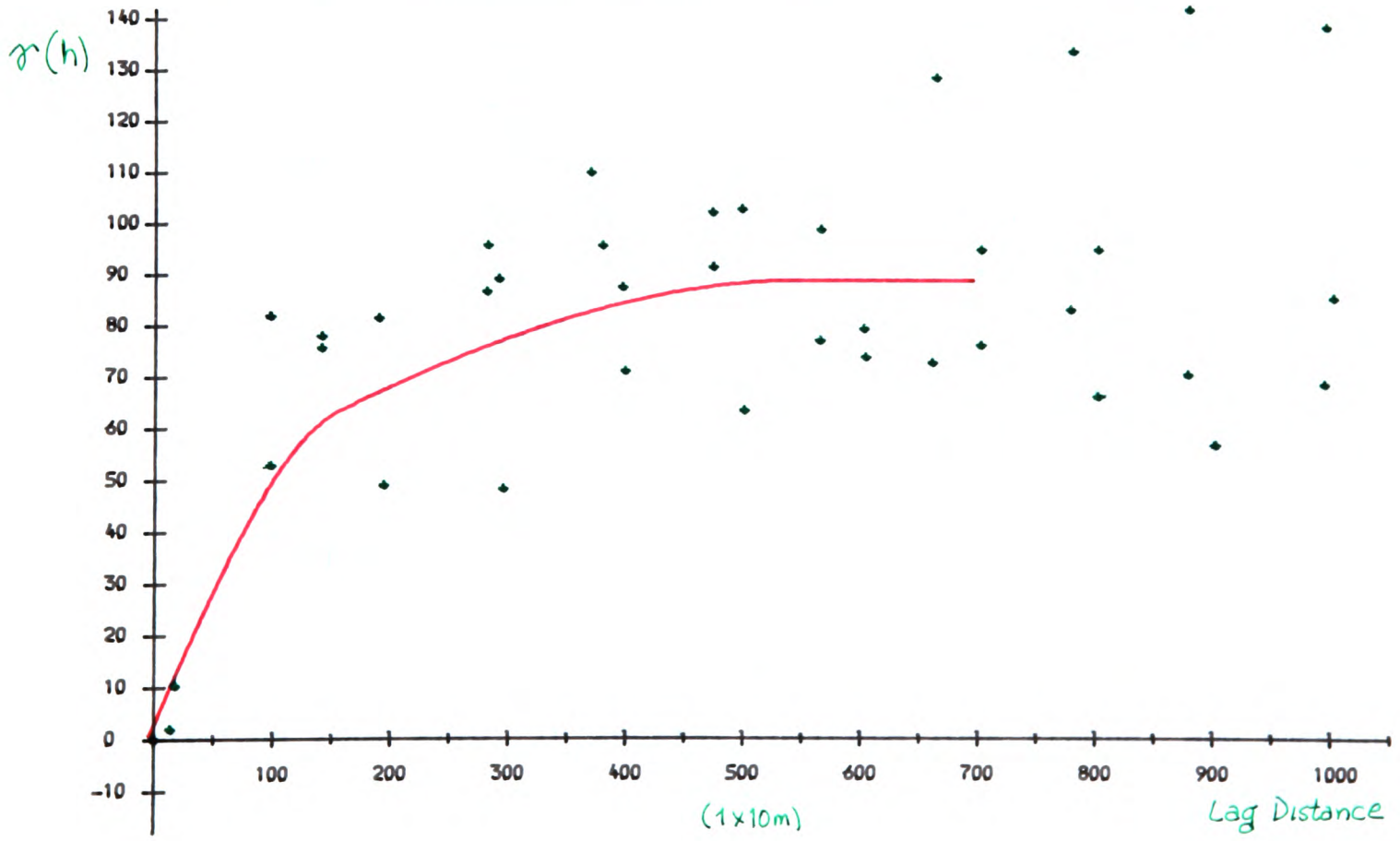


Fig 9.11d COARSE SAND, (NESTED STRUCTURES, 50, 100, 200 m grid)



$C_0 = 1.0$
$C_1 = 44$ $C_2 = 46$
$a_1 = 160$ $a_2 = 550$

Fig 9.11e COARSE SAND, 1 Km grid (NESTED STRUCTURES)



and presented clearly defined transitions and spatial dependence within equally clear ranges. The continuity in the transitional variation of this property within distances equal to the range (250m) in the 50m-grid semi-variogram, is demonstrated by the rather small nugget variance (fig 9.12a). The range of this semi-variogram corresponds roughly to the average inter-spacing of soil boundaries in the area covered by the grid.

Spherical models were found to fit well the semi-variograms for the 100m and 200m grid (figures 9.12b and 9.12c). In the former the fit is quite good, particularly in the region of the monotonic increase. Thus, the structure of the spatial variability of organic matter in these soils is not affected by drift and apparently nor by management. The ranges of the nested structures (fig 9.12d) are approximately the same as those for the two smallest grids, indicating the importance of variation at these sampling scales.

Some sort of spatial dependence in this property can be found within distances of 2 km, as given by the range of the second structure in the nested semi-variogram of fig 9.12e.

9.12.9 Concluding remarks about semi-variograms.

From the interpretation of the semi-variograms of the Yelford area some generalizations can be drawn.

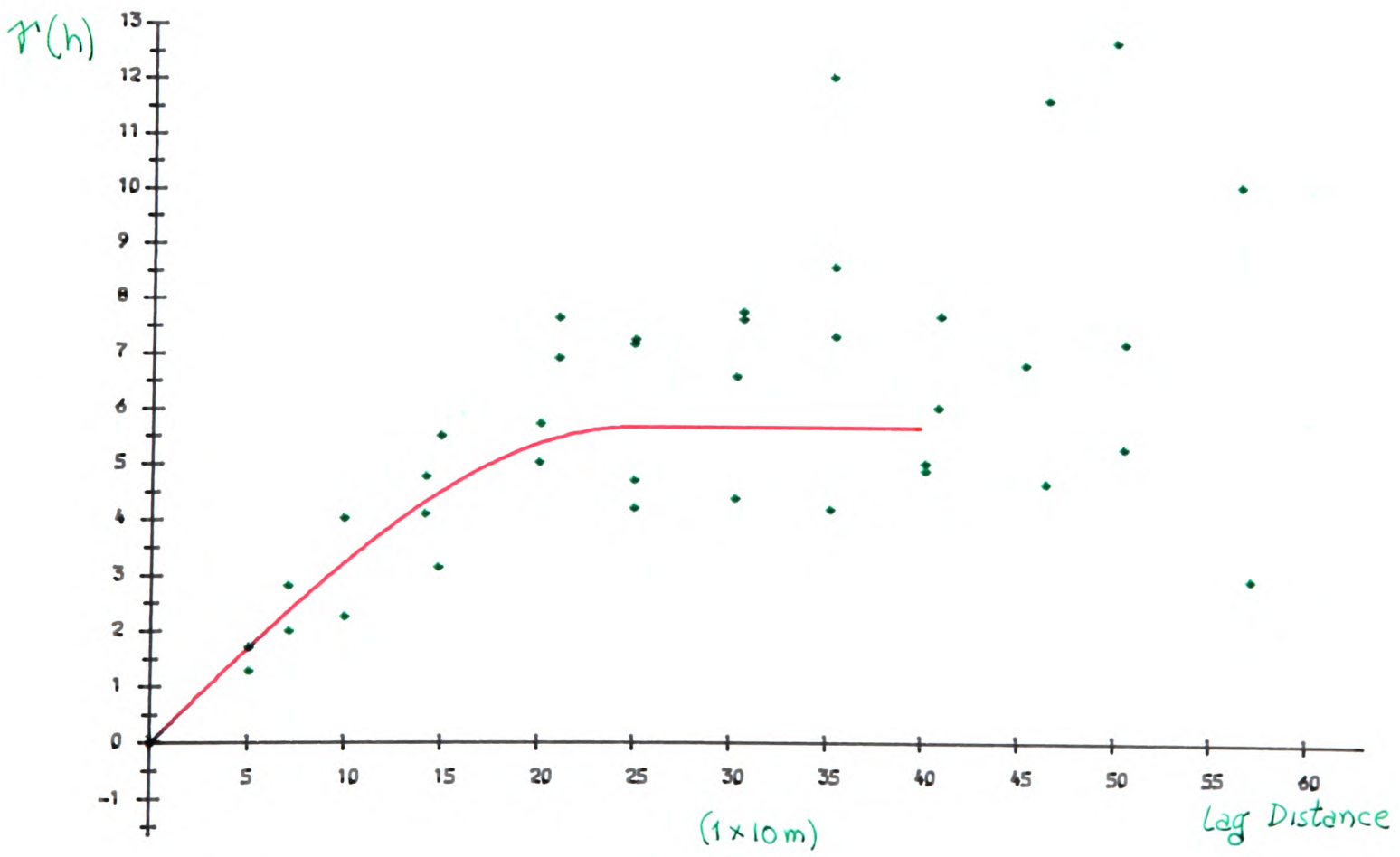
It was evident that the drift caused anisotropic variation in at least half of the properties studied in the direction NE-SW, from the terrace gravel and Oxford clay of Rickless hill down to the more plain lanscape of the older alluvium on the lower terrace. Anisotropic variation was more evident in the data from the 50m grid. This sampling grid covers particularly the hill slopes facing the SW direction, the direction of the drift. The presence of trends in the NE-SW direction are therefore explained by the landscape.

The exponential models found for semi-variograms calculated from the 200m grid, in general indicate long gradual and smooth changes as reflected on the lower frequency of soil boundaries and broader mapping units. Most of the 200m grid was laid on the older alluvium of the lower terrace.

$$\gamma(h) = 0.01 + 5.73\left[\frac{3}{2}\,\frac{h}{250} - \frac{1}{2}\left(\frac{h}{250}\right)^3\right]$$

Fig 9.12a

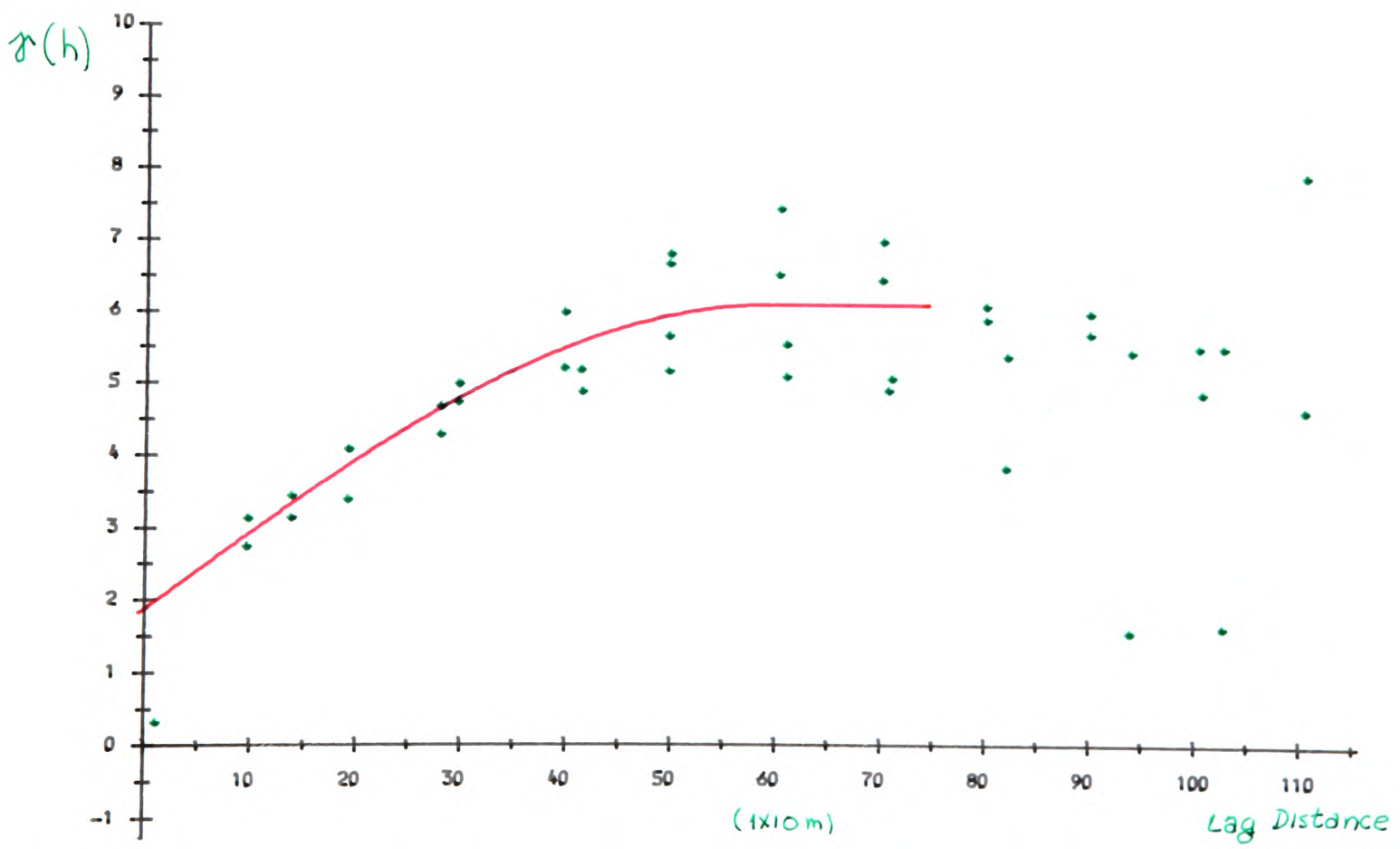
SEMI-VARIOGRAM of ORGANIC MATTER, 50 m grid, Yelford



$$\gamma(h) = 1.8 + 4.3\left[\frac{3}{2}\,\frac{h}{600} - \frac{1}{2}\left(\frac{h}{600}\right)^3\right]$$

Fig 9.12 b

SEMI-VARIOGRAM of ORGANIC MATTER, 100m grid, Yelford



$$\gamma(h) = 1.1 + 4.3\left[\frac{3}{2}\,\frac{h}{900} - \frac{1}{2}\left(\frac{h}{900}\right)^3\right]$$

Fig 9.12 c

SEMI-VARIOGRAM of ORGANIC MATTER, 200 m grid, Yelford

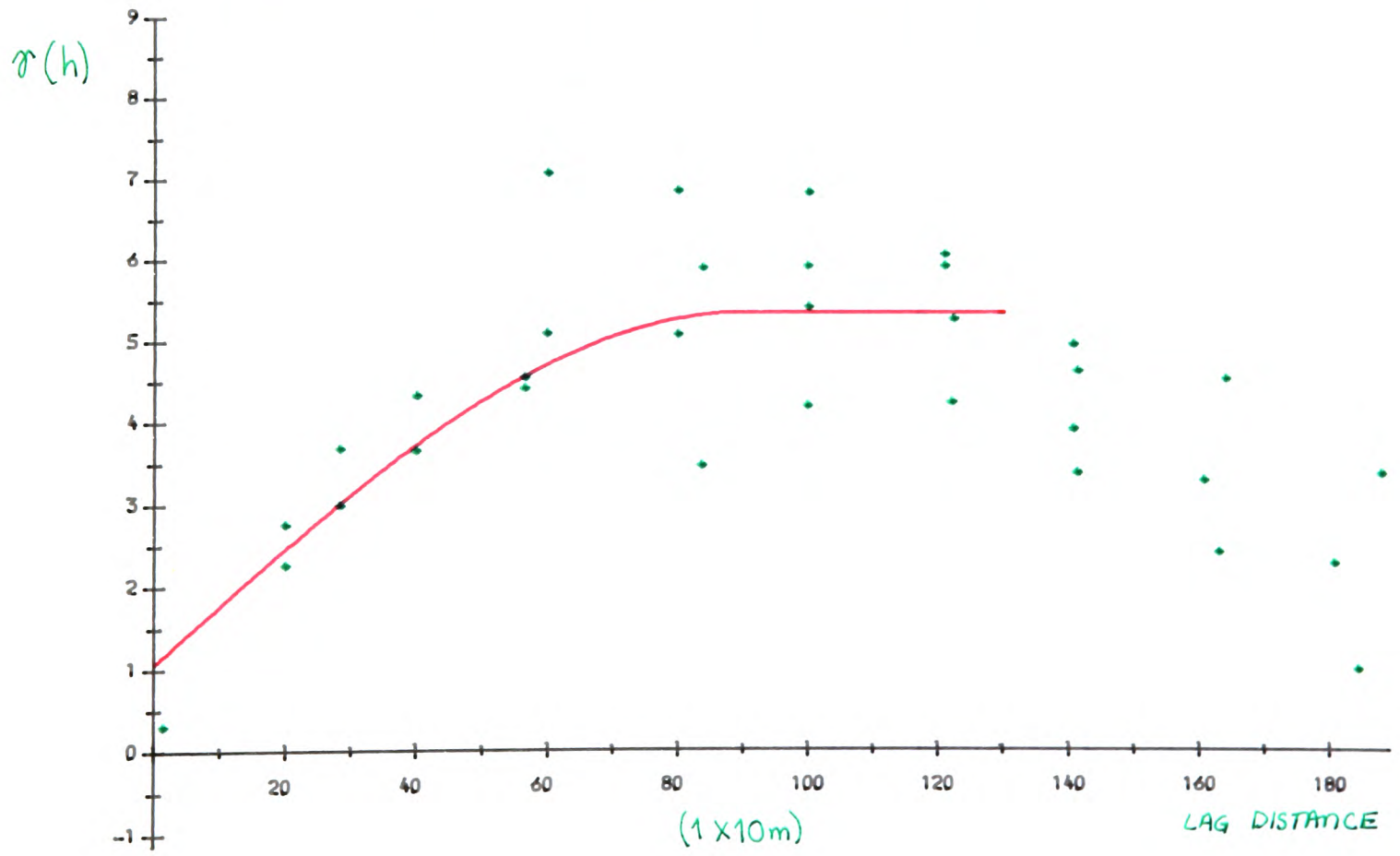


Fig. 9.12 d

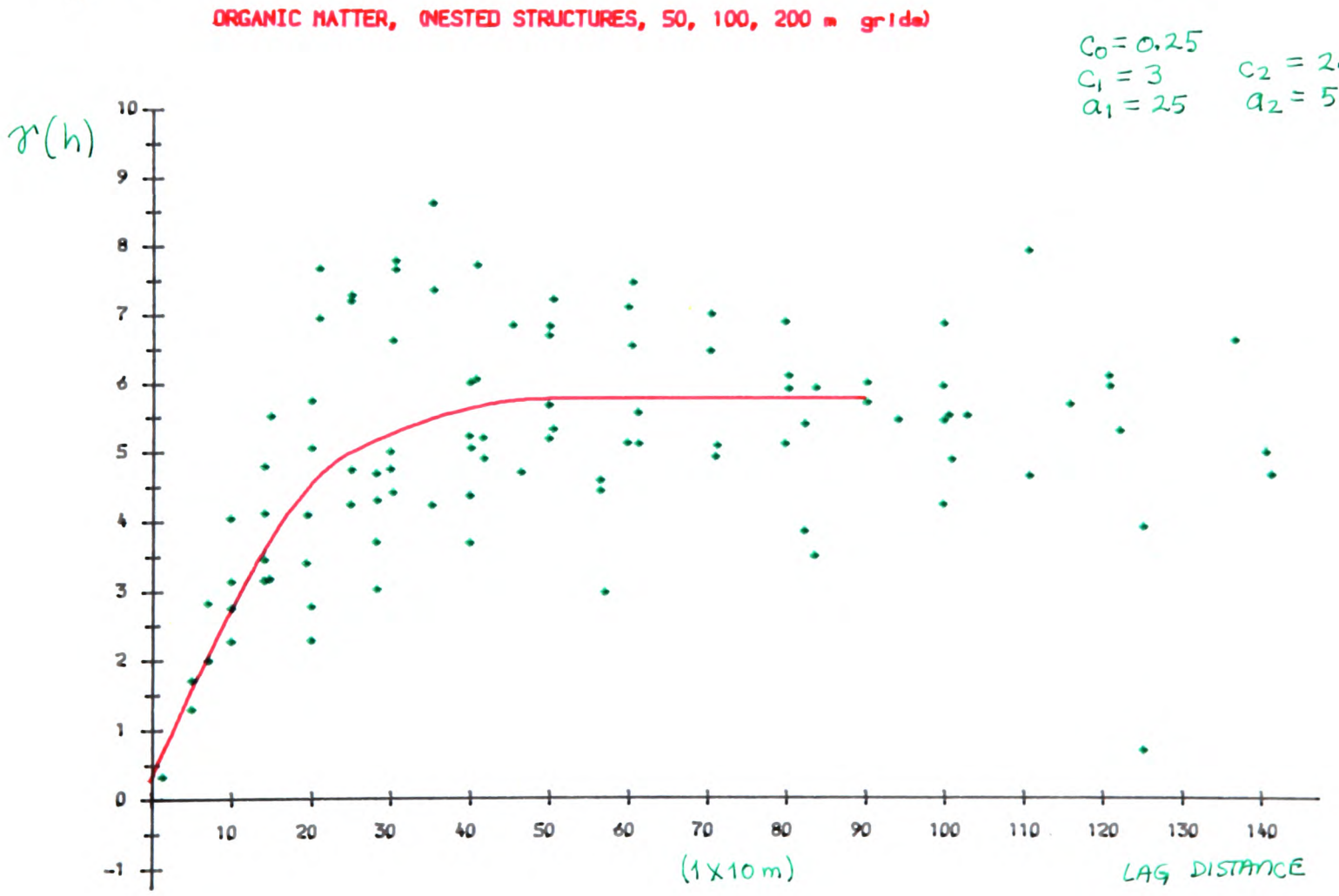
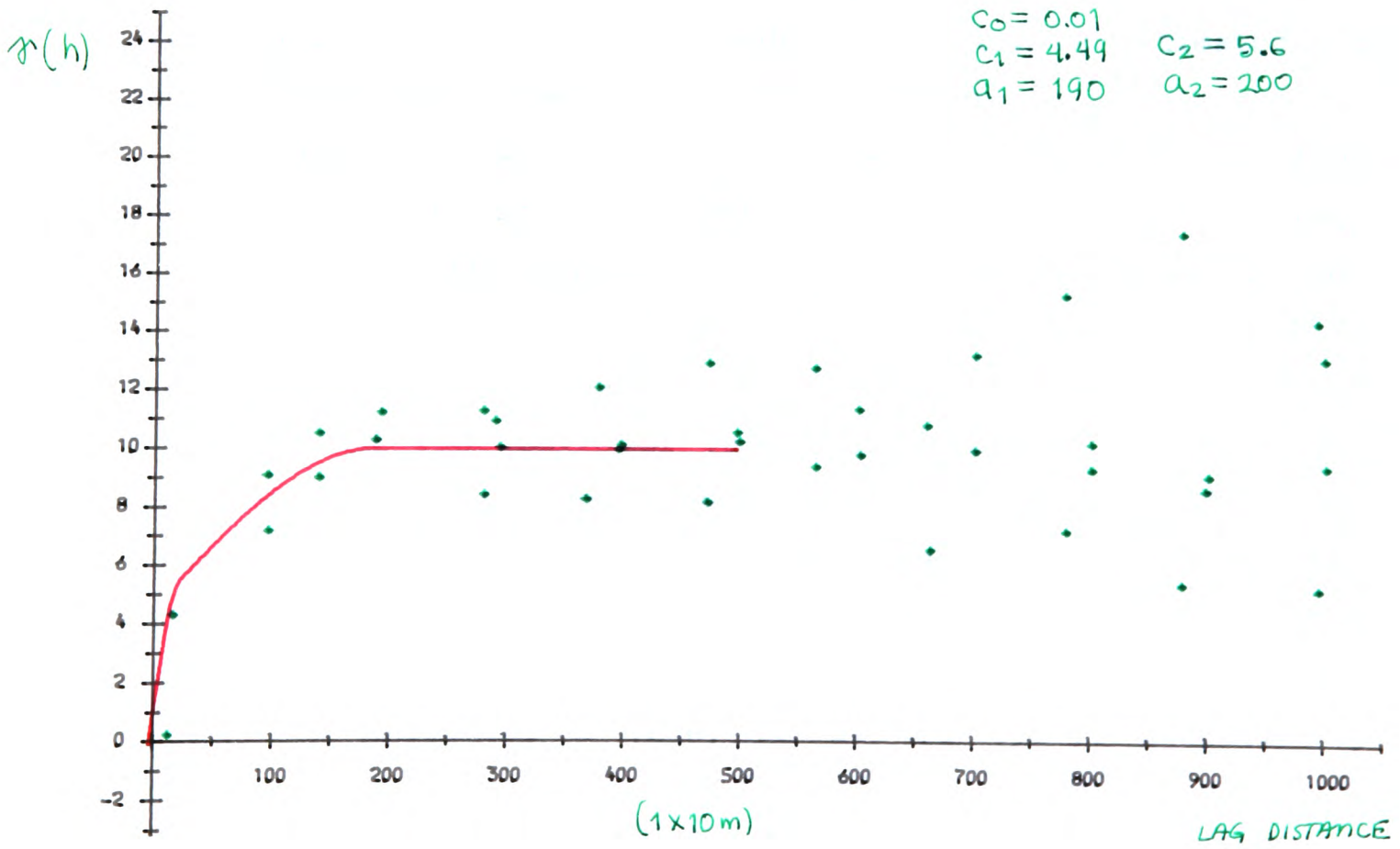

ORGANIC MATTER, (NESTED STRUCTURES, 50, 100, 200 m grids)
$C_0 = 0.25$
$C_1 = 3$
$a_1 = 25$
$C_2 = 2.5$
$a_2 = 50$
$\gamma(h)$
10
9
8
7
6
5
4
3
2
1
0
-1
10
20
30
40
50
60
70
80
90
100
110
120
130
140
(1 X 10 m)
LAG DISTANCE

Fig 9.12e

SEMI-VARIOGRAMS of ORGANIC MATTER, 1 Km grid, ISP30 SHEET



9.13 Fitting models for anisotropy.

Silt, Coarse Sand, Carbonates and Potassium vary anisotropically over the area covered by the 50m grid. The interpretation of the Yelford semi-variograms indicated that the anisotropies were induced by trends in the area which can be clearly explained from the landscape. They are part of the natural spatial distribution of soils in the area.

Nevertheless, attempts were made to fit models for anisotropy so that isotropic transformed semi-variograms, as in equation 7.46, could be obtained. The directional semi-variograms of any anisotropic property can be written as a single model. Thus recalling eq. (7.46):

$$\gamma(h,\theta) = Co + u(\theta)[h]$$

where θ is the direction of the vector [h] relative to the X-axis, and $u(\theta_i) = m_i$ is the slope of the semi-variogram in the direction θ_i. So, if a convenient estimator of the term $u(\theta)$ can be found, the single model (7.46) can be estimated and used as an isotropic model for interpolation.

Two functions were used in trials with the Silt data from the 50m grid, to estimate $u(\theta)$: An elliptical function (eq. 7.45) suggested by Burgess (1979), and a similar function (eq. 7.47) suggested by David (1977). The parameters A, B and ϕ in these equations were estimated by fitting the curves (7.45) and (7.47) to the four pairs of values of direction (radians) and slope of the known directional semi-variograms of Silt. The fitting was achieved by non-linear least squares using the Gauss-Newton iterative method. The directions were measured from the X-axis of the grid, so the direction N-S is at 0 (0°). The direction 2 (E-W) is at T/2 (90°), direction 3 (NW-SE) is at T /4 (45°), and direction 4 (NE-SW) is at 3T /4 (135°). Table 9.8 presents the four pairs of directions and slopes measured from the directional semi-variograms of Silt that are shown in figure 9.13.

The fitting of each curve proved to be lengthy (about 15 min CPU time = aprox. 1h 15 min real time), and expensive. The fitting of both equations (7.45 and 7.47) gave:

Fig. 9.13

MODEL FOR ANISOTROPY, SILT 50 m GRID, YELFORD
(DIRECTIONAL SEMI-VARIOGRAMS)

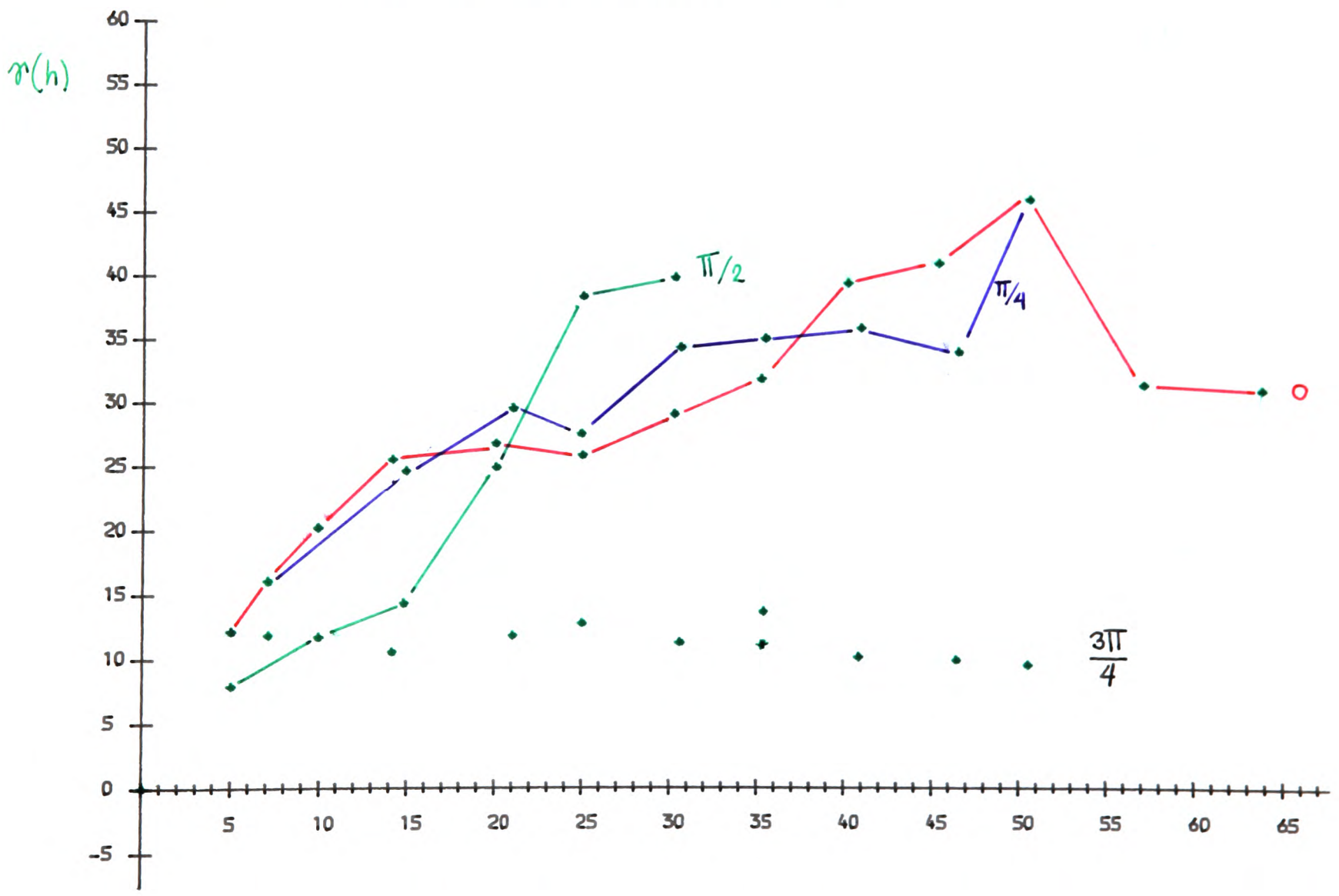


Table 9.8 Directions and slopes of the four semi-variograms of Silt at the 50 m grid, Yelford

Direction		θ_i	m_i
1	(N-S)	0	0.616
2	(E-W)	$\pi/2$	1.246
3	(NW-SE)	$\pi/4$	0.494
4	(NE-SW)	$3\pi/4$	0.050

Table 9.9 Fitted elliptical models and estimated values of $W(\theta_i)$ ($W(\theta_i) = m_i$) for the semi-variograms of silt at 50 m grid, Yelford

MODEL 1* = $W(\theta) = 0.2161 \cos^2(\theta-2.834) + 0.986 \sin^2(\theta-2.834)$

Direction	θ_i	m_i	$W(\theta_i)$
1	0	0.616	0.2867
2	$\pi/2$	1.246	0.8238
3	$\pi/4$	0.494	0.9161
4	$3\pi/4$	0.05	0.3788

MODEL 2**: $W(\theta) = [(0.95)^2 \cos^2(\theta-1.065) + (0.000164)^2 \sin^2(\theta - 1.065)]^{\frac{1}{2}}$

1	0	0.616	0.4603
2	$\pi/2$	1.246	0.8313
3	$\pi/4$	0.494	0.9132
4	$3\pi/4$	0.05	0.2621

* Burgess (1979)
** David (1977)

$$w(\theta) = 0.2161 \cos^2(\theta-2.834) + 0.986 \sin^2(\theta-2.834) \qquad (9.1)$$

for the elliptical function (Burgess, 1979), and:

$$w(\theta) = \{(0.95)^2\cos^2(\theta-1.06) + (0.00016)^2\sin^2(\theta-1.06)\}^{1/2} \qquad (9.2)$$

for the function by David (1977).

Table 9.9 shows the slopes estimated by $w(\theta)$ for the four directions. These are compared with the values of m_i, i=1,2, 3,4. It can be seen that the fit of equation (9.2) is marginally better than the fit of (9.1). However, none of the fits obtained was considered to be sufficiently good to proceed any further with modelling anisotropies and using the models for interpolation. The grounds for such decision are two. First, to mathematically "force" isotropy this way is to ignore the natural distribution of soils in the area (Yelford), their anisotropic tendency induced by drift and trends, and their relation to the landscape. Second, it is statistically unsound to fit a model which requires the estimation of 3 parameters (A, B, and ϕ), to only four points. Semi-variograms for 8 directions were calculated for the same sampling grid to increase the number of observations for the fitting to 8. However, this did not improve the fit of any of the equations either. Computer printouts with results for the fittings with 4 and 8 observations are included in appendix 9.2.

Faced with such evidence, the present approach to deal with anisotropy was abandoned in favour of exploring procedures that account for the trends.

9.14 Predicting values of soil properties from soil survey output.

The procedures that use soil survey output for prediction of soil properties have been discussed in section 4.9 (chapter 4). They concern the use of:

i. Means of soil classes.

ii. The typical profile of the mapping unit.

iii. Means of mapping units.

Notice that in 4.9.2 (chapter 4) the typical profile of the class is used to extract values as predictors. For these trials the typical profiles of the mapping units were used instead (in ii. above), since the retrospective identification of typical profiles of classes proved difficult.

The values of the 8 soil properties in the trial area were predicted at the check points of each corresponding grid by procedures i., ii., and iii., listed above. The means were calculated from the grid points available pertaining to each class or map unit. Records of the class identified for each grid point were available, and the mapping unit to which each point pertains could be read from the map. Within each mapping unit a "typical" profile was identified and records of the 8 properties extracted.

9.14.1 Class means as predictors.

It should be noted that the class was recorded for the grid points as well as for the check points. Hence the class at the check points did not have to be predicted as it would have been necessary in ordinary circumstances. This gives class means as predictors certain advantage over other procedures since the possibility of error involved in predicting the class was ommitted. Thus, calculating the class means involved, if anything, a laborious file-manipulating exercise. The "ALLOCATION" facility was used to allocate data blocks (sites) to their corresponding class as indicated by the records available. Then, means and other descriptive statistics for each class were calculated using the "STATISTICS" facility. Appendix 9.3 lists the soil class to which individual points in each grid were allocated. Then, mean values of the 8 soil properties within each class were calculated for each sampling grid.

Thus, the soil class for each of the check points had been determined (tables 9.3a to c). Using this information, the values of the 8 soil properties at the check points are predicted by applying to them the corresponding class mean. The predicted values of individual properties are shown in tables

9.10a-c. The tables containing the observed or measured values of the properties studied at all the random check points considered, are shown in table 9.11. The values in this table are the "ground truth", and serve as a yardstick aginst which the predicted values by all the procedures involved in this exercise were compared.

9.14.2 Values of the typical profiles within mapping units as predictors.

Following steps iii. to v. in 4.9.1 (chapter 4), profiles considered typical within the mapping units were selected. Apart from the ordinary, the selection presented three type of extra situations:

i. In some instances there were two sustantial delineations of the same mapping unit. The position in this case was to take the mean from the typical profiles selected for those delineations.

ii. Two profiles could be selected as representative of a probable range within the unit. The position in this case was also to take the mean from the two profiles representing the range.

iii. The mapping unit consisted of only a small delineation within the area corresponding to the sampling grid. In this case the typical profile was taken from this small delineation but this fact was noted as it may bear some weight in the accuracy of the prediction.

The values of the 8 properties were then predicted by extracting them from the typical profile, or from the average of typical profiles from delineations (as in i. or ii. above), and applying them to the check points laying within the same mapping unit. The list of mapping units with their corresponding typical profiles is presented in table 9.12. The values extracted from the typical profiles to serve as predictors for the 8 soil properties are listed in table 9.13. The predicted values at the check points for the three sampling grids are given in the tables 9.14a-c.

Table 9.10a. Values of soil properties estimated by CLASS MEANS at the random check points in the 50 m grid at Yelford.

SITE No	CLAY	K	OM	pH	CO3	Mg	SILT	SAND
95004	47.1600	1.1500	7.1800	7.320	5.250	1.3500	23.8300	12.5000
95024	38.0800	.3830	6.7900	6.530	1.940	1.3700	28.5000	4.2500
95027	31.8000	.9200	5.4800	7.180	9.480	1.0600	13.2000	34.0000
95028	44.5500	.5880	7.8600	6.870	.822	1.9300	30.8800	4.1100
95029	38.0800	.3830	6.7900	6.530	1.940	1.3700	28.5000	4.2500
95030	44.5500	.5880	7.8600	6.870	.822	1.9300	30.8800	4.1100
95031	35.1700	.5000	7.0400	6.570	.682	1.3200	28.2900	5.9400
95032	35.1700	.5000	7.0400	6.570	.682	1.3200	28.2900	5.9400
95033	44.1100	.7660	7.5000	7.210	3.880	1.3100	25.7700	9.6600
95034	36.8500	.3740	8.2000	6.490	.548	1.7000	28.5500	3.0000
95035	36.8500	.3740	8.2000	6.490	.548	1.7000	28.5500	3.0000
95036	36.8500	.3740	8.2000	6.490	.548	1.7000	28.5500	3.0000
95037	36.8500	.3740	8.2000	6.490	.548	1.7000	28.5500	3.0000
95047	36.8500	.3740	8.2000	6.490	.548	1.7000	28.5500	3.0000
95051	31.8000	.9200	5.4800	7.180	9.480	1.0600	13.2000	34.0000
95052	44.1100	.7660	7.5000	7.210	3.880	1.3100	25.7700	9.6600

Table 9.10b. Values of soil properties estimated by CLASS MEANS at the random check points in the 100 m grid at Yelford.

SITE No	CLAY	K	OM	pH	CO3	Mg	SILT	SAND
95002	38.7900	.4790	6.9700	6.600	.780	1.7300	27.2400	4.9300
95003	29.2500	.5870	4.3200	7.240	9.630	.8300	11.2500	39.3500
95004	39.6600	1.0860	6.6100	7.170	3.680	1.6200	24.8000	10.4600
95005	37.4100	.3940	7.4500	6.630	.620	1.5300	28.4700	3.8200
95006	32.2500	.3750	8.2200	7.130	1.220	.9800	25.7500	7.0000
95012	36.8300	.7250	6.1800	6.640	.920	1.6300	24.1600	5.5100
95013	36.8300	.7250	6.1800	6.640	.920	1.6300	24.1600	5.5100
95014	36.8300	.7250	6.1800	6.640	.920	1.6300	24.1600	5.5100
95015	36.8300	.7250	6.1800	6.640	.920	1.6300	24.1600	5.5100
95016	39.6600	1.0860	6.6100	7.170	3.680	1.6200	24.8000	10.4600
95017	39.6600	1.0860	6.6100	7.170	3.680	1.6200	24.8000	10.4600
95018	36.8300	.7250	6.1800	6.640	.920	1.6200	24.1600	5.5100
95019	46.3000	.6840	7.5200	6.810	.850	1.9600	29.6900	4.8400
95020	46.3000	.6840	7.5200	6.810	.850	1.9600	29.6900	4.8400
95021	38.7900	.4790	6.9700	6.600	.780	1.7300	27.2400	4.9300
95023	32.2500	.3750	8.2200	7.130	1.220	.9800	25.7500	7.0000
95024	38.7900	.4790	6.9700	6.600	.780	1.7300	27.2400	4.9300
95027	29.2500	.5870	4.3200	7.240	9.630	.8300	11.2500	39.3500
95028	46.3000	.6840	7.5200	6.810	.850	1.9600	29.6900	4.8400
95029	38.7900	.4790	6.9700	6.600	.780	1.7300	27.2400	4.9300
95030	46.3000	.6840	7.5200	6.810	.850	1.9600	29.6900	4.8400
95031	36.8300	.7250	6.1800	6.640	.920	1.6200	24.1600	5.5100
95032	36.8300	.7250	6.1800	6.640	.920	1.6200	24.1600	5.5100
95033	32.2500	.3750	8.2200	7.130	1.220	.9800	25.7500	7.0000
95034	37.4100	.3940	7.4500	6.630	.620	1.5300	28.4700	3.8200
95035	37.4100	.3940	7.4500	6.630	.620	1.5300	28.4700	3.8200
95036	37.4100	.3940	7.4500	6.630	.620	1.5300	28.4700	3.8200
95037	37.4100	.3940	7.4500	6.630	.620	1.5300	28.4700	3.8200
95041	42.4200	.7280	6.0700	7.230	3.260	1.4200	24.0700	10.4200
95042	32.2500	.3750	8.2200	7.130	1.220	.9800	25.7500	7.0000
95043	46.3000	.6840	7.5200	6.810	.850	1.9600	29.6900	4.8400
95044	42.4200	.7280	6.0700	7.230	3.260	1.4200	24.0700	10.4200
95045	37.4100	.3940	7.4500	6.630	.620	1.5300	28.4700	3.8200
95046	46.3000	.6840	7.5200	6.810	.850	1.9600	29.6900	4.8400
95047	37.4100	.3940	7.4500	6.630	.620	1.5300	28.4700	3.8200
95048	39.6600	1.0860	6.6100	7.170	3.620	1.6200	24.8000	10.4600
95049	39.6600	1.0860	6.6100	7.170	3.620	1.6200	24.8000	10.4600
95051	29.2500	.5870	4.3200	7.240	9.630	.8300	11.2500	39.3500
95052	42.4200	.7280	6.0700	7.230	3.260	1.4200	24.0700	10.4200

Table 9.10c. Values of soil properties estimated by class means at the random check points in the 200 m grid at Yelford.

SITE No	CLAY	K	OM	pH	CO3	Mg	SILT	SAND
95001	40.2500	.5000	6.1400	6.530	.780	1.5500	25.0000	5.0800
95002	40.2500	.5000	6.1400	6.530	.780	1.5500	25.0000	5.0800
95003	27.4000	.6400	4.2800	7.340	10.660	.9200	13.8000	32.6000
95004	36.8300	.6700	5.7000	7.270	4.870	1.1600	22.0000	15.9100
95005	37.9500	.4090	6.8900	6.560	.780	1.5800	26.2200	2.8600
95006	32.2500	.3750	8.2200	7.130	1.220	.9800	25.7500	7.0000
95007	55.0000	.6400	5.3400	6.790	.850	1.8200	20.2300	3.0000
95008	37.9500	.4090	6.8900	6.560	.780	1.5800	26.2200	2.8600
95009	37.2500	.4090	6.8900	6.560	.780	1.5800	26.2200	2.8600
95010	44.5000	.6600	5.0500	7.190	3.150	1.4800	22.8300	6.5000
95011	36.8300	.6700	5.7000	7.270	4.870	1.1600	22.0000	15.9100
95012	33.8000	.7600	6.6200	6.730	1.270	1.6800	26.8000	5.6000
95013	33.8000	.7600	6.6200	6.730	1.270	1.6800	26.8000	5.6000
95014	33.8000	.7600	6.6200	6.730	1.270	1.6800	26.8000	5.6000
95015	33.8000	.7600	6.6200	6.730	1.270	1.6800	26.8000	5.6000
95016	36.8300	.6700	5.7000	7.270	4.870	1.1600	22.0000	15.9100
95017	36.8300	.6700	5.7000	7.270	4.870	1.1600	22.0000	15.9100
95018	33.8000	.7600	6.6200	6.730	1.270	1.6800	26.8000	5.6000
95019	40.0000	.5160	7.8100	6.990	.950	1.7100	29.1600	5.3300
95020	40.0000	.5160	7.8100	6.990	.950	1.7100	29.1600	5.3300
95021	40.2500	.5000	6.1400	6.530	.780	1.5500	25.0000	5.0800
95022	40.2500	.5000	6.1400	6.530	.780	1.5500	25.0000	5.0800
95023	32.2500	.3750	8.2200	7.130	1.220	.9800	25.7500	7.0000
95024	40.2500	.5000	6.1400	6.530	.780	1.5500	25.0000	5.0800
95025	36.8300	.6700	5.7000	7.270	4.870	1.1600	22.0000	15.9100
95026	40.2500	.5000	6.1400	6.530	.780	1.5500	25.0000	5.0800
95027	27.4000	.6400	4.2800	7.340	10.660	.9200	13.8000	32.6000
95028	40.0000	.5160	7.8100	6.990	.950	1.7100	29.1600	5.3300
95029	40.2500	.5000	6.1400	6.530	.780	1.5500	25.0000	5.0800
95030	40.0000	.5160	7.8100	6.990	.950	1.7100	29.1600	5.3300
95031	33.8000	.7600	6.6200	6.730	1.270	1.6800	26.8000	5.6000
95032	33.8000	.7600	6.6200	6.730	1.270	1.6800	26.8000	5.6000
95033	32.2500	.3750	8.2200	7.130	1.220	.9800	25.7500	7.0000
95034	37.9500	.4090	6.8900	6.560	.780	1.5800	26.2200	2.8600
95035	37.9500	.4090	6.8900	6.560	.780	1.5800	26.2200	2.8600
95036	37.9500	.4090	6.8900	6.560	.780	1.5800	26.2200	2.8600
95037	37.9500	.4090	6.8900	6.560	.780	1.5800	26.2200	2.8600
95038	37.9500	.4090	6.8900	6.560	.780	1.5800	26.2200	2.8600
95039	56.0000	.9000	5.2000	7.280	1.200	1.5000	21.0000	8.0000
95040	40.0000	.5160	7.8100	6.990	.950	1.7100	29.1600	5.3300
95041	44.5000	.6600	5.0500	7.190	3.150	1.4800	22.8300	6.5000
95042	32.2500	.3750	8.2200	7.130	1.220	.9800	25.7500	7.0000
95043	40.0000	.5160	7.8100	6.990	.950	1.7100	29.1600	5.3300
95044	44.5000	.6600	5.0500	7.190	3.150	1.4800	22.8300	6.5000
95045	37.9500	.4090	6.8900	6.560	.780	1.5800	26.2200	2.8600
95046	40.0000	.5160	7.8100	6.990	.950	1.7100	29.1600	5.3300
95047	37.9500	.4090	6.8900	6.560	.780	1.5800	26.2200	6.5000
95048	36.8300	.6700	5.7000	7.270	4.870	1.1600	22.0000	15.9100
95049	36.8300	.6700	5.7000	7.270	4.870	1.1600	22.0000	15.9100
95050	27.4000	.6400	4.2800	7.340	10.660	.9200	13.8000	32.6000
95051	27.4000	.6400	4.2800	7.340	10.660	.9200	13.8000	32.6000
95052	44.5000	.6600	5.0500	7.190	3.150	1.4800	22.8300	6.5000

Table 9.11 . Observed or measured values of soil properties at the random check points in the area of the trials.

SITE	NORTHING	EASTING	SN	CLAY	K	OM	pH	CO3	Mg	SILT	SAND
A1	20563	43584	95001	46.0000	.6000	5.000	7.04	.400	1.2000	27.000	5.0000
A2	20540	43503	95002	36.0000	.9000	7.9000	6.670	.400	2.3000	32.0000	6.0000
A3	20533	43571	95003	29.0000	.5000	3.4000	7.310	5.500	.7000	10.0000	43.0000
A4	20511	43547	95004	47.0000	1.0000	8.4000	7.380	2.400	1.4000	32.0000	8.0000
A5	20502	43512	95005	34.0000	.4000	12.6000	5.750	.000	1.6000	34.0000	6.0000
A6	20466	43572	95006	25.0000	.3000	13.7000	6.620	.200	1.6000	32.0000	6.0000
A7	20446	43558	95007	65.0000	.6000	5.2000	6.330	.600	1.9000	20.0000	2.0000
A8	20415	43481	95008	45.0000	.4000	5.0000	5.980	.000	1.8000	25.0000	2.0000
A9	20397	43501	95009	55.0000	.3000	6.4000	6.580	.300	2.2000	24.0000	2.0000
A10	20382	43572	95010	33.0000	.4000	3.6000	7.290	3.700	1.0000	17.0000	15.0000
A12	20372	43563	95011	31.0000	.5000	4.2000	7.210	5.100	1.0000	19.0000	19.0000
B1	20544	43572	95012	34.0000	.4000	8.0000	5.740	.400	1.9000	26.0000	9.0000
B2	20531	43504	95013	30.0000	1.2000	5.9000	6.970	1.500	1.5000	24.0000	8.0000
B3	20521	43495	95014	28.0000	.6000	2.2000	5.500	.100	2.0000	29.0000	3.0000
B4	20532	43544	95015	51.0000	.5000	6.6000	7.100	.100	1.6000	29.0000	7.0000
B5	20532	43590	95016	46.0000	1.0000	5.6000	7.140	3.300	1.2000	22.0000	15.0000
B6	20524	43539	95017	48.0000	.8000	4.9000	7.090	2.200	.8000	24.0000	11.0000
B7	20508	43491	95018	24.0000	.2000	10.0000	6.670	2.100	2.1000	30.0000	4.0000
B8	20500	43579	95019	50.0000	1.5000	9.2000	5.090	.100	3.8000	30.0000	5.0000
B9	20491	43582	95020	53.0000	.8000	9.1000	4.530	.300	4.2000	27.0000	5.0000
B10	20461	43509	95021	33.0000	.2000	4.8000	5.580	.400	1.1000	26.0000	3.0000
B11	20475	43572	95022	30.0000	.2000	9.8000	5.940	.100	1.3000	33.0000	6.0000
B12	20455	43576	95023	45.0000	.5000	5.7000	7.080	.600	1.2000	23.0000	6.0000
B13	20513	43542	95024	49.0000	.7000	7.3000	7.330	.400	1.3000	31.0000	6.0000
B14	20501	43559	95025	52.0000	.9000	7.2000	7.240	3.700	1.3000	32.0000	8.0000
B15	20481	43538	95026	37.0000	.2000	3.8000	6.730	.300	1.1000	25.0000	6.0000
C1	20519	43560	95027	28.0000	.7000	3.8000	7.080	18.000	.6000	10.0000	44.0000
C2	20502	43538	95028	48.0000	.6000	8.3000	6.800	.200	2.0000	33.0000	3.0000
C3	20500	43542	95029	41.0000	.5000	6.1000	7.240	1.000	1.6000	30.0000	8.0000
C5	20494	43548	95030	41.0000	1.1000	12.3000	6.750	.600	2.8000	35.0000	5.0000
C6	20491	43538	95031	33.0000	.3000	4.9000	6.080	.200	.9000	28.0000	8.0000
C9	20487	43558	95032	33.0000	.3000	13.0000	6.040	.000	1.7000	37.0000	6.0000
C11	20468	43557	95033	27.0000	.3000	11.6000	6.780	.800	1.0000	29.0000	13.0000
C12	20468	43561	95034	31.0000	.5000	14.9000	6.020	.200	3.4000	33.0000	3.0000
C13	20462	43544	95035	54.0000	.4000	5.3000	6.260	.300	1.7000	25.0000	2.0000
C14	20461	43547	95036	54.0000	.4000	5.4000	5.980	.100	1.9000	23.0000	2.0000
C16	20458	43560	95037	59.0000	.5000	6.3000	6.530	.400	1.7000	25.0000	2.0000
E1	20418	43468	95038	45.0000	.4000	5.0000	6.230	.000	1.9000	26.0000	2.0000
E2	20567	43605	95039	43.0000	.7000	4.0000	7.330	25.000	1.5000	23.0000	23.0000
E3	20567	43578	95040	49.0000	.5000	4.5000	7.170	.500	1.3000	30.0000	5.0000
E4	20513	43575	95041	49.0000	1.1000	6.4000	7.390	6.100	1.5000	22.0000	16.0000
E5	20453	43572	95042	44.0000	.4000	5.9000	6.900	1.300	1.0000	26.0000	5.0000
E6	20560	43552	95043	49.0000	.6000	5.0000	6.840	.200	1.3000	32.0000	4.0000
E7	20450	43588	95044	26.0000	.3000	9.0000	6.780	.400	1.0000	30.0000	8.0000
E8	20528	43524	95045	42.0000	1.0000	9.6000	6.680	.600	2.1000	31.0000	8.0000
E9	20553	43541	95046	54.0000	.4000	5.4000	6.780	.000	2.0000	28.0000	5.0000
E10	20454	43544	95047	55.0000	.6000	5.4000	6.870	.400	2.1000	27.0000	1.0000
E11	20515	43567	95048	34.0000	.7000	4.3000	7.330	7.500	.8000	13.0000	33.0000
E12	20500	43563	95049	52.0000	.8000	8.1000	7.220	1.600	1.8000	34.0000	5.0000
E13	20515	43561	95050	33.0000	.4000	4.0000	7.250	8.900	.7000	1.0000	34.0000
E14	20511	43560	95051	32.0000	1.3000	7.7000	7.040	8.900	.8000	19.0000	26.0000
E15	20503	43545	95052	49.0000	1.1000	7.9000	7.030	1.400	1.5000	32.0000	6.0000

Table 9.12 - Selected typical profiles within mapping units in the three sampling grids at Yelford

Sampling Grid	Map Unit	Typical Profile
50m	BCT 4/5	15045
	BCg-3(2)	65015
	BEg-3(2)	65060/(65014)
	NSG-2	65032
	NGG1-2	65080/(65097)
	NGG2-2	65101
100m	BCg-3(2)	15043//55015 (15013)
	BCg-3/4	(45015)
	BEg-3(2)	15069//15066 (55007)
	NSG-2	15010//45020
	CGG-2	15101
	NGG1-2	15104 (15051)
	NGG2-2	15106
200m	BCT-4/5	55046
	BCg-3	55042
	BCg-3(2)	55015
	BCT-4/5	(55011)
	BEg-3(2)	55013/55020
	NSG-2	(55019)
	NGG1-2	55040//55028 (55022)
	NGG2-2	55033

A//B = Two substantial delineations. The mean was taken.

X/Y = Representative of probable range. The mean was taken.

( ) = Small delineation. If no other major delineation exists, this is taken.

Table 9.13 - Values of Soil Properties extracted from typical profiles within mapping units and used as predictors

Sampling Grid	Mapping Unit	PREDICTED VALUES FROM TYPICAL PROFILES pH	K	Mg	Carbonate	Clay	Silt	C.Sand	O.M.
50m	BCT-4/5	7.11	0.7	0.9	10.5	28.0	10.0	41.0	4.7
	BCg-3(2)	7.25	0.6	0.8	7.30	45.0	19.0	23.0	4.7
	BEg-3/(2)	6.85	0.95	1.0	1.05	41.5	25.0	9.0	5.7
	NSG-2	6.86	0.4	2.20	0.7	38.0	37.0	4.0	8.4
	NGG1-2	6.67	0.4	1.7	0.7	36.0	27.0	3.5	7.5
	NGG2-2	6.88	0.5	1.7	0.4	60.0	24.0	2.0	5.8
100m	BCg-3(2)	7.22	1.0	1.05	2.45	44.0	25.5	11.5	5.9
	BCg-3/4	7.44	0.8	1.7	2.1	48.0	27.0	8.0	7.6
	BEg-3(2)	6.63	0.45	1.55	2.75	39.5	23.5	5.0	6.3
	NSG-2	7.07	0.7	1.75	0.85	44.0	27.5	5.5	8.7
	CGG-2	6.79	0.4	1.30	0.4	42.0	29.0	6.0	6.1
	NGG1-2	6.81	0.2	1.3	0.6	35.0	28.0	2.0	4.9
	NGG2-2	7.26	0.6	2.1	0.6	66.0	21.0	1.0	6.6
200m	BCT-4/5	7.4	0.5	0.8	8.7	26.0	14.0	29.0	3.6
	BCg-3	7.34	0.7	1.0	22.0	27.0	16.0	28.0	4.2
	BCg-3(2)	7.32	1.0	1.3	3.6	46.0	28.0	12.0	6.6
	BEg-3(2)	6.8	0.35	8.5	0.6	27.5	26.5	5.5	4.3
	NSG-2	7.13	0.8	2.7	1.1	35.0	33.0	5.0	12.3
	NGG1-2	6.02	0.25	1.4	0.8	33.0	30.5	2.0	7.55
	NGG2-2	6.95	0.4	1.6	0.8	51.0	21.0	2.0	5.0

Table 9.14a. Values of soil properties estimated by values from TYPICAL PROFILE in the map unit at the random check points in the 50 m grid at Velford.

SITE	NORTH	EAST	SN	MAP UNIT	pH	K	Mg	CO3	CLAY	SILT	SAND	OM
A4	20511	43547	95004	BCg-3(2)	7.25	0.6	0.8	7.30	45.0	19.0	23.0	4.7
B13	20513	43542	95024	BCg-3(2)	7.25	0.6	0.8	7.30	45.0	19.0	23.0	4.7
C1	20519	43560	95027	BCT-4/5	7.11	0.7	0.9	10.50	28.0	10.0	41.0	4.7
C2	20502	43538	95028	NSG-2	6.86	0.4	2.20	0.7	38.0	37.0	4.0	8.4
C3	20500	43542	95029	NSG-2	6.86	0.4	2.20	0.7	38.0	37.0	4.0	8.4
C5	20494	43548	95030	NSG-2	6.86	0.4	2.20	0.7	38.0	37.0	4.0	8.4
C6	20491	43538	95031	BEg-3(2)	6.85	0.9	1.0	1.0	41.5	25.0	9.0	5.7
C9	20487	43558	95032	BEg-3(2)	6.85	0.9	1.0	1.0	41.5	25.0	9.0	5.7
C11	20468	43557	95033	NGG1-2	6.67	0.4	1.7	0.7	36.0	27.0	3.5	7.5
C12	20468	43561	95034	NGG1-2	6.67	0.4	1.7	0.7	36.0	27.0	3.5	7.5
C13	20462	43544	95035	NGG2-2	6.88	0.5	1.7	0.4	60.0	24.0	2.0	5.8
C14	20461	43547	95036	NGG2-2	6.88	0.5	1.7	0.4	60.0	24.0	2.0	5.8
C16	20458	43560	95037	NGG2-2	6.88	0.5	1.7	0.4	60.0	24.0	2.0	5.8
E10	20454	43544	95047	NGG1-2	6.67	0.4	1.7	0.7	36.0	27.0	3.5	7.5
E14	20511	43560	95051	BCT-4/5	7.11	0.7	0.9	10.5	28.0	10.0	41.0	4.7
E15	20503	43545	95052	BCg-3(2)	7.25	0.6	0.8	7.3	45.0	19.0	23.0	4.7

Table 9.14b. Values of soil properties estimated by values from TYPICAL PROFILE in the map unit at the random check points in the 100 m grid at Velford.

SITE	NORTHING	EASTING	SN	MAP UNIT	pH	K	Mg	CO3	CLAY	SILT	SAND	OM
A1	20563	43584	95001	NSG-2	7.07	0.7	1.75	0.85	44.0	27.0	5.5	8.7
A2	20540	43503	95002	BEg-3(2)	6.63	0.45	1.55	2.7	39.5	23.5	5.0	6.3
A3	20533	43571	95003	BCT-4/5	7.11	0.7	0.9	10.5	28.0	10.0	41.0	4.7
A4	20511	43547	95004	BCg-3(2)	7.22	1.0	1.05	2.45	44.0	25.5	11.5	5.9
A5	20502	43512	95005	BEg-3(2)	6.63	0.45	1.55	2.75	39.5	23.5	5.0	6.0
A6	20466	43572	95006	CGG-2	6.79	0.4	1.30	0.4	42.0	29.0	6.0	6.1
B1	20544	43572	95012	BEg-3(2)	6.63	0.45	1.55	2.75	39.5	23.5	5.0	6.3
B2	20531	43504	95013	BEg-3(2)	6.63	0.45	1.55	2.75	39.5	23.5	5.0	6.3
B3	20521	43595	95014	BEg-3(2)	6.63	0.45	1.55	2.75	39.5	23.5	5.0	6.3
B4	20532	43544	95015	BCg-3(2)	7.22	1.0	1.05	2.45	44.0	25.5	11.5	5.9
B5	20532	43590	95016	BCg-3(2)	7.22	1.0	1.05	2.45	44.0	25.5	11.5	5.9
B6	20524	43539	95017	BCg-3(2)	7.22	1.0	1.05	2.45	44.0	25.5	11.5	5.9
B7	20508	43491	95018	BEg-3(2)	6.63	0.45	1.55	2.75	39.5	23.5	5.0	6.3
B8	20500	43579	95019	NSG-2	7.07	0.7	1.75	0.85	44.0	27.5	5.5	8.7
B9	20491	43582	95020	NSG-2	7.07	0.7	1.75	0.85	44.0	27.5	5.5	8.7
B10	20461	43509	95021	NGG1-2	6.81	0.2	1.3	0.6	35.0	28.0	2.0	4.9
B12	20455	43576	95023	CGG-2	6.79	0.4	1.3	0.4	42.0	29.0	6.0	6.1
B13	20513	43542	95024	BCg-3(2)	7.22	1.0	1.05	2.45	44.0	25.5	11.5	5.9
C1	20519	43560	95027	BCT-4/5	7.11	0.7	0.9	10.5	28.0	10.0	41.0	4.7
C2	20502	43538	95028	BEg-3(2)	6.63	0.45	1.55	2.75	39.5	23.5	5.0	6.3
C3	20500	43542	95029	BEg-3(2)	6.63	0.45	1.55	2.75	39.0	23.5	5.0	6.3
C5	20494	43548	95030	BEg-3(2)	6.63	0.45	1.55	2.75	39.0	23.5	5.0	6.3
C6	20491	43538	95031	BEg-3(2)	6.63	0.45	1.55	2.75	39.0	23.5	5.0	6.3
C9	20487	43558	95032	BEg-3(2)	6.63	0.45	1.55	2.75	39.0	23.5	5.0	6.3
C11	20468	43557	95033	NGG1-2	6.81	0.2	1.3	0.6	35.0	28.0	2.0	4.9
C12	20468	43561	95034	NGG1-2	6.81	0.2	1.3	0.6	35.0	28.0	2.0	4.9
C13	20462	43544	95035	NGG1-2	6.81	0.2	1.3	0.6	35.0	28.0	2.0	4.9
C14	20461	43547	95036	NGG1-2	6.81	0.2	1.3	0.6	35.0	28.0	2.0	4.9
C16	20458	43560	95037	NGG2-2	7.26	0.6	2.1	0.6	66.0	21.0	1.0	6.6
E4	20513	43575	95041	BCg-3/4	7.44	0.8	1.7	2.1	48.0	27.0	8.0	7.6
E5	20453	43572	95042	CGG-2	6.79	0.4	1.3	0.4	42.0	29.0	6.0	6.1
E6	20560	43552	95043	NSG-2	7.07	0.7	1.75	0.85	44.0	27.5	5.5	8.7
E7	20450	43588	95044	NGG2-2	7.26	0.6	2.1	0.6	66.0	21.0	1.0	6.6
E8	20528	43524	95045	NGG1-2	6.81	0.2	1.3	0.6	35.0	28.0	2.0	4.9
E9	20553	43541	95046	NSG-2	7.07	0.7	1.75	0.85	44.0	27.5	5.5	8.7
E10	20454	43544	95047	NGG2-2	7.26	0.6	2.1	0.6	66.0	21.0	1.0	6.6
E11	20515	43567	95048	BCT-4/5	7.11	0.7	0.9	10.5	28.0	10.0	41.0	4.7
E12	20500	43563	95049	BCg-3/4	7.44	0.8	1.7	2.1	48.0	27.0	8.0	7.6
E14	20511	43560	95051	BCT-4/5	7.11	0.7	0.9	10.5	28.0	10.0	41.0	4.7
E15	20503	43545	95052	BCg-3(2)	7.22	1.0	1.05	2.45	44.0	25.5	11.5	5.9

Table 9.14c. Values of soil properties estimated by values of TYPICAL PROFILE in the map unit at the random check points in the 200 m grid at Yelford.

SITE	NORTHING	EASTING	SN	MAP UNIT	pH	K	Mg	CO3	CLAY	SILT	SAND	OM
A1	20563	43584	95001	BCg-3(2)	7.32	1.0	1.3	3.6	46.0	28.0	12.0	6.6
A2	20540	43503	95002	BEg-3(2)	6.8	0.35	8.5	0.6	27.5	26.5	5.5	4.3
A3	20533	43571	95003	BCg-3(2)	7.32	1.0	1.3	3.6	46.0	28.0	12.0	6.6
A4	20511	43547	95004	BEg-3(2)	6.8	0.35	8.5	0.6	27.5	26.5	5.5	4.3
A5	20502	43512	95005	BEg-3(2)	6.8	0.35	8.5	0.6	27.5	26.5	5.5	4.3
A6	20466	43572	95006	NGG2-2	6.95	0.4	1.6	0.8	51.0	21.0	2.0	5.0
A7	20446	43558	95007	NGG2-2	6.95	0.4	1.6	0.8	51.0	21.0	2.0	5.0
A8	20415	43481	95008	NGG2-2	6.95	0.4	1.6	0.8	51.0	21.0	2.0	5.0
A9	20397	43501	95009	NGG1-2	6.02	0.25	1.4	0.8	33.0	30.5	2.0	7.55
A10	20382	43572	95010	BCg-3	7.34	0.7	1.0	22.0	27.0	16.0	28.0	4.2
A12	20372	43563	95011	BCT-4/5	7.4	0.5	0.8	8.7	26.0	14.0	29.0	3.6
B1	20544	43572	95012	BEg-3(2)	6.8	0.35	8.5	0.6	27.5	26.5	5.5	4.3
B2	20531	43504	95013	BEg-3(2)	6.8	0.35	8.5	0.6	27.5	26.5	5.5	4.3
B3	20521	43495	95014	BEg-3(2)	6.8	0.35	8.5	0.6	27.5	26.5	5.5	4.3
B4	20532	43544	95015	BCg-3(2)	7.32	1.0	1.3	3.6	46.0	28.0	12.0	6.6
B5	20532	43590	95016	BCg-3(2)	7.32	1.0	1.3	3.6	46.0	28.0	12.0	6.6
B6	20524	43539	95017	NSG-2	7.13	0.8	2.7	1.1	35.0	33.0	5.0	12.3
B7	20508	43491	95018	BEg-3(2)	6.8	0.35	8.5	0.6	27.5	26.5	5.5	4.3
B8	20500	43579	95019	BEg-3(2)	6.8	0.35	8.5	0.6	27.5	26.5	5.5	4.3
B9	20491	43582	95020	BEg-3(2)	6.8	0.35	8.5	0.6	27.5	26.5	5.5	4.3
B10	20461	43509	95021	NGG1-2	6.02	0.25	1.4	0.8	33.0	30.0	2.0	7.55
B11	20475	43572	95022	NGG1-2	6.02	0.25	1.4	0.8	33.0	30.0	2.0	7.55
B12	20545	43576	95023	NGG2-2	6.95	0.4	1.6	0.8	51.0	21.0	2.0	5.0
B13	20513	43542	95024	BEg-3(2)	6.8	0.35	8.5	0.6	27.5	26.5	5.5	4.3
B14	20501	43559	95025	BEg-3(2)	6.8	0.35	8.5	0.6	27.5	26.5	5.5	4.3
B15	20481	43538	95026	NGG1-2	6.02	0.25	1.4	0.8	33.0	30.5	2.0	7.55
C1	20519	43560	95027	NSG-2	7.13	0.8	2.7	1.1	35.0	33.0	5.0	12.3
C2	20502	43538	95028	BEg-3(2)	6.8	0.35	8.5	0.6	27.5	26.5	5.5	4.3
C3	20500	43542	95029	NGG1-2	6.02	0.25	1.4	0.8	33.0	30.5	2.0	7.55
C5	20494	43548	95030	NGG1-2	6.02	0.25	1.4	0.8	33.0	30.5	2.0	7.55
C6	20491	43538	95031	NGG1-2	6.02	0.25	1.4	0.8	33.0	30.5	2.0	7.55
C9	20487	43558	95032	NGG1-2	6.02	0.25	1.4	0.8	33.0	30.5	2.0	7.55
C11	20468	43557	95033	NGG1-2	6.02	0.25	1.4	0.8	33.0	30.5	2.0	7.55
C12	20468	43557	95033	NGG1-2	6.02	0.25	1.4	0.8	33.0	30.5	2.0	7.55
C13	20462	43544	95035	NGG1-2	6.02	0.25	1.4	0.8	33.0	30.5	2.0	7.55
C14	20461	43547	95036	NGG1-2	6.02	0.25	1.4	0.8	33.0	30.5	2.0	7.55
C16	20458	43560	95037	NGG2-2	6.95	0.4	1.6	0.8	51.0	21.0	2.0	5.0
E1	20418	43468	95038	NGG1-2	6.02	0.25	1.4	0.8	33.0	30.5	2.0	7.55
E2	20567	43605	95039	BEg-3(2)	6.8	0.35	8.5	0.6	27.5	26.5	5.5	4.3
E3	20567	43578	95040	NSG-2	7.13	0.8	2.7	1.1	35.0	33.0	5.0	12.3
E4	20513	43575	95041	NSG-2	7.13	0.8	2.7	1.1	35.0	33.0	5.0	12.3
E5	20453	43572	95042	NGG2-2	6.95	0.4	1.6	0.8	51.0	21.0	2.0	5.0
E6	20560	43552	95043	NSG-2	7.13	0.8	2.7	1.1	35.0	33.0	5.0	12.3
E7	20450	43588	95044	NGG2-2	6.95	0.4	1.6	0.8	51.0	21.0	2.0	5.0
E8	20528	43524	95045	BEg-3(2)	6.8	0.35	8.5	0.6	27.5	26.5	5.5	4.3
E9	20553	43541	95046	BCg-3(2)	7.32	1.0	1.3	3.6	46.0	28.0	12.0	6.6
E10	20454	43544	95047	NGG1-2	6.02	0.25	1.4	0.8	33.0	30.5	2.0	7.55
E11	20515	43567	95048	NSG-2	7.13	0.8	2.7	1.1	35.0	33.0	5.0	12.3
E12	20500	43563	95049	BEg-3(2)	6.8	0.35	8.5	0.6	27.5	26.5	5.5	4.3
E13	20515	43561	95050	NSG-2	7.13	0.8	2.7	1.1	35.0	33.0	5.0	12.3
E14	20511	43560	95051	BEg-3(2)	6.8	0.35	8.5	0.6	27.5	26.5	5.5	4.3
E15	20503	43545	95052	BEg-3(2)	6.8	0.35	8.5	0.6	27.5	26.5	5.5	4.3

9.14.3 Mapping unit means as predictors.

The procedure to obtain the mapping unit means involved steps i. to iv. as listed in section 4.9.3 (chapter 4). In this case the site of interest referred to as X in 4.9.3 was each and every one of the check points at the three sampling grids.

Thus the position of profile data from the grids was located on the map. Their corresponding mapping unit read out from the map, and then data within the same mapping unit were grouped, and the mapping units means calculated for each property. The sites grouped within each of the mapping units are listed in appendix 9.4. The mapping unit means calculated in each sampling grid are included in tables 9.15a-c. The predicted values from mapping unit means were applied at the check points in the three grids. These are given in tables 9.15d-f.

9.15 Predictions based on spatial interpolation by Kriging.

The kriging technique was used to interpolate values of the 8 soil properties at the check points of the 50, 100 and 200m sampling grids, which were taken as targets.

Since Kriging requires the definition of a neighbourhood within which the property is spatially auto-correlated, and of the assignement of weights to neighbouring points according to the spatial relationship described by the semi-variogram model, the parameters of this model for the property and sampling grid in turn were provided by the semi-variograms of the Yelford area (figures 9.5-9.12).

Due to the nature of variability of the individual soil properties in the area, in not all the cases was it possible to find a model for the experimental semi-variogram. Anisotropic variation caused by the presence of strong trends precluded the fit of an isotropic model and the use of ordinary point Kriging to predict values at the target points. Cases of this sort were considered separately as indicated in section 9.13 above.

Where a semi-variogram model could be fitted satisfactorily to the experimental semi-variogram, spatial interpolation by

Table 9. 15a. Mapping unit means of 8 soil properties at the 50m grid, Yelford.

Map Unit	Mean values of soil properties within mapping units							
	pH	K	Mg	Carbonate	Clay	Silt	Coarse sand	Organic matter
BCT 4-5	7.18	0.92	1.06	9.48	31.80	13.20	34.00	5.48
BEg 3-2	6.57	0.55	1.05	0.35	42.00	26.50	9.00	5.35
BCg 3(2)	7.29	0.86	1.25	4.95	44.93	24.26	12.40	6.92
BCg 3/4	7.24	1.25	1.80	4.05	46.00	28.50	6.00	9.40
NSG-2	6.83	0.55	1.85	0.70	49.00	30.41	4.41	7.43
BEg-3(2)	6.52	0.41	1.31	1.32	35.45	28.66	4.75	7.23
NGG 1-2	6.52	0.40	1.71	0.57	37.62	28.59	2.88	8.47
NGG 2-2	6.81	0.56	1.90	0.74	59.18	22.90	1.63	5.81

Table 9. 15b. Mapping unit means of 8 soil properties at the 100m grid, Yelford.

Map unit	Mean values of soil properties within mapping units.							
	pH	K	Mg	Carbonate	Clay	Silt	Coarse sand	Organic matter
NSG-2	6.81	0.69	1.80	0.77	46.21	28.85	5.28	6.70
BCT-4/5	7.23	0.58	0.82	9.60	29.57	11.28	39.28	4.31
BEg-3(2)	6.63	0.60	1.71	1.01	37.58	26.24	4.90	6.30
BCg-3(2)	7.09	0.90	1.42	3.40	40.00	24.96	10.75	6.79
BCg-3/4	7.33	0.95	1.65	1.90	49.00	28.50	6.00	8.55
CSG-2	7.30	0.60	1.75	1.75	50.50	28.50	6.00	7.15
CGG-2	7.13	0.37	1.10	1.22	32.25	25.75	7.00	8.22
NGG 2-2	6.93	0.55	1.82	1.17	50.75	22.50	2.75	6.67
NGG 1-2	6.53	0.32	1.68	0.58	39.70	28.76	3.52	7.74

Table 9. 15c. Mapping unit means of 8 soil properties at the 200m grid, Yelford.

Map unit	Mean values of soil properties within mapping units.							
	pH	K	Mg	Carbonate	Clay	Silt	Coarse sand	Organic matter
BCT 4-5	7.34	0.64	0.92	10.66	27.40	13.80	32.60	4.28
BCg-3	7.06	0.46	1.05	3.56	29.50	20.00	13.33	6.01
NGG 1-2	6.54	0.42	1.53	0.76	35.13	26.66	3.00	7.00
NGG 2-2	6.76	0.64	1.79	0.77	53.66	21.33	2.20	5.34
BEg-3(2)	6.62	0.70	1.67	1.45	36.56	26.34	5.73	6.70
BCg-3(2)	7.27	0.71	1.35	2.86	47.36	24.72	10.00	5.49
NSG-2	7.25	0.30	1.60	0.90	39.00	33.00	4.00	8.10
BEg-2(3)	6.74	0.60	1.63	0.76	43.33	26.00	6.33	6.76
BC(g) 3-4	7.36	0.70	0.86	8.93	32.33	20.33	16.00	4.70

Table 9.15d. Values of soil properties estimated by MAP UNIT MEANS at the random check points in the 50 m grid at Yelford.

SITE	NORTH	EAST	SN	MAP UNIT	pH	K	Mg	CO3	CLAY	SILT	SAND	OM
A4	20511	43547	95004	BCg-3(2)	7.29	0.86	1.25	4.95	44.93	24.26	12.4	6.92
B13	20513	43542	95024	BCg-3(2)	7.29	0.86	1.25	4.95	44.93	24.26	12.4	6.92
C1	20519	43560	95027	BCT-4/5	7.18	0.92	1.06	9.48	31.8	13.2	34.0	5.48
C2	20502	43538	95028	NSG-2	6.83	0.55	1.85	0.7	49.0	30.41	4.41	7.43
C3	20500	43542	95029	NSG-2	6.83	0.55	1.85	0.7	49.0	30.41	4.41	7.43
C5	20494	43548	95030	NSG-2	6.83	0.55	1.85	0.7	49.0	30.41	4.41	7.43
C6	20491	43538	95031	BEg-3/2	6.57	0.55	1.05	0.35	42.0	26.50	9.0	5.35
C9	20487	43558	95032	BEg-3/2	6.57	0.55	1.05	0.35	42.0	26.50	9.0	5.35
C11	20468	43557	95033	NGG1-2	6.52	0.40	1.71	0.57	37.62	28.59	2.88	8.47
C12	20468	43561	95034	NGG1-2	6.52	0.40	1.71	0.57	37.62	28.59	2.88	8.47
C13	20462	43544	95035	NGG2-2	6.81	0.56	1.90	0.74	59.18	22.90	1.63	5.81
C14	20461	43547	95036	NGG2-2	6.81	0.56	1.90	0.74	59.18	22.90	1.63	5.81
C16	20458	43560	95037	NGG2-2	6.81	0.56	1.90	0.74	59.18	22.90	1.63	5.81
E10	20454	43544	95047	NGG1-2	6.52	0.40	1.71	0.57	37.62	28.59	2.88	8.47
E14	20511	43560	95051	BCT-4/5	7.18	0.92	1.06	9.48	31.80	13.20	34.0	5.48
E15	20503	43545	95052	BCg-3(2)	7.29	0.86	1.25	4.95	44.93	24.96	12.4	6.92

Table 9.15e. Values of soil properties estimated by MAP UNIT MEANS at the random check points in the 100 m grid at Yelford.

SITE	NORTH	EAST	SN	MAP UNIT	pH	K	Mg	CO3	CLAY	SILT	SAND	OM
A1	20563	43584	95001	NSG-2	6.81	0.69	1.80	0.77	46.21	28.85	5.28	6.70
A2	20540	43503	95002	BEg-3(2)	6.63	0.60	1.71	1.01	37.58	26.24	4.90	6.30
A3	20533	43571	95003	BCT-4/5	7.23	0.58	0.82	9.60	29.57	11.28	39.28	4.31
A4	20511	43547	95004	BCg-3(2)	7.09	0.90	1.42	3.40	40.0	24.96	10.75	6.79
A5	20502	43512	95005	BEg-3(2)	6.63	0.60	1.71	1.01	37.58	26.24	4.90	6.30
A6	20466	43572	95006	CGG-2	7.13	0.37	1.10	1.22	32.25	25.75	7.0	8.22
B1	20544	43572	95012	BEg-3(2)	6.63	0.60	1.71	1.01	37.58	26.24	4.90	6.30
B2	20531	43504	95013	BEg-3(2)	6.63	0.60	1.71	1.01	37.58	26.24	4.90	6.30
B3	20521	43495	95014	BEg-3(2)	6.63	0.60	1.71	1.01	37.58	26.24	4.90	6.30
B4	20532	43544	95015	BCg-3(2)	7.09	0.90	1.42	3.40	40.0	24.96	10.75	6.79
B5	20532	43590	95016	BCg-3(2)	7.09	0.90	1.42	3.40	40.0	24.96	10.75	6.79
B6	20524	43539	95017	BCg-3(2)	7.09	0.90	1.42	3.40	40.0	24.96	10.75	6.79
B7	20508	43491	95018	BEg-3(2)	6.63	0.60	1.71	1.01	37.58	26.24	4.90	6.30
B8	20500	43579	95019	NSG-2	6.81	0.69	1.80	0.77	46.21	28.85	5.28	6.70
B9	20491	43582	95020	NSG-2	6.81	0.69	1.80	0.77	46.21	28.85	5.28	6.70
B10	20461	43509	95021	NGG1-2	6.53	0.32	1.68	0.58	39.70	28.76	3.52	7.74
B12	20455	43576	95023	CGG-2	7.13	0.37	1.10	1.22	32.25	25.75	7.0	8.22
B13	20513	43542	95024	BCg-3(2)	7.09	0.90	1.42	3.40	40.0	24.96	10.75	6.79
C1	20519	43560	95027	BCT-4/5	7.23	0.58	0.82	9.60	29.57	11.28	39.28	4.31
C2	20502	43538	95028	BEg-3(2)	6.63	0.60	1.71	1.01	37.58	26.24	4.90	6.30
C3	20500	43542	95029	BEg-3(2)	6.63	0.60	1.71	1.01	37.58	26.24	4.90	6.30
C5	20494	43548	95030	BEg-3(2)	6.63	0.60	1.71	1.01	37.58	26.24	4.90	6.30
C6	20491	43538	95031	BEg-3(2)	6.63	0.60	1.71	1.01	37.58	26.24	4.90	6.30
C9	20487	43558	95032	BEg-3(2)	6.63	0.60	1.71	1.01	37.58	26.24	4.90	6.30
C11	20468	43557	95033	NGG1-2	6.53	0.32	1.68	0.58	39.70	28.76	3.52	7.74
C12	20468	43561	95034	NGG1-2	6.53	0.32	1.68	0.58	39.70	28.76	3.52	7.74
C13	20462	43544	95035	NGG1-2	6.53	0.32	1.68	0.58	39.70	28.76	3.52	7.74
C14	20461	43547	95036	NGG1-2	6.53	0.32	1.68	0.58	39.70	28.76	3.52	7.74
C16	20458	43560	95037	NGG2-2	6.93	0.55	1.82	1.17	50.75	22.50	2.75	6.67
E4	20513	43575	95041	BCg-3/4	7.33	0.95	1.65	1.90	49.0	28.5	6.0	8.55
E5	20453	43572	95042	CGG-2	7.13	0.37	1.10	1.22	32.25	25.75	7.0	8.22
E6	20560	43552	95043	NSG-2	6.81	0.69	1.80	0.77	46.21	28.85	5.28	6.70
E7	20450	43588	95044	NGG2-2	6.93	0.55	1.82	1.17	50.75	22.50	2.75	6.67
E8	20528	43524	95045	NGG1-2	6.53	0.32	1.68	0.58	39.70	28.76	3.52	7.74
E9	20553	43541	95046	NSG-2	6.81	0.69	1.80	0.77	46.21	28.85	5.28	6.70
E10	20454	43544	95047	NGG2-2	6.93	0.55	1.82	1.17	50.75	22.50	2.75	6.67
E11	20515	43567	95048	BCT-4/5	7.23	0.58	0.82	9.60	29.57	11.28	39.28	4.31
E12	20500	43563	95049	BCg-3/4	7.33	0.95	1.65	1.90	49.0	28.50	6.0	8.55
E14	20511	43560	95051	BCT-4/5	7.23	0.58	0.82	9.60	29.57	11.28	39.28	4.31
E15	20503	43545	95052	BCg-3(2)	7.09	0.90	1.42	3.40	40.0	24.96	10.75	6.79

Table 9.15f. Values of soil properties estimated by MAP UNIT MEANS at the random check points in the 200 m grid at Velford.

SITE	NORTH	EAST	SN	MAP UNIT	pH	K	Mg	CO3	CLAY	SILT	SAND	OM
A1	20563	43584	95001	BCg-3(2)	7.27	0.71	1.35	2.86	47.36	24.72	10.00	5.49
A2	20540	43503	95002	BEg-3(2)	6.62	0.70	1.67	1.45	36.56	26.34	5.73	6.70
A3	20533	43571	95003	BCg-3(2)	7.27	0.71	1.35	2.86	47.36	24.72	10.00	5.49
A4	20511	43547	95004	BEg-3(2)	6.62	0.70	1.67	1.45	36.56	26.34	5.73	6.70
A5	20502	43512	95005	BEg-3(2)	6.62	0.70	1.67	1.45	36.56	26.34	5.73	6.70
A6	20466	43572	95006	NGG2-2	6.76	0.64	1.79	0.77	53.66	21.33	2.20	5.34
A7	20446	43558	95007	NGG2-2	6.76	0.64	1.79	0.77	53.66	21.33	2.20	5.34
A8	20415	43418	95008	NGG2-2	6.76	0.64	1.79	0.77	53.66	21.33	2.20	5.34
A9	20397	43501	95009	NGG1-2	6.54	0.42	1.53	0.76	35.13	26.66	3.0	7.0
A10	20382	43572	95010	BCg-3(2)	7.27	0.71	1.35	2.86	47.36	24.72	10.0	5.49
A12	20372	43563	95011	BCT-4/5	7.34	0.64	0.92	10.66	27.40	13.80	32.6	4.28
B1	20544	43572	95012	BEg-3(2)	6.62	0.70	1.67	1.45	36.56	26.34	5.73	6.70
B2	20531	43504	95013	BEg-3(2)	6.62	0.70	1.67	1.45	36.56	26.34	5.73	6.70
B3	20521	43495	95014	BEg-3(2)	6.62	0.70	1.67	1.45	36.56	26.34	5.73	6.70
B4	20532	43544	95015	BCg-3(2)	7.27	0.71	1.35	2.86	47.36	24.72	10.00	5.49
B5	20532	43590	95016	BCg-3(2)	7.27	0.71	1.35	2.86	47.36	24.72	10.00	5.49
B6	20524	43539	95017	NSG-2	7.25	0.30	1.60	0.90	39.0	33.0	4.0	8.10
B7	20508	43491	95018	BEg-3(2)	6.62	0.7	1.67	1.45	36.56	26.34	5.73	6.70
B8	20500	43579	95019	BEg-3(2)	6.62	0.7	1.67	1.45	36.56	26.34	5.73	6.70
B9	20491	43582	95020	BEg-3(2)	6.62	0.7	1.67	1.45	36.36	26.34	5.73	6.70
B10	20461	43509	95021	NGG1-2	6.54	0.42	1.53	0.76	35.13	26.66	3.0	7.0
B11	20475	43572	95022	NGG1-2	6.54	0.42	1.53	0.76	35.13	26.66	3.0	7.0
B12	20455	43576	95023	NGG2-2	6.76	0.64	1.79	0.77	53.66	21.33	2.2	5.34
B13	20513	43542	95024	BEg-3(2)	6.62	0.7	1.67	1.45	36.56	26.34	5.73	6.70
B14	20501	43559	95025	BEg-3(2)	6.62	0.7	1.67	1.45	36.56	26.34	5.73	6.70
B15	20481	43538	95026	NGG1-2	6.54	0.42	1.53	0.76	35.13	26.66	3.0	7.0
C1	20519	43560	95027	NSG-2	7.25	0.30	1.60	0.90	39.0	33.00	4.0	8.1
C2	20502	43538	95028	BEg-3(2)	6.62	0.7	1.67	1.45	36.56	26.34	5.73	6.7
C3	20500	43542	95029	NGG1-2	6.54	0.42	1.53	0.76	35.13	26.66	3.0	7.0
C5	20494	43548	95030	NGG1-2	6.54	0.42	1.53	0.76	35.13	26.66	3.0	7.0
C6	20491	43538	95031	NGG1-2	6.54	0.42	1.53	0.76	35.13	26.66	3.0	7.0
C9	20487	43558	95032	NGG1-2	6.54	0.42	1.53	0.76	35.13	26.66	3.0	7.0
C11	20468	43557	95033	NGG1-2	6.54	0.42	1.53	0.76	35.13	26.66	3.0	7.0
C12	20468	43561	95034	NGG1-2	6.54	0.42	1.53	0.76	35.13	26.66	3.0	7.0
C13	20462	43544	95035	NGG1-2	6.54	0.42	1.53	0.76	35.13	26.66	3.0	7.0
C14	20461	43547	95036	NGG1-2	6.54	0.42	1.53	0.76	35.13	26.66	3.0	7.0
C16	20458	43560	95037	NGG2-2	6.76	0.64	1.79	0.77	53.66	21.33	2.2	5.34
E1	20418	43468	95038	NGG1-2	6.54	0.42	1.53	0.76	35.13	26.66	3.0	7.0
E2	20567	43605	95039	BEg-2(3)	6.74	0.60	1.63	0.76	43.33	26.00	6.33	6.76
E3	20567	43578	95040	NSG-2	7.25	0.30	1.60	0.90	39.0	33.00	4.0	8.1
E4	20513	43575	95041	NSG-2	7.25	0.30	1.60	0.90	39.0	33.00	4.0	8.1
E5	20453	43572	95042	NGG2-2	6.76	0.64	1.79	0.77	53.66	21.33	2.2	5.34
E6	20560	43552	95043	NSG-2	7.25	0.30	1.6	0.90	39.0	33.00	4.0	8.1
E7	20450	43588	95044	NGG2-2	6.76	0.64	1.79	0.77	53.66	21.33	2.2	5.34
E8	20528	43524	95045	BEg-3(2)	6.62	0.7	1.67	1.45	36.56	26.34	5.73	6.70
E9	20553	43541	95046	BCg-3(2)	7.27	0.71	1.35	2.86	47.36	24.72	10.0	5.49
E10	20454	43544	95047	NGG1-2	6.54	0.42	1.53	0.76	35.13	26.66	3.0	7.0
E11	20515	43567	95048	NSG-2	7.25	0.30	1.60	0.90	39.0	33.00	4.0	8.1
E12	20500	43563	95049	BEg-3(2)	6.62	0.7	1.67	1.45	36.56	26.34	5.73	6.70
E13	20515	43561	95050	NSG-2	7.25	0.30	1.6	0.90	39.0	33.00	4.0	8.10
E14	20511	43560	95051	BEg-3(2)	6.62	0.7	1.67	1.45	36.56	26.34	5.73	6.70
E15	20503	43545	95052	BEg-3(2)	6.62	0.7	1.67	1.45	36.56	26.34	5.73	6.70

ordinary point Kriging was performed. Spatial prediction by Kriging interpolation was, in these cases a relatively straight-forward operation. In a practical context, these cases would present no difficulties to a soil surveyor with some knowledge of the technique and in possesion of the appropriate software. His efforts would have to concentrate only on learning to manipulate the data files and run the Kriging algorithm in his own system before obtaining some predictions.

Thus, where spatial dependence was evident and a model for the experimental semi-variogram was found, the parameters of the model were fed into the algorithm of the "KRIGING" facility of OXSIS and run interactively. The point-data in the intersections of the sampling grids were used as the data for Kriging. The values at the random check points within each of the corresponding grids were thus interpolated by the procedures described in section 7.3.3.11 in chapter 7, using equation 7.51, and solving the Kriging system in eqs. 7.59 and 7.60. Estimates at those target points and their estimation variances were this way obtained. In all instances the neighbourhood for interpolation was limited to the closest 10 to 20 points.

The values of the soil properties predicted by interpolation at the check points for the three grids are shown in tables 9.16a-c. Three-dimensional plots of the Kriging estimates at the random check points were produced for visual inspection of the degree of approximation to the original observed values. These are included in appendix 9.4. The estimation variances for the same points were calculated using equation 7.61, and are given separately in appendix 9.5.

The way in which anisotropic variation caused by the presence of strong trends in the area affected spatial prediction was through preventing the use of ordinary Kriging interpolation. At the 50 m grid such type of variation was present in half of the properties studied. Potassium, Carbonate, Silt, and Coarse Sand could not be predicted by Kriging interpolation at the target points for they showed clear anisotropic variation, (figs. 9.6a, 9.8a, 9.10a, and 9.11a).

Table 9.16.a. Predicted values by kriging at the random check points in the 50 m grid.

NORTHING	EASTING	SITE No	pH	K*	Mg	CO_3*	CLAY	SILT*	SAND*	OM
205110	435470	95004	7.3041		1.3121		45.9403			6.6530
205130	435420	95024	7.1679		1.3628		43.8827			6.3610
205190	435600	95027	7.2075		0.9577		28.9801			5.1287
205020	435380	95028	6.7743		1.6935		44.9531			6.7377
205000	435420	95029	6.9227		1.6058		42.0004			6.7762
204940	435480	95030	6.9028		1.9022		44.0343			7.8589
204910	435380	95031	6.4513		1.2497		31.8327			5.5527
204870	435580	95032	6.7438		1.7539		34.4729			9.0518
204680	435570	95033	6.4120		2.1161		32.1655			10.5189
204680	435610	95034	6.6889		1.9624		27.8661			10.2154
204620	435440	95035	6.6910		1.9242		50.8074			7.0596
204610	435470	95036	6.6345		1.9716		55.6594			6.9965
204580	435600	95037	6.9344		1.8526		58.6968			6.6493
204540	435440	95047	6.9400		1.7339		49.9293			5.6015
205110	435600	95051	7.1888		1.1744		34.3286			7.0532
205030	435450	95052	7.0763		1.7205		45.0025			7.4107

*Values not estimated by kriging due to anisotropic variation at this sampling scale.

Table 9.16.b Predicted values by kriging at the random check points in the 100 m grid.

NORTHING	EASTING	SITE No	pH	K	Mg	CO_3	CLAY	SILT	SAND	OM
205400	435030	95002	6.8430	0.6490	1.8977	0.6988	31.7500	26.9348	5.4866	7.4039
205330	435710	95003	7.2578	0.4648	0.9629	9.2997	26.6954	9.3787	48.1732	3.9436
205110	435470	95004	7.1861	1.0551	1.3027	4.2188	44.3273	26.3065	11.7319	6.6219
205020	435120	95005	6.5325	0.5004	1.6609	0.3833	34.3670	28.6367	6.3985	7.5428
204660	435720	95006	6.9787	0.2952	1.6448	1.0993	34.6795	23.8741	5.5135	8.2304
205440	435720	95012	7.1556	0.5476	1.0888	5.5196	34.7670	14.2190	32.5690	4.4021
205310	435040	95013	6.7835	0.9936	1.9360	1.0954	29.1039	25.8526	5.5839	8.0549
205210	434950	95014	6.6143	1.2873	2.1273	1.3488	29.7120	24.5839	4.0840	8.1513
205320	435440	95015	7.0074	0.5010	1.3026	2.1251	49.2093	24.9986	9.6311	5.6574
205320	435900	95016	7.2266	0.7919	1.2660	5.2576	35.0395	21.7988	8.5964	5.1518
205240	435390	95017	7.0324	0.5998	1.3250	2.0133	45.9333	25.5791	9.2755	6.0208
205080	434910	95018	6.3389	0.5339	2.2430	0.5555	27.1133	27.9098	2.3814	8.2420
205000	435790	95019	6.5626	1.0514	2.4916	0.5163	49.0109	29.3038	3.6555	8.6102
204910	435820	95020	6.8385	0.7749	2.2889	0.7679	42.9373	28.5446	6.3801	8.8948
204610	435090	95021	6.6503	0.1657	1.2188	0.4456	37.3191	23.2481	3.8741	4.6603
204550	435760	95023	7.0508	0.2777	1.4650	0.7289	37.5359	23.3674	5.3861	7.5626
205130	435420	95024	7.1557	0.8056	1.3288	2.7049	44.2469	27.4410	8.6909	6.3908
205190	435600	95027	7.2099	0.7458	1.1088	10.6970	31.9088	13.5663	39.0150	5.1964
205020	435380	95028	6.9080	0.4424	1.3572	0.6393	41.2456	28.5820	5.4968	6.7493
205000	435420	95029	6.9055	0.5693	1.4097	0.6997	41.7424	29.2491	6.1431	7.1461
204940	435480	95030	6.8252	0.6652	1.5239	0.9492	39.6361	30.7660	4.6061	3.4166
204910	435380	95031	6.5893	0.2388	1.3029	0.3641	33.7925	28.0376	7.4839	6.8212
204870	435580	95032	6.8277	0.4350	1.6668	1.7800	34.7939	30.6999	4.5142	9.5148
204680	435570	95033	6.7613	0.3337	1.8224	0.7146	37.3711	25.9592	3.4951	9.0015
204680	435610	95034	6.8438	0.2735	1.7698	0.6235	35.2062	25.3535	3.6418	8.9256
204620	435440	95035	6.6947	0.4655	1.8479	0.7481	51.6581	26.8911	1.4385	7.2296
204610	435470	95036	6.7284	0.5420	1.8976	1.0238	53.1284	25.3021	1.6713	7.1961
204580	435600	95037	6.9836	0.5851	1.7870	0.7806	54.9891	21.9106	1.8355	6.6742
205130	435750	95041	7.0275	0.8242	1.7344	4.0845	44.1374	23.8402	12.6901	6.5122
204530	435720	95042	7.0050	0.3555	1.5640	0.6878	44.9910	22.4482	3.4643	6.9186
205600	435520	95043	6.8638	0.6226	1.5241	1.0049	50.8808	28.7123	3.5962	6.1202
204500	435880	95044	7.0931	0.4078	1.3651	1.2596	26.4461	26.8625	5.9379	8.4276
205280	435240	95045	6.9600	0.7901	1.6584	3.8988	40.9502	27.8467	10.5157	6.3710
205530	435410	95046	6.7288	0.4020	1.6532	0.6393	48.9725	32.1020	3.9709	7.3746
204540	435440	95047	6.7963	0.3890	1.7991	0.6305	49.8023	26.7520	1.5575	5.9269
205150	435670	95048	7.1609	0.7822	1.3851	7.3476	35.8929	18.9465	25.9182	6.0308
205000	435630	95049	7.0783	0.9513	1.7886	2.1284	45.5122	29.4270	5.1849	9.1849
205110	435600	95051	7.1799	0.8708	1.3322	10.0136	37.2378	20.9064	28.0415	6.9639
205030	435450	95052	7.0394	0.8604	1.4260	1.6675	44.3826	29.0809	5.6712	7.5349

Table 9.16.C Predicted values by kriging at the random check points in the 200m grid.

NORTHING	EASTING	SITE No	pH	K	Mg	CO_3	CLAY	SILT	SAND	OM
205630	435840	95001	6.85	0.5445	1.1725	3.0957	62.1220	23.5158	11.0612	4.6357
205400	435030	95002	6.85	0.8108	1.9271	1.2678	35.9815	26.2920	7.5051	7.6089
205330	435710	95003	6.85	0.6977	1.2764	4.3256	36.3972	22.9298	17.9039	4.7790
205110	435470	95004	6.85	0.7609	1.3638	2.5343	44.7373	27.5164	9.7054	6.6573
205020	435120	95005	6.85	0.5628	1.7048	1.0488	35.2827	26.5890	5.7955	6.8010
204660	435720	95006	6.85	0.7299	2.1735	2.3755	37.4341	25.7177	5.4469	9.3640
204460	435580	95007	6.85	0.4750	1.9598	0.8419	52.9318	24.2695	2.6597	6.2536
204150	434810	95008	6.85	0.5467	1.6992	0.8459	46.0113	23.6758	2.4351	5.0192
203970	435010	95009	6.85	0.7312	1.9961	0.7296	54.3189	22.3960	2.3808	6.5239
203820	435720	95010	6.85	0.4465	1.0915	6.6794	33.7530	18.7517	14.4651	4.8354
203720	435630	95011	6.85	0.4977	1.0515	8.9393	31.6767	17.9121	17.5236	4.6051
205440	435720	95012	6.85	0.6238	1.3064	2.7098	41.4698	23.3286	13.2834	4.9879
205310	435040	95013	6.85	0.8043	2.0072	1.2455	33.0225	26.4467	7.1388	8.0301
205210	434950	95014	6.85	0.7117	2.0060	1.0932	31.4246	26.2446	6.4250	7.7733
205320	435440	95015	6.85	0.6654	1.3756	2.6211	49.6337	26.2732	10.9634	5.6710
205320	435930	95016	6.85	0.7206	1.3879	4.6638	39.4986	23.7229	11.6096	5.7607
205240	435390	95017	6.85	0.6864	1.3728	2.3539	47.3721	27.0173	9.6153	6.1480
205080	434910	95018	6.85	0.6001	1.9185	1.0540	28.3886	26.1175	5.2631	7.7004
205000	435790	95019	6.85	0.9275	2.2192	2.0207	39.3632	27.9706	7.4414	10.5656
204910	435820	95020	6.85	0.9288	2.2875	2.3535	35.5031	27.8477	7.0542	11.4717
204610	435090	95021	6.85	0.3575	1.3120	0.7332	40.5770	25.2001	2.8518	5.0430
204750	435720	95022	6.85	0.7900	2.2677	2.3378	34.5387	26.8576	5.9278	10.5175
204550	435760	95023	6.85	0.6930	1.9689	2.2184	42.7546	24.3819	5.0153	7.8097
205130	435420	95024	6.85	0.7182	1.3557	2.2783	44.5018	27.5772	9.2039	6.5240
205010	435590	95025	6.85	0.8118	1.8384	1.9111	40.6543	28.1891	7.8892	8.8359
204810	435380	95026	6.85	0.4676	1.8278	0.7918	39.3790	27.3193	4.6857	7.6045
205190	435600	95027	6.85	0.7662	1.4421	3.3165	43.7226	25.8784	13.0873	6.0031
205020	435380	95028	6.85	0.6271	1.4862	1.5275	39.8173	27.9661	7.0357	7.0464
205000	435420	95029	6.85	0.6531	1.5464	1.5623	40.1544	28.1033	7.1765	7.4618
204940	435480	95030	6.85	0.6570	1.7794	1.3347	38.8857	28.4478	6.3109	8.7429
204910	435380	95031	6.85	0.5473	1.6280	1.0277	37.8759	27.6003	5.7259	7.2336
204870	435580	95032	6.85	0.7090	2.1524	1.3377	37.2794	28.5211	5.6117	10.4004
204690	435570	95033	6.85	0.5619	2.3030	1.1657	38.0380	26.7191	4.1892	9.7736
204690	435610	95034	6.85	0.6075	2.2612	1.4058	37.7585	26.4859	4.4463	9.6851
204620	435440	95035	6.85	0.4254	2.0777	0.7038	42.6628	26.4361	3.1742	7.7254
204610	435470	95036	6.85	0.4423	2.1294	0.7401	42.6162	26.3526	3.2099	7.9741
204530	435600	95037	6.85	0.5246	2.1554	1.1683	45.0813	25.3597	3.6063	7.9803
204190	434680	95038	6.85	0.5580	1.5892	0.8480	45.0639	23.7906	2.7140	4.0963
205670	436050	95039	6.85	0.5740	1.0846	9.8855	33.2864	21.2722	18.7936	4.6119
205670	435780	95040	6.85	0.5347	1.2659	2.3020	44.7571	24.4105	9.6473	4.7960
205130	435750	95041	6.85	0.8917	1.8850	2.5383	41.9671	26.6309	10.1386	8.0323
204530	435720	95042	6.85	0.6253	1.9883	1.7469	46.9870	24.1318	4.2157	7.2321
205500	435520	95043	6.85	0.5468	1.4713	0.8679	49.5342	26.0727	7.5535	5.2286
204500	435880	95044	6.85	0.7384	1.6967	2.4826	34.8504	23.9819	6.0975	7.9630
205230	435240	95045	6.85	0.7151	1.6348	1.6085	42.0488	27.4290	7.2937	7.2539
205530	435410	95046	6.85	0.5833	1.5856	1.2018	47.2319	27.2467	7.5574	6.7992
204540	435440	95047	6.85	0.3950	1.9316	0.6242	44.9504	25.9206	2.5832	6.5229
205150	435470	95048	6.85	0.8257	1.6844	2.8018	42.5165	26.3651	11.6176	7.0679
205020	435630	95049	6.85	0.8433	1.9661	1.8243	39.3497	28.2761	7.6052	9.3331
205150	435610	95050	6.85	0.7923	1.5406	2.9180	43.5309	26.4748	11.7315	6.5192
205110	435600	95051	6.85	0.7978	1.6101	2.5828	43.4939	27.1488	10.3781	7.1904
205030	435450	95052	6.85	0.6935	1.5163	1.3526	41.5133	28.0558	7.8897	7.3745

Kriging as a predictive procedure is at severe disadvantage in circumstances like these if it is compared against any other conventional predictive method. This is true unless more elaborate geostatistical procedures are followed to account for the anisotropies and trends, and to enable interpolation by Kriging. However, in practice, procedures of this sort are usually not within the grasp of a soil surveyor with no special training in geostatistics.

Attempts to fit semi-variogram models for anisotropies have been described in section 9.13 above. Their results lead to nothing and had to be abandoned.

The presence of strong trends in the area which reflect the local geology and soil/landscape morphogenesis lent themselves to suggest that the anisotropies encountered were only particular cases of a general variability situation. Thus the causality of the anisotropies encountered is explained here in terms of the general trend and the alluvium drift. The natural variability of such properties appears to challenge the assumptions in the intrinsic hypothesis. The trends can be considered strong enough to indicate that it might not be even safe to assume quasi-stationarity, and thence to require Universal Kriging to interpolate with non-stationary data. However, the complications of a detailed structural analysis as suggested by David (1977), and Burgess (1979) leading to Universal Kriging are such that they were considered not pertinent for the current cases, and were avoided in this thesis.

Thus, ignoring the trends meant that interpolated values by Kriging could not be obtained for those properties showing anisotropies since no isotropic semi-variogram model was found to describe adequately such variability.

On the other hand, the semi-variogram of pH at the 200m grid (fig 9.5c) was interpreted as Pure Nugget Effect and consequently the best estimate for this property at the target points in this grid was the general mean value of the sample data in the grid.

9.16 Comparison of predictive procedures ignoring trends.

The presence of trends in the variability of some of the properties studied in the Yelford area was sufficiently clear, and was reflected on anisotropic variations of those properties at the same sampling scale, i.e. the 50m grid. Spatial phenomena of this sort obscures the direct applicability of Kriging . Although elaborate, there are current procedures to permit Kriging under such circumstances, but they were discarded in this exercise and the predictive procedures compared as if the trends and their consequent anisotropies were not present. This was so in order to be able to explore the effect that ignoring the trends or accounting for them by modelling or by landscape stratification, could have in the results of the comparison.

So, first, the trends were ignored and the merits of the various predictive tools compared. The criterion of success in prediction that was used for the com parison was the Residual Sum of Squares of the predicted values by each procedure, from the values observed and measured at the check points in each of the three sampling grids. The squared difference between the observed and the predicted values by each technique were calculated and added up to obtain the RSS for each of the 8 properties studied at the 3 sampling grids. The RSS is taken as a measure of the error in prediction. The predictive merits of a technique increase as RSS decreases.

Thus the values predicted by conventional soil survey output, i.e. class means, mapping unit means, the typical profile in the mapping unit, and by Kriging are shown in tables 9.17a-c for soil pH, 9.18a-c for Exchangeable Potassium, 9.19a-c for Exchangeable Magnesium, 9.20a-c for Total Carbonate, 9.21a-c for Clay, 9.22a-c for Silt, 9.23a-c for Coarse Sand and 9.24a-c for Organic Matter content. These are termed the "tables of predictions", and they also give the values of RSS from the observed.

In order to analyze the implications of the results in the tables of predictions some observations are pertinent. First, the predictive success of each technique (RSS) must be related to the sampling effort involved in obtaining the predictions.

Table 9.17a Comparison of predicted values of pH by different predictive procedures at the random check points in the 50 m grid at Yelford.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING ESTIMATE
20511	43547	95004	7.38	7.32	7.29	7.25	7.30
20513	43542	95024	7.33	6.53	7.29	7.25	7.17
20519	43560	95027	7.08	7.18	7.18	7.11	7.20
20502	43538	95028	6.80	6.87	6.83	6.86	6.78
20500	43542	95029	7.24	6.53	6.83	6.86	6.93
20494	43548	95030	6.75	6.87	6.83	6.86	6.90
20491	43538	95031	6.08	6.57	6.57	6.85	6.45
20487	43558	95032	6.04	6.57	6.57	6.85	6.74
20468	43557	95033	6.78	7.21	6.52	6.67	6.41
20468	43561	95034	6.02	6.49	6.52	6.67	6.68
20462	43544	95035	6.26	6.49	6.81	6.88	6.70
20461	43547	95036	5.98	6.49	6.81	6.88	6.64
20458	43560	95037	6.53	6.49	6.81	6.88	6.93
20454	43544	95047	6.87	6.49	6.52	6.67	6.94
20511	43560	95051	7.04	7.18	7.18	7.11	7.15
20503	43545	95052	7.03	7.21	7.29	7.25	7.02
ERROR IN PREDICTION:							
		SUMS OF SQUARES =		2.614	2.313	3.278	2.179
		(Deviations from observed)					

Table 9.17.b. Comparison of predicted values of pH by different predictive procedures at the random check points of the 100 m grid.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING ESTIMATES
20540	43503	95002	6.67	6.60	6.63	6.63	6.84
20533	43571	95003	7.31	7.24	7.23	7.11	7.26
20511	43547	95004	7.38	7.17	7.09	7.22	7.19
20502	43512	95005	5.75	6.63	6.63	6.63	6.53
20466	43572	95006	6.62	7.13	7.13	6.79	6.98
20544	43572	95012	5.74	6.64	6.63	6.63	7.16
20531	43504	95013	6.97	6.64	6.63	6.63	6.78
20521	43495	95014	5.50	6.64	6.63	6.63	6.61
20532	43544	95015	7.10	6.64	7.09	7.22	7.01
20532	43590	95016	7.14	7.17	7.09	7.22	7.23
20524	43539	95017	7.09	7.17	7.09	7.22	7.03
20508	43491	95018	6.67	6.64	6.63	6.63	6.34
20500	43579	95019	5.09	6.81	6.81	7.07	6.56
20491	43582	95020	4.53	6.81	6.81	7.07	6.84
20461	43509	95021	5.58	6.60	6.53	6.81	6.65
20455	43576	95023	7.08	7.13	7.13	6.79	7.05
20513	43542	95024	7.33	6.60	7.09	7.22	7.16
20519	43560	95027	7.08	7.24	7.23	7.11	7.21
20502	43538	95028	6.80	6.81	6.63	6.63	6.91
20500	43542	95029	7.24	6.60	6.63	6.63	6.91
20494	43548	95030	6.75	6.81	6.63	6.63	6.83
20491	43538	95031	6.08	6.64	6.63	6.63	6.59
20487	43558	95032	6.04	6.64	6.63	6.63	6.83
20468	43557	95033	6.78	7.13	6.53	6.81	6.76
20468	43561	95034	6.02	6.63	6.53	6.81	6.84
20462	43544	95035	6.26	6.63	6.53	6.81	6.69
20461	43547	95036	5.98	6.63	6.53	6.81	6.73
20458	43560	95037	6.53	6.63	6.93	7.26	6.98
20513	43575	95041	7.39	7.23	7.33	7.44	7.03
20453	43572	95042	6.90	7.13	7.13	6.79	7.00
20560	43552	95043	6.84	6.81	6.81	7.07	6.86
20450	43588	95044	6.78	7.23	6.93	7.26	7.09
20528	43524	95045	6.68	6.63	6.53	6.81	6.96
20553	43541	95046	6.78	6.81	6.81	7.07	6.73
20454	43544	95047	6.87	6.63	6.93	7.26	6.80
20515	43567	95048	7.33	7.17	7.23	7.11	7.16
20500	43563	95049	7.22	7.17	7.33	7.44	7.08
20511	43560	95051	7.04	7.24	7.23	7.11	7.18
20503	43545	95052	7.03	7.23	7.09	7.22	7.04
ERROR IN PREDICTION:							
		SUMS OF SQUARES =		15.8868	14.5500	18.9805	15.9505
		(DEVIATIONS FROM THE OBSERVED)					

Table 9.17.c. Comparison of predicted values of pH by different predictive procedures at the random check points in the 200 m grid at Yelford.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING * ESTIMATE
20563	43584	95001	7.04	6.53	7.27	7.32	6.85
20540	43503	95002	6.67	6.53	6.62	6.80	6.85
20533	43571	95003	7.31	7.34	7.27	7.32	6.85
20511	43547	95004	7.38	7.27	6.62	6.80	6.85
20502	43512	95005	5.75	6.56	6.62	6.80	6.85
20466	43572	95006	6.62	7.13	6.76	6.95	6.85
20446	43558	95007	6.33	6.79	6.76	6.95	6.85
20415	43481	95008	5.98	6.56	6.76	6.95	6.85
20397	43501	95009	6.58	6.56	6.54	6.02	6.85
20382	43572	95010	7.29	7.19	7.27	7.34	6.85
20372	43563	95011	7.21	7.27	7.34	7.40	6.85
20544	43572	95012	5.74	6.73	6.62	6.80	6.85
20531	43504	95013	6.97	6.73	6.62	6.80	6.85
20521	43495	95014	5.50	6.73	6.62	6.80	6.85
20532	43544	95015	7.10	6.73	7.27	7.32	6.85
20532	43590	95016	7.14	7.27	7.27	7.32	6.85
20524	43539	95017	7.09	7.27	7.25	7.13	6.85
20508	43491	95018	6.67	6.73	6.62	6.80	6.85
20500	43579	95019	5.09	6.99	6.62	6.80	6.85
20491	43582	95020	4.53	6.99	6.62	6.80	6.85
20461	43509	95021	5.58	6.53	6.54	6.02	6.85
20475	43572	95022	5.94	6.53	6.54	6.02	6.85
20455	43576	95023	7.08	7.13	6.76	6.95	6.85
20513	43542	95024	7.33	6.53	6.62	6.80	6.85
20501	43559	95025	7.24	7.27	6.62	6.80	6.85
20481	43538	95026	6.73	6.53	6.54	6.02	6.85
20519	43560	95027	7.08	7.34	7.25	7.13	6.85
20502	43538	95028	6.80	6.99	6.62	6.80	6.85
20500	43542	95029	7.24	6.53	6.54	6.02	6.85
20494	43548	95030	6.75	6.99	6.54	6.02	6.85
20491	43538	95031	6.08	6.73	6.54	6.02	6.85
20487	43558	95032	6.04	6.73	6.54	6.02	6.85
20468	43557	95033	6.78	7.13	6.54	6.02	6.85
20468	43561	95034	6.02	6.56	6.54	6.02	6.85
20462	43544	95035	6.26	6.56	6.54	6.02	6.85
20461	43547	95036	5.98	6.56	6.54	6.02	6.85
20458	43560	95037	6.53	6.56	6.76	6.95	6.85
20418	43468	95038	6.23	6.56	6.54	6.02	6.85
20567	43605	95039	7.33	7.28	6.74	6.80	6.85
20567	43578	95040	7.17	6.99	7.25	7.13	6.85
20513	43575	95041	7.39	7.19	7.25	7.13	6.85
20453	43572	95042	6.90	7.13	6.76	6.95	6.85
20560	43552	95043	6.84	6.99	7.25	7.13	6.85
20450	43588	95044	6.78	7.19	6.76	6.95	6.85
20528	43524	95045	6.68	6.56	6.62	6.80	6.85
20553	43541	95046	6.78	6.99	7.27	7.32	6.85
20454	43544	95047	6.87	6.56	6.54	6.02	6.85
20515	43567	95048	7.33	7.27	7.25	7.13	6.85
20500	43563	95049	7.22	7.27	6.62	6.80	6.85
20515	43561	95050	7.25	7.34	7.25	7.13	6.85
20511	43560	95051	7.04	7.34	6.62	6.80	6.85
20503	43545	95052	7.03	7.19	6.62	6.80	6.85
Error in prediction:							
		SUMS OF SQUARES =		19.224	17.017	20.255	22.937

(Residuals from observed)

* Values estimated by the MEAN (6.85) due to PURE NUGGET EFFECT at this

Table 9.18a. Comparison of predicted values of POTASSIUM by different predictive procedures at the random check points in the 50 m grid at Yelford.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING **
20511	43547	95004	1.00	1.15	0.86	0.60	
20513	43542	95024	0.70	0.38	0.86	0.60	
20519	43560	95027	0.70	0.92	0.92	0.70	
20502	43538	95028	0.60	0.59	0.55	0.40	
20500	43542	95029	0.50	0.38	0.55	0.40	
20494	43548	95030	1.10	0.59	0.55	0.40	
20491	43538	95031	0.30	0.50	0.55	0.90	
20487	43558	95032	0.30	0.50	0.55	0.90	
20468	43557	95033	0.30	0.77	0.40	0.40	
20468	43561	95034	0.50	0.37	0.40	0.40	
20462	43544	95035	0.40	0.37	0.56	0.50	
20461	43547	95036	0.40	0.37	0.56	0.50	
20458	43560	95037	0.50	0.37	0.56	0.50	
20454	43544	95047	0.60	0.37	0.40	0.40	
20511	43560	95051	1.30	0.92	0.92	0.70	
20503	43545	95052	1.10	0.77	0.86	0.60	
		SUMS OF SQUARES = (Residuals from observed)		1.084	0.842	2.120	**

** Values predicted by ordinary KRIGING could not be obtained due to ANISOTROPIC variation of this property at this scale of sampling.

Table 9.18.b Comparison of predicted values of POTASSIUM by different predictive procedures at the random check points of the 100 m grid, Yelford.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING ESTIMATES
20540	43503	95002	0.90	0.48	0.60	0.45	0.65
20533	43571	95003	0.50	0.59	0.58	0.70	0.46
20511	43547	95004	1.00	1.09	0.90	1.00	1.06
20502	43512	95005	0.40	0.39	0.60	0.45	0.50
20466	43572	95006	0.30	0.38	0.37	0.40	0.30
20544	43572	95012	0.40	0.73	0.60	0.45	0.55
20531	43504	95013	1.20	0.73	0.60	0.45	0.99
20521	43495	95014	0.60	0.73	0.60	0.45	1.29
20532	43544	95015	0.50	0.73	0.90	1.00	0.50
20532	43590	95016	1.00	1.09	0.90	1.00	0.79
20524	43539	95017	0.80	1.09	0.90	1.00	0.60
20508	43491	95018	0.20	0.73	0.60	0.45	0.53
20500	43579	95019	1.50	0.68	0.69	0.70	1.05
20491	43582	95020	0.80	0.68	0.69	0.70	0.77
20461	43509	95021	0.20	0.48	0.32	0.20	0.17
20455	43576	95023	0.50	0.38	0.37	0.40	0.28
20513	43542	95024	0.70	0.48	0.90	1.00	0.81
20519	43560	95027	0.70	0.59	0.58	0.70	0.75
20502	43538	95028	0.60	0.68	0.60	0.45	0.44
20500	43542	95029	0.50	0.48	0.60	0.45	0.57
20494	43548	95030	1.10	0.68	0.60	0.45	0.67
20491	43538	95031	0.30	0.73	0.60	0.45	0.24
20487	43558	95032	0.30	0.73	0.60	0.45	0.44
20468	43557	95033	0.30	0.38	0.32	0.20	0.33
20468	43561	95034	0.50	0.39	0.32	0.20	0.27
20462	43544	95035	0.40	0.39	0.32	0.20	0.47
20461	43547	95036	0.40	0.39	0.32	0.20	0.54
20458	43560	95037	0.50	0.39	0.55	0.60	0.59
20513	43575	95041	1.10	0.73	0.95	0.80	0.82
20453	43572	95042	0.40	0.38	0.37	0.40	0.36
20560	43552	95043	0.60	0.68	0.69	0.70	0.62
20450	43588	95044	0.30	0.73	0.55	0.60	0.41
20528	43524	95045	1.00	0.39	0.32	0.20	0.79
20553	43541	95046	0.40	0.68	0.69	0.70	0.40
20454	43544	95047	0.60	0.39	0.55	0.60	0.39
20515	43567	95048	0.70	1.09	0.58	0.70	0.78
20500	43563	95049	0.80	1.09	0.95	0.80	0.95
20511	43560	95051	1.30	0.59	0.58	0.70	0.87
20503	43545	95052	1.10	0.73	0.90	1.00	0.86
ERROR IN PREDICTION:							
		SUMS OF SQUARES =		4.0617	3.3716	3.9175	1.8496
		(DEVIATIONS FROM OBSERVED)					

Table 9.18c. Comparison of predicted values of POTASSIUM by different predictive procedures at the random check points in the 200 m grid at Yelford.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING ESTIMATE
20563	43584	95001	0.60	0.50	0.71	1.00	0.54
20540	43503	95002	0.90	0.50	0.70	0.35	0.81
20533	43571	95003	0.50	0.64	0.71	1.00	0.70
20511	43547	95004	1.00	0.67	0.70	0.35	0.76
20502	43512	95005	0.40	0.41	0.70	0.35	0.56
20466	43572	95006	0.30	0.38	0.64	0.40	0.73
20446	43558	95007	0.60	0.64	0.64	0.40	0.48
20415	43481	95008	0.40	0.41	0.64	0.40	0.54
20397	43501	95009	0.30	0.41	0.42	0.25	0.73
20382	43572	95010	0.40	0.66	0.71	0.70	0.45
20372	43563	95011	0.50	0.67	0.64	0.50	0.50
20544	43572	95012	0.40	0.76	0.70	0.35	0.62
20531	43504	95013	1.20	0.76	0.70	0.35	0.80
20521	43495	95014	0.60	0.76	0.70	0.35	0.71
20532	43544	95015	0.50	0.76	0.71	1.00	0.67
20532	43590	95016	1.00	0.67	0.71	1.00	0.72
20524	43539	95017	0.80	0.67	0.30	0.80	0.69
20508	43491	95018	0.20	0.76	0.70	0.35	0.60
20500	43579	95019	1.50	0.52	0.70	0.35	0.93
20491	43582	95020	0.80	0.52	0.70	0.35	0.93
20461	43509	95021	0.20	0.50	0.42	0.25	0.36
20475	43572	95022	0.20	0.50	0.42	0.25	0.79
20455	43576	95023	0.50	0.38	0.64	0.40	0.69
20513	43542	95024	0.70	0.50	0.70	0.35	0.72
20501	43559	95025	0.90	0.67	0.70	0.35	0.81
20481	43538	95026	0.20	0.50	0.42	0.25	0.47
20519	43560	95027	0.70	0.64	0.30	0.80	0.77
20502	43538	95028	0.60	0.52	0.70	0.35	0.63
20500	43542	95029	0.50	0.50	0.42	0.25	0.65
20494	43548	95030	1.10	0.52	0.42	0.25	0.66
20491	43538	95031	0.30	0.76	0.42	0.25	0.55
20487	43558	95032	0.30	0.76	0.42	0.25	0.71
20468	43557	95033	0.30	0.38	0.42	0.25	0.56
20468	43561	95034	0.50	0.41	0.42	0.25	0.61
20462	43544	95035	0.40	0.41	0.42	0.25	0.43
20461	43547	95036	0.40	0.41	0.42	0.25	0.44
20458	43560	95037	0.50	0.41	0.64	0.40	0.52
20418	43468	95038	0.40	0.41	0.42	0.25	0.56
20567	43605	95039	0.70	0.90	0.60	0.35	0.57
20567	43578	95040	0.50	0.52	0.30	0.80	0.53
20513	43575	95041	1.10	0.66	0.30	0.80	0.89
20453	43572	95042	0.40	0.38	0.64	0.40	0.63
20560	43552	95043	0.60	0.52	0.30	0.80	0.55
20450	43588	95044	0.30	0.66	0.64	0.40	0.74
20528	43524	95045	1.00	0.41	0.70	0.35	0.72
20553	43541	95046	0.40	0.52	0.71	1.00	0.58
20454	43544	95047	0.60	0.41	0.42	0.25	0.40
20515	43567	95048	0.70	0.67	0.30	0.80	0.83
20500	43563	95049	0.80	0.67	0.70	0.35	0.84
20515	43561	95050	0.40	0.64	0.30	0.80	0.79
20511	43560	95051	1.30	0.64	0.70	0.35	0.80
20503	43545	95052	1.10	0.66	0.70	0.35	0.69
Error in Prediction:							
		SUMS OF SQUARES =		4.969	4.995	8.407	3.488
		(Deviations from observed)					

Table 9.19a. Comparison of predicted values of MAGNESIUM by different predictive procedures at the random check points in the 50 m grid at Yelford.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING ESTIMATE
20511	43547	95004	1.40	1.35	1.25	0.80	1.31
20513	43542	95024	1.30	1.37	1.25	0.80	1.36
20519	43560	95027	0.60	1.06	1.06	0.90	0.96
20502	43538	95028	2.00	1.93	1.85	2.20	1.69
20500	43542	95029	1.60	1.37	1.85	2.20	1.61
20494	43548	95030	2.80	1.93	1.85	2.20	1.90
20491	43538	95031	0.90	1.32	1.05	1.00	1.25
20487	43558	95032	1.70	1.32	1.05	1.00	1.75
20468	43557	95033	1.00	1.31	1.71	1.70	2.12
20468	43561	95034	3.40	1.70	1.71	1.70	1.96
20462	43544	95035	1.70	1.70	1.90	1.70	1.92
20461	43547	95036	1.90	1.70	1.90	1.70	1.97
20458	43560	95037	1.70	1.70	1.90	1.70	1.85
20454	43544	95047	2.10	1.70	1.71	1.70	1.73
20511	43560	95051	0.80	1.06	1.06	0.90	1.17
20503	43545	95052	1.50	1.31	1.25	0.80	1.72
ERROR IN PREDICTION:							
		SUMS OF SQUARES =		4.6443	5.3915	6.0400	4.8787
		(Deviations from observed)					

Table 9.19.b. Comparison of predicted values of MAGNESIUM by different predictive procedures at the random check points of the 100 m grid, Yelford.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING ESTIMATES
20540	43503	95002	2.30	1.73	1.71	1.55	1.90
20533	43571	95003	0.70	0.83	0.82	0.90	0.96
20511	43547	95004	1.40	1.62	1.42	1.05	1.30
20502	43512	95005	1.60	1.53	1.71	1.55	1.66
20466	43572	95006	1.60	0.98	1.10	1.30	1.64
20544	43572	95012	1.90	1.63	1.71	1.55	1.09
20531	43504	95013	1.50	1.63	1.71	1.55	1.94
20521	43495	95014	2.00	1.63	1.71	1.55	2.13
20532	43544	95015	1.60	1.63	1.42	1.05	1.30
20532	43590	95016	1.20	1.62	1.42	1.05	1.27
20524	43539	95017	0.80	1.62	1.42	1.05	1.32
20508	43491	95018	2.10	1.62	1.71	1.55	2.24
20500	43579	95019	3.80	1.96	1.80	1.75	2.49
20491	43582	95020	4.20	1.96	1.80	1.75	2.29
20461	43509	95021	1.10	1.73	1.68	1.30	1.22
20455	43576	95023	1.20	0.98	1.10	1.30	1.46
20513	43542	95024	1.30	1.73	1.42	1.05	1.33
20519	43560	95027	0.60	0.83	0.82	0.90	1.11
20502	43538	95028	2.00	1.96	1.71	1.55	1.36
20500	43542	95029	1.60	1.73	1.71	1.55	1.41
20494	43548	95030	2.80	1.96	1.71	1.55	1.52
20491	43538	95031	0.90	1.62	1.71	1.55	1.30
20487	43558	95032	1.70	1.62	1.71	1.55	1.67
20468	43557	95033	1.00	0.98	1.68	1.30	1.82
20468	43561	95034	3.40	1.53	1.68	1.30	1.77
20462	43544	95035	1.70	1.53	1.68	1.30	1.85
20461	43547	95036	1.90	1.53	1.68	1.30	1.90
20458	43560	95037	1.70	1.53	1.82	2.10	1.79
20513	43575	95041	1.50	1.42	1.65	1.70	1.73
20453	43572	95042	1.00	0.98	1.10	1.30	1.56
20560	43552	95043	1.30	1.96	1.80	1.75	1.52
20450	43588	95044	1.00	1.42	1.82	2.10	1.37
20528	43524	95045	2.10	1.53	1.68	1.30	1.66
20553	43541	95046	2.00	1.96	1.80	1.75	1.65
20454	43544	95047	2.10	1.53	1.82	2.10	1.80
20515	43567	95048	0.80	1.62	0.82	0.90	1.39
20500	43563	95049	1.80	1.62	1.65	1.70	1.79
20511	43560	95051	0.80	0.83	0.82	0.90	1.33
20503	43545	95052	1.50	1.42	1.42	1.05	1.43
ERROR IN PREDICTION:							
		SUMS OF SQUARES =		18.0958	18.2797	22.1125	14.4060
		(DEVIATIONS FROM OBSERVED)					

Table 9.19.c. Comparison of predicted values of MAGNESIUM by different predictive procedures at the random check points in the 200 m grid at Yelford.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING ESTIMATE
20563	43584	95001	1.20	1.55	1.35	1.30	1.17
20540	43503	95002	2.30	1.55	1.67	8.50	1.93
20533	43571	95003	0.70	0.92	1.35	1.30	1.28
20511	43547	95004	1.40	1.16	1.67	8.50	1.36
20502	43512	95005	1.60	1.58	1.67	8.50	1.70
20466	43572	95006	1.60	0.98	1.79	1.60	2.17
20446	43558	95007	1.90	1.82	1.79	1.60	1.96
20415	43481	95008	1.80	1.58	1.79	1.60	1.70
20397	43501	95009	2.20	1.58	1.53	1.40	2.00
20382	43572	95010	1.00	1.48	1.35	1.00	1.09
20372	43563	95011	1.00	1.16	0.92	0.80	1.05
20544	43572	95012	1.90	1.68	1.67	8.50	1.31
20531	43504	95013	1.50	1.68	1.67	8.50	2.01
20521	43495	95014	2.00	1.68	1.67	8.50	2.01
20532	43544	95015	1.60	1.68	1.35	1.30	1.38
20532	43590	95016	1.20	1.16	1.35	1.30	1.39
20524	43539	95017	0.80	1.16	1.60	2.70	1.37
20508	43491	95018	2.10	1.68	1.67	8.50	1.92
20500	43579	95019	3.80	1.71	1.67	8.50	2.22
20491	43582	95020	4.20	1.71	1.67	8.50	2.29
20461	43509	95021	1.10	1.55	1.53	1.40	1.31
20475	43572	95022	1.30	1.55	1.53	1.40	2.27
20455	43576	95023	1.20	0.98	1.79	1.60	1.97
20513	43542	95024	1.30	1.55	1.67	8.50	1.36
20501	43559	95025	1.30	1.16	1.67	8.50	1.84
20481	43538	95026	1.10	1.55	1.53	1.40	1.83
20519	43560	95027	0.60	0.92	1.60	2.70	1.44
20502	43538	95028	2.00	1.71	1.67	8.50	1.49
20500	43542	95029	1.60	1.55	1.53	1.40	1.55
20494	43548	95030	2.80	1.71	1.53	1.40	1.78
20491	43538	95031	0.90	1.68	1.53	1.40	1.63
20487	43558	95032	1.70	1.68	1.53	1.40	2.15
20468	43557	95033	1.00	0.98	1.53	1.40	2.30
20468	43561	95034	3.40	1.58	1.53	1.40	2.26
20462	43544	95035	1.70	1.58	1.53	1.40	2.08
20461	43547	95036	1.90	1.58	1.53	1.40	2.13
20458	43560	95037	1.70	1.58	1.79	1.60	2.16
20418	43468	95038	1.90	1.58	1.53	1.40	1.59
20567	43605	95039	1.50	1.50	1.63	8.50	1.08
20567	43578	95040	1.30	1.71	1.60	2.70	1.27
20513	43575	95041	1.50	1.48	1.60	2.70	1.88
20453	43572	95042	1.00	0.98	1.79	1.60	1.99
20560	43552	95043	1.30	1.71	1.60	2.70	1.47
20450	43588	95044	1.00	1.48	1.79	1.60	1.70
20528	43524	95045	2.10	1.58	1.67	8.50	1.63
20553	43541	95046	2.00	1.71	1.35	1.30	1.59
20454	43544	95047	2.10	1.58	1.53	1.40	1.93
20515	43567	95048	0.80	1.16	1.60	2.70	1.68
20500	43563	95049	1.80	1.16	1.67	8.50	1.97
20515	43561	95050	0.70	0.92	1.60	2.70	1.54
20511	43560	95051	0.80	0.92	1.67	8.50	1.61
20503	43545	95052	1.50	1.48	1.67	8.50	1.52
ERROR IN PREDICTION:							
		SUMS OF SQUARES =		20.861	26.510	73.300	21.137
		(Deviations from observed)					

Table 9.20.a. Comparison of predicted values of total CARBONATE by different predictive procedures at the random check points in the 50 m grid at Yelford.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING **
20511	43547	95004	2.40	5.25	4.95	7.30	
20513	43542	95024	0.40	1.94	4.95	7.30	
20519	43560	95027	18.00	9.48	9.48	10.50	
20502	43538	95028	0.20	0.82	0.70	0.70	
20500	43542	95029	1.00	1.94	0.70	0.70	
20494	43548	95030	0.60	0.82	0.70	0.70	
20491	43538	95031	0.20	0.68	0.35	1.00	
20487	43558	95032	0.00	0.68	0.35	1.00	
20468	43557	95033	0.80	3.88	0.57	0.70	
20468	43561	95034	0.20	0.55	0.57	0.70	
20462	43544	95035	0.30	0.55	0.74	0.40	
20461	43547	95036	0.10	0.55	0.74	0.40	
20458	43560	95037	0.40	0.55	0.74	0.40	
20454	43544	95047	0.40	0.55	0.57	0.70	
20511	43560	95051	8.90	9.48	9.48	10.50	
20503	43545	95052	1.40	3.88	4.95	7.30	
Error in prediction:							
		SUMS OF SQUARES =	101.502	114.166	167.680	**	
		(Residuals from observed)					

** Values predicted by ordinary KRIGING could not be obtained due to ANISOTROPIC variation of this property at this scale of sampling

Table 9.20.b. Comparison of predicted values of total CARBONATE by different predictive procedures at the random check points in the 100 m grid, Yelford.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING ESTIMATES
20540	43503	95002	0.40	0.78	1.01	2.70	0.70
20533	43571	95003	5.50	9.63	9.60	10.50	9.30
20511	43547	95004	2.40	3.68	3.40	2.45	4.22
20502	43512	95005	0.00	0.62	1.01	2.75	0.38
20466	43572	95006	0.20	1.22	1.22	0.40	1.10
20544	43572	95012	0.40	0.92	1.01	2.75	5.52
20531	43504	95013	1.50	0.92	1.01	2.75	1.10
20521	43495	95014	0.10	0.92	1.01	2.75	1.35
20532	43544	95015	0.10	0.92	3.40	2.45	2.13
20532	43590	95016	3.30	3.68	3.40	2.45	5.26
20524	43539	95017	2.20	3.68	3.40	2.45	2.01
20508	43491	95018	2.10	0.92	1.01	2.75	0.56
20500	43579	95019	0.10	0.85	0.77	0.85	0.52
20491	43582	95020	0.30	0.85	0.77	0.85	0.77
20461	43509	95021	0.40	0.78	0.58	0.60	0.45
20455	43576	95023	0.60	1.22	1.22	0.40	0.73
20513	43542	95024	0.40	0.78	3.40	2.45	2.70
20519	43560	95027	18.00	9.63	9.60	10.50	10.70
20502	43538	95028	0.20	0.85	1.01	2.75	0.64
20500	43542	95029	1.00	0.78	1.01	2.75	0.70
20494	43548	95030	0.60	0.85	1.01	2.75	0.95
20491	43538	95031	0.20	0.92	1.01	2.75	0.36
20487	43558	95032	0.00	0.92	1.01	2.75	1.78
20468	43557	95033	0.80	1.22	0.58	0.60	0.71
20468	43561	95034	0.20	0.62	0.58	0.60	0.62
20462	43544	95035	0.30	0.62	0.58	0.60	0.75
20461	43547	95036	0.10	0.62	0.58	0.60	1.02
20458	43560	95037	0.40	0.62	1.17	0.60	0.78
20513	43575	95041	6.10	3.26	1.90	2.10	4.08
20453	43572	95042	1.30	1.22	1.22	0.40	0.69
20560	43552	95043	0.20	0.85	0.77	0.85	1.00
20450	43588	95044	0.40	3.26	1.17	0.60	1.26
20528	43524	95045	0.60	0.62	0.58	0.60	3.90
20553	43541	95046	0.00	0.85	0.77	0.85	0.64
20454	43544	95047	0.40	0.62	1.17	0.60	0.63
20515	43567	95048	7.50	3.62	9.60	10.50	7.35
20500	43563	95049	1.60	3.62	1.90	2.10	2.13
20511	43560	95051	8.90	9.63	9.60	10.50	10.01
20503	43545	95052	1.40	3.26	3.40	2.45	1.67
ERROR IN PREDICTION:							
		SUMS OF SQUARES		=140.7833	148.1930	179.9100	139.9038
		(deviations from observed)					

Table 9.20c. Comparison of predicted values of total CARBONATE by different predictive procedures at the random check points in the 200 m grid at Yelford.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING ESTIMATES
20563	43584	95001	0.40	0.78	2.86	3.60	3.10
20540	43503	95002	0.40	0.78	1.45	0.60	1.27
20533	43571	95003	5.50	10.66	2.86	3.60	4.33
20511	43547	95004	2.40	4.87	1.45	0.60	2.53
20502	43512	95005	0.00	0.78	1.45	0.60	1.05
20466	43572	95006	0.20	1.22	0.77	0.80	2.38
20446	43558	95007	0.60	0.85	0.77	0.80	0.84
20415	43481	95008	0.00	0.78	0.77	0.80	0.85
20397	43501	95009	0.30	0.78	0.76	0.80	0.73
20382	43572	95010	3.70	3.15	2.86	22.00	6.68
20372	43563	95011	5.10	4.87	10.66	8.70	8.94
20544	43572	95012	0.40	1.27	1.45	0.60	2.71
20531	43504	95013	1.50	1.27	1.45	0.60	1.25
20521	43495	95014	0.10	1.27	1.45	0.60	1.09
20532	43544	95015	0.10	1.27	2.86	3.60	2.62
20532	43590	95016	3.30	4.87	2.86	3.60	4.66
20524	43539	95017	2.20	4.87	0.90	1.10	2.35
20508	43491	95018	2.10	1.27	1.45	0.60	1.05
20500	43579	95019	0.10	0.95	1.45	0.60	2.02
20491	43582	95020	0.30	0.95	1.45	0.60	2.35
20461	43509	95021	0.40	0.78	0.76	0.80	0.73
20475	43572	95022	0.10	0.78	0.76	0.80	2.34
20455	43576	95023	0.60	1.22	0.77	0.80	2.22
20513	43542	95024	0.40	0.78	1.45	0.60	2.28
20501	43559	95025	3.70	4.87	1.45	0.60	1.91
20481	43538	95026	0.30	0.78	0.76	0.80	0.79
20519	43560	95027	18.00	10.66	0.90	1.10	3.32
20502	43538	95028	0.20	0.95	1.45	0.60	1.53
20500	43542	95029	1.00	0.78	0.76	0.80	1.56
20494	43548	95030	0.60	0.95	0.76	0.80	1.33
20491	43538	95031	0.20	1.27	0.76	0.80	1.03
20487	43558	95032	0.00	1.27	0.76	0.80	1.34
20468	43557	95033	0.80	1.22	0.76	0.80	1.17
20468	43561	95034	0.20	0.78	0.76	0.80	1.41
20462	43544	95035	0.30	0.78	0.76	0.80	0.70
20461	43547	95036	0.10	0.78	0.76	0.80	0.74
20458	43560	95037	0.40	0.78	0.77	0.80	1.17
20418	43468	95038	0.00	0.78	0.76	0.80	0.85
20567	43605	95039	25.00	1.20	0.76	0.60	9.89
20567	43578	95040	0.50	0.95	0.90	1.10	2.30
20513	43575	95041	6.10	3.15	0.90	1.10	2.54
20453	43572	95042	1.30	1.22	0.77	0.80	1.75
20560	43552	95043	0.20	0.95	0.90	1.10	0.87
20450	43588	95044	0.40	3.15	0.77	0.80	2.48
20528	43524	95045	0.60	0.78	1.45	0.60	1.61
20553	43541	95046	0.00	0.95	2.86	3.60	1.20
20454	43544	95047	0.40	0.78	0.76	0.80	0.62
20515	43567	95048	7.50	4.87	0.90	1.10	2.80
20500	43563	95049	1.60	4.87	1.45	0.60	1.82
20515	43561	95050	8.90	10.66	0.90	1.10	2.92
20511	43560	95051	8.90	10.66	1.45	0.60	2.58
20503	43545	95052	1.40	3.15	1.45	0.60	1.88
Error in Prediction:							
		SUMS OF SQUARES =		724.379	1157.301	1490.600	652.252
		(Residuals from observed)					

Table 9.21.a. Comparison of predicted values of CLAY by different predictive procedures at the random check points in the 50 m grid at Yelford.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING ESTIMATE
20511	43547	95004	47.00	47.16	44.93	45.00	45.95
20513	43542	95024	49.00	38.08	44.93	45.00	43.86
20519	43560	95027	28.00	31.80	31.80	28.00	28.98
20502	43538	95028	48.00	44.55	49.00	38.00	45.04
20500	43542	95029	41.00	38.08	49.00	38.00	42.00
20494	43548	95030	41.00	44.55	49.00	38.00	44.00
20491	43538	95031	33.00	35.17	42.00	41.50	31.88
20487	43558	95032	33.00	35.17	42.00	41.50	34.45
20468	43557	95033	27.00	44.11	37.62	36.00	32.16
20468	43561	95034	31.00	36.85	37.62	36.00	27.86
20462	43544	95035	54.00	36.85	59.18	60.00	50.79
20461	43547	95036	54.00	36.85	59.18	60.00	55.57
20458	43560	95037	59.00	36.85	59.18	60.00	58.69
20454	43544	95047	55.00	36.85	37.62	36.00	49.89
20511	43560	95051	32.00	31.80	31.80	28.00	34.33
20503	43545	95052	49.00	44.11	44.93	45.00	44.94
Error in Prediction:							
		SUMS OF SQUARES =		1935.37	855.26	854.50	147.93
		(Deviations from observed)					

Table 9.21.b. Comparison of predicted values of CLAY by different predictive procedures at the random check points of the 100 m grid, Yelford.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING ESTIMATES
20540	43503	95002	36.00	38.79	37.58	39.50	34.39
20533	43571	95003	29.00	29.25	29.57	28.00	37.80
20511	43547	95004	47.00	39.66	40.00	44.00	40.81
20502	43512	95005	34.00	37.41	37.58	39.50	38.51
20466	43572	95006	25.00	32.25	32.25	42.00	38.10
20544	43572	95012	34.00	36.83	37.58	39.50	38.95
20531	43504	95013	30.00	36.83	37.58	39.50	34.53
20521	43495	95014	28.00	36.83	37.58	39.50	33.05
20532	43544	95015	51.00	36.83	40.00	44.00	43.00
20532	43590	95016	46.00	39.66	40.00	44.00	38.96
20524	43539	95017	48.00	39.66	40.00	44.00	42.63
20508	43491	95018	24.00	36.83	37.58	39.50	33.42
20500	43579	95019	50.00	46.30	46.21	44.00	45.48
20491	43582	95020	53.00	46.30	46.21	44.00	41.35
20461	43509	95021	33.00	38.79	39.70	35.00	39.26
20455	43576	95023	45.00	32.25	32.25	42.00	39.05
20513	43542	95024	49.00	38.79	40.00	44.00	41.85
20519	43560	95027	28.00	29.25	29.57	28.00	35.00
20502	43538	95028	48.00	46.30	37.58	39.50	40.01
20500	43542	95029	41.00	38.79	37.58	39.00	40.14
20494	43548	95030	41.00	46.30	37.58	39.00	39.20
20491	43538	95031	33.00	36.83	37.58	39.00	39.88
20487	43558	95032	33.00	36.83	37.58	39.00	37.75
20468	43557	95033	27.00	32.25	39.70	35.00	41.30
20468	43561	95034	31.00	37.41	39.70	35.00	39.25
20462	43544	95035	54.00	37.41	39.70	35.00	42.50
20461	43547	95036	54.00	37.41	39.70	35.00	43.42
20458	43560	95037	59.00	37.41	50.75	66.00	43.97
20513	43575	95041	49.00	42.42	49.00	48.00	40.67
20453	43572	95042	44.00	32.25	32.25	42.00	38.65
20560	43552	95043	49.00	46.30	46.21	44.00	45.38
20450	43588	95044	26.00	42.42	50.75	66.00	35.64
20528	43524	95045	42.00	37.41	39.70	35.00	40.17
20553	43541	95046	54.00	46.30	46.21	44.00	45.26
20454	43544	95047	55.00	37.41	50.75	66.00	43.90
20515	43567	95048	34.00	39.66	29.57	28.00	40.67
20500	43563	95049	52.00	39.66	49.00	48.00	39.92
20511	43560	95051	32.00	29.25	29.57	28.00	38.83
20503	43545	95052	49.00	42.42	40.00	44.00	39.05
ERROR IN PREDICTION:							
			SUMS OF SQUARES =	3327.80	2881.971	4054.750	2566.316
			(deviations from observed)				

Table 9.21c. Comparison of predicted values of CLAY by different predictive procedures at the randon check points in the 200 m grid at Yelford.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING ESTIMATE
20563	43584	95001	46.00	40.25	47.36	46.00	42.04
20540	43503	95002	36.00	40.25	36.56	27.50	35.15
20533	43571	95003	29.00	27.40	47.36	46.00	36.10
20511	43547	95004	47.00	36.83	36.56	27.50	44.52
20502	43512	95005	34.00	37.95	36.56	27.50	34.93
20466	43572	95006	25.00	32.25	53.66	51.00	36.89
20446	43558	95007	65.00	55.00	53.66	51.00	52.57
20415	43481	95008	45.00	37.95	53.66	51.00	45.85
20397	43501	95009	55.00	37.25	35.13	33.00	54.26
20382	43572	95010	33.00	44.50	47.36	27.00	33.58
20372	43563	95011	31.00	36.83	27.40	26.00	31.75
20544	43572	95012	34.00	33.80	36.56	27.50	41.08
20531	43504	95013	30.00	33.80	36.56	27.50	33.01
20521	43495	95014	28.00	33.80	36.56	27.50	31.64
20532	43544	95015	51.00	33.80	47.36	46.00	49.87
20532	43590	95016	46.00	36.83	47.36	46.00	39.29
20524	43539	95017	48.00	36.83	39.00	35.00	47.92
20508	43491	95018	24.00	33.80	36.56	27.50	28.49
20500	43579	95019	50.00	40.00	36.56	27.50	39.55
20491	43582	95020	53.00	40.00	36.36	27.50	36.43
20461	43509	95021	33.00	40.25	35.13	33.00	40.34
20475	43572	95022	30.00	40.25	35.13	33.00	34.36
20455	43576	95023	45.00	32.25	53.66	51.00	42.79
20513	43542	95024	49.00	40.25	36.56	27.50	44.34
20501	43559	95025	52.00	36.83	36.56	27.50	40.61
20481	43538	95026	37.00	40.25	35.13	33.00	38.85
20519	43560	95027	28.00	27.40	39.00	35.00	43.68
20502	43538	95028	48.00	40.00	36.56	27.50	40.21
20500	43542	95029	41.00	40.25	35.13	33.00	39.95
20494	43548	95030	41.00	40.00	35.13	33.00	39.03
20491	43538	95031	33.00	33.80	35.13	33.00	38.11
20487	43558	95032	33.00	33.80	35.13	33.00	37.32
20468	43557	95033	27.00	32.25	35.13	33.00	38.42
20468	43561	95034	31.00	37.95	35.13	33.00	37.68
20462	43544	95035	54.00	37.95	35.13	33.00	42.75
20461	43547	95036	54.00	37.95	35.13	33.00	43.11
20458	43560	95037	59.00	37.95	53.66	51.00	44.91
20418	43468	95038	45.00	37.95	35.13	33.00	45.11
20567	43605	95039	43.00	56.00	43.33	27.50	33.39
20567	43578	95040	49.00	40.00	39.00	35.00	44.78
20513	43575	95041	49.00	44.50	39.00	35.00	41.78
20453	43572	95042	44.00	32.25	53.66	51.00	47.36
20560	43552	95043	49.00	40.00	39.00	35.00	49.78
20450	43588	95044	26.00	44.50	53.66	51.00	34.82
20528	43524	95045	42.00	37.95	36.56	27.50	42.47
20553	43541	95046	54.00	40.00	47.36	46.00	47.21
20454	43544	95047	55.00	37.95	35.13	33.00	45.78
20515	43567	95048	34.00	36.83	39.00	35.00	42.36
20500	43563	95049	52.00	36.83	36.56	27.50	40.10
20515	43561	95050	33.00	27.40	39.00	35.00	43.52
20511	43560	95051	32.00	27.40	36.56	27.50	43.46
20503	43545	95052	49.00	44.50	36.56	27.50	41.66
Error in prediction:							
		SUMS OF SQUARES =		5208.234	6783.544	9847.25	2992.972
		(Residuals from observed)					

Table 9.22a. Comparison of predicted values of SILT by different predictive procedures at the random check points in the 50 m grid at Yelford.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING **
20511	43547	95004	32.00	23.83	24.26	19.00	
20513	43542	95024	31.00	28.50	24.26	19.00	
20519	43560	95027	10.00	13.20	13.20	10.00	
20502	43538	95028	33.00	30.88	30.41	37.00	
20500	43542	95029	30.00	28.50	30.41	37.00	
20494	43548	95030	35.00	30.88	30.41	37.00	
20491	43538	95031	28.00	28.29	26.50	25.00	
20487	43558	95032	37.00	28.29	26.50	25.00	
20468	43557	95033	29.00	25.77	28.59	27.00	
20468	43561	95034	33.00	28.55	28.59	27.00	
20462	43544	95035	25.00	28.55	22.90	24.00	
20461	43547	95036	23.00	28.55	22.90	24.00	
20458	43560	95037	25.00	28.55	22.90	24.00	
20454	43544	95047	27.00	28.55	28.59	27.00	
20511	43560	95051	19.00	13.20	13.20	10.00	
20503	43545	95052	32.00	25.77	24.96	19.00	
Error in prediction:							
		SUMS OF SQUARES =	344.00	370.195	828.00	**	
		(Residuals form observed)					

** Predicted values by ordinary KRIGING could not be obtained due to ANISOTROPIC variation of this property at this scale of sampling.

Table 9.22 b. Comparison of predicted values of SILT by different predictive procedures at the random check points in the 100m m grid, Yelford.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING ESTIMATES
20540	43503	95002	32.00	27.24	26.24	23.50	26.98
20533	43571	95003	10.00	11.25	11.28	10.00	9.38
20511	43547	95004	32.00	24.80	24.96	25.50	26.31
20502	43512	95005	34.00	28.47	26.24	23.50	28.64
20466	43572	95006	32.00	25.75	25.75	29.00	23.87
20544	43572	95012	26.00	24.16	26.24	23.50	14.22
20531	43504	95013	24.00	24.16	26.24	23.50	25.85
20521	43495	95014	29.00	24.16	26.24	23.50	24.58
20532	43544	95015	29.00	24.16	24.96	25.50	25.00
20532	43590	95016	22.00	24.80	24.96	25.50	21.80
20524	43539	95017	24.00	24.80	24.96	25.50	25.58
20508	43491	95018	30.00	24.16	26.24	23.50	27.91
20500	43579	95019	30.00	29.69	28.85	27.50	29.30
20491	43582	95020	27.00	29.69	28.85	27.50	28.54
20461	43509	95021	26.00	27.24	28.76	28.00	23.25
20455	43576	95023	23.00	25.75	25.75	29.00	23.37
20513	43542	95024	31.00	27.24	24.96	25.50	27.44
20519	43560	95027	10.00	11.25	11.28	10.00	13.57
20502	43538	95028	33.00	29.69	26.24	23.50	28.58
20500	43542	95029	30.00	27.24	26.24	23.50	29.25
20494	43548	95030	35.00	29.69	26.24	23.50	30.77
20491	43538	95031	28.00	24.16	26.24	23.50	28.04
20487	43558	95032	37.00	24.16	26.24	23.50	30.70
20468	43557	95033	29.00	25.75	28.76	28.00	25.96
20468	43561	95034	33.00	28.47	28.76	28.00	25.35
20462	43544	95035	25.00	28.47	28.76	28.00	26.89
20461	43547	95036	23.00	28.47	28.76	28.00	25.30
20458	43560	95037	25.00	28.47	22.50	21.00	21.91
20513	43575	95041	22.00	24.07	28.50	27.00	23.84
20453	43572	95042	26.00	25.75	25.75	29.00	22.45
20560	43552	95043	32.00	29.69	28.85	27.50	28.71
20450	43588	95044	30.00	24.07	22.50	21.00	26.86
20528	43524	95045	31.00	28.47	28.76	28.00	27.85
20553	43541	95046	28.00	29.69	28.85	27.50	32.10
20454	43544	95047	27.00	28.47	22.50	21.00	26.75
20515	43567	95048	13.00	24.80	11.28	10.00	18.95
20500	43563	95049	34.00	24.80	28.50	27.00	29.43
20511	43560	95051	19.00	11.25	11.28	10.00	20.91
20503	43545	95052	32.00	24.07	24.96	25.50	29.08
ERROR IN PREDICTION: SUMS OF SQUARES = (deviations from observed)				985.244	898.429	1321.250	669.105

Table 9.22c. Comparison of predicted values of SILT by different predictive procedures at the random check points in the 200 m grid at Yelford.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING ESTIMATE
20563	43584	95001	27.00	25.00	24.72	28.00	23.52
20540	43503	95002	32.00	25.00	26.34	26.50	26.29
20533	43571	95003	10.00	13.80	24.72	28.00	22.93
20511	43547	95004	32.00	22.00	26.34	26.50	27.52
20502	43512	95005	34.00	26.22	26.34	26.50	26.59
20466	43572	95006	32.00	25.75	21.33	21.00	25.72
20446	43558	95007	20.00	20.23	21.33	21.00	24.27
20415	43481	95008	25.00	26.22	21.33	21.00	23.68
20397	43501	95009	24.00	26.22	26.66	30.50	22.40
20382	43572	95010	17.00	22.83	24.72	16.00	18.75
20372	43563	95011	19.00	22.00	13.80	14.00	17.91
20544	43572	95012	26.00	26.80	26.34	26.50	23.33
20531	43504	95013	24.00	26.80	26.34	26.50	26.45
20521	43495	95014	29.00	26.80	26.34	26.50	26.24
20532	43544	95015	29.00	26.80	24.72	28.00	26.27
20532	43590	95016	22.00	22.00	24.72	28.00	23.72
20524	43539	95017	24.00	22.00	33.00	33.00	27.02
20508	43491	95018	30.00	26.80	26.34	26.50	26.12
20500	43579	95019	30.00	29.16	26.34	26.50	27.97
20491	43582	95020	27.00	29.16	26.34	26.50	27.85
20461	43509	95021	26.00	25.00	26.66	30.00	25.20
20475	43572	95022	33.00	25.00	26.66	30.00	26.86
20455	43576	95023	23.00	25.75	21.33	21.00	24.38
20513	43542	95024	31.00	25.00	26.34	26.50	27.58
20501	43559	95025	32.00	22.00	26.34	26.50	28.19
20481	43538	95026	25.00	25.00	26.66	30.50	27.32
20519	43560	95027	10.00	13.80	33.00	33.00	25.88
20502	43538	95028	33.00	29.16	26.34	26.50	27.97
20500	43542	95029	30.00	25.00	26.66	30.50	28.10
20494	43548	95030	35.00	29.16	26.66	30.50	28.45
20491	43538	95031	28.00	26.80	26.66	30.50	27.60
20487	43558	95032	37.00	26.80	26.66	30.50	28.52
20468	43557	95033	29.00	25.75	26.66	30.50	26.72
20468	43561	95034	33.00	26.22	26.66	30.50	26.49
20462	43544	95035	25.00	26.22	26.66	30.50	26.49
20461	43547	95036	23.00	26.22	26.66	30.50	26.35
20458	43560	95037	25.00	26.22	21.33	21.00	25.36
20418	43468	95038	26.00	26.22	26.66	30.50	23.79
20567	43605	95039	23.00	21.00	26.00	26.50	21.27
20567	43578	95040	30.00	29.16	33.00	33.00	24.41
20513	43575	95041	22.00	22.83	33.00	33.00	26.63
20453	43572	95042	26.00	25.75	21.33	21.00	24.13
20560	43552	95043	32.00	29.16	33.00	33.00	26.07
20450	43588	95044	30.00	22.83	21.33	21.00	23.98
20528	43524	95045	31.00	26.22	26.34	26.50	27.43
20553	43541	95046	28.00	29.16	24.72	28.00	27.25
20454	43544	95047	27.00	26.22	26.66	30.50	25.82
20515	43567	95048	13.00	22.00	33.00	33.00	26.37
20500	43563	95049	34.00	22.00	26.34	26.50	28.28
20515	43561	95050	1.00	13.80	33.00	33.00	26.47
20511	43560	95051	19.00	13.80	26.34	26.50	27.15
20503	43545	95052	32.00	22.83	26.34	26.50	28.06
Error in Prediction:							
		SUMS OF SQUARES =		1416.221	3490.620	3533.250	2034.909
		(Residuals from observed)					

Table 9.23a. Comparison of predicted values of COARSE SAND by different predictive procedures at the random check points in the 50 m grid at Yelford.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING **
20511	43547	95004	8.00	12.50	12.40	23.00	
20513	43542	95024	6.00	4.25	12.40	23.00	
20519	43560	95027	44.00	34.00	34.00	41.00	
20502	43538	95028	3.00	4.11	4.41	4.00	
20500	43542	95029	8.00	4.25	4.41	4.00	
20494	43548	95030	5.00	4.11	4.41	4.00	
20491	43538	95031	8.00	5.94	9.00	9.00	
20487	43558	95032	6.00	5.94	9.00	9.00	
20468	43557	95033	13.00	9.66	2.88	3.50	
20468	43561	95034	3.00	3.00	2.88	3.50	
20462	43544	95035	2.00	3.00	1.63	2.00	
20461	43547	95036	2.00	3.00	1.63	2.00	
20458	43560	95037	2.00	3.00	1.63	2.00	
20454	43544	95047	1.00	3.00	2.88	3.50	
20511	43560	95051	26.00	34.00	34.00	41.00	
20503	43545	95052	6.00	9.66	12.40	23.00	
Error in prediction:							
		SUMS OF SQUARES =		239.197	396.878	1161.75	**
		(Residuals from observed)					

** Predicted values by ordinary KRIGING could not be obtained due to ANISOTROPIC variation of this property at this scale of sampling.

Table 9.23b. Comparison of predicted values of COARSE SAND by different predictive procedures at the random check points in the 100 m grid at Yelford.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING ESTIMATES
20540	43503	95002	6.00	4.93	4.90	5.00	5.72
20533	43571	95003	43.00	39.35	39.28	41.00	48.06
20511	43547	95004	8.00	10.46	10.75	11.50	11.76
20502	43512	95005	6.00	3.82	4.90	5.00	6.24
20466	43572	95006	6.00	7.00	7.00	6.00	5.54
20544	43572	95012	9.00	5.51	4.90	5.00	32.59
20531	43504	95013	8.00	5.51	4.90	5.00	5.55
20521	43495	95014	3.00	5.51	4.90	5.00	4.07
20532	43544	95015	7.00	5.51	10.75	11.50	9.49
20532	43590	95016	15.00	10.46	10.75	11.50	8.67
20524	43539	95017	11.00	10.46	10.75	11.50	9.22
20508	43491	95018	4.00	5.51	4.90	5.00	2.93
20500	43579	95019	5.00	4.84	5.28	5.50	3.72
20491	43582	95020	5.00	4.84	5.28	5.50	6.24
20461	43509	95021	3.00	4.93	3.52	2.00	3.89
20455	43576	95023	6.00	7.00	7.00	6.00	5.22
20513	43542	95024	6.00	4.93	10.75	11.50	8.64
20519	43560	95027	44.00	39.35	39.28	41.00	38.98
20502	43538	95028	3.00	4.84	4.90	5.00	5.53
20500	43542	95029	8.00	4.93	4.90	5.00	6.20
20494	43548	95030	5.00	4.84	4.90	5.00	4.53
20491	43538	95031	8.00	5.51	4.90	5.00	7.44
20487	43558	95032	6.00	5.51	4.90	5.00	4.64
20468	43557	95033	13.00	7.00	3.52	2.00	3.44
20468	43561	95034	3.00	3.82	3.52	2.00	3.59
20462	43544	95035	2.00	3.82	3.52	2.00	1.44
20461	43547	95036	2.00	3.82	3.52	2.00	1.75
20458	43560	95037	2.00	3.82	2.75	1.00	1.84
20513	43575	95041	16.00	10.42	6.00	8.00	12.47
20453	43572	95042	5.00	7.00	7.00	6.00	3.40
20560	43552	95043	4.00	4.84	5.28	5.50	3.69
20450	43588	95044	8.00	10.42	2.75	1.00	5.94
20528	43524	95045	8.00	3.82	3.52	2.00	10.23
20553	43541	95046	5.00	4.84	5.28	5.50	4.59
20454	43544	95047	1.00	3.82	2.75	1.00	1.50
20515	43567	95048	33.00	10.46	39.28	41.00	26.35
20500	43563	95049	5.00	10.46	6.00	8.00	5.28
20511	43560	95051	26.00	39.35	39.28	41.00	27.97
20503	43545	95052	6.00	10.42	10.75	11.50	6.04
ERROR IN PREDICTION: SUMS OF SQUARES = (deviations from observed)				972.555	649.262	748.500	868.353

Table 9.23c. Comparison of predicted values of COARSE SAND by different predictive procedures at the random check points in the 200 m grid at Yelford.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING ESTIMATE
20563	43584	95001	5.00	5.08	10.00	12.00	11.06
20540	43503	95002	6.00	5.08	5.73	5.50	7.51
20533	43571	95003	43.00	32.60	10.00	12.00	17.90
20511	43547	95004	8.00	15.91	5.73	5.50	9.71
20502	43512	95005	6.00	2.86	5.73	5.50	5.80
20466	43572	95006	6.00	7.00	2.20	2.00	5.45
20446	43558	95007	2.00	3.00	2.20	2.00	2.66
20415	43481	95008	2.00	2.86	2.20	2.00	2.44
20397	43501	95009	2.00	2.86	3.00	2.00	2.38
20382	43572	95010	15.00	6.50	10.00	28.00	14.47
20372	43563	95011	19.00	15.91	32.60	29.00	17.52
20544	43572	95012	9.00	5.60	5.73	5.50	13.28
20531	43504	95013	8.00	5.60	5.73	5.50	7.14
20521	43495	95014	3.00	5.60	5.73	5.50	6.43
20532	43544	95015	7.00	5.60	10.00	12.00	10.96
20532	43590	95016	15.00	15.91	10.00	12.00	11.61
20524	43539	95017	11.00	15.91	4.00	5.00	9.62
20508	43491	95018	4.00	5.60	5.73	5.50	5.26
20500	43579	95019	5.00	5.33	5.73	5.50	7.44
20491	43582	95020	5.00	5.33	5.73	5.50	7.05
20461	43509	95021	3.00	5.08	3.00	2.00	2.85
20475	43572	95022	6.00	5.08	3.00	2.00	5.93
20455	43576	95023	6.00	7.00	2.20	2.00	5.02
20513	43542	95024	6.00	5.08	5.73	5.50	9.20
20501	43559	95025	8.00	15.91	5.73	5.50	7.89
20481	43538	95026	6.00	5.08	3.00	2.00	4.69
20519	43560	95027	44.00	32.60	4.00	5.00	13.09
20502	43538	95028	3.00	5.33	5.73	5.50	7.04
20500	43542	95029	8.00	5.08	3.00	2.00	7.18
20494	43548	95030	5.00	5.33	3.00	2.00	6.31
20491	43538	95031	8.00	5.60	3.00	2.00	5.73
20487	43558	95032	6.00	5.60	3.00	2.00	5.61
20468	43557	95033	13.00	7.00	3.00	2.00	4.19
20468	43561	95034	3.00	2.86	3.00	2.00	4.45
20462	43544	95035	2.00	2.86	3.00	2.00	3.17
20461	43547	95036	2.00	2.86	3.00	2.00	3.21
20458	43560	95037	2.00	2.86	2.20	2.00	3.61
20418	43468	95038	2.00	2.86	3.00	2.00	2.71
20567	43605	95039	23.00	8.00	6.33	5.50	18.79
20567	43578	95040	5.00	5.33	4.00	5.00	9.65
20513	43575	95041	16.00	6.50	4.00	5.00	10.14
20453	43572	95042	5.00	7.00	2.20	2.00	4.22
20560	43552	95043	4.00	5.33	4.00	5.00	7.58
20450	43588	95044	8.00	6.50	2.20	2.00	6.10
20528	43524	95045	8.00	2.86	5.73	5.50	7.28
20553	43541	95046	5.00	5.33	10.00	12.00	7.59
20454	43544	95047	1.00	6.50	3.00	2.00	2.58
20515	43567	95048	33.00	15.91	4.00	5.00	11.62
20500	43563	95049	5.00	15.91	5.73	5.50	7.61
20515	43561	95050	34.00	32.60	4.00	5.00	11.73
20511	43560	95051	26.00	32.60	5.73	5.50	10.38
20503	43545	95052	6.00	6.50	5.73	5.50	7.89
Error in Prediction:							
		SUMS OF SQUARES =		1416.85	5917.295	5776.25	3133.215
		(Residuals from observed)					

Table 9.24a. Comparison of predicted values of ORGANIC MATTER content by different predictive procedures at the random check points in the 50 m grid at Yelford.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING ESTIMATES
20511	43547	95004	8.40	7.18	6.92	4.70	6.65
20513	43542	95024	7.30	6.79	6.92	4.70	6.36
20519	43560	95027	3.80	5.48	5.48	4.70	5.13
20502	43538	95028	8.30	7.86	7.43	8.40	6.74
20500	43542	95029	6.10	6.79	7.43	8.40	6.78
20494	43548	95030	12.30	7.86	7.43	8.40	7.86
20491	43538	95031	4.90	7.04	5.35	5.70	5.55
20487	43558	95032	13.00	7.04	5.35	5.70	9.65
20468	43557	95033	11.60	7.50	8.47	7.50	10.52
20468	43561	95034	14.90	8.20	8.47	7.50	10.22
20462	43544	95035	5.30	8.20	5.81	5.80	7.06
20461	43547	95036	5.40	8.20	5.81	5.80	7.00
20458	43560	95037	6.30	8.20	5.81	5.80	6.65
20454	43544	95047	5.40	8.20	8.47	7.50	5.60
20511	43560	95051	7.70	5.48	5.48	4.70	7.05
20503	43545	95052	7.90	7.50	6.92	4.70	7.41
Error in Prediction:							
		SUMS OF SQUARES =		159.543	157.248	191.580	69.511
		(Deviations from observed)					

Table 9.24b. Comparison of predicted values of ORGANIC MATTER content by different predictive procedures at the random check points in the 100 m grid at Yelford.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING ESTIMATES
20540	43503	95002	7.90	6.97	6.30	6.30	7.40
20533	43571	95003	3.40	4.32	4.31	4.70	3.94
20511	43547	95004	8.40	6.61	6.79	5.90	6.62
20502	43512	95005	12.60	7.45	6.30	6.00	7.54
20466	43572	95006	13.70	8.22	8.22	6.10	8.23
20544	43572	95012	8.00	6.18	6.30	6.30	4.40
20531	43504	95013	5.90	6.18	6.30	6.30	8.05
20521	43495	95014	2.20	6.18	6.30	6.30	8.15
20532	43544	95015	6.60	6.18	6.79	5.90	5.66
20532	43590	95016	5.60	6.61	6.79	5.90	5.15
20524	43539	95017	4.90	6.61	6.79	5.90	6.02
20508	43491	95018	10.00	6.18	6.30	6.30	8.24
20500	43579	95019	9.20	7.52	6.70	8.70	8.61
20491	43582	95020	9.10	7.52	6.70	8.70	8.89
20461	43509	95021	4.80	6.97	7.74	4.90	4.66
20455	43576	95023	5.70	8.22	8.22	6.10	7.56
20513	43542	95024	7.30	6.97	6.79	5.90	6.39
20519	43560	95027	3.80	4.32	4.31	4.70	5.20
20502	43538	95028	8.30	7.52	6.30	6.30	6.75
20500	43542	95029	6.10	6.97	6.30	6.30	7.15
20494	43548	95030	12.30	7.52	6.30	6.30	8.42
20491	43538	95031	4.90	6.18	6.30	6.30	6.82
20487	43558	95032	13.00	6.18	6.30	6.30	9.51
20468	43557	95033	11.60	8.22	7.74	4.90	9.00
20468	43561	95034	14.90	7.45	7.74	4.90	8.93
20462	43544	95035	5.30	7.45	7.74	4.90	7.23
20461	43547	95036	5.40	7.45	7.74	4.90	7.20
20458	43560	95037	6.30	7.45	6.67	6.60	6.67
20513	43575	95041	6.40	6.07	8.55	7.60	6.51
20453	43572	95042	5.90	8.22	8.22	6.10	6.92
20560	43552	95043	5.00	7.52	6.70	8.70	6.12
20450	43588	95044	9.00	6.07	6.67	6.60	8.43
20528	43524	95045	9.60	7.45	7.74	4.90	6.87
20553	43541	95046	5.40	7.52	6.70	8.70	7.37
20454	43544	95047	5.40	7.45	6.67	6.60	5.93
20515	43567	95048	4.30	6.61	4.31	4.70	6.03
20500	43563	95049	8.10	6.61	8.55	7.60	9.18
20511	43560	95051	7.70	4.32	4.31	4.70	6.96
20503	43545	95052	7.90	6.07	6.79	5.90	7.53
ERROR IN PREDICTION:							
		SUMS OF SQUARES =		321.684	344.433	451.340	227.350
		(Deviations from observed)					

Table 9.24c. Comparison of predicted values of ORGANIC MATTER content by different predictive procedures at the random check points in the 200 m grid at Yelford.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING ESTIMATE
20563	43584	95001	5.00	6.14	5.49	6.60	4.64
20540	43503	95002	7.90	6.14	6.70	4.30	7.61
20533	43571	95003	3.40	4.28	5.49	6.60	4.78
20511	43547	95004	8.40	5.70	6.70	4.30	6.66
20502	43512	95005	12.60	6.89	6.70	4.30	6.80
20466	43572	95006	13.70	8.22	5.34	5.00	9.36
20446	43558	95007	5.20	5.34	5.34	5.00	6.25
20415	43481	95008	5.00	6.89	5.34	5.00	5.02
20397	43501	95009	6.40	6.89	7.00	7.55	6.52
20382	43572	95010	3.60	5.05	5.49	4.20	4.84
20372	43563	95011	4.20	5.70	4.28	3.60	4.61
20544	43572	95012	8.00	6.62	6.70	4.30	4.99
20531	43504	95013	5.90	6.62	6.70	4.30	8.03
20521	43495	95014	2.20	6.62	6.70	4.30	7.77
20532	43544	95015	6.60	6.62	5.49	6.60	5.67
20532	43590	95016	5.60	5.70	5.49	6.60	5.76
20524	43539	95017	4.90	5.70	8.10	12.30	6.15
20508	43491	95018	10.00	6.62	6.70	4.30	7.70
20500	43579	95019	9.20	7.81	6.70	4.30	10.57
20491	43582	95020	9.10	7.81	6.70	4.30	11.47
20461	43509	95021	4.80	6.14	7.00	7.55	5.04
20475	43572	95022	9.80	6.14	7.00	7.55	10.52
20455	43576	95023	5.70	8.22	5.34	5.00	7.81
20513	43542	95024	7.30	6.14	6.70	4.30	6.52
20501	43559	95025	7.20	5.70	6.70	4.30	8.84
20481	43538	95026	3.80	6.14	7.00	7.55	7.60
20519	43560	95027	3.80	4.28	8.10	12.30	6.00
20502	43538	95028	8.30	7.81	6.70	4.30	7.05
20500	43542	95029	6.10	6.14	7.00	7.55	7.46
20494	43548	95030	12.30	7.81	7.00	7.55	8.74
20491	43538	95031	4.90	6.62	7.00	7.55	7.23
20487	43558	95032	13.00	6.62	7.00	7.55	10.40
20468	43557	95033	11.60	8.22	7.00	7.55	9.77
20468	43561	95034	14.90	6.89	7.00	7.55	9.69
20462	43544	95035	5.30	6.89	7.00	7.55	7.73
20461	43547	95036	5.40	6.89	7.00	7.55	7.97
20458	43560	95037	6.30	6.89	5.34	5.00	7.98
20418	43468	95038	5.00	6.89	7.00	7.55	4.10
20567	43605	95039	4.00	5.20	6.76	4.30	4.61
20567	43578	95040	4.50	7.81	8.10	12.30	4.80
20513	43575	95041	6.40	5.05	8.10	12.30	8.03
20453	43572	95042	5.90	8.22	5.34	5.00	7.28
20560	43552	95043	5.00	7.81	8.10	12.30	5.23
20450	43588	95044	9.00	5.05	5.34	5.00	7.86
20528	43524	95045	9.60	6.89	6.70	4.30	7.26
20553	43541	95046	5.40	7.81	5.49	6.60	6.79
20454	43544	95047	5.40	6.89	7.00	7.55	6.52
20515	43567	95048	4.30	5.70	8.10	12.30	7.07
20500	43563	95049	8.10	5.70	6.70	4.30	9.33
20515	43561	95050	4.00	4.28	8.10	12.30	6.62
20511	43560	95051	7.70	4.28	6.70	4.30	7.19
20503	43545	95052	7.90	5.05	6.70	4.30	7.37
Error in Prediction:							
		SUMS OF SQUARES =		386.160	476.051	999.565	254.489
		(Residuals from observed)					

Second, in order to derive conclusions valid in a practical context, this should be seen against a background of the sampling effort that can be afforded in routine soil survey.

Amongst the soil survey output, the typical profile of the mapping unit was at a disadvantage in the comparison since the values predicted from them were based only on one observation, that of the typical profile itself. The predictive capacity of those values is also strongly dependent on the criteria used to select the typical profile.

Kriging, mapping unit means and class means used the same sampling effort (i.e. density of point-data), therefore the comparison between them can be said to be fair. However, in a practical context, only in exceptional circumstances class means and map unit means would be calculated from such a wealth of point-data as they were in these trials. Usually, in a practical situation in routine survey, mean values is all that is readily available from the survey's memoir or else the number of point-samples reported is considerably smaller to those used in these trials to calculate the means. It is in this sense that the comparison is unfair to Kriging, which in contrast, in an ordinary situation, it is known to require relatively high densities of point-samples to be applicable. Such densities are the ordinary situation associated with the use of Kriging, and this is assumed to be known by any surveyor undertaking its use in spatial prediction.

On the other hand, in most instances success in prediction by these techniques can be related to the structures of the semi-variogram. Hence, this can be used as a reconnaissance tool to decide, purely from the point of view of making the best predictions, in which cases to use Kriging or else to make a soil map or even to attempt a combination of both.

So, it can be seen from tables 9.17a-c that for soil pH Kriging was the best predictive tool at the sampling scale corresponding to the 50 m grid, although the improvement over map unit means is only marginal. This can be interpreted in terms of the shape of the semi-variogram (fig. 9.5a) of pH at this grid spacing. The semi-variogram clearly shows spatial dependence, but this is strong only in the first few lags and it weakens as lag distance increases even before reaching the sill

the points representing the semi-variances fluctuate wider. The success of map unit means as predictors also indicate that the criteria employed in the delineation of soil boundaries to partition variability of pH in the landscape, at this scale, were adequate.

At the 100 m grid (table 9.17b) means from map units were the best predictors of soil pH, confirming the adequacy of the criteria for delineation of such mapping units too.

Even class means were marginally better as predictors than Kriging. The apparent failure of Kriging can be explained by the semi-variogram (fig 9.5b). Only a relatively small portion of the total variability at this scale is spatially dependent which is weak. The random variation (Nugget variance) is almost as large as the spatially related variance. This considerably large random component makes the Kriging estimates less accurate.

Moreover, as the sampling spacing increased the random component in the variation of this property becomes larger. Thus, at the 200 m grid the semi-variogram showed "Pure Nugget effect" (fig. 9.5c) indicating a total absence of spatial dependence. The best estimate that can be found in this case is the general mean of the data from the corresponding grid. This is included as the Kriging estimate for all the target points in table 9.17c. This general mean is less accurate as a predictor than means from smaller strata as generated by classes or mapping units, where variability had been narrowed down.

The results of the comparison for soil pH are summarized graphically in figure 9.14. In order to make the RSS comparable between different sampling spacings which contained different number of predicted target points, the Residual Mean Squares (RMS = RSS/No of target points) were plotted instead, against the sampling spacings of the three grids. These are shown in figures 9.14 to 9.21 for the 8 soil properties studied. In the legends of these plots, where it reads "class" and "mapunit" it refers to predictions by class means and map unit means. "Profile" refers to predictions by values from the typical profile in the mapping unit.

So, on the whole, it appears that the best predictive tool for soil pH is the mean value from mapping units, and the worst

Fig. 9.14 pH

Comparison of the predictive value of Soil Survey output and spatial interpolation by Kriging

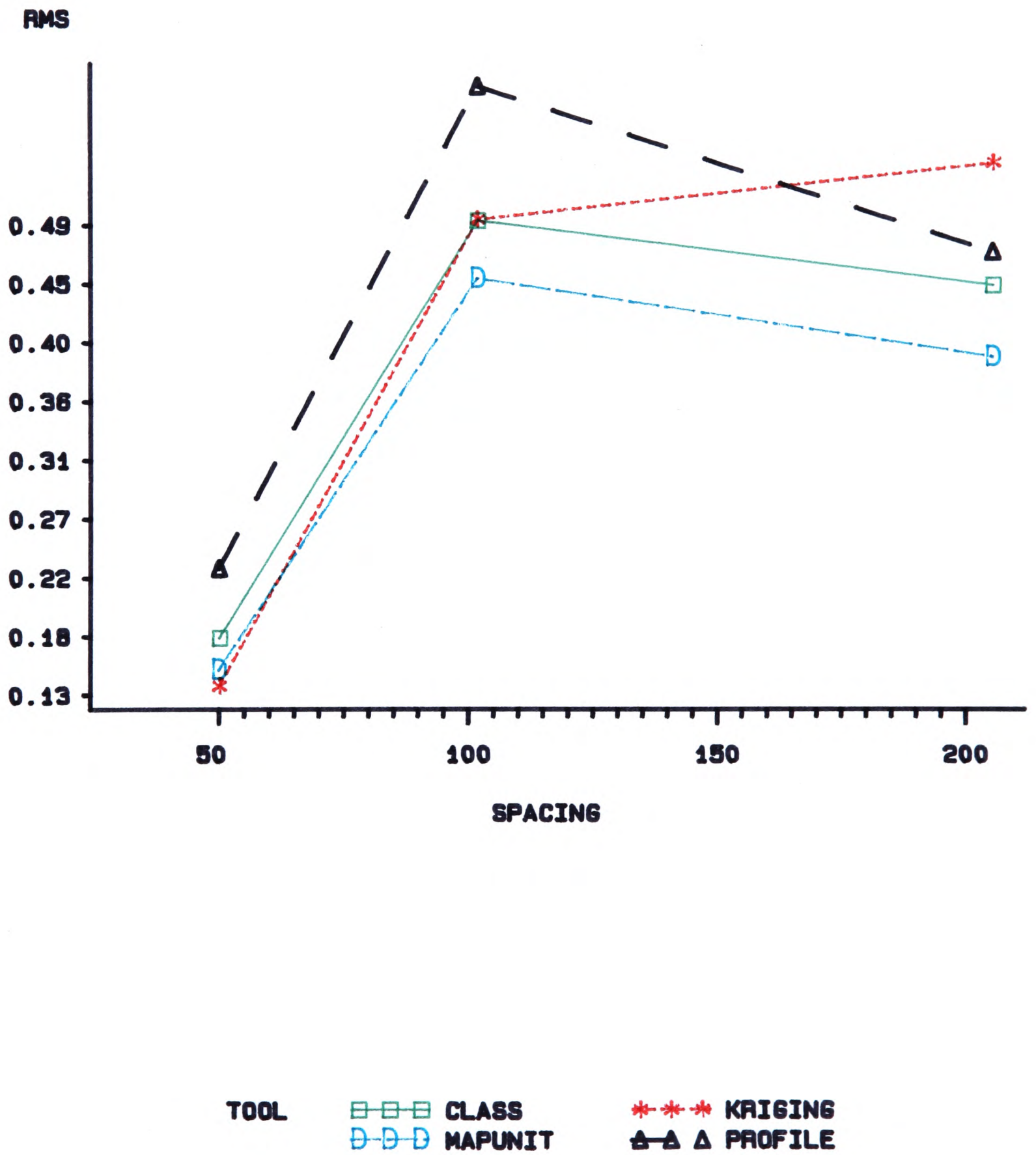


Fig. 9.15 POTASSIUM

Comparison of the predictive value of Soil Survey output and spatial interpolation by Kriging

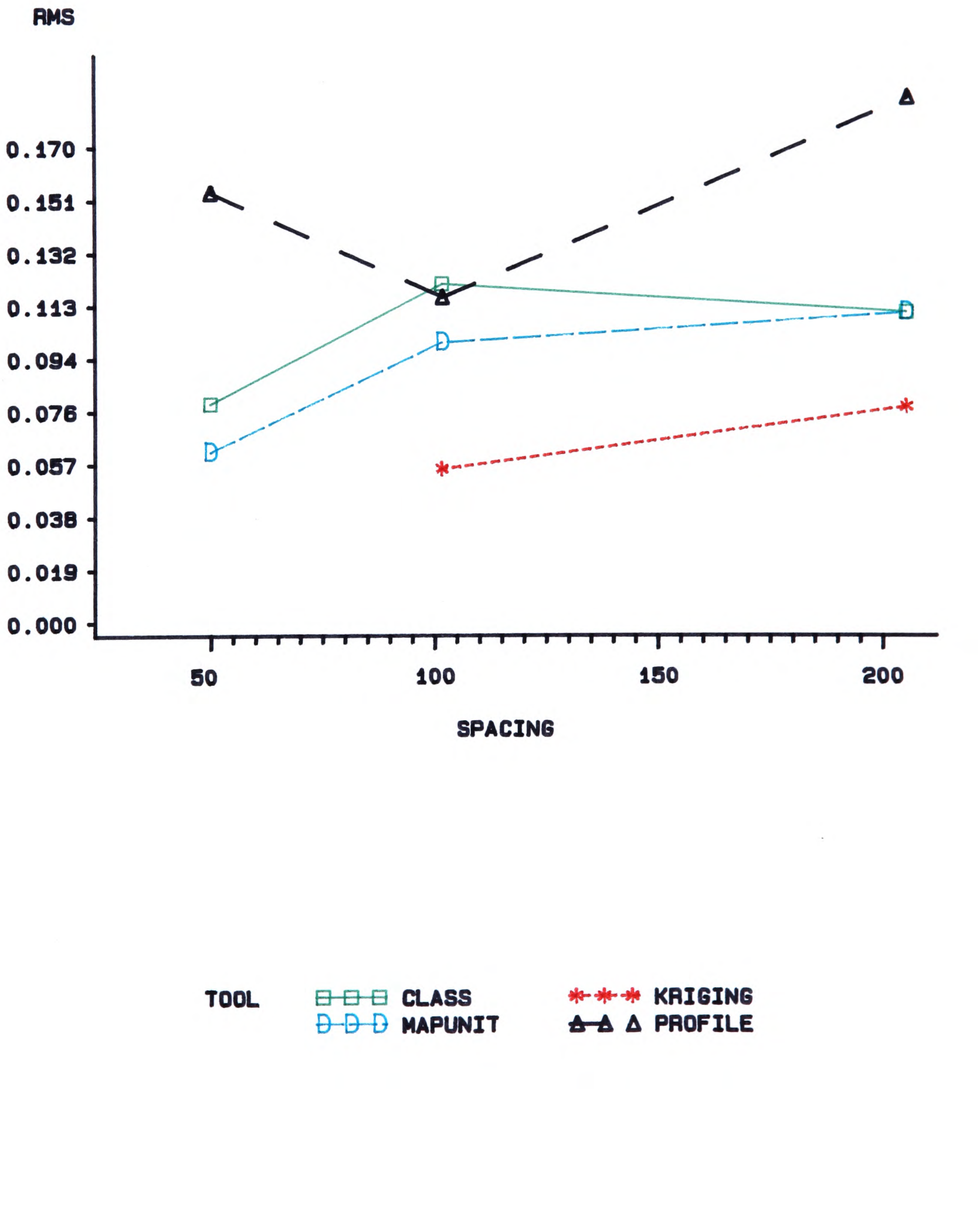


Fig. 9.16

MAGNESIUM

Comparison of the predictive value of Soil Survey output and spatial interpolation by Kriging

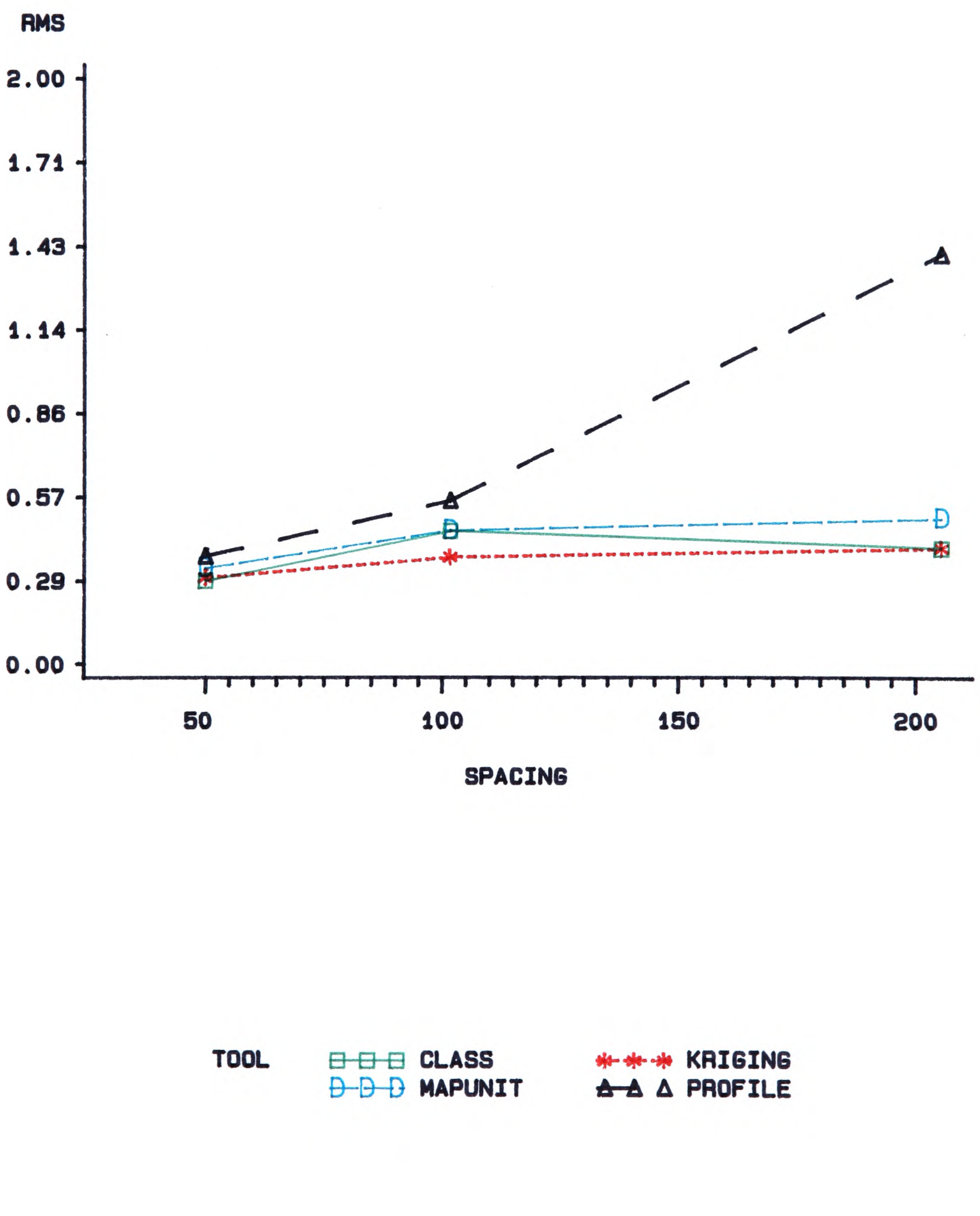


Fig. 9.17 CARBONATE

Comparison of the predictive value of Soil Survey output and spatial interpolation by Kriging

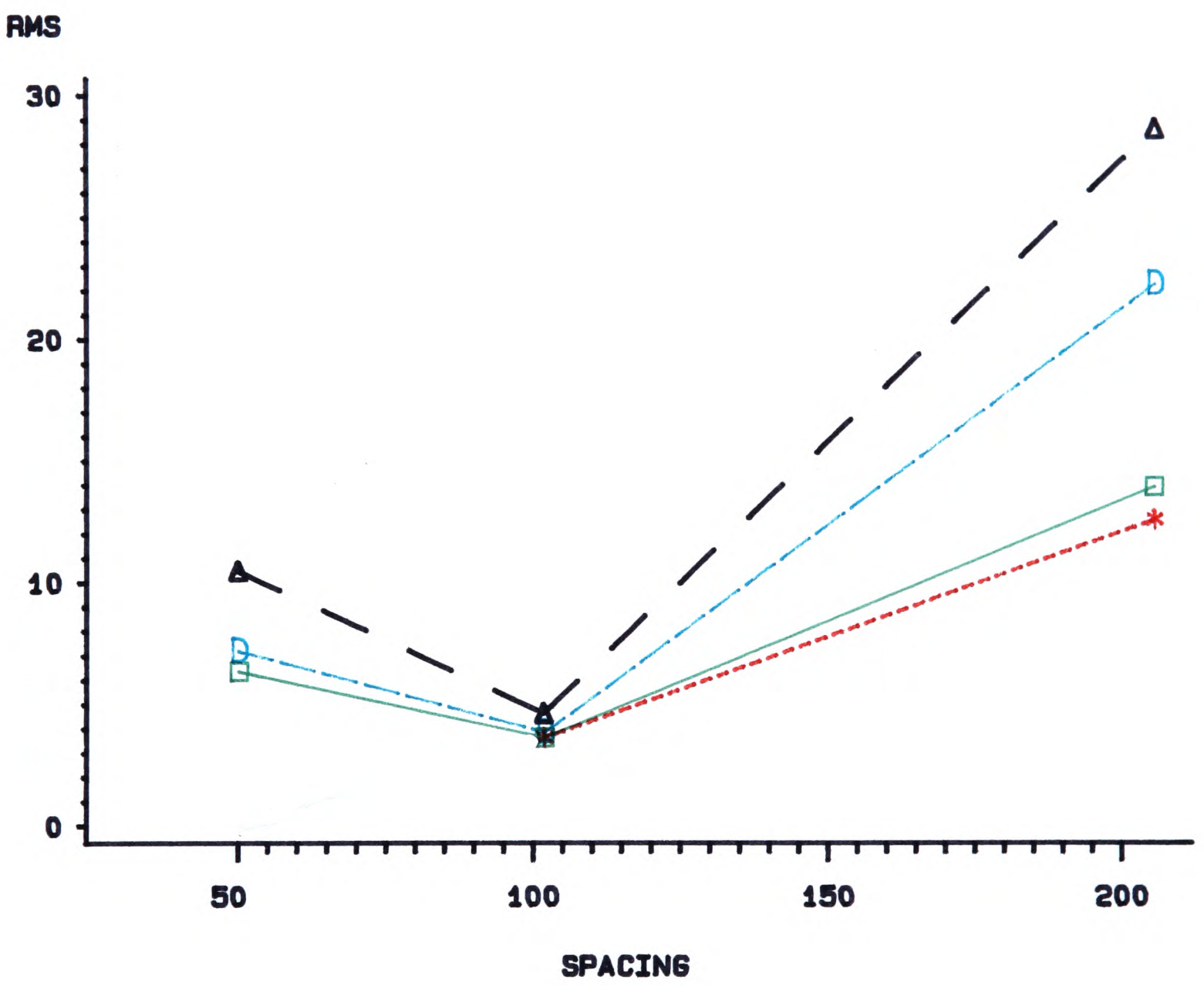


Fig. 9.18

CLAY

Comparison of the predictive value of Soil Survey output and spatial interpolation by Kriging

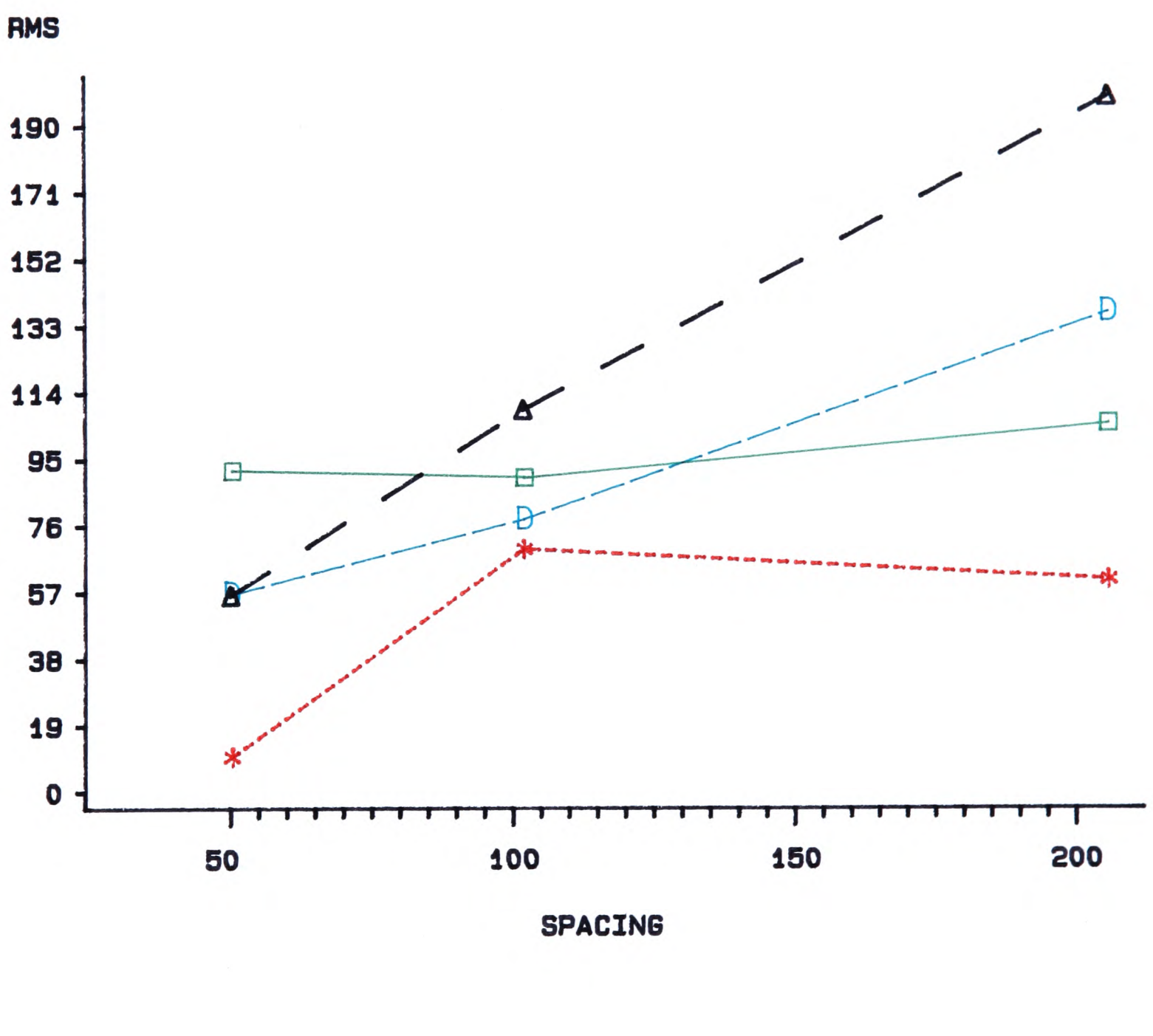


Fig. 9.19

SILT

Comparison of the predictive value of Soil Survey output and spatial interpolation by Kriging

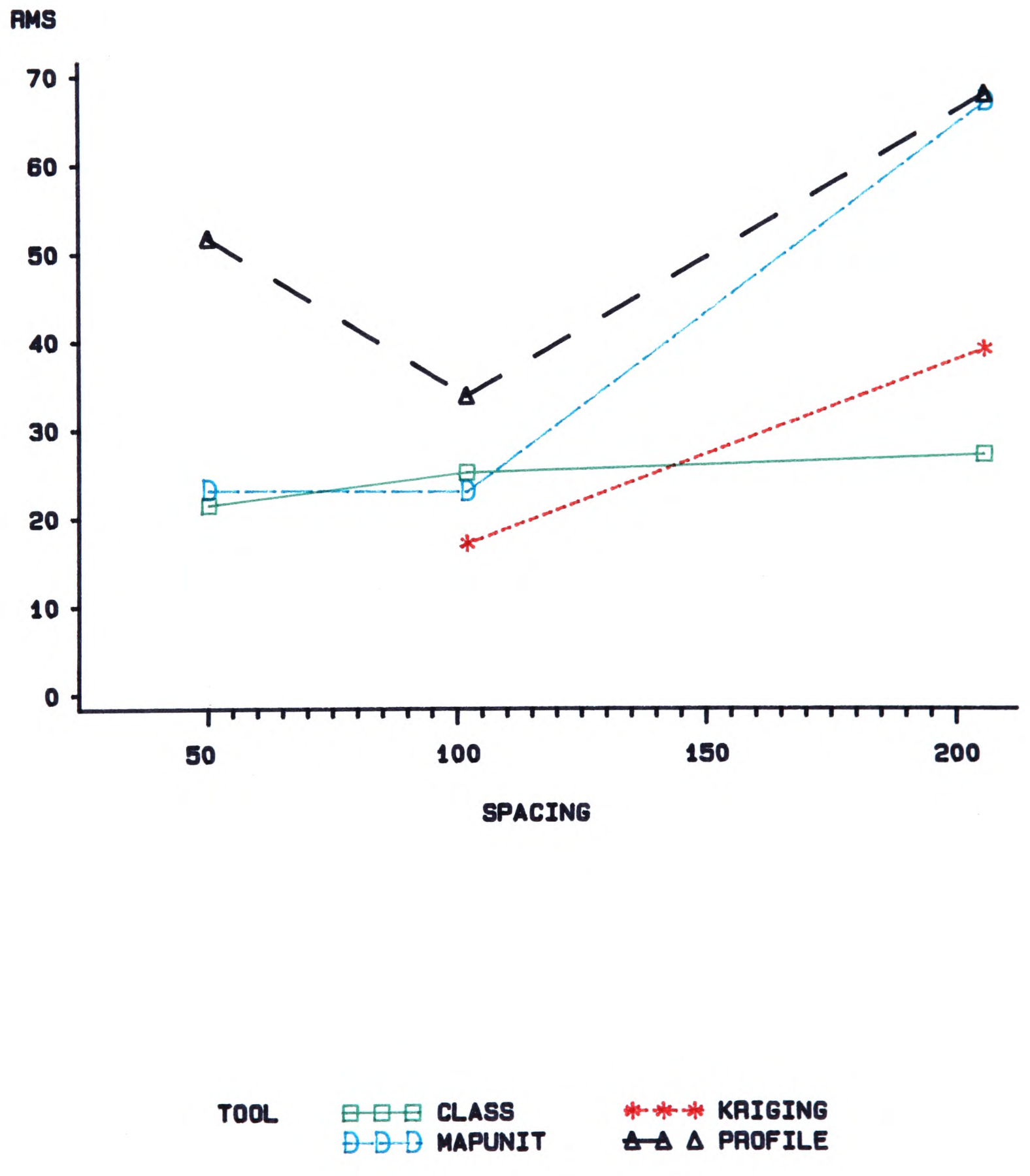


Fig. 9.20 COARSE SAND

Comparison of the predictive value of Soil Survey output and spatial interpolation by Kriging

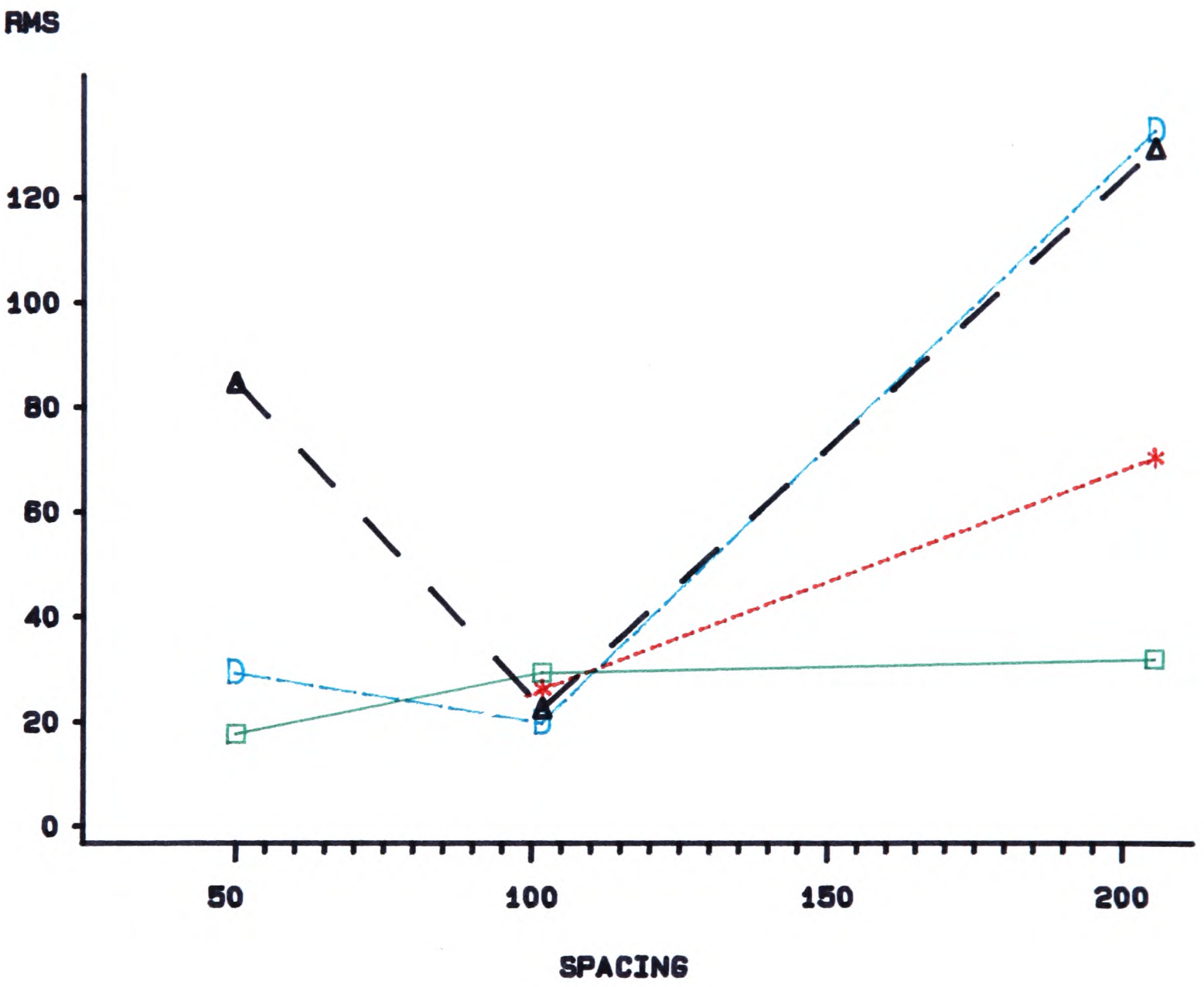
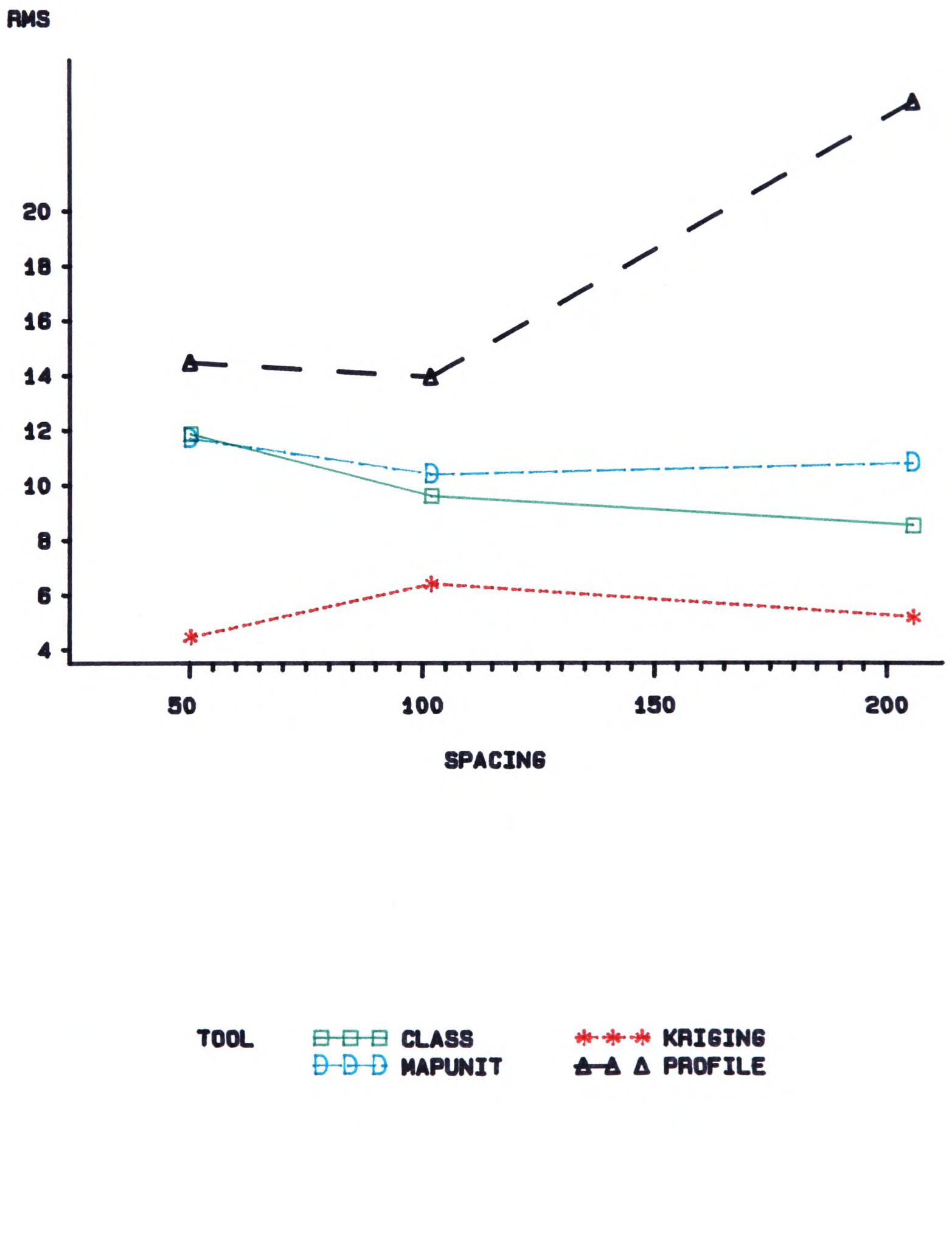


Fig. 9.21 ORGANIC MATTER

Comparison of the predictive value of Soil Survey output and spatial interpolation by Kriging



are the values estimated from the map unit's typical profile, but the poor performance of Kriging has its explanation in the nature of the variability as depicted by the structures of the semi-variograms. Success in prediction decreased with increasing sampling spacing, but the change is more pronounced when passing from 50 m to 100 m of grid spacing. From 100 m to 200 m the differences in accuracy in prediction are small. For this property increasing the sampling effort from a spacing of 200 m to a spacing of 100 m achieves nothing (fig. 9.14), if the products of soil survey are used as predictors, and very little is gained in accuracy if Kriging is the predictive tool.

Exchangeable Potassium showed anisotropic variation at a sampling scale corresponding to 50 m of intersampling spacing, and consequently no Kriging estimates were obtained by ignoring the presence of trends. For this reason there are no values of RSS for the 50 m grid (table 9.16a) and no values of RMS in the plot of figure 9.15. It can be seen, however, that where variation was isotropic as at the 100 and 200 m grids (figures 9.6b and 9.6c), predictions of Exchangeable Potassium by Kriging were better than by any other procedure using conventional soil survey output.

Means from mapping units made better predictions of Potassium (fig 9.15) than means from soil classes, presumably due to the less subjective element in mapping which is aided by external features in the subdivision of the landscape.

Like with pH, the success in prediction of Potassium (fig. 9.15) decreased with the increase in sampling spacing, being the decrease more pronounced when going from 50 m to 100 m except in the predictions by the typical profile and in the case of Kriging for which not predictions could be made at 50 m spacing.

Little is gained in accuracy of the predictions by Kriging and by soil survey output if the sampling effort is increased from 200 m to 100 m grid spacing. This can also be related to the structures of the semi-variograms of figs. 9.6b and 9.6c. Although both semi-variograms are of transition phenomena in the semi-variogram for the 200 m grid spatial dependence is weaker than that showed in the semi-variogram of the 100 m grid. In the 200 m grid semi-variogram, the spatially related variance is

actually smaller than the Nugget variance resulting in less accurate predictions and larger estimation variances.

Exchangeable Magnesium was predicted better by Kriging at the 50 m and 200 m sampling spacing (tables 9.19a and c) than by soil survey output, particularly class means (fig. 9.16), but the success in prediction at these scales was only marginally superior. At the 100 m grid spacing (table 9.19b) the advantage of Kriging was more evident (fig. 9.16) and map unit means and class means achieved quite similar results at this scale.

For this property success in prediction by Kriging was quite similar for the three sampling spacings. RMS did not decrease significantly with the increase in sampling effort from 200 m inter-sample spacing to 50 m spacing. This indicated that no gains could have been achieved in the accuracy of prediction at a considerable increase in survey cost, and 200 m grid spacing in this case was the most convenient sampling spacing.

This can also be explained by the semi-variograms of this property at the three sampling grids (figs. 9.7a,b and c). The semi-variogram at 200 m grid shows better structure in the region of the monotonic increase than the semi-variogram at 100 m grid. This is the region of importance for Kriging. So it can be interpreted that the possible gains in accuracy that would have occurred by passing from 200 m to 100 m spacings were ameliorated by more scanty variation of the transition phenomenon at the scale represented by the 100 m grid.

The closeness of the predictions of survey output to those of Kriging may be due to Exchangeable Magnesium being associated with other soil properties, notably, total carbonate, as in clacareous and non-calcareous soils, used as keys in the definition of classes and in the mapping of the soils in the Yelford area.

Total Carbonate showed anisotropic variation for samples taken at 50 m grid spacings (fig. 9.8a). No Kriging estimates were obtained (table 9.20a) by ignoring the trends causing the anisotropic variation and doing nothing about them. This is also shown in figure 9.17 where no RMS could be calculated at this spacing and missing in the plot.

The best predictive tool in these circumstances was the classs means but followed closely by mapping unit means. Predictions by the typical profile in the mapping unit as in all the other properties were the poorest.

The best predictions by all the procedures compared were obtained at the 100 m grid. The RSS (table 9.20b) were fairly close indeed for all the procedures involved but particularly for kriging and for means from classes and from mapping units, being Kriging the best of all but only marginally at this spacing.

In all the cases the success in prediction was better at this grid spacing than at the smaller spacing, i.e. 50 m. This fact indicates that even the predictive power of soil classes and map units can be affected by anisotropies and that the presence of these is reflected on a larger error in prediction. Here, stating the obvious, the structures of the anisotropic semi-variogram are definitely related to the success in prediction of the various techniques.

Errors in prediction of total Carbonate were higher at the 200 m grid, for all the techniques compared, than for the smaller grid spacings, but Kriging still made the best predictions. This can also be related to the structures of semi-variograms (figs. 9.8b and 9.8c). Spatial dependence is strong within a relatively short range of 350 m in the semi-variogram for the 100 m grid. A spherical model clearly fits and represents adequately such a well-defined experimental semi-variogram. In contrast, the spherical model fitted to the semi-variogram for the 200 m grid does not fit as neatly and the range of influence is far larger allowing for interactions with medium and long range sources of variation. This could be the causes of larger errors in prediction as measured by RSS and RMS.

In the case of total Carbonate increasing sampling effort from the 200 m spacing to the 100 m spacing would have brought considerable gains in accuracy of predicted values.

Table 9.21a shows that the predictions of Clay content by Kriging were sustantially better than predictions by soil survey output at the 50 m grid spacing. Kriging remained the best predictive technique of Clay at the 100 m grid (table 9.21b),

but at this sampling scale it was not considerably superior to the other predictive procedures. At 200 m of grid spacing (table 9.21c) Kriging again was the best predictive technique by a large margin. This can be appreciated in figure 9.18.

Predictions of Clay by mapping unit means and by the typical profile in the mapping units were almost identical at the 50 m grid and second best after Kriging. Mapping unit means made the best predictions after Kriging, at the 100 m grid but was replaced in this position by the class means as the grid spacing increased to 200 m (fig. 9.18).

The success of Kriging can be related to the semi-variograms of Clay (figs. 9.9a to c). All of them showed strong spatial dependence and well defined ranges of influence, and they were well represented by spherical models. In situations of this sort it can be advanced that Kriging almost invariably will make the best predictions, provided there is a good density of point-data.

Increasing sampling effort from 100 m inter-sampling spacing to 50 m (fig. 9.18) would improve predictions of Clay by Kriging, but an increase of sampling effort for the same area by reducing spacing from 200 m to 100 m would have no effect on the goodness of predictions. For map unit means and their typical profiles increasing sampling effort by reducing spacing from 200 m to 100 m and down to 50 m would improve predictions sustantially. The success in predictions of Clay by class means does not appear to change radically with the inter-sampling spacing (fig. 9.18). So, no improvement in predictions could be made with this technique by increasing sampling effort.

The variation of Silt at the 50 m grid was anisotropic (fig. 9.10a) and consequently no Kriging estimates were possible (table 9.22a). At this sampling scale class means made slightly better predictions than mapping unit means.

Kriging was the best technique to predict values of Silt at 100 m grid spacing (table 9.22b and fig 9.19), where variation was isotropic, but became second to class means when the grid spacing increased to 200 m. These results can be related to the structures of the corresponding semi-variograms (figs 9.10b and c). The semi-variogram at the 100 m grid shows a strong spatial dependence and a clearly defined range of influence (450m),

whereas at the 200 m grid Nugget variance is higher and the spatial dependence variance is weaker and relatively smaller. The range of influence stretches far longer to up to 1250 m. The exponential model fitted to this semi-variogram is characteristic of long transitions. Greater random fluctuations in a larger zone of influence makes a property which is not strongly spatally-auto-correlated more difficult to predict.

The success of predictions of Silt by class means seem to be unaffected by increasing sampling spacing, whereas for the other predictive procedures the step from 100 m to 200 m spacing decreases the goodness of predictions considerably.

Coarse Sand also could not be predicted by Kriging due to anisotropic variation of this property at the 50 m grid (fig. 9.11a). Class means made the best predictions at this scale (table 9.23a). In a similar situation as for Silt, the neatness and strength of the spatial dependence shown by the semi-variogram at the 100 m grid, combined with the rather small Nugget variance (fig 9.11b) help in explaining why Kriging makes better predictions of values of Coarse Sand at the 100 m grid than at the 200 m grid. However, at the 100 m grid the best predictions of this property were made by mapping unit means, followed closely by the values from their typical profiles. Kriging estimates were only marginally better than the predicted values by class means at this scale of sampling (fig 9.20). The predictive success of mapping units at this scale can be attributed to the medium range variations of Coarse Sand which are sharp and appear quite distinctly. This facilitates mapping enormously improving the quality of the map. In the case of Coarse Sand this is true particularly across the direction of the drift and in the intersect of the descent from Rickless hill with the older alluvium on the plains. Mapping units seem to stratify rather well the variation of this property at this scale.

A common feature of the predictive procedures compared for most of the properties studied is that almost invariably, at the 100 m grid their rates of success in prediction are not too different from each other, and for some properties the RMS are even quite close. This does not occur at the other grid spacings, and is true also for Coarse Sand (fig. 9.20). Here,

it can be seen that all predictive tools compared came remarkably close in their predictions at the 100 m grid. Hence, sampling spacings of 100 m or greater appear to reveal better the medium range structures of variation which are usually dealt with by soil mapping. Consequently, a soil map of better quality enhances the predictive power of conventional soil survey output.

Where the semi-variogram is isotropic, shows strong spatial dependence and its structures are clearly revealed by the experimental data, model fitting and estimation of parameters for Kriging are greately improved. This enhances sustantially the goodness of predictions by Kriging. In these circumstances Kriging is definitely the best predictive tool.

Variation of Organic Matter in the Yelford area is so spatially structured that is just as described above. Its three semi-variograms, i.e. 50 m, 100 m and 200 m grid, showed strong spatial auto-covariance, well defined ranges and relative small Nugget variances (figs. 9.12a, b and c). Furthermore, variation around the region of the monotonic increase, which is of great importance for Kriging, is fairly tight. The fit of models was quite good and the estimation of semi-variogram parameters for Kriging reasonably accurate.

With such a spatial structure, predictions of Organic Matter by Kriging were the best for the three sampling grids, by a large margin (tables 9.24a, b and c). So, Kriging was superior to conventional soil survey output as a predictive tool for this property.

The predictive success of Kriging did not change significantly with the increase of sampling spacing (fig. 9.21), although by increasing sampling effort from 100 m spacing to 50 m spacing results in some gains in accuracy of predictions. However, since predictions at a sampling spacing of 200 m were better than those at 100 m spacing, it is senseless to increase sampling effort by reducing spacing from 200 m to 100 m.

At the scale corresponding to the 50 m grid, map unit means and class means had practically the same success in prediction. Predictions by class means improved and were better than those by mapping unit means as the grid spacing increased.

As with the other properties, the predictions by the typical profile within mapping units were the poorest.

9.17 The semi-variogram as a reconnaissance tool when trends are ignored.

The predictive success (RSS and RMS) of the various procedures involved in the comparison , indeed, shown to be related to a lesser or grater extent, to the structures of the semi-variogram that corresponds to the sampling scale at which values were predicted. Hence, the structures of a carefully-planned reconnaissance semi-variogram can be used to ascertain in advance whether Kriging or whether any of the forms of conventional soil survey output will make the best predictions of the soil property at unvisited sites.

For a soil surveyor who wants to produce the best predictive tool possible for the minimum survey effort, the critical point in a practical situation is to decide whether to make a soil map by the conventional procedure or whether to use Kriging, or even to attempt an inexpensive combination of both.

If the reconnaissance semi-variogram shows strong spatial dependence and clearly defined structures (i.e. range, sill and transitional monotonic increase), with relatively small Nugget variances, (i.e. like the semi-variograms of Organic Matter at Yelford), it is almost safe to assume that Kriging will produce the best predictions and consequently this technique should be used with confidence, provided the sampling densities required can be afforded. If, on the contrary, the reconnaissance semi-variogram shows that the property has poor spatial structure or the fitting of a model and the estimation of its parameters is not satisfactory, the surveyor may decide that Kriging is not worthwhile and that similar or even better predictions can be made by soil classes and mapping units with possibly less effort.

Anisotropic variation is commonplace in soils. It can be detected in advance by the reconnaissance semi-variogram at the start of a survey, before resources have been committed to any

sampling intensity intended for the use of Kriging as a predictive tool.

The nature and strength of the spatial auto-covariance of the regionalized variable can also be judged a priori through the reconnaissance semi-variogram and helped to decide whether assumptions about stationarity and the intrinsic hypothesis are met. Thus, diagnosis work could be carried out about the existence of trends in the data or anisotropic variation which would preclude the immediate application of Kriging, requiring more elaborate procedures to account for such "anomalies" before predictions can be made.

For the sake of exploring the role of the semi-variogram as a reconnaissance tool, it was assumed that the semi-variograms of the properties studied in the Yelford area were calculated as if they were reconnaissance semi-variograms. Then, they were used for diagnosis and inference purposes. Thus, ignoring the results from the comparison of predictive procedures and based purely on the structures of the semi-variogram, the procedure which could have made the best predictions was inferred. So, it was decided whether it would have been better to have used Kriging or to have made a soil map and use soil classes and mapping units as predictive tools.

When the semi-variogram revealed anisotropies and the presence of trends, they were ignored. Kriging estimates were obtained wherever possible without having removed the trends. But, where it was not possible to obtain Kriging estimates, then it was assumed that the best option in such cases was to have made a soil map and to use it for prediction.

The choice of the best predictive technique, according to the reconnaissance semi-variogram, was then matched against the technique which the most successful from the comparison (i.e. section 9.16). This allowed to ascertain the extent to which, based on the reconnaissance semi-variogram, the correct inference was made, and its usefulness in the selection of either, Kriging or soil survey output, as the tool which will make the best predictions for a given condition of variability and sampling effort.

Thus, using the semi-variograms of pH (figs. 9.5a, b and c) as if they were calculated from reconnaissance, for a survey

effort corresponding to 50 m of inter-sampling spacing the Kriging technique should be used for prediction. For an effort of 100 m spacings between samples (fig 9.5b), it is not absolutely certain from the semi-variogram, whether Kriging or a map would be the best, but since spatial dependence was present at this scale and the model fitted well, the surveyor should feel inclined to prefer Kriging. The reason for uncertainty is a fairly large Nugget variance as compared with the size of the variance due to spatial dependence, which is weak at this scale.

For an effort represented by a 200 m spacing (fig 9.5c), it would have been better to have made a soil map and use it for prediction since there is total absence of spatial dependence at this sampling scale and the only predicted value that can be obtained is the general mean.

Matching these a priori inferences with the results in section 9.16, it can be seen (fig 9.14) that two out of the three inferences were correct. Indeed, as inferred, Kriging was a better predictor of soil pH at the 50 m grid, and soil classes and map units were superior to Kriging at the 200 m grid. The uncertainty regarding predictions at 100 m grid was justified. It would have been safer in these circumstances to have made a map and to avoid Kriging despite the spatial dependence shown by the semi-variogram. Such feature of the semi-variogram, in this case, may be misleading since mapping unit means made better predictions than Kriging at this sampling scale, possibly due to the presence of a large Nugget variance in this semi-variogram.

The semi-variogram of Potassium at the 50 m grid (fig. 9.6a), taken as reconnaissance, indicates that it might be better to undertake soil mapping than to sample for Kriging. There is almost total absence of spatial dependence and variability at this scale has a sustantial random component. This could make Kriging predictions not sufficiently accurate as to justify the sampling effort required.

For a 100 m of inter-sample spacing, the shape and structures of the semi-variogram (fig. 9.6b) would indicate that Kriging should be used as the predictive technique. Variation is isotropic, there is spatial dependence and the Nugget variance is not large.

At the 200 m grid the reconnaissance semi-variogram of Potassium shows a fairly large Nugget variance indicating a large random component. Spatial dependence is weak and in comparison too small. In these circumstances it would be too risky to undertake sampling for Kriging as it cannot be garanteed that it will make the best predictions. It would be safer then to make a soil map and use it to predict.

Matching the inferences made from the reconnaissance semi-variograms of Potassium with the results obtained from the comparison (section 9.16) of procedures (fig. 9.15), it can be seen that the reconnaissance semi-variogram helped to make the correct decisions at the 50 m grid where Kriging was not possible and a soil map is the best alternative, and at the 100 m grid where Kriging, indeed, made the best predictions. The choice, based on the semi-variogram, of making a map instead of using Kriging at the 200 m grid proved incorrect since Kriging made the best predictions in spite of the arguments used to infer the oposite.

The semi-variograms of Magnesium (figs. 9.7a, b and c), used for reconnaissance, would suggest that at the three sampling scales, 50 m, 100 m and 200 m grid spacing, Kriging would make better predictions than soil classes or a soil map. The three semi-variograms show a good degree of spatial auto-correlation of this property. However, Kriging predictions should not be expected to be considerably superior to those by map units and classes. This is so due to large Nugget variances and to sparse distribution of points along the the transition represented by the monotonic increase of the semi-variogram. This is so particularly for the semi-variograms of Magnesium at the 100 m and 200 m grids.

When these inferences were matched with the results of the comparison of procedures (fig. 9.16), it became clear that the reconnaissance semi-variograms of this property were quite useful in inferring accurately the predictive succes of Kriging and soil survey output, enabling one to select correctly the procedure to follow for spatial predictions. It goes without saying that no direct considerations of survey costs are made when choosinng the procedure, but the decision only concerns accuracy in prediction.

The semi-variograms of total Carbonate (figs. 9.8a, b and c) would have allowed to detect at the reconnaissance stage anisotropic variation of this property at the 50 m grid (fig. 9.8a). Kriging estimates could not have been obtained unless the trends were removed or a model for anisotropy fitted to the experimental semi-variogram. If the trends causing anisotropy are ignored then the surveyor should decide to produce a soil map to predict values of Carbonate at this sampling scale.

Judging from the structures of the semi-variogram at the 100 m grid (fig. 9.8b), the surveyor should use the Kriging technique for prediction. This semi-variogram shows Carbonate as having a good spatial structure, small Nugget variance and a clearly defined range of influence. These features should be sufficient to garantee that Kriging predictions would be the best at this sampling scale.

With the help of the semi-variogram at the 200 m grid as a reconnaissance tool (fig. 9.8c) the surveyor would have decided to undertake sampling for the use of Kriging technique to make predictions, since its structures given by a gradual transitional change in the values of this property at this scale, i.e. long range and an exponential model with good spatial dependence and low Nugget variance, would have made safe the use of Kriging.

The inferences made using the reconnaissance semi-variograms all match the facts from the results obtained from comparing the predictive value of Kriging and soil survey output (section 9.16, and fig. 9.17). Kriging resulted in the best predictive technique at 100 m and 200 m grid. At the 50 m grid the surveyor would be better using a soil map and classes for his predictions.

Observing and interpreting the structures of the semi-variograms of Clay at the 3 sampling grids (figs. 9.9a, b and c) as if they were from reconnaissance sampling, all of them suggested that Kriging was feasible and would predict better the values of Clay than soil survey output. Therefore, Kriging should be used to predict missing information at unsampled sites.

The grounds for such forecast rest on the fact that the three semi-variograms of Clay describe transition phenomena

which are strongly spatially-dependent, and with relatively small Nugget variances.

Because of a somewhat larger random component of the spatial variability of Clay at the 100 m grid, the predictive power of Kriging should be expected to come up somewhere closer to that of soil mapping units and classes, at this sampling spacing, than at the other samping grids.

Again these inferences on the success of Kriging, based purely on the structures of a priori semi-variograms, matched the results obtained from comparing the predicted procedures (fig 9.18). This demonstrates the value of pre-survey semi-variograms and the correct interpretation of their structures, as a basis for defining apropriate survey procedures to help improve spatial predictions. Kriging showed (fig. 9.18) to predict best the Clay values at the three sampling grids, although, as inferred from the semi-variogram, the improvement that it achieved over classical survey output at the 100 m grid was not considerable.

For Silt, the anisotropic variation depicted by the semi-variogram at the 50 m grid (fig. 9.10a), taken as if calculated from reconnaissance, permitted to infer that at this sampling scale, if trends were ignored, it would be better to make a map than to sample for Kriging, in order to predict missing information.

The strong spatial dependence in the variation of Silt at the 100 m grid (fig. 9.10b), and the relatively small Nugget variance and clearly defined range and sill of the semi-variogram suggested that it was not unreasonable to assume that Kriging would be the best predictive technique. Therefore, at this sampling scale the surveyor should prefer to use Kriging rather than classical output from conventional survey to predict missing information.

Although the transition is longer, and spatial dependence is somewhat weaker in the pre-survey exponential semi-variogram at the 200 m grid (fig. 9.10c), it would have been inferred that Kriging should also have predicted better the values of Silt than any other classical procedure at this sampling spacing.

When these forecasts were confronted with the results on the predictive performance of the various techniques (fig.

9.19), it turned out that the inference made by means of the semi-variogram at the 200 m grid was wrong. Kriging made inferior predictions than soil class means at this sampling scale, even when interpreting the semi-variogram in advance indicated the oposite. The interpretation of a long spatially-dependent transition was misleading. The effect of a large Nugget variance combined with a small spatial dependence variance in this semi-variogram (fig 9.10c) contributed to the inaccuracy of predictions. This was overlooked.

The ascertions made based on the semi-variograms at the other two sampling grids proved correct. At the 50 m grid class means is the best predictor. At the 100 m grid, as inferred, Kriging was superior.

The anisotropic variation of Coarse Sand at the 50 m grid shown by the semi-variogram (fig. 9.11a) did not allowed to predict by Kriging. The surveyor's best option in these circumstances would be to make a soil map.

Based on the semi-variogram of this property at the 100 m grid (fig. 9.11b) it could have been interpreted in advance that Kriging would have made better predictions than map units and classes at this sampling spacing. The surveyor should have sampled with Kriging in mind rather than for drawing soil boundaries. The semi-variogram shows strong spatial dependence, well defined structures and a low random component of variation. These features provide sufficient grounds for such decision.

At the 200 m grid the structures of the semi-variogram (fig 9.11c) would have permited to forecast that Kriging would have made better preditions than other methods. Variation appears to be of a long transition and spatially dependent. The Nugget variance is not large. The surveyor, at this sampling spacing, should have decided to predict by Kriging rather than to make a soil map and use it for such purpose.

Comparing the forecasts made using these assummed reconnaissance semi-variograms with the actual evidence on the predictive success of Kriging and classical survey output (fig. 9.20), it can be seen that for the 100 m and 200 m grid spacings the inferences made were proved wrong. Mapping unit means made better predictions than Kriging at the 100 m grid, although the

improvement is only marginal. Soil class means was also superior to Kriging as predictior at the 200 m grid. In this case the improvement is considerable.

Only at the 50 m grid the decision taken based on the reconnaissance semi-variogram for this property was the correct one, if trends were to be ignored. So, for this property, the recconaisance semi-variogram would have been useful only in cases of marked "accidents" in its spatial variation, which are obstacles to Kriging like anisotropies or trends. Even where the semi-variogram showed structures which were interpreted as favorable to Kriging, it was not possible to infer from that information how better a soil map and classes would have done for the inference focuses on Kriging. There was nothing in the structures of the semi-variograms on which to base inferences for classical survey output.

The semi-variograms of Organic Matter content at the three sampling grids (figs. 9.12a, b and c), taken as if they were from reconnaissance sampling, suggested on the basis of their structures, that Kriging was viable and would make better spatial predictions of this property than any other procedure. Thus, the surveyor in all cases should prefer to obtain Kriging estimates as predictors of Organic Matter content rather than to make a soil map and use classes and mapping units. The degree of spatial structure as depicted by the three semi-variograms is, for the properties studied, exceptionally good. The fit of the models was also good and Nugget variances relatively small.

Checking the validity of these inferences against the facts (fig. 9.21), it can be seen that all forecasts based on interpretations of the reconnaissance semi-variograms of this property proved correct. The Kriging technique was, indeed, superior to any other procedure that uses conventional soil survey output. The reconnaissance semi-variogram was, in the case of this property, quite useful.

Having used semi-variograms which were assummed to have been calculated as if they were from reconnaissance sampling, for all the properties involved in this study, the rate of success of the inferences and forecasts based on their interpretation was also recorded. This is an indirect measure of the utility of the reconnaissance semi-variogram to plan

survey strategies endevoured to produce the best predictive tools for soil missing information.

The rates of success are given in terms of "wrong/right" and are presented as % of correct inferences over the total made. They are shown in table 9.25.

Table 9.25. Rate of success with which the semi-variogram used as a reconnaissance tool, predicted correctly the adequacy of chosing between Kriging or mapping units and classes as predictors of soil missing information.

Soil property	INTER-SAMPLE SPACING (meters)		
	50	100	200
pH	R	W	R
K	R	R	W
Mg	R	R	R
Carbonate	R	R	R
Clay	R	R	R
Silt	R	R	W
Coarse Sand	R	W	W
Organic Matter	R	R	R
TOTAL RATE OF SUCCESS = 79.2% (80%)			

Inference proved: R = RIGHT, W = WRONG

From table 9.25 it can be seen that the success of the semi-variogram as a reconnaissance tool was relatively high, even when the trends present in the area were ignored.

On the other hand it shoud be stressed that the semi-variograms used here as reconnaissance tool were calculated with a wealth of point-data that could not have been obtained in ordinary circumstances from pre-survey reconnaissance. So, in this sense the semi-variogram estimation was perhaps better that it could have been with less point-data, unless the pre-survey

sampling were so carefully designed as to give an equally good estimation of the semi-variogram with less point-data.

Hence, from the above, it all indicates that the critical point is how best to sample in order to estimate well a reconnaissance semi-variogram so that it can be used in choosing the best predictive tool, and thence designing the appropriate survey strategy.

From table 9.25, it must be mention that there were some cases in which the reconnaissance semi-variogram showed such favorable spatial structures that the surveyor coud have been, and in fact was, misled by them and it could have been advanced with confidence that Kriging was the best predictive technique. Yet, soil mapping units or means from classes gave better predictions than Kriging, e.g. Coarse Sand at 100 m and 200 m grids, pH at 100 m grid, Silt at 200 m grid. This apparent failure of the semi-variogram as reconnaissance tool occurred in most cases where there was a clear external expresion of soil boundaries and a strong degree of association between soil varibility and landscape, e.g. Coarse Sand at 100 and 200 m grids.

So, where the location of soil boundaries was sufficiently accurate due to a clear external expression of soil/landscape relationships, mapping units and soil classes may give predictions which could be as good as those by Kriging. It depends on the quality of the map produced which is to be used as predictive tool.

This leads to the conclusion that although in many cases the semi-variogram alone is a powerful inference tool, there is no standard procedure to interpret reconnaissance semi-variograms and to use those interpretations for inferences involving the selection of the best predictive procedure of soil properties. The choice can not be based solely on the reconnaissance semi-variogram on its own without due attention to the landscape and its relation to soil variability.

9.18 Removing Trends.

The presence of trends in the data have a significant effect on the results of the comparison of efficacy in spatial prediction by the various procedures.

Exchangeable Potassium, total Carbonate, Silt and Coarse Sand showed anisotropic variations at the 50 m grid. This was associated with the presence of trends in the direction of the drift (NE-SW). These four properties, therefore, were selected for studying the effect of removing trends on the success of the predictive procedures.

In order to examine such effects, two alternative procedures for trend removal were explored: modelling and stratification.

9.18.1 Modelling of trends by bicubic spline surfaces.

Bicubic spline surfaces were calculated and fitted to the variation of the four properties selected to model the trends. The spline surfaces were fitted not only at the 50 m grid where anisotropy was evident but at the 100m and 200m grids too, in order to compare the effect of trend removal at all sampling scales.

The bicubic spline surfaces were fitted to the data on the grids by means of the procedures described in 7.3.2.2.i in chapter 7. Equations 7.19 and particularly 7.20 were applied for the computation of the spline coefficients. Software was written to interface the NAG library routines E02DAF, E02DBF and E02ZAF (NAG, 1982). First, the modelling of the trend surface was achieved interactively by guiding the fit of a weighted least-squares bicubic spline surface through the pairs of co-ordinates of a suggested set of knots, replacing any of the pairs or suggesting a new pair when necessary to improve the fit. At each step the achieved surface was visually inspected on a 3-D plot until a satisfactory fit was found. Unit weights were used in all cases. Then the spline surface was evaluated at the grid points returning the values of the surface, i.e. the

trend values of the property, at those points. This allowed for the calculation of residuals from the trend.

In tables 9.26, 9.27 and 9.28 the bicubic spline surfaces fitted to the variation of Potassium at the 50 m, 100 m, and 200 m grids correspondingly, are given to illustrate the nature of the output. They list the final set of suggested knots, the observed values at the grid points, the corresponding fit values on the surface and the residuals between them. The set of spline coefficients corresponding to each patch or panel that make up the whole surface are also included. For brevity, the tables listing similar information for the other three properties at the three sampling grids are given in appendix 9.6.

The final fit of the surfaces to model the trends were displayed in 3-D plots using SAS/Graph standard graphics package. The plots of the surfaces are shown together with corresponding 3-D plots of the observed raw values at the grid points to enable visual inspection of the fit of each surface. First, the plot of the observed raw values is given. Then, it is followed by the corresponding surface that models the trend.

So, figures 9.22a-b, 9.23a-b, and 9.24a-b show the plots of observed values and their trend surface for Potassium at the 50, 100, and 200 m grids respectively. In the same order the figures 9.25a-b, 9.26a-b and 9.27a-b show plots of observed values and the trend surfaces of Carbonate at the three grids. The plots of Silt are shown in figures 9.28a-b, 9.29a-b and 9.30a-b. Similarly, the plots for Coarse Sand are included in figures 9.31a-b, 9.32a-b, and 9.33a-b.

On the whole it can be said that the surface fits achieved were good. However, the fitting procedure was not as straight-forward as desired, and several iterations involving trials with the positioning of knots had to take place before achieving a satisfactory surface fit.

Nevertheless, it can be seen from the plots listed above that the surfaces exhibit the major features of the variation of the property at the corresponding scale. Extrema and inflection points are reasonably well represented.

The surfaces on the plots are not as smooth as they should really appear. Some of the roughness and accidents that appear

Table 9.26 . Modelling of the TREND by fitting a BICUBIC SPLINE SURFACE to the variation of POTASSIUM over the 50m grid at Yelford.

BICUBIC SPLINE SURFACE, POTASSIUM 50 m GRID.

INTERIOR X-KNOTS

2.0460E+04
2.0476E+04
2.0500E+04
2.0515E+04

INTERIOR Y-KNOTS

4.3539E+04
4.3550E+04
4.3557E+04

SUM OF SQUARES OF RESIDUAL RHS 3.0014E+00
RANK 49

NORTHING	EASTING	OBSERVED	SPLINE FIT	RESIDUALS
2.0445E+04	4.3535E+04	8.0000E-01	7.9474E-01	-5.2635E-03
2.0455E+04	4.3530E+04	3.0000E-01	2.3680E-01	-6.3199E-02
2.0455E+04	4.3535E+04	2.0000E-01	1.4433E-01	-5.5666E-02
2.0450E+04	4.3535E+04	2.0000E-01	2.0505E-01	5.0531E-03
2.0450E+04	4.3540E+04	2.0000E-01	2.6753E-01	6.7525E-02
2.0455E+04	4.3540E+04	4.0000E-01	3.3523E-01	-6.4775E-02
2.0455E+04	4.3545E+04	4.0000E-01	4.9683E-01	9.6830E-02
2.0451E+04	4.3546E+04	6.0000E-01	4.2057E-01	-1.7943E-01
2.0450E+04	4.3550E+04	4.0000E-01	5.0134E-01	1.0134E-01
2.0455E+04	4.3550E+04	5.0000E-01	5.5102E-01	5.1021E-02
2.0455E+04	4.3555E+04	5.0000E-01	5.2408E-01	2.4076E-02
2.0450E+04	4.3555E+04	7.0000E-01	6.6165E-01	-3.8350E-02
2.0450E+04	4.3560E+04	6.0000E-01	6.0587E-01	5.8733E-03
2.0455E+04	4.3560E+04	7.0000E-01	6.9221E-01	-7.7922E-03
2.0460E+04	4.3530E+04	1.0000E-01	2.0167E-01	1.0167E-01
2.0470E+04	4.3530E+04	6.0000E-01	3.7248E-01	-2.2752E-01
2.0475E+04	4.3530E+04	3.0000E-01	4.7583E-01	1.7583E-01
2.0475E+04	4.3535E+04	3.0000E-01	3.1942E-01	1.9416E-02
2.0470E+04	4.3535E+04	4.0000E-01	3.5523E-01	-4.4773E-02
2.0465E+04	4.3530E+04	2.0000E-01	2.5288E-01	5.2881E-02
2.0460E+04	4.3535E+04	2.0000E-01	2.8434E-01	8.4339E-02
2.0470E+04	4.3540E+04	4.0000E-01	5.3419E-01	1.3419E-01
2.0460E+04	4.3540E+04	4.0000E-01	4.7181E-01	7.1808E-02
2.0475E+04	4.3540E+04	7.0000E-01	4.9226E-01	-2.0774E-01
2.0475E+04	4.3545E+04	3.0000E-01	4.6333E-01	1.6333E-01
2.0470E+04	4.3545E+04	6.0000E-01	4.8854E-01	-1.1146E-01
2.0465E+04	4.3540E+04	7.0000E-01	5.3849E-01	-1.6151E-01
2.0464E+04	4.3544E+04	5.0000E-01	5.4313E-01	4.3131E-02
2.0460E+04	4.3545E+04	5.0000E-01	5.4174E-01	4.1743E-02
2.0460E+04	4.3550E+04	7.0000E-01	5.2913E-01	-1.7087E-01
2.0470E+04	4.3550E+04	3.0000E-01	3.6453E-01	6.4529E-02
2.0475E+04	4.3550E+04	7.0000E-01	3.3256E-01	-3.6744E-01
2.0475E+04	4.3555E+04	3.0000E-01	2.8308E-01	-1.6920E-02
2.0470E+04	4.3555E+04	5.0000E-01	3.5970E-01	-1.4030E-01
2.0466E+04	4.3551E+04	3.0000E-01	4.2265E-01	1.2265E-01
2.0465E+04	4.3555E+04	3.0000E-01	4.6759E-01	1.6759E-01
2.0460E+04	4.3555E+04	6.0000E-01	5.2913E-01	-7.0871E-02
2.0470E+04	4.3560E+04	2.0000E-01	5.4943E-01	3.4943E-01
2.0460E+04	4.3560E+04	6.0000E-01	7.0724E-01	1.0724E-01
2.0475E+04	4.3560E+04	3.0000E-01	4.6390E-01	1.6390E-01
2.0465E+04	4.3560E+04	1.0000E+00	6.4686E-01	-3.5314E-01

2.0480E+04	4.3530E+04	6.0000E-01	4.8992E-01	-1.1008E-01
2.0490E+04	4.3530E+04	2.0000E-01	3.4048E-01	1.4048E-01
2.0495E+04	4.3530E+04	3.0000E-01	2.7994E-01	-2.0060E-02
2.0495E+04	4.3535E+04	3.0000E-01	3.9754E-01	9.7539E-02
2.0490E+04	4.3535E+04	4.0000E-01	3.3352E-01	-6.6484E-02
2.0485E+04	4.3535E+04	3.0000E-01	3.0028E-01	2.7913E-04
2.0485E+04	4.3535E+04	3.0000E-01	3.0028E-01	2.7913E-04
2.0480E+04	4.3535E+04	3.0000E-01	2.9616E-01	-3.8436E-03
2.0490E+04	4.3540E+04	2.0000E-01	3.7822E-01	1.7822E-01
2.0480E+04	4.3540E+04	3.0000E-01	4.4381E-01	1.4381E-01
2.0495E+04	4.3540E+04	4.0000E-01	3.7537E-01	-2.4633E-02
2.0495E+04	4.3545E+04	6.0000E-01	6.8921E-01	8.9209E-02
2.0490E+04	4.3545E+04	7.0000E-01	6.1815E-01	-8.1847E-02
2.0485E+04	4.3540E+04	4.0000E-01	4.0341E-01	3.4118E-03
2.0485E+04	4.3545E+04	1.1000E+00	5.4199E-01	-5.5801E-01
2.0480E+04	4.3545E+04	2.0000E-01	4.8298E-01	2.8298E-01
2.0480E+04	4.3550E+04	1.0000E-01	4.0435E-01	3.0435E-01
2.0490E+04	4.3550E+04	6.0000E-01	7.4096E-01	1.4096E-01
2.0495E+04	4.3550E+04	9.0000E-01	8.9829E-01	-1.7092E-03
2.0495E+04	4.3555E+04	8.0000E-01	6.9442E-01	-1.0558E-01
2.0490E+04	4.3555E+04	4.0000E-01	5.5806E-01	1.5806E-01
2.0485E+04	4.3555E+04	4.0000E-01	4.1220E-01	1.2198E-02
2.0485E+04	4.3555E+04	6.0000E-01	4.1220E-01	-1.8780E-01
2.0480E+04	4.3555E+04	2.0000E-01	3.0465E-01	1.0465E-01
2.0490E+04	4.3560E+04	4.0000E-01	5.7109E-01	1.7109E-01
2.0480E+04	4.3560E+04	5.0000E-01	4.3462E-01	-6.5377E-02
2.0480E+04	4.3570E+04	3.0000E-01	3.0000E-01	4.7186E-06
2.0495E+04	4.3560E+04	5.0000E-01	7.2705E-01	2.2705E-01
2.0480E+04	4.3560E+04	8.0000E-01	4.3462E-01	-3.6538E-01
2.0500E+04	4.3530E+04	3.0000E-01	2.9764E-01	-2.3567E-03
2.0510E+04	4.3530E+04	3.0000E-01	4.8280E-01	1.8280E-01
2.0510E+04	4.3535E+04	5.0000E-01	5.7849E-01	7.8490E-02
2.0505E+04	4.3530E+04	6.0000E-01	4.0876E-01	-1.9124E-01
2.0500E+04	4.3535E+04	6.0000E-01	4.9402E-01	-1.0598E-01
2.0510E+04	4.3540E+04	7.0000E-01	5.7350E-01	-1.2650E-01
2.0500E+04	4.3540E+04	4.0000E-01	4.0197E-01	1.9682E-03
2.0510E+04	4.3545E+04	8.0000E-01	7.7841E-01	-2.1586E-02
2.0505E+04	4.3545E+04	5.0000E-01	7.4309E-01	2.4309E-01
2.0505E+04	4.3545E+04	7.0000E-01	7.4309E-01	4.3086E-02
2.0500E+04	4.3545E+04	9.0000E-01	7.3289E-01	-1.6711E-01
2.0500E+04	4.3550E+04	1.1000E+00	9.7653E-01	-1.2347E-01
2.0510E+04	4.3550E+04	1.3000E+00	9.7254E-01	-3.2746E-01
2.0510E+04	4.3555E+04	9.0000E-01	9.8296E-01	8.2958E-02
2.0505E+04	4.3555E+04	4.0000E-01	7.9911E-01	3.9911E-01
2.0505E+04	4.3550E+04	7.0000E-01	9.5837E-01	2.5837E-01
2.0505E+04	4.3555E+04	1.2000E+00	7.9911E-01	-4.0089E-01
2.0500E+04	4.3555E+04	8.0000E-01	7.7342E-01	-2.6581E-02
2.0510E+04	4.3560E+04	8.0000E-01	1.0741E+00	2.7413E-01
2.0500E+04	4.3560E+04	1.1000E+00	9.3492E-01	-1.6508E-01
2.0505E+04	4.3560E+04	1.4000E+00	1.1382E+00	-2.6179E-01
2.0520E+04	4.3530E+04	6.0000E-01	5.9785E-01	-2.1471E-03
2.0520E+04	4.3535E+04	8.0000E-01	8.2446E-01	2.4457E-02
2.0515E+04	4.3530E+04	4.0000E-01	3.5285E-01	-4.7155E-02
2.0515E+04	4.3535E+04	3.0000E-01	2.7759E-01	-2.2406E-02
2.0520E+04	4.3540E+04	7.0000E-01	6.1962E-01	-8.0380E-02
2.0520E+04	4.3545E+04	6.0000E-01	7.2012E-01	1.2012E-01
2.0515E+04	4.3540E+04	7.0000E-01	7.3469E-01	3.4687E-02
2.0520E+04	4.3550E+04	9.0000E-01	8.0939E-01	-9.0608E-02
2.0520E+04	4.3555E+04	6.0000E-01	6.3309E-01	3.3090E-02
2.0515E+04	4.3555E+04	1.2000E+00	1.5883E+00	3.8832E-01
2.0515E+04	4.3550E+04	1.1000E+00	1.1843E+00	8.4299E-02
2.0515E+04	4.3555E+04	2.0000E+00	1.5883E+00	-4.1168E-01
2.0520E+04	4.3560E+04	7.0000E-01	6.9541E-01	-4.5873E-03
2.0515E+04	4.3560E+04	5.0000E-01	4.2831E-01	-7.1686E-02

TABLE 9.26 (CONTINUATION)

SUM OF SQUARED RESIDUALS 3.0014E+00

	8.4292E-01	1.7000E-01
SPLINE COEFFS	9.8577E-01	1.0905E+00
	1.2269E-01	-9.5712E-03
-4.9517E-01		4.4654E+00
1.1919E+00		2.0723E-01
4.4647E-01	8.1275E-01	
-4.0388E-01	-1.3088E-01	
2.7919E+00	8.5237E-01	9.9302E-02
-3.0965E+00	-2.1257E-02	-1.0702E+00
-1.5477E-01	-1.5727E-01	1.8116E+00
	1.0183E+00	4.2043E-01
	3.9393E-01	4.6992E+00
4.5355E-01		-9.3378E+00
-3.8471E-01		-4.5644E-01
1.4039E-02	-5.8368E-02	
8.2044E-01	6.9289E-01	
-9.1025E-02	-4.2312E-01	5.9785E-01
1.7880E+00	1.8072E+00	1.1048E+00
8.0409E-02	4.2718E-01	2.6296E-01
	6.9973E-01	1.0693E+00
	1.5668E-01	2.4557E-01
-4.3895E-02		1.3321E+00
4.5227E-01		6.1166E-02
7.6041E-01	7.3002E-01	
5.4531E-01	1.1310E+00	

Table 9.27 . Modelling the TREND by fitting a BICUBIC SPLINE SURFACE to the variation of POTASSIUM over the 100m grid at Yelford.

BICUBIC SPLINE SURFACE, POTASSIUM 100m GRID.

INTERIOR X-KNOTS

2.0460E+04
2.0494E+04
2.0523E+04
2.0550E+04

INTERIOR Y-KNOTS

4.3498E+04
4.3534E+04
4.3571E+04

SUM OF SQUARES OF RESIDUAL RHS 8.4240E+00

RANK 56

NORTHING	EASTING	OBSERVED	SPLINE FIT	RESIDULAS
2.0450E+04	4.3480E+04	4.0000E-01	4.0223E-01	2.2253E-03
2.0450E+04	4.3490E+04	4.0000E-01	3.7465E-01	-2.5348E-02
2.0450E+04	4.3500E+04	3.0000E-01	3.3639E-01	3.6389E-02
2.0451E+04	4.3520E+04	2.0000E-01	-1.4627E-03	-2.0146E-01
2.0450E+04	4.3510E+04	2.0000E-01	2.5626E-01	5.6261E-02
2.0450E+04	4.3530E+04	1.0000E-01	1.6000E-01	5.9996E-02
2.0450E+04	4.3540E+04	2.0000E-01	2.4809E-01	4.8087E-02
2.0450E+04	4.3560E+04	6.0000E-01	5.1511E-01	-8.4894E-02
2.0450E+04	4.3550E+04	4.0000E-01	4.0292E-01	2.9220E-03
2.0450E+04	4.3570E+04	4.0000E-01	4.7347E-01	7.3472E-02
2.0450E+04	4.3580E+04	3.0000E-01	2.7511E-01	-2.4891E-02
2.0450E+04	4.3590E+04	5.0000E-01	5.0265E-01	2.6455E-03
2.0490E+04	4.3480E+04	7.0000E-01	5.5314E-01	-1.4686E-01
2.0480E+04	4.3480E+04	6.0000E-01	5.7028E-01	-2.9724E-02
2.0479E+04	4.3489E+04	5.0000E-01	4.4646E-01	-5.3535E-02
2.0470E+04	4.3480E+04	4.0000E-01	4.9594E-01	9.5945E-02
2.0460E+04	4.3480E+04	4.0000E-01	3.6119E-01	-3.8814E-02
2.0460E+04	4.3490E+04	3.0000E-01	3.7835E-01	7.8348E-02
2.0470E+04	4.3490E+04	4.0000E-01	4.3493E-01	3.4928E-02
2.0490E+04	4.3490E+04	5.0000E-01	5.0039E-01	3.8733E-04
2.0490E+04	4.3500E+04	5.0000E-01	4.7971E-01	-2.0289E-02
2.0490E+04	4.3520E+04	5.0000E-01	3.5495E-01	-1.4505E-01
2.0480E+04	4.3500E+04	4.0000E-01	4.8147E-01	8.1467E-02
2.0481E+04	4.3511E+04	1.0000E+00	5.1879E-01	-4.8121E-01
2.0480E+04	4.3520E+04	7.0000E-01	5.4467E-01	-1.5533E-01
2.0480E+04	4.3530E+04	6.0000E-01	5.0315E-01	-9.6847E-02
2.0470E+04	4.3500E+04	4.0000E-01	4.7290E-01	7.2898E-02
2.0469E+04	4.3521E+04	4.0000E-01	5.9010E-01	1.9010E-01
2.0460E+04	4.3500E+04	5.0000E-01	2.5706E-01	-2.4294E-01
2.0460E+04	4.3510E+04	1.0000E-01	1.7382E-01	7.3821E-02
2.0460E+04	4.3520E+04	2.0000E-01	1.6013E-01	-3.9867E-02
2.0460E+04	4.3530E+04	1.0000E-01	2.2788E-01	1.2788E-01
2.0490E+04	4.3510E+04	5.0000E-01	4.2068E-01	-7.9324E-02
2.0470E+04	4.3530E+04	6.0000E-01	5.7428E-01	-2.5719E-02
2.0490E+04	4.3530E+04	2.0000E-01	3.2800E-01	1.2800E-01
2.0470E+04	4.3510E+04	2.0000E-01	5.4950E-01	3.4950E-01
2.0490E+04	4.3540E+04	2.0000E-01	3.8209E-01	1.8209E-01
2.0490E+04	4.3560E+04	4.0000E-01	6.7152E-01	2.7152E-01
2.0480E+04	4.3540E+04	3.0000E-01	3.8184E-01	8.1840E-02

2.0480E+04	4.3552E+04	1.0000E-01	2.4568E-01	1.4568E-01
2.0480E+04	4.3560E+04	5.0000E-01	2.2642E-01	-2.7358E-01
2.0480E+04	4.3570E+04	3.0000E-01	4.5751E-01	1.5751E-01
2.0470E+04	4.3540E+04	4.0000E-01	4.0876E-01	8.7619E-03
2.0470E+04	4.3560E+04	2.0000E-01	7.2646E-02	-1.2735E-01
2.0460E+04	4.3540E+04	4.0000E-01	3.8426E-01	-1.5736E-02
2.0460E+04	4.3550E+04	7.0000E-01	5.6663E-01	-1.3337E-01
2.0460E+04	4.3560E+04	6.0000E-01	6.5851E-01	5.8515E-02
2.0460E+04	4.3570E+04	4.0000E-01	5.4208E-01	1.4208E-01
2.0490E+04	4.3550E+04	6.0000E-01	5.1106E-01	-8.8939E-02
2.0470E+04	4.3570E+04	2.0000E-01	2.5476E-01	5.4762E-02
2.0490E+04	4.3570E+04	8.0000E-01	8.1913E-01	1.9126E-02
2.0470E+04	4.3550E+04	3.0000E-01	1.8319E-01	-1.1681E-01
2.0490E+04	4.3580E+04	5.0000E-01	9.7849E-01	4.7849E-01
2.0479E+04	4.3581E+04	1.9000E+00	1.2411E+00	-6.5890E-01
2.0480E+04	4.3590E+04	3.0000E+00	2.6987E+00	-3.0127E-01
2.0470E+04	4.3580E+04	6.0000E-01	9.7303E-01	3.7303E-01
2.0460E+04	4.3580E+04	4.0000E-01	2.3173E-01	-1.6827E-01
2.0460E+04	4.3590E+04	4.0000E-01	4.5720E-01	5.7203E-02
2.0490E+04	4.3590E+04	6.0000E-01	1.5473E+00	9.4732E-01
2.0470E+04	4.3590E+04	2.9000E+00	2.8272E+00	-7.2803E-02
2.0520E+04	4.3480E+04	5.0000E-01	3.9878E-01	-1.0122E-01
2.0520E+04	4.3490E+04	1.6000E+00	1.0506E+00	-5.4942E-01
2.0510E+04	4.3480E+04	2.0000E-01	2.9652E-01	9.6518E-02
2.0500E+04	4.3480E+04	3.0000E-01	4.2277E-01	1.2277E-01
2.0500E+04	4.3490E+04	1.0000E+00	6.7596E-01	-3.2404E-01
2.0510E+04	4.3490E+04	4.0000E-01	9.1424E-01	5.1424E-01
2.0520E+04	4.3500E+04	1.5000E+00	1.1476E+00	-3.5238E-01
2.0520E+04	4.3520E+04	1.6000E+00	8.8863E-01	-7.1137E-01
2.0520E+04	4.3530E+04	6.0000E-01	7.5160E-01	1.5160E-01
2.0510E+04	4.3500E+04	3.0000E-01	9.4535E-01	6.4535E-01
2.0510E+04	4.3520E+04	4.0000E-01	6.4389E-01	2.4389E-01
2.0500E+04	4.3510E+04	6.0000E-01	5.1737E-01	-8.2628E-02
2.0500E+04	4.3520E+04	2.0000E-01	3.8381E-01	1.8381E-01
2.0500E+04	4.3530E+04	3.0000E-01	3.4315E-01	4.3147E-02
2.0510E+04	4.3530E+04	3.0000E-01	5.5568E-01	2.5568E-01
2.0520E+04	4.3540E+04	7.0000E-01	7.2711E-01	2.7108E-02
2.0520E+04	4.3550E+04	9.0000E-01	7.9482E-01	-1.0518E-01
2.0520E+04	4.3560E+04	7.0000E-01	8.5223E-01	1.5223E-01
2.0520E+04	4.3570E+04	6.0000E-01	7.9476E-01	1.9476E-01
2.0510E+04	4.3540E+04	7.0000E-01	6.3756E-01	-6.2444E-02
2.0510E+04	4.3560E+04	8.0000E-01	1.0108E+00	2.1084E-01
2.0500E+04	4.3540E+04	4.0000E-01	4.8067E-01	8.0672E-02
2.0500E+04	4.3550E+04	1.1000E+00	7.4491E-01	-3.5509E-01
2.0500E+04	4.3560E+04	1.1000E+00	9.7909E-01	-1.2091E-01
2.0500E+04	4.3570E+04	6.0000E-01	1.0237E+00	4.2375E-01
2.0510E+04	4.3570E+04	8.0000E-01	9.8534E-01	1.8534E-01
2.0510E+04	4.3550E+04	1.3000E+00	8.4163E-01	-4.5837E-01
2.0520E+04	4.3580E+04	6.0000E-01	6.2577E-01	2.5772E-02
2.0520E+04	4.3590E+04	6.0000E-01	9.3267E-01	3.3267E-01
2.0510E+04	4.3580E+04	6.0000E-01	7.1371E-01	1.1371E-01
2.0500E+04	4.3580E+04	1.2000E+00	8.1197E-01	-3.8803E-01
2.0500E+04	4.3590E+04	1.4000E+00	7.7750E-01	-6.2250E-01
2.0509E+04	4.3571E+04	1.6000E+00	9.7832E-01	-6.2168E-01
2.0510E+04	4.3590E+04	1.0000E+00	7.3125E-01	-2.6875E-01
2.0539E+04	4.3481E+04	1.1000E+00	1.1629E+00	6.2939E-02
2.0540E+04	4.3490E+04	1.1000E+00	8.4658E-01	-2.5342E-01
2.0530E+04	4.3480E+04	9.0000E-01	8.8952E-01	-1.0477E-02
2.0530E+04	4.3490E+04	8.0000E-01	9.5242E-01	1.5242E-01
2.0540E+04	4.3520E+04	9.0000E-01	6.3824E-01	-2.6176E-01
2.0540E+04	4.3500E+04	6.0000E-01	7.9137E-01	1.9137E-01
2.0540E+04	4.3510E+04	7.0000E-01	7.2811E-01	2.8111E-02
2.0540E+04	4.3520E+04	3.0000E-01	6.3824E-01	3.3824E-01
2.0540E+04	4.3530E+04	7.0000E-01	5.4685E-01	-1.5315E-01

2.0530E+04	4.3500E+04	6.0000E-01	1.0610E+00	4.6098E-01
2.0530E+04	4.3520E+04	6.0000E-01	8.8707E-01	2.8707E-01
2.0531E+04	4.3511E+04	1.6000E+00	9.9306E-01	-6.0694E-01
2.0530E+04	4.3530E+04	6.0000E-01	7.3439E-01	1.3439E-01
2.0540E+04	4.3540E+04	5.0000E-01	4.7807E-01	-2.1935E-02
2.0540E+04	4.3550E+04	5.0000E-01	4.4093E-01	-5.9066E-02
2.0540E+04	4.3560E+04	6.0000E-01	4.3294E-01	-1.6706E-01
2.0540E+04	4.3570E+04	4.0000E-01	4.5127E-01	5.1269E-02
2.0530E+04	4.3540E+04	4.0000E-01	6.4141E-01	2.4141E-01
2.0530E+04	4.3560E+04	6.0000E-01	6.0173E-01	1.7326E-03
2.0530E+04	4.3550E+04	6.0000E-01	6.1358E-01	1.3578E-02
2.0530E+04	4.3570E+04	5.0000E-01	5.5529E-01	5.5292E-02
2.0540E+04	4.3580E+04	5.0000E-01	5.1000E-01	1.0004E-02
2.0540E+04	4.3590E+04	9.0000E-01	7.1465E-01	-1.8535E-01
2.0530E+04	4.3580E+04	6.0000E-01	5.0573E-01	-9.4266E-02
2.0530E+04	4.3590E+04	8.0000E-01	9.2835E-01	1.2835E-01
2.0560E+04	4.3480E+04	4.0000E-01	4.0882E-01	8.8227E-03
2.0559E+04	4.3491E+04	2.2000E+00	1.9941E+00	-2.0593E-01
2.0550E+04	4.3480E+04	5.0000E-01	4.9330E-01	-6.7048E-03
2.0551E+04	4.3491E+04	1.1000E+00	1.2094E+00	1.0942E-01
2.0560E+04	4.3500E+04	7.0000E-01	9.8375E-01	2.8375E-01
2.0560E+04	4.3510E+04	3.0000E-01	3.3156E-01	3.1565E-02
2.0560E+04	4.3520E+04	7.0000E-01	3.1301E-01	-3.8699E-01
2.0560E+04	4.3530E+04	4.0000E-01	5.6557E-01	1.6557E-01
2.0550E+04	4.3500E+04	6.0000E-01	6.2242E-01	2.2423E-02
2.0550E+04	4.3520E+04	4.0000E-01	2.7680E-01	-1.2320E-01
2.0550E+04	4.3530E+04	3.0000E-01	3.7294E-01	7.2945E-02
2.0560E+04	4.3550E+04	6.0000E-01	7.4959E-01	1.4959E-01
2.0560E+04	4.3560E+04	7.0000E-01	6.7930E-01	-2.0704E-02
2.0560E+04	4.3570E+04	8.0000E-01	6.3241E-01	-1.6759E-01
2.0550E+04	4.3540E+04	4.0000E-01	4.5595E-01	5.5948E-02
2.0550E+04	4.3560E+04	8.0000E-01	5.6377E-01	-2.3623E-01
2.0550E+04	4.3570E+04	9.0000E-01	7.1485E-01	-1.8515E-01
2.0550E+04	4.3570E+04	3.0000E-01	7.1485E-01	4.1485E-01
2.0550E+04	4.3550E+04	4.0000E-01	4.9890E-01	9.8897E-02
2.0560E+04	4.3580E+04	6.0000E-01	6.9623E-01	9.6229E-02
2.0560E+04	4.3590E+04	9.0000E-01	8.8977E-01	-1.0228E-02
2.0550E+04	4.3580E+04	1.0000E+00	9.1230E-01	-8.7700E-02
2.0550E+04	4.3590E+04	5.0000E-01	5.5326E-01	5.3257E-02

SUM OF SQUARED RESIDUALS 8.4245E+00

	2.3666E+00		1.6835E-01
SPLINE COEFFS	5.9980E+00		8.5053E-02
4.0223E-01			2.2265E-01
3.7586E-01			4.3698E-01
3.8512E-01	7.5624E-01		
-1.1662E-01	2.5754E-01		
8.8966E-01	1.1041E-01		-9.1381E-01
5.9086E-02	-1.0679E+00		3.0524E+00
5.0265E-01	2.5435E+00		-1.0357E+00
	4.2304E-01		9.1823E-01
	-4.3748E-01		6.3420E-01
2.5321E-01			1.8963E+00
4.8750E-01			5.7491E-01
-2.4998E-01	-2.0488E-01		
-1.2841E+00	1.5654E+00		
3.5921E+00	1.8607E+00		4.0882E-01
-1.5962E+00	8.7458E-01		3.3548E+00
-2.9458E+00	9.9563E-01		-7.0983E-01
	4.7247E-01		1.1929E+00
	1.4828E+00		4.7005E-01
5.0153E-01			7.2695E-01
4.3625E-01			8.8977E-01
7.8521E-01	2.3170E+00		
2.3342E+00	-9.2357E-04		
-3.0605E+00	8.3466E-01		

Table 9.28. Modelling of the TREND by fitting a BICUBIC SPLINE SURFACE to the variation of POTASSIUM over the 200m grid at Yelford.

BICUBIC SPLINE SURFACE, POTASSIUM, 200m GRID.

INTERIOR X-KNOTS

2.0390E+04
2.0460E+04
2.0530E+04

INTERIOR Y-KNOTS

4.3483E+04
4.3525E+04
4.3569E+04

SUM OF SQUARES OF RESIDUAL RHS 8.7916E+00

RANK 49

NORTHING	EASTING	OBSERVED	SPLINE FIT	RESIDUALS
2.0350E+04	4.3450E+04	6.0000E-01	6.0157E-01	1.5741E-03
2.0370E+04	4.3470E+04	3.0000E-01	2.7398E-01	-2.6020E-02
2.0370E+04	4.3450E+04	4.0000E-01	3.7722E-01	-2.2779E-02
2.0350E+04	4.3470E+04	3.0000E-01	3.2408E-01	2.4081E-02
2.0370E+04	4.3510E+04	6.0000E-01	8.5442E-01	2.5442E-01
2.0350E+04	4.3490E+04	1.7000E+00	1.5928E+00	-1.0723E-01
2.0370E+04	4.3490E+04	2.0000E-01	4.1785E-01	2.1785E-01
2.0351E+04	4.3509E+04	8.0000E-01	9.0390E-01	1.0390E-01
2.0370E+04	4.3550E+04	2.0000E-01	4.4006E-02	-1.5599E-01
2.0350E+04	4.3530E+04	4.0000E-01	2.6572E-01	-1.3428E-01
2.0370E+04	4.3530E+04	5.0000E-01	7.2882E-01	2.2882E-01
2.0350E+04	4.3550E+04	8.0000E-01	8.6520E-01	6.5202E-02
2.0370E+04	4.3590E+04	5.0000E-01	6.6540E-01	1.6540E-01
2.0350E+04	4.3570E+04	8.0000E-01	8.0339E-01	3.3945E-03
2.0370E+04	4.3570E+04	7.0000E-01	6.1129E-01	-8.8708E-02
2.0350E+04	4.3590E+04	5.0000E-01	4.7933E-01	-2.0675E-02
2.0450E+04	4.3470E+04	1.0000E+00	6.5323E-01	-3.4677E-01
2.0430E+04	4.3450E+04	9.0000E-01	9.1999E-01	1.9987E-02
2.0410E+04	4.3470E+04	3.0000E-01	4.4677E-01	1.4677E-01
2.0390E+04	4.3450E+04	2.0000E-01	2.6661E-01	6.6615E-02
2.0430E+04	4.3470E+04	6.0000E-01	5.6709E-01	-3.2913E-02
2.0410E+04	4.3450E+04	8.0000E-01	7.4205E-01	-5.7954E-02
2.0390E+04	4.3470E+04	5.0000E-01	3.3408E-01	-1.6592E-01
2.0450E+04	4.3510E+04	2.0000E-01	2.4956E-01	4.9561E-02
2.0430E+04	4.3490E+04	4.0000E-01	5.1718E-01	1.1718E-01
2.0410E+04	4.3510E+04	6.0000E-01	8.1096E-01	2.1096E-01
2.0390E+04	4.3490E+04	5.0000E-01	6.3549E-01	1.3549E-01
2.0450E+04	4.3490E+04	4.0000E-01	2.8482E-01	-1.1518E-01
2.0430E+04	4.3510E+04	3.0000E-01	5.2286E-01	2.2286E-01
2.0411E+04	4.3491E+04	4.0000E-01	7.1493E-01	3.1493E-01
2.0390E+04	4.3510E+04	2.5000E+00	9.4483E-01	-1.5552E+00
2.0450E+04	4.3550E+04	4.0000E-01	1.7409E-01	-2.2591E-01
2.0430E+04	4.3530E+04	4.0000E-01	4.2970E-01	2.9699E-02
2.0410E+04	4.3550E+04	3.0000E-01	3.6675E-01	6.6748E-02
2.0390E+04	4.3530E+04	4.0000E-01	8.0225E-01	4.0225E-01
2.0450E+04	4.3530E+04	1.0000E-01	2.5775E-01	1.5775E-01
2.0430E+04	4.3550E+04	4.0000E-01	2.6465E-01	-1.3535E-01
2.0410E+04	4.3530E+04	4.0000E-01	6.5566E-01	2.5566E-01
2.0390E+04	4.3550E+04	4.0000E-01	2.7059E-01	-1.2941E-01
2.0450E+04	4.3590E+04	5.0000E-01	1.2222E+00	7.2215E-01

2.0430E+04	4.3570E+04	6.0000E-01	3.9248E-01	-2.0752E-01
2.0410E+04	4.3590E+04	2.0000E-01	1.5133E-01	-4.8673E-02
2.0390E+04	4.3570E+04	2.0000E-01	4.6507E-01	2.6507E-01
2.0450E+04	4.3570E+04	4.0000E-01	4.8833E-01	8.8335E-02
2.0430E+04	4.3590E+04	5.0000E-01	6.5030E-01	1.5030E-01
2.0410E+04	4.3570E+04	5.0000E-01	3.8797E-01	-1.1203E-01
2.0390E+04	4.3590E+04	4.0000E-01	1.3248E-01	-2.6752E-01
2.0490E+04	4.3470E+04	7.0000E-01	6.2104E-01	-7.8961E-02
2.0510E+04	4.3470E+04	4.0000E-01	5.9051E-01	1.9051E-01
2.0470E+04	4.3470E+04	3.0000E-01	6.6584E-01	3.6584E-01
2.0510E+04	4.3490E+04	4.0000E-01	6.4103E-01	2.4103E-01
2.0490E+04	4.3510E+04	5.0000E-01	3.3380E-01	-1.6620E-01
2.0470E+04	4.3490E+04	4.0000E-01	2.3330E-01	-1.6670E-01
2.0490E+04	4.3490E+04	5.0000E-01	4.0030E-01	-9.9696E-02
2.0470E+04	4.3510E+04	2.0000E-01	1.5679E-01	-4.3209E-02
2.0510E+04	4.3530E+04	3.0000E-01	6.0567E-01	3.0567E-01
2.0490E+04	4.3550E+04	6.0000E-01	5.9511E-01	-4.8938E-03
2.0470E+04	4.3530E+04	6.0000E-01	2.6684E-01	-3.3316E-01
2.0510E+04	4.3550E+04	1.3000E+00	8.1211E-01	-4.8789E-01
2.0490E+04	4.3530E+04	2.0000E-01	4.3789E-01	2.3789E-01
2.0470E+04	4.3550E+04	3.0000E-01	2.9478E-01	-5.2201E-03
2.0510E+04	4.3570E+04	8.0000E-01	1.0433E+00	2.4330E-01
2.0510E+04	4.3610E+04	7.0000E-01	7.6879E-01	6.8788E-02
2.0490E+04	4.3590E+04	6.0000E-01	1.3137E+00	7.1367E-01
2.0470E+04	4.3570E+04	2.0000E-01	6.8229E-01	4.8229E-01
2.0509E+04	4.3571E+04	1.6000E+00	1.0453E+00	-5.5468E-01
2.0510E+04	4.3590E+04	1.0000E+00	9.6071E-01	-3.9293E-02
2.0490E+04	4.3570E+04	8.0000E-01	9.1144E-01	1.1144E-01
2.0470E+04	4.3590E+04	2.9000E+00	1.4729E+00	-1.4271E+00
2.0570E+04	4.3470E+04	6.0000E-01	5.7810E-01	-2.1900E-02
2.0530E+04	4.3470E+04	1.0000E+00	6.4834E-01	-3.5166E-01
2.0550E+04	4.3470E+04	6.0000E-01	7.7520E-01	1.7520E-01
2.0570E+04	4.3510E+04	7.0000E-01	7.0639E-01	6.3913E-03
2.0551E+04	4.3491E+04	1.1000E+00	1.0484E+00	-5.1620E-02
2.0531E+04	4.3511E+04	1.6000E+00	1.5435E+00	-5.6458E-02
2.0570E+04	4.3490E+04	7.0000E-01	7.0634E-01	6.3444E-03
2.0530E+04	4.3490E+04	8.0000E-01	8.0270E-01	2.6987E-03
2.0570E+04	4.3550E+04	5.0000E-01	5.5262E-01	5.2620E-02
2.0550E+04	4.3530E+04	3.0000E-01	3.9460E-01	9.4603E-02
2.0530E+04	4.3550E+04	6.0000E-01	6.7278E-01	7.2780E-02
2.0570E+04	4.3530E+04	6.0000E-01	5.6997E-01	-3.0034E-02
2.0550E+04	4.3550E+04	4.0000E-01	1.9404E-01	-2.0596E-01
2.0530E+04	4.3530E+04	6.0000E-01	5.9858E-01	-1.4243E-03
2.0570E+04	4.3590E+04	3.0000E-01	3.0744E-01	7.4352E-03
2.0550E+04	4.3570E+04	9.0000E-01	5.9561E-01	-3.0439E-01
2.0550E+04	4.3610E+04	7.0000E-01	7.4129E-01	4.1291E-02
2.0530E+04	4.3590E+04	8.0000E-01	6.4350E-01	-1.5650E-01
2.0570E+04	4.3570E+04	5.0000E-01	4.8065E-01	-1.9346E-02
2.0570E+04	4.3610E+04	6.0000E-01	5.9383E-01	-6.1718E-03
2.0550E+04	4.3570E+04	3.0000E-01	5.9561E-01	2.9561E-01
2.0550E+04	4.3590E+04	5.0000E-01	5.0070E-01	6.9606E-04
2.0530E+04	4.3570E+04	5.0000E-01	9.4233E-01	4.4233E-01
2.0530E+04	4.3610E+04	7.0000E-01	6.2303E-01	-7.6974E-02

SUM OF SQUARED RESIDUALS 8.7987E+00

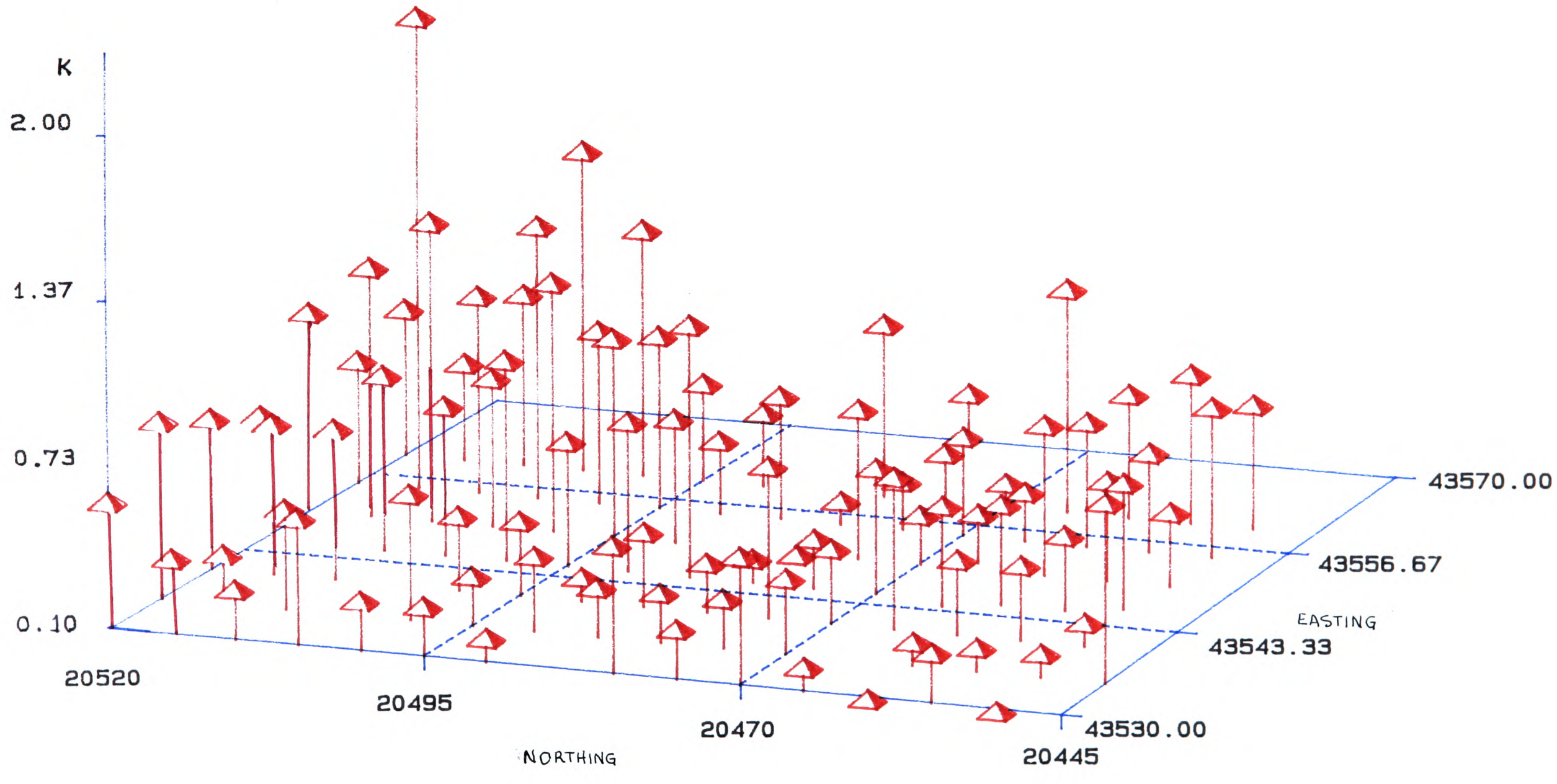
	-1.6643E+00		-2.3495E+00
SPLINE COEFFS	-2.8793E+00		1.0538E+00
6.0157E-01			3.0372E-01
-1.3211E+00			9.3652E-01
3.5054E+00	4.8020E+00		
-1.3446E+00	1.6251E+00		

2.9719E+00	-1.2722E+00	5.3196E+00
-2.4166E+00	3.8216E-01	-2.8595E-01
9.2356E+00	-1.6202E+00	1.1073E+00
	4.2684E+00	4.1552E-01
	2.3400E+00	6.9731E-01
5.9437E-01		8.3571E-02
9.5489E-01		5.9383E-01
-1.3884E+00	-2.4905E+01	
2.7799E+00	2.8539E+00	
-3.8355E+00	6.4205E-01	
8.5328E+00	6.0710E-01	
-2.2873E+01	2.6827E+00	
	-3.8462E-01	
	1.0633E-01	
-7.0420E-01		
-5.6563E-01		
2.1606E+00	2.4660E+02	
7.1430E-01	-2.9481E+01	
1.2267E+00	1.0916E+01	

Fig. 9.22a

POTASSIUM, EXCHANGE

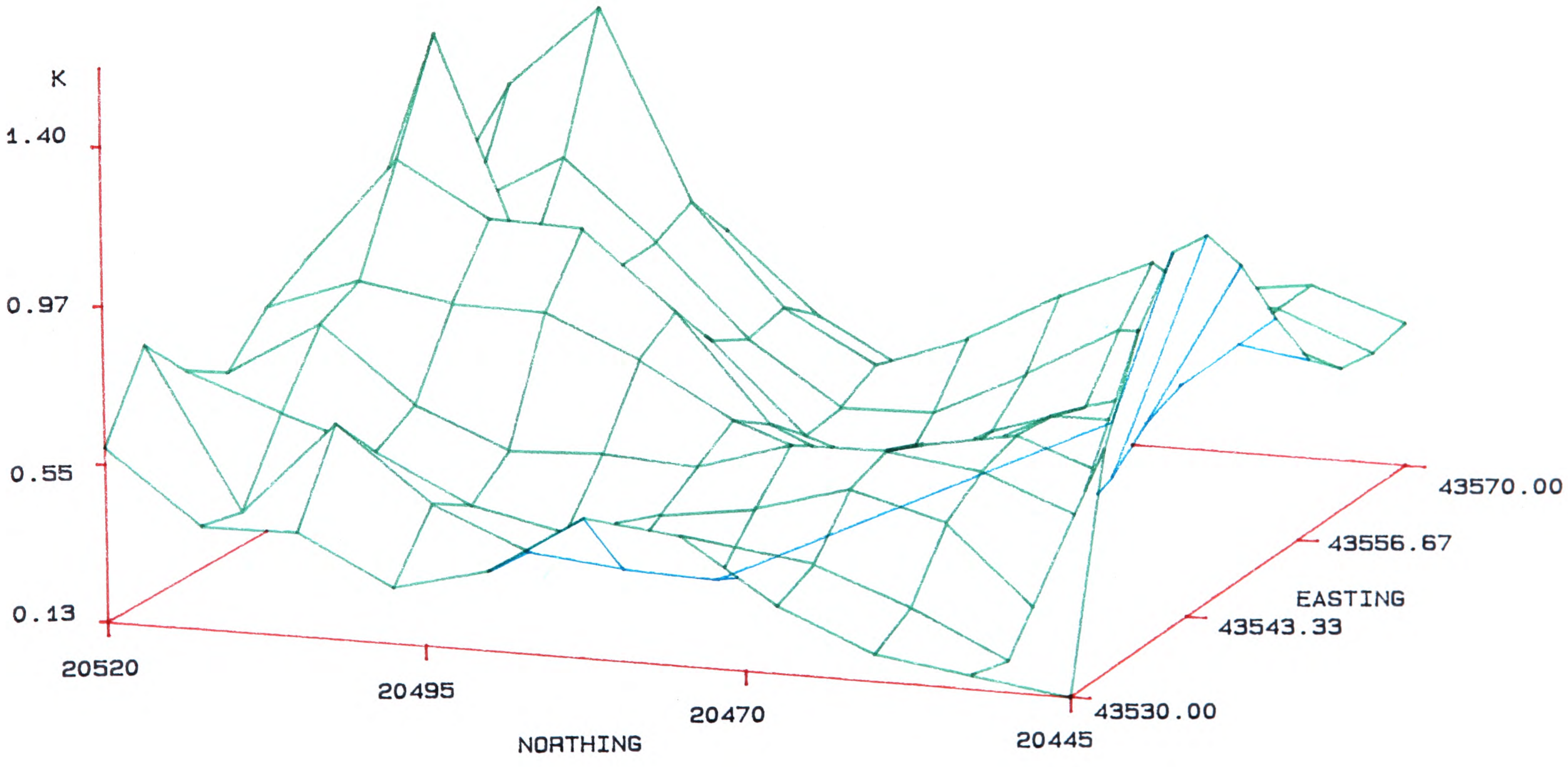
0–15cm.



YELFORD, 50m. GRID

Fig. 9.22 b

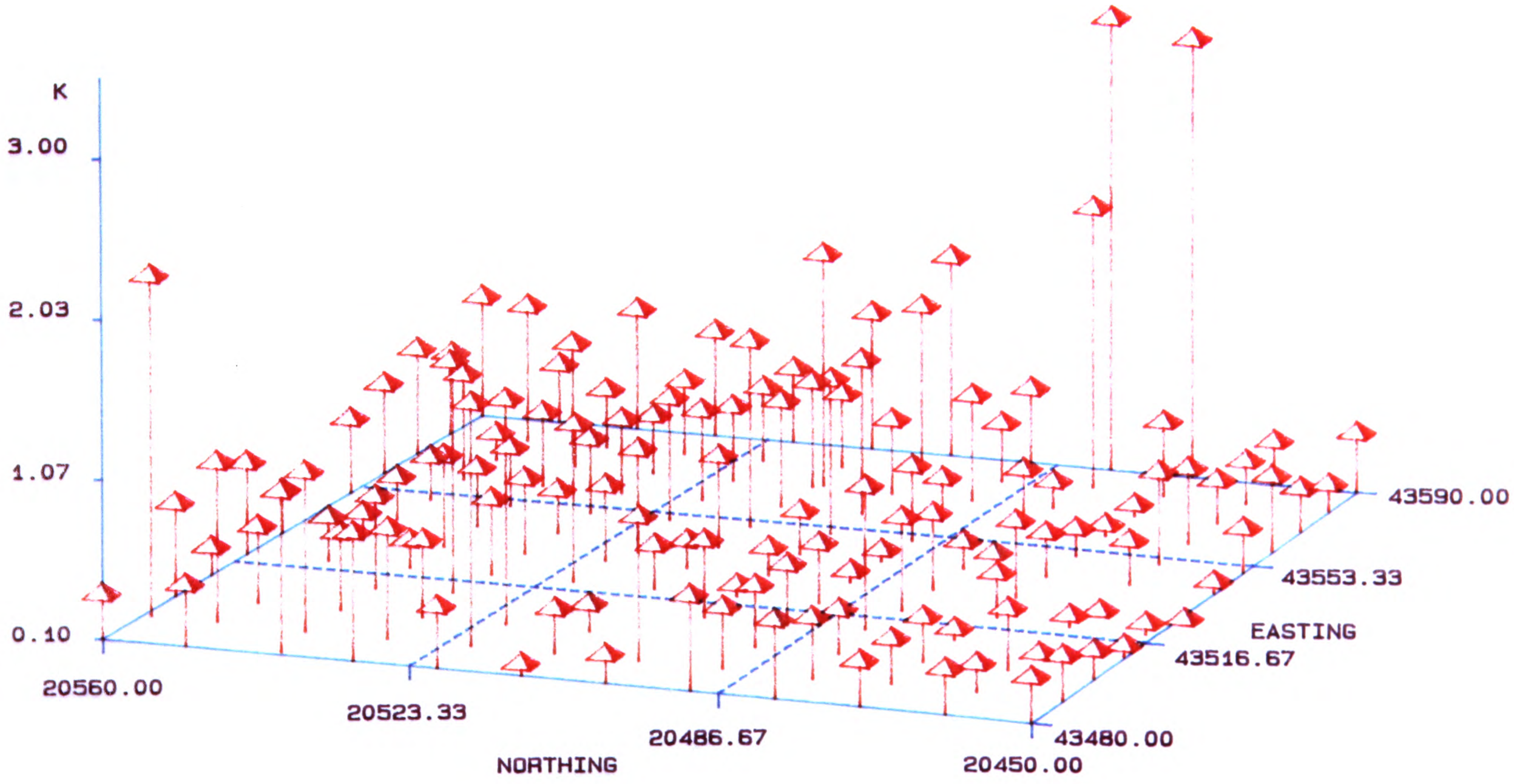
POTASSIUM
BICUBIC SPLINE SURFACE
0–15cm.



50 m GRID, YELFORD

Fig. 9.23a

POTASSIUM
0–15cm.



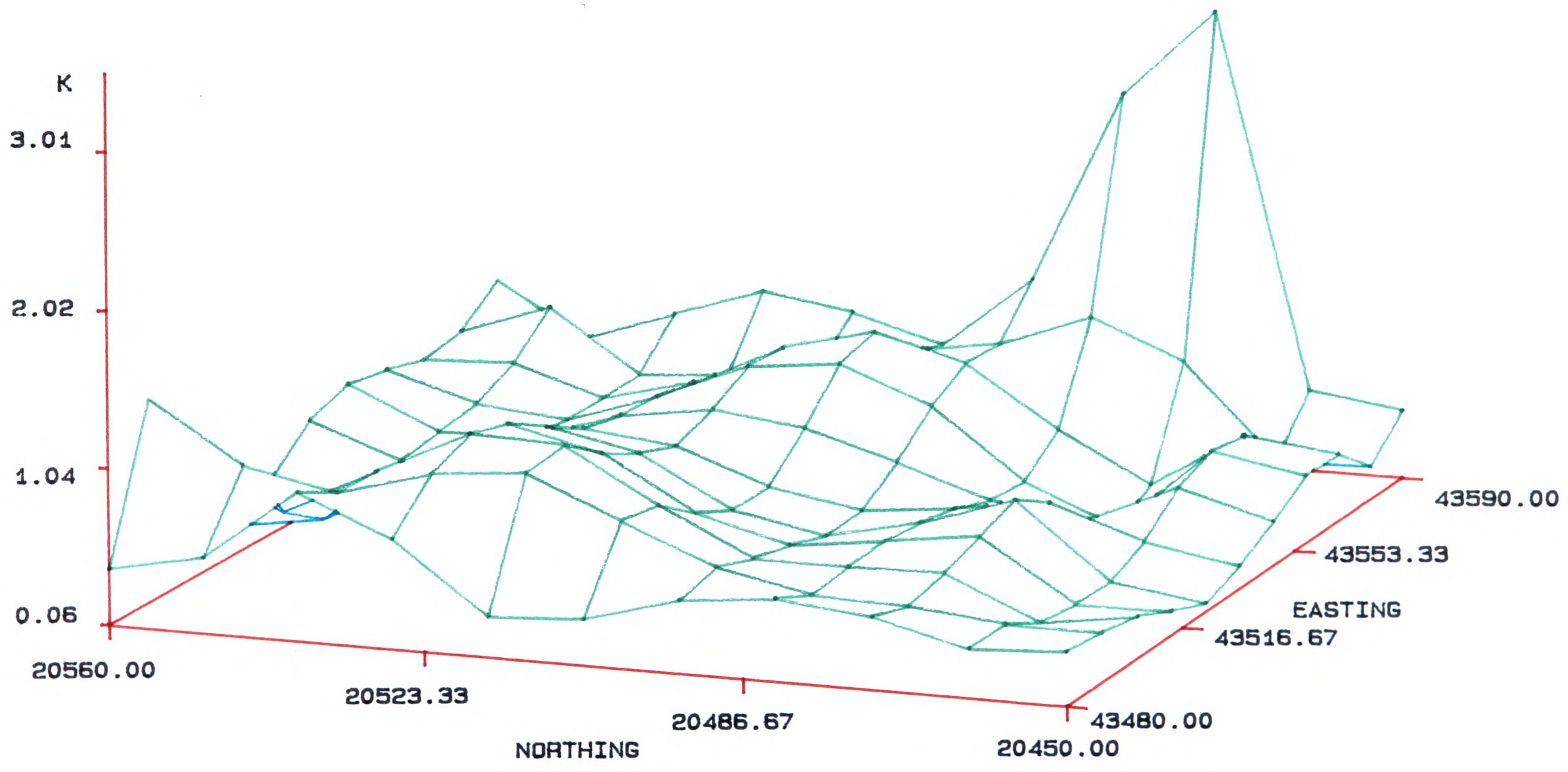
100 m GRID, YELFORD

Fig 9.23b

POTASSIUM

BICUBIC SPLINE SURFACE

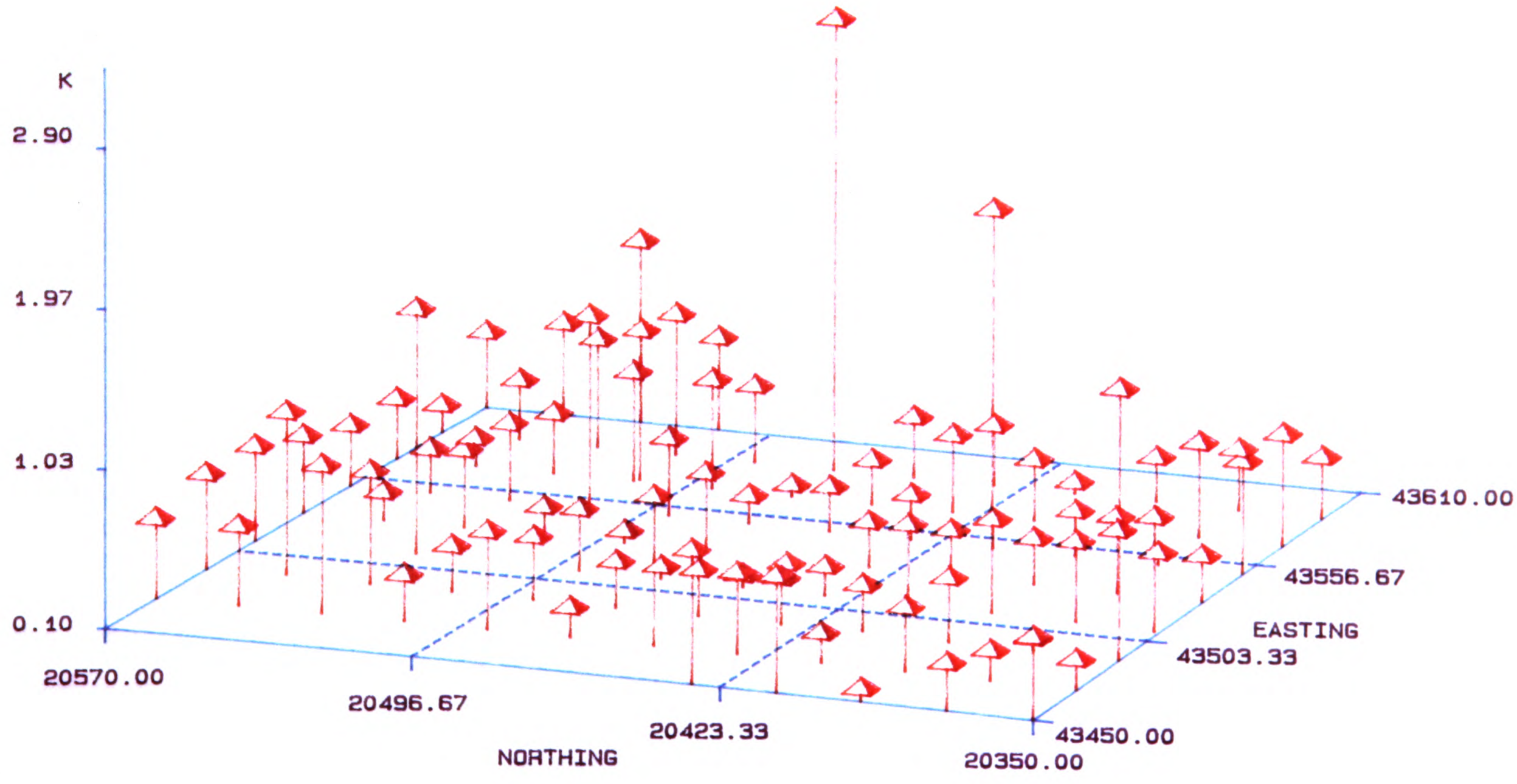
0–15cm.



100 m GRID, YELFORD

Fig. 9.24a

POTASSIUM
0–15cm.

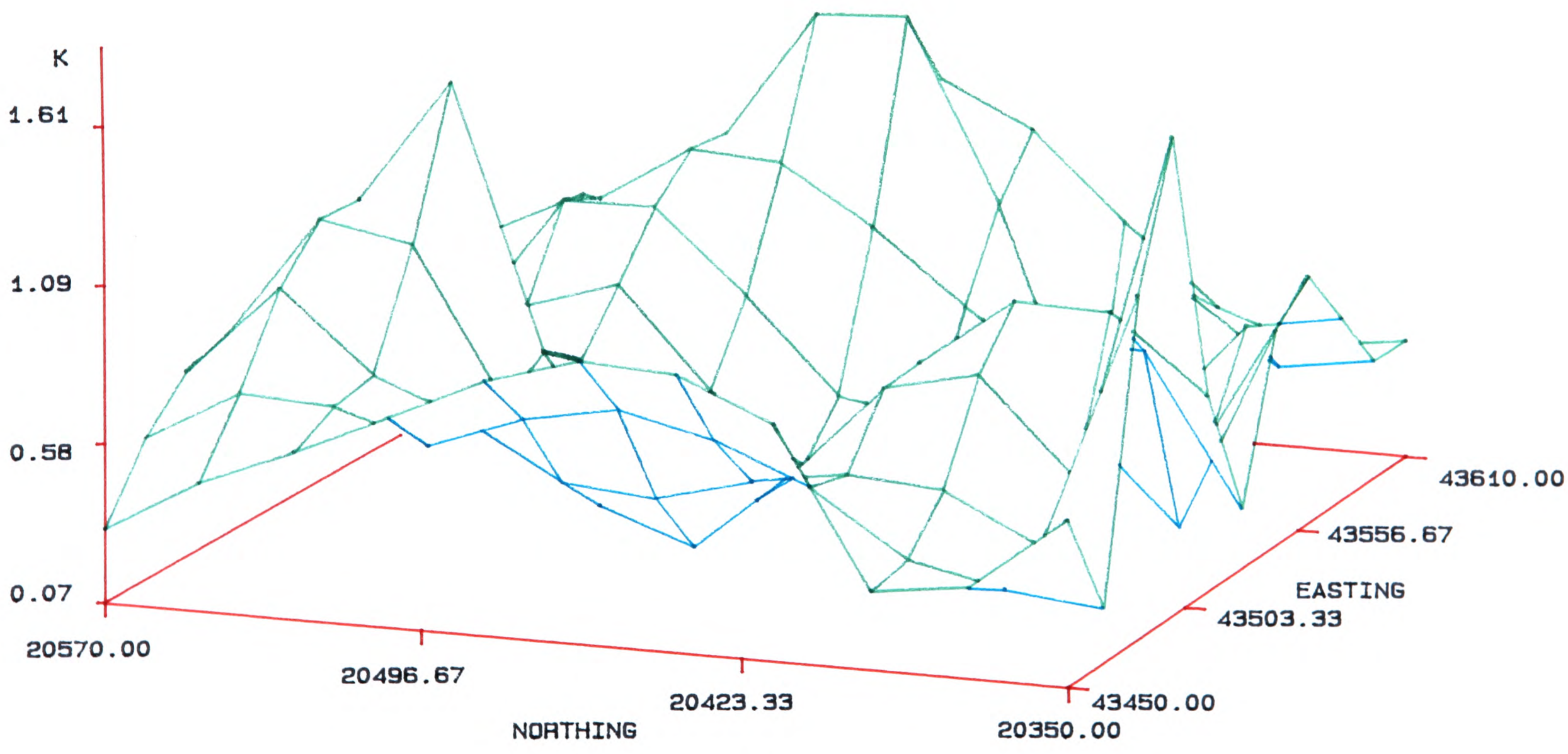


200 m GRID, YELFORD

Fig. 9.24 b. ***POTASSIUM***

BICUBIC SPLINE SURFACE

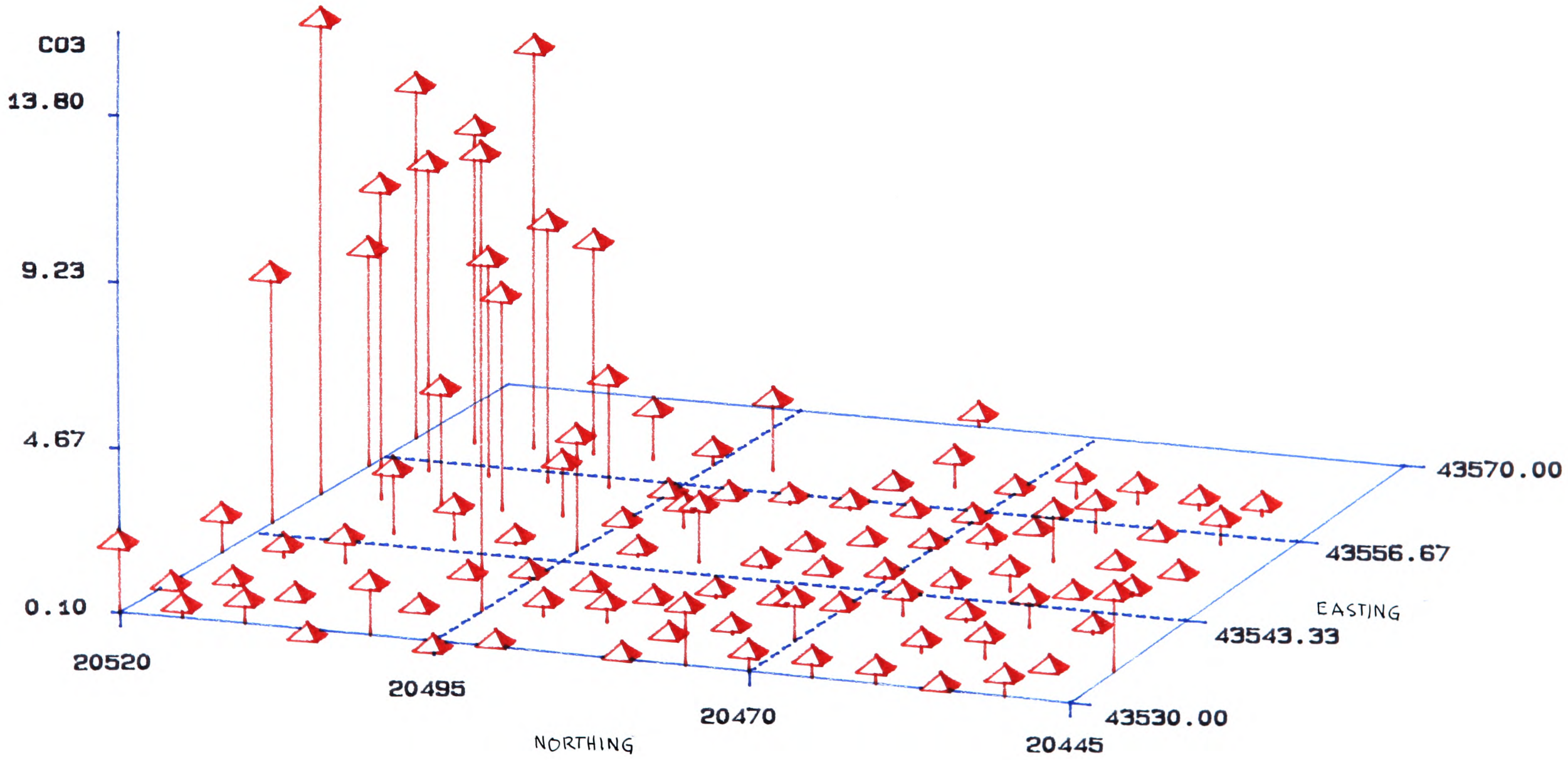
0–15cm.



200 m GRID, YELFORD

Fig 9.25a **CO3**

0–15cm.



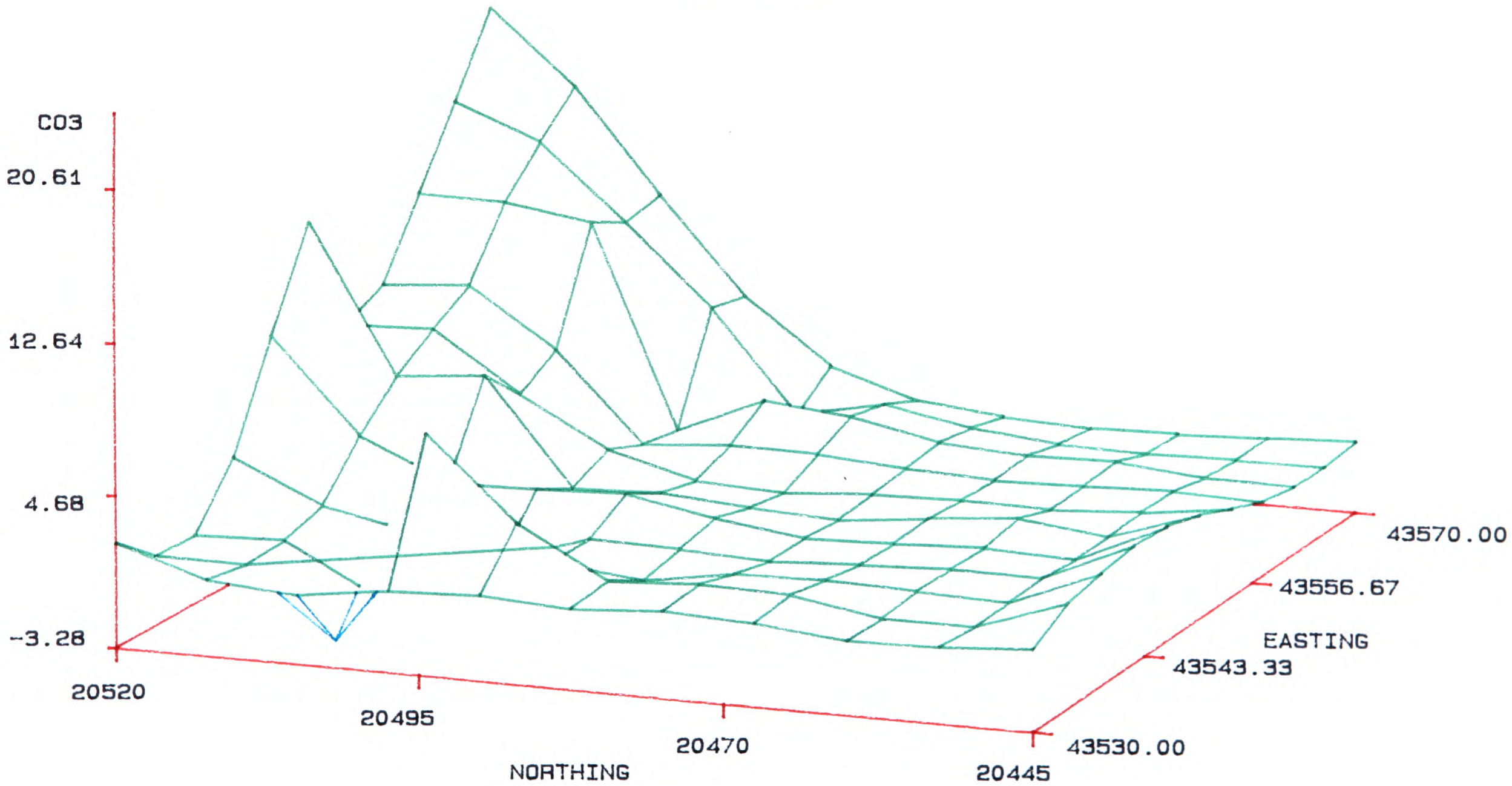
YELFORD 50 m GRID

Fig 9.25b

CARBONATE

BICUBIC SPLINE SURFACE

0–15cm.



50 m GRID, YELFORD

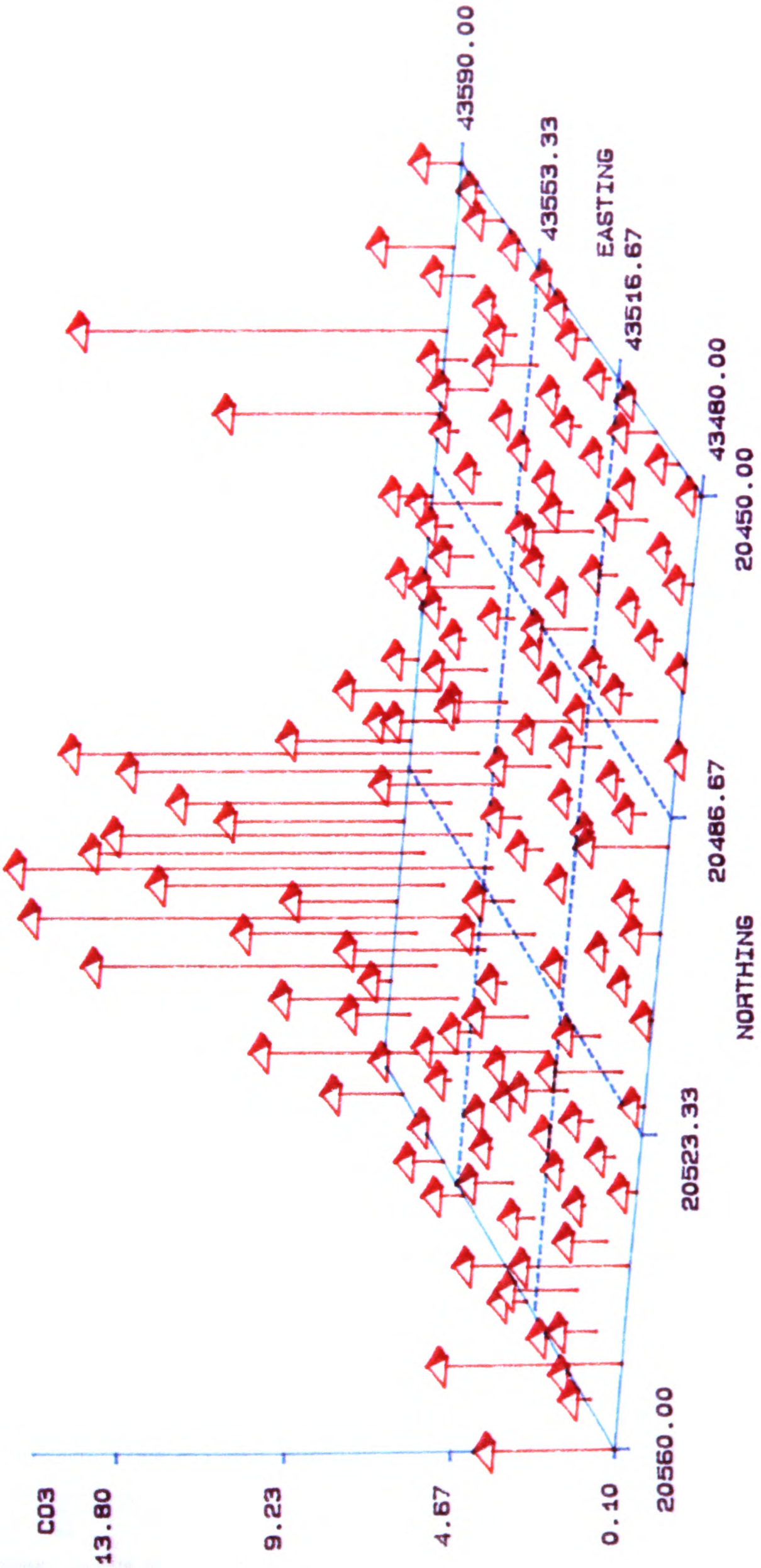

CARBONATE
0–15cm.
CO3
13.80
9.23
4.67
0.10
20560.00
20523.33
20486.67
20450.00
NORTHING
43480.00
43516.67
43553.33
43590.00
EASTING
100 m GRID, YELFORD

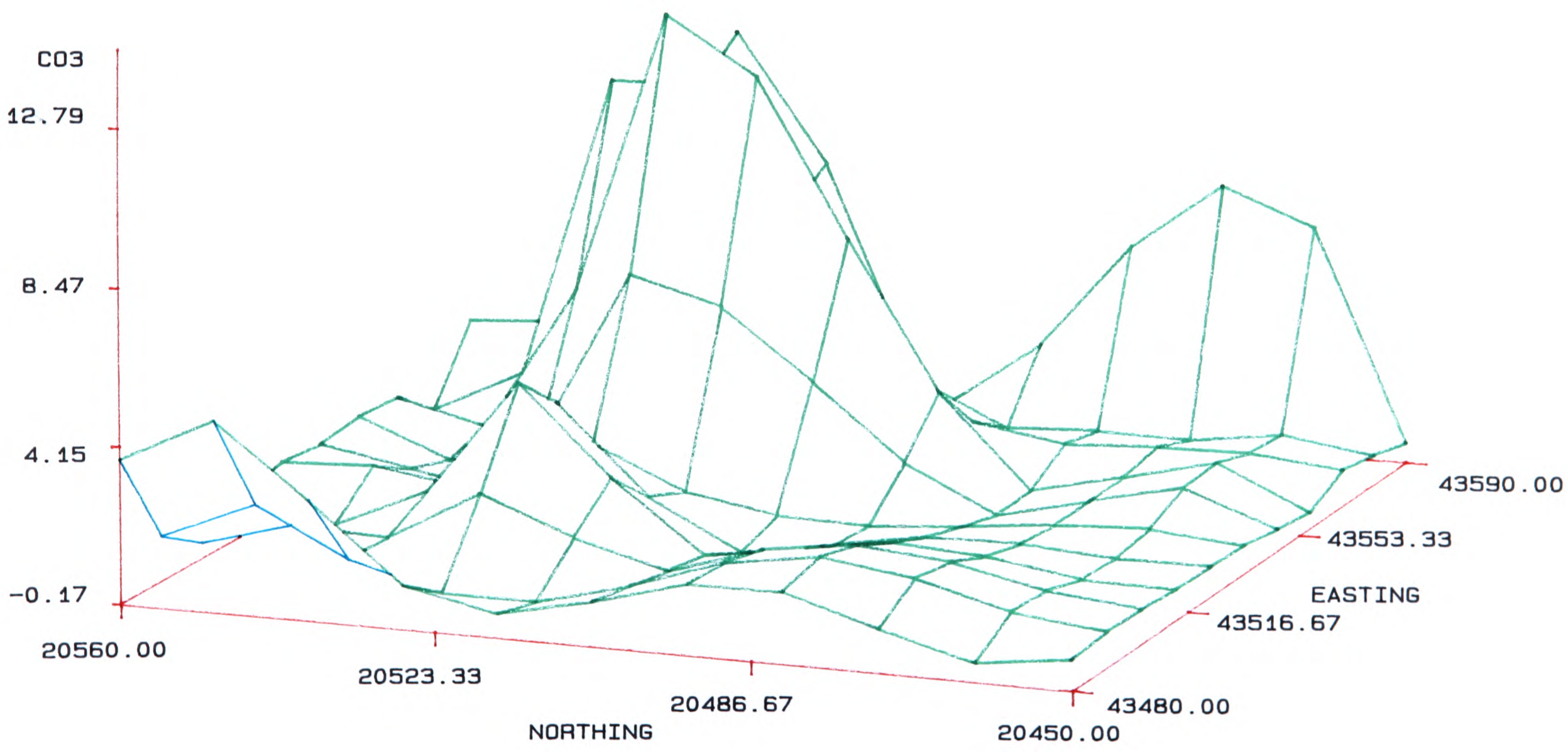
Fig. 9.26a

Fig. 9.26b

CARBONATE

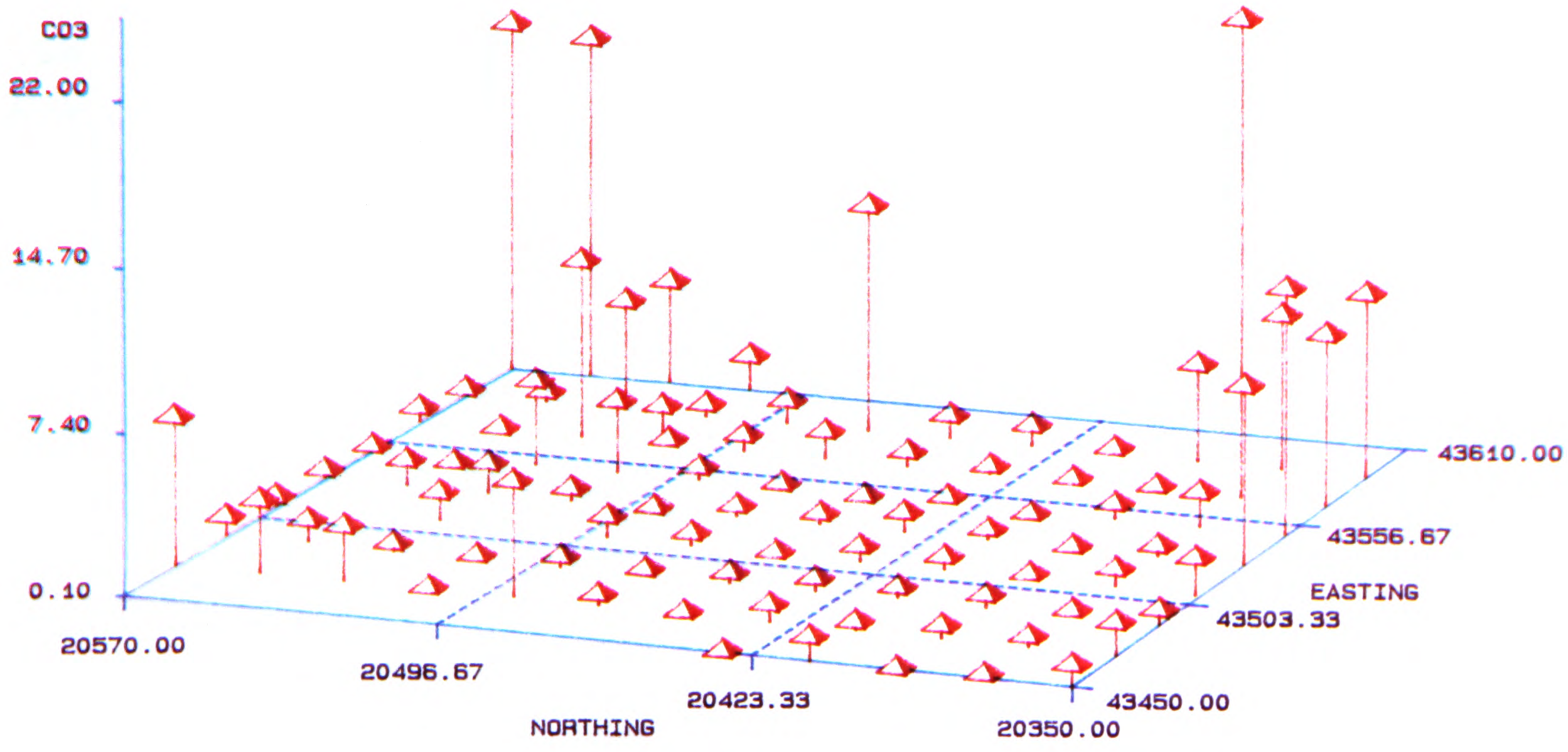
BICUBIC SPLINE SURFACE

0–15cm.



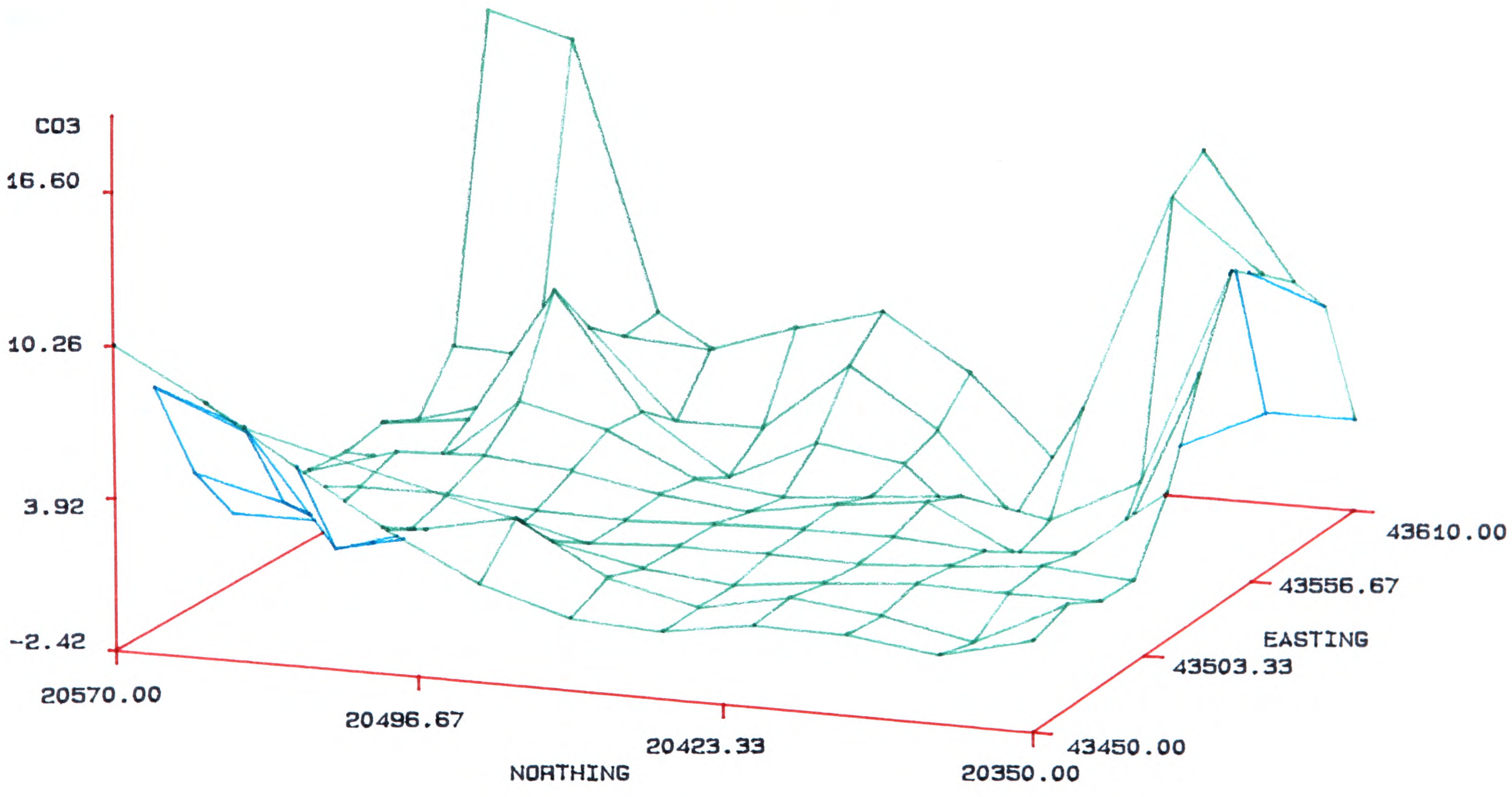
100 m GRID, YELFORD

Fig. 9.27a **CARBONATE**
0--15cm.



200 m GRID, YELFORD

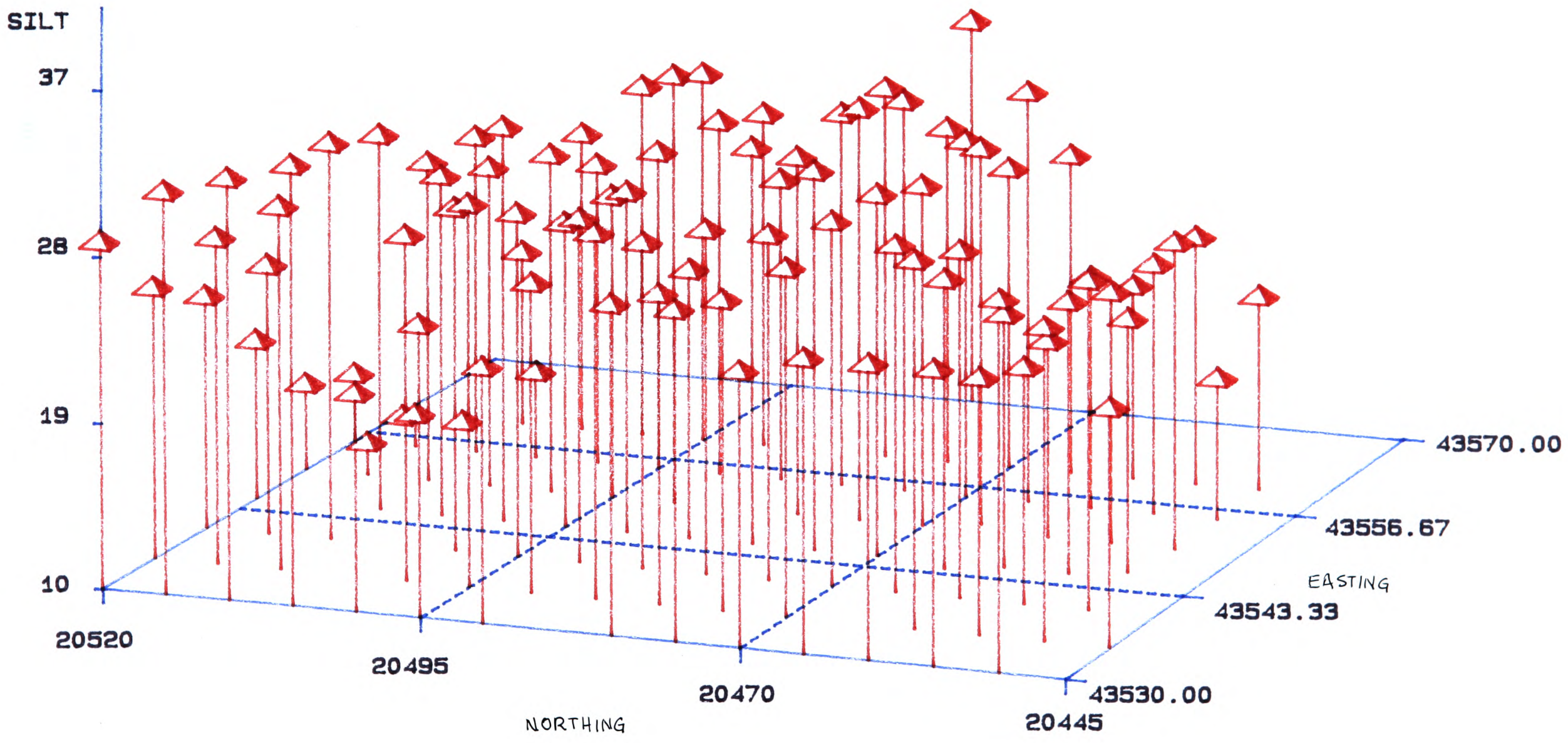
Fig. 9.27b



200 m GRID, YELFORD

Fig. 9.28a. ***SILT***

0–15cm.

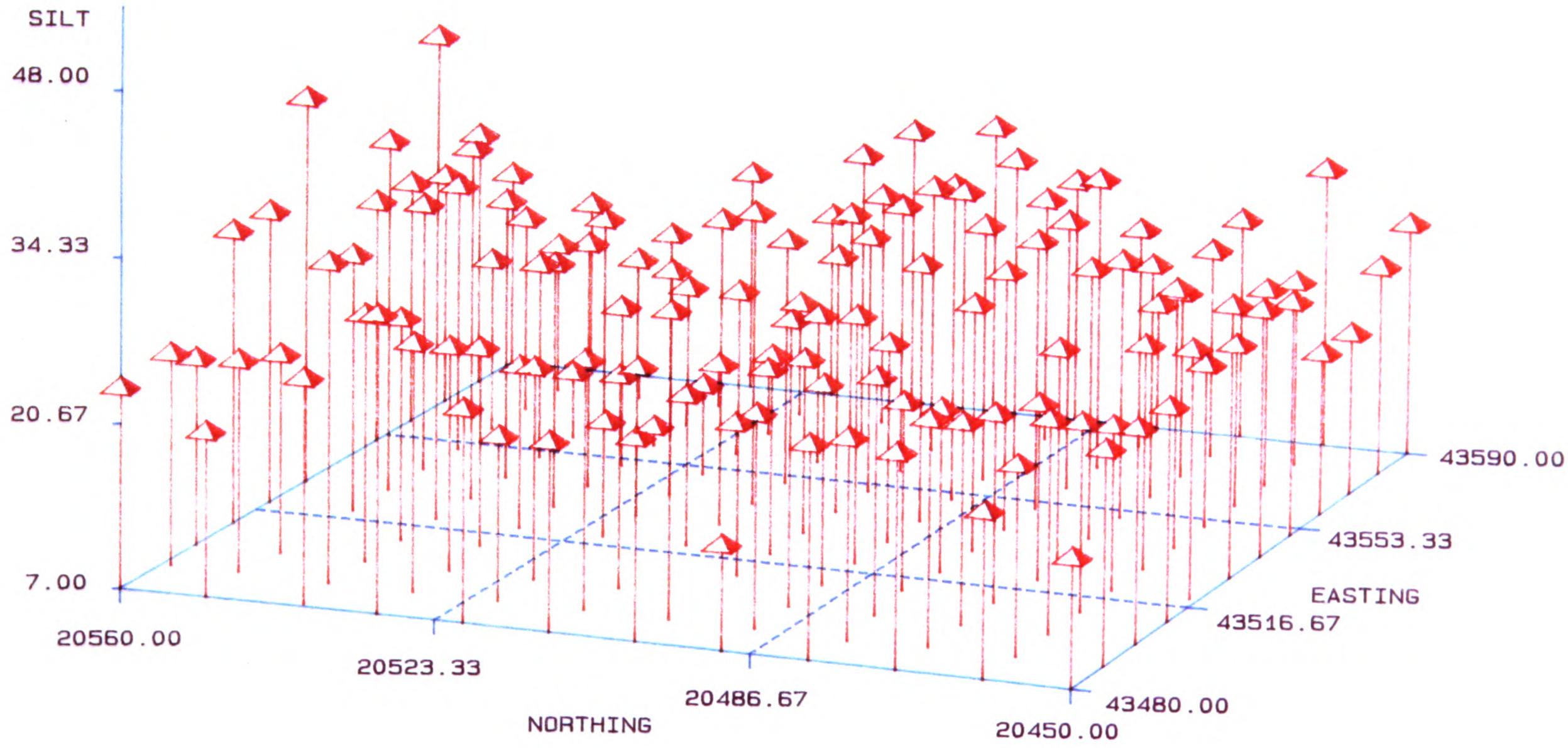


YELFORD, 50 m GRID

Fig. 9.29a

SILT

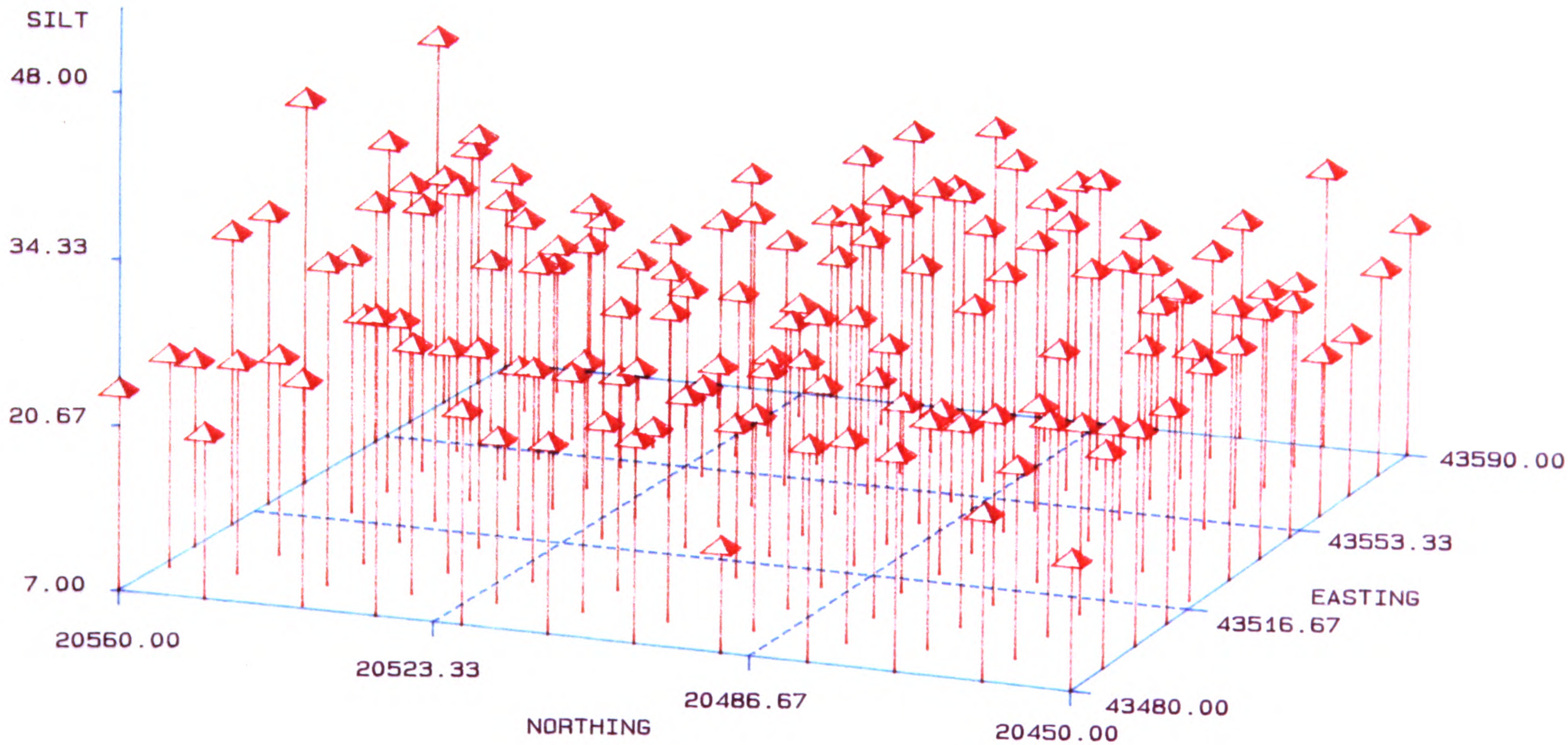
0–15cm.



100 m GRID, YELFORD

Fig. 9.29a

SILT
0–15cm.

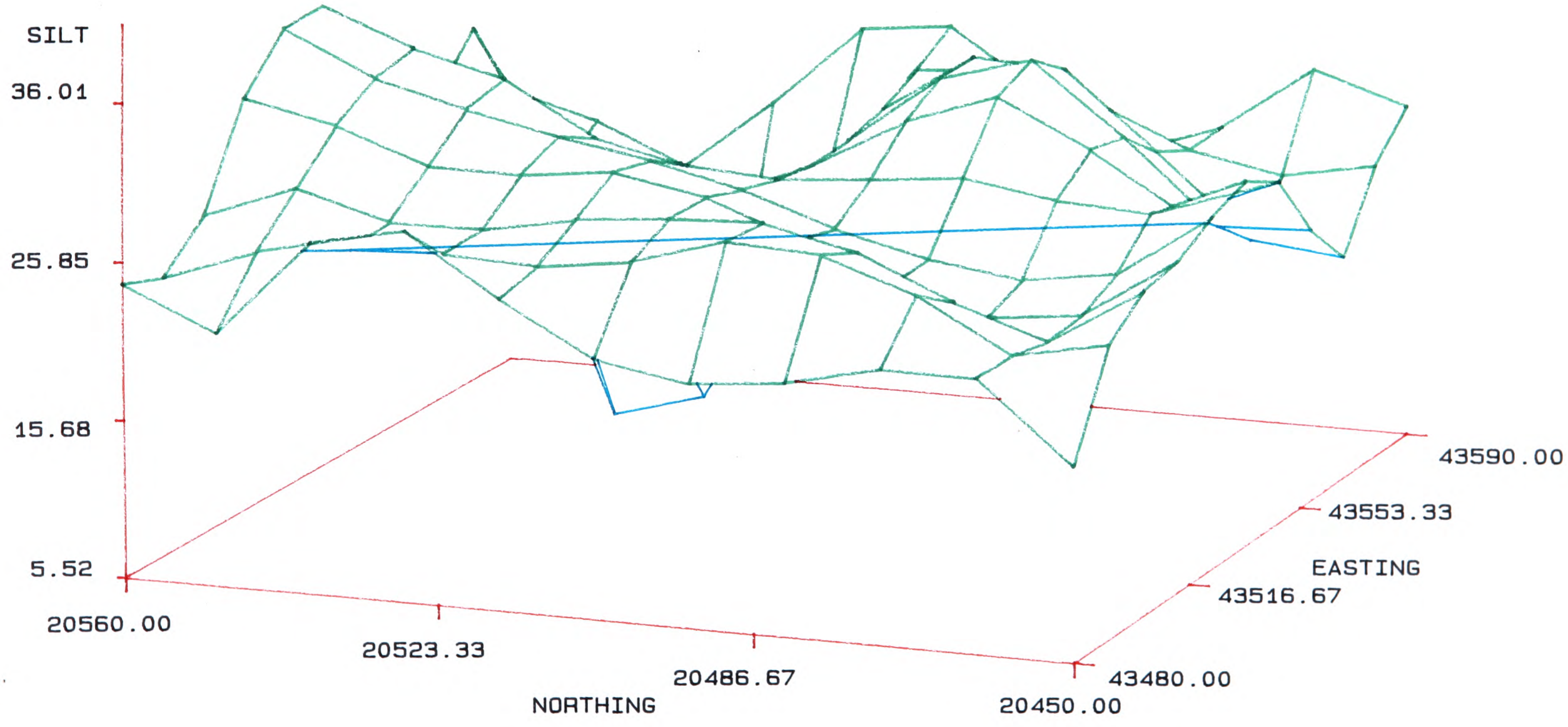


100 m GRID, YELFORD

Fig 9.29b *SILT*

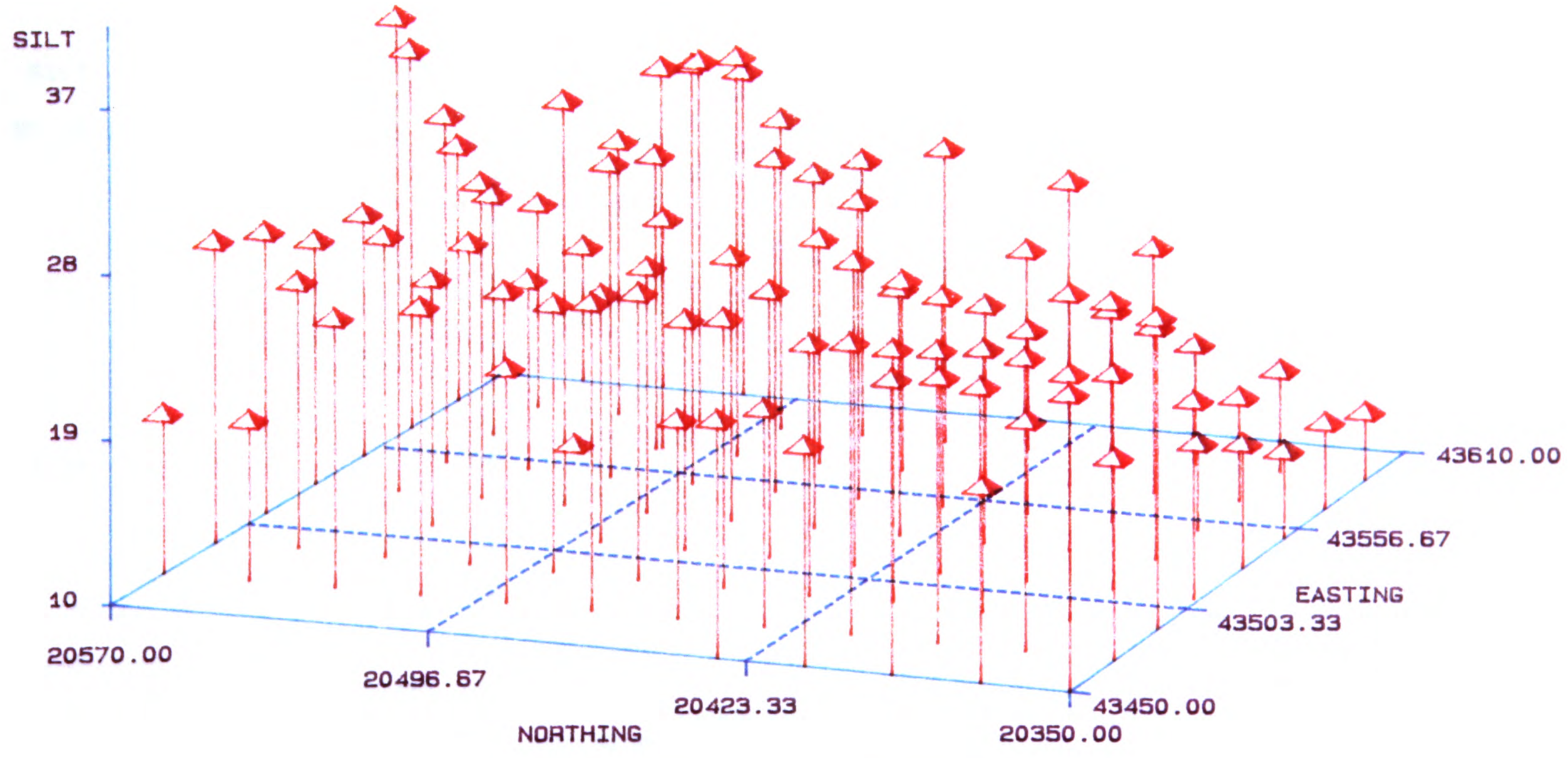
BICUBIC SPLINE SURFACE

0–15cm.



100 m GRID, YELFORD

Fig. 9.30a **SILT**
0-15cm.



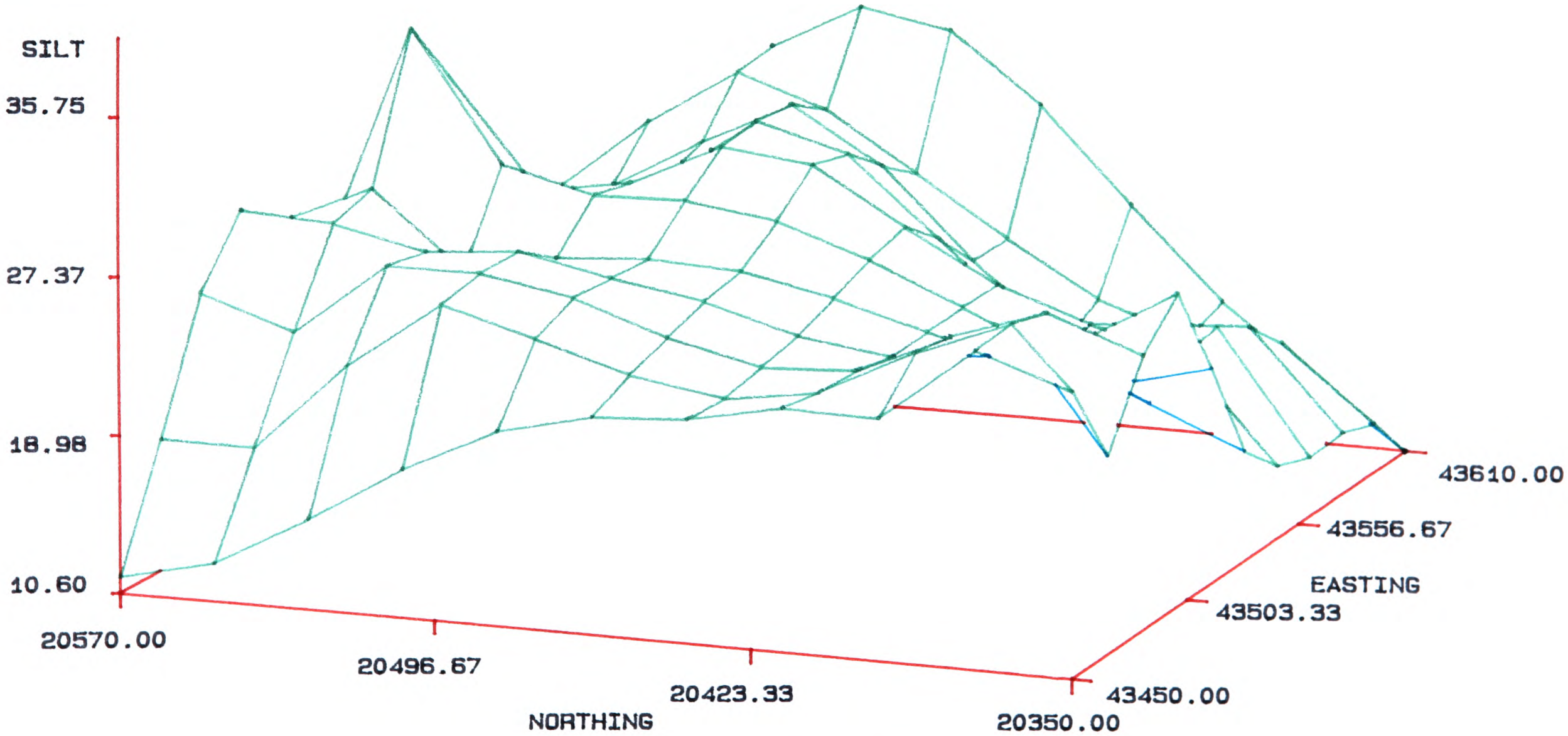
200 m GRID, YELFORD

Fig. 9.30b

SILT

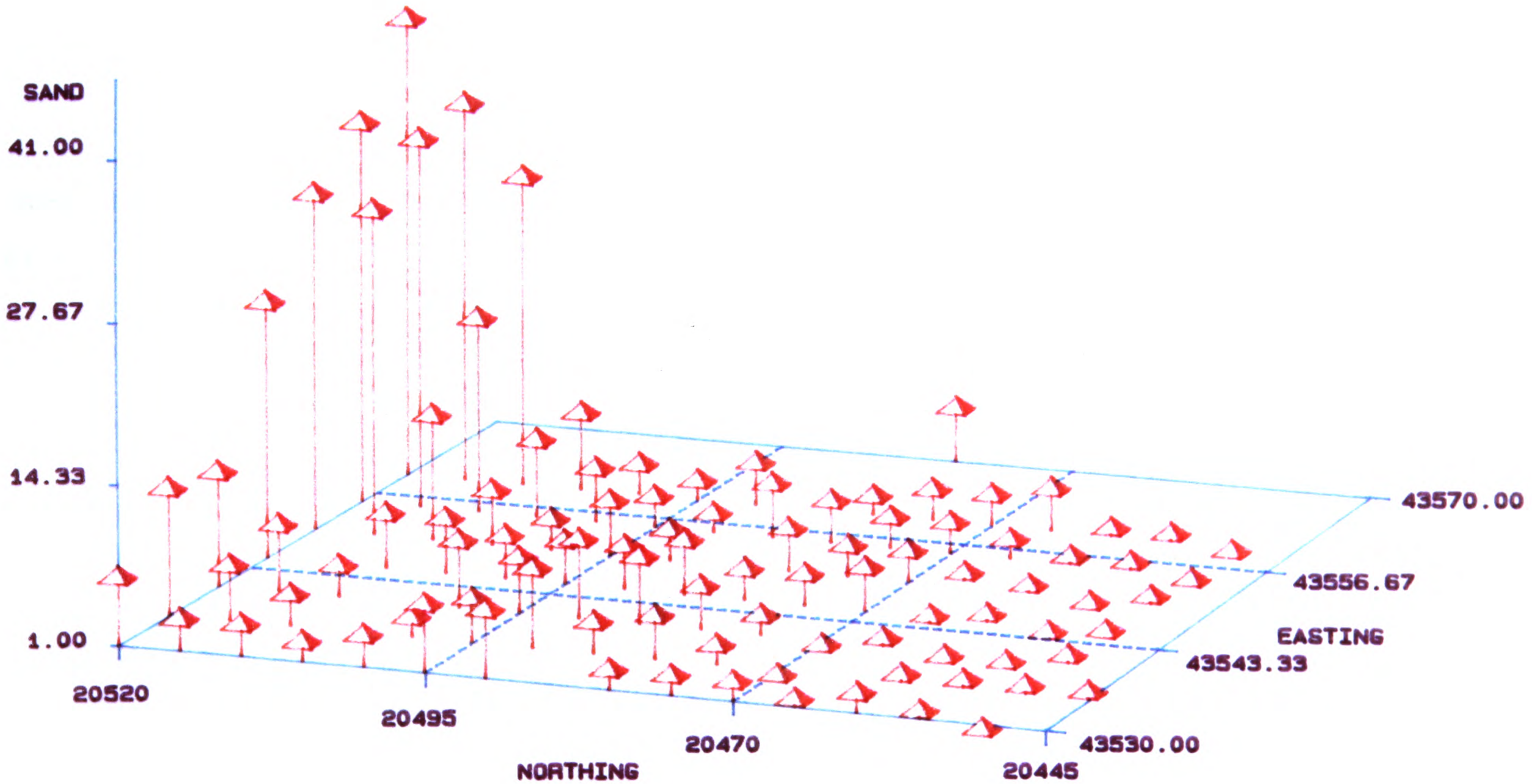
BICUBIC SPLINE SURFACE

0–15cm.



200 m GRID, YELFORD

Fig. 9.31a



SAND
0–15cm.

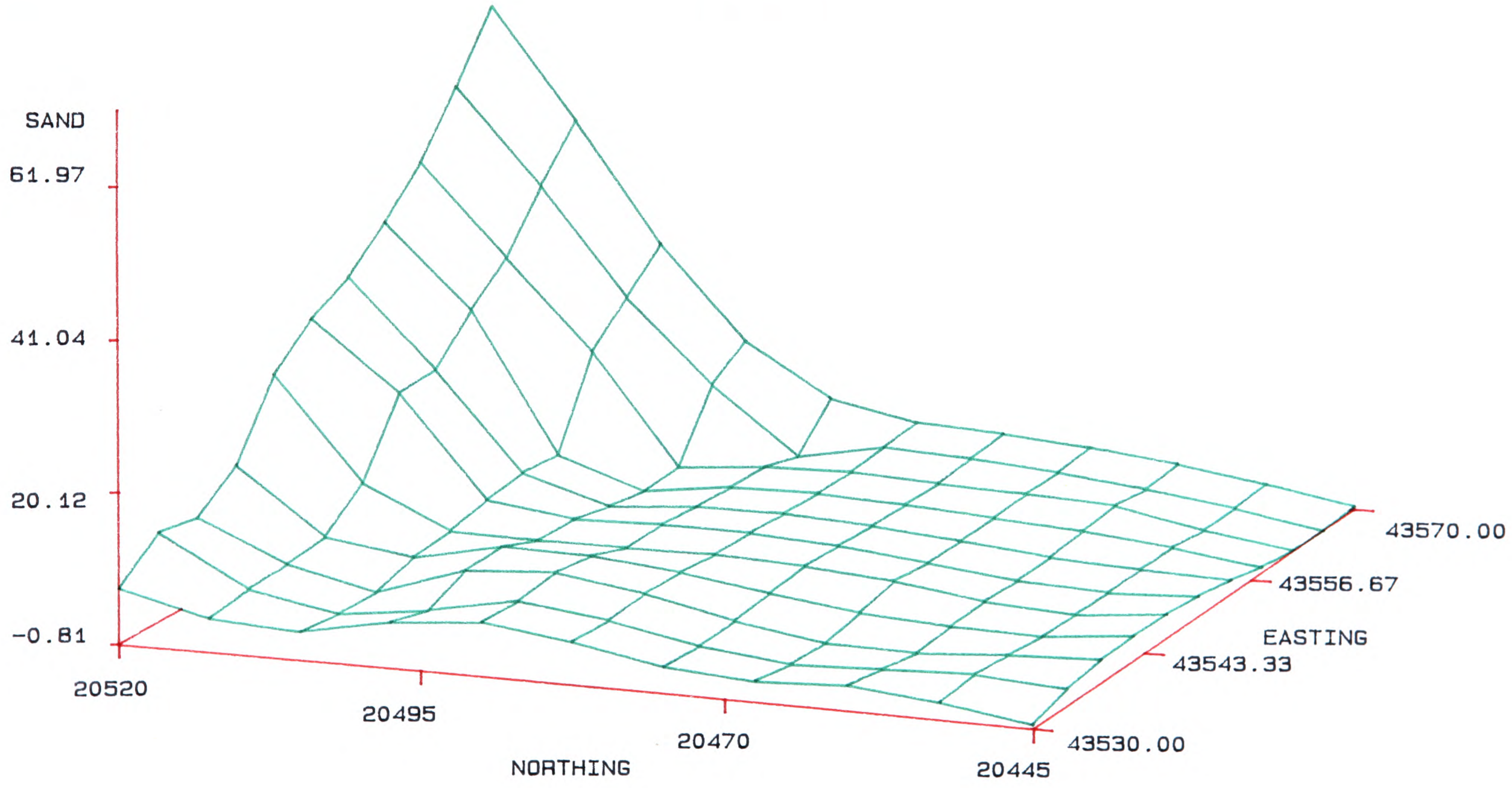
50 m GRID, YELFORD

Fig 9.31b

COARSE SAND

BICUBIC SPLINE SURFACE

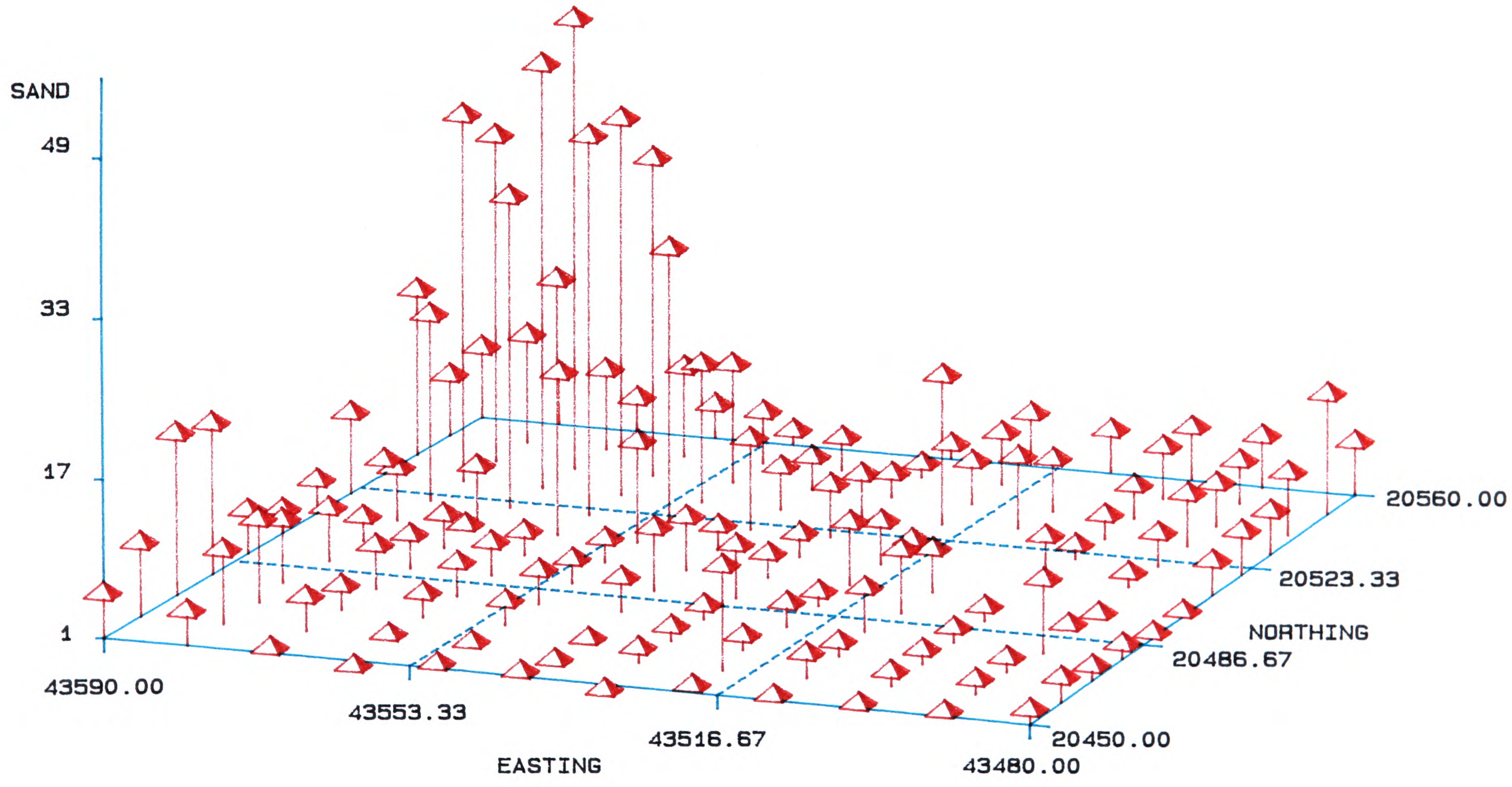
0–15cm.



50 m GRID, YELFORD

SAND

0–15cm.



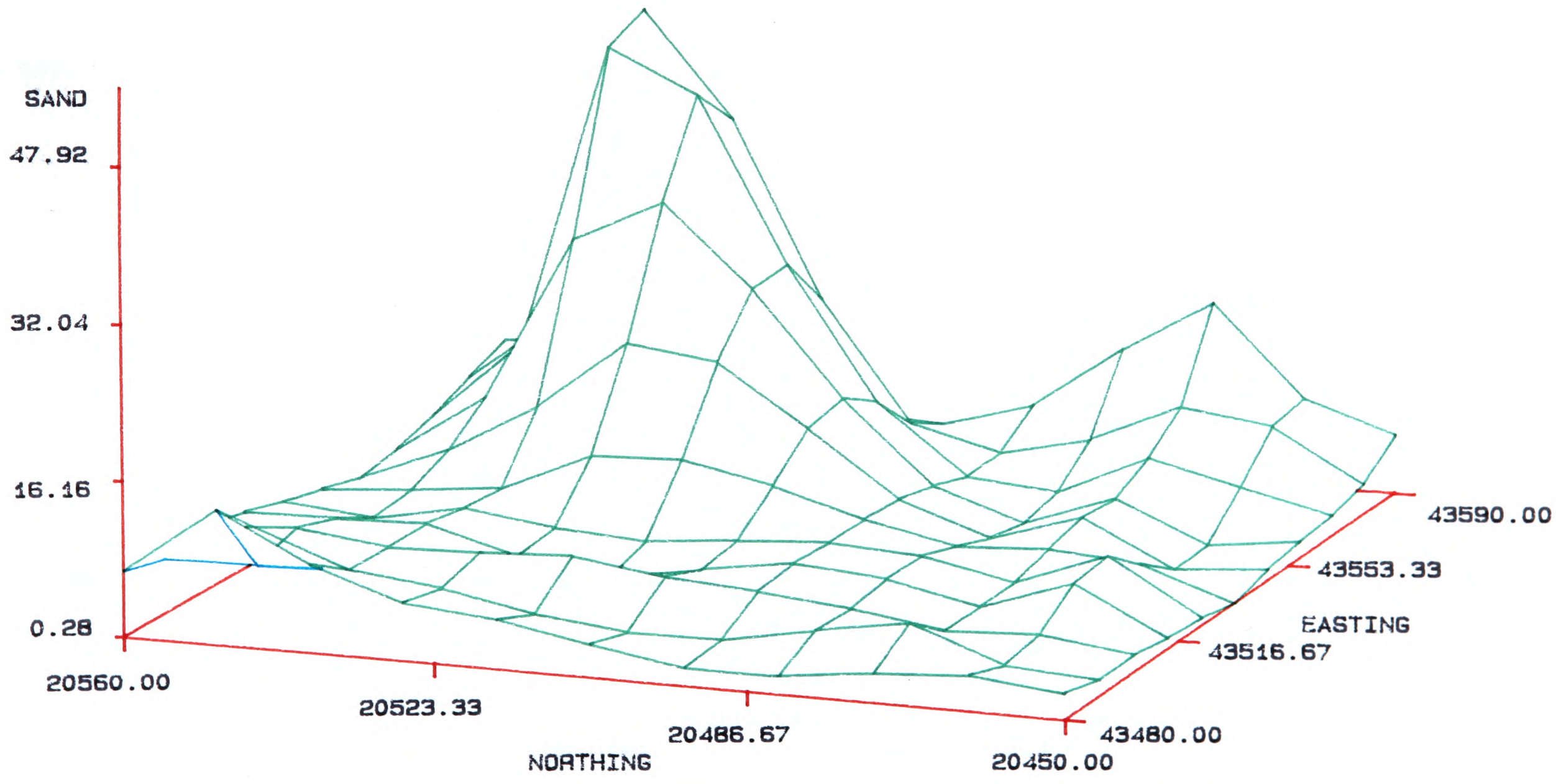
100 m GRID, YELFORD

Fig. 9.32b

COARSE SAND

BICUBIC SPLINE SURFACE

0–15cm.



100 m GRID, YELFORD

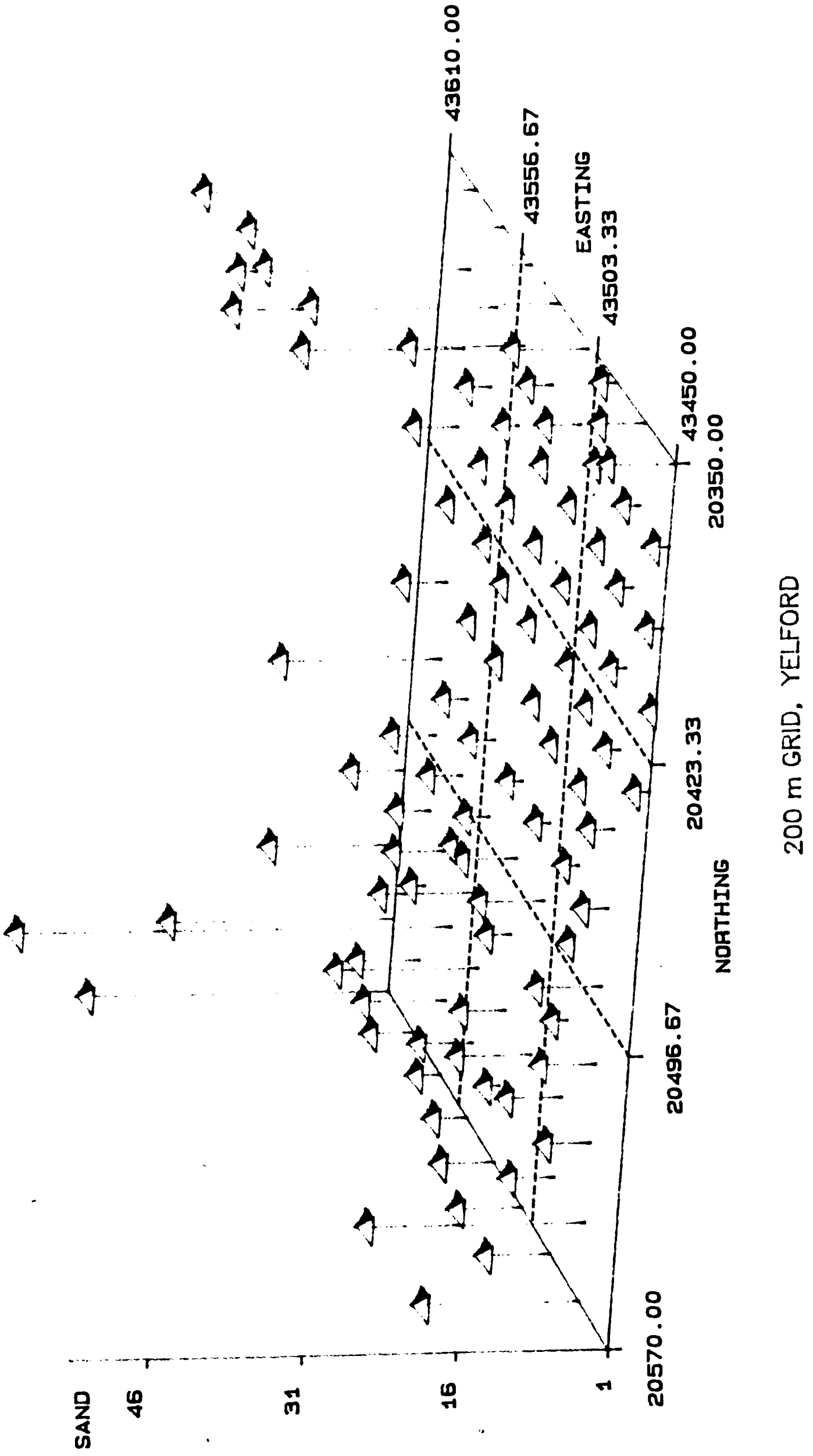


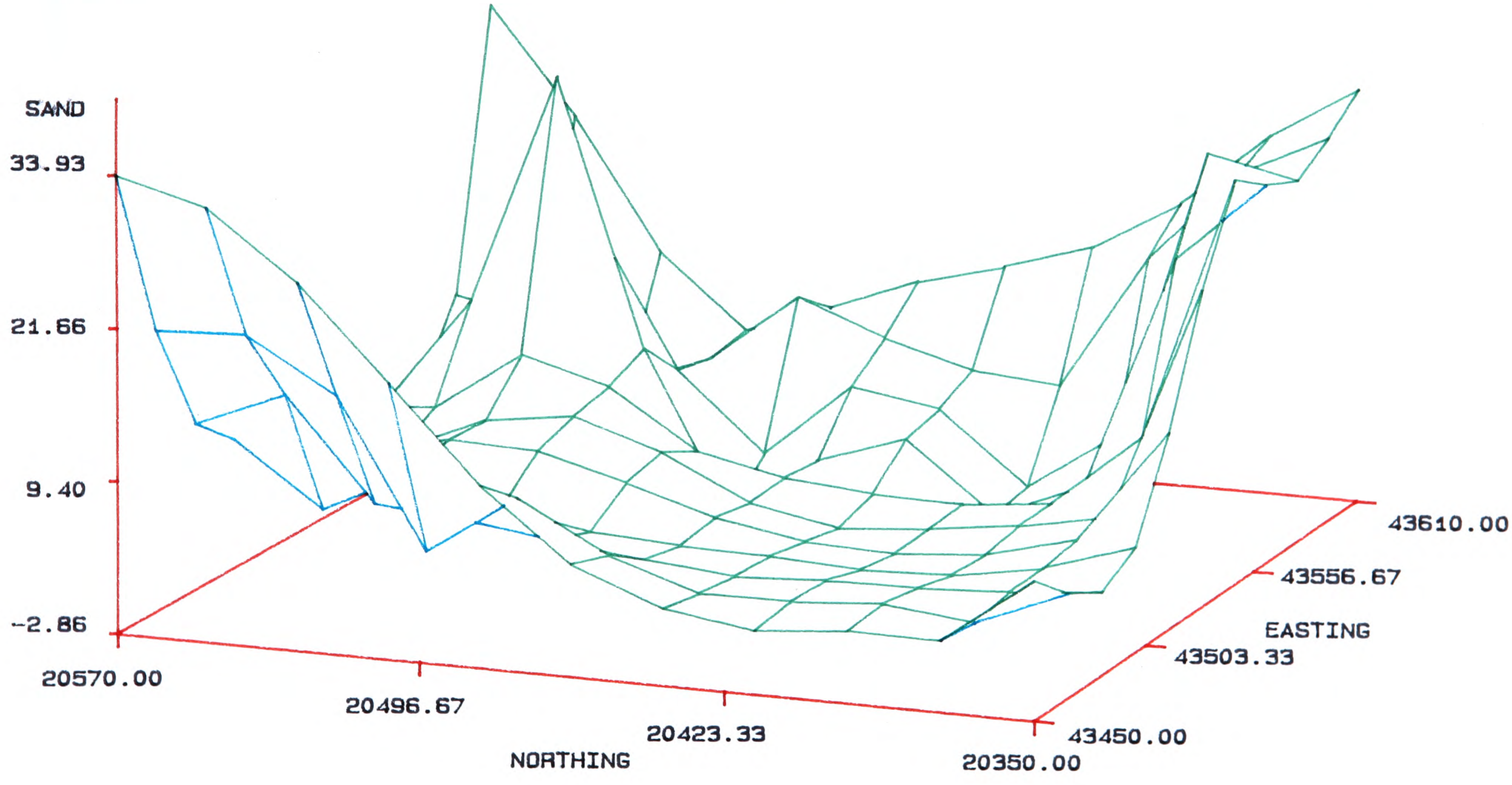
Fig. 9.33a **SAND** 0–15cm.

Fig. 9.33b

COARSE SAND

BICUBIC SPLINE SURFACE

0–15cm.



200 m GRID, YELFORD

in some of them are merely the product of the way the graphical representation of the surface takes place in the graphics package used to plot them, and not due to the computed spline surfaces themselves, which are, by definition, smooth.

In figure 9.22b, at the 50 m grid, the trend in the values of Potassium is clearly evident, but not strong. Its general slope runs in the direction of the drift (NE-SW). At the 100 and 200 m grids (figs. 9.23b and 9.24b) the trends of this property are not evident and variation at those scales appears to have a greater random component. It is for this reason that anisotropy only appears in the semi-variogram of Potassium at the 50m grid (fig. 9.6a) and the semi-variograms at the two other grids (figs. 9.6b and c) are isotropic.

There is a pattern in the variation of the other three properties which is reflected in the form of their trend surfaces and which is also related to the structures of their respective semi-variograms.

At the 50 m grid the trends of Carbonate, Silt and Coarse Sand are quite evident. Though, the trends shown by Carbonate and Coarse Sand are really strong, while that of Silt is weaker and with positive gradient, oposite to the gradient of the other two (figs. 9.25b, 9.28b and 9.31b). The direction of the general slope of these trends is the direction of the drift (NE-SW).

The location of the orign of these trends, the north-east corner of the surfaces in their plots, corresponds, in the landscape of the area, to the gravels and calcareous sands of the hill top, i.e. Richless hill, and from there the trend runs down the hill slope in the south-west direction (drift). The values of Carbonate and Coarse Sand, as shown by their surfaces, are the highest on that region and those of Silt, conversely, some of the lowest. So, the surfaces of Carbonate and Coarse Sand decrease steply in the direction of the drift, i.e. down hill in the landscape, as loamy and clayey soils begin to be more frequent, whilst that of Silt increases in the same direction. These features of the surfaces at this scale and their relation to the soil/landscape, provide a complete explanation for the anisotropies shown by the semi-variograms of

these properties at this sampling scale (figs. 9.8a, 9.11a and 9.10a).

At a smaller scale, i.e. 100 m grid, the surfaces of Carbonate, Silt and Coarse Sand (figs. 9.26b, 9.29b and 9.32b), all share a common feature. They show well localized areas with marked contrast between them in the values of the property (regions of high and low values) alternating over the whole surface but without a definite pattern or trend. The contrasts between rich and poor areas in the surfaces of Carbonate and Coarse Sand are more acute than those of Silt. These undulations represent the variation that occurs by localized regions in the area, within which changes in the value of the property are gradual and transitional, but between regions the changes are contrasting and with no pattern. Such a type of variation is indicative of a variable strongly regionalized. Variation of the property usually shows strong spatial dependence which is given by the transitions between tops and lows in the undulations, and their average length should aproximately correspond to the length of the range of influence of a semi-variogram.

Indeed, the semivariograms of these properties at the 100 m grid (figs 9.8b, 9.10b and 9.11b) showed strong spatial dependence with clearly defined structures (sill, range, Nugget), making sense of the variation depicted by the surfaces and the lack of evident trends at this scale.

Arguably, if the surfaces would have been given in advance, the semi-variograms probably could have been predicted.

At the 200 m grid Carbonate, Silt and Coarse Sand also have a common feature. Their surfaces (figs. 9.27b, 9.30b, and 9.33b) show smooth gradual changes, particularly in the central part of the surfaces. No strong trend is evident in any of them. These long gradual transitions are reflected in semi-variograms of the exponential type.

The semi-variograms af these properties at the 200 m grid (figs. 9.8c, 9.10c and 9.11c) were all, indeed, isotropic and exponential. This indicates again a good agreement between the information that can be extracted from the surfaces fit and the semi-variograms of the same area.

If the surfaces would have been given in advance, the long transitions in the center of the surface leading to stronger changes in the perifery, arguably, could have helped predicting an exponential semi-variogram. This type of variation can be seen as an extended case of the localized regions or undulations found in the surfaces at the 100 m grid, but with the regions within which the property bears some similarity and transitional change, far more extended and with less pronounced changes. There is a risk, however, that these long transitions could have been mistaken by weak trends. This raises the question again of how strong a trend has to be before it reflects itself as anisotropic variation in the semi-variogram, and before it can be recognized in advance from a plot or surface.

The trend surfaces were evaluated at the grid points in order to estimate the value of these points at the surface. Residuals of the observed values from the trend values were then calculated. Once the trend component has been removed the residuals would show random variation and possibly spatially dependent variation. The nature of the structures of such variation needed to be investigated by computing the semi-variogram of the residuals.

9.18.2 Stratifying trends by soil mapping units.

As an alternative approach to modelling, to account for the trends, a combination of spatial interpolation and intuition, of the form of that used to draw soil boundaries in conventional survey, was also explored.

The object of this approach was to find out if a simple and quick method based on intuition, i.e. a map, would help in solving the problem of accounting for trends and aid mathematical interpolation. Thus, soil mapping units were used in this context as a basis for the partitioning of the trends.

The soil mapping units defined during survey of the area by Marsman (table 9.2, and fig. 9.3) were employed for this purpose.

The aim was to stratify the trends and therefore remove the trend component by partitioning the area of each grid into segments which would both make sense to the soil/landscape

units, and provide sufficient point-data within the strata to allow for the computation of reliable within-stratum semi-variograms.

The use of individual mapping units as suitable strata for this purpose was immediately discarded since the within-unit point-data were not sufficient for reliable semi-variograms due to the relatively small areal extent of each individual unit.

Hence, soil mapping units had to be grouped on the basis of a pedogenetic criterion which could both be closely related to the variation of the properties involved, and which could find clear expression on the landscape units of the area.

Thus, the mapping units were grouped into calcareous soils and non-calcareous soils. These two strata formed the basis for the partition of the trends.

The list of the soil mapping units included in each group is given in table 9.29. It is worth mention that although the soils in the mapping units BEg (brown earths with gleying) and BEG (brown earths gley soils) are not calcareous in the strict sense, however, they were included within the calcareous soils for convenience and for pedogenetic and landscape association to calcareous soils.

For convenience in reference then, the strata were labelled so that calcareous soils were termed "BC-BE" and non-calcareous soils were termed "NS-NG".

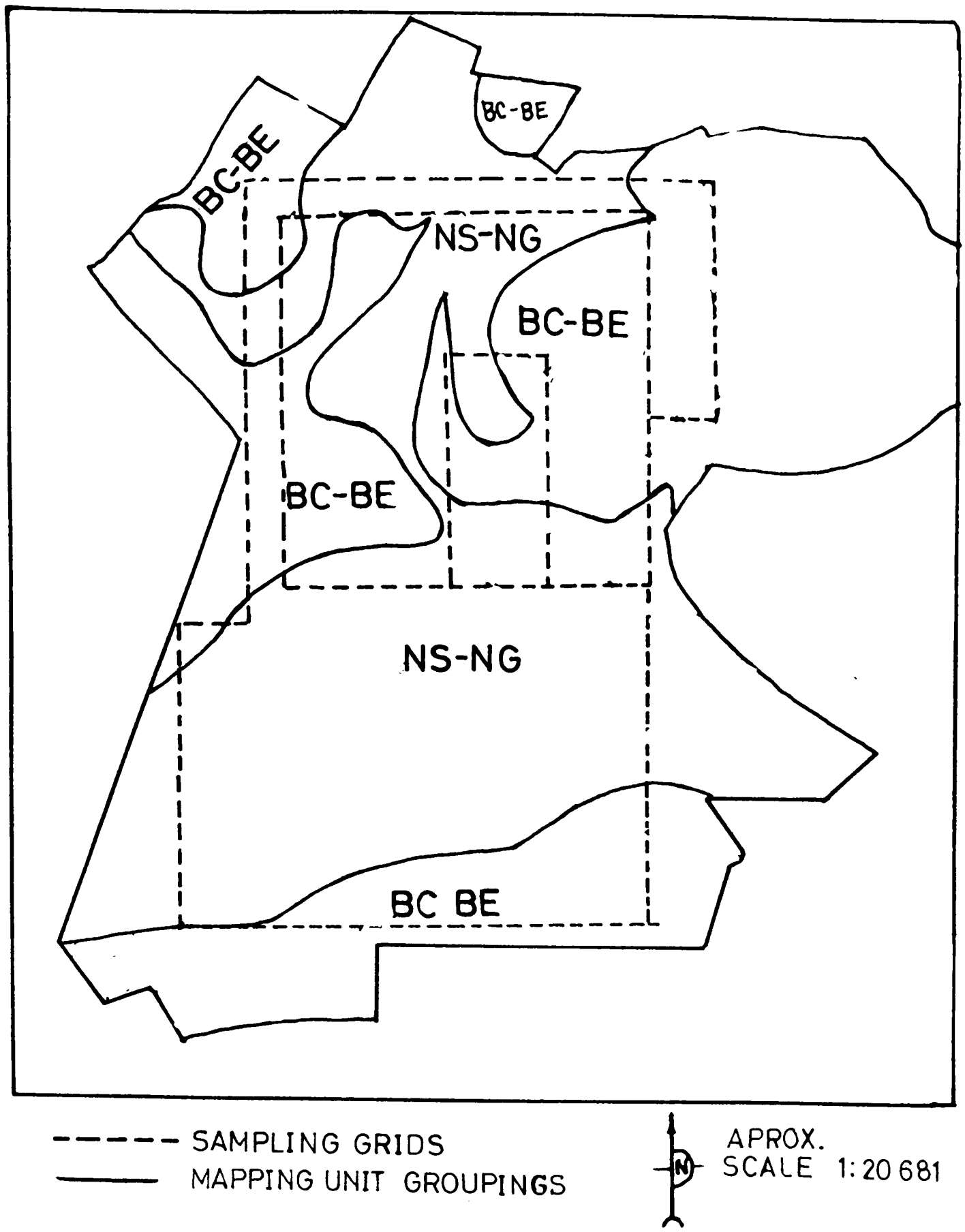
The delineations of the strata resulting from the groupings are shown on figure 9.34.

The list of sites within each of the mapping units listed in table 9.29 is included in appendix 9.7.

The data within the strata were manipulated in OXSIS to uniformize depth (0-15 cm). The two data-sets were ready then for the computation of within-stratum semi-variograms.

9.19 Semi-variograms of de-trended data.

Semi-variograms of de-trended data of Potassium, Carbonate, Silt and Coarse Sand were computed for the same 4 directions and exactly in the same way as for the semi-variograms of the original data (see section 9.10). Two groups of semi-variograms were obtained.



BC-BE Brown calcareous soils and brown earths

NS-NG Non-calcareous soils

Fig. 9.34 Grouping of soil mapping units into strata for partition of the trends in soil variation at Yelford.

9.19.1 Semi-variograms of residuals from trend.

The semi-variograms of residuals from trend are shown in figures 9.35a, 9.35b and 9.35c for Potassium at the 50, 100 and 200 m grids. In the same order, figs 9.36a, 9.36b, and 9.36c show those for Carbonate. Figures 9.37a, 9.37b and 9.37c show the semi-variograms of residuals from trend of Silt, and figures 9.38a, 9.38b, and 9.38c display similar semi-variogram plots for Coarse Sand.

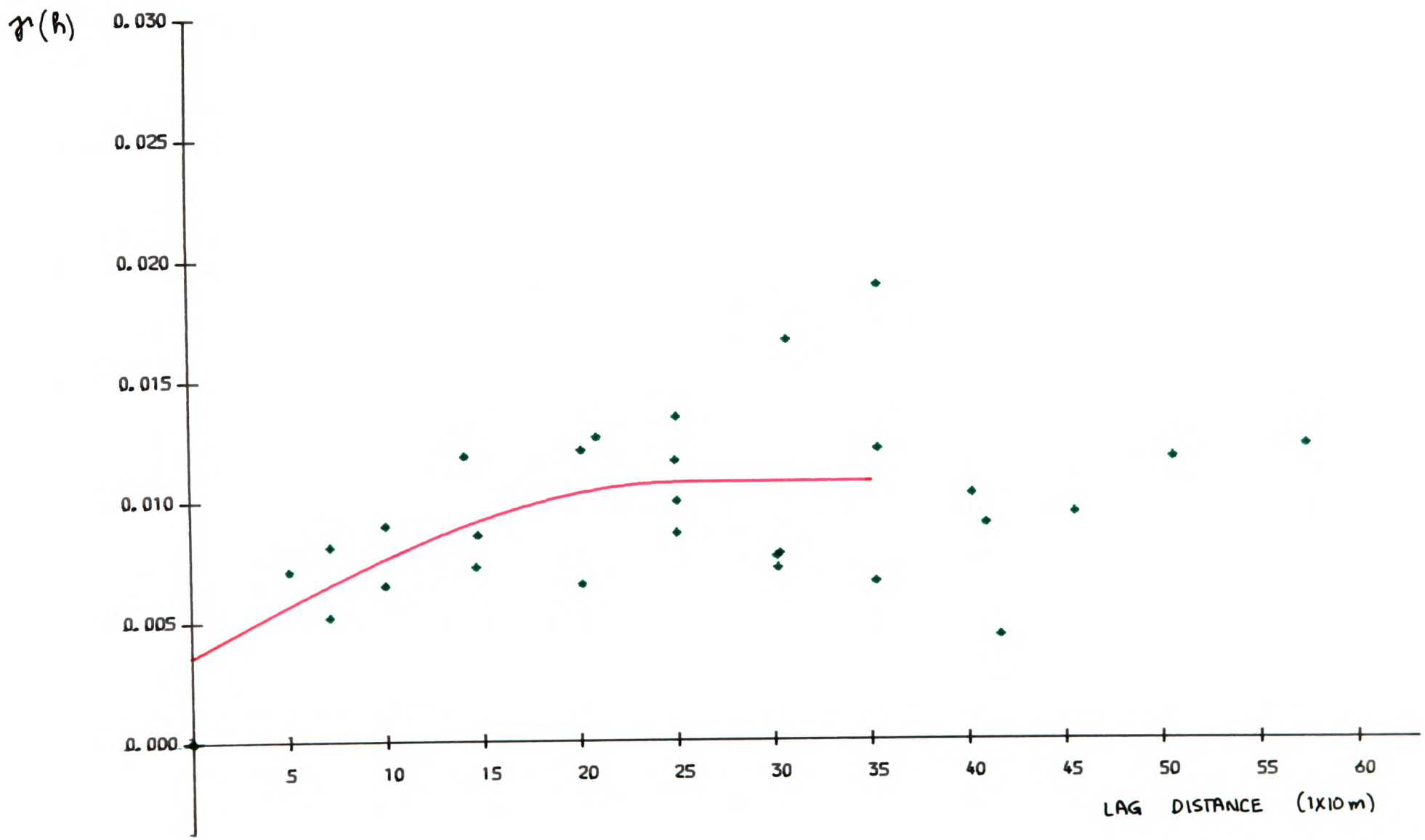
Table 9.29. Grouping of soil mapping units into strata for partitioning and removing trends in soil variation at Yelford.

	Stratum BC-BE: Calcareous Soils	Stratum NS-NG: Non-calcareous soils
50 m grid		
	BCT 4-5	NSG-2
	BCg 3(2)	NGG1-2
	BCg 3/4	NGG2-2
	BEg 3-2	
	BEg-3(2)	
100 m grid		
	BCT-4/5	NSG-2
	BEg-3(2)	NGG1-2
	BCg-3(2)	NGG2-2
	BCg-3/4	
	CSG-2	
	CGG-2	
200 m grid		
	BCT 4-5	NGG2-2
	BCg-3	NSG-2
	BCg-3(2)	NGG1-2
	BCg 3-4	
	BEg-3(2)	
	BEg-2(3)	

It can be seen from the plots of semi-variograms in the figures listed above, that only in 50% of them (6 out of 12) the residuals are shown to be spatially auto-correlated. The other 50% of the semi-variograms showed total random variation in the residuals from trend.

In those cases where the semi-variogram of residuals showed spatial dependence and structure, models for the experimental

Fig. 9.35a SEMI-VARIOGRAM of Residuals from TREND, POTASSIUM, 50m grid



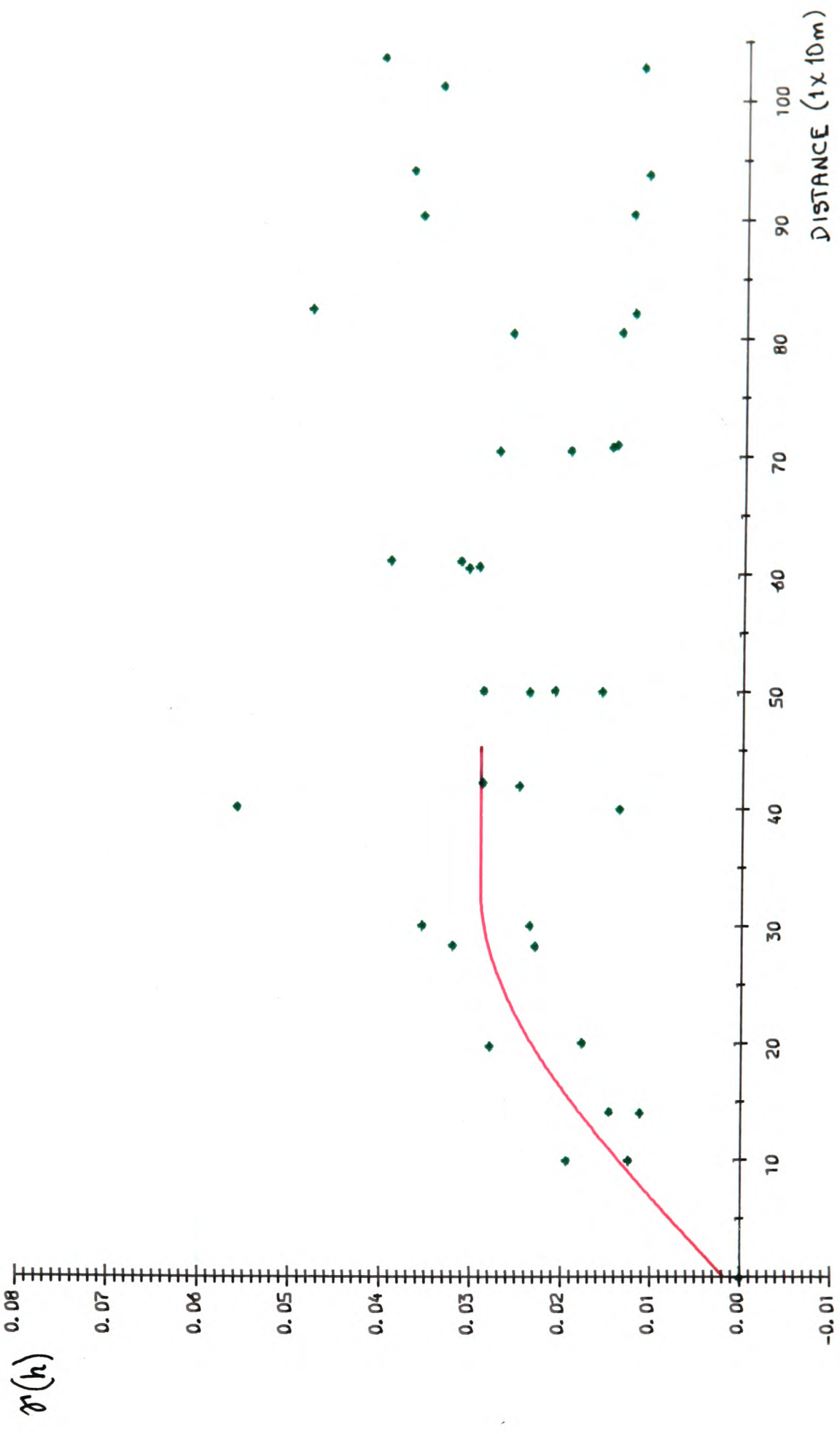


Fig 9.35 b. SEMI-VARIOGRAM of RESIDUALS from TREND, POTASSIUM, 100m grd.

Fig. 9.35c. SEMI-VARIOGRAM of RESIDUALS from TREND, POTASSIUM, 200m grid

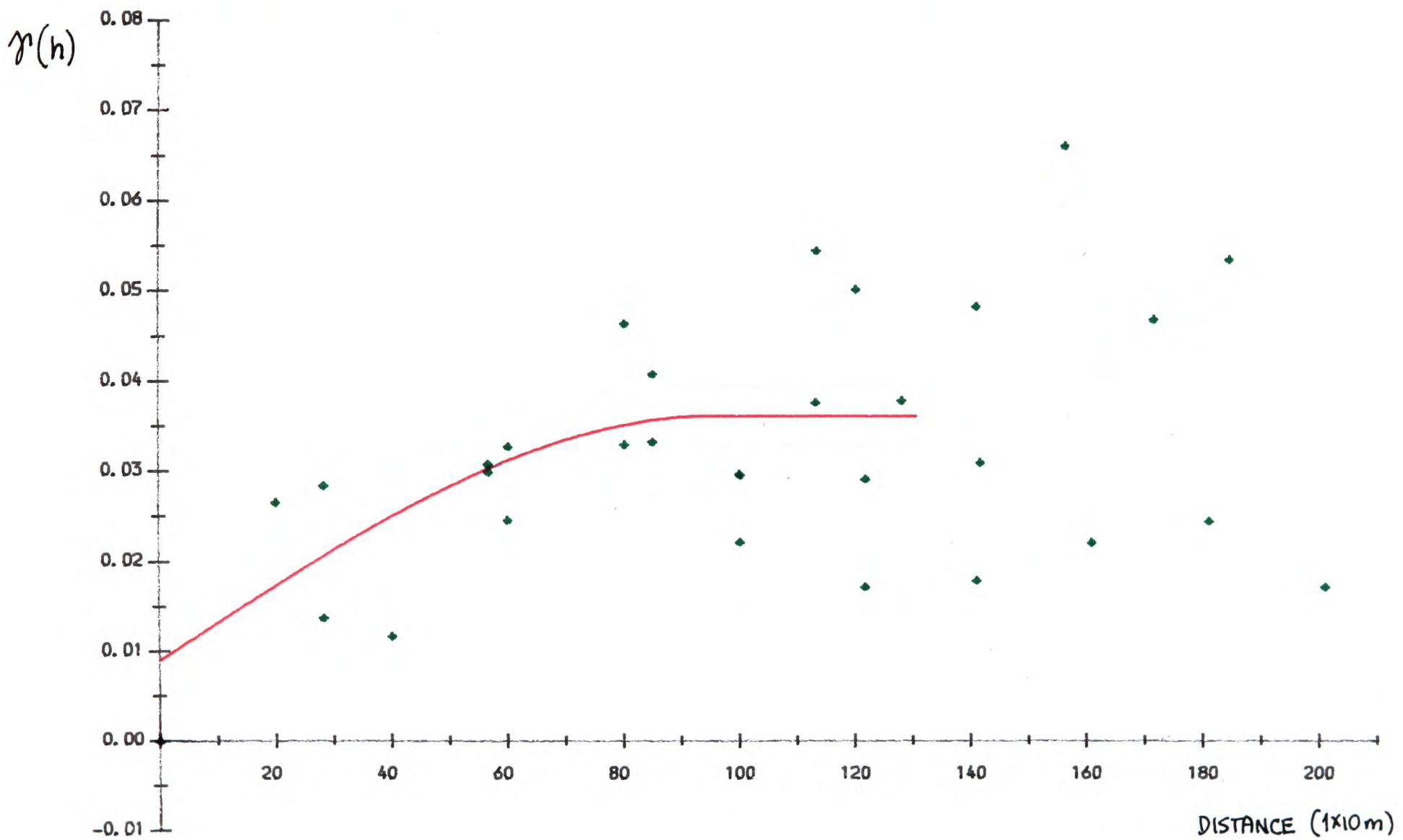


Fig 9.36 a. SEMI-VARIOGRAM of RESIDUALS from TREND, CARBONATE, 50m grid

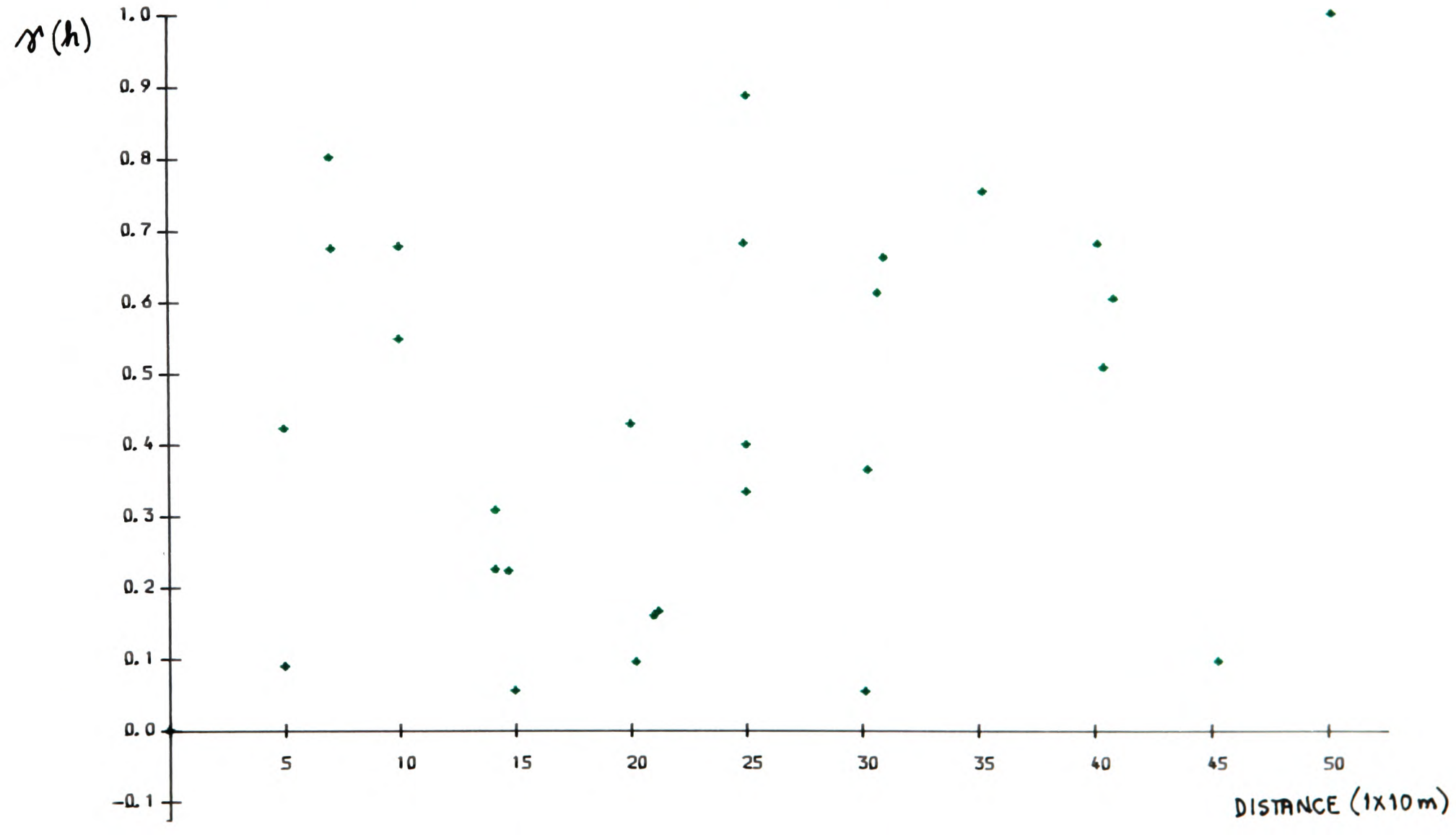
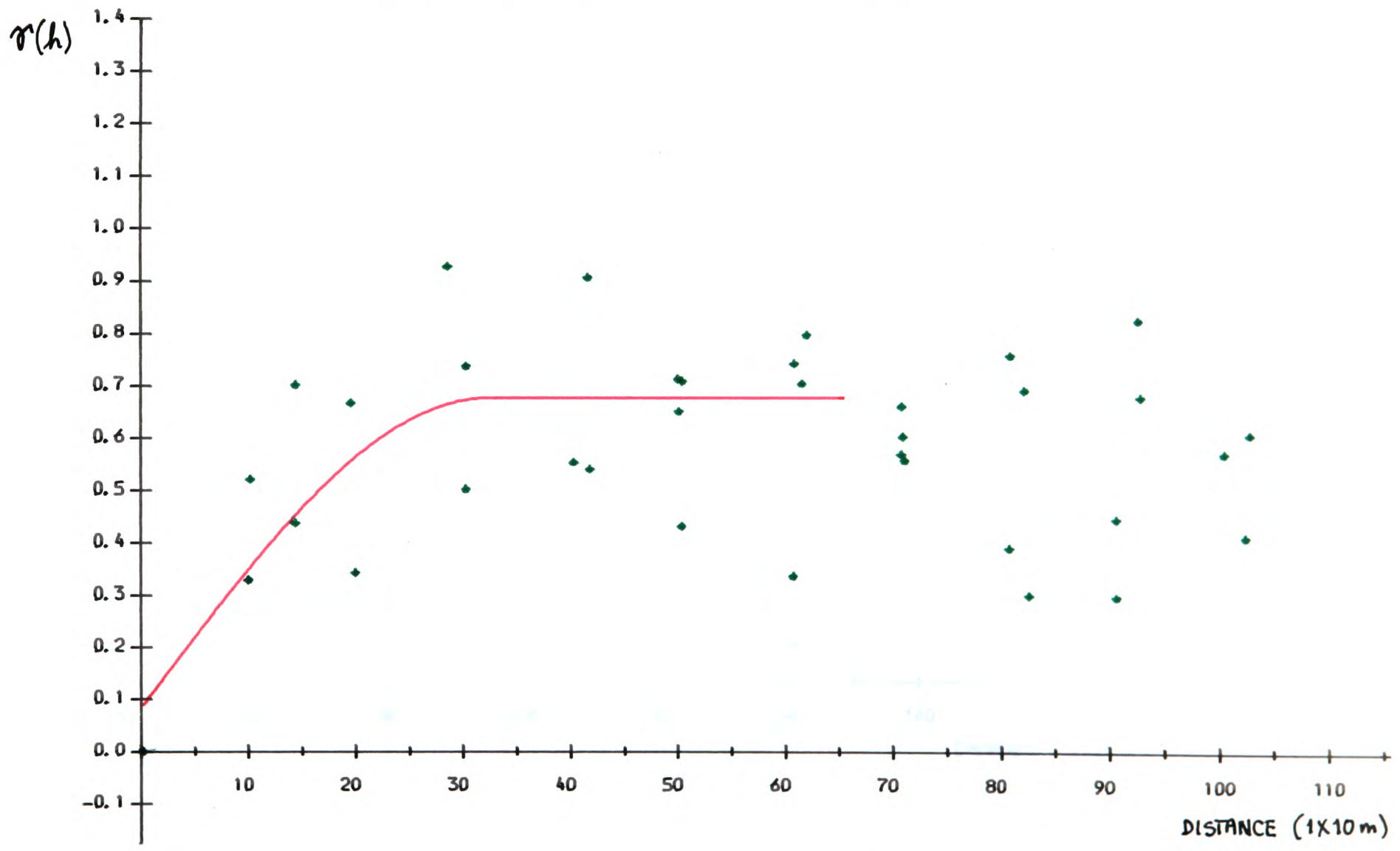


Fig. 9.36.b. SEMI-VARIOGRAM of RESIDUALS from TREND, CARBONATE, 100m grid



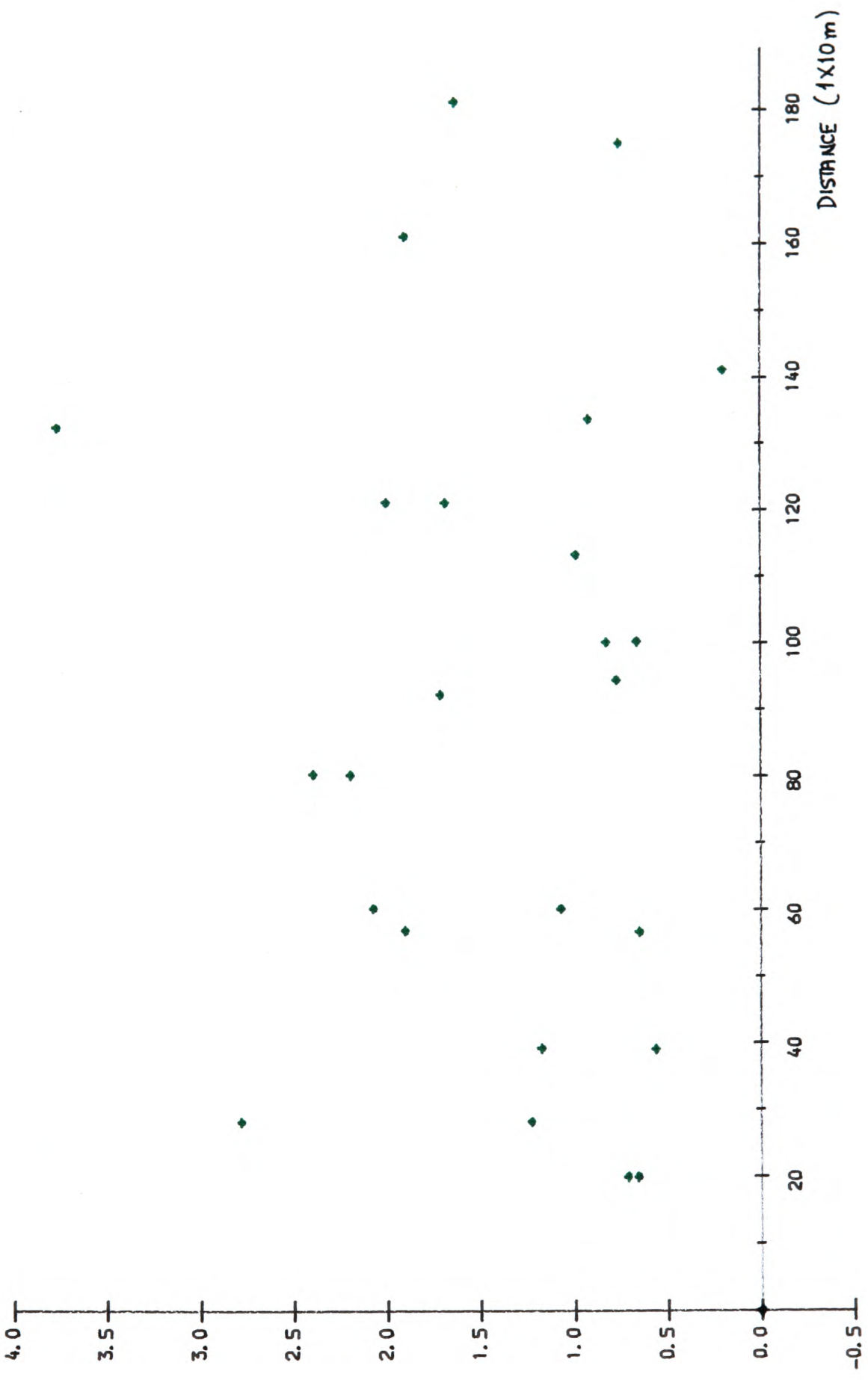

SEMI-VARIOGRAM of RESIDUALS from TREND, CARBONATE, 200m grid
γ(h)
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
20
40
60
80
100
120
140
160
180
DISTANCE (1X10m)

Fig. 9.36c

Fig. 9.37 b. SEMI-VARIOGRAM of RESIDUALS from TREND, SILT, 100m grid

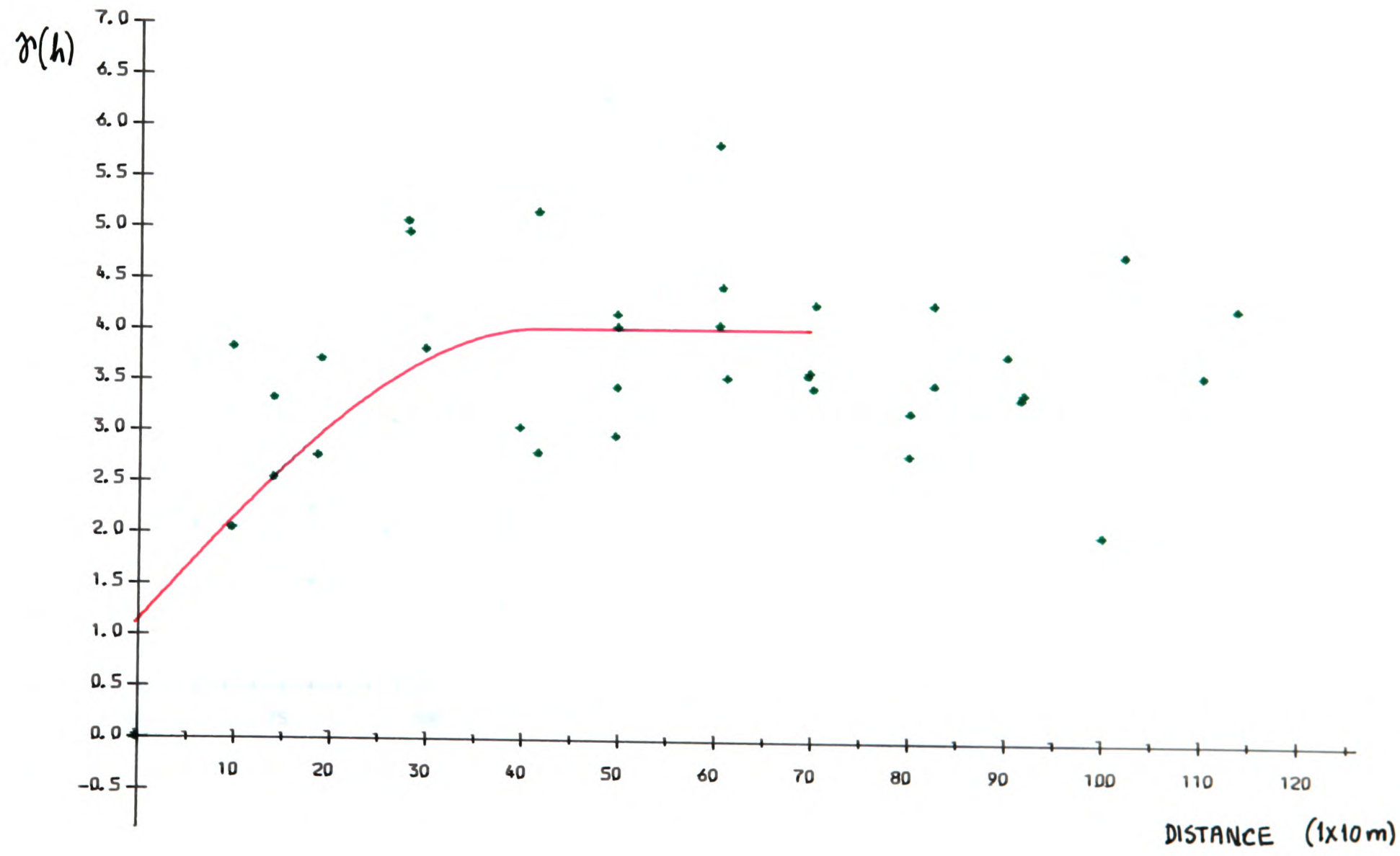


Fig. 9.37 b. SEMI-VARIOGRAM of RESIDUALS from TREND, SILT, 100m grid

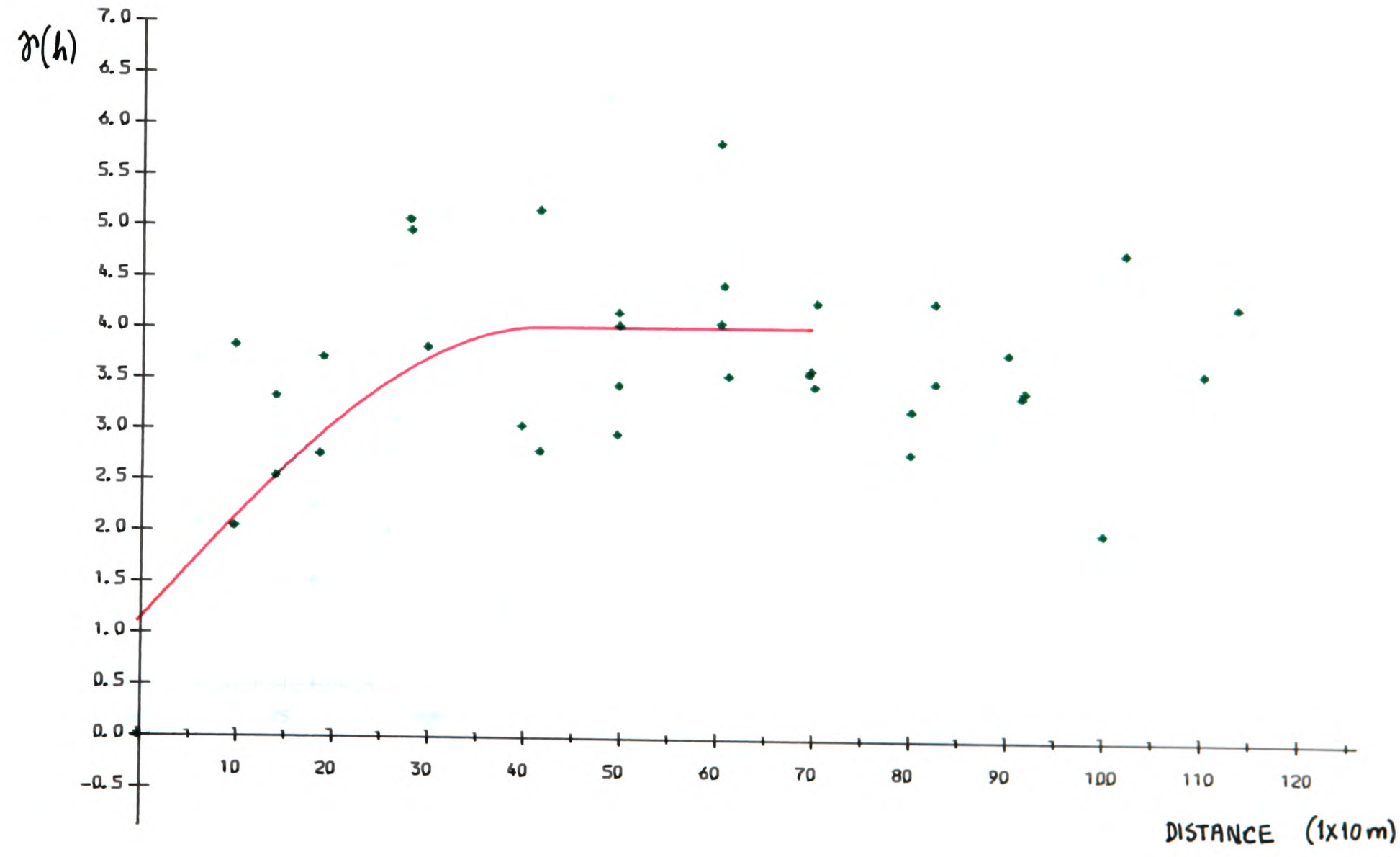


Fig 9.38 a. SEMI-VARIOGRAM of RESIDUALS from TREND, SAND, 50m grid

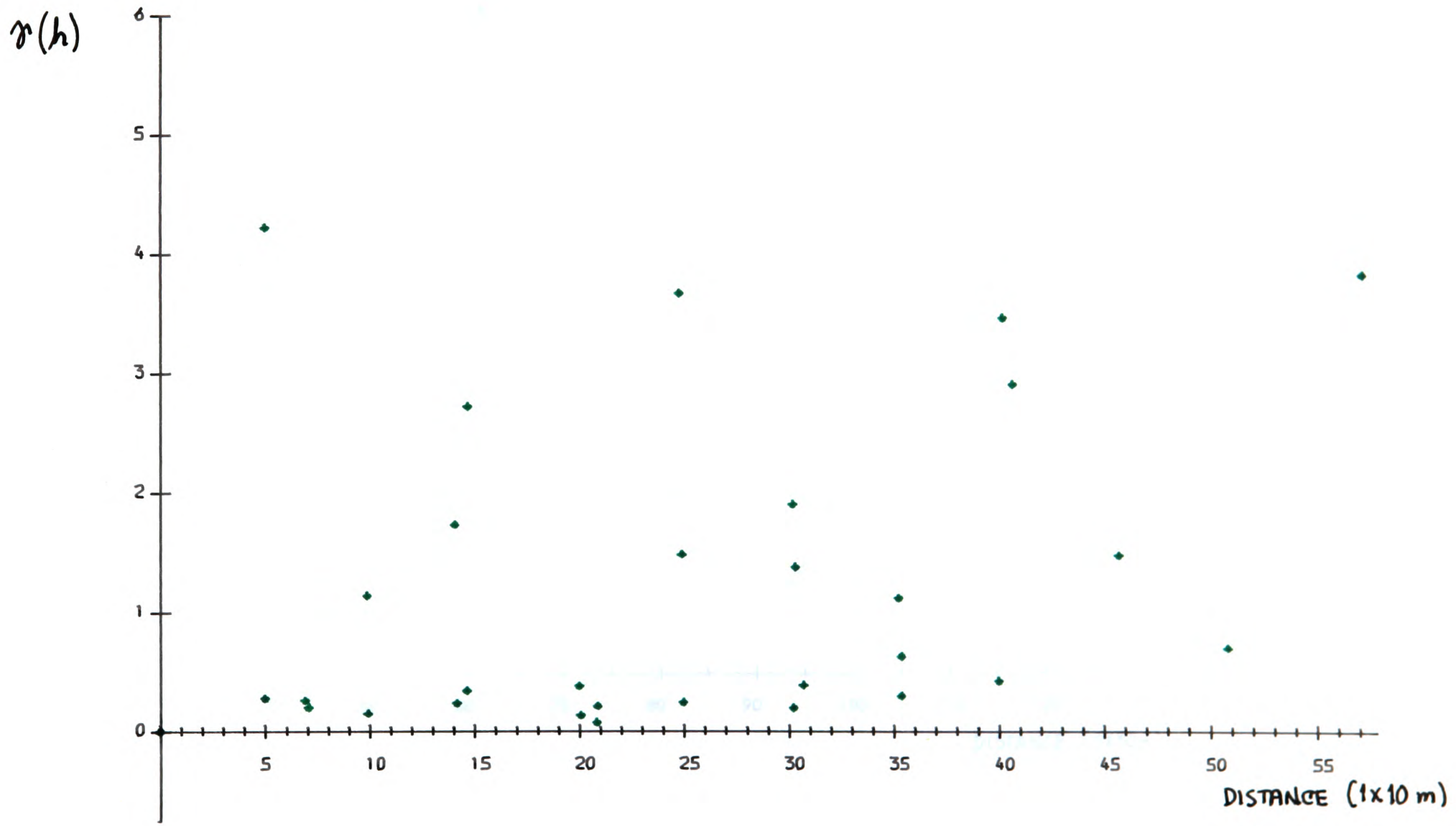


Fig 9.38 a. SEMI-VARIOGRAM of RESIDUALS from TREND, SAND, 50m grid

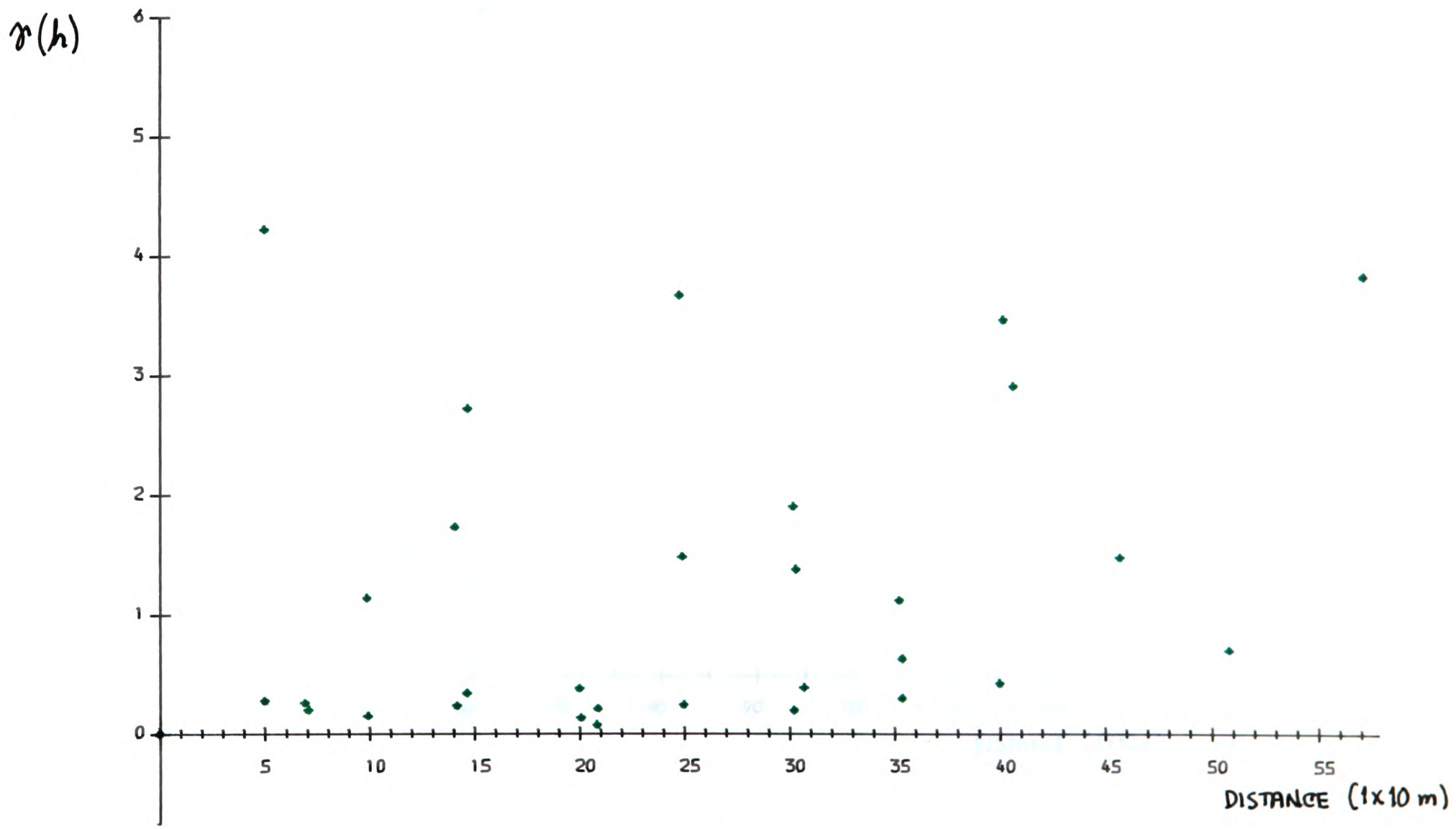


Fig 9.38b. SEMI-VARIOGRAM of RESIDUALS from TREND, SAND, 100m grid

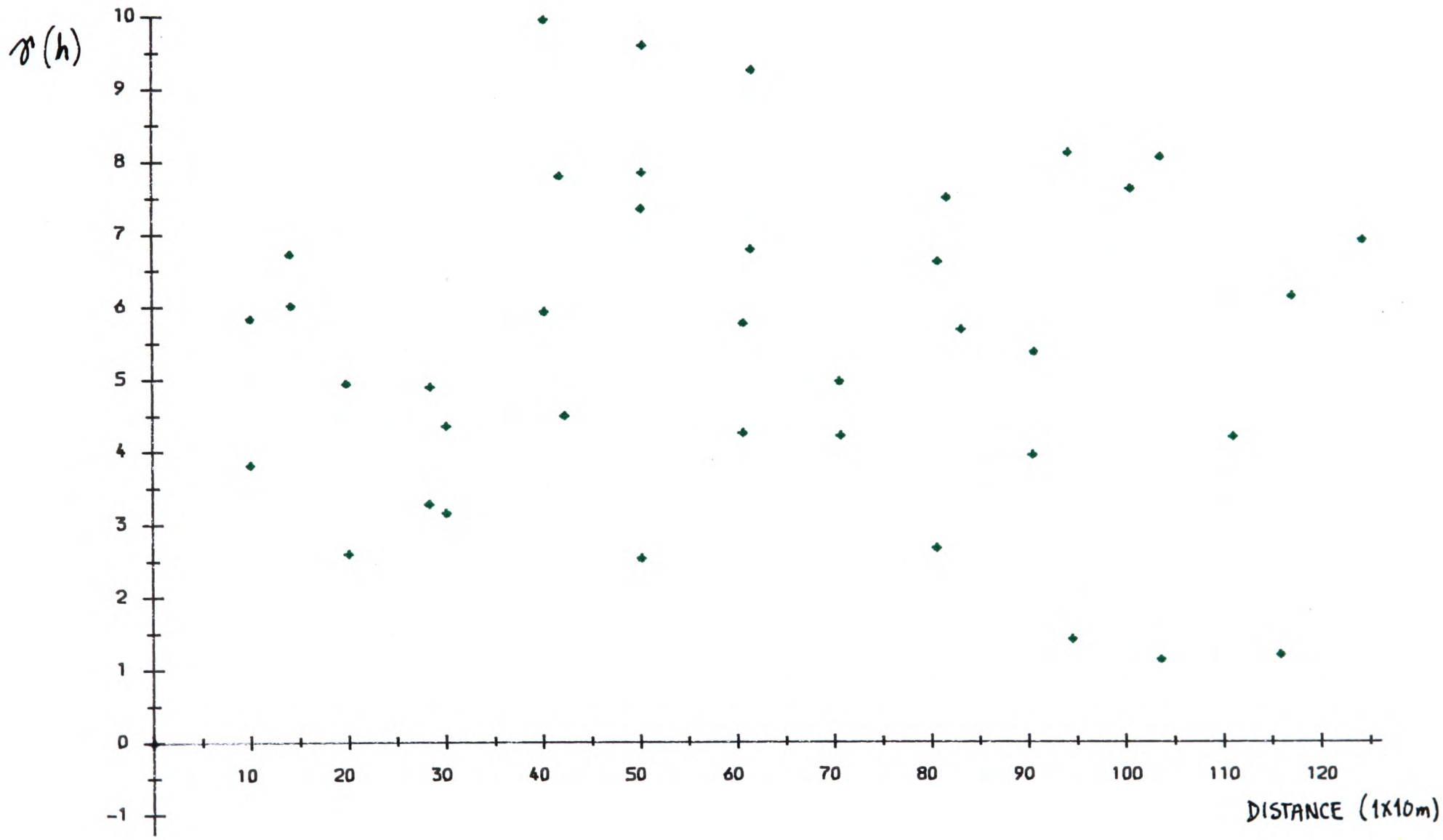
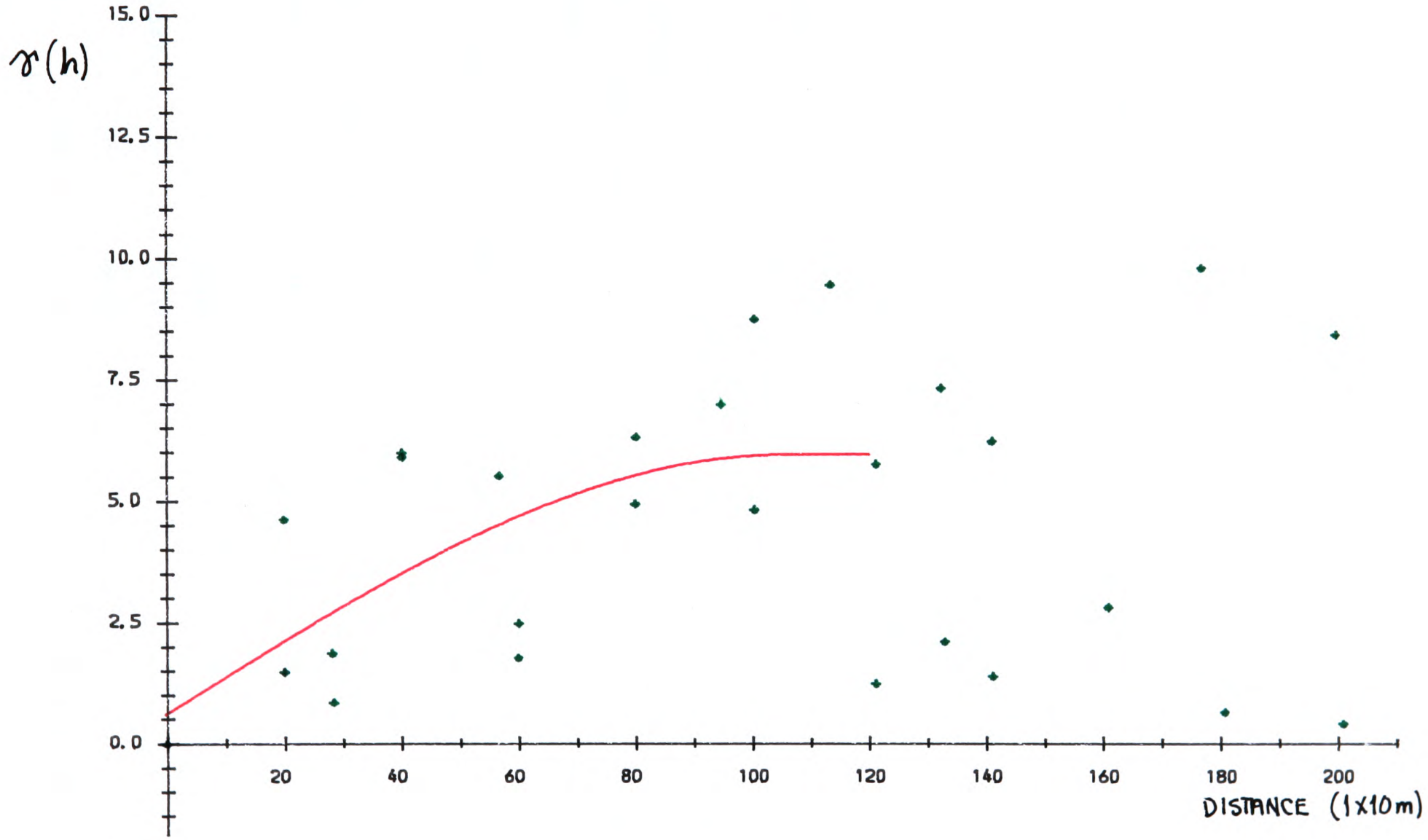


Fig. 9.38c. S EMI-VARIOGRAM of RESIDUALS from TREND, SAND, 200m grid



semi-variograms were fitted following similar procedures as in section 9.11 above. Table 9.30 gives information about the type of model found in each case, and the parameters of the model estimated during the fitting.

Only in the case of Potassium semi-variograms the residuals from the trend at the three sampling grids were spatially auto-correlated. The three isotropic semi-variograms were fitted by spherical models (figs. 9.35a-c, and table 9.30). However, spatial dependence was not strong at the 50 m grid, and at the 200 m grid the fit of the model was poor and remains conjectural.

A sampling spacing of 100 m seems to be the most appropriate to pick up spatial structure in the residuals of Carbonate, since only at this scale the semi-variogram showed spatial dependence, and a spherical model was fitted. At the 50m grid the spacing seems too short to allow for spatial dependence. Most of the variation is picked up by the surface fitted to the trend, leaving only random variation in the residuals.

The situatuion of the residuals of Silt from trend is identical to that of Carbonate, though the points in the plot of the semi-variogram at the 100m grid are more disperse round the curve of the spherical model fitted. A spacing of 100 m seems suitable to remove trends of this property by this method.

Table 9.30. Semi-variograms of residuals from trend of four soil properties, models fitted and their parameters, at Yelford.

Property	Grid (m)	Model	Nugget	Sill	Range (m)
Potassium	50	spherical	0.0035	0.0107	250
"	100	spherical	0.004	0.028	350
"	200	spherical	0.009	0.036	950
Carbonate	100	spherical	0.092	0.68	320
Silt	100	spherical	1.10	4.0	420
Sand	200	spherical	0.65	6.0	105

All the semi-variograms of residuals from trend of Coarse Sand, but that at the 200 m grid, showed total absence of spatial dependence. However, the fit of the spherical model at

the 200 m grid is not quite satisfactory. Variation of the residuals at the 50 and 100 m grids showed no spatial structure, which may have been accounted for also by the surfaces that model the trends.

9.19.2 Semi-variograms within mapping unit groupings.

Semi-variograms within the strata formed by grouping map units of calcareous and non-calcareous soils were calculated, plotted and a model was fitted to them.

The data in each of the sampling grids was therefore split into two sub-sets: the BC-BE soils (calcareous) and the NS-NG soils (non-calcareous), giving two semi-variograms at each sampling grid for each property of the four studied.

The BC-BE semi-variograms for Potassium at the three sampling grids are shown in figures 9.39a,b and c, whereas the NS-NG semi-variograms for the same property are shown in figures 9.40a,b and c.

In similar order figs. 9.41a-c, and 9.42a-c display the semi-variograms of Carbonate within both strata. The BC-BE semi-variograms of Silt are shown in figures 9.43a-c, and the NS-NG of the same property in figures 9.44a-c, while figures 9.45a-c and 9.46a-c show the BC-BE and NS-NG semi-variograms of Coarse Sand.

On the whole, in more than a half of these semi-variograms (about 60%), the properties examined appeared to be spatially auto-correlated. The rest are semi-variograms showing total random variation.

The division of the landscape on the basis of mapping unit groupings, in terms of calcareous and non-calcareous soils, proved useful in stratifying trends, particularly at the 50m grid, where the anisotropic variation of all the properties studied here, ocurred. So, from four anisotropic semi-variograms which could not be used directly for Kriging interpolation, by partitioning the trend into two strata, five isotropic semi-variograms out of eight were obtained. The remaining three showed total absence of spatial dependence in the form of Pure Nugget effect. Predictions on these cases were based on the mean values of the stratum. The availability of

Fig. 9.39a SEMI-VARIOGRAM within MAP UNITS BC-BE, POTASSIUM, 50m grid

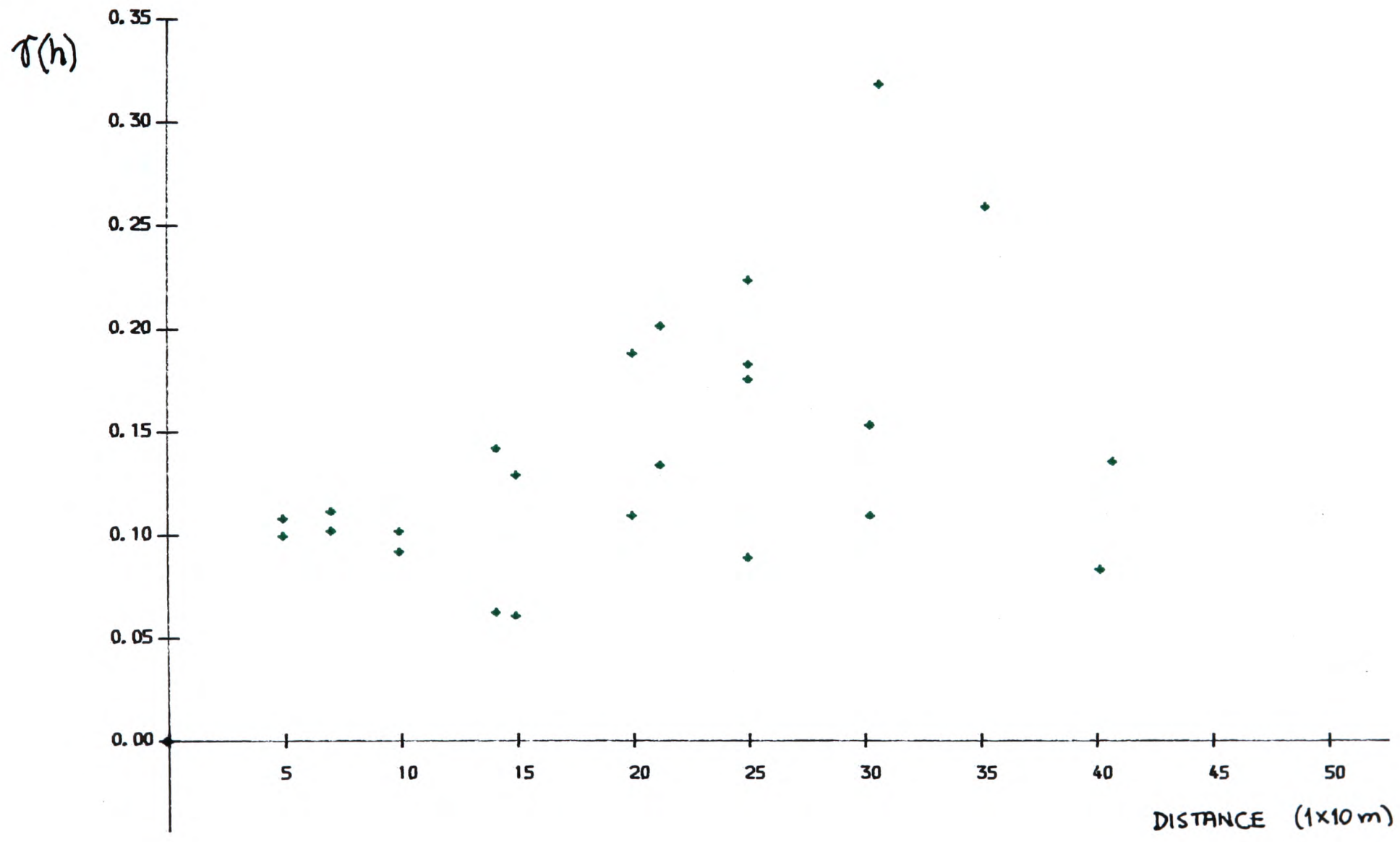


Fig 9.39b SEMI-VARIOGRAM within MAP UNITS BC-BE, POTASSIUM, 100m grid

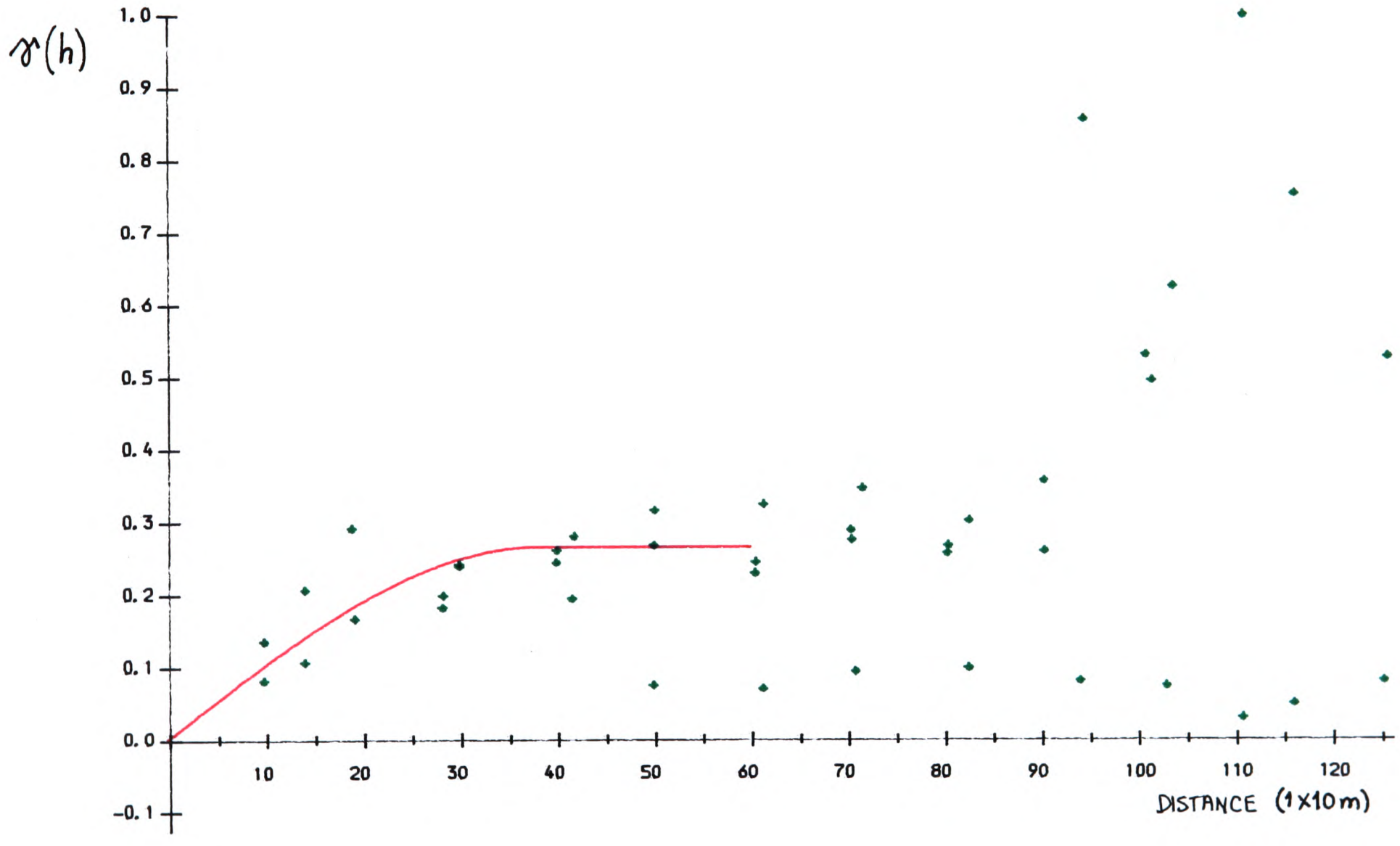


Fig. 9.39c. SEMI-VARIOGRAM within MAP UNITS, BC-BE, POTASSIUM, 200m grid

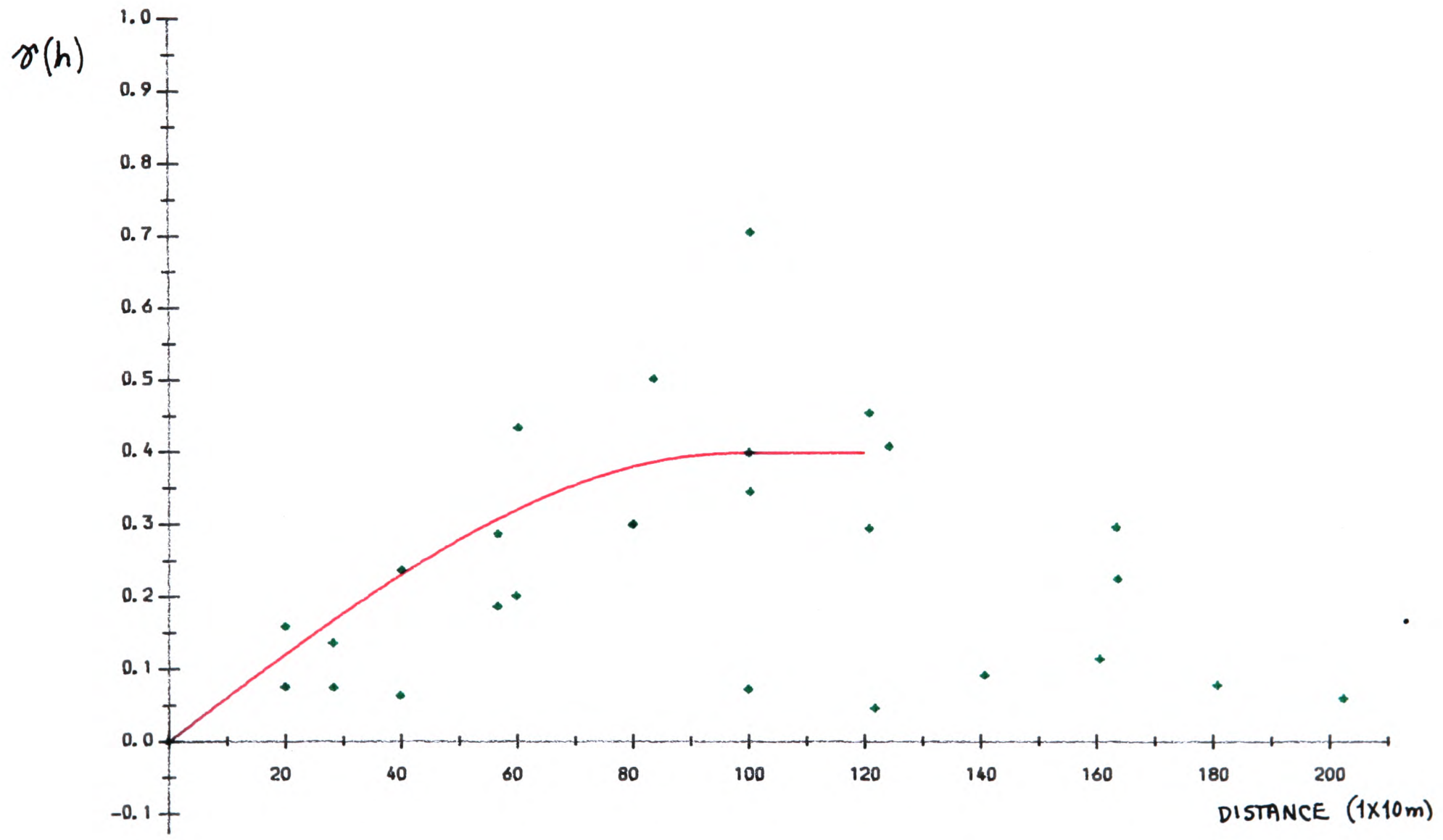


Fig 9.40a. SEMI-VARIOGRAM of DATA in MAP UNITS NS-NG, POTASSIUUM, 50m g

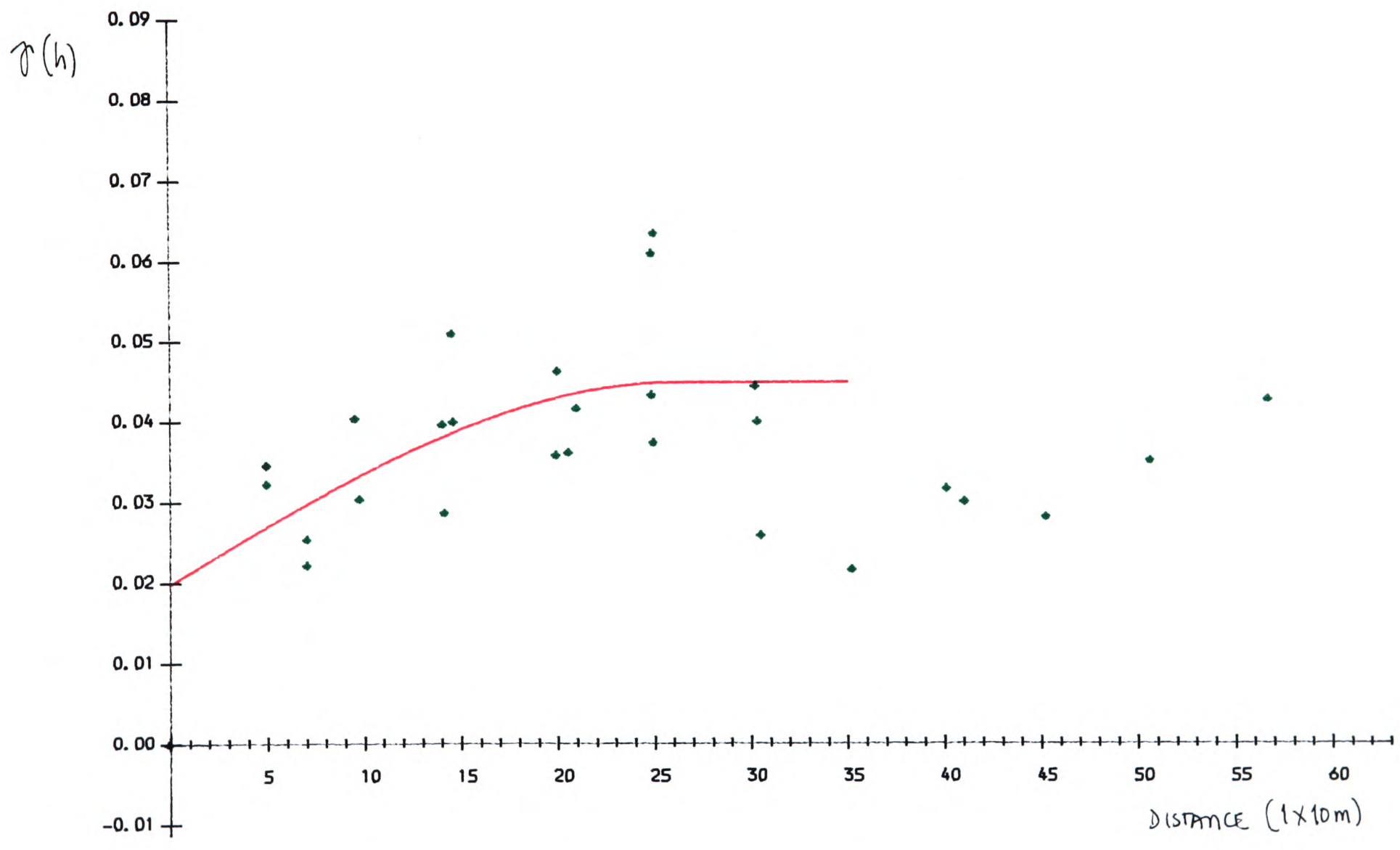


Fig. 9.40b. SEMI-VARIOGRAM within MAP UNITS NS-NG, POTASSIUM, 100m grid

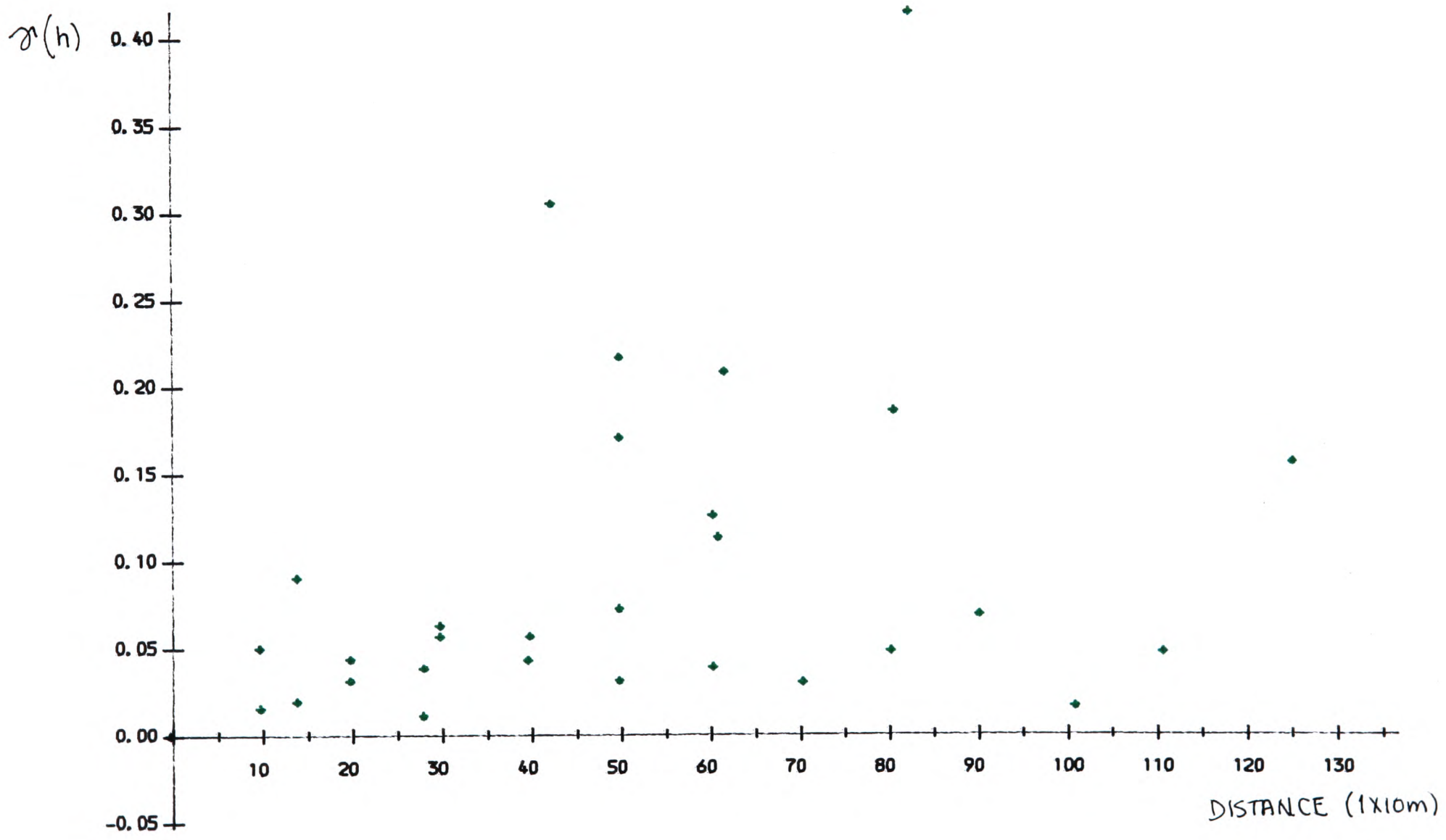


Fig 9.40c SEMI-VARIOGRAM within MAP UNITS, POTASSIUM, 200m grid

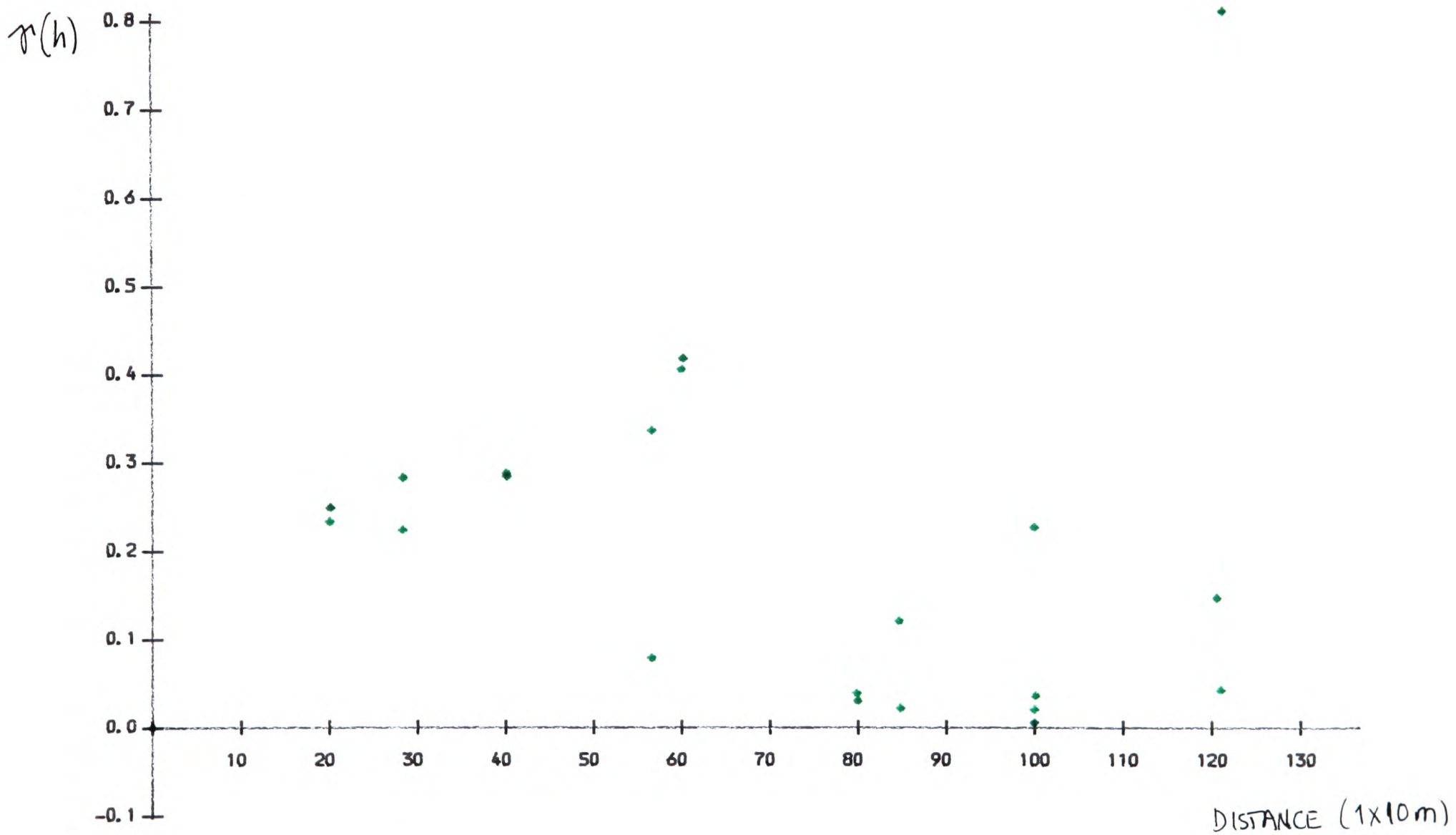


Fig. 9.41 a SEMI-VARIOGRAM of DATA IN MAP UNITS BC-BE, CARBONATE, 50m gr

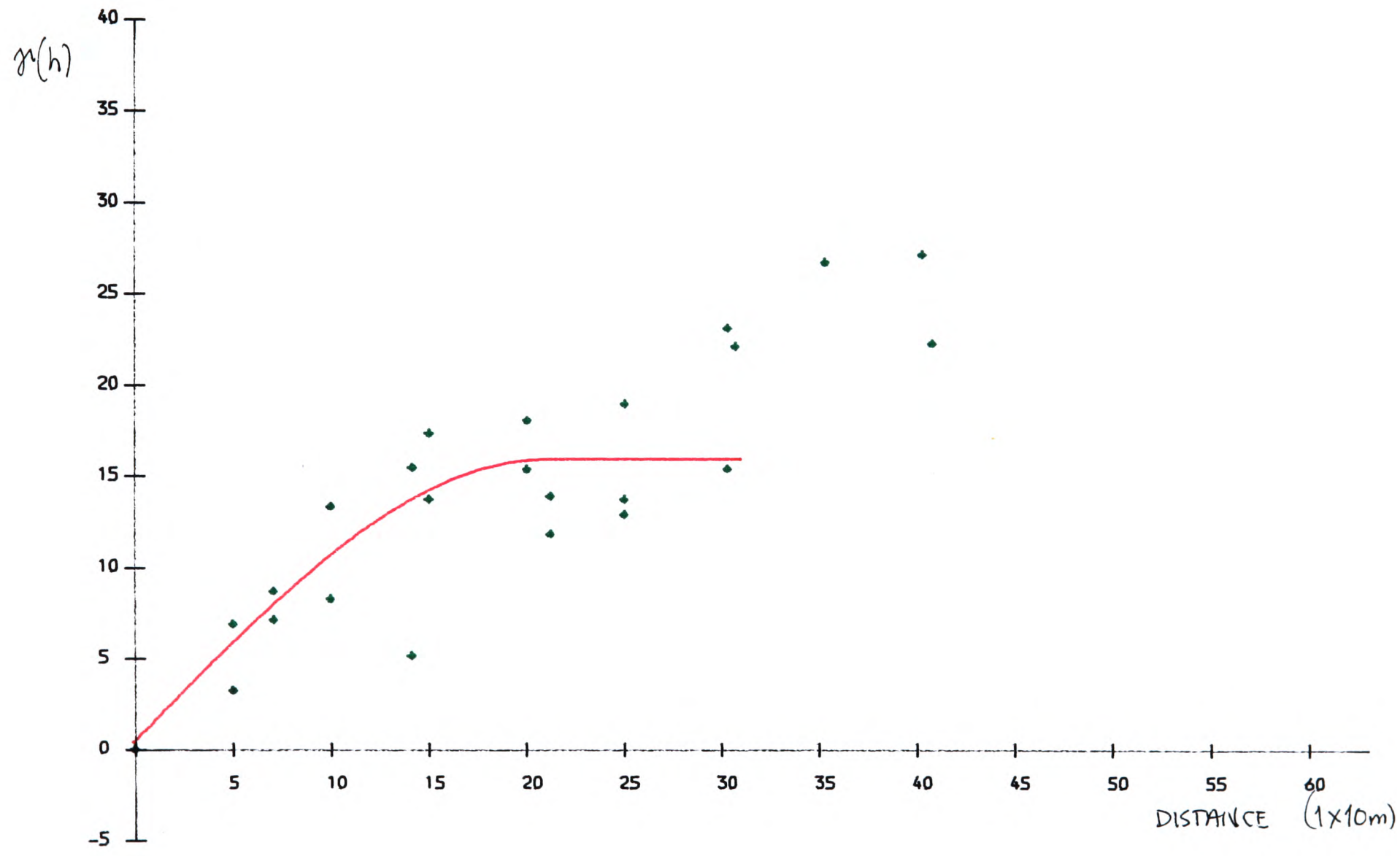


Fig 9.41 b SEMI-VARIOGRAM within MAP UNITS BC–BE, CARBONATE, 100m grid

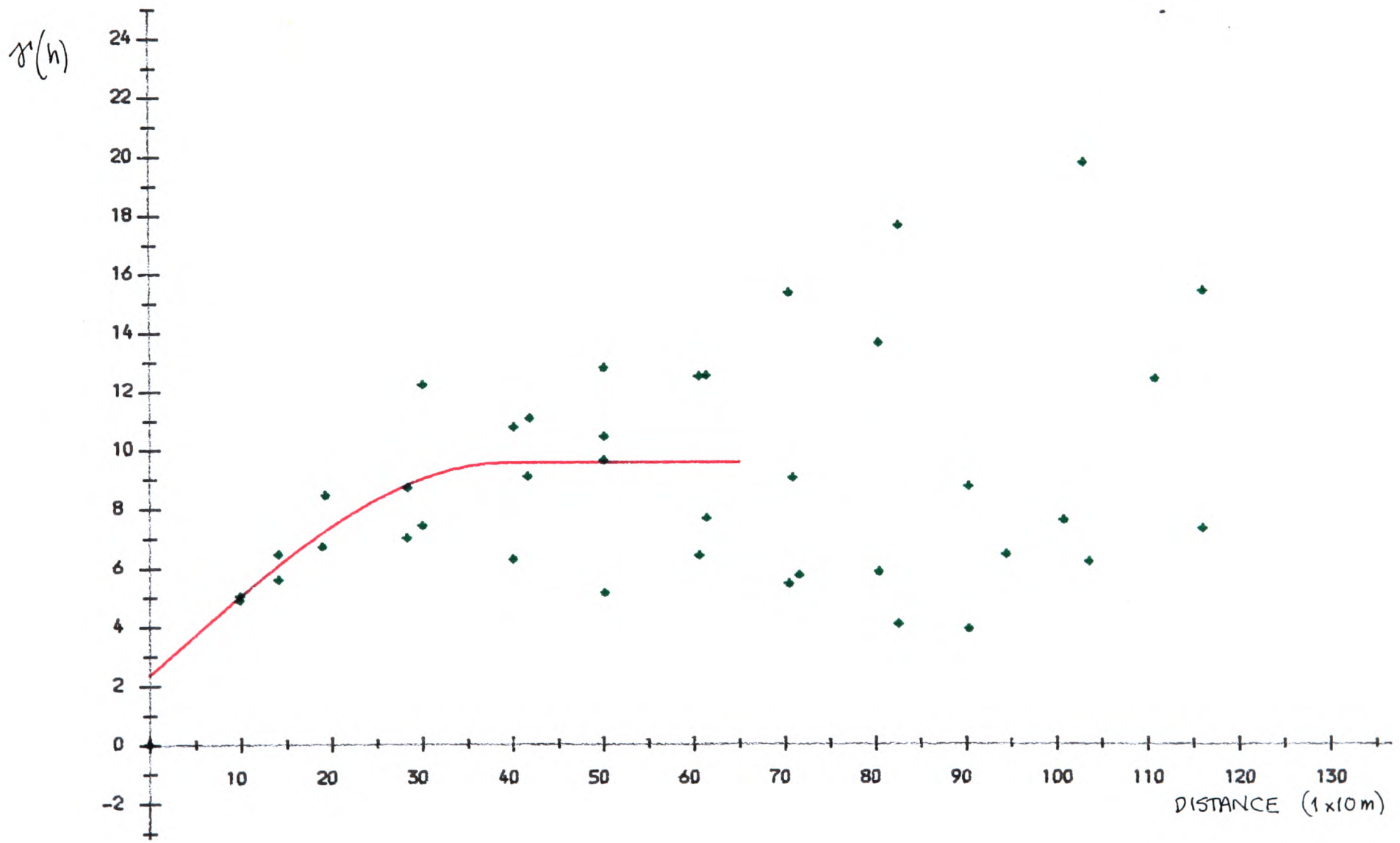


Fig 9.41 c SEMI-VARIOGRAM within MAP UNITS BC-BE, CARBONATE, 200m grid

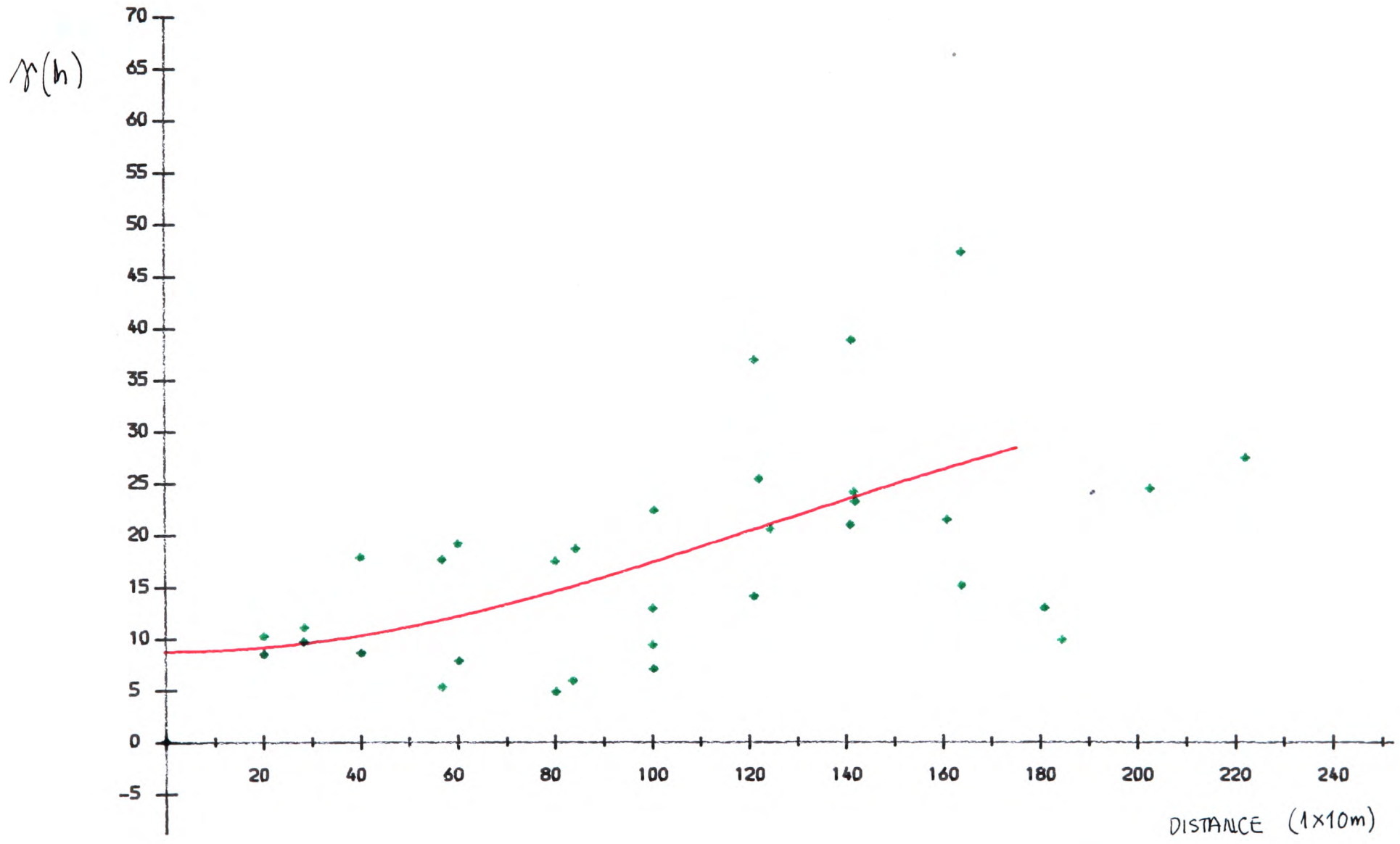


Fig 9.42 a SEMI-VARIOGRAM within MAP UNITS NS-NG, CARBONATE, 50m grid

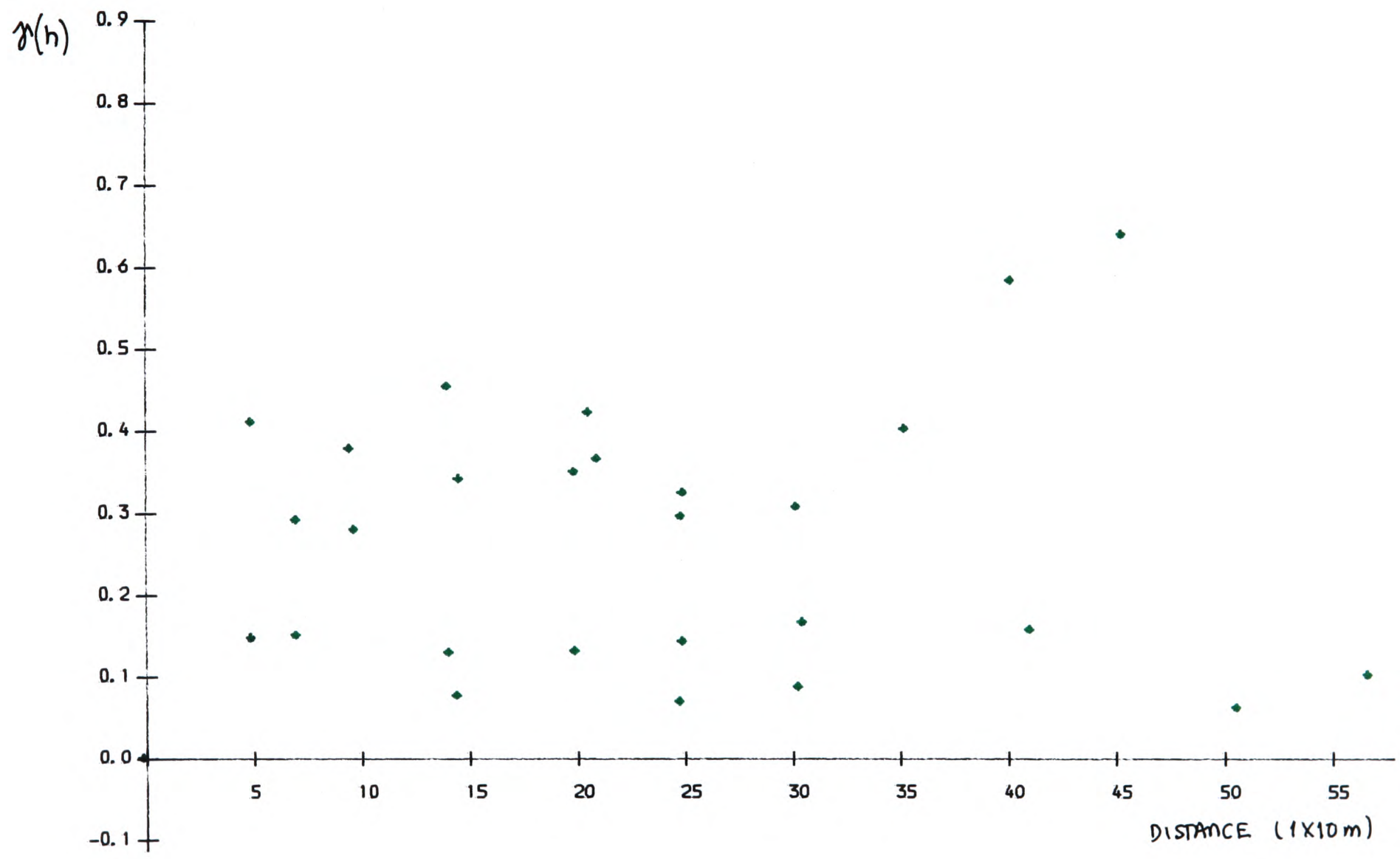


Fig 9.42 b SEMI-VARIOGRAM within MAP UNITS NS-NG, CARBONATE, 100m grid

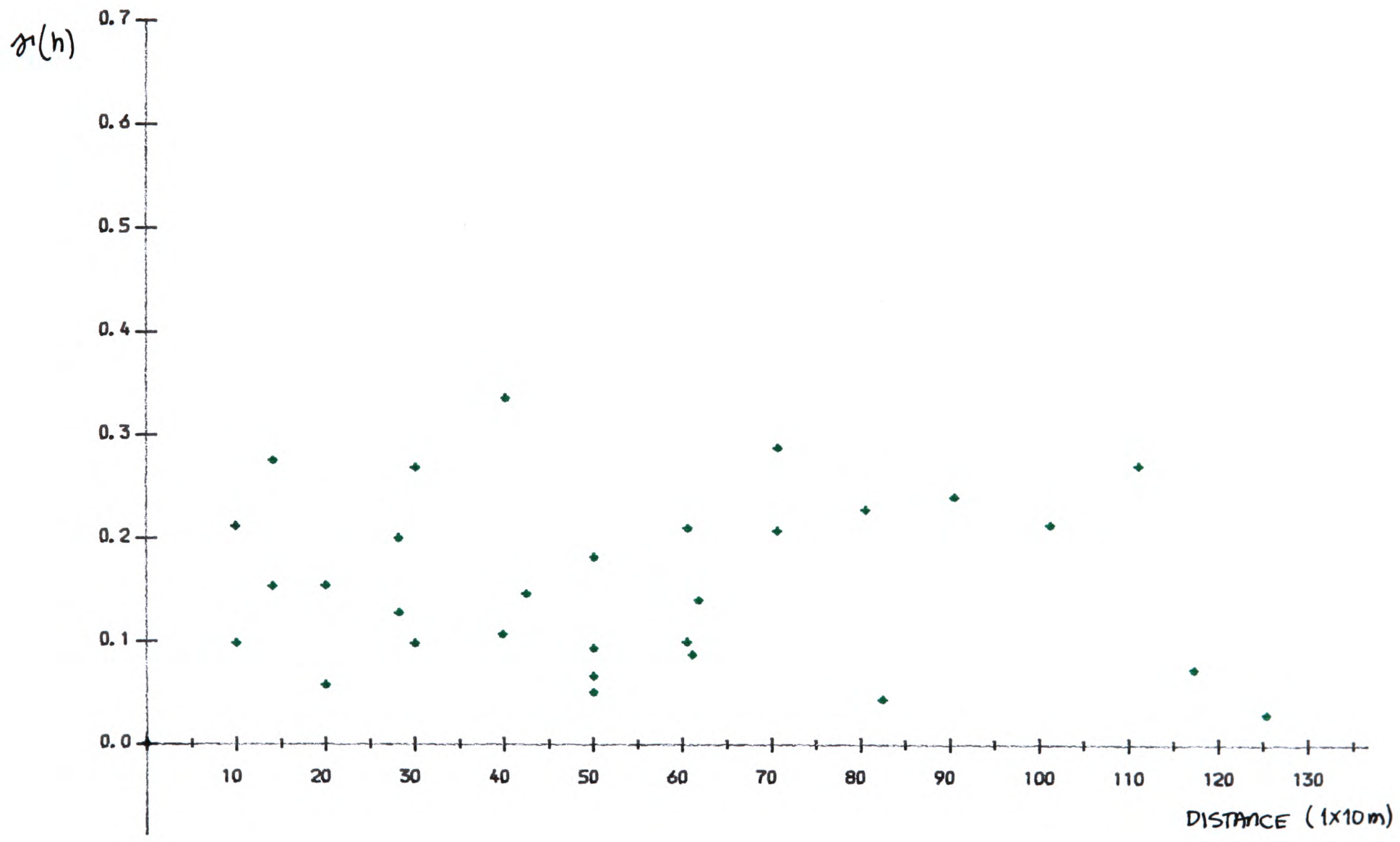


Fig. 9.42c SEMI-VARIOGRAM within MAP UNITS NS–NG, CARBONATE, 200m grid

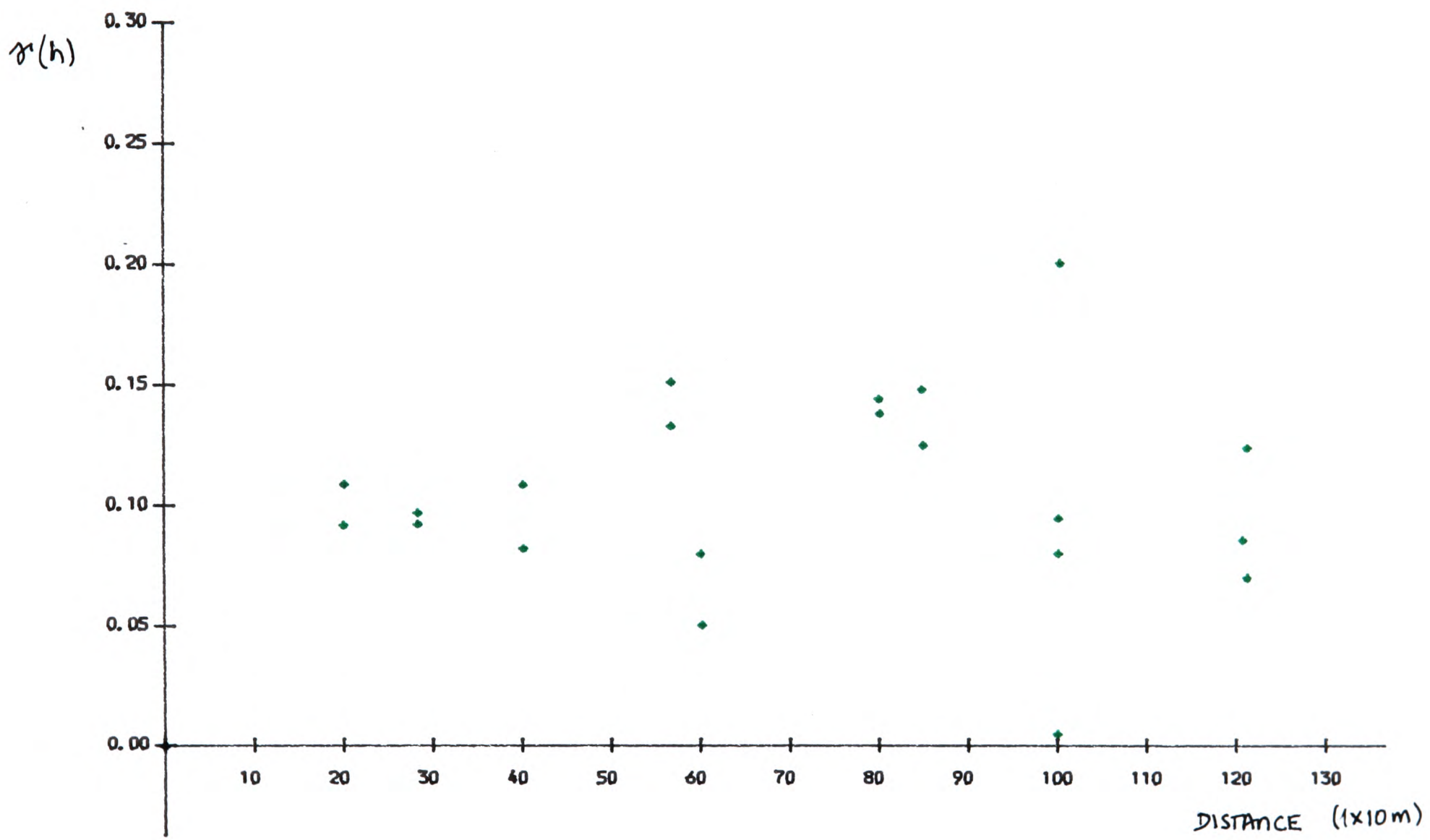


Fig. 9.43a SEMI-VARIOGRAM within MAP UNITS BC-BE, SILT 50m grid

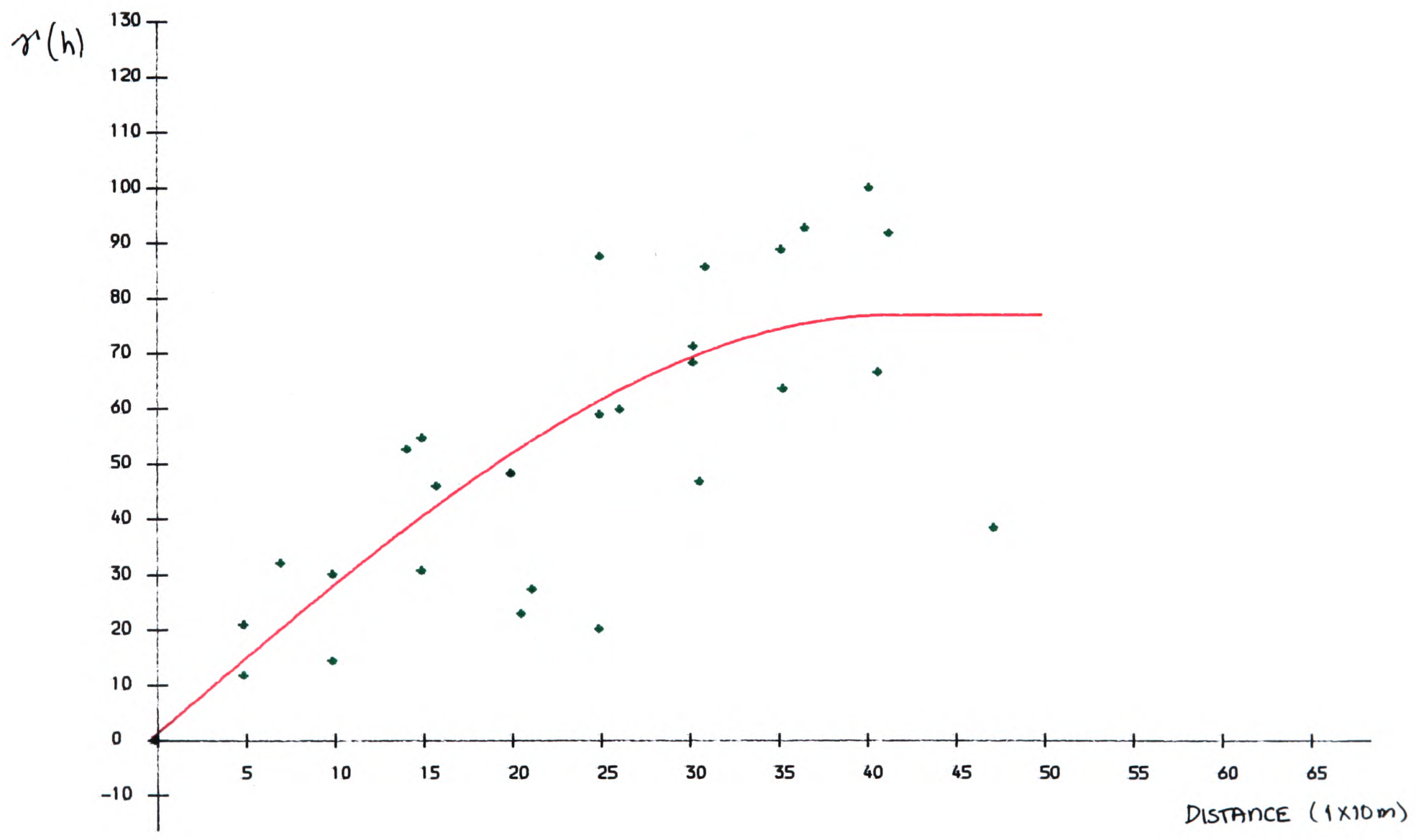
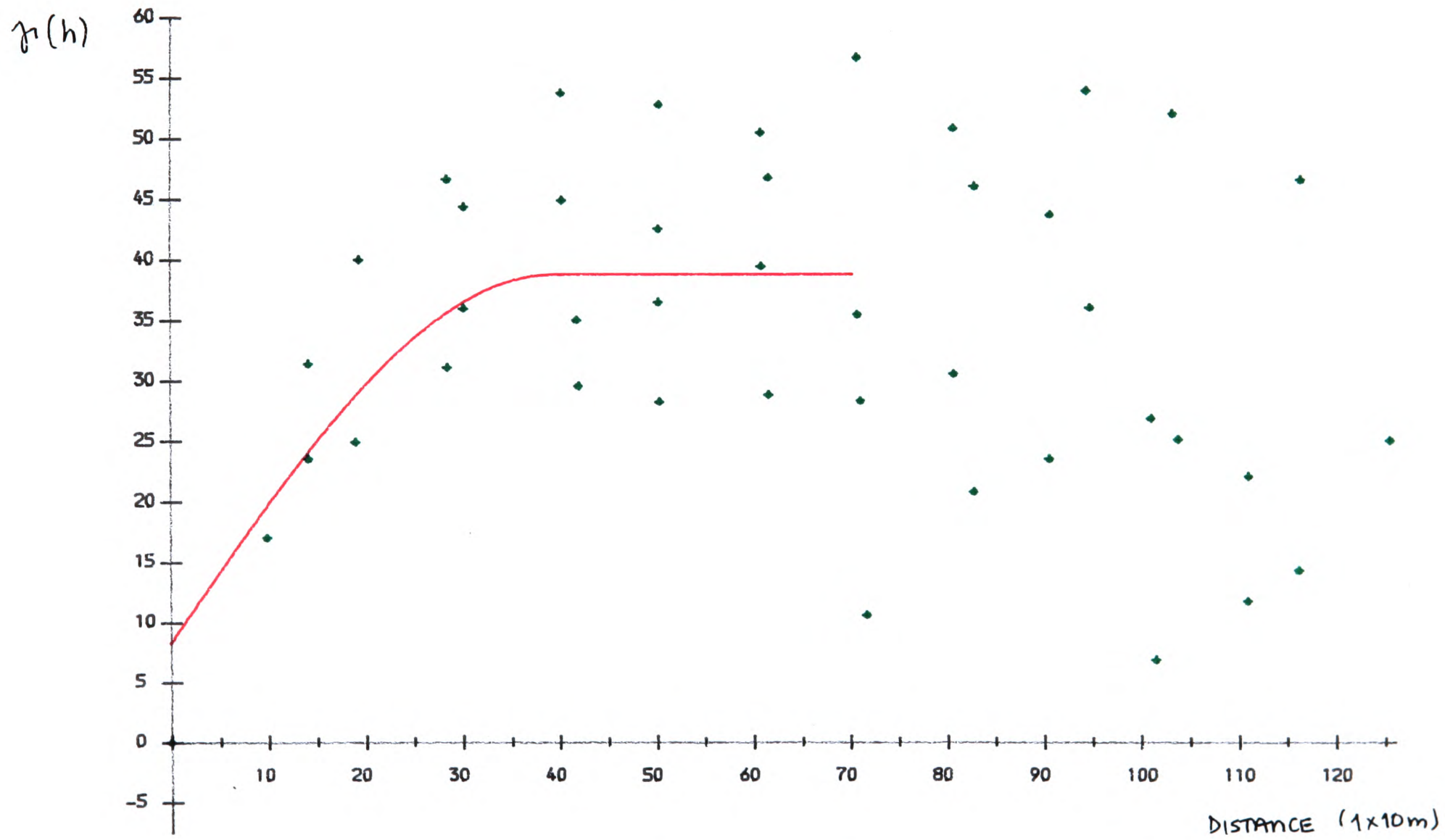


Fig. 9.43b SEMI-VARIOGRAM within MAP UNITS BC-BE, SILT, 100m grid



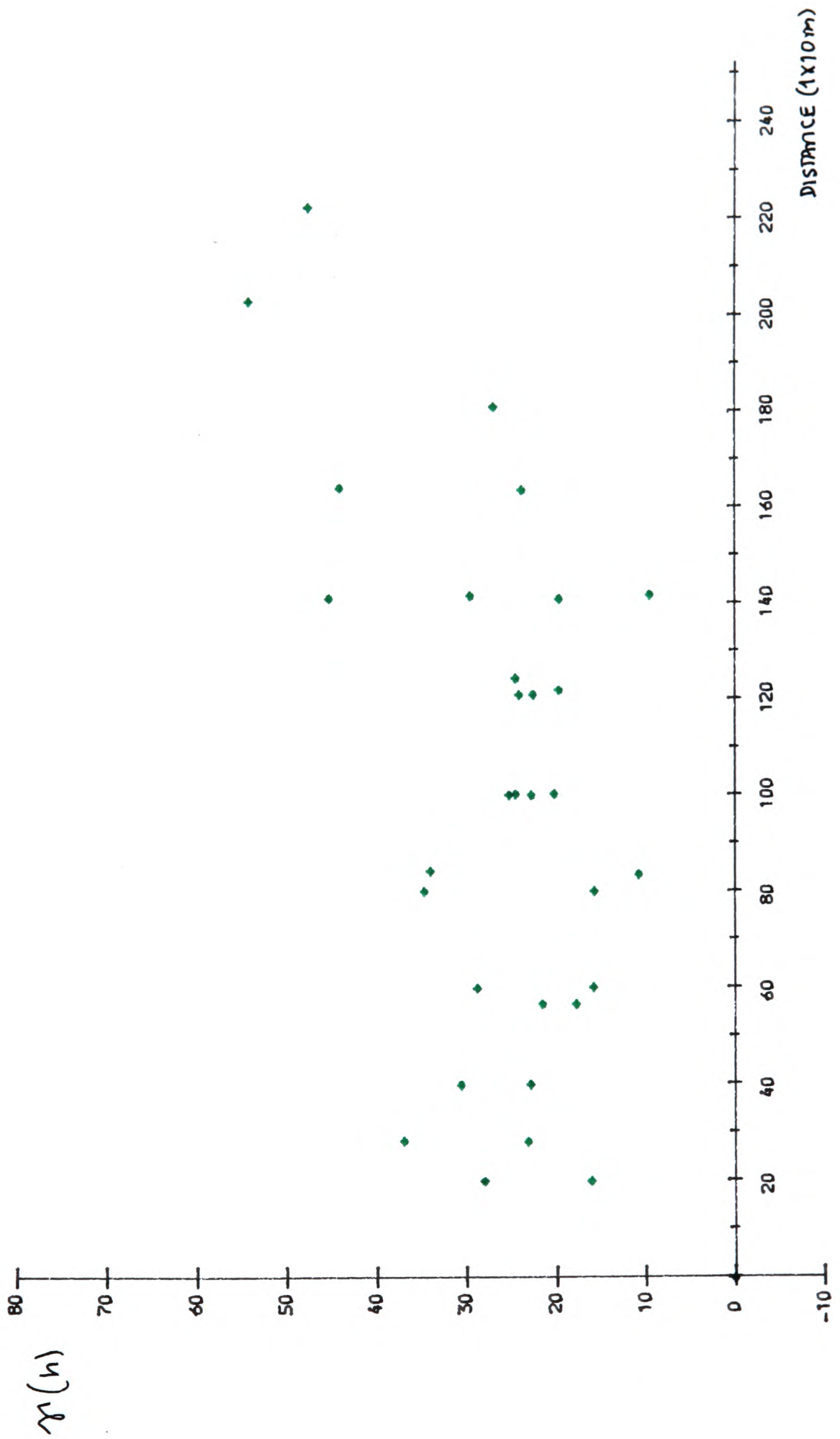


Fig 9.43c SEMI-VARIOGRAM within MAP UNITS BC-BE, SILT, 200m grid

Fig. 9 44a SEMI-VARIOGRAM within MAP UNITS NS-NG, SILT, 50m grid

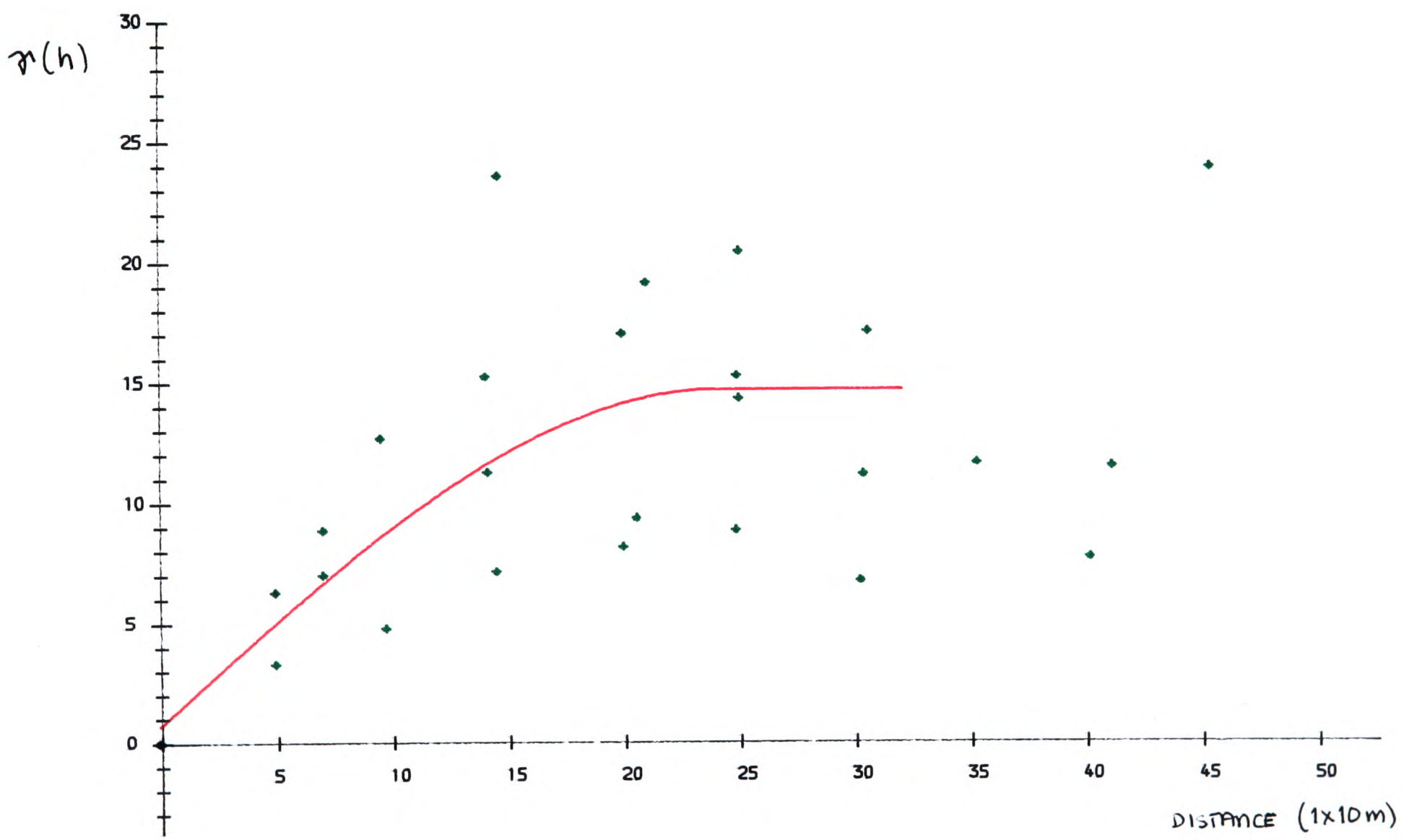


Fig. 9.44b SEMI-VARIOGRAM within MAP UNITS NS–NG, SILT, 100m grid

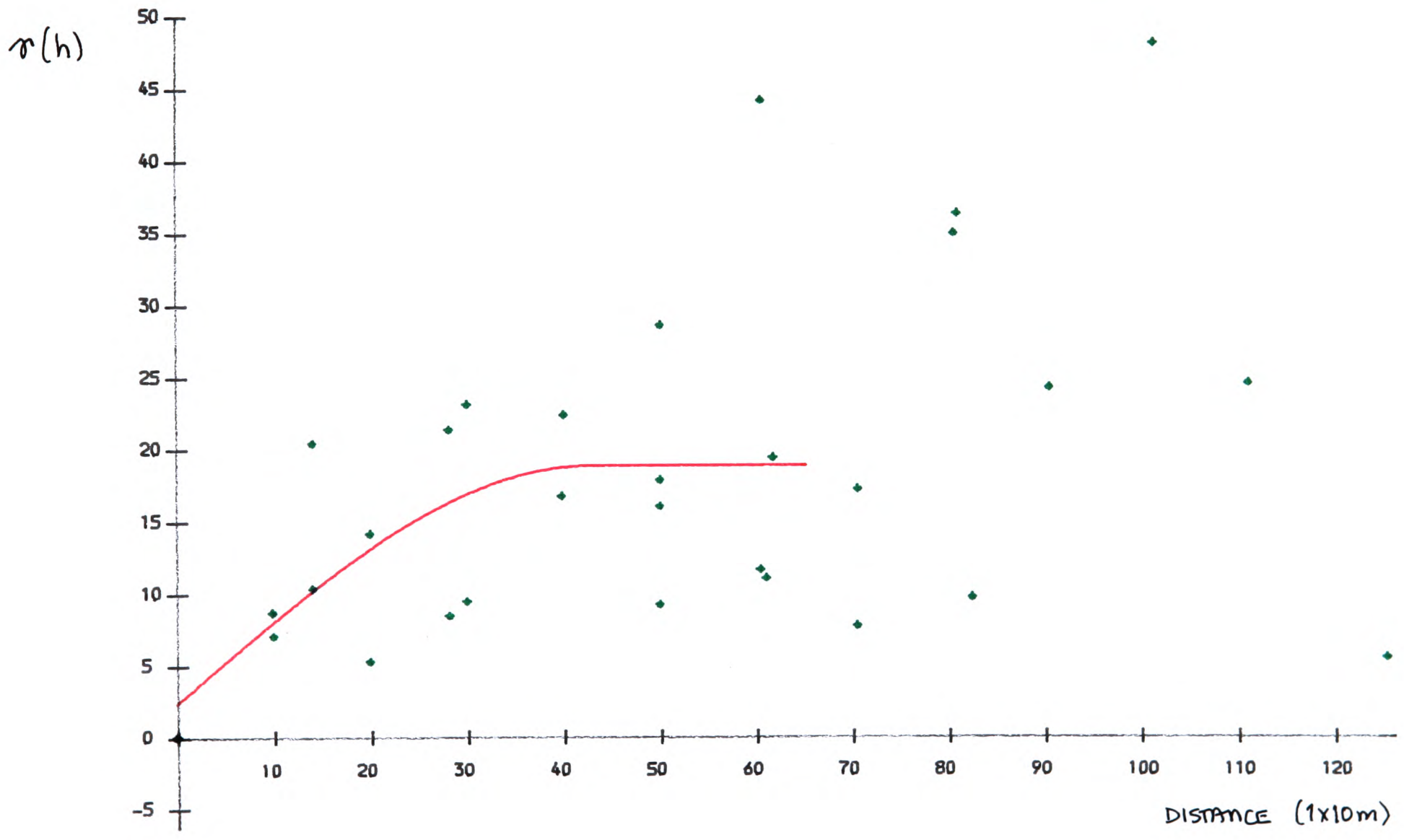
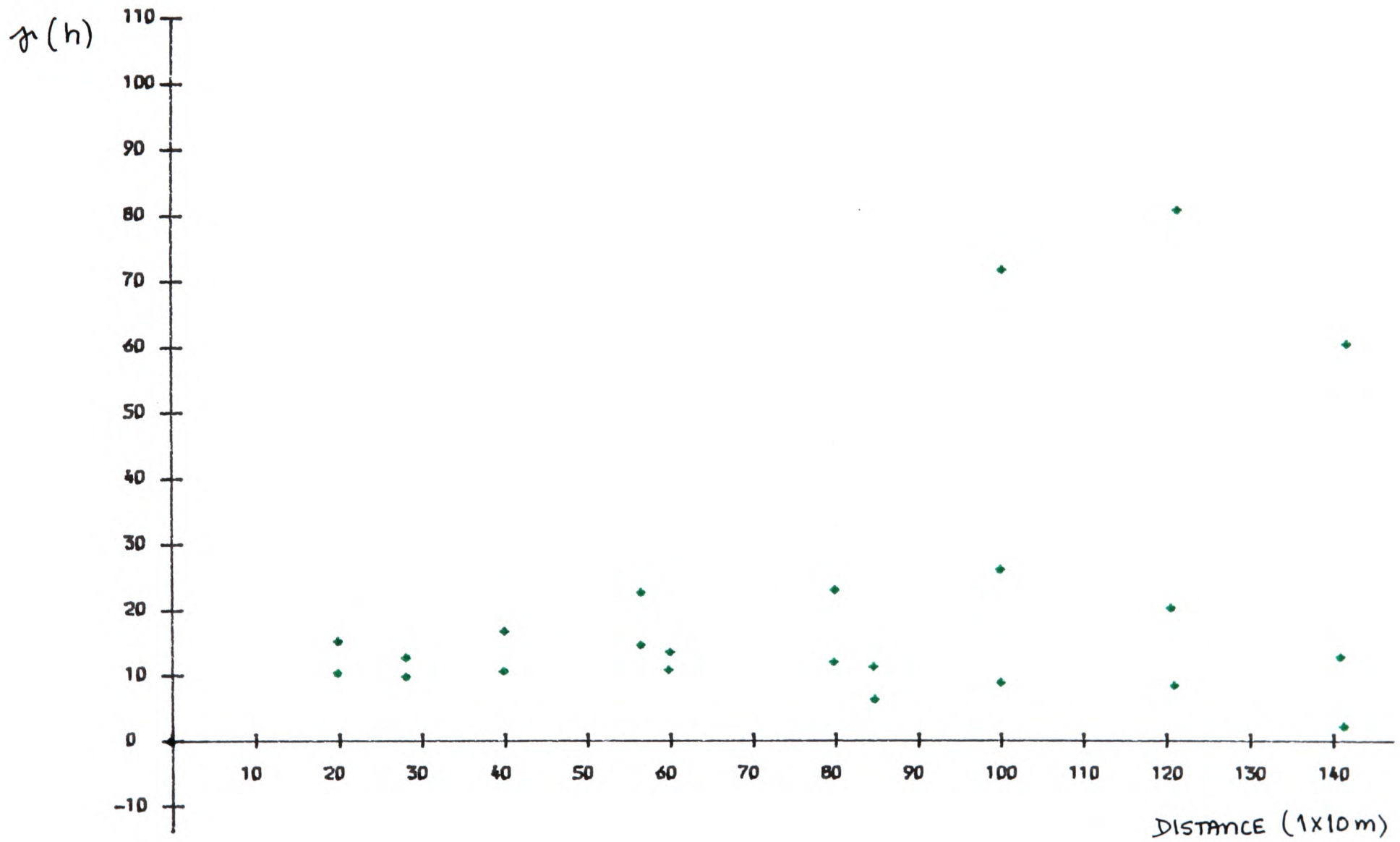


Fig 9.44 e. SEMI-VARIOGRAM within MAP UNITS NS-NG, SILT, 200m grid



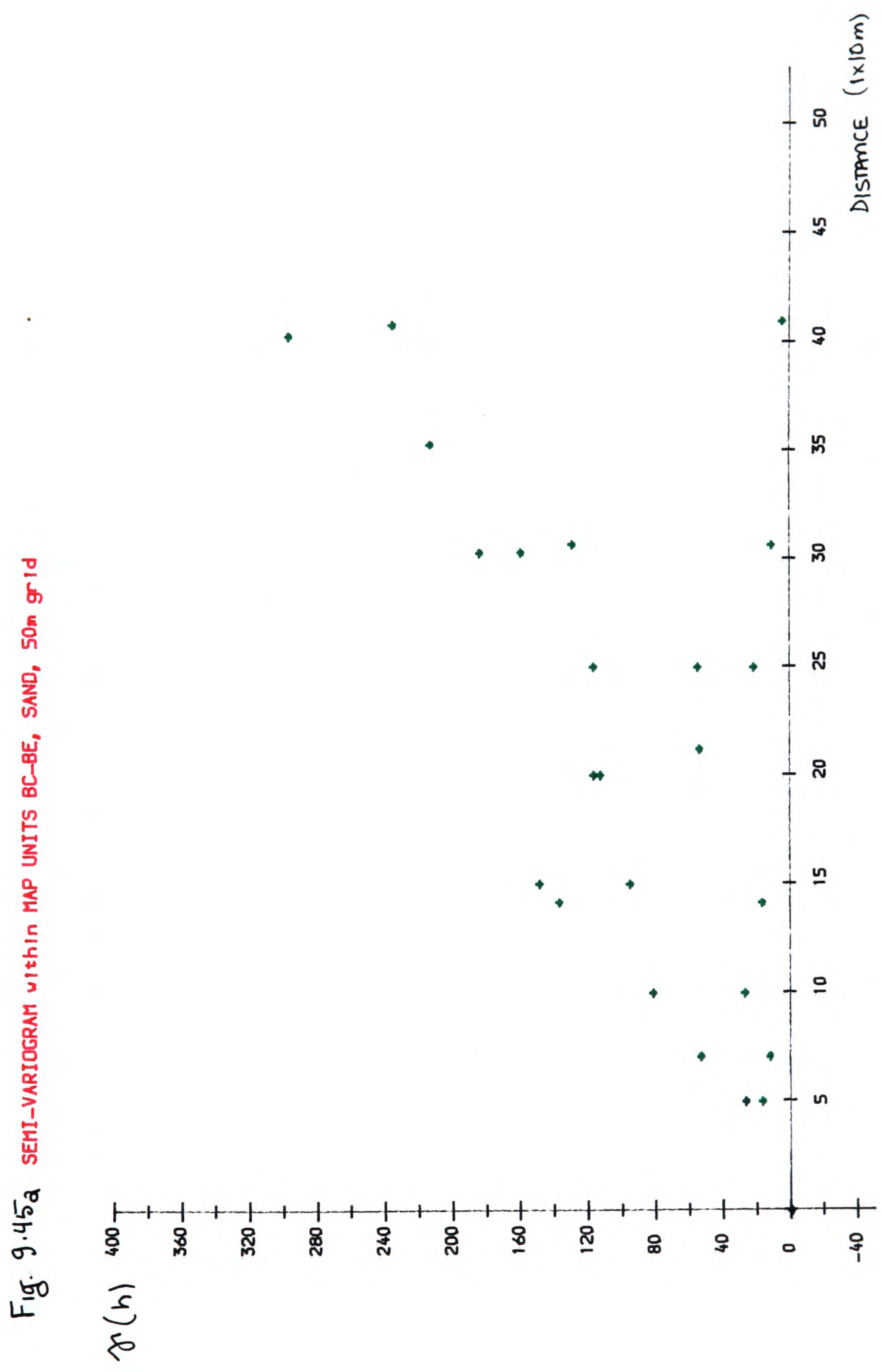

SEMI-VARIOGRAM within MAP UNITS BC-BE, SAND, 50m grid
γ(h)
400
360
320
280
240
200
160
120
80
40
0
-40
5
10
15
20
25
30
35
40
45
50
DISTANCE (1x10m)

Fig. 9.45a

Fig. 9.45c SEMI-VARIOGRAM within MAP UNITS BC-BE, SAND, 200m grid

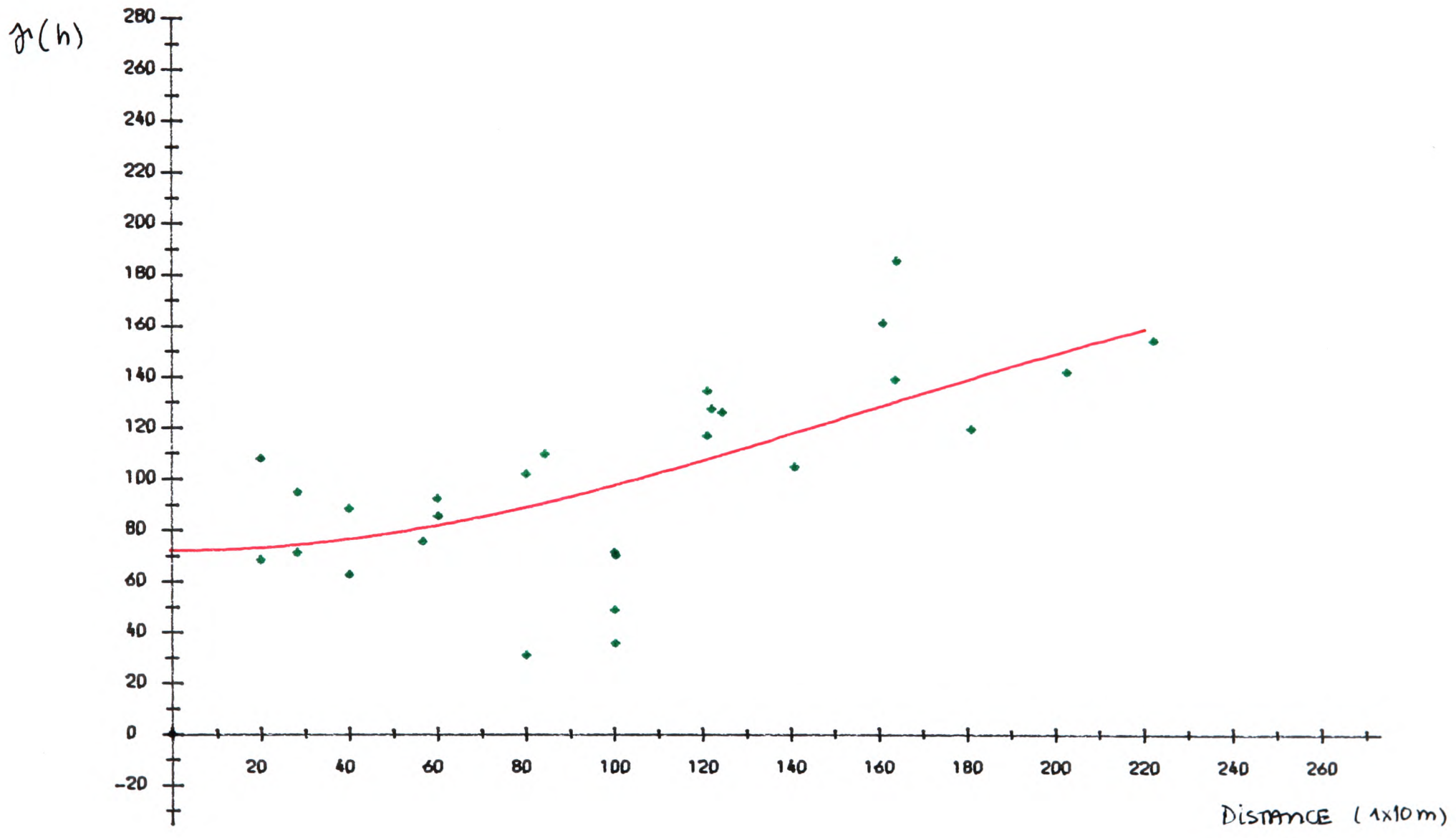
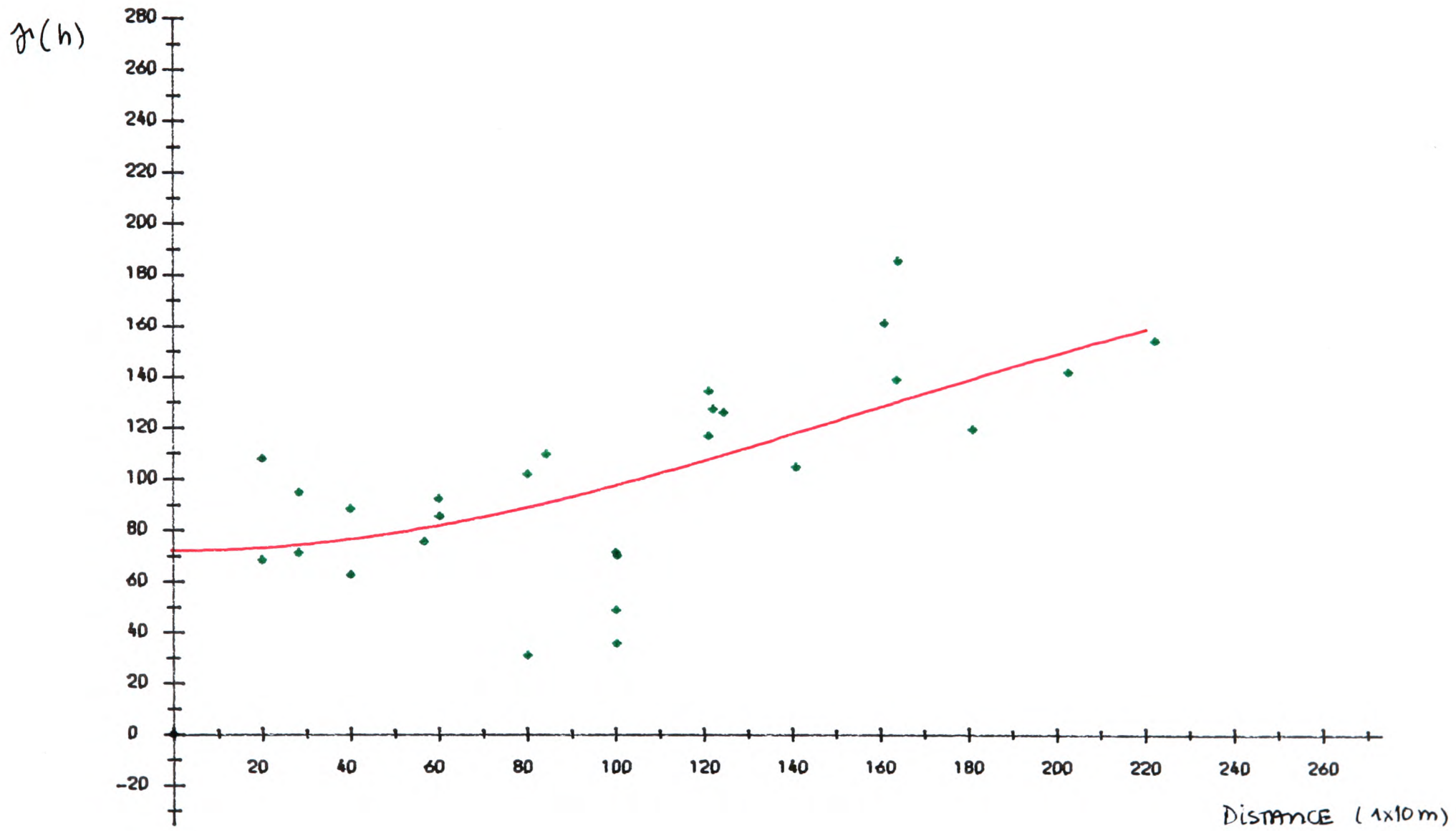


Fig. 9.45c SEMI-VARIOGRAM within MAP UNITS BC-BE, SAND, 200m grid



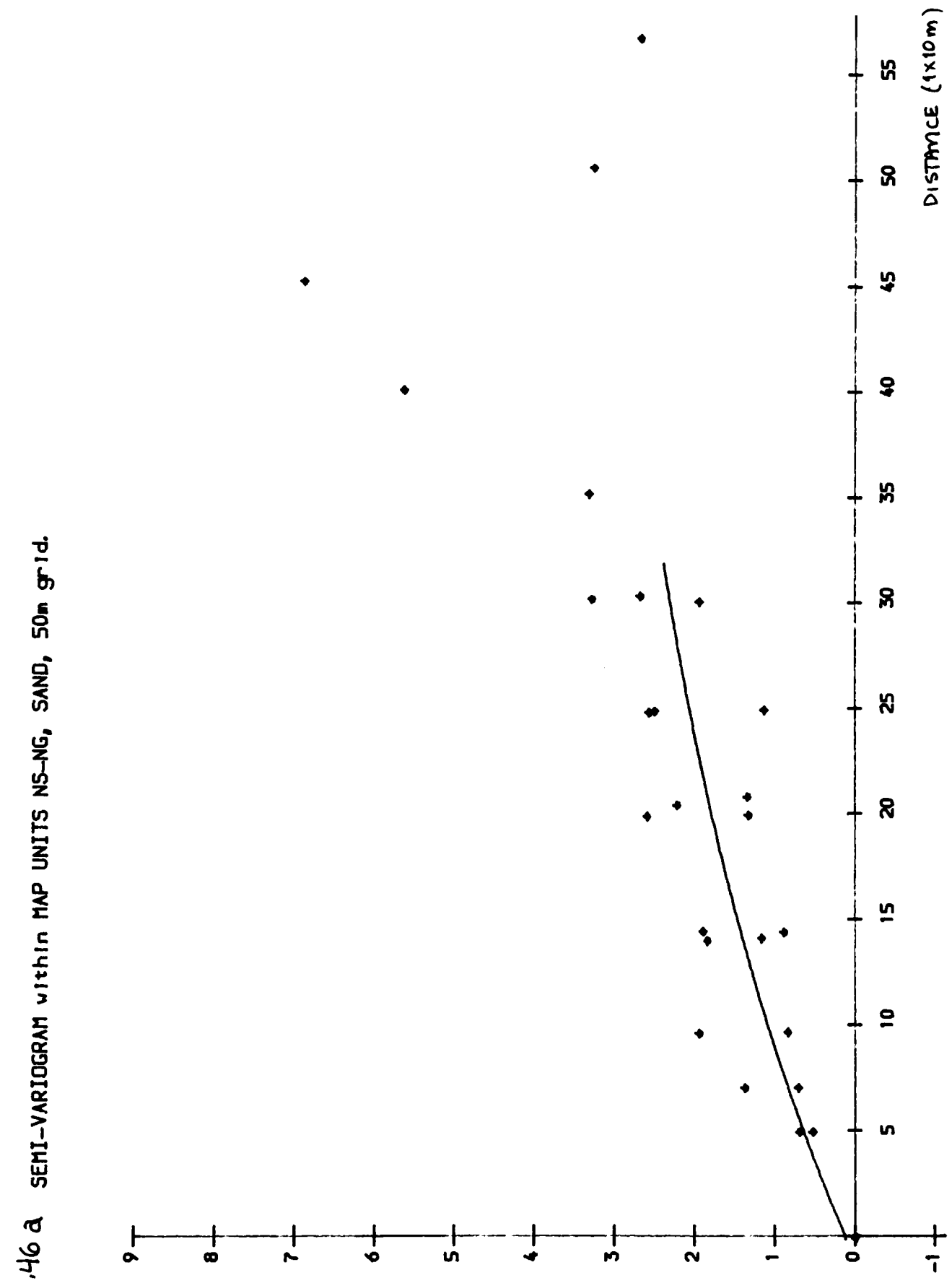


Fig 9.46 a SEMI-VARIOGRAM within MAP UNITS NS-NG, SAND, 50m grid.

Fig. 9.46b SEMI-VARIOGRAM within MAP UNITS NS-NG, SAND, 100m grid

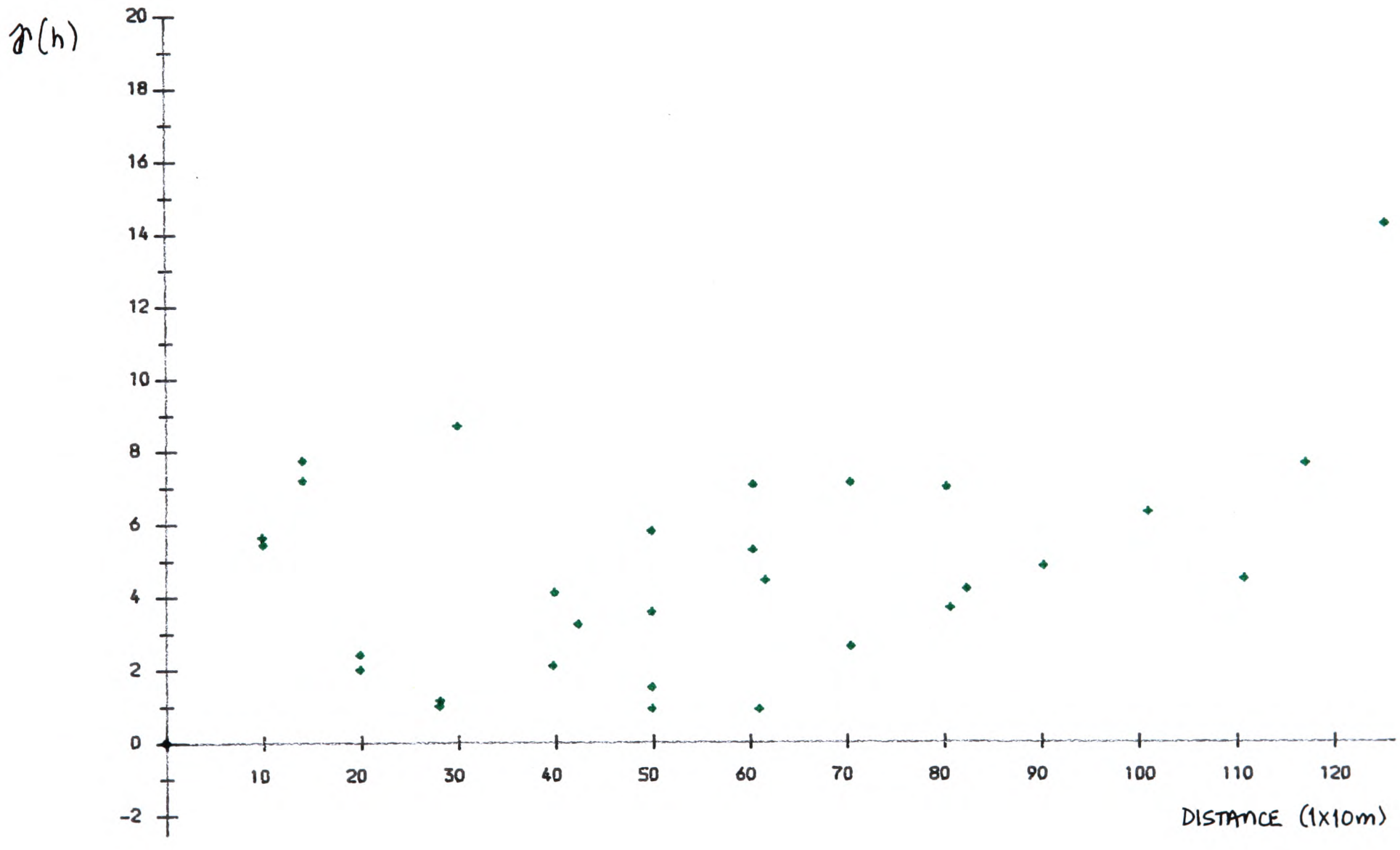
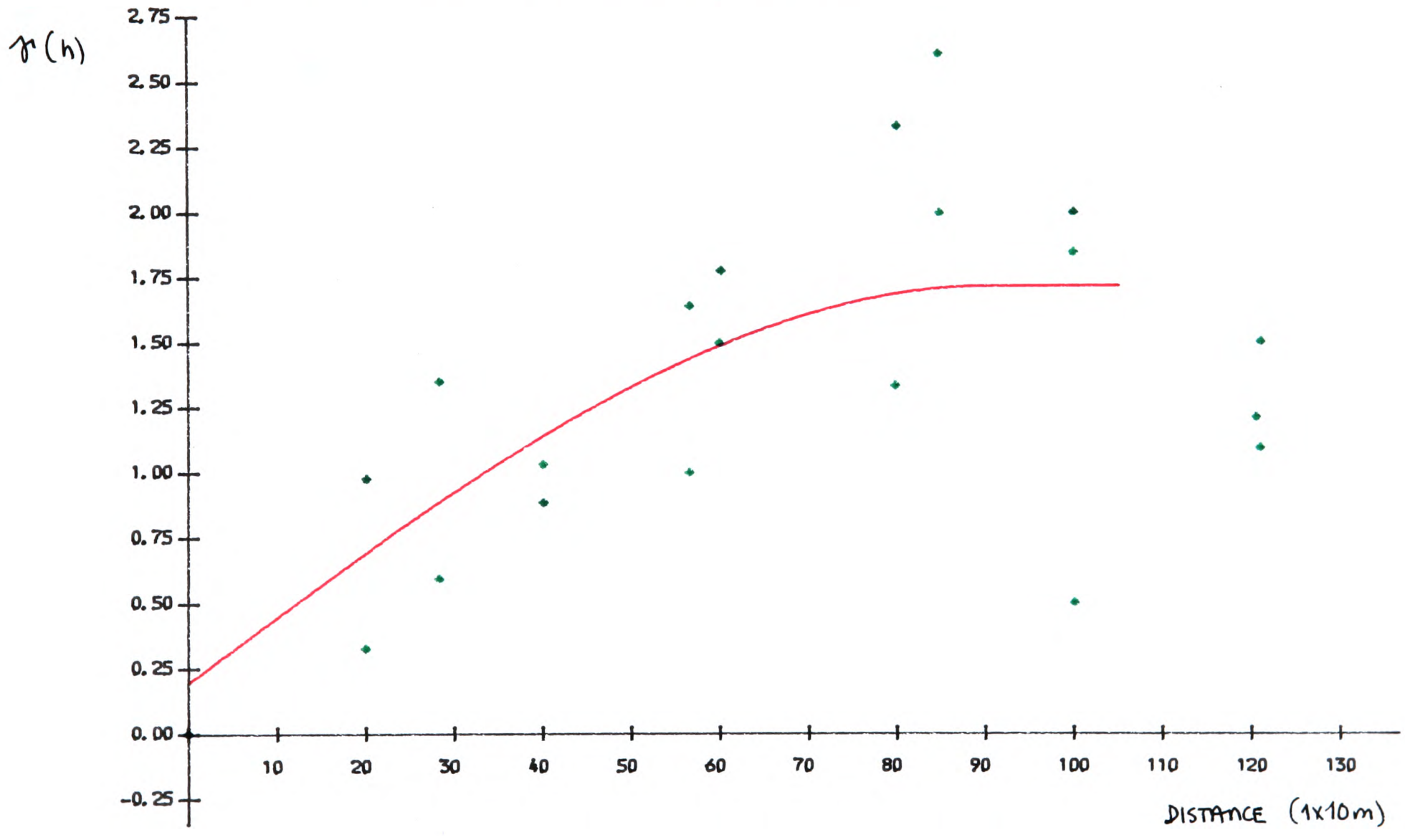


Fig. 9.46c SEMI-VARIOGRAM within MAP UNITS NS-NG, SAND, 200m grid



these semi-variograms allowed the application of Kriging within the stratum of their domain, which could not have been possible at this sampling scale without stratification.

The variation of Potassium in the calcareous soils, i.e. BC-BE, is spatially dependent only at sampling spacings of 100m or more (figs. 9.39b and c). The two semi-variograms were fitted by spherical models. At 50 m grid only the non-calcareous soils shown spatial dependence (fig. 9.40a) but with relatively high Nugget variance. Variation of Potassium at smaller sampling scales (figs. 9.40b-c) for these soils (NS-NG) was totally random.

Expectedly, Carbonate showed spatial structure at all sampling scales in the stratum (BC-BE) of calcareous soils (figs. 9.41a,b and c). This is an indirect indication of the usefulness of the grouping criterion employed. In contrast, Carbonate in the non-calcareous soils (NS-NG) showed (figs. 9.42a-c) variation totally random or Pure Nugget effect, which can be due to the low concentration of this mineral in this type of soils.

At sampling spacings smaller or as long as 100 m, Silt shows spatial structure (figs. 9.43a-b and 9.44a-b) in both calcareous and non-calcareous soils. When the spacing increases, i.e. 200m grid, variation of Silt in both type of soils is totally random (figs. 9.43c and 9.44c). Hence, spatial dependence in this property after stratification is restricted to 50 m and 100 m sampling spacings.

Coarse Sand showed spatial dependence only in the non-calcareous (NS-NG) soils at the 50 m grid (fig. 9.46a). Variation was very transitional and described by an exponential model. For the BC-BE soils at this scale variation of Coarse Sand (fig.9.45a) was random.

The type of variation displayed by Coarse Sand in the calcareous soils (BC-BE) at the 100 m and 200 m grids (figs. 9.45b and c) is quite similar to that of Carbonate in the same soils and at the same grids. On the whole, through this study, the spatial variability of these two properties has shown to be strongly correlated, and their association finds an explanation in the soil/landscape relationships. So, at the 100 m grid a spherical model described variation in both cases (figs. 9.41b

and 9.45b), and at the 200 m grid a Gaussian model with different parameters described also the variation of both properties (figs. 9.41c and 9.45c).

The model type and the set of semi-variogram parameters estimated for each of the BC-BE and NS-NG semi-variograms described above are given in table 9.31. The semi-variograms not listed in table 9.31 showed total random variation of the property (Nugget effect) at the scale of sampling. The best predicted value of the property involved in these cases, at any point within the corresponding stratum, is the general mean of the stratum which should be more accurate as a predictor than the general mean over the whole grid. This is important and should be born in mind while comparing predictive procedures after trend removal.

The parameters listed in table 9.31 were then used in the next stage to compute the interpolated values by Kriging at the random check points within the same stratum as the semi-variogram.

Table 9.31 Model type and parameters of semi-variorams computed within the strata defined by grouping soil mapping units in order to stratify trends.

Property	Grid (m)	Stratum	Model	Nugget V.	Sill	Range
Potassium	50	NS-NG	spherical	0.02	0.045	260
"	100	BC-BE	spherical	0.005	0.270	350
"	200	BC-BE	spherical	0.0001	0.40	980
Carbonate	50	BC-BE	spherical	0.05	16.0	210
"	100	BC-BE	spherical	2.4	9.6	390
"	200	BC-BE	gaussian	9.0	40.0	1750
Silt	50	BC-BE	spherical	0.5	77.0	415
"	50	NS-NG	spherical	0.7	14.7	240
"	100	BC-BE	spherical	8.5	39.0	390
"	100	NS-NG	spherical	2.5	19.0	420
Sand	50	NS-NG	exponential	0.1	3.7	320
"	100	BC-BE	spherical	4.5	102.00	310
"	200	BC-BE	gaussian	73.00	210.00	2200
"	200	NS-NG	spherical	0.20	1.72	900

9.20 Spatial interpolation of de-trended data by Kriging.

The semi-variograms of de-trended data displayed the spatial structures left in the data after the trends in the properties studied were removed by any of the approaches tried. The parameters of such semi-variograms (tables 9.30 and 9.31) provided the information needed by the Kriging algorithm so that predictions by spatial interpolation could take place.

The random check points found in the domain of the data specific to each semi-variogram were used as target points to interpolate by Kriging. The object was to obtain predicted values of the soil properties studied after the trend had been removed by the two alternative procedures and to compare the predictive success of Kriging and classical procedures of de-trended data against the predictive success of the same techniques when trends are ignored. It was also possible to compare the effectiveness of approaches used for removing the trends, in terms of their effect on accuracy on predictions by Kriging.

9.20.1 Kriging predictions using data de-trended by spline surfaces.

Values of the four soil properties involved in trend removal by spline surfaces were predicted by Kriging interpolation at the target points. The predicted values were calculated indirectly, since the semi-variograms used for Kriging were semi-variograms of residuals from trend, and not those of the original data. Obtaining the Kriging predictions at the check points in each grid with de-trended data involved, then, the following operations:

9.20.1.1 Kriging of the residuals from trend.

The semi-variograms of the residuals from the trend, at the grid points, were used to provide the parameters for Kriging. The random check points were the targets at which residuals,

rather than property values, were interpolated. The result were the predicted residuals of the four properties from their respective trend surfaces at the random check points. These are given in tables 9.32, 9.33 and 9.34 for the three sampling grids.

9.20.1.2 Evaluation of the trend surface at the random check points (estimation of trend values).

The computed bicubic spline surfaces used to model trends, needed to be evaluated at the random check points to obtain the value of the surface at such locations.

Under the intrinsic hypothesis (recalling equation 7.31 in chapter 7) the values of the property Z(x) are given by the values of the trend **m** and the residual terms **e(x)**:

$$Z(x) = m + e(x).$$

So, in order to obtain the predicted values Z(x) of the soil property, the values of **m** accounting for the trend were added to the interpolated residuals **e(x)** by Kriging.

In effect, the trend surface is providing, as a result of this evaluation, interpolated values of the trend, that is, the values of the "drift" **m** in equation 7.31 (see chapter 7), at the random check points. Tables 9.35a,b and c give the trend values of Potassium evaluated on the surface at the random check points in the 50 m, 100 m and 200 m grids respectively. In the same order tables 9.36a,b and c give the trend values of Carbonate, tables 9.37a,b and c list the trend values of Silt, and tables 9.38a,b and c give the values of Coarse Sand.

9.20.1.3 Kriging predictions from addition of trend values to the residuals interpolated by Kriging.

The values of the four soil properties predicted by Kriging, after the trends were accounted for, were obtained by adding up the values of trend from the evaluation of the surface shown in tables 9.35 to 9.38, to the values of the kriged residuals in tables 9.32, 9.33 and 9.34. The predicted values resulting from this operation are listed in tables 9.39a-c for the three sampling grids.

Table 9.32 . Values of residuals from trend predicted by Kriging interpolation at the random check points in the 50 m grid.

NORTHING	EASTING	K	CARBONATE	SILT	COARSE SAND
20511	43547	-0.030811	0.030801	0.0002959	-0.0000068
20513	43542	-0.017616	0.030801	0.0002959	-0.0000068
20519	43560	-0.003573	0.030801	0.0002959	-0.0000068
20502	43538	-0.028546	0.030801	0.0002959	-0.0000068
20500	43542	-0.011036	0.030801	0.0002959	-0.0000068
20494	43548	-0.003086	0.030801	0.0002959	-0.0000068
20491	43538	0.030459	0.030801	0.0002959	-0.0000068
20487	43558	0.036616	0.030801	0.0002959	-0.0000068
20468	43557	0.193848	0.030801	0.0002959	-0.0000068
20468	43561	0.377671	0.030801	0.0002959	-0.0000068
20462	43544	0.950912	0.030801	0.0002959	-0.0000068
20461	43547	-0.307163	0.030801	0.0002959	-0.0000068
20458	43560	0.575887	0.030801	0.0002959	-0.0000068
20454	43544	0.648357	0.030801	0.0002959	-0.0000068
20511	43560	0.628116	0.030801	0.0002959	-0.0000068
20503	43545	0.261402E	0.030801	0.0002959	-0.0000068

Table 9.33. Values of residuals from trend predicted by Kriging interpolation at the random check points in the 100m grid.

Northing	Easting	K	Carbonate	Silt	Coarse Sand
205400	435030	0.1236		2.760	-0.0000068
205330	435710	0.0334		0.9459	-0.0000068
205110	435470	-0.285		3.082	-0.0000068
205020	435120	0.0343		2.214	-0.0000068
204660	435720	0.101		0.7334	-0.0000068
205440	435720	0.0337		0.3416	-0.0000068
205310	435040	0.0267		2.659	-0.0000068
205210	434950	-0.220		2.834	-0.0000068
205320	435440	0.0923		0.3065	-0.0000068
205320	435900	0.0556		1.356	-0.0000068
205240	435390	0.113		1.706	-0.0000068
205080	434910	0.326		2.770	-0.0000068
205000	435790	-0.240		1.040	-0.0000068
204910	435820	0.349		1.154	-0.0000068
204610	435090	0.0741		0.1993	-0.0000068
204550	435760	0.0216		0.2792	-0.0000068
205130	435420	-0.103		2.614	-0.0000068
205190	435600	0.128		4.250	-0.0000068
205020	435380	0.0441		2.202	-0.0000068
205000	435420	-0.0301		1.794	-0.0000068
204940	435480	-0.102		1.046	-0.0000068
204910	435380	0.145		1.323	-0.0000068
204870	435580	0.0707		1.829	-0.0000068
204680	435570	-0.106		0.7209	-0.0000068
204680	435610	-0.0645		0.6589	-0.0000068
204620	435440	-0.0523		0.9882	-0.0000068
204610	435470	-0.0860		1.377	-0.0000068
204580	435600	0.0256		0.8453	-0.0000068
205130	435750	-0.0143		1.489	-0.0000068
204530	435720	0.0594		0.09788	-0.0000068
205600	435520	0.0977		1.592	-0.0000068
204500	435580	0.00739		0.4590	-0.0000068
205280	435240	0.0552		4.276	-0.0000068
205530	435410	0.104		2.194	-0.0000068
204540	435440	-0.0104		0.9242	-0.0000068
205150	435670	0.146		2.566	-0.0000068
205000	435630	0.0439		1.053	-0.0000068
205110	435600	0.140		3.327	-0.0000068
205030	435450	-0.205		1.890	-0.0000068

Table 9.34. Values of residuals from trend predicted by Kriging interpolation at the random check points in the 200m grid.

Northing	Easting	K	Carbonate	Silt	Coarse Sand
20563	43584	0.00431	0.0000039	-0.00483	0.2427
20540	43503	0.00183	0.0000039	-0.00483	0.4651
20533	43571	0.130	0.0000039	-0.00483	2.490
20511	43547	-0.127	0.0000039	-0.00483	2.853
20502	43512	0.0450	0.0000039	-0.00483	1.277
20466	43572	0.341	0.0000039	-0.00483	3.480
20446	43558	-0.0458	0.0000039	-0.00483	2.173
20415	43481	0.133	0.0000039	-0.00483	-0.01442
20397	43501	0.230	0.0000039	-0.00483	0.02948
20382	43572	0.0175	0.0000039	-0.00483	0.9104
20372	43563	-0.0275	0.0000039	-0.00483	0.6954
20544	43572	0.0504	0.0000039	-0.00483	1.664
20531	43504	0.0130	0.0000039	-0.00483	0.7876
20521	43495	0.0456	0.0000039	-0.00483	0.7749
20532	43544	0.00627	0.0000039	-0.00483	0.3520
20532	43590	-0.0392	0.0000039	-0.00483	1.444
20524	43539	0.00160	0.0000039	-0.00483	1.390
20508	43491	0.0999	0.0000039	-0.00483	1.105
20500	43579	0.104	0.0000039	-0.00483	5.285
20491	43582	0.309	0.0000039	-0.00483	4.494
20461	43509	-0.0416	0.0000039	-0.00483	0.3001
20475	43572	0.332	0.0000039	-0.00483	3.650
20455	43576	0.304	0.0000039	-0.00483	3.746
20513	43542	-0.0432	0.0000039	-0.00483	2.531
20501	43559	-0.0668	0.0000039	-0.00483	4.346
20481	43538	-0.0153	0.0000039	-0.00483	0.5399
20519	43560	-0.0479	0.0000039	-0.00483	3.871
20502	43538	0.00934	0.0000039	-0.00483	2.105
20500	43542	-0.0260	0.0000039	-0.00483	1.898
20494	43548	-0.0336	0.0000039	-0.00483	1.368
20491	43538	0.0204	0.0000039	-0.00483	0.8270
20487	43558	0.0405	0.0000039	-0.00483	2.533
20468	43557	0.0903	0.0000039	-0.00483	2.313
20468	43561	-0.163	0.0000039	-0.00483	2.731
20462	43544	-0.0572	0.0000039	-0.00483	1.050
20461	43547	-0.0452	0.0000039	-0.00483	1.240
20458	43560	0.0778	0.0000039	-0.00483	2.755
20418	43468	0.0435	0.0000039	-0.00483	-0.1143
20567	43605	0.00559	0.0000039	-0.00483	-0.008242
20567	43578	-0.00153	0.0000039	-0.00483	0.09988
20513	43575	-0.0319	0.0000039	-0.00483	6.180
20453	43572	0.206	0.0000039	-0.00483	3.981
20560	43552	-0.0238	0.0000039	-0.00483	-0.001727
20450	43588	0.477	0.0000039	-0.00483	2.606
20528	43524	0.0249	0.0000039	-0.00483	1.075
20553	43541	-0.0194	0.0000039	-0.00483	0.05777
20454	43544	-0.0739	0.0000039	-0.00483	0.9624
20515	43567	-0.0448	0.0000039	-0.00483	5.858
20500	43563	-0.0376	0.0000039	-0.00483	5.076
20515	43561	-0.0786	0.0000039	-0.00483	4.751
20511	43560	-0.108	0.0000039	-0.00483	4.958
20503	43545	-0.0783	0.0000039	-0.00483	2.358

Table 9.35a. Values of POTASSIUM evaluated at the RANDOM CHECK POINTS on the TREND SURFACE over the 50 m grid at Yelford.

POTASSIUM, values at CHECK POINTS on BISPLINE SURFACE, 50m grid

INTERIOR X-KNOTS

2.0460E+04
2.0476E+04
2.0500E+04
2.0515E+04

INTERIOR Y-KNOTS

4.3539E+04
4.3550E+04
4.3557E+04

SUM OF SQUARES OF RESIDUAL RHS 3.0014E+00

RANK 49

RANDOM CHECK POINTS

NORTHING	EASTING	OBSERVED	SURFACE FIT	RESIDUALS
2.0454E+04	4.3544E+04	3.0000E-01	4.5033E-01	1.5033E-0
2.0458E+04	4.3560E+04	6.0000E-01	7.1072E-01	1.1072E-0
2.0462E+04	4.3544E+04	2.0000E-01	5.4410E-01	3.4410E-0
2.0461E+04	4.3547E+04	4.0000E-01	5.3588E-01	1.3588E-0
2.0468E+04	4.3557E+04	1.1000E+00	4.5896E-01	-6.4104E-0
2.0468E+04	4.3561E+04	1.1000E+00	6.3254E-01	-4.6746E-0
2.0491E+04	4.3538E+04	3.0000E-01	3.3231E-01	3.2312E-0
2.0494E+04	4.3548E+04	8.0000E-01	8.4039E-01	4.0387E-0
2.0487E+04	4.3558E+04	4.0000E-01	4.4713E-01	4.7128E-0
2.0502E+04	4.3538E+04	7.0000E-01	4.2459E-01	-2.7541E-0
2.0511E+04	4.3547E+04	6.0000E-01	8.8361E-01	2.8361E-0
2.0513E+04	4.3542E+04	7.0000E-01	7.3839E-01	3.8388E-0
2.0500E+04	4.3542E+04	7.0000E-01	5.0157E-01	-1.9843E-0
2.0503E+04	4.3545E+04	5.0000E-01	7.4018E-01	2.4018E-0
2.0511E+04	4.3560E+04	1.0000E-01	1.0016E+00	9.0156E-0
2.0519E+04	4.3560E+04	9.0000E-01	2.0778E-01	-6.9222E-0

SUM OF SQUARED RESIDUALS 2.3529E+00

Table 9.35b. values of POTASSIUM evaluated the RANDOM CHECK POINTS on the TREND SURFACE over the 100m grid at Yelford.

INTERIOR X-KNOTS
2.0460E+04
2.0494E+04
2.0523E+04
2.0550E+04

INTERIOR Y-KNOTS
4.3498E+04
4.3534E+04
4.3571E+04

RANDOM CHECK POINTS

NORTHING	EASTING	SURFACE FIT
2.0458E+04	4.3560E+04	8.8856E-01
2.0454E+04	4.3544E+04	5.3287E-01
2.0455E+04	4.3576E+04	2.4602E-01
2.0453E+04	4.3572E+04	6.5348E-01
2.0450E+04	4.3588E+04	3.8565E-01
2.0461E+04	4.3509E+04	2.4114E-01
2.0491E+04	4.3538E+04	3.6421E-01
2.0487E+04	4.3558E+04	5.0692E-01
2.0468E+04	4.3557E+04	1.2741E-01
2.0468E+04	4.3561E+04	1.1732E-01
2.0462E+04	4.3544E+04	4.1718E-01
2.0461E+04	4.3547E+04	4.7126E-01
2.0466E+04	4.3572E+04	3.4318E-01
2.0491E+04	4.3582E+04	1.0073E+00
2.0521E+04	4.3495E+04	1.1419E+00
2.0508E+04	4.3491E+04	8.8732E-01
2.0502E+04	4.3512E+04	5.3616E-01
2.0511E+04	4.3547E+04	7.8233E-01
2.0513E+04	4.3542E+04	7.0573E-01
2.0519E+04	4.3560E+04	8.7406E-01
2.0502E+04	4.3538E+04	4.6897E-01
2.0500E+04	4.3542E+04	5.2827E-01
2.0494E+04	4.3548E+04	5.7381E-01
2.0515E+04	4.3567E+04	9.4189E-01
2.0500E+04	4.3563E+04	1.0196E+00
2.0511E+04	4.3560E+04	1.0020E+00
2.0503E+04	4.3545E+04	6.5195E-01
2.0500E+04	4.3579E+04	8.3510E-01
2.0513E+04	4.3575E+04	8.1828E-01
2.0540E+04	4.3503E+04	7.7668E-01
2.0531E+04	4.3504E+04	1.0406E+00
2.0528E+04	4.3524E+04	8.4624E-01
2.0532E+04	4.3544E+04	5.8830E-01
2.0524E+04	4.3539E+04	7.1672E-01
2.0533E+04	4.3571E+04	4.9376E-01
2.0544E+04	4.3572E+04	5.1972E-01
2.0532E+04	4.3590E+04	8.9415E-01
2.0560E+04	4.3552E+04	7.3845E-01
2.0553E+04	4.3541E+04	5.1829E-01
SUM OF SQUARED RESIDUALS		7.3524E+00

Table 9.35c. Values of POTASSIUM evaluated at the RANDOM CHECK POINTS on the TREND SURFACE over the 200m grid at Yelford.

INTERIOR X-KNOTS
2.0390E+04
2.0460E+04
2.0530E+04
INTERIOR Y-KNOTS
4.3483E+04
4.3525E+04
4.3569E+04

RANDOM CHECK POINTS

NORTHING	EASTING	SURFACE FIT
2.0372E+04	4.3563E+04	1.2938E-01
2.0382E+04	4.3572E+04	6.6828E-01
2.0415E+04	4.3481E+04	5.8575E-01
2.0418E+04	4.3468E+04	4.9485E-01
2.0397E+04	4.3501E+04	8.5951E-01
2.0446E+04	4.3558E+04	2.1783E-01
2.0458E+04	4.3560E+04	2.8280E-01
2.0454E+04	4.3544E+04	1.7907E-01
2.0455E+04	4.3576E+04	7.8027E-01
2.0453E+04	4.3572E+04	5.8800E-01
2.0450E+04	4.3588E+04	1.1757E+00
2.0502E+04	4.3512E+04	5.6651E-01
2.0521E+04	4.3495E+04	1.0988E+00
2.0508E+04	4.3491E+04	6.4658E-01
2.0461E+04	4.3509E+04	1.6177E-01
2.0528E+04	4.3524E+04	7.8607E-01
2.0511E+04	4.3547E+04	7.7061E-01
2.0524E+04	4.3539E+04	6.1529E-01
2.0513E+04	4.3542E+04	7.0358E-01
2.0501E+04	4.3559E+04	8.6421E-01
2.0481E+04	4.3538E+04	3.8363E-01
2.0519E+04	4.3560E+04	9.6746E-01
2.0502E+04	4.3538E+04	6.1004E-01
2.0500E+04	4.3542E+04	6.3456E-01
2.0494E+04	4.3548E+04	6.3121E-01
2.0491E+04	4.3538E+04	5.0160E-01
2.0487E+04	4.3558E+04	6.4663E-01
2.0468E+04	4.3557E+04	3.3626E-01
2.0468E+04	4.3561E+04	4.0330E-01
2.0462E+04	4.3544E+04	2.0363E-01
2.0461E+04	4.3547E+04	1.9797E-01
2.0515E+04	4.3567E+04	1.0362E+00
2.0500E+04	4.3563E+04	9.0683E-01
2.0515E+04	4.3561E+04	9.8396E-01
2.0511E+04	4.3560E+04	9.6060E-01
2.0503E+04	4.3545E+04	6.9686E-01
2.0466E+04	4.3572E+04	7.1697E-01
2.0500E+04	4.3579E+04	1.0949E+00
2.0491E+04	4.3582E+04	1.1788E+00
2.0475E+04	4.3572E+04	8.1285E-01
2.0513E+04	4.3575E+04	1.0409E+00
2.0540E+04	4.3503E+04	2.2155E+00
2.0531E+04	4.3504E+04	1.7573E+00
2.0532E+04	4.3544E+04	5.3046E-01
2.0560E+04	4.3552E+04	2.0491E-01
2.0553E+04	4.3541E+04	8.6058E-02
2.0563E+04	4.3584E+04	4.2050E-01
2.0533E+04	4.3571E+04	8.9699E-01
2.0544E+04	4.3572E+04	7.0808E-01
2.0532E+04	4.3590E+04	6.2249E-01
2.0567E+04	4.3605E+04	5.0065E-01
2.0567E+04	4.3578E+04	4.1779E-01
SUM OF SQUARED RESIDUALS		1.2472E+01

Table 9.36a Values of CARBONATE evaluated at the RANDOM CHECK POINTS on the TREND SURFACE over the 50m grid at Yelford.

CARBONATE at CHECK POINTS on BISPLINE SURFACE, 50m grid

INTERIOR X-KNOTS

2.0480E+04
2.0490E+04
2.0500E+04
2.0508E+04
2.0517E+04

INTERIOR Y-KNOTS

4.3536E+04
4.3543E+04
4.3549E+04
4.3556E+04

RANK 63

RANDOM CHECK POINTS

NORTHING	EASTING	SURFACE FIT
2.0462E+04	4.3544E+04	5.9820E-01
2.0461E+04	4.3547E+04	9.9846E-01
2.0454E+04	4.3544E+04	5.9295E-01
2.0468E+04	4.3557E+04	5.2610E-01
2.0468E+04	4.3561E+04	7.4480E-01
2.0458E+04	4.3560E+04	7.6102E-01
2.0487E+04	4.3558E+04	1.2065E+00
2.0491E+04	4.3538E+04	4.4963E-01
2.0494E+04	4.3548E+04	8.1955E-01
2.0502E+04	4.3538E+04	-2.5647E+00
2.0500E+04	4.3542E+04	9.7495E-01
2.0503E+04	4.3545E+04	9.2911E-01
2.0513E+04	4.3542E+04	1.0869E+01
2.0511E+04	4.3547E+04	3.2887E+00
2.0511E+04	4.3560E+04	1.2443E+01
2.0519E+04	4.3560E+04	2.3550E-01

SUM OF SQUARED RESIDUALS 7.3220E+02

Table 9.36c. Values of CARBONATE evaluated at the RANDOM CHECK POINTS on the TREND SURFACE over the 200 m grid at Yelford.

INTERIOR X-KNOTS

2.0410E+04
2.0450E+04
2.0496E+04
2.0534E+04

INTERIOR Y-KNOTS

4.3503E+04
4.3556E+04
4.3585E+04

RANDOM CHECK POINTS

NORTHING	EASTING	SURFACE FIT
2.0397E+04	4.3501E+04	9.7615E-01
2.0382E+04	4.3572E+04	1.2939E+01
2.0372E+04	4.3563E+04	1.1649E+01
2.0415E+04	4.3481E+04	9.1567E-01
2.0418E+04	4.3468E+04	1.1459E+00
2.0446E+04	4.3558E+04	7.6551E-01
2.0461E+04	4.3509E+04	8.7839E-01
2.0481E+04	4.3538E+04	7.5136E-01
2.0494E+04	4.3548E+04	1.2011E+00
2.0491E+04	4.3538E+04	1.2504E+00
2.0462E+04	4.3544E+04	2.5681E-01
2.0461E+04	4.3547E+04	2.9764E-01
2.0454E+04	4.3544E+04	3.3995E-01
2.0466E+04	4.3572E+04	3.1854E+00
2.0491E+04	4.3582E+04	1.6900E+00
2.0475E+04	4.3572E+04	1.2834E+00
2.0455E+04	4.3576E+04	7.0080E+00
2.0487E+04	4.3558E+04	-1.7389E-01
2.0468E+04	4.3557E+04	3.5007E-01
2.0468E+04	4.3561E+04	5.0593E-01
2.0458E+04	4.3560E+04	1.0536E+00
2.0453E+04	4.3572E+04	4.5150E+00
2.0450E+04	4.3588E+04	1.6981E+01
2.0521E+04	4.3495E+04	4.6656E-01
2.0508E+04	4.3491E+04	6.3805E-01
2.0511E+04	4.3547E+04	3.0447E+00
2.0502E+04	4.3512E+04	4.5276E-01
2.0531E+04	4.3504E+04	8.9122E-01
2.0532E+04	4.3544E+04	3.1102E+00
2.0524E+04	4.3539E+04	2.8424E+00
2.0513E+04	4.3542E+04	2.8458E+00
2.0502E+04	4.3538E+04	1.9075E+00
2.0500E+04	4.3542E+04	1.8616E+00
2.0528E+04	4.3524E+04	1.4599E+00
2.0503E+04	4.3545E+04	2.1602E+00
2.0533E+04	4.3571E+04	6.9529E+00
2.0500E+04	4.3579E+04	4.0719E-01
2.0501E+04	4.3559E+04	1.2223E+00
2.0519E+04	4.3560E+04	4.9981E+00
2.0513E+04	4.3575E+04	3.6364E+00
2.0515E+04	4.3567E+04	4.4095E+00
2.0500E+04	4.3563E+04	5.7408E-01
2.0515E+04	4.3561E+04	4.2890E+00
2.0511E+04	4.3560E+04	3.3717E+00
2.0532E+04	4.3590E+04	3.1543E+00
2.0540E+04	4.3503E+04	1.7881E+00
2.0560E+04	4.3552E+04	-1.4746E+00
2.0553E+04	4.3541E+04	1.9929E-01
2.0563E+04	4.3584E+04	-4.3476E-01
2.0544E+04	4.3572E+04	3.6956E+00
2.0567E+04	4.3578E+04	-2.5678E-01
2.0567E+04	4.3605E+04	1.4785E+01

SUM OF SQUARED RESIDUALS 9.0735E+02

Table 9.37a. Values of SILT evaluated on the TREND SURFACE at the RANDOM CHECK POINTS in the 50 m grid at Yelford.

SILT at CHECK POINTS on BISPLINE SURFACE, 50m grid

INTERIOR X-KNOTS

2.0470E+04
2.0476E+04
2.0495E+04
2.0508E+04

INTERIOR Y-KNOTS

4.3539E+04
4.3550E+04
4.3556E+04

RANDOM CHECK POINTS

NORTHING	EASTING	SURFACE FIT
2.0462E+04	4.3544E+04	2.5103E+01
2.0461E+04	4.3547E+04	2.3238E+01
2.0454E+04	4.3544E+04	2.4179E+01
2.0468E+04	4.3557E+04	2.4155E+01
2.0468E+04	4.3561E+04	4.1859E+01
2.0458E+04	4.3560E+04	2.3166E+01
2.0491E+04	4.3538E+04	2.8001E+01
2.0494E+04	4.3548E+04	2.8539E+01
2.0487E+04	4.3558E+04	3.2167E+01
2.0502E+04	4.3538E+04	3.0569E+01
2.0500E+04	4.3542E+04	3.0291E+01
2.0503E+04	4.3545E+04	3.1455E+01
2.0511E+04	4.3547E+04	2.4235E+01
2.0513E+04	4.3542E+04	2.6860E+01
2.0519E+04	4.3560E+04	8.7046E+00
2.0511E+04	4.3560E+04	1.7636E+01
SUM OF SQUARED RESIDUALS		1.0363E+03

Table 9.37b. Values of SILT evaluated on the TREND SURFACE at the RANDOM CHECK POINTS in the 100m grid at Yelford.

INTERIOR X-KNOTS
2.0479E+04
2.0511E+04
2.0551E+04
INTERIOR Y-KNOTS
4.3516E+04
4.3545E+04
4.3580E+04

RANDOM CHECK POINTS

NORTHING	EASTING	SURFACE FIT
2.0461E+04	4.3509E+04	2.2559E+01
2.0462E+04	4.3544E+04	2.6291E+01
2.0454E+04	4.3544E+04	2.6566E+01
2.0466E+04	4.3572E+04	2.2733E+01
2.0455E+04	4.3576E+04	2.1535E+01
2.0468E+04	4.3557E+04	2.6267E+01
2.0468E+04	4.3561E+04	2.5344E+01
2.0461E+04	4.3547E+04	2.5699E+01
2.0458E+04	4.3560E+04	2.1236E+01
2.0453E+04	4.3572E+04	2.0040E+01
2.0450E+04	4.3588E+04	2.9785E+01
2.0502E+04	4.3512E+04	2.7983E+01
2.0508E+04	4.3491E+04	2.6596E+01
2.0502E+04	4.3538E+04	2.7001E+01
2.0500E+04	4.3542E+04	2.7661E+01
2.0491E+04	4.3538E+04	2.8521E+01
2.0500E+04	4.3579E+04	2.9648E+01
2.0494E+04	4.3548E+04	3.0348E+01
2.0487E+04	4.3558E+04	3.2858E+01
2.0500E+04	4.3563E+04	2.9967E+01
2.0503E+04	4.3545E+04	2.7071E+01
2.0491E+04	4.3582E+04	2.8040E+01
2.0540E+04	4.3503E+04	2.6939E+01
2.0531E+04	4.3504E+04	2.6956E+01
2.0521E+04	4.3495E+04	2.6382E+01
2.0532E+04	4.3544E+04	2.5686E+01
2.0524E+04	4.3539E+04	2.6948E+01
2.0513E+04	4.3542E+04	2.5296E+01
2.0528E+04	4.3524E+04	3.0253E+01
2.0533E+04	4.3571E+04	4.3250E+00
2.0511E+04	4.3547E+04	2.4595E+01
2.0544E+04	4.3572E+04	9.2472E+00
2.0519E+04	4.3560E+04	1.6915E+01
2.0513E+04	4.3575E+04	2.2776E+01
2.0515E+04	4.3567E+04	1.9372E+01
2.0511E+04	4.3560E+04	2.3054E+01
2.0532E+04	4.3590E+04	2.1936E+01
2.0553E+04	4.3541E+04	3.3801E+01
2.0560E+04	4.3552E+04	2.9953E+01

SUM OF SQUARED RESIDUALS	3.2759E+03

Table 9.37c. Values of SILT evaluated on the TREND SURFACE at the RANDOM CHECK POINTS in the 200m grid at Yelford.

SILT evaluated at CHECK POINTS on BISPLINE SURFACE 200m grid

INTERIOR X-KNOTS

2.0410E+04
2.0483E+04
2.0557E+04

INTERIOR Y-KNOTS

4.3476E+04
4.3529E+04
4.3582E+04

RANDOM CHECK POINTS

NORTHING	EASTING	SURFACE FIT
2.0397E+04	4.3501E+04	2.2713E+01
2.0382E+04	4.3572E+04	1.9709E+01
2.0372E+04	4.3563E+04	2.0802E+01
2.0418E+04	4.3468E+04	2.3074E+01
2.0415E+04	4.3481E+04	2.2909E+01
2.0461E+04	4.3509E+04	2.5023E+01
2.0466E+04	4.3572E+04	2.8315E+01
2.0446E+04	4.3558E+04	2.6612E+01
2.0475E+04	4.3572E+04	2.9507E+01
2.0455E+04	4.3576E+04	2.4576E+01
2.0481E+04	4.3538E+04	2.8611E+01
2.0468E+04	4.3557E+04	2.9977E+01
2.0468E+04	4.3561E+04	2.9934E+01
2.0462E+04	4.3544E+04	2.8647E+01
2.0461E+04	4.3547E+04	2.8834E+01
2.0458E+04	4.3560E+04	2.8649E+01
2.0453E+04	4.3572E+04	2.5435E+01
2.0454E+04	4.3544E+04	2.7857E+01
2.0450E+04	4.3588E+04	2.0006E+01
2.0540E+04	4.3503E+04	2.3265E+01
2.0502E+04	4.3512E+04	2.5418E+01
2.0531E+04	4.3504E+04	2.3735E+01
2.0521E+04	4.3495E+04	2.5006E+01
2.0508E+04	4.3491E+04	2.6544E+01
2.0528E+04	4.3524E+04	2.8442E+01
2.0533E+04	4.3571E+04	1.9938E+01
2.0511E+04	4.3547E+04	2.6524E+01
2.0544E+04	4.3572E+04	1.9298E+01
2.0532E+04	4.3544E+04	2.7472E+01
2.0524E+04	4.3539E+04	2.7760E+01
2.0500E+04	4.3579E+04	2.7721E+01
2.0513E+04	4.3542E+04	2.7030E+01
2.0501E+04	4.3559E+04	2.7245E+01
2.0519E+04	4.3560E+04	2.3348E+01
2.0502E+04	4.3538E+04	2.7749E+01
2.0500E+04	4.3542E+04	2.7871E+01
2.0494E+04	4.3548E+04	2.8566E+01
2.0491E+04	4.3538E+04	2.8309E+01
2.0487E+04	4.3558E+04	2.9758E+01
2.0513E+04	4.3575E+04	2.4493E+01
2.0553E+04	4.3541E+04	3.5998E+01
2.0515E+04	4.3567E+04	2.3555E+01
2.0500E+04	4.3563E+04	2.7369E+01
2.0515E+04	4.3561E+04	2.3984E+01
2.0511E+04	4.3560E+04	2.4948E+01
2.0503E+04	4.3545E+04	2.7520E+01
2.0532E+04	4.3590E+04	2.6753E+01
2.0491E+04	4.3582E+04	2.8528E+01
2.0567E+04	4.3578E+04	2.4929E+01
2.0560E+04	4.3552E+04	3.5267E+01
2.0563E+04	4.3584E+04	1.9144E+01
2.0567E+04	4.3605E+04	1.2189E+01

SUM OF SQUARED RESIDUALS 2.2935E+03

Table 9.38a. Values of COARSE SAND evaluated on the TREND SURFACE at the RANDOM CHECK POINTS in the 50m grid at Yelford.

COARSE SAND at CHECK POINTS on BISPLINE SURFACE, 50m grid

INTERIOR X-KNOTS

2.0470E+04
2.0495E+04
2.0508E+04
2.0518E+04

INTERIOR Y-KNOTS

4.3535E+04
4.3543E+04
4.3556E+04

RANDOM CHECK POINTS

NORTHING	EASTING	SURFACE FIT
2.0462E+04	4.3544E+04	1.9923E+00
2.0461E+04	4.3547E+04	1.9608E+00
2.0454E+04	4.3544E+04	1.1793E+00
2.0468E+04	4.3557E+04	3.9300E+00
2.0468E+04	4.3561E+04	2.5205E+00
2.0458E+04	4.3560E+04	2.7676E+00
2.0491E+04	4.3538E+04	8.1740E+00
2.0494E+04	4.3548E+04	4.7864E+00
2.0487E+04	4.3558E+04	3.7779E+00
2.0502E+04	4.3538E+04	5.7755E+00
2.0500E+04	4.3542E+04	5.4484E+00
2.0503E+04	4.3545E+04	4.4737E+00
2.0513E+04	4.3542E+04	9.3039E+00
2.0511E+04	4.3547E+04	9.8710E+00
2.0511E+04	4.3560E+04	3.1835E+01
2.0519E+04	4.3560E+04	-3.7299E+00
SUM OF SQUARED RESIDUALS		5.9486E+03

Table 9.38 b Values of COARSE SAND evaluated on the TREND SURFACE at the RANDOM CHECK POINTS in the 100m grid at Yelford.

INTERIOR X-KNOTS
2.0459E+04
2.0486E+04
2.0523E+04
2.0548E+04
0INTERIOR Y-KNOTS
4.3490E+04
4.3516E+04
4.3553E+04

RANDOM CHECK POINTS

NORTHING	EASTING	SURFACE FIT
2.0454E+04	4.3544E+04	-4.3296E+00
2.0455E+04	4.3576E+04	3.2785E+00
2.0458E+04	4.3560E+04	-1.3724E+00
2.0453E+04	4.3572E+04	8.5593E-01
2.0450E+04	4.3588E+04	5.7337E+00
2.0461E+04	4.3509E+04	4.7964E+00
2.0462E+04	4.3544E+04	2.3357E+00
2.0461E+04	4.3547E+04	1.2049E+00
2.0466E+04	4.3572E+04	8.0927E+00
2.0468E+04	4.3557E+04	5.4444E+00
2.0468E+04	4.3561E+04	5.8662E+00
2.0502E+04	4.3512E+04	4.7311E+00
2.0521E+04	4.3495E+04	6.5970E+00
2.0508E+04	4.3491E+04	3.3186E+00
2.0511E+04	4.3547E+04	1.6477E+01
2.0513E+04	4.3542E+04	1.4685E+01
2.0502E+04	4.3538E+04	8.2754E+00
2.0500E+04	4.3542E+04	8.1266E+00
2.0494E+04	4.3548E+04	4.7928E+00
2.0491E+04	4.3538E+04	4.4792E+00
2.0503E+04	4.3545E+04	1.0475E+01
2.0500E+04	4.3579E+04	4.1265E+00
2.0491E+04	4.3582E+04	4.2037E+00
2.0519E+04	4.3560E+04	2.7242E+01
2.0487E+04	4.3558E+04	1.3164E+00
2.0513E+04	4.3575E+04	1.6014E+01
2.0515E+04	4.3567E+04	2.2373E+01
2.0500E+04	4.3563E+04	7.3086E+00
2.0511E+04	4.3560E+04	1.8920E+01
2.0540E+04	4.3503E+04	6.4839E+00
2.0531E+04	4.3504E+04	6.0440E+00
2.0532E+04	4.3544E+04	1.4371E+01
2.0524E+04	4.3539E+04	1.3365E+01
2.0528E+04	4.3524E+04	4.5565E+00
2.0533E+04	4.3571E+04	4.9139E+01
2.0544E+04	4.3572E+04	3.9972E+01
2.0532E+04	4.3590E+04	9.7784E+00
2.0560E+04	4.3552E+04	3.3063E+00
2.0553E+04	4.3541E+04	5.6195E+00

SUM OF SQUARED RESIDUALS 4.8539E+03

Table 9.38c Values of COARSE SAND evaluated on the TREND SURFACE at the RANDOM CHECK POINTS in the 200m grid at Yelford.

SAND evaluated at CHECK POINTS on BISPLINE SURFACE 200m grid

INTERIOR X-KNOTS

2.0387E+04
2.0459E+04
2.0496E+04
2.0532E+04

INTERIOR Y-KNOTS

4.3503E+04
4.3556E+04
4.3583E+04

RANDOM CHECK POINTS

NORTHING	EASTING	SURFACE FIT
2.0382E+04	4.3572E+04	1.4297E+01
2.0372E+04	4.3563E+04	2.0538E+01
2.0415E+04	4.3481E+04	2.5211E+00
2.0397E+04	4.3501E+04	1.7413E+00
2.0418E+04	4.3468E+04	2.8109E+00
2.0454E+04	4.3544E+04	2.2551E+00
2.0446E+04	4.3558E+04	5.3378E+00
2.0455E+04	4.3576E+04	2.8220E+00
2.0458E+04	4.3560E+04	5.6926E+00
2.0453E+04	4.3572E+04	6.2616E+00
2.0450E+04	4.3588E+04	-1.0464E+01
2.0461E+04	4.3509E+04	2.9037E+00
2.0481E+04	4.3538E+04	5.0531E+00
2.0494E+04	4.3548E+04	5.0832E+00
2.0491E+04	4.3538E+04	6.5673E+00
2.0462E+04	4.3544E+04	2.6863E+00
2.0461E+04	4.3547E+04	2.8821E+00
2.0466E+04	4.3572E+04	1.9880E+00
2.0491E+04	4.3582E+04	-4.8206E+00
2.0475E+04	4.3572E+04	-2.9555E+00
2.0487E+04	4.3558E+04	-1.0280E-01
2.0468E+04	4.3557E+04	3.0127E+00
2.0468E+04	4.3561E+04	3.0294E+00
2.0521E+04	4.3495E+04	6.4336E+00
2.0508E+04	4.3491E+04	4.2287E+00
2.0511E+04	4.3547E+04	1.0262E+01
2.0502E+04	4.3512E+04	6.6748E+00
2.0531E+04	4.3504E+04	6.9569E+00
2.0524E+04	4.3539E+04	9.3559E+00
2.0513E+04	4.3542E+04	9.6956E+00
2.0502E+04	4.3538E+04	8.0565E+00
2.0500E+04	4.3542E+04	7.5037E+00
2.0528E+04	4.3524E+04	6.5459E+00
2.0503E+04	4.3545E+04	7.9096E+00
2.0500E+04	4.3579E+04	9.7046E-01
2.0501E+04	4.3559E+04	4.9271E+00
2.0519E+04	4.3560E+04	2.0899E+01
2.0513E+04	4.3575E+04	1.9377E+01
2.0515E+04	4.3567E+04	2.1328E+01
2.0500E+04	4.3563E+04	3.0310E+00
2.0515E+04	4.3561E+04	1.8213E+01
2.0511E+04	4.3560E+04	1.3891E+01
2.0540E+04	4.3503E+04	5.7608E+00
2.0532E+04	4.3544E+04	1.0799E+01
2.0560E+04	4.3552E+04	2.7103E+00
2.0553E+04	4.3541E+04	4.4561E+00
2.0533E+04	4.3571E+04	3.3767E+01
2.0544E+04	4.3572E+04	2.0902E+01
2.0567E+04	4.3578E+04	7.7349E-01
2.0563E+04	4.3584E+04	3.3544E+00
2.0532E+04	4.3590E+04	2.1767E+00
2.0567E+04	4.3605E+04	2.1741E+01
SUM OF SQUARED RESIDUALS		5.8719E+03

Table 9.39a Values of soil properties at the random check points in the 50 m grid predicted by Kriging of de-trended data after trend modelling by bicubic spline surfaces

Northing	Easting	K	Carbonate	Silt	Coarse Sand
20511	43547	0.8528	3.3195	24.2353	9.8710
20513	43542	0.7208	10.8998	26.8603	9.3039
20519	43560	0.2042	0.2663	8.7049	-3.7299
20502	43538	0.3960	-2.5339	30.5693	5.7755
20500	43542	0.4905	1.0058	30.2913	5.4484
20494	43548	0.8373	0.8504	28.5393	4.7864
20491	43538	0.3628	0.4804	28.0013	8.1740
20487	43558	0.4837	1.2373	32.1673	3.7779
20468	43557	0.4783	0.5569	24.1553	3.9300
20468	43561	0.6703	0.7756	41.8593	2.5205
20462	43544	0.5536	0.6290	25.1033	1.9923
20461	43547	0.5328	1.0293	23.2383	1.9608
20458	43560	0.7165	0.7918	23.1663	2.7676
20454	43544	0.4568	0.6238	24.1793	1.1793
20511	43560	1.0644	12.4738	17.6363	31.8350
20503	43545	0.7663	0.9599	31.4553	4.4737

Table 9.39b Values of soil properties at the random check points in the 100 m grid predicted by Kriging of de-trended data after trend modelling by bicubic spline surfaces

Northing	Easting	K	Carbonate	Silt	Coarse Sand
20540	43503	0.9003	*	29.6997	6.4839
20533	43571	0.5272		5.2710	49.1390
20511	43547	0.4965		27.6780	16.4770
20502	43512	0.5706		30.1980	4.7311
20466	43572	0.4445		23.4665	8.0927
20544	43572	0.5535		9.5889	39.9720
20531	43504	1.0673		29.6156	6.0440
20521	43495	0.9217		29.2169	6.5970
20532	43544	0.6807		25.9925	14.3710
20532	43590	0.9498		23.2924	9.7784
20524	43539	0.8304		28.6540	13.3650
20508	43491	1.2137		29.3662	3.3186
20500	43579	0.5941		30.6885	4.1265
20491	43582	1.3565		29.1948	4.2037
20461	43509	0.3152		22.7583	4.7964
20455	43576	0.2676		21.8143	3.2785
20513	43542	0.6018		27.9107	14.6850
20519	43560	1.0028		21.1651	27.2420
20502	43538	0.5131		29.2040	8.2754
20500	43542	0.4981		29.4551	8.1266
20494	43548	0.4709		31.3940	4.7928
20491	43538	0.5092		29.8447	4.4792
20487	43558	0.5776		34.6870	1.3164
20468	43557	0.0208		26.9880	5.4444
20468	43561	0.0527		26.0030	5.8662
20462	43544	0.3648		27.2793	2.3357
20461	43547	0.3852		27.0768	1.2049
20458	43560	0.9142		22.0813	-1.3724
20513	43575	0.8039		24.2659	16.0140
20453	43572	0.7130		20.1379	0.8559
20560	43552	0.8362		31.5459	3.3063
20450	43588	0.3930		30.2441	5.7337
20528	43524	0.9015		34.5296	4.5565
20553	43541	0.6229		35.9951	5.6195
20454	43544	0.5224		27.4903	-4.3296
20515	43567	1.0885		21.9390	22.3730
20500	43563	1.0636		31.0208	7.3086
20511	43560	1.1428		26.3813	18.9200
20503	43545	0.4460		28.9617	10.4750

* No estimates were obtained for this property at this sampling scale

Table 9.39c Values of soil properties at the random check points in the 200 m grid predicted by Kriging of de-trended data after trend modelling by bicubic spline surfaces

Northing	Easting	K	Carbonate	Silt	Coarse Sand
20563	43584	0.4248	-0.4348	19.1392	3.5971
20540	43503	2.2173	1.7881	23.2602	6.2259
20533	43571	1.0270	6.9529	19.9332	36.2572
20511	43547	0.6432	3.0447	26.5192	13.1152
20502	43512	0.6115	0.4528	25.4132	7.9528
20466	43572	1.0588	3.1854	28.3102	5.4686
20446	43558	0.1720	0.7655	26.6072	7.5114
20415	43481	0.7188	0.9157	22.9042	2.5067
20397	43501	1.0904	0.9762	22.7082	1.7708
20382	43572	0.6858	12.9390	19.7042	15.2074
20372	43563	0.1018	11.6490	20.7972	21.2335
20544	43572	0.7585	3.6956	19.2932	22.5665
20531	43504	1.7703	0.8912	23.7302	7.7445
20521	43495	1.1444	0.4666	25.0012	7.2086
20532	43544	0.5367	3.1102	27.4672	11.1511
20532	43590	0.5832	3.1543	26.7482	3.6215
20524	43539	0.6155	2.8424	27.7552	10.7460
20508	43491	0.7465	0.6381	26.5392	5.3344
20500	43579	1.1997	0.4072	27.7162	6.2557
20491	43582	1.4879	1.6900	28.5232	-0.3262
20461	43509	0.1201	0.8784	25.0182	3.2039
20475	43572	1.1456	1.2834	29.5022	0.6946
20455	43576	1.0846	7.0080	24.5712	6.5681
20513	43542	0.6603	2.8458	27.0252	12.2268
20501	43559	0.7974	1.2223	27.2402	9.2734
20481	43538	0.3683	0.7514	28.6062	5.5931
20519	43560	0.9195	4.9981	23.3432	24.7701
20502	43538	0.6194	1.9075	27.7442	10.1620
20500	43542	0.6086	1.8616	27.8662	9.4025
20494	43548	0.5975	1.2011	28.5612	6.4518
20491	43538	0.5220	1.2504	28.3042	7.3943
20487	43558	0.6872	-0.1739	29.7532	2.4306
20468	43557	0.4266	0.3501	29.9722	5.3262
20468	43561	0.5672	0.5059	29.9292	5.7608
20462	43544	0.1464	0.2568	28.6422	3.7369
20461	43547	0.1527	0.2976	28.8292	4.1231
20458	43560	0.3607	1.0536	28.6442	8.4481
20418	43468	0.5384	1.1459	23.0692	2.6965
20567	43605	0.5062	14.7850	12.1842	21.7328
20567	43578	0.4163	-0.2568	24.9242	0.8734
20513	43575	1.0090	3.6364	24.4882	25.5571
20453	43572	0.7946	4.5150	25.4302	10.2430
20560	43552	0.1811	-1.4746	35.2622	2.7086
20450	43588	1.6530	16.9810	20.0012	-7.8577
20528	43524	0.8110	1.4599	28.4372	7.6219
20553	43541	0.0666	0.1993	35.9932	4.5139
20454	43544	0.1051	0.3400	27.8522	3.2176
20515	43567	0.9913	4.4095	23.5502	27.1865
20500	43563	0.8692	0.5741	27.3642	8.1079
20515	43561	0.9053	4.2890	23.9792	22.9647
20511	43560	0.8524	3.3717	24.9432	18.8496
20503	43545	0.6185	2.1602	27.5152	10.2683

9.20.2 Predictions by Kriging of de-trended data from stratification based on soil mapping units.

Having partitioned trends by grouping the data within mapping units into larger strata, and having calculated the within-stratum semi-variograms (figs. 9.39a to 9.46c) the information about the parameters of these semi-variograms, in table 9.31 were used to compute the interpolated values by Kriging at the random check points in each sampling grid.

First, it was ascertained to what mapping unit stratum the individual check points pertained and then they were allocated into such stratum. Then, the predicted values of soil properties by Kriging interpolation were calculated. The Kriging interpolations were computed using the same procedures already described in chapter 7. Tables 9.40a, b and c show the predicted values of the four soil properties at the check points in the three sampling grids. They also include the mapping unit stratum corresponding to each point.

After the partitioning, for some strata the semi-variogram showed total absence of spatial dependence. These were considered as cases of total random variation within the stratum and the predicted values at the check points within that stratum were given by its mean value.

From observation of tables 9.40a-c one fact is immediately evident. No within-stratum semi-variogram showed anisotropic variation, particularly at the scale corresponding to the 50 m grid, as it was the case with the original semi-variograms of these four properties at the same scale, when trends were ignored. This allowed for ordinary Kriging interpolation which was not possible before the trends were partitioned, but only by having to resort to the complications of detailed structural analysis and Universal Kriging. It is in this sense that the removal of the trend by stratification based on mapping units was successful.

At the 50 m grid Carbonate and Silt sowed spatial structure in the calcareous soils, but Carbonate showed no spatial structure in the non-calcareous soils, as it could have been

Table 9.40a Values of soil properties predicted at the random check points by Kriging of de-trended data based on stratification by soil mapping units in the 50 m grid at Yelford.

NORTHING	EASTING	SITE No	STRATUM	K	CARBONATE	SILT	SAND
20511	43547	95004	BC-BE	* 0.650	3.624	27.15	* 10.416
20513	43542	95024	BC-BE	* 0.650	2.150	27.01	* 10.416
20519	43560	95027	BC-BE	* 0.650	10.104	10.10	* 10.416
20502	43538	95028	NS-NG	0.519	* 0.640	29.52	4.845
20500	43542	95029	NS-NG	0.525	* 0.640	28.78	5.874
20494	43548	95030	NS-NG	0.641	* 0.640	27.15	5.471
20491	43538	95031	BC-BE	* 0.650	3.640	29.33	* 10.416
20487	43558	95032	BC-BE	* 0.650	1.239	29.66	* 10.416
20468	43557	95033	NS-NG	0.465	* 0.640	29.88	4.027
20468	43561	95034	NS-NG	0.505	* 0.640	32.85	3.022
20462	43544	95035	NS-NG	0.502	* 0.640	25.14	1.926
20461	43547	95036	NS-NG	0.511	* 0.640	23.93	1.841
20458	43560	95037	NS-NG	0.626	* 0.640	22.35	2.662
20454	43544	95047	NS-NG	0.432	* 0.640	24.62	1.083
20511	43560	95051	BC-BE	* 0.650	11.270	18.46	* 10.416
20503	43545	95052	BC-BE	* 0.650	1.938	29.40	* 10.416

* Values predicted by the mean value within the stratum due to total random variation (Pure Nugget Effect) as shown by the semi-variogram

Table 9.40b. Values of soil properties predicted at the random check points by Kriging of de-trended data based on stratification by soil mapping units in the 100 m grid at Yelford.

NORTHING	EASTING	SITE No	STRATUM	K	CARBONATE	SILT	SAND
20540	43503	95002	BC-BE	0.622	0.776	27.20	5.66
20533	43571	95003	BC-BE	0.455	8.568	9.65	46.78
20511	43547	95004	BC-BE	1.119	4.786	24.95	12.75
20502	43512	95005	BC-BE	0.517	0.695	27.89	6.56
20466	43572	95006	BC-BE	0.406	1.466	24.17	5.38
20544	43572	95012	BC-BE	0.645	5.310	13.46	31.48
20531	43504	95013	BC-BE	0.966	0.979	25.90	5.64
20521	43495	95014	BC-BE	1.370	1.258	24.23	4.16
20532	43544	95015	BC-BE	0.599	5.289	21.27	17.96
20532	43590	95016	BC-BE	0.796	5.060	22.14	8.49
20524	43539	95017	BC-BE	0.647	3.754	23.64	11.82
20508	43491	95018	BC-BE	0.492	0.634	28.03	2.90
20500	43579	95019	NS-NG	* 0.497	* 0.728	28.68	* 4.14
20491	43582	95020	NS-NG	* 0.497	* 0.728	28.77	* 4.14
20461	43509	95021	NS-NG	* 0.497	* 0.728	23.15	* 4.14
20455	43576	95023	BC-BE	0.298	0.903	24.48	5.95
20513	43542	95024	BC-BE	0.869	3.752	24.61	10.95
20519	43560	95027	BC-BE	0.728	10.246	13.37	38.78
20502	43538	95028	BC-BE	0.450	1.066	27.15	6.23
20500	43542	95029	BC-BE	0.554	1.102	28.59	6.28
20494	43548	95030	BC-BE	0.678	1.060	30.91	4.20
20491	43538	95031	BC-BE	0.213	0.369	27.88	7.33
20487	43558	95032	BC-BE	0.405	1.634	30.63	4.48
20468	43557	95033	NS-NG	* 0.497	* 0.728	26.41	* 4.14
20468	43561	95034	NS-NG	* 0.497	* 0.728	26.05	* 4.14
20462	43544	95035	NS-NG	* 0.497	* 0.728	25.98	* 4.14
20461	43547	95036	NS-NG	* 0.497	* 0.728	24.57	* 4.14
20458	43560	95037	NS-NG	* 0.497	* 0.728	22.69	* 4.14
20513	43575	95041	BC-BE	0.841	4.169	24.00	12.33
20453	43572	95042	BC-BE	0.368	0.646	24.43	5.41
20560	43552	95043	NS-NG	* 0.497	* 0.728	28.39	* 4.14
20450	43588	95044	NS-NG	* 0.497	* 0.728	26.45	* 4.14
20528	43524	95045	NS-NG	* 0.497	* 0.728	32.37	* 4.14
20553	43541	95046	NS-NG	* 0.497	* 0.728	31.23	* 4.14
20454	43544	95047	NS-NG	* 0.497	* 0.728	26.21	* 4.14
20515	43567	95048	BC-BE	0.727	6.999	19.08	25.97
20500	43563	95049	BC-BE	0.932	2.233	29.48	5.14
20511	43560	95051	BC-BE	0.823	8.710	20.91	27.59
20503	43545	95052	BC-BE	0.871	2.179	28.02	6.61

* Values predicted by mean value within the stratum due to random variation (Pure Nugget Effect) as indicated by the within-stratum semi-variogram.

Table 9.40c. Values of soil properties predicted at the random check points by Kriging of de-trended data based on stratification by soil mapping units in the 200 m grid at Yelford.

NORTHING	EASTING	SITE No	STRATUM	K	CARBONATE	SILT	SAND
20563	43584	95001	BC-BE	0.402	4.880	* 23.64	12.69
20540	43503	95002	BC-BE	1.233	1.994	* 23.64	8.78
20533	43571	95003	BC-BE	0.518	3.351	* 23.64	11.19
20511	43547	95004	BC-BE	1.168	1.932	* 23.64	8.30
20502	43512	95005	BC-BE	0.622	1.622	* 23.64	6.17
20466	43572	95006	NS-NG	* 0.525	* 0.770	* 24.29	4.33
20446	43558	95007	NS-NG	* 0.525	* 0.770	* 24.29	2.28
20415	43418	95008	NS-NG	* 0.525	* 0.770	* 24.29	2.44
20397	43501	95009	NS-NG	* 0.525	* 0.770	* 24.29	2.07
20382	43572	95010	BC-BE	0.388	6.385	* 23.64	18.44
20372	43563	95011	BC-BE	0.511	6.418	* 23.64	18.72
20544	43572	95012	BC-BE	0.562	3.590	* 23.64	11.48
20531	43504	95013	BC-BE	1.323	1.982	* 23.64	8.54
20521	43495	95014	BC-BE	0.797	2.098	* 23.64	8.22
20532	43544	95015	BC-BE	0.564	1.819	* 23.64	8.87
20532	43590	95016	BC-BE	0.757	5.293	* 23.64	13.09
20524	43539	95017	NS-NG	* 0.525	* 0.770	* 24.29	9.01
20508	43491	95018	BC-BE	0.429	1.796	* 23.64	6.82
20500	43579	95019	BC-BE	1.477	4.072	* 23.64	10.71
20491	43582	95020	BC-BE	1.601	4.306	* 23.64	10.47
20461	43509	95021	NS-NG	* 0.525	* 0.770	* 24.29	2.25
20475	43572	95022	NS-NG	* 0.525	* 0.770	* 24.29	4.55
20455	43576	95023	NS-NG	* 0.525	* 0.770	* 24.29	4.15
20513	43542	95024	BC-BE	0.993	1.858	* 23.64	8.23
20501	43559	95025	BC-BE	1.070	2.729	* 23.64	9.19
20481	43538	95026	NS-NG	* 0.525	* 0.770	* 24.29	3.62
20519	43560	95027	NS-NG	* 0.525	* 0.770	* 24.29	12.89
20502	43538	95028	BC-BE	0.750	1.635	* 23.64	7.80
20500	43542	95029	NS-NG	* 0.525	* 0.750	* 24.29	4.08
20494	43548	95030	NS-NG	* 0.525	* 0.750	* 24.29	4.19
20491	43538	95031	NS-NG	* 0.525	* 0.750	* 24.29	3.82
20487	43558	95032	NS-NG	* 0.525	* 0.750	* 24.29	4.38
20468	43557	95033	NS-NG	* 0.525	* 0.750	* 24.29	3.81
20468	43561	95034	NS-NG	* 0.525	* 0.750	* 24.29	3.94
20462	43544	95035	NS-NG	* 0.525	* 0.750	* 24.29	2.80
20461	43547	95036	NS-NG	* 0.525	* 0.750	* 24.29	2.86
20458	43560	95037	NS-NG	* 0.525	* 0.750	* 24.29	3.26
20418	43468	95038	NS-NG	* 0.525	* 0.750	* 24.29	2.77
20567	43605	95039	BC-BE	0.542	6.797	* 23.64	14.32
20567	43578	95040	NS-NG	* 0.525	* 0.750	* 24.29	8.98
20513	43575	95041	NS-NG	* 0.525	* 0.750	* 24.29	10.01
20453	43572	95042	NS-NG	* 0.525	* 0.750	* 24.29	3.65
20560	43552	95043	NS-NG	* 0.525	* 0.750	* 24.29	7.40
20560	43552	95043	NS-NG	* 0.525	* 0.750	* 24.29	5.04
20528	43524	95045	BC-BE	0.887	1.514	* 23.64	8.25
20553	43541	95046	BC-BE	0.344	1.572	* 23.64	8.72
20454	43544	95047	NS-NG	* 0.525	* 0.750	* 24.29	2.17
20515	43567	95048	NS-NG	* 0.525	* 0.750	* 24.29	10.89
20500	43563	95049	BC-BE	1.141	2.921	* 23.64	9.45
20515	43561	95050	NS-NG	* 0.525	* 0.750	* 24.29	11.21
20511	43560	95051	BC-BE	0.999	2.654	* 23.64	9.57
20503	43545	95052	BC-BE	0.953	1.819	* 23.64	7.89

* Values predicted by the mean value within the stratum due to random variation (Pure Nugget Effect) as shown by the within-stratum semi-variogram

expected for variation of this property in these soils is random. The same is true at the 200 m grid for Carbonate and Potassium which in consequence were predicted by means. Silt showed no structure in nither calcareous or non-calcareous soils at this scale, and Coarse Sand did the oposite allowing for Kriging in both type of soils.

At the 100 m grid all properties but Silt showed random variation in the non-calcareous soils after de-trending, and therefore were predicted by the stratum means. So these properties in the calcareous soils still showed spatial structure after trend partitioning. This makes sense to the landscape since variation of calcareous soils, at this scale, is reflected in rapid changes with distance.

Although predictions by Kriging were possible by partitioning the trends, however, it was not certain how accurate were those predictions and how did they compare with predictions by other methods for removing trends, i.e. spline surfaces, and with predictions from the products of soil survey. They needed to be compared.

9.21 Comparison of predictive procedures when trends are accounted for.

Ignoring the trends present in the area of Yelford had a definite effect on the predictive success by Kriging and its relative merits as compared to those of soil classes and mapping units. So, to find out the effect that accounting for trends had on the relative predictive success of Kriging with respect to that of the conventional soil survey products, a comparison similar to that described in section 9.16 was carried out, but this time accounting for the trends by the two methods described above.

The criterion of success was again the Residual Sum of Squares (RSS) from the observed values at all check points within the corresponding sampling grid.

Since trends are common-place in soil spatial variability, the comparison was also aimed at discovering what method for

trend removal, i.e. modelling by spline surfaces or partitioning by stratification, would be the best and in which circumstances.

9.21.1 Comparison of predictive procedures when trends are accounted for by modelling with bicubic spline surfaces.

The values of soil properties predicted by Kriging of de-trended data after modelling of the trend by bicubic spline surfaces were compared against the predicted values by conventional soil survey output at the random check points in the three sampling grids. The results of the comparison are shown in tables 9.41a-c for Potassium, 9.42a and c for Carbonate (table 9.42b for the 100 m grid was not calculated), 9.43a-c for Silt and 9.44a-c for Coarse Sand.

The results shown in tables 9.41 to 9.44 indicate that Kriging was not the best predictive procedure after trends were accounted for by spline surfaces. The best predictions were made, in most of the cases by soil class means. Kriging came second.

Kriging made the best predictions in three of the eleven comparisons involving four properties and three sampling scales (tables 9.41a, 9.43a and 9.43b), and came second after class means in four of them (tables 9.42c, 9.43c, 9.44b and 9.44c). This accounts for rates of predictive success of 27.2% as best predictive technique and 36.3% as second best.

On the other hand, class means made the best predictions in seven of the eleven comparisons. This means a rate of predictive success of 63.6%. Class means came second best in three of the eleven comparisons (27.2%).

Kriging was at its best in predicting values of Silt (50m and 100m grids, tables 9.43a and b), and Potassium (50m grid, table 9.41a).

The success in prediction by Kriging and by maps and classes can only be vaguely related, in this type of analysis, to the structures of the semi-variograms (figs. 9.35a-c to 9.38a-c), for these are the semi-variograms of residuals from trend, and not of the raw data. The structures they depict do

Table 9.41a Comparison of predicted values of Potassium at the random check points in the 50 m grid, by conventional soil survey output and by Kriging of de-trended data after trend modelling by bicubic spline surfaces.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING
20511	43547	95004	1.00	1.15	0.86	0.60	0.85
20513	43542	95024	0.70	0.38	0.86	0.60	0.72
20519	43560	95027	0.70	0.92	0.92	0.70	0.20
20502	43538	95028	0.60	0.59	0.55	0.40	0.40
20500	43542	95029	0.50	0.38	0.55	0.40	0.49
20494	43548	95030	1.10	0.59	0.55	0.40	0.84
20491	43538	95031	0.30	0.50	0.55	0.90	0.36
20487	43558	95032	0.30	0.50	0.55	0.90	0.48
20468	43557	95033	0.30	0.77	0.40	0.40	0.48
20468	43561	95034	0.50	0.37	0.40	0.40	0.67
20462	43544	95035	0.40	0.37	0.56	0.50	0.55
20461	43547	95036	0.40	0.37	0.56	0.50	0.53
20458	43560	95037	0.50	0.37	0.56	0.50	0.72
20454	43544	95047	0.60	0.37	0.40	0.40	0.46
20511	43560	95051	1.30	0.92	0.92	0.70	1.06
20503	43545	95052	1.10	0.77	0.86	0.60	0.77
Error in prediction: SUMS OF SQUARES =				1.0847	0.8429	2.1200	0.7526

Table 9.41b Comparison of predicted values of Potassium at the random check points in the 100 m grid, by soil survey output and by Kriging of de-trended data after trend modelling by bicubic spline surfaces.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING
20540	43503	95002	0.90	0.48	0.60	0.45	0.90
20533	43571	95003	0.50	0.59	0.58	0.70	0.53
20511	43547	95004	1.00	1.09	0.90	1.00	0.50
20502	43512	95005	0.40	0.39	0.60	0.45	0.57
20466	43572	95006	0.30	0.38	0.37	0.40	0.44
20544	43572	95012	0.40	0.73	0.60	0.45	0.55
20531	43504	95013	1.20	0.73	0.60	0.45	1.07
20521	43495	95014	0.60	0.73	0.60	0.45	0.92
20532	43544	95015	0.50	0.73	0.90	1.00	0.68
20532	43590	95016	1.00	1.09	0.90	1.00	0.95
20524	43539	95017	0.80	1.09	0.90	1.00	0.83
20508	43491	95018	0.20	0.73	0.60	0.45	1.21
20500	43579	95019	1.50	0.68	0.69	0.70	0.59
20491	43582	95020	0.80	0.68	0.69	0.70	1.36
20461	43509	95021	0.20	0.48	0.32	0.20	0.32
20475	43572	95022	0.20	0.38	0.37	0.40	0.27
20455	43576	95023	0.50	0.48	0.90	1.00	0.60
20513	43542	95024	0.70	0.59	0.58	0.70	1.00
20519	43560	95027	0.70	0.68	0.60	0.45	0.51
20502	43538	95028	0.60	0.48	0.60	0.45	0.50
20500	43542	95029	0.50	0.68	0.60	0.45	0.47
20494	43548	95030	1.10	0.73	0.60	0.45	0.51
20491	43538	95031	0.30	0.73	0.60	0.45	0.58
20487	43558	95032	0.30	0.38	0.32	0.20	0.02
20468	43557	95033	0.30	0.39	0.32	0.20	0.05
20468	43561	95034	0.50	0.39	0.32	0.20	0.36
20462	43544	95035	0.40	0.39	0.32	0.20	0.39
20461	43547	95036	0.40	0.39	0.55	0.60	0.91
20458	43560	95037	0.50	0.73	0.95	0.80	0.80
20513	43575	95041	1.10	0.38	0.37	0.40	0.71
20453	43572	95042	0.40	0.68	0.69	0.70	0.84
20560	43552	95043	0.60	0.73	0.55	0.60	0.39
20450	43588	95044	0.30	0.39	0.32	0.20	0.90
20528	43524	95045	1.00	0.68	0.69	0.70	0.62
20553	43541	95046	0.40	0.39	0.55	0.60	0.52
20454	43544	95047	0.60	1.09	0.58	0.70	1.09
20515	43567	95048	0.70	1.09	0.95	0.80	1.06
20500	43563	95049	0.80	0.59	0.58	0.70	1.14
20511	43560	95051	1.30	0.73	0.90	1.00	0.45
Error in prediction:							
		SUMS OF SQUARES = (from the observed)		3.6981	3.4116	3.7875	5.8690

Table 9.41c Comparison of predicted values of Potassium at the random check points in the 200 m grid, by soil survey output and by Kriging of de-trended data after trend modelling by bicubic spline surfaces.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING
20563	43584	95001	0.60	0.50	0.71	1.00	0.42
20540	43503	95002	0.90	0.50	0.70	0.35	2.22
20533	43571	95003	0.50	0.64	0.71	1.00	1.03
20511	43547	95004	1.00	0.67	0.70	0.35	0.64
20502	43512	95005	0.40	0.41	0.70	0.35	0.61
20466	43572	95006	0.30	0.38	0.64	0.40	1.06
20446	43558	95007	0.60	0.64	0.64	0.40	0.17
20415	43481	95008	0.40	0.41	0.64	0.40	0.72
20397	43501	95009	0.30	0.41	0.42	0.25	1.09
20382	43572	95010	0.40	0.66	0.71	0.70	0.69
20372	43563	95011	0.50	0.67	0.64	0.50	0.10
20544	43572	95012	0.40	0.76	0.70	0.35	0.76
20531	43504	95013	1.20	0.76	0.70	0.35	1.77
20521	43495	95014	0.60	0.76	0.70	0.35	1.14
20532	43544	95015	0.50	0.76	0.71	1.00	0.54
20532	43590	95016	1.00	0.67	0.71	1.00	0.58
20524	43539	95017	0.80	0.67	0.30	0.80	0.62
20508	43491	95018	0.20	0.76	0.70	0.35	0.75
20500	43579	95019	1.50	0.52	0.70	0.35	1.20
20491	43582	95020	0.80	0.52	0.70	0.35	1.49
20461	43509	95021	0.20	0.50	0.42	0.25	0.12
20475	43572	95022	0.20	0.50	0.42	0.25	1.15
20455	43576	95023	0.50	0.38	0.64	0.40	1.08
20513	43542	95024	0.70	0.50	0.70	0.35	0.66
20501	43559	95025	0.90	0.67	0.70	0.35	0.80
20481	43538	95026	0.20	0.50	0.42	0.25	0.37
20519	43560	95027	0.70	0.64	0.30	0.80	0.92
20502	43538	95028	0.60	0.52	0.70	0.35	0.62
20500	43542	95029	0.50	0.50	0.42	0.25	0.61
20494	43548	95030	1.10	0.52	0.42	0.25	0.60
20491	43538	95031	0.30	0.76	0.42	0.25	0.52
20487	43558	95032	0.30	0.76	0.42	0.25	0.69
20468	43557	95033	0.30	0.38	0.42	0.25	0.43
20468	43561	95034	0.50	0.41	0.42	0.25	0.57
20462	43544	95035	0.40	0.41	0.42	0.25	0.15
20461	43547	95036	0.40	0.41	0.42	0.25	0.15
20458	43560	95037	0.50	0.41	0.64	0.40	0.36
20418	43468	95038	0.40	0.41	0.42	0.25	0.54
20567	43605	95039	0.70	0.90	0.60	0.35	0.51
20567	43578	95040	0.50	0.52	0.30	0.80	0.42
20513	43575	95041	1.10	0.66	0.30	0.80	1.01
20453	43572	95042	0.40	0.38	0.64	0.40	0.79
20560	43552	95043	0.60	0.52	0.30	0.80	0.18
20450	43588	95044	0.30	0.66	0.64	0.40	1.65
20528	43524	95045	1.00	0.41	0.70	0.35	0.81
20553	43541	95046	0.40	0.52	0.71	1.00	0.07
20454	43544	95047	0.60	0.41	0.42	0.25	0.11
20515	43567	95048	0.70	0.67	0.30	0.80	0.99
20500	43563	95049	0.80	0.67	0.70	0.35	0.87
20515	43561	95050	0.40	0.64	0.30	0.80	0.91
20511	43560	95051	1.30	0.64	0.70	0.35	0.85
20503	43545	95052	1.10	0.66	0.70	0.35	0.62
Error in prediction:							
		SUMS OF SQUARES = (from the observed)		4.9698	4.9950	8.4075	11.125

Table 9.42a Comparison of predicted values of CARBONATE at the random check points in the 50 m grid, by survey output and by Kriging of de-trended data after trend modelling by bicubic spline surfaces.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING
20511	43547	95004	2.40	5.25	4.95	7.30	3.32
20513	43542	95024	0.40	1.94	4.95	7.30	10.90
20519	43560	95027	18.00	9.48	9.48	10.50	0.27
20502	43538	95028	0.20	0.82	0.70	0.70	-2.53
20500	43542	95029	1.00	1.94	0.70	0.70	1.01
20494	43548	95030	0.60	0.82	0.70	0.70	0.85
20491	43538	95031	0.20	0.68	0.35	1.00	0.48
20487	43558	95032	0.00	0.68	0.35	1.00	1.24
20468	43557	95033	0.80	3.88	0.57	0.70	0.56
20468	43561	95034	0.20	0.55	0.57	0.70	0.78
20462	43544	95035	0.30	0.55	0.74	0.40	0.63
20461	43547	95036	0.10	0.55	0.74	0.40	1.03
20458	43560	95037	0.40	0.55	0.74	0.40	0.79
20454	43544	95047	0.40	0.55	0.57	0.70	0.62
20511	43560	95051	8.90	9.48	9.48	10.50	12.47
20503	43545	95052	1.40	3.88	4.95	7.30	0.96
Error in prediction: SUMS OF SQUARES = (from the observed)				101.5021	114.166	167.68	449.2535

Table 9.42c Comparison of predicted values of CARBONATE at the random check points in the 200 m grid, by survey output and by Kriging of de-trended data after trend modelling by bicubic spline surfaces.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGIN
20563	43584	95001	0.40	0.78	2.86	3.60	-0.43
20540	43503	95002	0.40	0.78	1.45	0.60	1.79
20533	43571	95003	5.50	10.66	2.86	3.60	6.95
20511	43547	95004	2.40	4.87	1.45	0.60	3.04
20502	43512	95005	0.00	0.78	1.45	0.60	0.45
20466	43572	95006	0.20	1.22	0.77	0.80	3.19
20446	43558	95007	0.60	0.85	0.77	0.80	0.77
20415	43481	95008	0.00	0.78	0.77	0.80	0.92
20397	43501	95009	0.30	0.78	0.76	0.80	0.98
20382	43572	95010	3.70	3.15	2.86	22.00	12.94
20372	43563	95011	5.10	4.87	10.66	8.70	11.65
20544	43572	95012	0.40	1.27	1.45	0.60	3.70
20531	43504	95013	1.50	1.27	1.45	0.60	0.89
20521	43495	95014	0.10	1.27	1.45	0.60	0.47
20532	43544	95015	0.10	1.27	2.86	3.60	3.11
20532	43590	95016	3.30	4.87	2.86	3.60	3.15
20524	43539	95017	2.20	4.87	0.90	1.10	2.84
20508	43491	95018	2.10	1.27	1.45	0.60	0.64
20500	43579	95019	0.10	0.95	1.45	0.60	0.41
20491	43582	95020	0.30	0.95	1.45	0.60	1.69
20461	43509	95021	0.40	0.78	0.76	0.80	0.88
20475	43572	95022	0.10	0.78	0.76	0.80	1.28
20455	43576	95023	0.60	1.22	0.77	0.80	7.01
20513	43542	95024	0.40	0.78	1.45	0.60	2.85
20501	43559	95025	3.70	4.87	1.45	0.60	1.22
20481	43538	95026	0.30	0.78	0.76	0.80	0.75
20519	43560	95027	18.00	10.66	0.90	1.10	5.00
20502	43538	95028	0.20	0.95	1.45	0.60	1.91
20500	43542	95029	1.00	0.78	0.76	0.80	1.86
20494	43548	95030	0.60	0.95	0.76	0.80	1.20
20491	43538	95031	0.20	1.27	0.76	0.80	1.25
20487	43558	95032	0.00	1.27	0.76	0.80	-0.17
20468	43557	95033	0.80	1.22	0.76	0.80	0.35
20468	43561	95034	0.20	0.78	0.76	0.80	0.51
20462	43544	95035	0.30	0.78	0.76	0.80	0.26
20461	43547	95036	0.10	0.78	0.76	0.80	0.30
20458	43560	95037	0.40	0.78	0.77	0.80	1.05
20418	43468	95038	0.00	0.78	0.76	0.80	1.15
20567	43605	95039	25.00	1.20	0.76	0.60	14.78
20567	43578	95040	0.50	0.95	0.90	1.10	-0.26
20513	43575	95041	6.10	3.15	0.90	1.10	3.64
20453	43572	95042	1.30	1.22	0.77	0.80	4.51
20560	43552	95043	0.20	0.95	0.90	1.10	-1.47
20450	43588	95044	0.40	3.15	0.77	0.80	16.98
20528	43524	95045	0.60	0.78	1.45	0.60	1.46
20553	43541	95046	0.00	0.95	2.86	3.60	0.20
20454	43544	95047	0.40	0.78	0.76	0.80	0.34
20515	43567	95048	7.50	4.87	0.90	1.10	4.41
20500	43563	95049	1.60	4.87	1.45	0.60	0.57
20515	43561	95050	8.90	10.66	0.90	1.10	4.29
20511	43560	95051	8.90	10.66	1.45	0.60	3.37
20503	43545	95052	1.40	3.15	1.45	0.60	2.16
Error in prediction:							
		SUMS OF SQUARES =		724.379	1157.301	1490.60	863.0
(deviations from the observed)							

Table 9.43a Comparison of predicted values of SILT at the random check points in the 50 m grid, by conventional soil survey output and by Kriging of de-trended data after trend modelling by bicubic spline surfaces.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING
20511	43547	95004	32.00	23.83	24.26	19.00	24.24
20513	43542	95024	31.00	28.50	24.26	19.00	26.86
20519	43560	95027	10.00	13.20	13.20	10.00	8.70
20502	43538	95028	33.00	30.88	30.41	37.00	30.57
20500	43542	95029	30.00	28.50	30.41	37.00	30.29
20494	43548	95030	35.00	30.88	30.41	37.00	28.54
20491	43538	95031	28.00	28.29	26.50	25.00	28.00
20487	43558	95032	37.00	28.29	26.50	25.00	32.17
20468	43557	95033	29.00	25.77	28.59	27.00	24.16
20468	43561	95034	33.00	28.55	28.59	27.00	41.86
20462	43544	95035	25.00	28.55	22.90	24.00	25.10
20461	43547	95036	23.00	28.55	22.90	24.00	23.24
20458	43560	95037	25.00	28.55	22.90	24.00	23.17
20454	43544	95047	27.00	28.55	28.59	27.00	24.18
20511	43560	95051	19.00	13.20	13.20	10.00	17.64
20503	43545	95052	32.00	25.77	24.96	19.00	31.46
Error in prediction:							
		SUMS OF SQUARES = (from the observed)		344.004	370.195	828.00	265.69

Table 9.43b Comparison of predicted values of SILT at the random check points in the 100m grid, by conventional soil survey output and by Kriging of de-trended data after modelling by bicubic spline surfaces.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING
20540	43503	95002	32.00	27.24	26.24	23.50	29.70
20533	43571	95003	10.00	11.25	11.28	10.00	5.27
20511	43547	95004	32.00	24.80	24.96	25.50	27.68
20502	43512	95005	34.00	28.47	26.24	23.50	30.20
20466	43572	95006	32.00	25.75	25.75	29.00	23.47
20544	43572	95012	26.00	24.16	26.24	23.50	9.59
20531	43504	95013	24.00	24.16	26.24	23.50	29.62
20521	43495	95014	29.00	24.16	26.24	23.50	29.22
20532	43544	95015	29.00	24.16	24.96	25.50	25.99
20532	43590	95016	22.00	24.80	24.96	25.50	23.29
20524	43539	95017	24.00	24.80	24.96	25.50	28.65
20508	43491	95018	30.00	24.16	26.24	23.50	29.37
20500	43579	95019	30.00	29.69	28.85	27.50	30.69
20491	43582	95020	27.00	29.69	28.85	27.50	29.19
20461	43509	95021	26.00	27.24	28.76	28.00	22.76
20475	43572	95022	33.00	25.75	25.75	29.00	21.81
20455	43576	95023	23.00	27.24	24.96	25.50	27.91
20513	43542	95024	31.00	11.25	11.28	10.00	21.17
20519	43560	95027	10.00	29.69	26.24	23.50	29.20
20502	43538	95028	33.00	27.24	26.24	23.50	29.46
20500	43542	95029	30.00	29.69	26.24	23.50	31.39
20494	43548	95030	35.00	24.16	26.24	23.50	29.84
20491	43538	95031	28.00	24.16	26.24	23.50	34.69
20487	43558	95032	37.00	25.75	28.76	28.00	26.99
20468	43557	95033	29.00	28.47	28.76	28.00	26.00
20468	43561	95034	33.00	28.47	28.76	28.00	27.28
20462	43544	95035	25.00	28.47	28.76	28.00	27.08
20461	43547	95036	23.00	28.47	22.50	21.00	22.08
20458	43560	95037	25.00	24.07	28.50	27.00	24.27
20513	43575	95041	22.00	25.75	25.75	29.00	20.14
20453	43572	95042	26.00	29.69	28.85	27.50	31.55
20560	43552	95043	32.00	24.07	22.50	21.00	30.24
20450	43588	95044	30.00	28.47	28.76	28.00	34.53
20528	43524	95045	31.00	29.69	28.85	27.50	36.00
20553	43541	95046	28.00	28.47	22.50	21.00	27.49
20454	43544	95047	27.00	24.80	11.28	10.00	21.94
20515	43567	95048	13.00	24.80	28.50	27.00	31.02
20500	43563	95049	34.00	11.25	11.28	10.00	26.38
20511	43560	95051	19.00	24.07	24.96	25.50	28.96
Error in prediction:							
		SUMS OF SQUARES =		2233.62	2398.05	2745.25	1921.9728
		(deviations from the observed)					

Table 9.43c Comparison of predicted values of SILT at the random check points in the 200m grid, by conventional soil survey output and by Kriging of de-trended data after trend modelling by bicubic spline surfaces.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING
20563	43584	95001	27.00	25.00	24.72	28.00	19.14
20540	43503	95002	32.00	25.00	26.34	26.50	23.26
20533	43571	95003	10.00	13.80	24.72	28.00	19.93
20511	43547	95004	32.00	22.00	26.34	26.50	26.52
20502	43512	95005	34.00	26.22	26.34	26.50	25.41
20466	43572	95006	32.00	25.75	21.33	21.00	28.31
20446	43558	95007	20.00	20.23	21.33	21.00	26.61
20415	43481	95008	25.00	26.22	21.33	21.00	22.90
20397	43501	95009	24.00	26.22	26.66	30.50	22.71
20382	43572	95010	17.00	22.83	24.72	16.00	19.70
20372	43563	95011	19.00	22.00	13.80	14.00	20.80
20544	43572	95012	26.00	26.80	26.34	26.50	19.29
20531	43504	95013	24.00	26.80	26.34	26.50	23.73
20521	43495	95014	29.00	26.80	26.34	26.50	25.00
20532	43544	95015	29.00	26.80	24.72	28.00	27.47
20532	43590	95016	22.00	22.00	24.72	28.00	26.75
20524	43539	95017	24.00	22.00	33.00	33.00	27.76
20508	43491	95018	30.00	26.80	26.34	26.50	26.54
20500	43579	95019	30.00	29.16	26.34	26.50	27.72
20491	43582	95020	27.00	29.16	26.34	26.50	28.52
20461	43509	95021	26.00	25.00	26.66	30.00	25.02
20475	43572	95022	33.00	25.00	26.66	30.00	29.50
20455	43576	95023	23.00	25.75	21.33	21.00	24.57
20513	43542	95024	31.00	25.00	26.34	26.50	27.03
20501	43559	95025	32.00	22.00	26.34	26.50	27.24
20481	43538	95026	25.00	25.00	26.66	30.50	28.61
20519	43560	95027	10.00	13.80	33.00	33.00	23.34
20502	43538	95028	33.00	29.16	26.34	26.50	27.74
20500	43542	95029	30.00	25.00	26.66	30.50	27.87
20494	43548	95030	35.00	29.16	26.66	30.50	28.56
20491	43538	95031	28.00	26.80	26.66	30.50	28.30
20487	43558	95032	37.00	26.80	26.66	30.50	29.75
20468	43557	95033	29.00	25.75	26.66	30.50	29.97
20468	43561	95034	33.00	26.22	26.66	30.50	29.93
20462	43544	95035	25.00	26.22	26.66	30.50	28.64
20461	43547	95036	23.00	26.22	26.66	30.50	28.83
20458	43560	95037	25.00	26.22	21.33	21.00	28.64
20418	43468	95038	26.00	26.22	26.66	30.50	23.07
20567	43605	95039	23.00	21.00	26.00	26.50	12.18
20567	43578	95040	30.00	29.16	33.00	33.00	24.92
20513	43575	95041	22.00	22.83	33.00	33.00	24.49
20453	43572	95042	26.00	25.75	21.33	21.00	25.43
20560	43552	95043	32.00	29.16	33.00	33.00	35.26
20450	43588	95044	30.00	22.83	21.33	21.00	20.00
20528	43524	95045	31.00	26.22	26.34	26.50	28.44
20553	43541	95046	28.00	29.16	24.72	28.00	35.99
20454	43544	95047	27.00	26.22	26.66	30.50	27.85
20515	43567	95048	13.00	22.00	33.00	33.00	23.55
20500	43563	95049	34.00	22.00	26.34	26.50	27.36
20515	43561	95050	1.00	13.80	33.00	33.00	23.98
20511	43560	95051	19.00	13.80	26.34	26.50	24.94
20503	43545	95052	32.00	22.83	26.34	26.50	27.52
Error in prediction:							
		SUMS OF SQUARES =		1416.22	3490.62	3533.25	2054.82
	(deviations from the observed)						

Table 9.44a Comparison of predicted values of COARSE SAND at the random check points in the 50m grid, by conventional soil survey output and by Kriging of de-trended data after trend modelling by bicubic spline surfaces.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING
20511	43547	95004	8.00	12.50	12.40	23.00	9.87
20513	43542	95024	6.00	4.25	12.40	23.00	9.30
20519	43560	95027	44.00	34.00	34.00	41.00	-3.73
20502	43538	95028	3.00	4.11	4.41	4.00	5.78
20500	43542	95029	8.00	4.25	4.41	4.00	5.45
20494	43548	95030	5.00	4.11	4.41	4.00	4.79
20491	43538	95031	8.00	5.94	9.00	9.00	8.17
20487	43558	95032	6.00	5.94	9.00	9.00	3.78
20468	43557	95033	13.00	9.66	2.88	3.50	3.93
20468	43561	95034	3.00	3.00	2.88	3.50	2.52
20462	43544	95035	2.00	3.00	1.63	2.00	1.99
20461	43547	95036	2.00	3.00	1.63	2.00	1.96
20458	43560	95037	2.00	3.00	1.63	2.00	2.77
20454	43544	95047	1.00	3.00	2.88	3.50	1.18
20511	43560	95051	26.00	34.00	34.00	41.00	31.83
20503	43545	95052	6.00	9.66	12.40	23.00	4.47
Error in prediction:							
		SUMS OF SQUARES =		239.19	396.87	1161.75	2431.2820
(deviations from the observed)							

Table 9.44b Comparison of predicted values of COARSE SAND at the random check points in the 100m grid, by conventional soil survey output and by Kriging of de-trended data after trend modeling by bicubic spline surfaces.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING
20540	43503	95002	6.00	4.93	4.90	5.00	6.48
20533	43571	95003	43.00	39.35	39.28	41.00	49.14
20511	43547	95004	8.00	10.46	10.75	11.50	16.48
20502	43512	95005	6.00	3.82	4.90	5.00	4.73
20466	43572	95006	6.00	7.00	7.00	6.00	8.09
20544	43572	95012	9.00	5.51	4.90	5.00	39.97
20531	43504	95013	8.00	5.51	4.90	5.00	6.04
20521	43495	95014	3.00	5.51	4.90	5.00	6.60
20532	43544	95015	7.00	5.51	10.75	11.50	14.37
20532	43590	95016	15.00	10.46	10.75	11.50	9.78
20524	43539	95017	11.00	10.46	10.75	11.50	13.36
20508	43491	95018	4.00	5.51	4.90	5.00	3.32
20500	43579	95019	5.00	4.84	5.28	5.50	4.13
20491	43582	95020	5.00	4.84	5.28	5.50	4.20
20461	43509	95021	3.00	4.93	3.52	2.00	4.80
20475	43572	95022	6.00	7.00	7.00	6.00	3.28
20455	43576	95023	6.00	4.93	10.75	11.50	14.69
20513	43542	95024	6.00	39.35	39.28	41.00	27.24
20519	43560	95027	44.00	4.84	4.90	5.00	8.28
20502	43538	95028	3.00	4.93	4.90	5.00	8.13
20500	43542	95029	8.00	4.84	4.90	5.00	4.79
20494	43548	95030	5.00	5.51	4.90	5.00	4.48
20491	43538	95031	8.00	5.51	4.90	5.00	1.32
20487	43558	95032	6.00	7.00	3.52	2.00	5.44
20468	43557	95033	13.00	3.82	3.52	2.00	5.87
20468	43561	95034	3.00	3.82	3.52	2.00	2.34
20462	43544	95035	2.00	3.82	3.52	2.00	1.20
20461	43547	95036	2.00	3.82	2.75	1.00	-1.37
20458	43560	95037	2.00	10.42	6.00	8.00	16.01
20513	43575	95041	16.00	7.00	7.00	6.00	0.86
20453	43572	95042	5.00	4.84	5.28	5.50	3.31
20560	43552	95043	4.00	10.42	2.75	1.00	5.73
20450	43588	95044	8.00	3.82	3.52	2.00	4.56
20528	43524	95045	8.00	4.84	5.28	5.50	5.62
20553	43541	95046	5.00	3.82	2.75	1.00	-4.33
20454	43544	95047	1.00	10.46	39.28	41.00	22.37
20515	43567	95048	33.00	10.46	6.00	8.00	7.31
20500	43563	95049	5.00	39.35	39.28	41.00	18.92
20511	43560	95051	26.00	10.42	10.75	11.50	10.48
Error in prediction:							
		SUMS OF SQUARES =		5082.93	6603.42	6954.50	5227.36
(deviations from the observed)							

Table 9.44c Comparison of predicted values of COARSE SAND at the random check points in the 200m grid, by conventional soil survey output and by Kriging of de-trended data after modelling by bicubic spline surfaces.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING
20563	43584	95001	5.00	5.08	10.00	12.00	3.60
20540	43503	95002	6.00	5.08	5.73	5.50	6.23
20533	43571	95003	43.00	32.60	10.00	12.00	36.26
20511	43547	95004	8.00	15.91	5.73	5.50	13.12
20502	43512	95005	6.00	2.86	5.73	5.50	7.95
20466	43572	95006	6.00	7.00	2.20	2.00	5.47
20446	43558	95007	2.00	3.00	2.20	2.00	7.51
20415	43481	95008	2.00	2.86	2.20	2.00	2.51
20397	43501	95009	2.00	2.86	3.00	2.00	1.77
20382	43572	95010	15.00	6.50	10.00	28.00	15.21
20372	43563	95011	19.00	15.91	32.60	29.00	21.23
20544	43572	95012	9.00	5.60	5.73	5.50	22.57
20531	43504	95013	8.00	5.60	5.73	5.50	7.74
20521	43495	95014	3.00	5.60	5.73	5.50	7.21
20532	43544	95015	7.00	5.60	10.00	12.00	11.15
20532	43590	95016	15.00	15.91	10.00	12.00	3.62
20524	43539	95017	11.00	15.91	4.00	5.00	10.75
20508	43491	95018	4.00	5.60	5.73	5.50	5.33
20500	43579	95019	5.00	5.33	5.73	5.50	6.26
20491	43582	95020	5.00	5.33	5.73	5.50	-0.33
20461	43509	95021	3.00	5.08	3.00	2.00	3.20
20475	43572	95022	6.00	5.08	3.00	2.00	0.69
20455	43576	95023	6.00	7.00	2.20	2.00	6.57
20513	43542	95024	6.00	5.08	5.73	5.50	12.23
20501	43559	95025	8.00	15.91	5.73	5.50	9.27
20481	43538	95026	6.00	5.08	3.00	2.00	5.59
20519	43560	95027	44.00	32.60	4.00	5.00	24.77
20502	43538	95028	3.00	5.33	5.73	5.50	10.16
20500	43542	95029	8.00	5.08	3.00	2.00	9.40
20494	43548	95030	5.00	5.33	3.00	2.00	6.45
20491	43538	95031	8.00	5.60	3.00	2.00	7.39
20487	43558	95032	6.00	5.60	3.00	2.00	2.43
20468	43557	95033	13.00	7.00	3.00	2.00	5.33
20468	43561	95034	3.00	2.86	3.00	2.00	5.76
20462	43544	95035	2.00	2.86	3.00	2.00	3.74
20461	43547	95036	2.00	2.86	3.00	2.00	4.12
20458	43560	95037	2.00	2.86	2.20	2.00	8.45
20418	43468	95038	2.00	2.86	3.00	2.00	2.70
20567	43605	95039	23.00	8.00	6.33	5.50	21.73
20567	43578	95040	5.00	5.33	4.00	5.00	0.87
20513	43575	95041	16.00	6.50	4.00	5.00	25.56
20453	43572	95042	5.00	7.00	2.20	2.00	10.24
20560	43552	95043	4.00	5.33	4.00	5.00	2.71
20450	43588	95044	8.00	6.50	2.20	2.00	-7.86
20528	43524	95045	8.00	2.86	5.73	5.50	7.62
20553	43541	95046	5.00	5.33	10.00	12.00	4.51
20454	43544	95047	1.00	6.50	3.00	2.00	3.22
20515	43567	95048	33.00	15.91	4.00	5.00	27.19
20500	43563	95049	5.00	15.91	5.73	5.50	8.11
20515	43561	95050	34.00	32.60	4.00	5.00	22.96
20511	43560	95051	26.00	32.60	5.73	5.50	18.85
20503	43545	95052	6.00	6.50	5.73	5.50	10.27
Error in prediction:							
		SUMS OF SQUARES =		1416.85	5917.29	5776.25	1747.68
		(deviations from the observed)					

not necessarily corresponds to the underlying spatial structures of the original property values.

The final interpolated values have two components, i.e. the trend and the residuals, of which, the trend is in most instances the largest and contributes more to the definition of the predicted value. Yet, trend values depend rather, on the goodness of the surface fit achieved.

Not only the structures of the semi-variogram can explain the predictive success of the procedures compared. The quality of the map made and the criteria for the definition of the classes and their breadth explains their predictive success.

So, there is no apparent relationship between structures in the semi-variograms of residuals from trend and the predictive success of the various techniques compared.

Residuals from the trend of Potassium at the 50m grid (fig. 9.35a) showed spatial structure and a good fit of the model. These features may explain why Kriging made the best predictions at this scale. But, on the other hand, the semi-variogram of residuals of Silt (fig. 9.37a) at the same 50m grid showed no spatial structures, i.e. a total absence of spatial dependence, and random variation. Yet, the best predictions in this case were made by Kriging, even if the interpolated values of residuals came only from the mean of residuals from trend.

It can be shown that in situations of trend removal by modelling with surfaces the predictive success of Kriging can be explained by both the goodness of the trend surface fit and by the structures of the semi-variograms of residuals from trend.

The most important practical implication of the results obtained here is that the modelling of the trends by spline surfaces and the subsequent spatial analysis of the residuals from the trend, made possible the application of spatial interpolation and predictions by Kriging, in cases where anisotropic variation caused by trends prevented their application.

Focusing on the 50m grid, where these anisotropies and trends were present, it can be seen that the exercise of trend removal by modelling with spline surfaces was a success, in the sense that not only made possible predictions by Kriging but

also in that Kriging provided the best predictions of Potassium (table 9.41a) and Silt (9.43a) at such sampling scale.

9.21.2 Comparison of predictive procedures when trends are partitioned by stratification based on soil mapping units.

The results of the comparison are given in tables 9.45a-c for Potassium at the three sampling scales respectively. In the same order tables 9.46a-c give those for Carbonate, tables 9.47a-c list the results for Silt and tables 9.48a-c for Coarse Sand.

The partitioning of the trend by means of stratification based on soil mapping units did indeed allow for predictions by Kriging interpolation in cases where it was not possible without accounting for the trends. However, Kriging was not the best technique in all situations, and in some of them did surprisingly badly (tables 9.45b, 9.46b and 9.48b). This performance, nevertheless, can be explained by the structures of the within-strata semi-variograms (figs. 9.39a to 9.46c) which led to Kriging interpolation.

In almost half of all the cases (42%) given by the combination of properties and sampling scales, i.e. 10 out of 24 semi-variograms, Kriging was not applied due to total lack of spatial dependence showed by the corresponding within-stratum semi-variogram and means from the stratum were applied instead.

So, strictly speaking, Kriging was only applied in just over half of the cases were a semi-variogram was calculated, yet the means from strata used as predictors are taken in the comparison as if they were from Kriging. In this sense the comparison is unfair to Kriging. This argument bears a heavy weight in explaining the apparent poor predictive performance of this technique.

Chemical properties, i.e. Potassium and Carbonate, in the non-calcareous soils (NS-NG stratum) showed no spatial structure as evidenced by the random variation depicted by the semi-variograms (figs. 9.40b and c, and 9.42a,b and c). In consequence, the Kriging predictions of these properties were based on the mean values within strata which are the result of

Table 9.45a Comparison of the success in predicting POTASSIUM by Kriging of de-trended data and by other procedures at the 50 m grid when TRENDS are accounted for by stratification based on mapping units.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING
20511	43547	95004	1.00	1.15	0.86	0.60	0.65
20513	43542	95024	0.70	0.38	0.86	0.60	0.65
20519	43560	95027	0.70	0.92	0.92	0.70	0.65
20502	43538	95028	0.60	0.59	0.55	0.40	0.52
20500	43542	95029	0.50	0.38	0.55	0.40	0.52
20494	43548	95030	1.10	0.59	0.55	0.40	0.64
20491	43538	95031	0.30	0.50	0.55	0.90	0.65
20487	43558	95032	0.30	0.50	0.55	0.90	0.65
20468	43557	95033	0.30	0.77	0.40	0.40	0.47
20468	43561	95034	0.50	0.37	0.40	0.40	0.50
20462	43544	95035	0.40	0.37	0.56	0.50	0.50
20461	43547	95036	0.40	0.37	0.56	0.50	0.51
20458	43560	95037	0.50	0.37	0.56	0.50	0.63
20454	43544	95047	0.60	0.37	0.40	0.40	0.43
20511	43560	95051	1.30	0.92	0.92	0.70	0.65
20503	43545	95052	1.10	0.77	0.86	0.60	0.65
Error in prediction: SUMS OF SQUARES = (from the observed)				1.0847	0.8429	2.1200	1.3094

Table 9.45b Comparison of the success in predicting POTASSIUM by Kriging of de-trended data and by other procedures from survey at the 100 m grid, when TRENDS are accounted for by stratification based on mapping units.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING
20563	43584	95001	0.60	0.48	0.69	0.70	0.62
20540	43503	95002	0.90	0.59	0.60	0.45	0.46
20533	43571	95003	0.50	1.09	0.58	0.70	1.12
20511	43547	95004	1.00	0.39	0.90	1.00	0.52
20502	43512	95005	0.40	0.38	0.60	0.45	0.41
20466	43572	95006	0.30	0.73	0.37	0.40	0.64
20544	43572	95012	0.40	0.73	0.60	0.45	0.97
20531	43504	95013	1.20	0.73	0.60	0.45	1.37
20521	43495	95014	0.60	0.73	0.60	0.45	0.60
20532	43544	95015	0.50	1.09	0.90	1.00	0.80
20532	43590	95016	1.00	1.09	0.90	1.00	0.65
20524	43539	95017	0.80	0.73	0.90	1.00	0.49
20508	43491	95018	0.20	0.68	0.60	0.45	0.50
20500	43579	95019	1.50	0.68	0.69	0.70	0.50
20491	43582	95020	0.80	0.48	0.69	0.70	0.50
20461	43509	95021	0.20	0.38	0.32	0.20	0.30
20475	43572	95022	0.20	0.48	0.37	0.40	0.87
20455	43576	95023	0.50	0.59	0.90	1.00	0.73
20513	43542	95024	0.70	0.68	0.58	0.70	0.45
20519	43560	95027	0.70	0.48	0.60	0.45	0.55
20502	43538	95028	0.60	0.68	0.60	0.45	0.68
20500	43542	95029	0.50	0.73	0.60	0.45	0.21
20494	43548	95030	1.10	0.73	0.60	0.45	0.41
20491	43538	95031	0.30	0.38	0.60	0.45	0.50
20487	43558	95032	0.30	0.39	0.32	0.20	0.50
20468	43557	95033	0.30	0.39	0.32	0.20	0.50
20468	43561	95034	0.50	0.39	0.32	0.20	0.50
20462	43544	95035	0.40	0.39	0.32	0.20	0.50
20461	43547	95036	0.40	0.73	0.55	0.60	0.84
20458	43560	95037	0.50	0.38	0.95	0.80	0.37
20513	43575	95041	1.10	0.68	0.37	0.40	0.50
20453	43572	95042	0.40	0.73	0.69	0.70	0.50
20560	43552	95043	0.60	0.39	0.55	0.60	0.50
20450	43588	95044	0.30	0.68	0.32	0.20	0.50
20528	43524	95045	1.00	0.39	0.69	0.70	0.50
20553	43541	95046	0.40	1.09	0.55	0.60	0.73
20454	43544	95047	0.60	1.09	0.58	0.70	0.93
20515	43567	95048	0.70	0.59	0.95	0.80	0.82
20500	43563	95049	0.80	0.73	0.58	0.70	0.87
Error in prediction:							
		SUMS OF SQUARES = (from the observed)		4.7947	3.2597	3.7075	5.191

Table 9.45c Comparison of the success in predicting POTASSIUM by Kriging o de-trended data and other predictive procedures at the 200 m grid, when trends are accounted for by stratification based on mapping units.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING
20563	43584	95001	0.60	0.50	0.71	1.00	0.40
20540	43503	95002	0.90	0.50	0.70	0.35	1.23
20533	43571	95003	0.50	0.64	0.71	1.00	0.52
20511	43547	95004	1.00	0.67	0.70	0.35	1.17
20502	43512	95005	0.40	0.41	0.70	0.35	0.62
20466	43572	95006	0.30	0.38	0.64	0.40	0.52
20446	43558	95007	0.60	0.64	0.64	0.40	0.52
20415	43481	95008	0.40	0.41	0.64	0.40	0.52
20397	43501	95009	0.30	0.41	0.42	0.25	0.52
20382	43572	95010	0.40	0.66	0.71	0.70	0.39
20372	43563	95011	0.50	0.67	0.64	0.50	0.51
20544	43572	95012	0.40	0.76	0.70	0.35	0.56
20531	43504	95013	1.20	0.76	0.70	0.35	1.32
20521	43495	95014	0.60	0.76	0.70	0.35	0.80
20532	43544	95015	0.50	0.76	0.71	1.00	0.56
20532	43590	95016	1.00	0.67	0.71	1.00	0.76
20524	43539	95017	0.80	0.67	0.30	0.80	0.52
20508	43491	95018	0.20	0.76	0.70	0.35	0.43
20500	43579	95019	1.50	0.52	0.70	0.35	1.48
20491	43582	95020	0.80	0.52	0.70	0.35	1.60
20461	43509	95021	0.20	0.50	0.42	0.25	0.52
20475	43572	95022	0.20	0.50	0.42	0.25	0.52
20455	43576	95023	0.50	0.38	0.64	0.40	0.52
20513	43542	95024	0.70	0.50	0.70	0.35	0.99
20501	43559	95025	0.90	0.67	0.70	0.35	1.07
20481	43538	95026	0.20	0.50	0.42	0.25	0.52
20519	43560	95027	0.70	0.64	0.30	0.80	0.52
20502	43538	95028	0.60	0.52	0.70	0.35	0.75
20500	43542	95029	0.50	0.50	0.42	0.25	0.52
20494	43548	95030	1.10	0.52	0.42	0.25	0.52
20491	43538	95031	0.30	0.76	0.42	0.25	0.52
20487	43558	95032	0.30	0.76	0.42	0.25	0.52
20468	43557	95033	0.30	0.38	0.42	0.25	0.52
20468	43561	95034	0.50	0.41	0.42	0.25	0.52
20462	43544	95035	0.40	0.41	0.42	0.25	0.52
20461	43547	95036	0.40	0.41	0.42	0.25	0.52
20458	43560	95037	0.50	0.41	0.64	0.40	0.52
20418	43468	95038	0.40	0.41	0.42	0.25	0.52
20567	43605	95039	0.70	0.90	0.60	0.35	0.54
20567	43578	95040	0.50	0.52	0.30	0.80	0.52
20513	43575	95041	1.10	0.66	0.30	0.80	0.52
20453	43572	95042	0.40	0.38	0.64	0.40	0.52
20560	43552	95043	0.60	0.52	0.30	0.80	0.52
20450	43588	95044	0.30	0.66	0.64	0.40	0.52
20528	43524	95045	1.00	0.41	0.70	0.35	0.89
20553	43541	95046	0.40	0.52	0.71	1.00	0.34
20454	43544	95047	0.60	0.41	0.42	0.25	0.52
20515	43567	95048	0.70	0.67	0.30	0.80	0.52
20500	43563	95049	0.80	0.67	0.70	0.35	1.14
20515	43561	95050	0.40	0.64	0.30	0.80	0.52
20511	43560	95051	1.30	0.64	0.70	0.35	1.00
20503	43545	95052	1.10	0.66	0.70	0.35	0.95
Error in prediction:							
		SUMS OF SQUARES =		4.9698	4.9950	8.4075	3.0052

Table 9.46a Comparison of the success in predicting CARBONATE by Kriging of de-trended data and other procedures from conventional survey at the 50 m grid, when TRENDS are accounted for by stratification based on mapping units.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING
20511	43547	95004	2.40	5.25	4.95	7.30	3.62
20513	43542	95024	0.40	1.94	4.95	7.30	2.15
20519	43560	95027	18.00	9.48	9.48	10.50	10.10
20502	43538	95028	0.20	0.82	0.70	0.70	0.64
20500	43542	95029	1.00	1.94	0.70	0.70	0.64
20494	43548	95030	0.60	0.82	0.70	0.70	0.64
20491	43538	95031	0.20	0.68	0.35	1.00	3.64
20487	43558	95032	0.00	0.68	0.35	1.00	1.24
20468	43557	95033	0.80	3.88	0.57	0.70	0.64
20468	43561	95034	0.20	0.55	0.57	0.70	0.64
20462	43544	95035	0.30	0.55	0.74	0.40	0.64
20461	43547	95036	0.10	0.55	0.74	0.40	0.64
20458	43560	95037	0.40	0.55	0.74	0.40	0.64
20454	43544	95047	0.40	0.55	0.57	0.70	0.64
20511	43560	95051	8.90	9.48	9.48	10.50	11.27
20503	43545	95052	1.40	3.88	4.95	7.30	1.94
Error in prediction:							
		SUMS OF SQUARES =		101.502	114.166	167.680	87.2490
		(from the observed)					

Table 9.46b Comparison of the success in predicting CARBONATE by Kriging on de-trended data and other procedures from conventional survey at the 100 m grid, when TRENDS are accounted for by stratification based on mapping units.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING
20563	43584	95001	0.40	0.78	0.77	0.85	0.78
20540	43503	95002	0.40	9.63	1.01	2.70	8.57
20533	43571	95003	5.50	3.68	9.60	10.50	4.79
20511	43547	95004	2.40	0.62	3.40	2.45	0.69
20502	43512	95005	0.00	1.22	1.01	2.75	1.47
20466	43572	95006	0.20	0.92	1.22	0.40	5.31
20544	43572	95012	0.40	0.92	1.01	2.75	0.98
20531	43504	95013	1.50	0.92	1.01	2.75	1.26
20521	43495	95014	0.10	0.92	1.01	2.75	5.29
20532	43544	95015	0.10	3.68	3.40	2.45	5.06
20532	43590	95016	3.30	3.68	3.40	2.45	3.75
20524	43539	95017	2.20	0.92	3.40	2.45	0.63
20508	43491	95018	2.10	0.85	1.01	2.75	0.73
20500	43579	95019	0.10	0.85	0.77	0.85	0.73
20491	43582	95020	0.30	0.78	0.77	0.85	0.73
20461	43509	95021	0.40	1.22	0.58	0.60	0.90
20475	43572	95022	0.10	0.78	1.22	0.40	3.75
20455	43576	95023	0.60	9.63	3.40	2.45	10.25
20513	43542	95024	0.40	0.85	9.60	10.50	1.07
20519	43560	95027	18.00	0.78	1.01	2.75	1.10
20502	43538	95028	0.20	0.85	1.01	2.75	1.06
20500	43542	95029	1.00	0.92	1.01	2.75	0.37
20494	43548	95030	0.60	0.92	1.01	2.75	1.63
20491	43538	95031	0.20	1.22	1.01	2.75	0.73
20487	43558	95032	0.00	0.62	0.58	0.60	0.73
20468	43557	95033	0.80	0.62	0.58	0.60	0.73
20468	43561	95034	0.20	0.62	0.58	0.60	0.73
20462	43544	95035	0.30	0.62	0.58	0.60	0.73
20461	43547	95036	0.10	3.26	1.17	0.60	4.17
20458	43560	95037	0.40	1.22	1.90	2.10	0.65
20513	43575	95041	6.10	0.85	1.22	0.40	0.73
20453	43572	95042	1.30	3.26	0.77	0.85	0.73
20560	43552	95043	0.20	0.62	1.17	0.60	0.73
20450	43588	95044	0.40	0.85	0.58	0.60	0.73
20528	43524	95045	0.60	0.62	0.77	0.85	0.73
20553	43541	95046	0.00	3.62	1.17	0.60	7.00
20454	43544	95047	0.40	3.62	9.60	10.50	2.23
20515	43567	95048	7.50	9.63	1.90	2.10	8.71
20500	43563	95049	1.60	3.26	9.60	10.50	2.18
Error in prediction:							
		SUMS OF SQUARES =		566.939	630.393	666.100	652.0
		(from the observed)					

Table 9.46b Comparison of the success in predicting CARBONATE by Kriging of de-trended data and other procedures from conventional survey at the 100 m grid, when TRENDS are accounted for by stratification based on mapping units.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING
20563	43584	95001	0.40	0.78	0.77	0.85	0.78
20540	43503	95002	0.40	9.63	1.01	2.70	8.57
20533	43571	95003	5.50	3.68	9.60	10.50	4.79
20511	43547	95004	2.40	0.62	3.40	2.45	0.69
20502	43512	95005	0.00	1.22	1.01	2.75	1.47
20466	43572	95006	0.20	0.92	1.22	0.40	5.31
20544	43572	95012	0.40	0.92	1.01	2.75	0.98
20531	43504	95013	1.50	0.92	1.01	2.75	1.26
20521	43495	95014	0.10	0.92	1.01	2.75	5.29
20532	43544	95015	0.10	3.68	3.40	2.45	5.06
20532	43590	95016	3.30	3.68	3.40	2.45	3.75
20524	43539	95017	2.20	0.92	3.40	2.45	0.63
20508	43491	95018	2.10	0.85	1.01	2.75	0.73
20500	43579	95019	0.10	0.85	0.77	0.85	0.73
20491	43582	95020	0.30	0.78	0.77	0.85	0.73
20461	43509	95021	0.40	1.22	0.58	0.60	0.90
20475	43572	95022	0.10	0.78	1.22	0.40	3.75
20455	43576	95023	0.60	9.63	3.40	2.45	10.25
20513	43542	95024	0.40	0.85	9.60	10.50	1.07
20519	43560	95027	18.00	0.78	1.01	2.75	1.10
20502	43538	95028	0.20	0.85	1.01	2.75	1.06
20500	43542	95029	1.00	0.92	1.01	2.75	0.37
20494	43548	95030	0.60	0.92	1.01	2.75	1.63
20491	43538	95031	0.20	1.22	1.01	2.75	0.73
20487	43558	95032	0.00	0.62	0.58	0.60	0.73
20468	43557	95033	0.80	0.62	0.58	0.60	0.73
20468	43561	95034	0.20	0.62	0.58	0.60	0.73
20462	43544	95035	0.30	0.62	0.58	0.60	0.73
20461	43547	95036	0.10	3.26	1.17	0.60	4.17
20458	43560	95037	0.40	1.22	1.90	2.10	0.65
20513	43575	95041	6.10	0.85	1.22	0.40	0.73
20453	43572	95042	1.30	3.26	0.77	0.85	0.73
20560	43552	95043	0.20	0.62	1.17	0.60	0.73
20450	43588	95044	0.40	0.85	0.58	0.60	0.73
20528	43524	95045	0.60	0.62	0.77	0.85	0.73
20553	43541	95046	0.00	3.62	1.17	0.60	7.00
20454	43544	95047	0.40	3.62	9.60	10.50	2.23
20515	43567	95048	7.50	9.63	1.90	2.10	8.71
20500	43563	95049	1.60	3.26	9.60	10.50	2.18
Error in prediction:							
		SUMS OF SQUARES =		566.939	630.393	666.100	652.0
		(from the observed)					

Table 9.46c Comparison of the success in predicting CARBONATE by Kriging of de-trended data and other procedures from conventional survey at the 200 m grid, when trends are accounted for by stratification based on mapping units.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING
20563	43584	95001	0.40	0.78	2.86	3.60	4.88
20540	43503	95002	0.40	0.78	1.45	0.60	1.99
20533	43571	95003	5.50	10.66	2.86	3.60	3.35
20511	43547	95004	2.40	4.87	1.45	0.60	1.93
20502	43512	95005	0.00	0.78	1.45	0.60	1.62
20466	43572	95006	0.20	1.22	0.77	0.80	0.77
20446	43558	95007	0.60	0.85	0.77	0.80	0.77
20415	43481	95008	0.00	0.78	0.77	0.80	0.77
20397	43501	95009	0.30	0.78	0.76	0.80	0.77
20382	43572	95010	3.70	3.15	2.86	22.00	6.39
20372	43563	95011	5.10	4.87	10.66	8.70	6.42
20544	43572	95012	0.40	1.27	1.45	0.60	3.59
20531	43504	95013	1.50	1.27	1.45	0.60	1.98
20521	43495	95014	0.10	1.27	1.45	0.60	2.10
20532	43544	95015	0.10	1.27	2.86	3.60	1.82
20532	43590	95016	3.30	4.87	2.86	3.60	5.29
20524	43539	95017	2.20	4.87	0.90	1.10	0.77
20508	43491	95018	2.10	1.27	1.45	0.60	1.80
20500	43579	95019	0.10	0.95	1.45	0.60	4.07
20491	43582	95020	0.30	0.95	1.45	0.60	4.31
20461	43509	95021	0.40	0.78	0.76	0.80	0.77
20475	43572	95022	0.10	0.78	0.76	0.80	0.77
20455	43576	95023	0.60	1.22	0.77	0.80	0.77
20513	43542	95024	0.40	0.78	1.45	0.60	1.86
20501	43559	95025	3.70	4.87	1.45	0.60	2.73
20481	43538	95026	0.30	0.78	0.76	0.80	0.77
20519	43560	95027	18.00	10.66	0.90	1.10	0.77
20502	43538	95028	0.20	0.95	1.45	0.60	1.63
20500	43542	95029	1.00	0.78	0.76	0.80	0.75
20494	43548	95030	0.60	0.95	0.76	0.80	0.75
20491	43538	95031	0.20	1.27	0.76	0.80	0.75
20487	43558	95032	0.00	1.27	0.76	0.80	0.75
20468	43557	95033	0.80	1.22	0.76	0.80	0.75
20468	43561	95034	0.20	0.78	0.76	0.80	0.75
20462	43544	95035	0.30	0.78	0.76	0.80	0.75
20461	43547	95036	0.10	0.78	0.76	0.80	0.75
20458	43560	95037	0.40	0.78	0.77	0.80	0.75
20418	43468	95038	0.00	0.78	0.76	0.80	0.75
20567	43605	95039	25.00	1.20	0.76	0.60	6.80
20567	43578	95040	0.50	0.95	0.90	1.10	0.75
20513	43575	95041	6.10	3.15	0.90	1.10	0.75
20453	43572	95042	1.30	1.22	0.77	0.80	0.75
20560	43552	95043	0.20	0.95	0.90	1.10	0.75
20450	43588	95044	0.40	3.15	0.77	0.80	0.75
20528	43524	95045	0.60	0.78	1.45	0.60	1.51
20553	43541	95046	0.00	0.95	2.86	3.60	1.57
20454	43544	95047	0.40	0.78	0.76	0.80	0.75
20515	43567	95048	7.50	4.87	0.90	1.10	0.75
20500	43563	95049	1.60	4.87	1.45	0.60	2.92
20515	43561	95050	8.90	10.66	0.90	1.10	0.75
20511	43560	95051	8.90	10.66	1.45	0.60	2.65
20503	43545	95052	1.40	3.15	1.45	0.60	1.82
Error in prediction:							
		SUMS OF SQUARES = (from the observed)		724.37	1157.30	1490.60	917.9

Table 9.47a Comparison of the success in predicting SILT by Kriging of de-trended data and by other procedures from conventional survey at the 50 m grid , when TRENDS are accounted for by stratification based on mapping units.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING
20511	43547	95004	32.00	23.83	24.26	19.00	27.15
20513	43542	95024	31.00	28.50	24.26	19.00	27.01
20519	43560	95027	10.00	13.20	13.20	10.00	10.10
20502	43538	95028	33.00	30.88	30.41	37.00	29.52
20500	43542	95029	30.00	28.50	30.41	37.00	28.78
20494	43548	95030	35.00	30.88	30.41	37.00	27.15
20491	43538	95031	28.00	28.29	26.50	25.00	29.33
20487	43558	95032	37.00	28.29	26.50	25.00	29.66
20468	43557	95033	29.00	25.77	28.59	27.00	29.88
20468	43561	95034	33.00	28.55	28.59	27.00	32.85
20462	43544	95035	25.00	28.55	22.90	24.00	25.14
20461	43547	95036	23.00	28.55	22.90	24.00	23.93
20458	43560	95037	25.00	28.55	22.90	24.00	22.35
20454	43544	95047	27.00	28.55	28.59	27.00	24.62
20511	43560	95051	19.00	13.20	13.20	10.00	18.46
20503	43545	95052	32.00	25.77	24.96	19.00	29.40
Error in prediction:							
		SUMS OF SQUARES =		344.00	370.14	828.00	191.73

Table 9.47b Comparison of the success in predicting SILT by Kriging of de-trended data and by procedures from conventional survey output at the 100 m grid, when TRENDS are accounted for by stratification based on soil mapping uits.

NORTHING	EASTING	SIT No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING
20563	43584	95001	27.00	27.24	28.85	27.00	27.20
20540	43503	95002	32.00	11.25	26.24	23.50	9.65
20533	43571	95003	10.00	24.80	11.28	10.00	24.95
20511	43547	95004	32.00	28.47	24.96	25.50	27.89
20502	43512	95005	34.00	25.75	26.24	23.50	24.17
20466	43572	95006	32.00	24.16	25.75	29.00	13.46
20544	43572	95012	26.00	24.16	26.24	23.50	25.90
20531	43504	95013	24.00	24.16	26.24	23.50	24.23
20521	43495	95014	29.00	24.16	26.24	23.50	21.27
20532	43544	95015	29.00	24.80	24.96	25.50	22.14
20532	43590	95016	22.00	24.80	24.96	25.50	23.64
20524	43539	95017	24.00	24.16	24.96	25.50	28.03
20508	43491	95018	30.00	29.69	26.24	23.50	28.68
20500	43579	95019	30.00	29.69	28.85	27.50	28.77
20491	43582	95020	27.00	27.24	28.85	27.50	23.15
20461	43509	95021	26.00	25.75	28.76	28.00	24.48
20475	43572	95022	33.00	27.24	25.75	29.00	24.61
20455	43576	95023	23.00	11.25	24.96	25.50	13.37
20513	43542	95024	31.00	29.69	11.28	10.00	27.15
20519	43560	95027	10.00	27.24	26.24	23.50	28.59
20502	43538	95028	33.00	29.69	26.24	23.50	30.91
20500	43542	95029	30.00	24.16	26.24	23.50	27.88
20494	43548	95030	35.00	24.16	26.24	23.50	30.63
20491	43538	95031	28.00	25.75	26.24	23.50	26.41
20487	43558	95032	37.00	28.47	28.76	28.00	26.05
20468	43557	95033	29.00	28.47	28.76	28.00	25.98
20468	43561	95034	33.00	28.47	28.76	28.00	24.57
20462	43544	95035	25.00	28.47	28.76	28.00	22.69
20461	43547	95036	23.00	24.07	22.50	21.00	24.00
20458	43560	95037	25.00	25.75	28.50	27.00	24.43
20513	43575	95041	22.00	29.69	25.75	29.00	28.39
20453	43572	95042	26.00	24.07	28.85	27.50	26.45
20560	43552	95043	32.00	28.47	22.50	21.00	32.37
20450	43588	95044	30.00	29.69	28.76	28.00	31.23
20528	43524	95045	31.00	28.47	28.85	27.50	26.21
20553	43541	95046	28.00	24.80	22.50	21.00	19.08
20454	43544	95047	27.00	24.80	11.28	10.00	29.48
20515	43567	95048	13.00	11.25	28.50	27.00	20.91
20500	43563	95049	34.00	24.07	11.28	10.00	28.02
Error in prediction:							
		SUMS OF SQUARES =		1787.98	2365.95	2703.00	2336.96

Table 9.47c Comparison of the success in predicting SILT by Kriging of de-trended data and by procedures from conventional survey output at the 200 m grid, when trends are accounted for by stratification based on mapping units.

NORHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING
20563	43584	95001	27.00	25.00	24.72	28.00	23.64
20540	43503	95002	32.00	25.00	26.34	26.50	23.64
20533	43571	95003	10.00	13.80	24.72	28.00	23.64
20511	43547	95004	32.00	22.00	26.34	26.50	23.64
20502	43512	95005	34.00	26.22	26.34	26.50	23.64
20466	43572	95006	32.00	25.75	21.33	21.00	24.29
20446	43558	95007	20.00	20.23	21.33	21.00	24.29
20415	43481	95008	25.00	26.22	21.33	21.00	24.29
20397	43501	95009	24.00	26.22	26.66	30.50	24.29
20382	43572	95010	17.00	22.83	24.72	16.00	23.64
20372	43563	95011	19.00	22.00	13.80	14.00	23.64
20544	43572	95012	26.00	26.80	26.34	26.50	23.64
20531	43504	95013	24.00	26.80	26.34	26.50	23.64
20521	43495	95014	29.00	26.80	26.34	26.50	23.64
20532	43544	95015	29.00	26.80	24.72	28.00	23.64
20532	43590	95016	22.00	22.00	24.72	28.00	23.64
20524	43539	95017	24.00	22.00	33.00	33.00	24.29
20508	43491	95018	30.00	26.80	26.34	26.50	23.64
20500	43579	95019	30.00	29.16	26.34	26.50	23.64
20491	43582	95020	27.00	29.16	26.34	26.50	23.64
20461	43509	95021	26.00	25.00	26.66	30.00	24.29
20475	43572	95022	33.00	25.00	26.66	30.00	24.29
20455	43576	95023	23.00	25.75	21.33	21.00	24.29
20513	43542	95024	31.00	25.00	26.34	26.50	23.64
20501	43559	95025	32.00	22.00	26.34	26.50	23.64
20481	43538	95026	25.00	25.00	26.66	30.50	24.29
20519	43560	95027	10.00	13.80	33.00	33.00	24.29
20502	43538	95028	33.00	29.16	26.34	26.50	23.64
20500	43542	95029	30.00	25.00	26.66	30.50	24.29
20494	43548	95030	35.00	29.16	26.66	30.50	24.29
20491	43538	95031	28.00	26.80	26.66	30.50	24.29
20487	43558	95032	37.00	26.80	26.66	30.50	24.29
20468	43557	95033	29.00	25.75	26.66	30.50	24.29
20468	43561	95034	33.00	26.22	26.66	30.50	24.29
20462	43544	95035	25.00	26.22	26.66	30.50	24.29
20461	43547	95036	23.00	26.22	26.66	30.50	24.29
20458	43560	95037	25.00	26.22	21.33	21.00	24.29
20418	43468	95038	26.00	26.22	26.66	30.50	24.29
20567	43605	95039	23.00	21.00	26.00	26.50	23.64
20567	43578	95040	30.00	29.16	33.00	33.00	24.29
20513	43575	95041	22.00	22.83	33.00	33.00	24.29
20453	43572	95042	26.00	25.75	21.33	21.00	24.29
20560	43552	95043	32.00	29.16	33.00	33.00	24.29
20450	43588	95044	30.00	22.83	21.33	21.00	24.29
20528	43524	95045	31.00	26.22	26.34	26.50	23.64
20553	43541	95046	28.00	29.16	24.72	28.00	23.64
20454	43544	95047	27.00	26.22	26.66	30.50	24.29
20515	43567	95048	13.00	22.00	33.00	33.00	24.29
20500	43563	95049	34.00	22.00	26.34	26.50	23.64
20515	43561	95050	1.00	13.80	33.00	33.00	24.29
20511	43560	95051	19.00	13.80	26.34	26.50	23.64
20503	43545	95052	32.00	22.83	26.34	26.50	23.64
Error in prediction:							
		SUMS OF SQUARES =		1416.22	3490.62	3533.25	2752.10

Table 9.48a Comparison of the success in predicting COARSE SAND by Kriging of de-trended data and classical survey output at the 50 m grid, when TRENDS are accounted for by stratification based on mapping units.

NORTHING	EASTING	SITE NO	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING
20511	43547	95004	8.00	12.50	12.40	23.00	10.42
20513	43542	95024	6.00	4.25	12.40	23.00	10.42
20519	43560	95027	44.00	34.00	34.00	41.00	10.42
20502	43538	95028	3.00	4.11	4.41	4.00	4.84
20500	43542	95029	8.00	4.25	4.41	4.00	5.87
20494	43548	95030	5.00	4.11	4.41	4.00	5.47
20491	43538	95031	8.00	5.94	9.00	9.00	10.42
20487	43558	95032	6.00	5.94	9.00	9.00	10.42
20468	43557	95033	13.00	9.66	2.88	3.50	4.03
20468	43561	95034	3.00	3.00	2.88	3.50	3.02
20462	43544	95035	2.00	3.00	1.63	2.00	1.93
20461	43547	95036	2.00	3.00	1.63	2.00	1.84
20458	43560	95037	2.00	3.00	1.63	2.00	2.66
20454	43544	95047	1.00	3.00	2.88	3.50	1.08
20511	43560	95051	26.00	34.00	34.00	41.00	10.42
20503	43545	95052	6.00	9.66	12.40	23.00	10.42
Error in prediction:							
		SUMS OF SQUARES =		239.19	396.87	1161.75	1530.06
		(from the observed)					

Table 9.48b Comparison of the success in predicting COARSE SAND by Kriging of de-trended data and by conventional survey output at the 100 m grid, when TRENDS are accounted for by stratification based on soil mapping units.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING
20563	43584	95001	5.00	4.93	5.28	5.50	5.66
20540	43503	95002	6.00	39.35	4.90	5.00	46.78
20533	43571	95003	43.00	10.46	39.28	41.00	12.75
20511	43547	95004	8.00	3.82	10.75	11.50	6.56
20502	43512	95005	6.00	7.00	4.90	5.00	5.38
20466	43572	95006	6.00	5.51	7.00	6.00	31.48
20544	43572	95012	9.00	5.51	4.90	5.00	5.64
20531	43504	95013	8.00	5.51	4.90	5.00	4.16
20521	43495	95014	3.00	5.51	4.90	5.00	17.96
20532	43544	95015	7.00	10.46	10.75	11.50	8.49
20532	43590	95016	15.00	10.46	10.75	11.50	11.82
20524	43539	95017	11.00	5.51	10.75	11.50	2.90
20508	43491	95018	4.00	4.84	4.90	5.00	4.14
20500	43579	95019	5.00	4.84	5.28	5.50	4.14
20491	43582	95020	5.00	4.93	5.28	5.50	4.14
20461	43509	95021	3.00	7.00	3.52	2.00	5.95
20475	43572	95022	6.00	4.93	7.00	6.00	10.95
20455	43576	95023	6.00	39.35	10.75	11.50	38.78
20513	43542	95024	6.00	4.84	39.28	41.00	6.23
20519	43560	95027	44.00	4.93	4.90	5.00	6.28
20502	43538	95028	3.00	4.84	4.90	5.00	4.20
20500	43542	95029	8.00	5.51	4.90	5.00	7.33
20494	43548	95030	5.00	5.51	4.90	5.00	4.48
20491	43538	95031	8.00	7.00	4.90	5.00	4.14
20487	43558	95032	6.00	3.82	3.52	2.00	4.14
20468	43557	95033	13.00	3.82	3.52	2.00	4.14
20468	43561	95034	3.00	3.82	3.52	2.00	4.14
20462	43544	95035	2.00	3.82	3.52	2.00	4.14
20461	43547	95036	2.00	10.42	2.75	1.00	12.33
20458	43560	95037	2.00	7.00	6.00	8.00	5.41
20513	43575	95041	16.00	4.84	7.00	6.00	4.14
20453	43572	95042	5.00	10.42	5.28	5.50	4.14
20560	43552	95043	4.00	3.82	2.75	1.00	4.14
20450	43588	95044	8.00	4.84	3.52	2.00	4.14
20528	43524	95045	8.00	3.82	5.28	5.50	4.14
20553	43541	95046	5.00	10.46	2.75	1.00	25.97
20454	43544	95047	1.00	10.46	39.28	41.00	5.14
20515	43567	95048	33.00	39.35	6.00	8.00	27.59
20500	43563	95049	5.00	10.42	39.28	41.00	6.61
Error in prediction:							
		SUMS OF SQUARES =		5505.27	6370.93	6744.50	6973.24
		(from the observed)					

Table 9.48c Comparison of the success in predicting COARSE SAND by Kriging of de-trended data and conventional survey output at the 200 m grid, when TRENDS are accounted for by stratification based on soil mapping units.

NORTHING	EASTING	SITE No	OBSERVED	CLASS MEANS	MAP UNIT MEANS	TYPICAL PROFILE	KRIGING
20563	43584	95001	5.00	5.08	10.00	12.00	12.69
20540	43503	95002	6.00	5.08	5.73	5.50	8.78
20533	43571	95003	43.00	32.60	10.00	12.00	11.19
20511	43547	95004	8.00	15.91	5.73	5.50	8.30
20502	43512	95005	6.00	2.86	5.73	5.50	6.17
20466	43572	95006	6.00	7.00	2.20	2.00	4.33
20446	43558	95007	2.00	3.00	2.20	2.00	2.28
20415	43481	95008	2.00	2.86	2.20	2.00	2.44
20397	43501	95009	2.00	2.86	3.00	2.00	2.07
20382	43572	95010	15.00	6.50	10.00	28.00	18.44
20372	43563	95011	19.00	15.91	32.60	29.00	18.72
20544	43572	95012	9.00	5.60	5.73	5.50	11.48
20531	43504	95013	8.00	5.60	5.73	5.50	8.54
20521	43495	95014	3.00	5.60	5.73	5.50	8.22
20532	43544	95015	7.00	5.60	10.00	12.00	8.87
20532	43590	95016	15.00	15.91	10.00	12.00	13.09
20524	43539	95017	11.00	15.91	4.00	5.00	9.01
20508	43491	95018	4.00	5.60	5.73	5.50	6.82
20500	43579	95019	5.00	5.33	5.73	5.50	10.71
20491	43582	95020	5.00	5.33	5.73	5.50	10.47
20461	43509	95021	3.00	5.08	3.00	2.00	2.25
20475	43572	95022	6.00	5.08	3.00	2.00	4.55
20455	43576	95023	6.00	7.00	2.20	2.00	4.15
20513	43542	95024	6.00	5.08	5.73	5.50	8.23
20501	43559	95025	8.00	15.91	5.73	5.50	9.19
20481	43538	95026	6.00	5.08	3.00	2.00	3.62
20519	43560	95027	44.00	32.60	4.00	5.00	12.89
20502	43538	95028	3.00	5.33	5.73	5.50	7.80
20500	43542	95029	8.00	5.08	3.00	2.00	4.08
20494	43548	95030	5.00	5.33	3.00	2.00	4.19
20491	43538	95031	8.00	5.60	3.00	2.00	3.82
20487	43558	95032	6.00	5.60	3.00	2.00	4.38
20468	43557	95033	13.00	7.00	3.00	2.00	3.81
20468	43561	95034	3.00	2.86	3.00	2.00	3.94
20462	43544	95035	2.00	2.86	3.00	2.00	2.80
20461	43547	95036	2.00	2.86	3.00	2.00	2.86
20458	43560	95037	2.00	2.86	2.20	2.00	3.26
20418	43468	95038	2.00	2.86	3.00	2.00	2.77
20567	43605	95039	23.00	8.00	6.33	5.50	14.32
20567	43578	95040	5.00	5.33	4.00	5.00	8.98
20513	43575	95041	16.00	6.50	4.00	5.00	10.01
20453	43572	95042	5.00	7.00	2.20	2.00	3.65
20560	43552	95043	4.00	5.33	4.00	5.00	7.40
20450	43588	95044	8.00	6.50	2.20	2.00	5.04
20528	43524	95045	8.00	2.86	5.73	5.50	8.25
20553	43541	95046	5.00	5.33	10.00	12.00	8.72
20454	43544	95047	1.00	6.50	3.00	2.00	2.17
20515	43567	95048	33.00	15.91	4.00	5.00	10.89
20500	43563	95049	5.00	15.91	5.73	5.50	9.45
20515	43561	95050	34.00	32.60	4.00	5.00	11.21
20511	43560	95051	26.00	32.60	5.73	5.50	9.57
20503	43545	95052	6.00	6.50	5.73	5.50	7.89
Error in prediction:							
		SUMS OF SQUARES =		1416.85	5917.29	5776.25	3809.17

grouping mapping units. Hence the mean values from strata are expected to give less accurate predictions than means from individual map units. This also serves as an explanation for the relatively high predictive performance of mapping unit means as compared to that of Kriging.

In contrast, the semi-variograms of these two properties in the calcareous soils showed considerable spatial dependence contributing favorably to the accuracy in predictions by Kriging.

For more stable properties like Silt, Kriging was the best predictive technique at the 50 m grid (table 9.47a), which is the relevant scale where anisotropies showed up. It came close to class means at the 100 m and 200 m grids (tables 9.47b and c). The definition of soil classes and mapping units is made by exploiting the medium range structures of variability. These are revealed particularly at the 100 m grid spacing. Hence, at these scales mapping units and classes should be at their best as predictive tools. This argument and those stated above combine to explain the relative poor predictive performance of Kriging for all the four properties at the 100 m grid.

However, attention should be focused mainly on the results obtained at the 50 m grid, which is the scale at which anisotropies were revealed and where predictions by Kriging were not possible due to such type of variation. The results at this scale showed that partitioning the trends by stratification made it possible to obtain, for all anisotropic properties, predictions by Kriging. It is in this sense that partitioning trends into strata was useful.

On the other hand, in terms of accuracy of predictions, in the case of Potassium trend removal by this approach made little sense since Kriging predictions were not superior to those from conventional soil survey output (table 9.45a). This could have been anticipated by examining the semi-variogram within the stratum of calcareous soils (fig. 9.39a). It showed Pure Nugget Effect and the best Kriging predictions in this case are given by the mean of the stratum. It goes without saying that Potassium is a property most affected by management.

For Carbonate and Silt the partitioning of the trends by stratification was most useful. Kriging made better predictions

than soil classes and mapping units at the same scale (50 m grid) where there were no predictions at all before the stratification. Although, due to their nature, these results are preliminary, it can be said that partitioning the trends by stratification based on soil mapping units is a useful tool for dealing with spatial phenomena which depart from the assumptions of the intrinsic hypothesis, such as anisotropy and trends in the variation of soil properties, thus allowing for Kriging interpolation and improving predictions.

9.22 The semi-variogram as a reconnaissance tool when trends are accounted for.

In section 9.17 of this chapter the role of the semi-variogram as a reconnaissance tool when trends are ignored was discussed. Similarly, it can be assummed, for the sake of exploring the utility of a reconnaissance semi-variogram, that the semi-variograms calculated after trend removal by both procedures, i.e. stratification and modelling, were estimated in advance during pre-survey reconnaissance.

It can be assumed that trends can be detected or expected simply from the background geology and landscape information or from interpretation of the stereo pair of air-photographs. Also, if it is assumed that those trends are modelled or stratified in advance from reconnaissance sampling, and with the same reconnaissance data semi-variograms of de-trended data are calculated. Then the role of such semi-variograms as a tool for decision making and survey design can be explored once the trends have been accounted for.

The semi-variograms of residuals from trend (figs. 9.35 to 9.38) and from strata after trend partition (figs. 9.39 to 9.46) can be considered then as calculated from reconnaissance work, and inferences from the interpretation of their structures can be made to guide the choice of the best tool or combination of tools for predicting missing information.

9.22.1 The semi-variogram of residuals from trend as a reconnaissance tool after trend modelling by bicubic spline surfaces.

It can be said in general that the semi-variograms of residuals from trend are not particularly suitable to be used as tools for reconnaissance work. They only permit to infer, on the basis of their structures, how accurate the predictions of the residuals by Kriging can be. However, in almost all instances the residuals are only a minute component of the predicted values of the property which result from the addition of the interpolated residuals to the trend values at the target points. Hence, since the predictions at the random check points are defined mostly by the values of the surface and with only a fairly small contribution from the residuals, then the utility of reconnaissance semi-variograms of residuals from trend is rather limited.

Bearing this in mind, however, it can be seen that the semi-variograms of residuals of Potassium (figs 9.35a-c) would have permited to infer in advance that the predictions after trend modelling and Kriging residuals for this property would have been the best of all the other three properties. All three semi-variograms show spatial structure, but particularly the semi-variogram at the 100 m grid (fig. 9.35b) shows a fairly small Nugget variance. This could have served to infer better predicted residuals by Kriging and this reflects itself in the final predicted values after addition to the trend at the same points. Indeed, table 9.41b indicates that although Kriging predictions were the least accurate for all properties and scales, in this case Kriging was closest to the best predictive technique than at any other scale and property.

No more could have been inferred and said confidently about the predictive success of Kriging of de-trended data, or of soils survey output, by taken the rest of the semi-variograms of residuals as reconnaisance tools. Most of them display total absence of spatial dependence making it necessary to add the means of residuals, and not the interpolated residuals, to the

values of the trend at the check points. In these circumstances the success in prediction of values of the properties depends even more on the success of the trend surface fit, which is not directly related to the structures of the semi-variograms of residuals. Again, the use of a semi-variogram of residuals from trend as a reconnaissance tool proved rather limited.

9.22.2 The within-stratum semi-variogram, after trend partitioning, as a reconnaissance tool.

Assuming that the within-stratum semi-variograms were calculated in advance from pre-survey reconnaissance sampling, there are now two semi-variograms,i.e. BC-BE and NS-NG strata, for each sampling grid, and the predictive success of Kriging at a such sampling scale depends on the structures of both semi-variograms at that scale. Hence they have to be used together to infer whether Kriging will make the best predictions ar whether the best predictions will come from soil classes or mapping units.

Since the quantity of semi-variograms to interpret (figs. 9.39a-c to 9.46a-c) was considerable, the final outcome of the inferences made based on their structures has been summarized and compiled in table 9.49. The decisions that a surveyor would reach are displayed in table 9.49 in terms of "Map" or "Krig" meaning in the former that the surveyor would prefer to make a map for he judges it would make the best predictions, or in the latter, he would interpolate by Kriging for he expects this technique to make the best predictions.

For the purpose of drawing inferences the decisive structures of the reconnaissance semi-variogram that were used are whether the property shows spatial dependence or not, if so, the size of the Nugget variance and its relation to the size of the variance of spatial dependence (Sill - Nugget), the degree of the scatter of the points around the curve of the semi-variogram model, the goodness of the model fit and the range of influence. These spatial structures were examined in all reconnaissance semi-variograms (figs. 9.39a-c to 9.46a-c) to decide between techniques on the basis of their likely

predictive success. So, only when it is of real merit, the interpretation of these structures for any given case are discussed in more detail.

Table 9.49. Decisions reached by a soil surveyor based on the structures of within-strata semi-variograms used as reconnaissance tools. The strata result from grouping mapping units for partitioning trends.

Property	Sampling Grid (metres) 50		100		200	
	BC-BE	NS-NG	BC-BE	NS-NG	BC-BE	NS-NG
Potassium	Map	Krig	Krig	Map	Krig	Map
	<Map >		<Krig>		< Krig >	
Carbonate	Krig	Map	Krig	Map	Map	Map
	<Krig>		<Krig>		<Map>	
Silt	Krig	Krig	Krig	Krig	Map	Map
	<Krig>		<Krig>		<Map>	
Coarse Sand	Map	Krig	Krig	Map	Map	Krig
	<Krig>		<Krig>		<Krig>	

< > Indicates the final decision of the two from both strata.

Although the interpretations of the reconnaissance semi-variograms in table 9.49 have a strong subjective component, nevertheless, the inferences that a soil surveyor can make based on the same tools cannot be devoid of a great contribution from intuition. In this sense these results can be said to be representative of a typical situation in practice.

To illustrate the reasoning behind the final choice between the two inferences, Potassium at the 50 m grid serves the purpose. The BC-BE semi-variogram of this property (fig 9.39a) showed no spatial dependence and the best estimate therefore is the mean from the stratum. In such a case the mean from mapping units are preferred since the stratum is generated from grouping mapping units and is likely to yield less accurate predictions. So, the surveyor would prefer to make a map as oposed to interpolate (see table 9.49).

The semi-variogram from the NS-NG stratum (fig 9.40a) showed that this property has spatial structure in these soils,

and the spherical model fits well. The surveyor would infer that Kriging will give the best predictions and that would be his choice. However, the relatively large Nugget variance as compared with the variance of spatial dependence in this semi-variogram makes a soil map the best choice (table 9.49). This sort of arguments were used for all the outcomes displayed in this table. It can be shown that each choice in table 9.49 has been made based on premises of this kind.

The inferences in table 9.49 were matched against the actual results of the comparison of predictive techniques after trend stratification which show what predictive technique was the best and are given in tables 9.45a-c to 9.48a-c. The correctness of each choice based on the reconnaissance semi-variograms (table 9.49) was recorded in terms of "right" or "wrong". These records are given in table 9.50.

It can be seen in table 9.50 that the efficacy of the within-stratum semi-variogram as reconnaissance tool was of 50%. The failures occured mainly in the inferences made based on the semi-variograms at the 100 m grid for all the properties. This is precisely the scale at which all the predictive procedures compared, ignoring trends, gave similar rates of predictive success, and it does not make much difference which of those procedures is employed for prediction. This scale is also that which reveals the medium range structures of variation and which losely corresponds to soil classes and mapping units. This is of particular significance for in all cases at the 100 m grid it was inferred that Kriging would make the best predictions (table 9.49). Yet, tables 9.45b, 9.46b, 9.47b and 9.48b indicate that the best predictors at this scale are means from soil classes and mapping units.

Such evidence suggests that, indeed, there may be a correspondence between medium range structures of variation and soil classes. This correspondence, if existing, enhances the definition of classes and the delineation of mapping units at the 100 m grid. Arguably, these factors contribute to the superiority of conventional survey output over Kriging for predicting soil missing information at this scale, despite the spatial structure in the variation of the properties within the

strata as depicted by the within-stratum semi-variograms (figs. 9.39b, 9.41b, 9.43b, 9.44b and 9.45b) at this scale.

Table 9.50. Correctness of the inferences made to chose the best predictive technique based on the within-strata semi-variograms when used as reconnaissance tools at Yelford.

Soil Property	Inter-sample spacing (metres) 50	100	200
Potassium	R	W	R
Carbonate	R	W	R
Silt	R	W	R
Coarse Sand	W	W	W

R = RIGHT W = WRONG

Thus, on the whole, it is not easy to make the correct inferences from simultaneous interpretation of reconnaissance semi-variograms calculated within the strata resulting from partitioning the trends. Some of the structures of the semi-variograms within individual strata turned out to be misleading, and the final choice of technique from the two candidates from strata is to a great extent subjective and difficult.

Nevertheless, the results tend to suggest that for the short range structures of variation (50 m grid) which are the structures exploited by Kriging, this type of reconnaissance semi-variogram after trend stratification, may be useful (see table 9.49 for Potassium, Carbonate and Silt, and table 9.50 for the same properties at the 50 m grid) in designing the survey strategy for best spatial predictions of the soil information that is missing.

9.23 Concluding remarks to this chapter.

i. Computerized soil information sysytems are an improved alternative to conventional soil survey methods to store and manipulate soil data. They are a powerful tool for prediction of the missing soil information necessary for

land-use planning.

ii. The use of a Soil Information System (OXSIS) and its built-in capabilities for predicting missing soil information has been demonstrated by its application as the tool for the computations and data manipulation involved in these trials.

iii. The Yelford area exhibits a complexity of variation in soils and an array of hierarchical sampling structures and wealth of data which presented a range of problems regarding soil variability and prediction and which may be common to ordinary survey practice. This also allowed the exploring of a range of techniques and procedures for their solution.

iv. The semi-variogram is an excellent tool for investigating and elucidating the structures of soil spatial variability. This aspect of its use has been sufficiently demonstrated in these trials.

v. Trends are commonplace in soil variability and they were associated in the Yelford area with its medium range structures.

vi. In the Yelford area the presence of trends was so evident that they were inferred from background information and landscape associations. In other contexts it may not be clear when trends are sufficiently strong as to merit forms of interpolation other than ordinary Kriging.

vii. The anisotropic variation of Potassium, Carbonate, Silt and Coarse Sand at the scale of the 50m grid was induced by strong trends. These run in the direction of the geological drift (NE-SW), and were detected quite evidently in the semi-variograms of these properties at that scale.

viii. The fit of elliptical functions to data from slopes of directional semi-variograms to obtain models for anisotropy did not provide an adequate solution on both statistical and natural grounds, to the problem of anisotropies.

ix. The presence of anisotropies evident at 50m of inter-samples spacing precluded the estimation of the

parameters of the semi-variogram, and the applicability of ordinary Kriging. The best predictive procedure of the properties affected in these circumstances was the class mean.

x. The best predictor of pH at all sampling scales were the mapping unit means.

xi. When trends were ignored, Kriging was the best predictor of Potassium at the 100m and 200m grids, Magnesium at the 100m grid, Carbonate at 100m and 200m grids, Clay at all grids, Silt at the 100m grid and Organic Matter at all sampling grids. Coarse Sand was best predicted by map unit means and class means.

xii. Ignoring trends affected negatively Kriging predictions, but even so, on the whole, Kriging made the best predictions.

xiii. Predictions by the typical profile in the mapping unit were the poorest of all the procedures compared. But the comparison was unfair to the typical profile in that its predictions were only based on one observation.

xiv. In general, the results from the comparison suggest that where the semi-variogram is isotropic, shows clear spatial dependence and its structures are clearly revealed and well modelled, Kriging is definitely the best predictive tool.

xv. The efficacy of conventional soil survey output in predicting values of soil properties at the 100m grid is quite similar to that of Kriging, and it makes no much difference which of those procedures is used for predictions. Survey output was at its best at this scale at which the medium range structures of variation are revealed and correspond loosly to soil classes.

xvi. Success in prediction by Kriging decreased with increasing sampling spacing from 50m to 100m grid for most properties. There is little point in increasing sampling effort from 200m to 100m sampling spacing since this brings no gains in accuracy of predictions.

xvii. For a limited budget 200m appears to be the optimal sampling spacing, but Kriging is still feasible even at 1km of inter-sample distances for some properties. If

a greater survey effort can be afforded, the 50m spacing is the best for Kriging predictions when variation is isotropic.

xviii. The success in prediction by the procedures compared proved to be related to the structures of the semi-variogram, in particular, to the form of the monotonic increase and to the size of Nugget variance which was a good indicator whether Kriging would be better than conventional survey procedures.

xix. In a practical context, the decision between making a map or interpolating by Kriging to obtain the best predictive tool is related to the structures of a carefully planned reconnaissance semi-variogram.

xx. Diagnosis work about the presence of anisotropies and obedience of stationarity assumptions, i.e. significant trends, can be carried out by a reconnaissance semi-variogram, helping also decision-making concerning survey design for best predictions.

xxi. A critical point, in practice, is how best to sample in order to estimate well a reconnaissance semi-variogram for designing surveys which will furnish the best predictive tools.

xxii. There are no established procedures or guidelines for interpreting reconnaissance semi-variograms and drawing inferences from their structures. However, they cannot go without paying due attention to the landscape and its relation to soil variability.

xxiii. Trends in the variation of soil properties need to be accounted for to improve predictions by the application of Kriging. Procedures for dealing with non-stationarity caused by trends, i.e. detailed structural analysis and universal Kriging are too complicated to be used by an average surveyor or user wishing to interpolate.

xxiv. It has been shown that trends can be adequately accounted for by their partition into strata or by mathematical modelling with surfaces.

xxv. The surfaces fitted by bicubic splines were particularly suitable for modelling the trends and for removing anisotropic variation present at the 50m grid as shown by

the semi-variograms of the residuals from trend.

xxvi. It can be shown that if the surface would have been given in advance, some features of the semi-variogram of the same data could have been predicted. This is particularly true for the surfaces fitted at the 100m grid where the pattern characteristic of certain types of variation can be recognized, e.g. the pattern typical of transition phenomena spatially auto-correlated.

xxvii. Although only in a few instances Kriging was a better predictor than class means after trend modelling by bicubic spline surfaces, trend removal by this method was successful in that not only it allowed for Kriging where it was not possible before modelling, but also Kriging made the best predictions of Potassium and Silt at the 50m grid were anisotropic variation precluded any estimates by spatial interpolation.

xxviii. The explanation of the predictive performance of Kriging could not be related to the structures of the semi-variograms of residuals from the trend since they are only a small part of the predicted values. In only half of the cases the residuals were spatially auto-correlated and in the rest they showed a tendency to vary randomly.

xxix. The semi-variograms of residuals from the trend have a very limited value as a reconnaissance tool, for they only permit inferences about predictions of residuals themselves.

xxx. Partitioning the trends by dividing the landscape in terms of calcareous and non-calcareous soil mapping units proved successful in removing anisotropic variation allowing for Kriging interpolation where it could not have been possible without stratification.

xxxi. After partitioning the trend into strata Kriging was not the best technique in all situations, but, however it gave the best predictions of Carbonate and Silt at the 50m grid, where anisotropies were present, and for Potassium at the 200m grid. In this sense this way of accounting for the trends was useful in prediction. In most other comparisons Kriging made the second best

predictions after those of soil classes.

xxxii. The predictive success of the procedures compared was related to the structures of the within-strata semi-variogram.

xxxiii. Although these results are not conclusive, problems of anisotropy and non-stationarity caused by trends may be tackled by removing the trends by any of the procedures explored in this thesis, so that interpolation can take place.

xxxiv. Of the two methods for trend removal explored, partition of the trend by division of the soil landscape based on mapping units is a simpler, cheaper and gives moderately better results in the goodness of predictions of values of soil properties by Kriging than modelling of the trends by bicubic spline surfaces.

xxxv. The structures of the within-stratum semi-variogram can be used for reconnaissance work within the limits of the stratum. To infer which predictive technique will be the best for the whole area, the final decision has a strong subjective component and a great contribution from intuition.

xxxvi. The use of soil mapping units as a basis for the division of the landscape and partitioning of the trends, combined with spatial analysis by the semi-variogram and Kriging interpolation, demonstrate that mathematical techniques and intuition can be combined to advantage to solve the problem of missing information.

xxxvii. These trials have demonstrated that mathematical interpolation by Kriging is, indeed, a superior technique for spatial prediction provided certain conditions of variability and sampling effort exist. Such conditions are very fragile. In a practical situation, where the presence of other disturbing phenomena of such conditions can be anticipated by a carefully planned reconnaissance semi-variogram, Kriging can be convenintly supported by maps or other forms of intuitive output.

X. CONCLUSIONS.

This thesis is essentially on soil data and on the provision of soil information necessary for land-use planning. It examined the nature of soil variability and the conventional procedures for acquiring, storing and providing soil information giving account of its complex variability in all directions. This work also examined the role of the new alternative procedures based on computer technology for storing, retrieving and providing soil information. The state of the art in computerized soil information systems was reviewed and the structure and capabilities of a soil information system developed at Oxford were introduced. The contribution to the development of such a system in terms of a package for data manipulation capabilities was also presented.

A range of fundamental problems related to the distribution of soils, and to providing soil information by means of generation and prediction, were analyzed and procedures and techniques for their solution explored.

The work in this thesis was essentially oriented towards solving the problems of predicting or generating the soil information necessary for land-use planning and that was not recorded from survey and considered as missing. The best combination of tools and techniques for solving this problems are proposed. The findings of this work can be summarized in the following general conclusions:

1. **On the importance of soil data for land-use planning.**

Soil data is essential for land-use planning. Current procedures for land suitability assessment are based on data of physical attributes of the land, and soil data is the core of such resource base. On the other hand, land suitability assessment for a given use can only be performed at sites where soil information has been recorded. It is the available point-data (i.e. sites) that are apprised. Hence, the importance of counting with soil information from the density of point-data required to expand the assessment to cover the whole area of interest. The information

at unvisited sites that need to be apprised can be predicted from the recorded sites by interpolation.

2. On soil variability.

The soil mantle is a rather heterogeneous body which varies in all directions at different rates and at multiple spatial and temporal scales. Soil variability is complex and many are the sources that determine the complexity of the spatial distribution of soils. Yet, these can only be observed and their properties measured at a limited number of sites. Soil variability makes predictions of soil properties at unvisited sites necessary. The improvement of predictions requires the elucidation of components of variability.

The semi-variogram is an excellent tool for revealing the spatial structures of soil variability and improving predictions. This has been demostrated with the semi-variograms of the soil properties studied in the Yelford area.

The components of soil variability detected in this study were three: a spatially dependent component, a random component, which includes variation within the shortest sampling interval, and a systematic component or trend. The elucidation of these components was a function of scale and density of observations.

3. On the tools for providing soil information.

Soil survey output, i.e. soil map, its legend and the survey's memoir, have been the conventional way of transmitting soil information to the user. These are fixed products in the sense that once decisions have been made about them these decisions are fossilized and it is not possible to revert back for changes. A considerable amount of valuable original information from point-data is lost by generalization, which is inherent in classificaton and mapping. The soil map is an inadequate retrieval system for the volume of point-data collected during survey.

Computerized soil information systems (SIS) are a better alternative to conventional soil survey output for providing the soil information needed for planning. SIS can perform the same tasks as in routine survey but faster and even better. SIS provide excellent capabilities for information storage and retrieval. More importantly, SIS retain the ungeneralized point-data with the values of soil properties as obtained from survey. This is new, and it represents an enormous advantage over conventional soil

survey output. Expensive information is not lost and ungeneralized point-data can be re-apprised or re-classified enabling updates, changes and improvements on past decisions.

Some of the advantages of SIS have been demonstrated in this thesis by retrieving ungeneralized point-data from the data bank of OXSIS which could not have been obtained from maps or survey memoir and performing manipulations and data processing with a flexibility which would not have been possible without a computerized soil information system.

The aim of any soil survey or any information system is to provide means for predicting information for any property, at any depth and at any site within the area surveyed. Its sucess depends on the extent to which these predictions are correct.

It has been demonstrated, with the development of the software package for the Manipulation stage in the information system at Oxford (OXSIS), that the inclusion of built-in capabilities in SIS for new geostatistical techniques for spatial analysis of soil data and for mathematical interpolation, makes SIS a powerful tool for prediction of soil properties at any site and any depth within a surveyed area, by exploiting on-line the point-data stored in OXSIS. These are capabilities which the conventional soil survey tools for providing soil information do not posses and which make computerized soil information systems superior.

With the development of the Manipulation package of OXSIS, users are now equiped with a powerful "user-friendly" tool for predicting values of soil properties at any site and depth in an area.

4. On the problem of missing soil information.

The total soil information in an area can never be known. Only a minimal part of the total soil information can be provided by soil survey. Information at unvisited sites, unsampled depths, and for unrecorded properties is considered as missing information.

4.1 Missing wanted soil property.

It has been shown that the values of missing soil properties can be generated by means of functions of the missing property on other recorded properties acting as its surrogates. The surrogate functions can be estimated by multiple regression analysis or curve fitting based on least squares techniques. There are a sufficient number of statistical procedures for constructing a surrogate

function. A strategy for the best use of these procedures to tackle a variety of problems during the model development stage was proposed in this thesis as an approach to the solution of this problem.

4.2 **Missing records at the wanted depth**.

This problem was solved by depth interpolation of values by means of a suitable algorithm based on equal-area spline curves.

The equal-area spline algorithm proved the best method, so far, for reconstructing soil depth functions which are used for interpolation of values of either, soil properties or soil depths in a range of formats. A package of software for depth functions by the equal-area spline solution was developed and included in the "user-friendly" Manipulaton package of OXSIS.

4.3 **Missing records at wanted sites**.

The provision of missing information at unvisited sites requires necessarily of the existence of some point-data. The density of point-data, determined by the inter-sample interval, influences to a large extent the predictive tool that is to be used to provide the information missing. This is so, since the nature of the structures of variation that can be revealed by spatial analysis are a function of the shortest inter-sample spacing.

Thus, the provision of soil information at unvisited sites by means of spatial prediction is seen in the context of survey effort in terms of inter-sample spacings.

It is worth mentioning that these conclusions can only reflect the exploratory character of this study.

In general, for a sampling effort corresponding to a relatively high density of point-samples, (e.g. 50m of inter-sample spacing or less), if the semi-variogram of the property is isotropic, shows spatial dependence, its structures are clearly revealed and a model fits well, spatial interpolation by Kriging is definitely the best technique for predicting the missing values of soil properties at any site in the area covered by the data.

For the same sampling effort, when variation of the property to be predicted is anisotropic and the parameters of one of the isotropic models cannot be estimated, or else the presence of strong trends is evident then, if no measures to account for the anisotropies or the trends are taken, the mean value of the class or the mapping unit which the site belongs to, provide the best

predictions of the information missing since interpolation by Kriging is not possible in these circumstances.

It was not clear from the evidence obtained in this work whether the structures of a reconnaissance semi-variogram which showed anisotropic variation would indicate adequately from the point of view of improving predictions, whether there is any sense in accounting for the anisotropies by removing the trends and to interpolate by Kriging of de-trended data, or whether to, simply, ignore the trends and rely on means from classes or mapping units.

The results obtained in this work offered split evidence. For some properties an improvement in predictions over those of conventional survey output can be achieved by accounting for the trends and predicting by Kriging of de-trended data. In contrast, for some others the predictions by Kriging of de-trended data did not achieve any improvement and the effort spent was, in this sense, wasted. The attempts to fit models for anisotropy by the use of elliptical functions will not work for all possible situations of variability. This approach was considered unsound on statistical grounds, and in situations where it ignores the natural distribution of the soils in the area.

However, if it is decided to account for the anisotropies by trend removal, the boundaries of soil mapping units can be used for partitioning the trends and for stratification of anisotropic variation. This approach proved to be better than trend modelling from the point of view of predictions.

After stratification, the within-stratum semi-variogram should show isotropic variation at the same sampling scale where it was anisotropic, allowing for Kriging interpolation.

At this stage, the structures of a reconnaissance semi-variogram within each stratum will provide some guidance as to whether to predict by means of Kriging interpolation or whether to use the mean of the class to predict the information needed.

For a survey effort which gives moderate sampling densities (e.g. 100m inter-sample spacing), predictions by conventional soil survey output can be almost as accurate as those by Kriging, if the medium range structures of variation correspond loosely to soil classes, or if they correspond loosely to soil mapping units when there are clear external expressions of the soil/landscape relationships.

The structures of a reconnaissance semi-variogram may or may not reflect this association, and intuition comes to play an important role here.

Kriging would still predict best the missing information, but means from soil mapping units or classes could provide just as good predictions too, at this scale. The criterion as to whether to use one or the other should not be, in situations like this, a matter of predictive accuracy.

For a sampling effort corresponding to medium to low densities of observations (e.g. 200m inter-sample spacing), the Kriging technique for most properties predicts with more accuracy the values of soil properties than conventional survey output, if the semi-variogram is isotropic and shows strong spatial dependence. If spatial dependence is weak, as shown by the small size of the variance of spatial dependence as compared to the Nugget variance in a reconnaissance semi-variogram, or the property changes in a long gradual transition, e.g. as described by an exponential model or by Pure Nugget Effect, it might be better to use means of soil mapping units or classes to provide the values missing. The evidence in this work showed that when variation is of that kind at this scale, conventional soil survey output predicted better than Kriging.

Thus, it has been shown that spatial analysis by the semi-variogram and mathematical interpolation by Kriging are fundamental tools for solving the problem of predicting the missing information at sites unvisited during survey or at any site within the survey area.

It has also been demonstrated that in situations where the occurrence of certain spatial phenomena invalidate the assumptions necessary for the applicability of Kriging, intuition, in the form of boundaries of soil mapping units and class concepts can be successfully combined with mathematical techniques to improve the goodness of the predicted information.

5. <u>On the trials and the comparison of predictive procedures</u>.

The trials in spatial prediction of soil properties in this thesis were carried out in exceptional conditions of wealth of data. It would be difficult, in practice, to count on such an abundance of point-data and at varying sampling densities. The results, however exploratory, provide valuable information about

the comparative predictive value of the various techniques explored for solving the problem of missing information.

Kriging is the best predictive technique, but its success depends on the structures of the semi-variogram. In situations where the applicability of Kriging is either, not possible, or where spatial dependence is weak or the Nugget variance is considerable, class means and mapping unit means, in that order, are the best tools for predicting the missing soil information, unless measures are taken to bring about the conditions that make Kriging applicable.

At sampling scales which reveal the medium range structures of variation class means an mapping unit means, in that order, provide as good predictions as Kriging, provided those structures are reflected somehow in the definition of the classes or in the delineation of mapping units boundaries.

The values extracted from the typical profile in the mapping unit make poor predictions of soil properties at unvisited sites within the same mapping unit. However the comparison is unfair to this tool since the predictive values are obtained from only one site.

6. <u>On the semi-variogram</u>.

It has been demonstrated that the semi-variogram is an excellent tool for elucidating the structures of soil spatial variability. Since the information on those structures, in the form of parameters, enters Kriging interpolation, the success in prediction of this technique bears a strong relation to them. These structures may also provide some explanations about the predictive success of soil classes and mapping units, particularly when the sampling spacing allows to discover the medium range structures of variation.

It has been shown that, other things being equal, to a considerable extent, the decision between chosing Kriging interpolation or a soil map and its legend to predict the soil data needed, is related to the size of the Nugget variance and to the proportions of the total variation which are attributable to spatial dependence variance and to random variation or Nugget variance.

The semi-variogram is a most useful reconnaissance tool. It permits detection of preferential directions of regionalization of

the property which give anisotropic variation, and the presence of strong trends which preclude the applicability of ordinary Kriging unless measures are taken to account for them.

If it is well estimated, with a rate of success of about 80% a reconnaissance semi-variogram will enable to infer correctly, based on its structures, which predictive tool, i.e. Kriging or a soil map and soil classes, will yield the best predicted values of soil properties at any point within the field within which the semi-variogram was estimated. This figure needs reconsidering in a practical situation since the semi-variograms calculated at Yelford and assum as if from reconnaissance sampling, were estimated from a wealth of point-data which is very unusual for pre-survey sampling.

The rate of inferential success of the reconnaissance semi-variogram calculated within the strata resulting from partitioning trends based on soil mapping units, decreases down to 50%. This relatively low rate may be due to a poorer estimation of the semi-variogram from a smaller number of samples, but, more fundamentally, to the subjective element in choosing the correct inference from two alternative made from semi-variograms of different strata. In any case, it is not easy to make simultaneous interpretation of two reconnaissance semi-variograms each from individual strata.

The semi-variograms of residuals from trend, after trend modelling by spline surfaces, are not useful for reconnaissance and inferential work since they can only be interpreted in terms of residuals from trend.

There are no established procedures or guidelines as to how to interpret a reconnaissance semi-variogram based purely on its structures. Since information about the landscape must be included during any interpretation, there is a strong intuitive element which may preclude reproducibility of interpretations.

Owing to the fact that the semi-variogram as tool for reconnaissance, plays an important role in selecting the best predictive technique of missing information and consequently in survey design, a critical point for future work is how best to sample in order to estimate as well as possible, and with the minimum of samples, the structures of a reconnaissance semi-variogram.

7. On dealing with anisotropies and trends.

The fitting of models for anisotropy by the use of elliptical functions on the slopes of directional semi-variograms was found unsound on statistical grounds and, for situations of landscape and variability similar to that of the Yelford area, incongruous with the natural distribution of soils.

Anisotropic variation in the form of preferential directions of regionalization can be induced by strong trends which may be originated by geological phenomena like drifts.

It is possible, from a careful study of the landscape and background information to anticipate the likelihood of the presence of anisotropic variation or of trends. But a carefully estimated reconnaissance semi-variogram should reveal or confirm their ocurrence.

Accounting for the trends by both modelling with surfaces and by partitioning the trends into strata, removed the anisotropic variation of properties. This is evidenced by the isotropic semi-variograms of de-trended data.

Partitioning the trends by stratification based on soil mapping units proved to be a better method for accounting for the trends and anisotropies than the modelling of the trends by a sophisticated bicubic spline surface fitting. This judgement is based on the goodness of predictions after de-trending by both methods, on simplicity and on the effort involved in each procedure. Modelling the trends by bicubic spline surfaces, since it is an iterative process, is lengthy and expensive.

It is not clear how strong a trend has to be in order to challenge the assumptions of the intrinsic hypothesis, invalidating the application of Kriging.

8. On the combination of Kriging interpolation and intuition.

The use of the intuitive ouput contained in the boundaries of soil mapping units as a basis for division of the landscape and partitioning of the trends, combined with spatial analysis by the semi-variogram and Kriging interpolation after de-trending, demonstrate that mathematical techniques for interpolation and intuition can be combined to advantage to solve the problem of missing information. It is not unlikely that other forms of combination of these tools could improve predictions even further.

9. About the optimal sampling spacing for prediction.

For a limited budget and sampling effort, 200m appears to be the optimal sampling spacing. There is little point in increasing sampling effort from 200m to 100m of inter-samples spacings since this brings no gains in accuracy of predictions.

For a less restricted budget, if a sampling effort of 50m of sampling spacing can be afforded, this is the optimal sampling spacing.

10. On the area of the trials.

The area selected for the trials, i.e. the Yelford area, exhibits a complexity of variability and a wealth of hierarchical structures of data wich permited to be confronted with a range of problems regarding soil variability which may be common in ordinary survey practice. This also allowed to explore the applicability and virtues of a range of techniques for their solution.

11. Solution.

Finally, the procedures and techniques proposed and applied to the data from the Yelford area permited to, sufficiently demonstrate, that the problems of missing information can be solved, and that the information required for land evaluation and land-use planning can be provided.

REFERENCES

Ahlberg, J.H., Nilson, E.N., and Walsh, J.L. (1967) The theory of splines and their applications. Academic Press, N.Y. and London.

Akaike, H. (1974) A new look at statistical model identified IEEE Transactions on automatic control. Vol. AC-19:716-723.

Alagic, S., Circic, M. and Zaciragic, B. (1984) Relational Pascal Model of soil database. In: Soil Information Systems Technology, Proceedings of the Sixth Meeting of the ISSS Working Group on Soil Information Systems. Bolkesjo, Norway. 28 February - 4 March 1983, PUDOC, Wageningen, The Netherlands.

Babalola, O. (1978) Spatial variability of soil water properties in tropical soils of Nigeria. Soil Sci. 126:269-79.

Banister, P., Tomlinson, P.R., and Beckett, P.H.T. (1974) A simple procedure for soil mechanical analysis. Pl. Soil 40:211-15.

Bauer, H.P., Beckett, P.H.T., and Bie, S.W. (1972) A rapid gravimetric method for estimating calcium carbonate in soils. Pl. Soil 37: 689-90.

Bayane, S., and Legros, J.P. (1978) Soil data processing international network project (SDPIN). In: Developments in Soil Information Systems, Proceedings of the Second Meeting of the ISSS Working Group on Soil Information Systems. Varna/Sofia, Bulgaria. 30 May - 4 June 1977. PUDOC, Wageningen, The Netherlands.

Beckett, P.H.T. (1975) The cost-effectiveness of terrain evaluation. Vol. 2. Review of field work 1971-74. Report on Contract No. DAJA-37-R-0029. For U.S. Army Research Office.

Beckett, P.H.T. (1981) The cost-benefit relationships of soil surveys. In: Soil Resource Inventories and Development Planning. Proceedings of Workshops at Cornell University, 1977-78. Technical Monograph No. 1. Soil Management Support Services. Soil Conservation Service, U.S.D.A, Washington, D.C.

Beckett, P.H.T. (1984) Soil Information Systems: The problem of developing a capacity of reapprising data for single purposes. In: Soil Information Systems Technology, Proceedings of the Sixth Meeting of the ISSS Working Group on Soil Information Systems. Bolkesjo, Norway. 28 February - 4 March 1983, PUDOC, Wageningen, The Netherlands.

Beckett, P.H.T., and Burrough, P.A. (1971). The relation between the cost and utility in soil survey. V. The cost effectiveness of different soil survey procedures. J. Soil Sci. 22:481-489.

Beckett, P.H.T., and Webster, R. (1971). Soil variability: a review. Soils Fert. 34:1-15.

Beckett, P.H.T., Webster, R., McNeil, G.M., and Mitchell, C.W. (1972) Terrain evaluation by means of a data bank. Geogrl. J. 138:430-56.

Beek, K.J. (1978) Land evaluation for agricultural development. International Institute for Land Reclamation and Development (ILRI). Publication No. 23. Wageningen, The Netherlands.

Behar, A., and Evtimov, K. (1978) Clay minerals in soils and relevant analytical results as components of a specialized data bank. In: Developments in Soil Information Systems, Proceedings of the Second Meeting of the ISSS Working Group on Soil Information Systems. Varna/Sofia, Bulgaria. 30 May - 4 June 1977. PUDOC, Wageningen, The Netherlands.

Bidwell, O.W., and Hole, F.D. (1964) Numerical taxonomy and soil classification. Soil Sci. 97:58-62.

Bie, S.W. (1984) Soil data in digital space. In: Soil Information Systems Technology, Proceedings of the Sixth Meeting of the ISSS Working Group on Soil Information Systems. Bolkesjo, Norway. 28 February - 4 March 1983, PUDOC, Wageningen, The Netherlands.

Bie, S.W., and Beckett, P.H.T. (1970) The costs of soil survey. Soils Fertil. 32:203-17.

Bie, S.W., and Beckett, P.H.T. (1971a) Quality control in soil survey introduction: I. The choice of mapping unit. J. Soil Sci. 22, 1:32-49.

Bie, S.W., and Beckett, P.H.T. (1971b) Quality control in soil survey: II. The costs of soil survey. J. Soil Sci. 22,4:453-465.

Bie, S.W., Lieftinck, J.R.E., van Lynden, K.R., and Waenink, A.W. (1976) Computer-aided interactive soil suitability classification: A simple Bayesian Approach. The Netherlands. J. Agric. Sci. 24:179-186.

Bie, S.W., and Schelling, J. (1975) Soil Information Systems in 1975. In: Soil Information Systems, Proceedings of the Meeting of the ISSS Working Group on Soil Information Systems. PUDOC, Wageningen, The Netherlands.

Bie, S.W., and Schelling, J. (1978) An integrated information system for soil survey. In: Developments in Soil Information Systems, Proceedings of the Second Meeting of the ISSS Working Group on Soil Information Systems. Varna/Sofia, Bulgaria. 30 May - 4 June 1977. PUDOC, Wageningen, The Netherlands.

Biggar, J.W. (1978) Spatial variability of nitrogen in soil. In: Nitrogen Behaviour in Field Soils (eds. D.R. Nielson and T.G. Macdonald). Vol. 1, pp. 201-11. Acad. Press, New York, San Francisco, London.

Bluhm, G., Carlis, J., and Decker, G. (1984) Going database in IRIS: a status report. In: Soil Information Systems Technology, Proceedings of the Sixth Meeting of the ISSS Working Group on Soil Information Systems. Bolkesjo, Norway. 28 February - 4 March 1983, PUDOC, Wageningen, The Netherlands.

Burgess, T.M. (1979) Optimizing procedures for soil survey at large to medium scales. D.Phil. Thesis, University of Oxford.

Burgess, T.M., and Webster, R. (1980a) Optimal interpolation and isarithmic mapping of soil properties. I. The semi-variogram and punctual kriging. J. Soil Sci. 31:315-331.

Burgess, T.M., and Webster, R. (1980b) Optimal interpolation and isarithmic mapping of soil properties. II. Block kriging. J. Soil Sci. 31:333-341.

Burgess, T.M., Webster, R., and McBratney, A.B. (1981) Optimal interpolation and isarithmic mapping of soil properties. IV. Sampling strategy. J. Soil Sci. 32:643-659.

Buringh, P., Steur, G.G.L., and Vink, A.P.A. (1962) Some techniques and methods of soil survey in The Netherlands. Neth. J. Agric. Sci. 10:157-172.

Burrough, P.A. (1969) Studies in soil survey methodology. D.Phil. Thesis, Oxford.

Burrough, P.A. (1983a) Multiscale sources of spatial variation in soil. I. The application of fractal concepts to nested levels of soil variation. J. Soil Sci. 34:577-597.

Burrough, P.A. (1983b) Multiscale sources of spatial variation in soil. II. A non-Brownian fratal model and its application in soil survey. J. Soil Sci. 34:599-560.

Burrough, P.A., Beckett, P.H.T., and Jarvis, M.G. (1971) The relation between cost and utility in soil survey (I-III). J. Soil Sci. 22, 3:359-393.

Butler (1980) Soil classification for soil survey. Clarendon Press, Oxford.

Calderbank, V.J., and Prior, W.A.J. (1980) GHOST graphical output system. User manual, Culham Laboratory, Oxford.

Campbell, J.B. (1978) Spatial variation of sand content and pH within single contiguous delineations of two soil mapping units. Soil Sci. Am. J. 42:460-464.

Canarache, A., Anca Mielscescu, and Popescu, M. (1978) A data bank on soil moisture properties. In: Developments in Soil Information Systems, Proceedings of the Second Meeting of the ISSS Working Group on Soil Information Systems. Varna/Sofia, Bulgaria. 30 May - 4 June 1977. PUDOC, Wageningen, The Netherlands.

Cassel, D.K. (1983) Spatial and temporal variability of soil physical properties following tillage of Norfolk loamy sand. Soil Sci. Soc. Am. J. 39:247-50.

Cassel, D.K., and Nelson, D.A. (1981) Selection of variables to be used in statistical analysis of field-measured soil water content. Soil Sci. Soc. Am. J. 45:1007-12.

Ciric, M. (1976) Foreword. In: Bie, S.W. (ed), Soil Information Systems, PUDOC, Wageningen, The Netherlands.

Ciric, M., and Palc, J. (1978) Development of soil information systems in Yugoslavia. In: Developments in Soil Information Systems, Proceedings of the Second Meeting of the ISSS Working Group on Soil Information Systems. Varna/Sofia, Bulgaria. 30 May - 4 June 1977. PUDOC, Wageningen, The Netherlands.

Cline, M.G. (1963) Logic of the new system of soil classification. Soil Sci. 96:17-22.

Cook, B.G. (1977) Land resource information systems: Use and display. In: Uses of Soil Information Systems, Proceedings of the Australian Meeting of the ISSS Working Group on Soil Information Systems. Canberra, Australia. March 2-4, 1976. PUDOC, Wageningen, The Netherlands.

Cormack, R.S., and Moore, A.W. (1981) Modelling of land resource data for computer storage and retrieval. In: Information Systems for Soil and Related Data, Proceedings of the Second Australian Meeting of the ISSS Working Group on Soil Information Systems, Canberra, Australia. 19-21 February 1980. PUDOC, Wageningen, The Netherlands.

Courtin, P., Feller, M.C., and Klinka, K. (1983) Lateral variability in some properties of disturbed forest soils in south-western British Columbia. Can. J. Soil Sci. 63:529-39.

Craze, B. (1981) Development of a soils information system for soil conservation use. In: Information Systems for Soil and Related Data, Proceedings of the Second Australian Meeting of the ISSS Working Group on Soil Information Systems, Canberra, Australia. 19-21 February 1980. PUDOC, Wageningen, The Netherlands.

Cuanalo de la Cerda, H.E. (1972) Numerical studies in soil distribution. D.Phil. These. University of Oxford.

Cuanalo de la Cerda, H.E. (1970) A comparative study of numerical classification and ordination of soil profiles in a locality near Oxford. J. Soil Sci. 21:340-42.

Dahiya, I.S., Richter, J., and Malik, R.S. (1984) Soil spatial variability: A review. Intern. J. Trop. Agri. Vol. II, 1:1-102.

David, M. (1977) Geostatistical ore reserve estimation. Elsevier, Scientific Publishing Co., Amsterdam.

De Boor, C. (1962) Bicubic spline interpolation. J. Maths. Phys. 41:212-218.

Delfiner, P. (1976). Linear estimation of non-stationary spatial phenomena. In: M. Guarascio, M.David, and C. Huijbregts (eds.), Advanced geostatistics in the mining industry. Reidel, Dordrecht. pp. 49-68.

Delfiner, P., and Delhomme, J.P. (1975) Optimum interpolation by Kriging. In: Display and analysis of spatial data (eds. J.C.Davis and M.J. McCullagh). pp. 96-114. London, Wiley.

Draper, N.R., and Smith, H. (1981) Applied regression analysis. 2nd Edition, Wiley and Sons, N.Y.

Dumanski, J., and Kloosterman, B. (1978) Soil information systems in perspective. In: Developments in Soil Information Systems, Proceedings of the Second Meeting of the ISSS Working Group on Soil Information Systems. Varna/Sofia, Bulgaria. 30 May - 4 June 1977. PUDOC, Wageningen, The Netherlands.

Dumanski, J., Kloosterman, B., and Brandon, S.E. (1975) Concepts, objectives and structure of the Canada Soil Information System. Can. J. Soil Science 55:181-187.

Emery, K.A. (1981) Practical uses of two information systems developed for land resource data. In: Information Systems for Soil and Related Data, Proceedings of the Second Australian Meeting of the ISSS Working Group on Soil Information Systems, Canberra, Australia. 19-21 February 1980. PUDOC, Wageningen, The Netherlands.

Faith, R., and Sheshinski, R. (1979) Misspecification of trend in spatial random-function interpolation with application to oxident mapping. Technical Report No. 28, Dept. of Statistics, Standford University, Standford, Calif., USA.

FAO (1974) Soil survey in irrigation investigation. Soils Bulletin, FAO,Rome.

FAO (1976) A framework for land evaluation. Soils Bulletin 32, FAO, Rome.

FAO (1981) Report on the agro-ecological zones project. Vol. 3, methodology and results for South and Central America. World Soil Resources Report 48/3, FAO, Rome.

Garbouchev, I., Sadovski, A., and Benevski, M. (1978) Crop monitoring based on soil information systems. In: Developments in Soil Information Systems, Proceedings of the Second Meeting of the ISSS Working Group on Soil Information Systems. Varna/Sofia, Bulgaria. 30 May - 4 June 1977. PUDOC, Wageningen, The Netherlands.

Giltrap, D.J. (1981) A practical application of spatial interpolation for the rapid production of a wide range of special purpose soild maps. In: Information systems for soil and related data. Proceedings 2nd Australian meeting. ISSS Workshop, Canberra, Australia.

Giltrap, D.H. (1984) MIDGE, a microcomputer soil information system. In: Soil Information Systems Technology, Proceedings of the Sixth Meeting of the ISSS Working Group on Soil Information Systems. Bolkesjo, Norway. 28 February - 4 March 1983, PUDOC, Wageningen, The Netherlands.

Green, P.J., and Sibson, R. (1978) Computing Dirichlet tesselations in the plane. The Computer Journal 21,2:168-73.

Hatje, J., and Lamp, J. (1984) Construction of a method base system for evaluation and planning of land and soil resources. In: Soil Information Systems Technology, Proceedings of the Sixth Meeting of the ISSS Working Group on Soil Information Systems. Bolkesjo, Norway. 28 February - 4 March 1983, PUDOC, Wageningen, The Netherlands.

Hayes, J.G., and Halliday, J. (1974) The least-squares fitting of cubic spline surfaces to general data sets. J. Inst. Maths. Applics. 14:89-103.

Hazelden, J., Beckett, P.H.T., and Jarvis, M.G. (1976) A computer* compatible proforma for field soil records. Geoderma, 15:21-29.

Hodgson, J.M. (1978) Soil sampling and soil description. Oxford University Press, Oxford.

Hoerl, A.E., and Kennard, R.W. (1970) Ridge regression: biased estimation for non-orthogonal problems. Technometrics 12:55-67.

Hole, F.D., and Hironaka, M. (1960) An experiment in ordination of some soil profiles. Soil Sci. Soc. Am. proc. 24:309-12.

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. J. Ed. Psych. 24:417-441, 489-520.

Hunter, R.B., Romney, E.M., and Wallace, A. (1982) Nitrate distribution in Mojave Desert soils. Soil Sci. 134:22-30.

Jarvis, M.G., and Hazelden, J. (1982) Soils in Oxfordshire I. Sheet SP30 (Witney south). Soil Survey Record No. 77. Soil Survey of England and Wales, Rothamsted, Harpenden, Herts.

Jenny, H. (1941). Factors of soil formation. McGraw-Hill, New York.

Journel, A.G., and Huijbregts, C.J. (1978) Mining geostatistics. Academic Press, London.

Kloosterman, B., and Dumanski, J. (1978) Data management capabilities of the Canada Soil Information System (CanSIS). In: Developments in Soil Information Systems, Proceedings of the Second Meeting of the ISSS Working Group on Soil Information Systems. Varna/Sofia, Bulgaria. 30 May - 4 June 1977. PUDOC, Wageningen, The Netherlands.

Krige, D.G. (1966) Two dimensional weighted moving average trend surfaces for ore-evaluation. J. South Africa Inst. Min. Metal. 66:13-38.

Lawes Agricultural Trust (1983) GENSTAT: A general statistical program. Rothamsted Experimental Station.

Lynch, L.G. (1977) Input methods and facilities available for land survey data. In: Uses of Soil Information Systems, Proceedings of the Australian Meeting of the ISSS Working Group on Soil Information Systems. Canberra, Australia. March 2-4, 1976. PUDOC, Wageningen, The Netherlands.

Lynch, L.G. (1981) The representation of spatial data by grid cell or polygon. In: Information Systems for Soil and Related Data, Proceedings of the Second Australian Meeting of the ISSS Working Group on Soil Information Systems, Canberra, Australia. 19-21 February 1980. PUDOC, Wageningen, The Netherlands.

Mackenzie, H.G., and Smith, J.L. (1977) Data storage and retrieval. In: Uses of Soil Information Systems, Proceedings of the Australian Meeting of the ISSS Working Group on Soil Information Systems. Canberra, Australia. March 2-4, 1976. PUDOC, Wageningen, The Netherlands.

Malthus, T.R. (1798) Essay on the principle of population. In: Samuelson, P.A., 1976. Economics. McGraw-Hill Kogakusha Ltd, N.Y.

Marsman, B.A. (1973) Soils of sheet SP30 (Witney), part II. Long and medium scale survey of three sample areas. Publ. No. 1107, Soil Survey Institute, Wageningen, The Netherlands.

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

Matheron, G. (1971) The theory of regionalised variables and its applications. Cahiers du Centre de Morphologie Mathematique, Fintainebleau, No. 5.

McBratney, A.B., and Webster, R. (1981a) Spatial dependence and classification of the soil along a transect in northest Scotland. Geoderma 26:63-82.

McBratney, A.B., and Webster, R. (1981b) The design of optimal sampling schemes for local estimation and mapping of regionalized variables. II. Program and examples. Comp. Geosci. 7:335-365.

McBratney, A.B., and Webster, R. (1983) Optimal interpolation and isarithmic mapping of soil properties. V. Co-regionalization and multiple sampling strategy. J. Soil Sci. 34:137-162.

McBratney, A.B., Webster, R., and Burgess, T.M. (1981) The design of optimal sampling schemes for local estimation and mapping of regionalized variables. I. Theory and method. Comp. Geosci. 7:331-334.

McBratney, A.B., Webster, R., McLaren, R.G., and Spiers, R.B. (1982) Regional variation of estractable copper and cobalt in the topsoil of south-east Scotland. Agronomie 2:969-982.

McCloy, K.R. (1981) Interfacing soil and remotely sensed data types:implictions and benefits. In: Information Systems for Soil and Related Data, Proceedings of the Second Australian Meeting of the ISSS Working Group on Soil Information Systems, Canberra, Australia. 19-21 February 1980. PUDOC, Wageningen, The Netherlands.

McDonald, R.C. (1981) Proposed field description sheet associated with the Australian Soil and Land Survey Handbook. In: Information Systems for Soil and Related Data, Proceedings of the Second Australian Meeting of the ISSS Working Group on Soil Information Systems, Canberra, Australia. 19-21 February 1980. PUDOC, Wageningen, The Netherlands.

Modenesi, M.Ch. (1983) Weathering and morphogenesis in a tropical plateau. Catena 10:237-51.

Moore, A.W. (1978) Classification and soil information systems. In: Developments in Soil Information Systems, Proceedings of the Second Meeting of the ISSS Working Group on Soil Information Systems. Varna/Sofia, Bulgaria. 30 May - 4 June 1977. PUDOC, Wageningen, The Netherlands.

Moore, A.W., Russell, J.S., and Ward, W.T. (1972) Numerical analysis of soils: a comparison of three soil profile models with field classification. J. Soil Sci. 23:193-209.

NAG (1982) NAG Library Manual. Numerical Algorithms Group, Mayfield House, Oxford, UK.

Nakano, Y., Cho, T., and Hillel, D. (1983) Effect of transient, partial area-shading on evaporation from a bare soil. Soil Sci. 135:282-95.

Nielsen, D.R., Biggar, J.W., and Erh, K.T. (1973) Spatial variability of field-measured soil-water properties. Hilgardia 42:215-60.

Norris, J.M. (1971) The application of multivariate analysis to soil studies. I. Grouping of soils using different properties. J. Soil Sci. 22, 1:69-80.

Norris, J.M. (1972) The application of multivariate analysis to soil studies. III. Soil Variation.

Nortcliff, S. (1978) Soil variability and reconnaissance soil mapping: a statistical study in Norfolk. J. Soil Sci. 29:403-418.

Oliver, M.A. (1984) Soil variation in the Wyre Forest: its elucidation and measurement. Ph.D. Thesis, Birmingham University.

Olea, R.A. (1975) Optimum mapping techniques using regionalized variable theory. Series on Spatial Analysis No. 2. Kansas Geological Survey, Lawrence.

Olea, R.A. (1977) Measuring spatial dependence with semi-variograms. Series on Spatial Analysis No. 3. Kansas Geological Survey, Lawrence.

Pachepsky, Y.A. (1978) Computer storage and retrieval of environmental managment data for irrigated territories. In: Developments in Soil Information Systems, Proceedings of the Second Meeting of the ISSS Working Group on Soil Information Systems. Varna/Sofia, Bulgaria. 30 May - 4 June 1977. PUDOC, Wageningen, The Netherlands.

Platou, S.W. (1984) The ADK data base systems (PDS and LADS) used in Denmark. In: Soil Information Systems Technology, Proceedings of the Sixth Meeting of the ISSS Working Group on Soil Information Systems. Bolkesjo, Norway. 28 February - 4 March 1983, PUDOC, Wageningen, The Netherlands.

Ponce-Hernandez, R. (1975) Analisis de tendencia de la variacion de propiedades del suelo por metodos graficos y numericos. Tesis profesional, E.N.A., Chapingo, Mexico.

Ponce-Hernandez, R., Marriott, F.H.C., and Beckett, P.H.T. (1986) An improved method for reconstructing a soil profile from analyses of a small number of samples. J. Soil Sci. 37:455-467.

Price, G.H. (1977) The use of a computer in a commercial soil analysis service. In: Uses of Soil Information Systems, Proceedings of the Australian Meeting of the ISSS Working Group on Soil Information Systems. Canberra, Australia. March 2-4, 1976. PUDOC, Wageningen, The Netherlands.

Rayner, J.H. (1966). Classification of soils by numerical methods. J. Soil Sci. 17:79-92.

Rhind, D. (1975) A skeletal overview of spatial interpolation techniques. Computer Applications 2:293-309.

Riquier, J., and Chidley, T.R.E. (1978) The FAO international soil data bank. In: Developments in Soil Information Systems, Proceedings of the Second Meeting of the ISSS Working Group on Soil Information Systems. Varna/Sofia, Bulgaria. 30 May - 4 June 1977. PUDOC, Wageningen, The Netherlands.

Rourke, J.D. (1981) Legend design for various kinds of soil surveys. In: Soil Resource Inventories and Development Planning. Proceedings of Workshops at Cornell University, 1977-78. Technical Monograph No. 1. Soil Management Support Services. Soil Conservation Service, U.S.D.A, Washington, D.C.

Russell, J.S., and Moore, A.W. (1968) Comparison of the different depth weightings in the numerical analysis of anisotropic soil profile data. Transactions of the 9th International Congress of Soil Science, 4:205-213.

Ryan, P.J., and McGarity, J.W. (1983a) The nature of spatial variability of soil properties adjacent to large forest eucalyptus. Soil Sci. Soc. Am. J. 47:286-93.

Sadovski, A.N. (1984) Soil data and decision making in agriculture. In: Soil Information Systems Technology, Proceedings of the Sixth Meeting of the ISSS Working Group on Soil Information Systems. Bolkesjo, Norway. 28 February - 4 March 1983, PUDOC, Wageningen, The Netherlands.

SAS (1981) SAS User's Guides: Statistics and SAS/Graph. SAS Institute Inc., Cary, N.C. 27511-8000, USA.

Schafer, B.M., Nielsen, G.A., and Nettleton, W.D. (1979) Morphology of a paralithic contact in a soil over soft sandstone. Soil Sci. Soc. Am. J. 43:383-86.

Schut, G.H. (1974) Evaluation of some interpolation methods. In: Proceedings of the Symposium of Commission III of the TSP, Stuttgart, West Germany.

Schut, G.H. (1976) Review of interpolation methods for digital terrain models. The Canadian Surveyor 30,5:389-411.

Scott, R.M. (1981) The South Australian environmental data banks. In: Information Systems for Soil and Related Data, Proceedings of the Second Australian Meeting of the ISSS Working Group on Soil Information Systems, Canberra, Australia. 19-21 February 1980. PUDOC, Wageningen, The Netherlands.

Sisson, J.B., and Wierenga, P.H. (1981) Spatial variability of steady state infiltration rates as a stochastic process. Soil Sci. Soc. Am. J. 45:699-704.

Sleeman, J.R. (1977) Use of a storage and retrieval system in soil fabric analysis. In: Uses of Soil Information Systems, Proceedings of the Australian Meeting of the ISSS Working Group on Soil Information Systems. Canberra, Australia. March 2-4, 1976. PUDOC, Wageningen, The Netherlands.

Smith, J.L. (1981) A tutorial introduction to FORDATA. In: Information Systems for Soil and Related Data, Proceedings of the Second Australian Meeting of the ISSS Working Group on Soil Information Systems, Canberra, Australia. 19-21 February 1980. PUDOC, Wageningen, The Netherlands.

Soil Survey Staff (1951) Soil survey manual. U.S. Dept. Agr. Handb. 18. U.S. Govt. Printing Office, Washington, D.C.

Starks, T.H., and Fang, J.M. (1982) The effect of drift on the experimental semivariogram. J. Int. Assoc. Math. Geol. 14:309-319.

Steur, G.G.L. (1961) Methods of soil surveying in use at The Netherlands Soil Survey Institute. Boor Spade 11:59-77.

Tobler, W.R. (1975) Linear operators applied to areal data. In: Davis, J.C., and McCullagh, M.J. (eds.), Display and Analysis of Spatial Data. NATO Advance Studies Institute 14-13. John Wiley and Sons, London.

Tomlinson, P.R., Bannister, P., Marsden, R., and Beckett, P.H.T. (1977) Simplified procedures for routine soil analysis. J. Appl. Ecol. 14:253-260.

Trangmar, B.B., Yost, R.S., and Uehara, G. (1985) Application of geostatistics to spatial studies of soil properties. In: Advances in Agronomy, vol. 38, 45-94. Academic Press, London and N.Y.

Van Berkel, P.R., and Eyles, G.O. (1981) The New Zealand resource inventory database for the national water and soil conservation organisation. In: Information Systems for Soil and Related Data, Proceedings of the Second Australian Meeting of the ISSS Working Group on Soil Information Systems, Canberra, Australia. 19-21 February 1980. PUDOC, Wageningen, The Netherlands.

Van der Driessche, R., Garcia Gomez, A.M., Aubury, A.M., and Giey, A. (1978) POSEIDON. In: Developments in Soil Information Systems, Proceedings of the Second Meeting of the ISSS Working Group on Soil Information Systems. Varna/Sofia, Bulgaria. 30 May - 4 June 1977. PUDOC, Wageningen, The Netherlands.

Venkataratnam, L. (1984) Remotely sensed multiband digital data for soil information systems: Problems and prospects in India. In: Soil Information Systems Technology, Proceedings of the Sixth Meeting of the ISSS Working Group on Soil Information Systems. Bolkesjo, Norway. 28 February - 4 March 1983, PUDOC, Wageningen, The Netherlands.

Vevers, E.V. (1978) Soil chemical testing data bank in the Latvian SSR. In: Developments in Soil Information Systems, Proceedings of the Second Meeting of the ISSS Working Group on Soil Information Systems. Varna/Sofia, Bulgaria. 30 May - 4 June 1977. PUDOC, Wageningen, The Netherlands.

Vieira, S.R., Nielsen, D.R., and Biggar, J.W. (1981) Spatial variability of field-measured infiltration rate. Soil Sci. Soc. Am. J. 45:1040-1048.

Vink, A.P.A. (1963) Planning of soil surveys in land development. International Institute for Land Reclamation and Improvement (ILRI) Publication No. 23. Wageningen, The Netherlands.

Wagenet, R.J., and Jurinak, J.J. (1978) Spatial variability of soluble salt content in a Mancos watershed. Soil Sci. 126:342-49.

Walker, P.H., Hall, G.F., and Protz, R. (1961) Soil trends and variability across selected landscapes in Iowa. Soil Sci. Soc. Am. Proc. 32:97-101.

Webster, R. (1968) Fundamental objections to the 7th Approximation. J. Soil Sci. 19:354-66.

Webster, R. (1971) Wilk's criterion: A measure for comparing the value of general purpose soil classifications. J. Soil Sci. 22:254-60.

Webster, R. (1973) Automatic soil boundary location from transect data. Math. Geol. 5:27-37.

Webster, R. (1976) The nature of soil variation. Classification Society Bull. 3:43-55.

Webster, R. (1977) Quantitative and numerical methods in soil classification and survey. Oxford University Press.

Webster, R. (1978) Optimally partitioning soil transects. J. Soil. Sci. 29:388-402.

Webster, R. (1981) Spatial analysis of soil and its application to soil mapping. In: Computer Applications in Geology, Misc. papers, Geological Society of London.

Webster, R. (1982) Spatial analysis of soil and its application to soil mapping. In: Computer Applications in Geology. I and II. Miscellaneous Paper No. 14, pp. 103-136. Geological Society of London.

Webster, R. (1985) Quantitative spatial analysis of soil in the field. In: Advances in Soil Science, Vol. 3. Springer-Verlay, N.Y.

Webster, R. (1986) Statistical methods for analysing soil variation. In: Transactions of the XIII Congress of The International Society of Soil Science. 13-20 August, 1986. Hamburg, West Germany.

Webster, R., and Beckett, P.H.T. (1968) Quality and usefulness of soil maps. Nature (London) 219:680-682.

Webster, R., and Burgess, T.M. (1980c) Optimal interpolation and isarithmic mapping of soil properties. III. Changing drift and universal kriging. J. Soil Sci. 31:505-524.

Webster, R., and Burgess, T.M. (1983) Spatial variation in soil and the role of kriging. Agricultural Water Management 6:111-122.

Webster, R., and Burrough, P.A. (1972a) Computer-based soil mapping of small areas from sample data. I. Multivariate classification and ordination. J. Soil Sci. 23:210-21.

Webster, R., and Cuanalo de la Cerda, H.E. (1975) Soil transect correlograms of north Oxfordshire and their interpretation. J. Soil Sci. 26:176-194.

Webster, R., and Nortcliff, S. (1984) Improved estimation of micro nutrients in hectare plots of the Sonning series. J. Soil Sci. 35:667-672.

Western, S. (1978) Soil survey contracts and quality control. Oxford University Press.

Wilding, L.P., and Drees, L.R. (1983) Spatial variability and pedology. In: Wilding, L.P., N.E. Smeck and G.F. Hall (eds.), 1983. Pedogenesis and Soil Taxonomy.

Wold, S. (1974) Spline functions in data analysis. Technometrics 16:1-11.

Xu Jiyan, and Webster, R. (1984) A geostatistical study of topsoil properties in Zhangwu County, China, Catena 11:13-26.

Young, A. (1976) Tropical soils and soil survey. Cambridge University Press.

Yost, R.S., Uehara, G., and Fox, R.L. (1982a) Geostatistical analysis of soil chemical properties of large land areas. I. Semivariograms. Soil Sci. Soc. Am. J. 46:1028-1032.

Yost, R.S., Hehara, G., and Fox, R.L. (1982b) Geostatistical analysis of soil chemical properties of large land areas. II. Kriging. Soil Sci. Soc. Am. J. 46:1033-1037.

Zinke, P.J., and Crocker, R.L. (1962) The influence of giant sequoia on soil properties. Forest Sci. 8:2-11.